

**Word order and language change.
On the interface between syntax and morphology**

Habilitationsschrift

eingereicht im Fachbereich Neuere Philologien (FB 10)
der Johann-Wolfgang-Goethe Universität
zu Frankfurt am Main

vorgelegt von

Eric Fuß
aus Frankfurt am Main

2008

Table of contents

Acknowledgments.....	1
Chapter 1: Language change – a generative perspective	3
1 Introduction	3
2 Foundations of a formal approach to language change	7
2.1 On the notion of 'language'.....	10
2.2 'Language change' vs. 'grammar change'	11
2.3 Possible causes of grammar change	13
3 Goals of historical linguistics.....	18
4 Outline of this work.....	19
Chapter 2: The syntax-morphology interface	22
1 Introduction	22
2 A realizational model of grammar	24
2.1 Blocking effects and the workings of Vocabulary Insertion	26
2.2 L-morphemes, f-morphemes, and the syntactic determination of lexical categories	28
2.3 Post-syntactic morphological operations.....	29
3 Cyclic Spell-Out and the domain of post-syntactic operations	33
3.1 Phases, Spell-Out domains, and the scope of PF operations.....	34
3.2 Complementizer agreement in Germanic.....	40
3.3 Cyclic Spell-Out and phonological domains	48
3.4 Section summary.....	55
4 The creation of linear order	57
4.1 Symmetric c-command and a phonological Head Parameter	59
4.1.1 The status of specifiers.....	65
4.2 Cyclic linearization and Vocabulary Insertion.....	68
4.2.1 LIN and X^0 -structures	68
4.2.2 LIN and XP-structures	73
4.2.3 LIN and the relative order of phonological domains.....	78
4.3 Section summary.....	87
5 The distribution of sentential complements in German.....	88
6 Possible pathways for word order change	105
6.1 Possible and impossible grammars	106
6.1.1 <i>Edge Replacement</i> vs. FOFC	118
6.2 Word order variation and change in the history of English.....	123
6.2.1 FOFC again: A pied-piping analysis of OE word order facts.....	130
6.2.2 <i>Edge Replacement</i> and OE word order	137
6.2.3 Some remarks on the rise of VO in the Middle English period	149
7 Summary	157
Chapter 3: V2 in Early Germanic.....	163
1 Introduction.....	163
2 Verb fronting and 'operator V2' in Gothic	167
2.1 Apparent counterexamples.....	170
2.1.1 Pronoun placement in <i>wh</i> -questions.....	170
2.1.2 Second position particles	172
2.1.2.1 The coordinating particle <i>-uh</i>	174
2.1.2.2 The modal particles <i>þan</i> , <i>nu</i> and <i>auk</i>	176
2.2 Further contexts that trigger V2.....	179
2.3 Section summary.....	183
3 Old English: operator V2 + 'pseudo V2'	184

3.1	Systematic deviations from V2 in Old English.....	185
3.1.1	V-first and V-final main clauses.....	185
3.1.2	V3 orders involving non-pronominal DP subjects	187
3.1.3	V2 and pronoun placement.....	187
3.2	Toward an analysis.....	192
3.3	The syntax-discourse interface and inversion in OE	199
3.3.1	The temporal interpretation of clause-initial <i>þa/þonne</i>	200
3.3.2	The distribution of sequential <i>þa/þonne</i> and subject pronouns in OE... <td>204</td>	204
3.3.3	Discourse-configurationality and the nature of SpecTP in OE	207
4	The 'loss' of V2 in the Middle English period	214
4.1	The loss of discourse-configurationality and the rise of	220
4.1	'fossilized' movement.....	220
5	V2 in Old High German	224
5.1	Introduction.....	224
5.2	Generalized V2?	226
5.3	Deviations from V2 in main declaratives.....	236
5.3.1	V-first order in main clauses	236
5.3.2	V3 orders.....	240
5.4	V2 and pronoun placement in Old High German	249
5.4.1	V2 and pronoun placement in the OHG <i>Isidor</i>	249
5.4.2	V2 and pronoun placement in the <i>Tatian</i>	257
5.4.3	Section summary	259
5.5	V2 and the placement of the adverb <i>thô</i>	260
5.6	Some remarks on the analysis of V2 in Old High German	265
5.6.1	A multiple specifiers approach to V2 in Old High German.....	270
5.6.2	The loss of V3 orders in OHG	282
5.7	Section summary: V2 in Old High German.....	294
6	Summary	295
Chapter 4: On the cyclic nature of language change.....		300
1	Introduction.....	300
2	Blocking and the grammaticalization of verbal inflection	303
2.1	The grammaticalization of agreement markers in German/Bavarian.....	304
2.1.1	Properties of blocking-induced change.....	308
3	Analological change.....	310
3.1	The development of <i>Einheitsplural</i> in Alemannic	311
3.1.1	The extension of 3pl <i>-nt</i> to 2pl.....	313
3.1.2	The rise of a general plural marker: Extension to 1pl	317
3.2	Analological change vs. blocking-induced change	321
3.3	Section summary.....	324
4	The rise and fall of null subjects	325
4.1	The rise and loss of partial pro-drop.....	328
4.1.1	Deblocking the zero spell-out of weak pronouns.....	334
4.1.2	Blocking the zero spell-out of weak pronouns: Colloquial Finnish.....	338
4.2	The rise of full-fledged pro-drop: Colloquial French.....	342
5	Summary	349
Appendix: Alternative paths toward null arguments.....		353
	Mauritian Creole.....	353
	Chabacano	359
Chapter 5: Concluding summary.....		365
References		373

Acknowledgments

My thinking on the subjects explored in this work has been shaped by collaborations and discussions with a variety of people, to whom I would like to express my gratitude.

First of all, thanks go to Günther Grewendorf for his support over the years; many ideas present in this work has been shaped (and reshaped) by discussions with him. In addition, I want to thank the other members of my committee, who immediately agreed to act as reviewers for this thesis (despite the fact that my request came at short notice): Ian Roberts, Alessandra Tomaselli, Helmut Weiß, and Ede Zimmermann.

I am also indebted to a number of colleagues and friends with whom I discussed various aspects of this work: Werner Abraham, Adam Albright, Katrin Axel, Josef Bayer, Ermenegildo Bidese, Theresa Biberauer, Gisella Ferraresi, Gaetano Fiorin, Susann Fischer, Jürg Fleischer, Werner Frey, Hans-Martin Gärtner, Kleanthes Grohmann, Eric Haeberli, Liliane Haegeman, Fabian Heck, Roland Hinterhölzl, Jiro Inaba, Agnes Jäger, Dianne Jonas, Hans Kamp, Shin-Sook Kim, Joost Kremers, Elisabeth Leiss, Gereon Müller, Cécile Meier, Ad Neeleman, Andrew Nevins, Damaris Nübling, Albert Ortmann, Svetlana Petrova, Cecilia Poletto, Monika Rathert, Marga Reis, Marc Richards, Antje Roßdeutscher, Joachim Sabel, Oliver Schallert, Eva Schlachter, Hans-Christian Schmitz, Peter Sells, Halldór Sigurðsson, Jochen Trommer, Anne Vainikka, Tonjes Veenstra, and Melani Wratil. (To be sure, this name list is not exhaustive, and I apologize to those who I should have mentioned here, but failed to do so).

Special thanks go to Patrick Brandt, Matthias Schulze-Bünte and Carola Trips, who read and commented on portions of the text. Of course, all remaining errors are mine. In addition, I want to thank Eberhard Fahlke for handling the administrative side of things as smoothly and professionally as possible.

A large portion of this work grew out of a series of lectures that I delivered in September 2007 at the University of Verona. I thank the organizers for inviting me, in particular Alessandra Tomaselli, Denis Delfitto, and Giorgio Graffi, who provided many insightful comments that led me to rethink various aspects of this work. In addition, I would like to thank the members of the DFG networks on “Historical syntax” and “Core mechanisms of exponence” as well as audiences at various other occasions where I presented parts and pieces of the present ideas, including the

annual meeting of the LAGB at the University of Cambridge (2005), the *Zentrum für allgemeine Sprachwissenschaft* in Berlin (2005), GGS 2006 and CGSW 2007 at the University of Stuttgart, the 2006 Workshop on Null Pronouns at the University of Jena, the *Linguistische Arbeitskreis* at the University of Cologne (2007), WOTM 2007 at the University of Leipzig, the symposium on “Language change and variation” at the University of Munich (2007), DGfS 2008 at the University of Bamberg, and the 44th Annual Meeting of the Chicago Linguistic Society at the University of Chicago (2008).

I would also like to give my warmest thanks to my parents and my grandma. Last but not least, I want to thank Anne, Anton & Adrian for their patience and support during the last five months. Without them, this work would not have been possible.

Editorial note:

Some parts of this work has already been published, albeit in different form. Section 3 of chapter 2 (“The syntax-morphology interface”) is a revised and shortened version of Fuß (2007). Sections 2 to 4 (except of section 2.2) of chapter 3 (“V2 in early Germanic”) are based on Fuß (2003) and Trips & Fuß (2008), but differ from the latter works in major aspects (in particular concerning specifics of the analysis proposed for Old English and the loss of V2 in the Middle English period).

Chapter 1: Language change – a generative perspective

1 Introduction

It is a long-standing conjecture that core aspects of the phenomenon of language change are deeply rooted in the process of first language acquisition. Compare the following quote taken from Paul (1880[⁸1968]: 34):

“Es liegt auf der Hand, dass die Vorgänge bei der Spracherlernung von der allerhöchsten Wichtigkeit für die Erklärung der Veränderung des Sprachusus sind, dass sie die wichtigste Ursache für diese Veränderungen abgeben.”

‘It is obvious that the processes of language acquisition are of the greatest importance for the explanation of changes of language use, that they constitute the most important cause of these changes.’ (translation: EF)

However, this claim seems to contradict the widely-held assumption that children always succeed in acquiring the target grammar that generates the linguistic data they are exposed to, even if this data is apparently flawed and insufficient (sometimes called the “the logical problem of language acquisition”, cf. Chomsky 1986a for discussion). This leads to the “logical problem of language change” as Niyogi and Berwick (1998) choose to call it (see also Roberts 2007a: 230f. for discussion):

“After all, if all children successfully attain the grammars of their parents and they continue to do this generation after generation, then the linguistic composition of every generation would look exactly like the linguistic

composition of the previous generation and languages would not change with time. Yet they do." (Niyogi and Berwick 1998: 192f.)

Thus, more has to be said to reconcile the fact that languages change with the idea that language acquisition leads to 'perfect' results. Under the plausible assumption that language acquisition is a deterministic process (that is, two different sets of input data give rise to two different grammars), the possibility of language change can be attributed to changes in the Primary Linguistic Data (PLD), that is the set of (partially parsed) linguistic signals on the basis of which the learner constructs a grammar (Chomsky 1965: 25). In other words, it is usually assumed that for some reason, the PLD the learner is confronted with differs from the PLD that gave rise to the target grammar, due to factors such as language contact, (morpho-) phonological erosion or reanalyses that blur the evidence for certain properties of the target grammar in the linguistic input the learner receives (cf. e.g. Lightfoot 1979, 1991, 1999, Hale 2007, Roberts 2007a). Of course, this raises a number of further questions, in particular concerning the way the language acquisition device (LAD) converts information conveyed by the PLD into a grammar G with a set of properties $\{P_1 \dots P_n\}$. Thus, in the course of language acquisition, children have to engage in the non-trivial task of detecting underlying structural properties of the target grammar via inspecting the linguistic output of that grammar, that is, linear sequences of sounds (or signs, in the case of sign languages). As a first working definition, we may say that language change occurs if the structural descriptions that the learner assigns to the input data differ from the relevant structural descriptions that are part of the target grammar. It is a central goal of any theoretical approach to language change to develop a restrictive theory of such mismatches between the target grammar and the grammar acquired by the learner.

This work adopts the view that for a certain language L , at any point in time, the set of possible changes is restricted by (i) general properties of grammar (in the sense that the outcome of a change must be a possible human language as defined by properties of UG), and (ii) factors governing the process of language acquisition, in particular acquisition strategies that the learner applies to the data to detect underlying properties of linguistic categories from the auditive/gestual input he/she receives (e.g., via paradigmatic oppositions that can be discerned in the Primary Linguistic Data). Accordingly, it is one goal of this work to explore aspects of grammar and its acquisition that delimit the set of possible changes and therefore provide us with a first approximation of a theory of language change.

Apart from developing a better understanding of the factors that delimit the set of possible changes, we may also ask whether there are properties of the acquisition device that may promote changes or determine the direction of change in case the evidence contained in the PLD is ambiguous or insufficient. It is commonly assumed that in this case, the eventual outcome is determined by acquisition principles that select the most economical grammar compatible with the PLD. Relevant proposals in the literature are based either on the notion of markedness or on the notion of derivational/representational economy. The former line of thinking has it that there are marked and unmarked (or default) parameter values and that the learner assigns a given parameter the unmarked value if no decision can be made based on the evidence available in the input (cf. Wexler and Culicover 1980, Berwick 1985, Wexler and Manzini 1987; see O'Grady 1997 for an overview and discussion). Approaches based on derivational/representational economy assume that the learner assigns a given input string the most economical representation/derivation that is compatible with the input data in case the evidence provided by the PLD is unclear (Clark and Roberts 1993, Roberts 1993a, 1993b, Roberts and Roussou 2003, Roberts 2007a; cf. Roberts 1999 for an account that combines markedness considerations with the notion of derivational economy).¹

In this work, I examine how the course of language change is shaped by the interaction of (i) universal properties of grammar that delimit the set of possible changes and (ii) mechanisms of the language acquisition device that may promote certain changes in case the relevant properties are underdetermined by the PLD.

Drawing on data from the history of the Germanic languages (in particular German and English), I will consider a selection of changes that affected phenomena at the interface between syntax and morphology/PF, focusing on word order (in particular V2 and the OV-VO parameter), the inventory, shape, and featural specifications of inflectional markers, and the availability of null arguments.

Adopting a realizational model of grammar (i.e., Distributed Morphology, Halle and Marantz 1993), I am going to argue that the set of possible mismatches

¹ A different approach is developed in Lightfoot (1999) who argues for a 'cue-based' theory of language acquisition (cf. Dresher and Kaye 1990, Dresher 1999 for earlier cue-based models of the acquisition of phonological properties). The basic assumption is that UG contains not only a set of parameters, but also specifies for each parameter a *cue* that serves to switch the parameter one way or other (cf. Fodor 1998 for a related approach). If the learner detects a cue that is attested robustly in these (initially incomplete) parses, this will activate a given parameter or syntactic operation in the learner's grammar. Language change results either if a given linguistic feature fails to be cued or if it starts to be cued, in contrast to the target grammar.

between the target grammar and the grammar eventually acquired by the learner is delimited by ‘hard-wired’ interface properties that determine the mapping from hierarchical syntactic structures to linear strings of sounds (i.e., phonological exponents). In particular, the mapping between syntax and morphology/PF is governed by a set of hardwired algorithms that the learner can apply in a backwards fashion to the linguistic input (i.e., strings of words/sounds) to ‘undo’ the workings of the linearization procedure, reconverting linear orderings into hierarchical structures. This delimits the set of structural hypotheses the learner must entertain when he/she is confronted with a certain surface string. In the same way, this kind of hard-wired knowledge restricts the set of possible misanalyses when applied to a certain data set and therefore determines possible pathways of change.

In addition, I assume that language change is shaped by acquisition strategies that the learner applies to the input he/she receives in case the linguistic evidence is ambiguous or not sufficient to trigger a certain property of the grammar. With respect to the acquisition of syntactic properties, I am going to propose that the learner may resort to endowing functional heads with semantically vacuous EPP-features in order to mimic dislocation phenomena the original semantic/pragmatic trigger of which has become unclear (cf. Simpson 2004). The effects of this ‘learning strategy’ are illustrated with changes affecting the V2 property in the history of English and German, and the loss of OV structures in English. Furthermore, we will see that the acquisition of phonological exponents of inflectional categories is shaped by another set of learning strategies which involve the notion of morphological blocking and a tendency to minimize the number of features/elements stored in the lexicon. It will become clear that the cyclic nature of changes affecting the null realization of subjects and the inventory and shape of inflectional markers in various varieties of German can be directly attributed to the workings of these acquisition strategies.

This chapter is organized as follows. Section 2 is concerned with a set of general issues that arise when language change is studied in a more formal way. In particular, I am going to argue that the proper object of a formal study of language change should be identified as ‘grammar change’, that is, a set of discrete differences between the target grammar and the grammar acquired by the learner. Section 3 summarizes the issues discussed so far. An overview of the chapters to come is provided in section 4.

2 Foundations of a formal approach to language change²

This section argues that we have to sharpen our definitions of 'language' and 'language change' if we want to deepen our understanding of structural restrictions on (and motivations of) language change. Even in theoretically informed work in historical linguistics, we often come across statements like the following:

(1) "in Middle High German, change X began in the early 12th century and was completed by the end of the 14th century"

Upon closer inspection, it turns out that (1) seems to involve a set of different, and conflicting notions of 'language' and 'change'. At first sight, the notion of "Middle High German" appears to be an idealization similar to the one we use when we talk about the grammars of present-day languages such as 'German', 'English' or 'Italian' (although we are usually aware of the fact that the linguistic features of individual speakers of this language may differ to some extent). However, a second look reveals that in the case at hand, this idealization has a different quality. Speaking of "Middle High German" suggests a uniformity which presumably never existed in any real sense. Rather, the notion "Middle High German" refers to a set of different grammars with different properties. Note that there was a huge amount of dialectal diversification in this period, which is not reflected by the term at hand. Moreover, as the "Middle High German" period extended over more than 300 years, there is also non-uniformity in the temporal dimension. That is, "Middle High German" is in fact a cover term that includes a large number of different grammars. Actually, this fact is already implicit in statement (1): Note that it refers to "Middle High German", whilst stating that there is a change in which the MHG of the 16th century differs from that of the 13th century. This statement makes only sense if we recognize that we deal with different grammars here – obviously two different instances of MHG.

So it turns out that the term "Middle High German" is actually a very vague notion. This of course does not mean that we should never use the term 'Middle High German'. It is a helpful simplification that can be put to work usefully if we want to refer to a certain period in the history of German with certain cultural *and* linguistic properties (and of course, I will use notions such as MHG in this work as

² The following discussion owes much to the work of Mark Hale, in particular Hale (1996), (1998) and (2007).

well) . But we should be aware of the fact that it is a simplification and that it can impede our understanding of the very nature of language change if we forget that.

Second, let's take a closer look at the notion of 'change' that is involved in (1). At closer inspection, it becomes clear that it involves at least two different notions of 'change': First, (1) seems to refer to an innovation that first occurred in the 12th century in the grammars of some speakers, that is, an instance of a 'real' linguistic change, in which the grammars of certain speakers came to differ from the previous grammar (the target grammar).³ In addition, (1) describes a process in which this change spread through the speaker community. This process was completed in the 14th century, when the grammars of (more or less) all speakers were identical with respect to the property in question. 'Change' in the latter sense is sometimes referred to as an instance of *diffusion*, which is a sociolinguistic notion. Note that diffusion does not necessarily involve a real 'change', in the sense that speakers acquired a grammar different from their target grammar. Rather, it refers to a development in which more and more speakers acquire a property that is already part of the input they receive (see also Hale 2007):

- (2) a. *Innovation*
 - (i) The target grammar that generates the PLD has properties A, B, C.
 - (ii) The grammar acquired by the learner has properties A, B, X.
- b. *Diffusion*
 - (i) There is a 'mixed' PLD generated by a grammar with properties A, B, C and another grammar with properties A, B, X.
 - (ii) The grammar acquired by the learner has properties A, B, X.

The confusion of different notions of 'change' in statements such as (1) arises at least partially from the fact that what we perceive as 'language change' is normally only the result of diffusion. In fact, there are presumably myriads of changes that never show up in the records since they were confined to a single speaker and never spread to others speakers, let alone to the whole community of speakers.

³ In addition, note that (1) obviously refers to the chronology of the change as it appears in the historical records and not to the actual change in the speaker community. As is well-known, written language is quite conservative. Accordingly, there is usually quite some difference between the chronology of the actual change, which typically occurs in the spoken language, and its first appearance in the historical records reflecting this change.

Still, it seems that we must focus on the first kind of change, that is, innovations if we aim at developing a restrictive theory of language change. Consider the following pair of words (the example is taken from Hale 2007: 39):

(3) a. Middle English *lutter* 'pure'
 b. Modern English *pure* 'pure'

Of course, the change from *lutter* to *pure* cannot be explained in terms of a restrictive theory of possible sound changes (e.g., /l/ → /p/ is a very unlikely type of sound change). Rather, the change from (3a) to (3b) is an example of borrowing due to language contact with French. As a result, the original English word meaning 'pure' was replaced by the loanword *pure*. As has repeatedly been pointed out in the literature, there are presumably no linguistic constraints on borrowing (cf. e.g. Thomason and Kaufman 1988, Harris and Campbell 1995, Curnow 2001).⁴ Furthermore, note that borrowing represents an instance of diffusion (the input contained both *pure* and *lutter*, and over time, more and more learners acquired *pure* instead of *lutter* as the realization of the concept 'pure'). Thus, while we can formulate a constrained theory of possible sound changes (ruling out a change from (3a) to (3b)), it seems quite unlikely that we can develop a restrictive linguistic theory of possible diffusion/borrowing events. Any change can diffuse, and there are presumably no strong linguistic constraints at work here.

Following Lightfoot (1999) and Hale (2007), I thus assume that the formal study of language change should focus on innovations, that is, discrete differences between the target grammar and the grammar acquired by the learner since only these can be captured by formal linguistic analyses.⁵ Thus, taking a closer look at a seemingly innocent statement such as (1) suggests that we must state more clearly what the proper scientific object of historical linguistics (and linguistics in general) should be.

⁴ However, see Heine (2008), Heine and Kuteva (2005) for an opposing view.

⁵ Of course, the study of diffusion can also reveal important insights, in that it tells us something about social aspects of language, for example the factors that govern the diffusion of forms, the social stratification of speech communities, social factors that govern linguistic variation etc. Crucially, however, it does not tell us much about language itself.

2.1 On the notion of ‘language’

Today, most (generative) linguists agree that the proper object of formal linguistic study is a special notion of ‘language’, namely the linguistic knowledge of an idealized speaker/hearer, cf. Chomsky (1965: 3) for a classic statement:

“Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogenous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance.”

In more recent work of Chomsky’s (1986a), this notion is refined in terms of the distinction between I(nternalized)-language and E(xternalized)-language:

- (4) a. *E-language*: “‘externalized language’ [...], in the sense that the construct is understood independently of the properties of the mind/brain.”
(Chomsky 1986a: 20)
 - (i) Sometimes defined as the set of actual or potential speech events/expressions that are in use in a speech community.
 - (ii) Fairly close to the common-sense (sociopolitical) notion of ‘language’.
- b. *I-language*: “‘internalized language’ [...] some element of the mind of the person who knows the language, acquired by the learner, and used by the speaker-hearer.”
(Chomsky 1986a: 22)
 - (i) Knowledge state in the mind of a particular human being.
 - (ii) Internalized production system (=grammar) that generates a potentially infinite range of linguistic (output) representations.⁶

⁶ Note that ‘linguistic output representations’ are not to be equated with ‘linguistic output’. The former are abstract linguistic representations (i.e., structured sets of features) produced by the grammar that are mapped to sequences of sounds/signs (the actual linguistic output, a piece of E-language) by the speaker’s production system. See below for some discussion of how a less than perfect mapping ‘linguistic output representations’ to ‘linguistic output’ introduces ‘noise in the channel’ that can lead to grammar change.

Chomsky (1986a) argues forcefully that the proper object of linguistic study must be I-language, sometimes also referred to as linguistic *competence* or simply *grammar*. In other words, any principle or rule of grammar that is posited by the linguist is to be seen as a piece of I-language. *Universal Grammar* (UG) is then conceived as a theory of formal universals of human language:

"UG now is construed as the theory of human I-languages, a system of conditions deriving from the human biological endowment that identifies the I-languages that are humanly accessible under normal conditions."

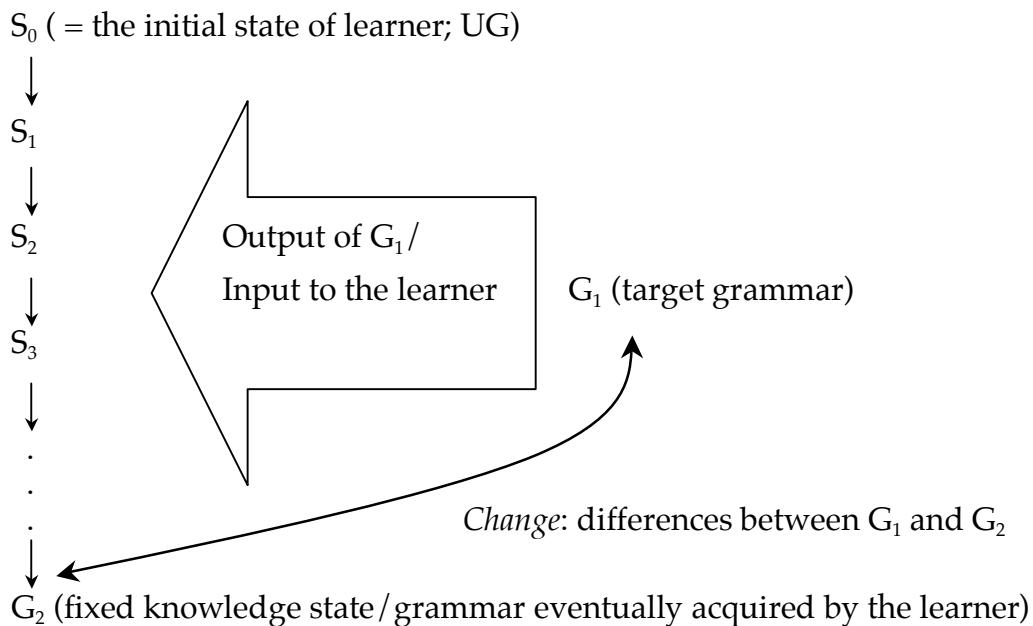
(Chomsky (1986a: 23)

Of course, this means that the proper object of formal *historical linguistics* is I-language(s) – or grammar(s) – as well. In the following I will briefly review some consequences of this position.

2.2 'Language change' vs. 'grammar change'

If we accept the notion that the proper object of the formal study of language change is I-language or *grammar*, then 'language change' must be redefined as a change between (individual) grammars, that is, *grammar change* (cf. Hale 1998, 2007, Lightfoot 1999). This change of perspective implies a particular model of the relationship between acquisition and change (cf. Hale 1998, 2007):

(5) A model of language acquisition and change



Starting out from S_0 (the initial state of grammar, presumably an expression of the genes, which can be modeled in terms of a system of abstract principles, UG), the learner constructs a number of intermediate knowledge stages during the acquisition process based on the evidence provided by the input (where S_n is revised to S_{n+1} if the learner becomes aware of the relevant evidence necessary to trigger a certain property of the grammar (the so-called *triggering experience*)).⁷ Eventually, the process of grammar construction gives rise to a fixed knowledge state which represents the grammar acquired by the learner in the course of language acquisition. This perspective implies that grammar change is necessarily an abrupt phenomenon, namely a clearly identifiable difference between the target grammar G_1 and the acquirer's grammar G_2 . In other words, innovations result from cognitive processes that determine the process of language acquisition, resulting in a grammar in the mind of the individual speaker that differs from the target grammar. Under this approach, 'language change' is to be identified as a rather sociolinguistic notion,

⁷ Note that each of the intermediate stages presumably represents a possible human grammar, cf. the following quote taken from Chomsky (2002: 130f.): "[...] there are no dead ends in language acquisition. You can't set parameters in such a way that you get a system that will fail to have an infinite satisfaction of the interface conditions. [...] the language faculty just has states; one state is the initial state; others are the stable states that people reach somehow, and then there are all kinds of states in between, which are also real states, just other languages. If the strong No Dead End Condition is met, then the minimalist thesis would say that all states have to satisfy the condition of infinite legibility at the interface – and to do so in an optimal manner, to the extent that the strong minimalist thesis holds."

referring to the diffusion of a given change in a speaker community. Thus, we must differentiate between the following aspects of any given change:

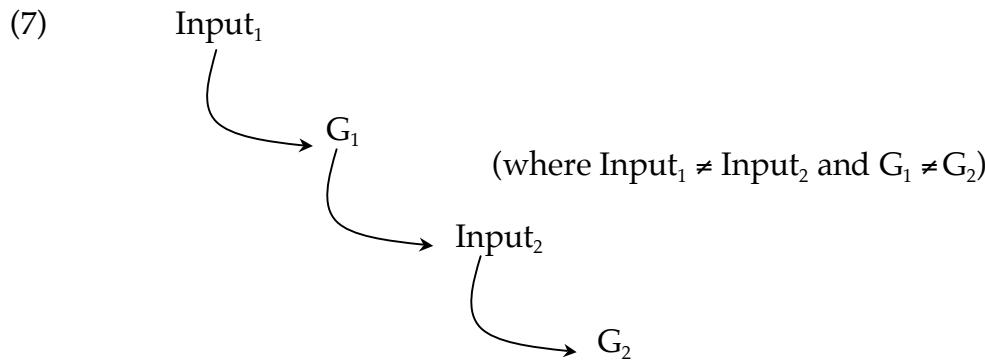
- (6) a. *Innovation* (grammar change, abrupt)
- b. *Diffusion* (a grammar change gradually gaining a wider distribution in a speech community, often perceived as 'language change')

Following Lightfoot (1999) and Hale (2007), I assume that only under the restricted interpretation of language change as grammar change can we hope to develop a restrictive theory that delimits the set of possible changes that can occur in human language. Of course, this raises the question of how the output of the target grammar can trigger a grammar with properties that differ from the properties of the target grammar. This problem is discussed in some more detail in the following section.

2.3 Possible causes of grammar change

As already noted at the outset of this chapter, it is by now widely assumed that there is an intimate connection between the phenomenon of language change and the nature and workings of language acquisition (but see e.g. Janda and Joseph 2003 for some critical discussion). This section explores how the connection between acquisition and change can be made more precise.

It seems likely that the possibility of change is linked to the fact that language transmission is necessarily discontinuous, that is, the fact that each time children engage in the task of first language acquisition, language is created afresh in the mind of each individual. During this process, the learner constructs a grammar based on the linguistic input he/she receives. Under the assumption that the process of language acquisition is highly deterministic (i.e., the same input presented in the same order gives rise to the same grammar, cf. e.g. Lightfoot 1999, Hale 1998, 2007), change can only occur if the acquirer is exposed to a linguistic input that differs in some way from the input that gave rise to the target grammar:



However, this scenario raises *the logical problem of language change*: How can a grammar G_1 produce an output (i.e., Input_2) that differs from the input that led to the acquisition of G_1 (assuming that G_1 was construed to match Input_1)? In particular, we must ask which factors might blur the evidence for certain properties of G_1 in Input_2 , which has been generated by G_1 .

Of course, one likely source of differences between Input_1 and Input_2 are grammar-external factors such as language contact or ‘conscious’ changes adopted by the adult speaker of G_1 , for example the use of linguistic features associated with a prestige dialect, or the avoidance of features that are not part of the prestige dialect (e.g., via prescriptive rules, as in the loss of double negation in the history of English and other standardized European languages). However, as already mentioned above, there are presumably no strong linguistic constraints on contact-induced change. The same goes for changes triggered by sociolinguistic factors. In what follows, I will therefore focus on triggers of change that can be studied by using methods of formal linguistics.

In general, it seems that there is a close link between the possibility of change and the notion that language acquisition is a highly deterministic process. At closer inspection, it is actually quite unlikely that two learners are confronted with exactly the same input. If factors such as the order of data presentation, or the frequency (and thus robustness) of various linguistic forms influence the shape of the emerging grammar, then it seems inevitable that there are at least slight differences between the target grammar and the grammar acquired by the learner (see Hale 2007: 33 on this point).

Let’s now take a closer look at the ways in which Input_2 , which is generated by G_1 , can differ from the data set that gave rise to G_1 (i.e., Input_1), focusing on stylistic changes that change the make-up of the triggering experience, the role of linguistic variation, and ‘noise in the channel’.

First of all, it has repeatedly been pointed out in the literature that stylistically motivated changes in language use may lead to significant changes in the make-up of

the triggering experience (cf. e.g. Lightfoot 1991). As a result, the evidence necessary to trigger a certain property of the grammar may cease to be robustly expressed in the input data. When a certain threshold (e.g., in frequency) is crossed, this may lead to grammar change. Examples discussed in the literature include the rise of VO via massive extraposition (cf. Stockwell 1977 on English, see also chapter 2 below), the rise of ergative/absolutive case marking via a reanalysis of frequently used passive constructions (cf. Anderson 1977), or the loss of V2 due to an increased frequency of subject-initial clauses (cf. Lightfoot 1991, 1997).

In addition, the evidence for a certain property of the target grammar may be obscured by the fact that the child usually receives input from different speakers with possibly different grammars. Note that (7) is actually a gross oversimplification. The input a child receives is usually not generated by a single grammar, but rather by a set of grammars with different properties (as for example is typical in environments where both dialects and the standard language are used). The learner must determine which output string is generated by which grammar, which is a non-trivial task. It is at least conceivable that in such a situation, the child may mistakenly attribute a certain output string to the wrong grammar, which in turn may give rise to a new grammar with properties that differ from those of the target grammar (see Hale 2007: 38f. for discussion). A relevant example comes from Kroch and Taylor's (1997) analysis of the loss of V2 in the Middle English period. Kroch and Taylor attribute the loss of V2 to a mixed dialect situation where speakers of a northern V2 grammar came into contact with speakers of a southern variety in which subject pronouns regularly intervened between a fronted XP and the finite verb. According to Kroch and Taylor, the resulting mixed input (in particular, the violations of the V2 constraint generated by the southern grammars) led to the acquisition of a grammar that also generated V3 patterns, leading to the loss of V2 in the northern variety:

(8) a. Output string generated by southern grammar: XP – pronoun – V_{fin}
 b. Output string generated by northern grammar: XP – V_{fin} – pronoun

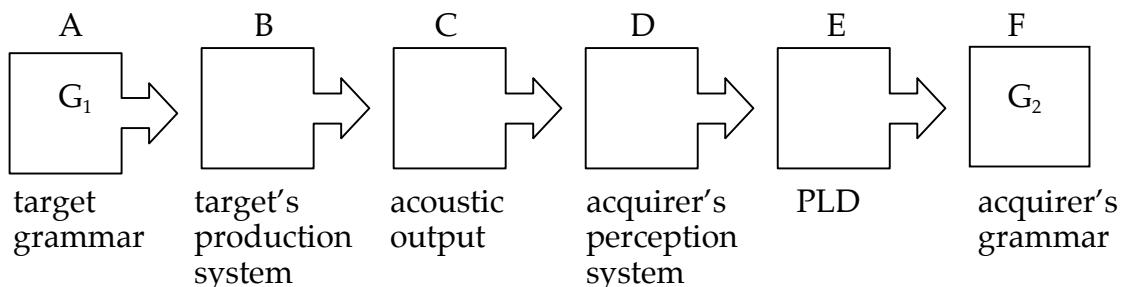
The relevant grammar change would then result from a misanalysis in which learners mistakenly attributed output string (8a) to the northern grammar, which originally was a strict V2 grammar (see chapter 3 for further factors that led to the loss of surface V2 patterns in English).

Perhaps the most common cause of grammar change involves the fact that drawing conclusions about properties of the target grammar from an often messy

and chaotic acoustic input is a highly complex and difficult task, which may involve misanalyses, missed triggers etc. This is what Hale (2007: 53ff.) calls “noise in the channel”.

It is quite obvious that the transmission of the features of the target grammar G_1 to the acquirer's grammar G_2 is neither direct nor instantaneous. Rather, it is mediated by a number of intermediate steps and stages that may hinder a flawless acquisition of properties of the target grammar. In particular, note that the relevant pieces of information are not directly accessible from the input. Rather, they are part of the structures that the learner posits when he/she parses the input he/she is exposed to. As discussed in Hale (1998, 2007), ‘grammar transmission’ actually involves a number of different mapping relationships that may each be affected by errors that obscure the evidence for a certain property of the target grammar (see Hale 2007: 54 for an even more complex picture of the relevant mapping relationships):

(9) Mapping relationships in the transmission from target grammar to acquirer's grammar (Hale 1998: 8)



Box A corresponds to the target grammar G_1 . It is a system of rules, stored information etc. in the mind of the speaker. G_1 generates output representations (structured sets of features) that are uniform for any given sentence. The latter are mapped to B, the target's production system. Thus, looking at spoken languages only, box B represents the articulatory/perceptual performance system that generates the actual acoustic output of G_1 .

As is well-known, the output of B may be highly variable, even for a single individual (due to random factors such as speed of pronunciation, a cold, or the general fact that there is a huge amount of variation in the way phonological units are actually pronounced by a single speaker). The variation inherent in the target realization (i.e., the mapping from A to B) may obscure properties of G_1 . It can be shown (e.g., by methods of instrumental phonetics such as spectrographic analyses) that even a single speaker seldom realizes one and the same linguistic sign (sounds,

morphemes, words, sentences) in exactly the same way. If the range of variation in the phonological realization of a given underlying target structure crosses a certain threshold, this may cause the learner to posit an underlying form that differs from the relevant structure in the target grammar (see Hale 2003 and Ohala 2003 for relevant considerations concerning aspects of phonological change, in particular changes affecting the quality (i.e., fundamental frequency) of vowels).

Box C represents the actual sound waves that float through the air. Again, C may be highly variable, due to environmental conditions, background noise etc.

A random subset of the acoustic output (i.e., those utterances that the learner happens to hear) is then mapped to D, the acquirer's perception system. D processes the incoming sounds and generates structural representations. The messy character of the incoming linguistic data may give rise to mispareses/misanalyses that lead to wrong conclusions concerning properties of the target grammar. If these conclusions fail to be corrected (e.g. by further evidence to the contrary), they will become part of the acquirer's steady state grammar, an instance of grammar change. Typical examples come from grammaticalization processes, where the learner misanalyzes the syntactic category of a certain element (see also Ohala 2003 for mispareses that set off sound change).

It is commonly assumed that not all incoming data is considered as equally relevant for the purposes of grammar construction. Rather, the language acquisition device selects a subset of the incoming data and maps it to the PLD that is considered to be relevant for the acquisition task. This is depicted in box E.

Finally, there is a learning algorithm that maps the PLD to properties of the emerging grammar, as illustrated in F. We may assume that both E and F are universal processes which are not subject to change.

Under the assumption that language acquisition is a deterministic process, it is fairly clear that the complex mapping from G_1 to properties of G_2 introduces quite a number of random factors ("noise in the channel") that may prevent a flawless transmission of features from G_1 to G_2 . Moreover, from the fact that the make-up of the input seems to be different for each individual speaker we may conclude that change is not a rare phenomenon (as suggested by the *logical problem of language change*), but rather a necessary consequence of the way human languages are acquired.

3 Goals of historical linguistics

From the perspective on language change laid out in the previous sections, the central goals and questions of a theoretically informed study of the historical development of languages can be summarized as follows:

(10) A. **Formulation of precise formal analyses of individual changes**

- (i) Precise analyses of two successive historical stages A (prior to change) and A^{+1} (after change) of a language.
- (ii) An explanation of diachronic variation in formal terms that are also used to explain synchronic variation (uniformitarianism).⁸

B. **Formulation of descriptive generalizations**

- (i) Which changes can be observed in the history of languages and which can't?
- (ii) Are there universals of language change that can be detected in many different historical developments in genetically unrelated languages?

C. **Formulation of an explanatory theory of language change**

Based on a conception of the principles that govern language change that is in line with a theory of the formal universals of human language (i.e., Universal Grammar, UG), we should be able to provide principled answers to the following questions:

- (i) What constitutes a possible change (cross-linguistically, and in the case of a particular grammar G with the properties a, b, c)
- (ii) What are the causes of language change?

Note that C (the formulation of a separate theory of language change) can possibly be reduced to independently established concepts, namely a (i) formal theory of (Universal) grammar and a (ii) theory of language acquisition, in the sense that restrictions on language change can be attributed to the way variable properties of grammars are fixed on the basis of the linguistic data children are exposed to in the course of language acquisition (cf. e.g. Lightfoot 2003).

⁸ Most work in historical linguistics adopts the so-called uniformitarian principle, which states that “the forces which operated to produce the historical record are the same as those which can be seen operating today.” (Labov 1978: 281), or somewhat more cautious, a principle which states “that proposals regarding the past are to be seen as independently motivated if they invoke processes known from the present.” (Janda and Joseph 2003: 30; see also Labov 1994: 21).

4 Outline of this work

From the previous discussion, it becomes clear that one central goal of a theory of language change consists in the formulation of constraints on change, that is, defining the ways in which languages can possibly change. Note that it is quite unlikely that these constraints range over the diffusion of a given change in a speech community (which is a sociolinguistic and often unpredictable phenomenon). Rather, the relevant restrictions must concern grammar change, that is, the ways in which an emerging grammar G_2 can possibly deviate from the target grammar G_1 . This work explores how grammar change is shaped and restricted by the process of language acquisition, focusing on phenomena at the interface between syntax and morphology/PF, in particular changes affecting word order, the make-up of inflectional paradigms, and the null-subject property in the history of various Germanic languages. I am going to argue that the course of the relevant historical developments is shaped by a set of universal, 'hard-wired' properties of the grammar/UG and learning strategies that are part of the language acquisition device. Obviously, the formulation of constraints on possible (grammar) changes should be based on a restrictive theory of grammar. This work presupposes familiarity with Principles and Parameters theory and core properties of minimalist syntax (cf. Chomsky 1995, 2000; see Grewendorf 2002 and Radford 2004 for overviews).

Chapter 2 is concerned with restrictions on grammar change imposed by the nature of the syntax-morphology interface. In particular, it will become clear that the (hard-wired) workings of the mapping from syntax to PF restrict the set of possible hypotheses about the underlying syntactic structures that the learner must entertain when he/she is confronted with a certain surface string of phonological exponents. To see this, I examine the syntax-morphology/PF interface in some detail. Adopting basic assumptions of Distributed Morphology (Halle and Marantz 1993), the chapter develops a detailed analysis of the interaction of syntax and the post-syntactic phonological component, focusing on the question how the hierarchical structures assembled in the syntax are mapped to linear orderings by the workings of Vocabulary Insertion, the process in which syntactic terminal nodes are associated with phonological exponents. In particular, I take a closer look at the assumption of cyclic Spell-Out and its implications for the realization of linear order, proposing that the individual chunks of structure (corresponding to syntactic Spell-Out domains) that are transferred to PF must be reassembled in the phonological component. By

assumption, PF first creates *phonological domains* from the cyclic output of syntax, which are isomorphic to syntactic phases. The edges of successive phonological domains are taken to overlap, which facilitates establishing a linear ordering between neighboring phonological domains. The specifics of the procedure that linearizes phonological domains impose a set of restrictions on possible orderings. These restrictions explain not only the cross-linguistic absence of certain word orders (such as *VO-Aux), but also give rise to a typology of possible and impossible grammars, which can be exploited for the analysis of word order change. Relevant empirical phenomena discussed include the distribution of finite complement clauses in German and the analysis of word order variation and change (OV-VO) in the history of English.

Chapter 3 takes a closer view at a particular syntactic phenomenon, focusing on the history of the V2 property in Germanic. We will see that “V2” is not a unitary phenomenon and that there are a couple of different structural configurations that may give rise to surface V2 orders in early Germanic: (i) The historical core of the V2 phenomenon seems to be a configuration called ‘operator V2’ in which verb fronting takes place in contexts such as interrogatives, imperatives, and neg-fronting. (ii) In addition, we can observe systematic verb fronting in clauses introduced by certain temporal adverbs roughly meaning ‘then’ in all early Germanic languages. (iii) In Old English, surface V2 orders may result from a configuration in which there is no spec-head-relation between the finite verb and the fronted XP. Rather, the two elements are merely linearly adjacent (so-called ‘pseudo V2’). (iv) Early forms of ‘generalized V2’ are found in Old High German, where the verb occupies C in all main clauses, accompanied by moving a single XP to clause-initial position. I am going to argue that the loss of surface V2 patterns in Middle English and the development of generalized V2 in the course of Old High German are linked to the loss of discourse-configurationality (i.e., the loss of discourse-related triggers for movement processes) and an acquisition strategy that enables learners to cope with dislocation phenomena the original semantic/pragmatic trigger of which is not any longer transparent (via positing semantically vacuous EPP-features in order to mimic word order patterns for which no substantial trigger can be detected). I will pay special attention to the development of generalized V2 in the course of Old High German, which is analyzed as resulting from the loss of a parametric option licensing multiple specifiers in the C-domain.

Chapter 4 is concerned with the cyclic course of morphological change, that is, the observation that the loss of distinctions via phonological erosion and analogical

leveling of paradigms is often compensated for by grammaticalization processes that provide new and more distinctive phonological exponents of underlying inflectional categories. I am going to argue that the cyclic nature of these changes is determined by (apparently) conflicting acquisition strategies employed by the learner when he/she acquires the elements of the lexicon (i.e., individual Vocabulary items) and the feature specifications they are associated with. One such strategy is based on the notion of morphological blocking, favoring the acquisition of more distinctive Vocabulary items. In addition, there is a conflicting strategy that aims at minimizing the number of elements/features stored in the lexicon. While the former typically gives rise to grammaticalization processes, the latter is a driving force in what is traditionally called analogical leveling, in which a less distinctive form gains a wider distribution in a paradigm. I show that processes of blocking and deblocking of exponents are also at play in another phenomenon at the interface between syntax and morphology, namely the historical development of null arguments. More precisely, I propose that a null realization of weak pronouns may emerge when competing overt exponents are lost (deblocking), for example due to a reanalysis as agreement markers. In addition, we will see that null subjects may disappear when a language develops new overt realizations of weak pronouns that realize more morphosyntactic features than the null-spell-out (blocking).

Chapter 5 provides a concluding summary.

Chapter 2: The syntax-morphology interface

1 Introduction

Any formal study of word order change and other diachronic phenomena that concern morphosyntactic properties of grammar (e.g., the inventory, shape, and specification of inflectional markers) should be based on a restrictive theory of the mapping from syntax to morphology/PF. Assuming that basic aspects of the mechanisms that accomplish this mapping are uniform across languages (i.e., hard-wired properties of grammar) eases the burden on the child during language acquisition via reducing the number of hypotheses about possible underlying structures that must be entertained. Since it is the task of the learner to detect underlying properties of syntactic structure (hierarchical relations, feature content of functional heads etc.) by inspecting properties of the string of words he/she is confronted with, a restrictive theory of the mapping from syntax to PF delimits the set of possible misanalyses and therefore reduces the number of possible deviations from the target grammar (i.e., grammar change). Furthermore, any such theory can inform the work of the historical linguist via reducing the number of possible changes and historical pathways that must be taken under consideration, in particular if we accept the notion that the proper object of study of historical linguistics is grammar change, that is, discrete differences between the target grammar and the grammar acquired by the learner.

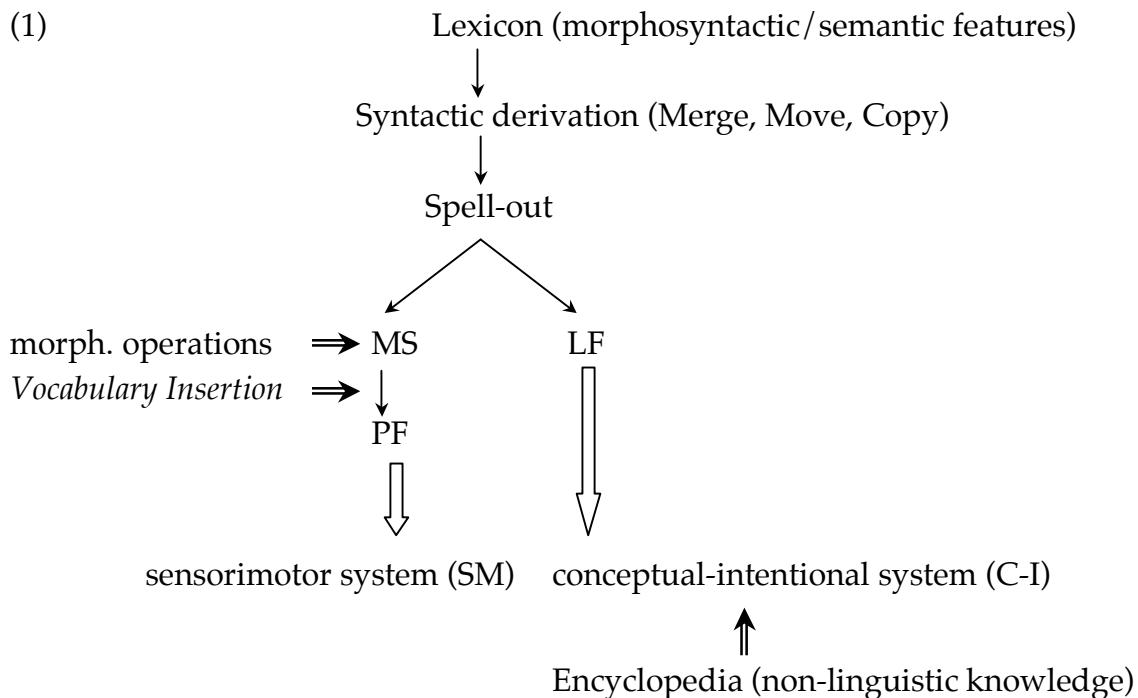
This chapter sets out to develop a restrictive theory of the mapping from syntax to PF that provides the theoretical foundation for the analyses of particular diachronic phenomena that are carried out later in this work.

The chapter is organized as follows. Section 2 gives an overview of Distributed Morphology, the architecture of grammar (and the morphological framework)

adopted here. Section 3 focuses on issues surrounding the idea that the structure(s) generated by the syntactic computational system are spelled out not as a single representation, but rather in a piecemeal fashion (“cyclic Spell-Out”, Epstein et al. 1998, Uriagereka 1999, Chomsky 2000, 2001, 2004, 2005). I propose that the phonological component reassembles the cyclic output of the syntax to create larger, overlapping domains for the application of phonological operations, including the insertion of phonological exponents and the creation of linear orderings. Section 4 is devoted to the question of how the morpho-phonological component of grammar converts hierarchical structures generated in the syntax into linear sequences of words. I develop a model of cyclic linearization which assumes that the phonological representation of a syntactic structure S is built incrementally via successive applications of the operation Vocabulary Insertion (following cyclic Transfer, Chomsky 2004), each time adding phonological material to the linear string of word/segments assembled by previous applications of Vocabulary Insertion. In addition, I address the question of how individual Spell-Out domains are linearized relative to each other. The relevant proposal is based on the notion of overlapping phonological domains developed in section 3. I am going to argue that the linear order between neighboring phonological domains is established via a process called *Edge Replacement* that substitutes the right edge of the higher domain with the string of exponents inserted to the lower domain. Section 5 illustrates the workings of this approach to linearization, focusing on the distribution of finite complement clauses in German. I demonstrate that the assumptions laid out in section 4 make available a new account of the fact that finite complement clauses obligatorily appear in postverbal position in many OV languages, in contrast to other types of complements. Section 6 takes a closer look at the typological and diachronic implications of the model developed so far, focusing on the restrictions that this theory of linearization imposes on the set of licit combinations of parametric choices. I argue that the proposed approach to linearization provides new explanations for cross-linguistic generalizations on possible linear orderings such as the absence of *VO-Aux orders, or the connection between the position of complementizers and preverbal/postverbal placement of complement clauses. In addition, I examine the correlation between word order variation and language change in the transition from OV to VO in the history of English, arguing that the theory of linearization developed in this chapter makes available a new analysis of Old English (OE) word order facts and the relevant changes that took place in the Middle English period. Section 7 provides a concluding summary of the findings reached in this chapter.

2 A realizational model of grammar

This work adopts a *realizational* model of the interplay between syntax and morphology/phonology in which word building operations are distributed over several components of grammar (Distributed Morphology, henceforth DM; Halle and Marantz 1993, 1994). DM assumes that the morphological component operates post-syntactically, interpreting the output of the syntactic derivation. In other words, the morphological component mediates between the syntactic and the phonological modules of grammar. Accordingly, the architecture of the grammar looks as in (1).



In this model of grammar, the syntactic operations Merge and Move manipulate bundles of morphosyntactic features that correspond to syntactic terminal nodes (i.e., heads in traditional terminology). The syntactic terminal nodes are commonly referred to as *morphemes*.¹ It is a central assumption of DM that all internally complex forms are assembled by one and the same computational system. Thus, the structure of words reduces to syntactic structure, in the sense that word-internal relationships among morphemes are structurally identical to the relationships that hold among words. This hypothesis is sometimes referred to as “syntactic hierarchical structure

¹ In some works, the notion *abstract morpheme* is used to refer to syntactic terminals in order to avoid confusion with the traditional usage of the term ‘morpheme’ as part-of-speech. Note that throughout this work, the notion *morpheme* is used in the technical sense introduced in the main text, that is, as referring to syntactic heads/terminal nodes.

all the way down" (cf. e.g. Harley and Noyer 1999). In the phonological component,² (i) the constituent structure derived in the syntax may be modified by a set of operations (readjustment, Morphological Merger, Fusion, Fission etc.). Subsequently, the terminal nodes are associated with phonological exponents in a process called *Vocabulary Insertion*. It is commonly assumed that up to the process of Vocabulary Insertion, post-syntactic operations have access to (and are restricted by) the hierarchical structure assembled in the syntax. The idea that phonological content is added after syntax is also known as *Late Insertion*.³ Thus, DM is an 'anti-lexicalist' model in the sense that no complete lexical elements (fully specified for semantic, morphosyntactic and phonological features) are inserted into the syntactic derivation (cf. Marantz 1997 for discussion). Instead, the information that is traditionally assumed to be part of a single lexical entry is divided into separate lexical entries that are part of distinct subparts of the lexicon: one part of the lexicon contains abstract morphemes, that is, only bundles of semantic and morphosyntactic features which are subject to 'early', that is pre-syntactic insertion (such as e.g. $C_{[+wh]}$, $T_{[+past]}$ etc.). Another part contains the set of *Vocabulary items* which link phonological exponents with morphosyntactic features (i.e., insertion contexts). Finally, there is the so-called *Encyclopedia* which contains the non-linguistic (idiomatic) information associated with lexical items (e.g. *dog*: 'four legs, canine, pet, sometimes bites etc.').

The form of Vocabulary items is illustrated in (2), which captures the English verbal inflection 3sg.pres.indic. /-z/. (2) can be read as an insertion rule: "the phonological exponent /-z/ is inserted in the context [3, sg, pres., indic.]":

² In what follows, I will use the terms "Morphological Structure", "phonological component" and "PF" interchangeably, as if they were synonymous, to refer to the PF branch of grammar, that is, that part of the computation that processes the (cyclic) output of the syntactic component for the purposes of the interface to the sensorimotor systems (SM), possibly deriving an interface representation traditionally called Phonological Form (PF).

³ Most current morphological theories accept the notion that the phonological form of a given affix should be separated from its morphosyntactic function. Thus, they adopt (in some form or other) the *Separation Hypothesis* (Beard 1988, 1995; cf. Stump 1998, Borer 1998 for some discussion), according to which grammatical operations (which manipulate the constituency and feature content of lexical items) are discrete from the morphological or phonological operations that control the spell-out of words. Accordingly, it is widely held that affixes or other forms of morphological marking (such as *Umlaut* or the modification of tone or accent) are merely the *exponents* (Matthews 1991) of abstract morphological features such as [future], [plural], [1st person]. Note that this distinction follows automatically from the overall structure of DM.

(2) [3, sg, pres., indic.] \leftrightarrow /-z/

In what follows, I will review some basic concepts of DM and some recent refinements of the theory that will be made use of in the diachronic analyses to come.

2.1 Blocking effects and the workings of Vocabulary Insertion

The procedure inserting phonological exponents of morphosyntactic feature bundles requires that the feature specification of the Vocabulary item is nondistinct from the features of the insertion site (i.e., a certain morpheme). Usually, this requirement is met by several items, which then enter into a competition. The item that realizes the greatest subset of features is chosen for insertion. In the case of English verbal inflection, the availability of the Vocabulary item in (2) blocks the insertion of the less specified exponent /-Ø/, which is found in all other contexts. The latter is commonly referred to as the ‘elsewhere’ case (cf. Kiparsky 1973, 1982; Aronoff 1976, Anderson 1986, among many others, for the workings of the *Elsewhere Condition* in phonology/morphology). Thus, the paradigm in (3) is captured most economically by positing merely two Vocabulary items (listed in (4)) for the realization of present tense INFL in English.

(3) a. I/you/we/you-PL/they sleep.
 b. He/she/it sleep-s.
 c. *He/she/it sleep.

(4) a. [3, sg, pres., indic.] \leftrightarrow /-z/
 b. elsewhere \leftrightarrow -Ø

This approach implies that Vocabulary items may be *underspecified* for the feature complexes they realize. In (4), the exponent -Ø is maximally underspecified since it is not linked to any feature specification at all (the ‘elsewhere’ case). The basic properties of the insertion procedure are captured by Halle’s (1997: 428) *Subset Principle* (not to be confused with the Subset Principle of Wexler and Manzini 1987):⁴

⁴ Note that the Subset Principle alone is not sufficient to determine the winner of the competition if two Vocabulary items realize the same number of inflectional features. For such cases, two different solutions are proposed in the literature. The winner is either simply stipulated by an extrinsic rule ordering (Halle and Marantz 1993) or follows from a universal hierarchy of morphosyntactic features

(5) *The Subset Principle*

The phonological exponent of a Vocabulary item is inserted into a morpheme in the terminal string if the item matches all of a subset of the grammatical features specified in the terminal morpheme. Insertion does not take place if the Vocabulary item contains features not present in the morpheme. Where several Vocabulary items meet the conditions for insertion, the item matching the greatest number of features specified in the terminal morpheme must be chosen.

I assume that the operation of Vocabulary Insertion leads to discharge of the morphosyntactic features present in a syntactic terminal node (cf. e.g. Noyer 1997, Trommer 2001, 2003):

(6) *Vocabulary Insertion* (Trommer 2003)

If M is a VI [Vocabulary item] with syntactic features α and phonological features β , and S is a head with features γ , where α is a subset of γ , then delete the features α in γ and add β to the phonological representation associated with S .

As a result, features that have been discharged by the application of Vocabulary Insertion can not trigger any further insertion processes. Furthermore, (6) implies that Vocabulary Insertion can apply repeatedly to a given syntactic terminal node, as long as there are still features present that can be discharged (giving rise to the same effects as Fission, see below).

Moreover, following Embick and Noyer (2001: 562), I assume that the assignment of precedence relations is another function of the operation of Vocabulary Insertion:

(7) *The Late Linearization Hypothesis*

The elements of a phrase-marker are linearized at Vocabulary Insertion.

In other words, the linearization of hierarchical syntactic structures is taken to be a late process at the PF branch of grammar (applying after other operations such as Morphological Merger, Fission etc.) which converts the output of the syntactic

(cf. e.g. Noyer 1997; see Harley 1994, Harley and Ritter 2002 for the use of structured feature geometries instead of simple hierarchies).

computation into a linear representation that can be read off at the serial interface to the sensorimotor (SM) systems. The fact that the hierarchical structure generated in the syntax may be modified by operations of the PF-branch implies that linear order is not fully determined by the syntactic structure, contra *Kayne (1994)*. We will return to this point in more detail in section 4.

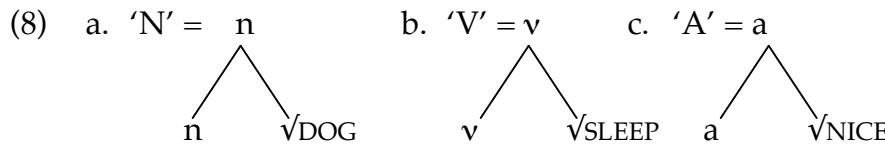
2.2 L-morphemes, f-morphemes, and the syntactic determination of lexical categories

Most work in DM acknowledges a systematic difference between lexical and functional categories (e.g. *Halle 1990*: ‘concrete’ vs. ‘abstract’ morphemes; *Harley and Noyer 1999*: ‘l-morphemes’ vs. ‘f-morphemes’). Adopting the terminology of *Harley and Noyer (1999)*, f-morphemes constitute a closed class and contain features relevant for the syntactic computation (e.g. [plural], [+past], [+wh]). In addition, they presumably carry semantic features associated with a certain set of logical meanings (cf. *von Fintel 1995*, *Roberts and Roussou 2003*). A property specific to f-morphemes is that their spell-out is deterministic, that is, the feature content of a given f-morpheme serves to determine a unique phonological realization. Thus, the insertion of Vocabulary items realizing f-morphemes is guided by the Subset Principle which ensures that the most specific candidate wins out over its competitors.

In contrast to f-morphemes, l-morphemes constitute an open class. They do not carry grammatical features driving the syntactic computation, but they may be specified for syntactically relevant semantic features such as [+animate], [+count] etc. Importantly, the spell-out of l-morphemes is not constrained by a competition between compatible Vocabulary items.

Current work in DM (cf. e.g. *Marantz 1995, 1997, Embick 1997, Harley and Noyer 1999*) assumes that the conventional lexical categories such as noun, verb, or adjective are not syntactic (or, morphological) primitives. Instead, lexical categories are decomposed into more basic combinations of l-morphemes and f-morphemes. L-morphemes are taken to be category-neutral roots that pair sequences of complexes of phonological features with certain semantic features. In other words, “roots are language-specific combinations of sound and meaning” (*Embick & Noyer 2007: 295*). The categorial properties formerly associated with the labels N, V, or A are determined syntactically via the structural configuration in which roots occur. Under this view, f-morphemes are taken to assume a category-defining role. For example, a ‘verb’ corresponds to a root that is merged with (and locally c-commanded by) a v-

head.⁵ Marantz (1995) extends this approach to all lexical categories, arguing that (lexical) categorial information is added by merging roots with a closed class of category-defining syntactic heads (n, v, and a; see also Embick 2000, Harley 2007, Embick and Noyer 2007).⁶ Adopting the notation “√” to indicate roots (following Pesetsky 1995) and words in small caps as shortcuts for the relevant sound-meaning pairings, this can be illustrated as follows:



Importantly, one and the same root may show up as a ‘verb’ or a ‘noun’ depending on the syntactic context where it is merged. For example, the Vocabulary item *destroy* surfaces as a ‘noun’ *destruction* when it is locally c-commanded by n, while it becomes a verbal element (*destroy* or a participle *destroyed/destroying*) when the closest category-defining element is a v-head.⁷

2.3 Post-syntactic morphological operations

Prior to Vocabulary Insertion, a set of morphological operations may apply to the output of the syntactic component, modifying the constituent structure and the content of morphemes. The most important of these are the insertion of so-called *dissociated morphemes*, (morphological) *Merger* or *Fusion*, *Fission* and *Impoverishment*. In the following, these mechanisms are only briefly introduced; they are discussed in more detail when they have a concrete bearing on the issues dealt with in this work.

The hierarchical structure derived in the syntax can be modified by the post-syntactic insertion of *dissociated morphemes* which may attach to other functional morphemes. Following Embick (1997), these morphemes are called ‘dissociated’, since they are not present in the syntactic derivation and merely reflect properties

⁵ Embick (1997) distinguishes between verbs and participles by assuming that a ‘verb’ is a root locally c-commanded by v, Aspect and Tense, while a ‘participle’ lacks a c-commanding T node.

⁶ See e.g. Harley and Noyer (1999) for an alternative proposal where a nominal element (a noun or a nominalization) corresponds to a root that is locally c-commanded and licensed by a D-head.

⁷ In Fuß (2005: 46), I argue that the systematic differences between l-morphemes and f-morphemes can be employed to explain certain characteristics of grammaticalization processes which are usually assumed to involve a transition from lexical to functional categories, that is, from Vocabulary items realizing l-morphemes to Vocabulary items realizing f-morphemes in the approach outlined here.

expressed by structural configurations in the syntax proper. Within DM, this mechanism is often used to account for case and agreement phenomena. For example, Marantz (1992), Halle and Marantz (1993) and Halle (1997) analyze subject-verb agreement in terms of the post-syntactic adjunction of an [Agr] morpheme to T (cf. Embick 1997 for a detailed discussion of the insertion of dissociated morphemes). See Fuß (2005), (2008), and section 3 below for an analysis of complementizer agreement in Germanic in terms of dissociated Agr-morphemes.⁸

Another morphological operation that may change the constituent structure generated by the syntax is *Morphological Merger* of terminal nodes that do not form a constituent in the syntactic output (cf. Marantz 1984, 1988). For our purposes, the most important instance of Merger are cases of apparent syntactic lowering, that is, syntactic heads that are not joined together via head movement, but still are spelled out as a unit. A prominent example of this type of Merger is affix-hopping in English, that is, the post-syntactic affixation of Tense/ Agr to the main verb that on standard assumptions fails to undergo overt head movement in English (cf. Chomsky 1957, Lasnik 2000, Bobaljik 2002). Another set of phenomena which is frequently analyzed in terms of post-syntactic readjustment is the positioning of clitics, for example second position clitics of the Wackernagel type (see e.g. Schütze 1994 on second position clitics in Serbo-Croatian). Embick and Noyer (2001) distinguish between local PF readjustment rules that apply before and after Vocabulary Insertion. The former operate in terms of hierarchical structure (Morphological Merger), while the latter operate in terms of linear order, switching the position of two linearly adjacent elements (so-called *Local Dislocation*).

A related operation is *Fusion* which may create a mismatch between the number of underlying morphemes and the number of inserted Vocabulary items, in the sense that two (or more) syntactic nodes are fused into a single terminal node which is then realized by a single phonological exponent. Again, the verbal inflection of English is a case in point: it is commonly assumed that AGR and T fuse into a single morpheme prior to Vocabulary Insertion (Halle and Marantz 1993, Halle 1997). Instances of Fusion are traditionally referred as *cumulative exponence* and may involve quite a number of different inflectional features as in the verbal inflections of Latin (cf. Matthews 1991: 233): in a 1st person singular present indicative active form like *am-ō*

⁸ In Fuß (2005) I argue that dissociated Agr-morphemes play an important role in the historical rise of new agreement markers as an intermediate step on the grammaticalization path from clitics to syntactic agreement markers.

'I love', the single exponent $-\bar{o}$ serves to simultaneously realize five inflectional features, namely person, number, tense, mood, and voice.

Similar to Fusion, *Fission* (Halle and Marantz 1993, Halle 1997, Noyer 1997) also disturbs the isomorphism between syntax and morphology, albeit with contrary effects: while Fusion leads to the amalgamation of two separate syntactic terminals, Fission results in a situation where a single syntactic terminal node is realized by more than one Vocabulary item. The concept of Fission is related to the notion that the insertion of Vocabulary items discharges the inflectional features present in the morpheme (Noyer 1997, Trommer 2001, 2003; see (6) above). In standard cases, the insertion procedure stops after a phonological exponent is inserted, even if this exponent discharges only a subset of the inflectional features present in the morpheme. However, when a morpheme is marked for undergoing Fission, the inflectional features that are not discharged (or 'matched', Halle 1997) by the first insertion operation are copied into an additional morpheme which is generated by the insertion procedure. Subsequently, this additional morpheme is then itself subject to Vocabulary Insertion. Typical examples of Fission come from languages where agreement is marked by a combination of prefixes and suffixes, as in many Afro-Asiatic languages (cf. Noyer 1997 for extensive discussion).⁹ This can be illustrated by the following example from Tamazight Berber, where the agreement morpheme "splits into three positions of exponence" (Noyer 1997: 89) which are realized by successive Fission (of a single Agr-morpheme) and insertion of the Vocabulary items in (10):

(9) *t-dawa-n-t*
 2-cure-PL-FEM
 'you (pl, fem) cured'
 (Noyer 1997: 89)

(10) a. [2] \leftrightarrow /t-/
 b. [pl] \leftrightarrow /-n/
 c. [fem] \leftrightarrow /-t/

⁹ See Arregi (1999) for an analysis of the person and number inflection of Basque and Halle (1997) for a analysis of the Latin noun declension in terms of Fission.

Due to the fact that morphemes created by Fission contain only a subset of the features contained in the original morpheme, Fission often results in the insertion of less marked or ‘elsewhere’ Vocabulary items, that is, “Fission extends the domain of less marked exponents” (Halle 1997: 432). See chapter 4 for an analysis of changes that affected the set of agreement markers in Alemannic which makes use of the concept of Fission.

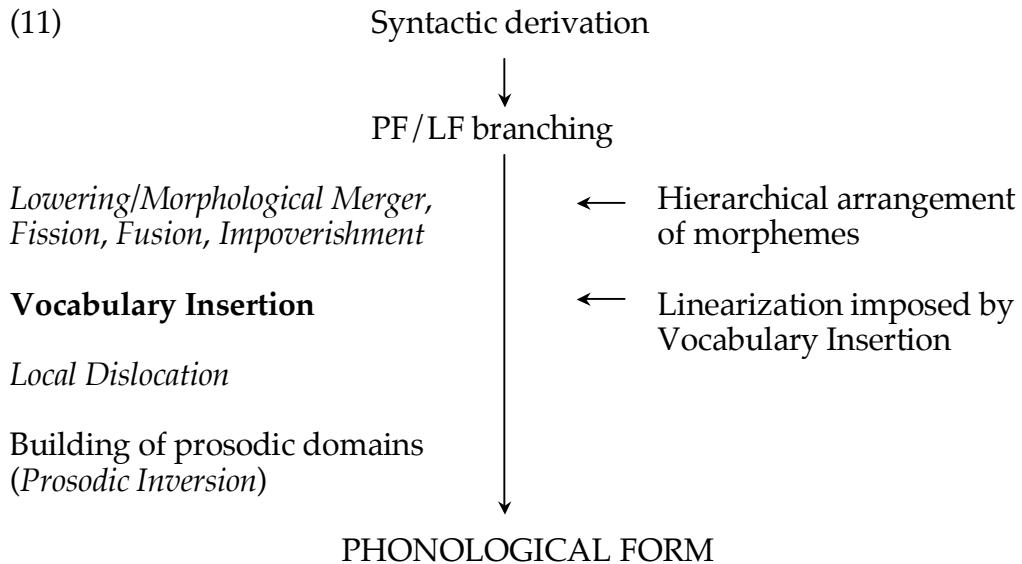
The insertion of less marked exponents can also follow from context-sensitive *Impoverishment* rules which delete morphosyntactic features from morphemes prior to Vocabulary Insertion (cf. Bonet 1991, Halle 1997, Noyer 1997). As a result, Vocabulary items that require the presence of those features cannot be inserted and a less specified exponent must be used to realize the morpheme affected by Impoverishment (which Halle and Marantz 1994 call the “Retreat to the General Case”). Thus, Impoverishment typically gives rise to systematic syncretism, in which – in a certain morphosyntactic context or generally (e.g., syncretism of 1pl and 3pl in the verbal inflection of German, cf. Müller 2006a, 2006b) – different cells of a paradigm fall together.¹⁰

The notion of Impoverishment can be put to use to describe instances of morphological change in which a less specified Vocabulary item gains a wider distribution in a paradigm, leading to (systematic) syncretism (cf. Noyer 1997: lxxx-lxxxi; see chapter 5 below for further discussion). Instead of invoking the workings of analogy, such processes of paradigm leveling can be attributed to the extension (or, development) of Impoverishment rules that delete features of syntactic terminal nodes and thus give rise to an “expansion of the domain of the unmarked exponent” (Halle 1997: 431).

At some point of the post-syntactic computation, the operation of Vocabulary Insertion applies, inserting phonological information into syntactic terminal nodes and converting hierarchic structures into a string of phonological exponents/words. Operations applying after Vocabulary Insertion include *Local Dislocation* and *Prosodic Inversion* (Halpern 1992) which may serve to switch the position of linearly adjacent

¹⁰ In ‘Minimalist DM’ (Trommer 2001, 2003) it is assumed that there is in fact only a single post-syntactic operation. More precisely, it is claimed that the effects of the different operations postulated in standard DM can all be derived from the workings of Vocabulary Insertion as defined in (6). For example, Fission is then reanalyzed as multiple insertion of Vocabulary items into a single syntactic terminal node (being only restricted by feature deletion via discharge), while Impoverishment effects are taken to result from the insertion of zero exponents that discharge features but do not contain any phonological features.

exponents (e.g., in the case of clitics, cf. Embick and Noyer 2001, 2007 for discussion) and the building of prosodic domains. Crucially, these operations do not any longer have access to the hierarchical syntactic structure after it has been linearized by the workings of Vocabulary Insertion. The structure of the PF-branch of grammar can thus be schematized as in (11) (cf. Embick and Noyer 2001: 566).



This section has served to establish basic properties of the grammar model adopted in this work. In the following sections, I will discuss a couple of more specific issues that arise under this model, focusing on the workings of the operation Spell-out/TRANSFER in a phase-based approach to syntax, and the interaction of syntax and morphology in the linearization of hierarchic syntactic structures.

3 Cyclic Spell-Out and the domain of post-syntactic operations

In current investigations into the nature of the mapping between syntax and the interpretative components of grammar, it is widely assumed that the structure(s) generated by the syntactic computation are spelled out not as a single representation, but rather in a piecemeal fashion (cf. e.g. Epstein et al. 1998, Uriagereka 1999, Chomsky 2000, 2001, 2004, 2005). That is, in contrast to previous grammar models, there is no single designated point at which the output of the syntactic derivation is handed over to the phonological and the semantic components. Instead, Spell-Out applies repeatedly during the syntactic derivation, each time transferring a subpart of the phrasemarker constructed so far to the post-syntactic components of grammar. The latter are assumed to operate in a strictly cyclic fashion as well, in the sense that a syntactic object transferred to the interpretative components is directly mapped to

the interfaces to the (language-external) sensorimotor (SM) and conceptual-intentional (C-I) systems (cf. Chomsky 2004, 2005). Thus, the language-external systems do not interpret complete representations derived from the output of narrow syntax, but rather structural chunks which correspond to the individual derivational cycles of narrow syntax. According to this model, then, post-syntactic operations cannot access pieces of information which are part of different Spell-Out domains.

Focusing on the phonological/morphological module of grammar, this section argues that this restriction on the workings of the post-syntactic components is too strong. More specifically, I show that post-syntactic operations may cut across the Spell-Out domains defined in Chomsky (2000, 2001, 2004, 2005), paying special attention to the phenomenon of complementizer agreement in Germanic. Based on the observation that this form of multiple agreement is subject to restrictions that exhibit an unmistakable phonological character (adjacency effects, sensitivity to PF processes such as sluicing or elision of the finite verb), I claim that complementizer agreement is established by a post-syntactic operation that copies ϕ -features from T to C (cf. Fuß 2005, 2008). Under the assumption that Spell-Out affects the complement of a (strong) phase head, handing TP and VP to the interfaces,¹¹ this morphological copy operation requires access to pieces of information that are distributed over different Spell-Out domains. We must therefore allow for an extension of the scope of PF processes, either via a redefinition of Spell-Out domains (with considerable consequences for the syntactic computation), or by assuming that phonological domains in fact differ from Spell-Out domains, keeping to the theory of phases devised by Chomsky. This section develops a proposal in the latter direction, arguing that the phonological component constructs from the cyclic output of narrow syntax larger units which consist of a Spell-Out domain Σ_n and the right edge of a subsequent Spell-Out domain Σ_{n+1} (see Dobashi 2003 for a related proposal).

3.1 Phases, Spell-Out domains, and the scope of PF operations

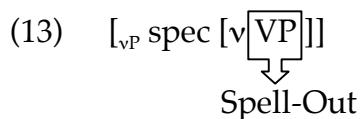
According to recent work by Chomsky, Spell-Out domains are associated with derivational cycles (of narrow syntax) which are referred to as *phases*. Chomsky identifies these phases as CP and (transitive/agentive) vP and assumes that the domain (i.e., the complement) of a phase head is transferred to the phonological

¹¹ For expository reasons, I will use the more traditional notions “VP” and “V” (instead of “ \sqrt{P} ” and “ \sqrt{V} ”) to refer to the complement of v and its head.

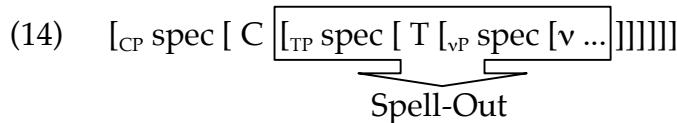
component when the phase is completed (cf. e.g. 2001: 13, 2004: 108).¹² This yields a strong version of cyclicity, since after Spell-Out, only the specifier(s) of a lower phase head and the phase head itself remain accessible to further syntactic operations (the so-called *Phase Impenetrability Condition*, PIC):

(12) At the phase ZP containing phase HP, the domain of H is not accessible to operations, but only the edge of HP. (Chomsky 2004: 108)

For example, the VP complement of a phase head v is sent to the post-syntactic components when vP has been completed. As a result, VP and everything contained in VP are no longer accessible to the ongoing syntactic computation:



Likewise, the TP complement of the phase head C is spelled out once CP is created (note that it must be possible to spell out root CP in full, presumably together with its TP complement, cf. Chomsky 2004: 108):



This set of assumptions introduces an asymmetry between the notions of *phase* and *Spell-Out domain*: while Spell-Out is associated with the phase level, the actual phrasal units sent to the interpretative components (i.e., the Spell-Out domains) correspond to VP and TP (with the edge of vP spelled out together with TP and the

¹² Still, Chomsky assumes that T can access a quirky nominative object in the domain of vP (cf. 2001: 13, Chomsky 2004: 108). At first sight, this seems to imply that in this case, Spell-Out of the domain of a lower phase head (here: VP) is actually delayed until a higher phase head (C , in the case at hand) is merged. However, Chomsky seems to stick to the notion that “the sister of H can be spelled out at HP” (see 2004: 125, n. 19), assuming that T may probe into VP (which has already been subject to Spell-Out at vP) if this operation does not have any visible effects on the (already spelled out) nominative object (i.e., it may neither raise nor undergo any phonetic change, *ibid.*). Furthermore, the assumption that the complement of the head of phase PH is spelled out when PH is completed is required for independent reasons in the model outlined in Chomsky (2004, 2005), where it is assumed that T inherits its feature content from C . As a result, operations triggered by C and T take place simultaneously (so-called “parallel probing”). Under these assumptions, “late” spell out of VP (i.e., when C is merged) would presumably facilitate violations of the PIC, since besides T , C should also be able to probe into VP.

edge of embedded CP spelled out together with matrix VP). The asymmetry between phases and Spell-Out domains raises an issue with respect to the original characterization of phases in terms of interface properties, cf. Chomsky (2004: 124):

“Ideally, phases should have a natural characterization in terms of IC [interface conditions]: they should be semantically and phonologically coherent and independent. At SEM, vP and CP (but not TP) are propositional constructions: vP has full argument structure, and CP is the minimal construction that includes Tense and event structure and (at the matrix, at least) force.”

However, if it is not the phase itself that is transferred to the SM and C-I interfaces, but merely the complement of the phase head, then the (syntactic) object that reaches the interfaces is actually *not* a “semantically and phonologically coherent and independent” unit. In other words, we may ask how interface conditions can identify CP and vP as phases if the units that are actually interpreted at the interfaces are significantly smaller, corresponding to TP and VP (see also Boeckx and Grohmann 2007, Epstein 2007; see section 3.4 below for more discussion and an answer based on the assumption that phasal units are restored in the phonological component).

A general question raised by the assumption of cyclic Spell-Out concerns the way the post-syntactic components deal with the parts and pieces handed to them in the course of the derivation. In particular, we may ask how the post-syntactic components map the syntactic output to interface representations usable by the language external sensorimotor (SM) and conceptual-intentional (C-I) systems “that enter into thought and action” (Chomsky 2004: 106). According to Chomsky (2004, 2005), the idea of cyclic Spell-Out entails that the interfaces are accessed in a cyclic fashion as well.¹³ More precisely, he assumes that the output of each derivational cycle is separately transferred to the interfaces. It follows that the post-syntactic components do not construct a single unified interface representation from the individual Spell-Out domains created by narrow syntax, cf. the following quote taken from Chomsky (2005: 8f.):

“[...] the final internal level LF is eliminated [...] at various stages of computation there are Transfer operations: one hands the SO already constructed to the

¹³ Put differently, we may ask whether it is possible for the language external systems to interpret the output of the syntactic computation in a piecemeal, phase-by-phase fashion. See Stechow (2005) for some discussion of this question with respect to the C-I interface.

phonological component, which maps it to the SM interface (“Spell-Out”); the other hands SO [syntactic object] to the semantic component, which maps it to the C-I interface.”

The assumption of cyclic Spell-Out implies that the operations carried out in the post-syntactic components of grammar are equally constrained by the cyclic nature of the syntactic derivation. Accordingly, processes such as prosodic phrasing, stress assignment, or the creation of linear order proceed in parallel with the syntactic computation and are thus assumed to be subject to the same kind of strict cyclicity as narrow syntax (i.e., the PIC):

“Φ [the phonological component] proceeds in parallel with the NS [Narrow Syntax] derivation. Φ is greatly simplified if it can “forget about” what has been transferred to it at earlier phases; otherwise, the advantages of cyclic computation are lost.” (Chomsky 2004: 107)

As a result, we expect that post-syntactic operations may not cut across phase boundaries to access information which is contained in two different phases/Spell-Out domains. Thus, at first sight, the assumption that the interfaces to SM and C-I are accessed at each phase level seems to establish strict isomorphism between the cycles of the syntactic and the post-syntactic computation, in the sense that both may use only information which is part of a single phase. Note, however, that the locality conditions imposed by this phase-driven model of Spell-Out are in fact slightly more restrictive for the post-syntactic components than they are for narrow syntax. For example, while a T head may enter into an Agree relation with material which is contained in the domain of v (e.g., with a quirky nominative object in Icelandic, see fn. 12 above), no such dependency can be created during the post-syntactic computation between elements that are part of different Spell-Out domains: once a subpart of a phrasemarker (say, VP) is spelled out, it is directly mapped to the interfaces and thus no longer accessible to operations involving material from other Spell-Out domains. In this sense, then, there is no strict parallelism between narrow syntax and the post-syntactic computation: syntactic operations may (minimally) cut across Spell-Out domains, while post-syntactic operations may not.

This should lead us to expect that there are empirical phenomena which reflect this kind of asymmetry, in the sense that there are instances in which the scope of post-syntactic processes is confined to the relevant Spell-Out domains, that is, VP or

TP. However, glossing over a small set of relevant empirical phenomena, it seems that it is rather the other way around. In other words, there are clear cases of “non-local” post-syntactic operations that cut across Spell-Out domains. For example, it is a well-known fact that the size of phonological phrases is dependent on, but not identical to the size of syntactic phrases (cf. e.g. Selkirk 1984; Truckenbrodt 1995, 1999). Yet, in contrast to what is expected under the above assumptions, prosodic phrases may include material from different Spell-Out domains. Consider the example from Italian in (15) and the corresponding prosodic structure in (16), in which phonological phrases are marked by φ (examples taken from Nespor et al. 1996: 9):

(15) [_{TP} Gianni avrà [_{vP} già mangiato [le belle mele]]].
 ‘Gianni will have already eaten the good apples.’

(16) (Gianni) φ (avrà già mangiato) φ (le belle mele) φ

In (16), the object *le belle mele* constitutes the rightmost phonological phrase. The next φ starts at the main verb and includes further material on its left, extending to the left edge of the relevant syntactic phrase (here: vP) Furthermore, the finite auxiliary *avrà* is integrated into the same φ as the main verb, since auxiliaries do not count as separate heads for purposes of prosodic phonology (cf. Selkirk 1984, Nespor and Vogel 1986). It is immediately clear that the kind of prosodic phrasing exhibited by examples such as (16) raises a problem for the assumptions (i) that cyclic Spell-Out affects VP and TP and (ii) that the phonological component “forgets about” what has been transferred to it at earlier stages of the derivation. Under this set of assumptions, we would not expect that the auxiliary (presumably located in T) forms a prosodic phrase together with material contained in the VP, which has been spelled out separately at the vP phase.¹⁴ These facts can be taken to indicate that the post-syntactic components must have access to pieces of information which are

¹⁴ Possibly, Chomsky (2004: 108) has similar cases in mind when he speculates “that global properties of phonology (e.g., intonation contour) are superimposed on the outcome of the cyclic operation of Φ [the phonological component].” However, it is not clear to me at which point of the post-syntactic computation the relevant operations should apply. Certainly, it is not very attractive to assume the existence of a phonological component that is associated with “global properties of phonology” and operates on the output of the “normal” phonological component, since this would void the advantages of cyclic computation (reduction of memory load etc.).

distributed over more than one Spell-Out domain (see Selkirk and Kratzer 2005 for related problems raised by properties of stress assignment and a solution based on a redefinition of phases).

The often non-local character of Binding phenomena constitutes another potential challenge for the assumption of cyclic Spell-Out.¹⁵ Fischer (2004) provides convincing arguments that the licensing and (phonetic) realization of bound pronouns represents another case where the domains of syntactic and post-syntactic operations differ. More precisely, she shows that the feature content of bound elements can be determined by operations of narrow syntax in a cyclic fashion, while the phonetic realization of these elements requires access to the complete syntactic chain created during the syntactic derivation.¹⁶ As noted by Fischer (2004: 204),

¹⁵ This is particularly clear for the unbounded character of Principle C of traditional Binding Theory (see Bouchard 2002: 351f. for discussion and some considerations concerning the extension of local domains for the purposes of binding theory). Note that the points raised here carry over to a non-syntactic approach in which Binding is analyzed as an interface phenomenon and attributed to properties of the C-I system (cf. e.g. Jackendoff and Culicover 2005; see Chomsky 2000: 146, n. 65 for a related suggestion). That is, in particular if Binding is treated as a post-syntactic phenomenon, its non-local character raises a problem for the assumption that the interpretative components are accessed in a phase-by-phase fashion. However, see Chomsky (2005) for some speculations on ways in which a syntactic analysis of Binding in terms of feature checking (along the lines proposed in Reuland 2001) can be integrated into a phase model.

¹⁶ Fischer (2004) develops a strictly derivational account of anaphoric relations in which the phonetic form and semantic interpretation of bound elements (including pronouns and anaphors) are determined as a result of syntactic operations. More precisely, she assumes that a bound pronoun β starts out in the syntax as a list of features {SE, PRON, SELF} which contains all possible realizations of β . During the syntactic derivation, the bound pronoun moves up in the structure, looking for its antecedent. When the bound element β reaches the edge of a cyclic domain (which corresponds to a phrase in Fischer's approach) and remains unbound, the feature matrix of β may be subject to an optimization process which reduces the number of features in the matrix. More specifically, the features corresponding to the most anaphoric realization of β may be deleted. In this way, the distance between the antecedent and the bound element is tracked, which captures the effects of the (representational) principles of traditional Binding Theory in a purely derivational approach. When the bound element eventually locates an antecedent with matching features, it does not move any further and enters into a checking relation with its antecedent. At this point, the concrete realization of β is determined (corresponding to the most anaphoric feature specification that remains in the matrix) and spelled out in the appropriate position (i.e., mapped to the interfaces). Crucially for our present purposes, this position may be located in a different Spell-Out domain than the antecedent where the form and interpretation of the bound element is ultimately determined.

“This constitutes a minimal violation of the Phrase Impenetrability Condition and the Strict Cycle Condition, but apparently this is what we have to accept if we want to integrate such a non-local phenomenon as binding into a local derivational approach.”

Note that this violation of cyclicity/locality is confined to the interpretative components of grammar, while in narrow syntax, the relevant derivation is strictly cyclic.¹⁷ Again, this suggests that post-syntactic operations may cut across Spell-Out domains. In the next section, I add further evidence from complementizer agreement in Germanic that such violations of strict cyclicity are characteristic of the post-syntactic components of grammar, that is, the mapping to the SM interface.

3.2 Complementizer agreement in Germanic

Before we can turn to the issues raised by complementizer agreement for the theory of cyclic Spell-Out, a closer look at the phenomenon in question is in order. It is well-known fact that in many non-standard varieties of Germanic, the subject’s ϕ -features are reflected not only on the verb, but also on the complementizer:¹⁸

(17) a. da-n=k ik werk-en *West Flemish*
 that-1SG=CLIT.1SG I work-1SG
 ‘that I work’

¹⁷ In a similar vein, Stechow (2005) argues that the LF (i.e., the interface to C-I) of an expression can be constructed cyclically from the output of narrow syntax (via the post-syntactic re-combination of individual phases/Spell-Out domains), but cannot be interpreted cyclically. That is, the proper interpretation of e.g. operator-variable chains in a syntactic structure X requires access to a single complete LF representation which contains all variables (plus indices) used during the syntactic derivation of X.

¹⁸ Cf. Bayer (1984), Altmann (1984), Weiß (1998, 2005), on Bavarian; Bennis and Haegeman (1984), Haegeman (1990), (1992), Shlonsky (1994), de Vogelaer et al. (2002) on (West) Flemish; de Haan and Weerman (1986), Hoekstra and Marácz (1989) on Frisian; Zwart (1993a), (1993b), (1997), van Koppen (2005) on various Dutch dialects; Hoekstra and Smits (1999) for an overview. Note that only West Flemish exhibits a full paradigm; in other varieties complementizer agreement is usually restricted to certain person/number combinations (Bavarian: 2nd person (and 1pl in some varieties), eastern dialects of Dutch: 1pl, southern dialects: 1pl and 3pl, Frisian: 2sg (plus 2pl in some varieties). See Fuß (2004, 2005) for a diachronic explanation of the person/number restrictions found in Bavarian.

As will become clear shortly, certain properties of complementizer agreement strongly suggest that this form of multiple agreement is accomplished by operations which are part of the post-syntactic components of grammar, that is, the mapping to PF (cf. Ackema and Neeleman 2004, Fuß 2005, 2008). Under the assumption that Spell-Out domains are to be defined as TP and VP (see above), this requires that the relevant post-syntactic operations may cut across Spell-Out domains, due to the fact that the source of the agreement features in C (either the subject or T) is spelled out prior to the Spell-Out domain (matrix VP) that contains C.

The first set of relevant data comes from dialects in which the shape of complementizer agreement differs from the shape of verbal agreement (called “double agreement” dialects in Zwart 1993a). In the Dutch dialect Hellendoorn, for example, the 1pl inflection found on the complementizer is /-ə/, while the verb carries the ending /-t/ (cf. Ackema and Neeleman 2003, 2004).¹⁹ This is shown in (20). In inversion contexts, the regular verbal agreement ending is replaced by the inflectional formative associated with complementizer agreement, cf. (21b):

¹⁹ Similar “double agreement” phenomena can be observed in Dutch dialects spoken in the Eastern Netherlands and Brabants, and in some Lower Bavarian dialects (on the latter see Bayer 1984, Kollmer 1987, Weiß 1998, 2005, and Fuß 2008).

(20) datt-e wiej noar't park loop-t
 that-1PL we to-the park walk-1PL
 'that we are walking to the park'

(21) a. Wiej loop-t noar't park.
 we walk-1PL to-the park
 'We are walking to the park.'

b. Volgens miej lop-e wiej noar't park.
 according-to me walk-1PL we to-the park
 'According to me we are walking to the park.'

However, the realization of complementizer agreement is subject to an adjacency requirement (cf. van Craenenbroeck and van Koppen 2002, Carstens 2003, Ackema and Neeleman 2004): the presence of an (scrambled) adjunct which intervenes between C and the subject blocks the availability of complementizer agreement. This restriction holds for both main and embedded clauses:

(22) a. dat/*datt-e [op den wärmsten dag van't joar]
 that/that-1PL on the warmest day of-the year
 wiej tegen oonze wil ewärkt hebt.
 we against our will worked have
 'that on the warmest day of the year we have worked against our will'

b. Volgens miej loop-t/*lop-e [op den wärmsten dag
 according-to me walk-1PL/walk-1PL on the warmest day
 van't joar] ook wiej noar't park.
 of-the year also we to-the park
 'According to me we are also walking to the park on the warmest day of
 the year.'

(22a) shows that the complementizer must appear without an inflectional ending if a (scrambled) PP intervenes between C and the subject. As illustrated by (22b), a similar adjacency effect can be observed in main clauses where the presence of an intervening XP blocks replacement of the regular verbal agreement ending /-t/ with /-ə/, the inflectional formative associated with complementizer agreement. Similar

adjacency effects can be observed in other Germanic varieties which exhibit complementizer agreement, cf. the following examples from Bavarian:²⁰

(23) a. obwoi-st du ins Kino ganga bist
 although-2SG you to-the movies gone are
 ‘although you went to the movies’

 b. *obwoi-st [woartscheints] du ins Kino ganga bist
 although-2SG probably you to-the movies gone are
 ‘although you probably went to the movies’

 c. obwoi [woartscheints] du ins Kino ganga bist
 although probably you to-the movies gone are
 ‘although you probably went to the movies’
 (Günther Grewendorf, p. c.)

Again, the presence of an XP (here a sentential adverb) that intervenes between C and the subject prevents the realization of inflection on the complementizer, as illustrated by the contrast between (23b) and (23c). As shown in Fuß (2005, 2008), this adjacency effect raises a problem for purely syntactic approaches to complementizer agreement.²¹ Moreover, under the assumption that adjacency effects are in fact more naturally accounted for in terms of post-syntactic operations/the mapping to PF (cf. e.g. Halle and Marantz 1993; Bobaljik 1994, 1995, 2002; Lasnik 1999, 2000 on affix hopping/Morphological Merger in English), these data suggest that complementizer agreement should be taken to operate in the phonological component of grammar.

²⁰ West Flemish and Frisian always require strict adjacency between the (inflected) complementizer and the subject. That is, violations of the adjacency requirement lead to ungrammaticality and not to non-inflected complementizers (Liliane Haegeman, Germen de Haan, p.c.). See Haeberli (2002) for detailed discussion and a syntactic analysis of the strict adjacency requirement.

²¹ The adjacency requirement in question is completely unexpected under approaches which attribute complementizer agreement to INFL-to-C movement (Hoekstra and Marácz 1989, Zwart 1993a, 1993b, 1997). Under the assumption that complementizer agreement results from a specifier-head relation between a separate AgrC-head and the subject (plus further AgrC-to-C movement, Shlonsky 1994), the adjacency effect can only be modeled by a stipulation that rules out adjunction of intervening XPs to AgrCP. Finally, approaches of complementizer agreement that posit an AGREE-relation between C and the subject (Carstens 2003, van Koppen 2005) must resort to the assumption that the intervening adjunct XP acts as an intervening goal (possibly due to the presence of a Case feature). As shown in Fuß (2005, 2008), this falsely predicts that adverbials intervening between T and the base position of the subject (SpecvP) should block the realization of verbal agreement.

Further support for a post-syntactic account of complementizer agreement comes from sluicing data from Bavarian:

(24) a. I woass dass-ts ihr a Madl gseng hoab-ts,
 I know that-2PL you a girl seen have-2PL
 owa I woass net wo-ts ihr a Madl gseng hoab-ts.
 but I know not where-2PL you a girl seen have-2PL)
 b. I woass dass-ts ihr a Madl gseng hoabts,
 I know that-2PL you a girl seen has-2PL
 owa I woass net wo (*-ts) ~~ihr a Madl gseng hoabts~~.
 but I know not where -2PL (you a girl seen have-2PL)
 'I know that you've seen a girl, but I don't know where (you've seen a girl).'
 (Günther Grewendorf, p.c.)

(24a) shows that in the absence of a complementizer, the C-related inflection can attach to the fronted *wh*-element in the embedded clause (*wo* → *wo-ts*).²² (24b) shows that complementizer agreement is blocked under sluicing, that is, in cases where an IP within a *wh*-CP is elided (cf. Lobeck 1995: 59; similar facts can be observed in Dutch dialects, cf. e.g. van Craenenbroeck 2004). Under the common assumption that sluicing is to be analyzed in terms of PF-deletion (Ross 1969, Lasnik 2001, Merchant 2001), the data in (24) indicate that the realization of complementizer agreement is sensitive to post-syntactic operations. This fact cannot be accounted for if it is assumed that complementizer agreement is established by purely syntactic

²² Similar facts can be observed in all cases lacking an overt complementizer, cf. the example in (i) (Bayer 1984: 235):

(i) a. Du soll-st song [CP[an wäichan Schuah]-st [IP du wui-st]].
 you should-2SG say which shoe-2SG you want-2SG
 'You should say which shoe you want.'
 b. [CP [Wia oit]-ts [IP ihr/ees sei-ts]] is mir wurscht.
 how old-2PL you are-2PL is me.DAT not-important
 'How old you are makes no difference to me.'
 c. [CP [Wia schnäi]-ts [IP ihr/es fahr-ts]]!
 how fast-2PL you.pl drive-2PL
 'How fast you drive!'

The fact that the exponents of complementizer agreement exhibit some optionality with respect to the host they can attach to can be taken as further indication that this kind of multiple agreement is not established in the syntax, but rather by post-syntactic readjustment operations that are sensitive to factors such as linear adjacency.

mechanisms. Rather, it suggests that the inflection carried by the complementizer is presumably the result of a post-syntactic mechanism as well. Importantly, the fact that the potential sources of C's agreement features (the subject or T⁰) are located in a Spell-Out domain separate from C (i.e., TP) seems to indicate that the relevant post-syntactic process ranges over more than a single Spell-Out domain. A final piece of evidence that helps to clarify the issues raised by complementizer agreement, in particular concerning the 'source' of C's agreement features, comes from comparatives in Bavarian:

(25) a. D'Resl is gresser [als wia-st du bist].
 the-Resl is taller than as-2SG you are
 'Resl is taller than you are.'
 b. *D'Resl is gresser [als wia-st du].
 the-Resl is taller than as-2SG you
 c. D'Resl is gresser [als wia du].
 the-Resl is taller than as you
 (Bayer 1984: 269)

(25b) shows that in comparatives, overt agreement on C leads to ungrammaticality if the finite verb is absent from the structure. The sentence becomes acceptable when the complementizer bears no inflection, as illustrated in (25c). This contrast shows that agreement between the complementizer and the subject cannot be implemented in terms of a checking relation between the set of φ-features in C and the subject in SpecTP. Otherwise one would expect examples such as (25b) to be grammatical (the φ-set of C should be able to enter into a checking relation with the subject's φ-set).²³ This conclusion holds for a syntactic analysis in terms of AGREE (e.g., Carstens 2003) as well as for an account involving PF checking rules as proposed by Ackema and

²³ As shown in (i), the second conjunct can be modified by adverbs such as 'yesterday' (with complementizer agreement still ruled out in the absence of a finite verb, Günther Grewendorf, Helmut Weiß, p.c.). This clearly shows that the bracketed part of (25c) is clausal in nature. Accordingly, the absence of complementizer agreement in (25c) cannot be accounted for by assuming that the second conjunct of (25c) is merely a PP and thus has a structure completely different from (25a).

(i) dass d'Resi heit schnella woar [als wia(*-st) du gesdan]
 that the-Resi today faster was than as(-2SG) you yesterday
 'that Resi was faster today than you were yesterday'

Neeleman (2004).²⁴ Moreover, the data in (25) suggest that it is the presence/absence of the inflected verb which is crucial for the availability of complementizer agreement. Thus, it seems that the inflection found in the C-domain is mediated by the finite verb. In other words, it seems that complementizer agreement is parasitic on the presence of a set of agreement features that has been evaluated during the syntactic derivation.²⁵ In Fuß (2005, 2008), I develop an analysis based on this idea, assuming that the agreement features giving rise to complementizer agreement are introduced by a post-syntactic operation which creates a copy of the relevant ϕ -set on T and inserts it to C at the level of Morphological Structure.

Following common practice in DM, I assume that the relevant set of agreement features in T (and C) corresponds to a separate agreement head or *morpheme* that is adjoined to a functional head with independent content (cf. e.g. Halle and Marantz 1993). The agreement morpheme on C (i.e., Agr-on-C) is analyzed as a post-syntactically inserted *copy* of the agreement morpheme located in T (Agr-on-T, which

²⁴ Focusing on the adjacency effects observed above, Ackema and Neeleman (2004) propose an analysis of complementizer agreement in terms of a PF feature checking rule which applies if C and the subject are part of the same prosodic phrase (marked by braces in (i)):

(i) *Germanic complementizer agreement*

$$\{[C \text{ (Prt)} \text{ (Add)} \text{ (Plr)}] \text{ [D} \text{ (Prt)} \text{ (Add)} \text{ (Plr)}]\} \rightarrow$$

$$\{[C \text{ (Prt}_i\text{)} \text{ (Add}_j\text{)} \text{ (Plr}_k\text{)}] \text{ [D} \text{ (Prt}_i\text{)} \text{ (Add}_j\text{)} \text{ (Plr}_k\text{)}]\}$$

(Ackema and Neeleman 2004: 241)

The rule in (i) serves to identify the set of ϕ -features associated with C (Prt = Participant, Add = Addressee, Plr = Plural) with the relevant (interpretable) ϕ -features of the subject. The adjacency effect is then attributed to a difference in prosodic phrasing caused by material adjoined to IP/TP: due to the presence of an XP that intervenes between C and the subject, rule (i) cannot apply since the complementizer and the subject are in two different prosodic domains (marked by braces):

(ii) a. $[\text{CP} \text{ C} \text{ [IP} \text{ XP} \text{ [IP} \text{ subject} \dots [\text{VP} \dots \text{V} \dots]]]]$
b. $\{C \text{ XP}\} \text{ {subject}} \{ \dots \} \{ \dots \text{V} \dots \}$

²⁵ Chomsky (2005: 9, fn. 23) considers the overt expression of inflectional features on C as further support for his proposal that T inherits its feature content from C, which is assumed to be the genuine host of ϕ -features. However, the fact that complementizer agreement is parasitic on verbal agreement seems to indicate that it is rather the other way around: the expression of inflectional features on C appears to depend on properties of T (i.e., its overt realization). The assumption that complementizer agreement is parasitic on verbal agreement is further supported by the observation that across Germanic, there appear to be no languages with complementizer agreement but without verbal agreement, while there are many languages that exhibit verbal agreement in the absence of complementizer agreement (Hoekstra and Smits 1999). Thus, it seems that cross-linguistically, the availability of complementizer agreement is dependent on the overt realization of verbal agreement morphology.

has been valued by a syntactic AGREE relation), giving rise to multiple agreement where the subject's ϕ -set is reflected on both the verb and the complementizer. Note that in realizational models of grammar such as DM, the post-syntactic insertion of inflectional heads/features (so-called *dissociated morphemes*, Embick 1997) is widely used to account for phenomena which involve features that do not receive an interpretation at the C-I interface (i.e., agreement and structural case, but not Tense or Aspect, cf. e.g. Marantz 1992, Halle and Marantz 1993, Embick 1997, Halle 1997, Noyer 1997, Harbour 2003). This approach to the phenomenon of multiple agreement preserves the idea that *syntactic* agreement is a unique dependency between two elements (cf. e.g. the Agreement Criterion, Uriagereka 1999: 270). Thus, inflectional features are absent from the C-domain during the core syntactic computation. This enables us to maintain a strict division of labor between the individual parts of clause structure in which the C-domain hosts features related to clause type or subordination, while inflectional features are confined to the IP/TP domain, at least in the syntax proper.

Under these assumptions, the restrictions on complementizer agreement observed above (adjacency effects, absence in sluicing and comparatives) can be accounted for in terms of constraints on the application of post-syntactic operations. While the absence of complementizer agreement in comparatives and sluicing constructions can presumably be attributed to the sequence of PF processes,²⁶ the adjacency effects illustrated in (22) and (23) suggest an explanation in terms of locality conditions on post-syntactic operations. In Fuß (2005, 2008) it is proposed that the copying/insertion procedure giving rise to Agr-on-C operates in a strictly local fashion, requiring structural adjacency between C^0 and T^0 (as is typical of morphological rules in DM, cf. e.g. Halle and Marantz 1993). As a result, the relevant post-syntactic copy operation can only create a dependency between C and T if no maximal projection intervenes between CP and TP (see Fuß 2005, 2008 for details). The adjacency effect can then be accounted for by assuming that scrambled XPs do not adjoin to IP/TP but occupy the specifier of a functional projection (TopP/FocP above TP; cf. Frey 2004, Grewendorf 2005; see Jayaseelan 2001, Belletti 2002, and Haeberli 2002 for related ideas) that is projected only if it serves to implement certain

²⁶ For example, we may assume that at MS, the insertion of morphological Agr-morphemes applies after the deletion of the syntactic terminal node which corresponds to the inflected verb (cf. Fuß 2005, 2008, for discussion; see e.g. Embick and Noyer 2001 and Ackema and Neeleman 2004 for the ordering relations between different types of MS/PF operations).

information-structural distinctions (and the head of which disrupts structural adjacency of C and T).²⁷

Summing up, I have argued that the phenomenon of complementizer agreement is established by operations which are part of the phonological component of grammar (basically following Fuß 2005, 2008). In particular, we have seen that certain properties of complementizer agreement suggest that the relevant post-syntactic operations may cut across Spell-Out domains, due to the fact that T, the source of the agreement features in C, is spelled out prior to the Spell-Out domain (matrix VP) that contains C. The next section comes back to the set of questions raised at the beginning of section 3, discussing the implications of complementizer agreement for the theory of the syntax-morphology/phonology interface.

3.3 Cyclic Spell-Out and phonological domains

If we accept the notion that complementizer agreement is accomplished in the post-syntactic components of grammar, we face another instance where morphological/phonological operations cut across the Spell-Out domains defined in Chomsky (2000) and subsequent work. Note that this conclusion holds not only for the above analysis in terms of a copy operation that transfers agreement features from T to C, but also for the model developed in Ackema and Neeleman (2004), which posits a PF checking relation between C and the subject. In both cases, there is a dependency between C (which is spelled out together with the matrix VP) and an element contained in a different Spell-Out domain (TP) which has already been sent to the interfaces. This raises the question of how the implementation of agreement on C can access information which is already gone from the computation. Three accounts seem likely on the basis of what has been discussed so far. First, we may assume that in the course of the derivation, a single unified interface representation (PF or Morphological Structure) is cyclically constructed from the successive output of narrow syntax (cf. e.g. Bouchard 2002: 343; see Stechow 2005 for a related proposal concerning the C-I interface). As a result, post-syntactic operations would have access to the whole structure of an expression generated by narrow syntax. In the case at hand, then, a dependency between C and T (or C and the subject) could easily be created, leading to complementizer agreement. Similarly, no problems would

²⁷ See Fuß (2008) for a discussion of the phenomenon of First Conjunct Agreement (van Koppen 2005) that raises a number of issues for a post-syntactic account of complementizer agreement.

arise in connection with prosodic phrasing or the realization of bound pronouns (see 3.1 above). However, note that this 'global' solution entails that PF is capable of storing the output of the individual cycles of narrow syntax until the final representation is assembled. Accordingly, we would completely lose the advantages of cyclic computation (i.e., reduction of memory load) for the phonological component.

Second, the problematic facts may be taken to indicate that the domain of Transfer operations should be redefined, in a way which warrants that C and T (or C and the subject) are part of the same Spell-Out domain (for concrete proposals to this effect, cf. e.g. Uriagereka 1999, Dobashi 2003, Fox and Pesetsky 2005; see Grohmann and Putnam 2007 for an alternative definition of Spell-Out domains based on the notion of 'Prolific domains' (Grohmann 2003)). In this way, we could ensure that the (post-syntactic) operations giving rise to complementizer agreement do not cross Spell-Out domains. For the sake of concreteness, we may assume that it is not the domain/complement of a phase head that is subject to Transfer, but rather the whole phase, including its edge (cf. e.g. Fox and Pesetsky 2003, 2005). Given Chomsky's motivation of phases in terms of interface conditions, this seems to be a natural move. Furthermore, this assumption would eliminate the asymmetry between phases and Spell-Out domains noticed in section 3.1 and in this way warrant strict isomorphism between the syntactic and post-syntactic cycles. However, note that a redefinition of Spell-Out domains along these lines is again a far-reaching proposal, which has serious consequences not only for the post-syntactic computation, but also for narrow syntax, in particular with respect to locality conditions (i.e., the PIC). For example, Spell-Out of a complete vP/CP phase including its edge should render impossible movement of elements contained in a phase, as for example in successive-cyclic *wh*-movement.²⁸ Moreover, it is questionable whether a redefinition of Spell-

²⁸ A possible solution to this problem would be to assume that the relevant phase is not really gone from the syntax after Spell-Out has applied (see also fn. 1 above). For example, Chomsky (2005: 9) hints at the possibility that the syntactic effects of the PIC actually follow from intervention effects: "Note that for narrow syntax, probe into an earlier phase will almost always be blocked by intervention effects. [...] It may be, then, that PIC holds only for the mappings to the interface, with the effects for narrow syntax automatic."

In other words, we might assume that after Spell-Out of phase PH, PH is still present in the syntax, with locality/cyclicity effects attributed to intervention phenomena. Movement to the edge of a phase would then have to be motivated as a necessary step to avoid intervention effects. In this sense, then, the point of Transfer would define only the size of the structural units that are dealt with in the post-syntactic components and are eventually interpreted at the interfaces.

Out domains alone suffices to capture other instances where post-syntactic operations cut across Spell-Out domains. For example, let's take a second look at the case of prosodic phrasing discussed in (15) and (16) above (repeated here as (26) and (27)). The assumption that there is complete isomorphism between syntactic and prosodic domains should lead us to expect that there is a derivational cycle in narrow syntax that corresponds to the prosodic phrase *(avrà già mangiato)φ*, which contains the auxiliary plus a participle but excludes the subject and object:

(26) [_{TP} Gianni avrà [_{VP} già mangiato [le belle mele]]].
 'Gianni will have already eaten the good apples.'

(27) (Gianni)φ (avrà già mangiato)φ (le belle mele)φ

However, it is very unlikely that there is a syntactic domain/cycle that includes *avrà già mangiato* (presumably corresponding to T'), but excludes the subject (presumably located in SpecTP). Thus, while it seems feasible to handle the problems raised by complementizer agreement by 'large-scale' solutions such as a redefinition of Spell-Out domains or the assumption that post-syntactic processes operate on a single, cyclically constructed PF representation, the conceptual and empirical issues raised by the relevant proposals cast into question whether this is actually the way we should proceed.

Alternatively, we may pursue a more conservative approach and assume that the relevant modification should concern not the definition of Spell-Out domains, but rather the domain of post-syntactic operations. In Fuß (2007), I propose the following definition of the scope of post-syntactic operations in the phonological component:

(28) *The domain of phonological operations*

Operations of the phonological component may access a single Spell-Out domain Σ_n and the right edge of the following Spell-Out domain Σ_{n+1} .

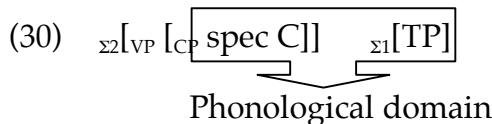
According to this definition of phonological domains, phonological/morphological operations may minimally cut across Spell-Out domains, accessing material which is part of two adjacent Spell-Out domains. Crucially, however, the range of this extension is sharply defined. In Fuß (2007), I claim that the "right edge" in (28) includes only the phase head (i.e., C in the case of $\Sigma_1 = \text{TP}$ and ν in the case of $\Sigma_1 = \text{VP}$) and the specifier(s) of CP and vP:

(29) *Definition of “right edge” in (28)*

The right edge of a Spell-Out domain Σ consists of a phase head (C or v) and its specifiers.

In this way, the advantages of cyclic computation in terms of a reduction of memory load are largely preserved in the phonological component. In the case of complementizer agreement, then, the phonological/ morphological component may create a dependency between material which is part of the TP Spell-Out domain (= Σ_1) and C⁰, which is located at the right edge of the next Spell-Out domain that is transferred to the phonological component (matrix VP = Σ_2). As a result, post-syntactic operations may copy agreement features from T to C, giving rise to an inflected complementizer.²⁹

In addition, the definition of “right edge” in (29) serves to repair the asymmetry between phases (CP, vP) and Spell-Out domains (TP, VP) created by the operation Transfer, which separates the domain of a phase head from the other parts of the phase (cf. section 3.1). Recall that this asymmetry has been deemed conceptually problematic, since it eliminates the parallelism between syntactic and phonological cycles and undermines the original motivation for phases in terms of interface conditions (due to the fact that the chunks reaching e.g. the SM interface do not correspond to CP and vP, but rather to VP and TP). My proposal that phonological operations are not confined to a single Spell-Out domain Σ_1 but may extend to the right edge of the next Spell-Out domain Σ_2 restores the isomorphism between the cycles of the syntactic and phonological computation by recombining those parts of a phase spelled out at Σ_1 (the complement of a phase head, e.g. TP in (30)) with the elements spelled out later (the phase head itself and its specifier(s)):



Positive evidence in favor of this characterization of phonological domains comes from complementizer agreement in Bavarian. As already noted in fn. 22 above, the

²⁹ Note that this definition of PF-domains works not only for the account proposed in this paper, but also for the analysis by Ackema and Neeleman (2004) in which C enters into a checking relation with the subject in SpecTP (if C and the subject are part of the same prosodic phrase).

relevant agreement formatives may also attach to elements occupying SpecCP, giving rise to an inflected *wh*-phrase, for example (Bayer 1984: 235):

(31) Du soll-st song [CP[an wäichan Schuah]-st[_{IP} du wui-st]]].
 you should-2SG say which shoe-2SG you want-2SG
 'You should say which shoe you want.'

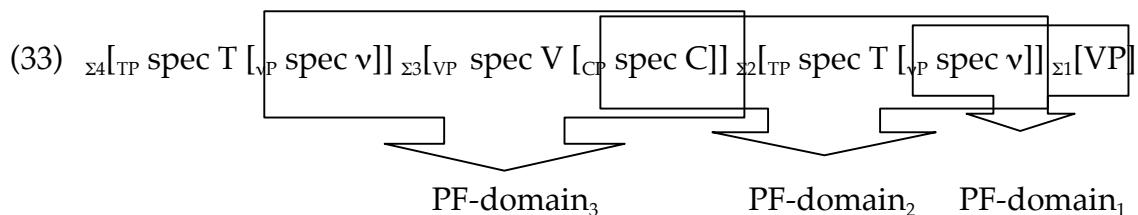
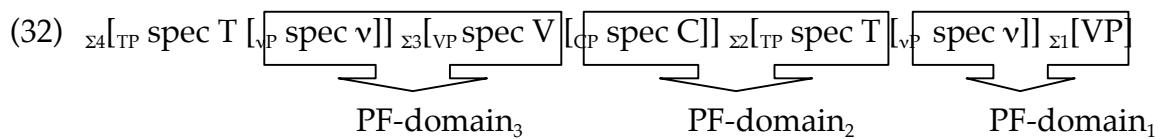
Thus, if C^0 does not contain any overt material the agreement formative can attach to, material in the specifier of CP is available as a host for the relevant inflectional ending.

While this set of assumptions removes the asymmetry between phases and Spell-Out domains for the purposes of the phonological computation, we may still wonder whether it also helps to overcome the very same asymmetry with respect to the mapping to the SM interface. Tentatively, I assume that this in fact the case. Thus, I propose that the phonological component recompletes the phasal units disrupted by the application of Transfer in narrow syntax. As a result, what surfaces at the SM interface is actually a unit which again corresponds to a syntactic phase.³⁰ In this way, Chomsky's original characterization of phases in terms of interface conditions (a phase is a coherent/independent phonological unit) can be maintained. Note that this 'repair' does not affect the way the syntactic derivation proceeds. Thus, in contrast to an approach that posits a redefinition of syntactic cycles, the present proposal does not require us to rethink basic syntactic concepts such as locality (i.e., the PIC).

So far, we have focused on the way syntactic Spell-Out domains are mapped to phonological domains. In what follows, I discuss how phonological domains relate to each other, focusing on the question of whether the right edge of a Spell-Out domain Σ_{n+1} forms a phonological domain only with the previous Spell-Out domain Σ_n or whether it is also part of the phonological domain based on Σ_{n+1} . In other words: Are phonological domains discrete units or do they overlap? In what follows I show that the latter is the more promising approach, and that the assumption of overlapping phonological domains makes correct predictions for a number of empirical phenomena.

³⁰ Conceivably, a similar mechanism is required in the semantic component, e.g. to create a complete predication structure that can be interpreted by the C-I-system – including the VP, (the copy of) the subject, and eventually temporal information (this was pointed out to me by Patrick Brandt).

A very general issue that arises under any approach that assumes a cyclic mapping from syntax to phonology concerns the question of how subsequent Spell-Out domains are successively mapped to phonological domains. In particular, if we assume that phases are recompleted in the phonological component, more has to be said about the affiliation of right edges (of Spell-Out domains) with respect to phonological domains. More specifically, it is unclear whether phonological operations may only affect the domain resulting from the union of a Spell-Out domain Σ_n and the right edge of Σ_{n+1} , cf. (32), or whether the phonological domains created by this process may overlap (i.e., whether the right edge of Σ_{n+1} is also accessible for operations at Σ_{n+1}), as illustrated in (33):



In Fuß (2007), I suggest that the more restrictive variant (32) is more adequate, since it captures the fact that CP generally constitutes a separate intonational domain, which at first seems to be somewhat at odds with the alternative in (33). In what follows, it will become clear that the more permissive variant (33) is actually to be preferred over (32) on both empirical and conceptual grounds.

First of all, the assumption of overlapping phonological domains guarantees that the output of the phonological component surfacing at the interface to SM is a coherent unit (consisting of cyclically constructed phonological domains which are linked together via the overlapping right edges of Spell-Out domains). Accordingly, it is not necessary to postulate late phonological processes that accomplish this task (as assumed by e.g. Chomsky 2004: 108).³¹ Note that this proposal still preserves the advantages of cyclic computation as long as we assume that phonological operations are confined to the individual phonological domains created post-syntactically. For

³¹ Alternatively, one might entertain the (quite unattractive) assumption that there is a computational process outside the computational system of grammar that reassembles the discrete phonological domains sent to the SM-interface.

example, phonological operations cannot create a dependency between T and V, since they are contained in different phonological domains.

An empirical argument in favor of overlapping domains comes from affix hopping in English, where by assumption T and v (hosting the verb) combine post-syntactically via Morphological Merger to produce an inflected verb (cf. e.g. Chomsky 1957, Lasnik 2000, Bobaljik 2002):³²

(34) ... [_{TP} John [_{T'} T [_{vP} t_{John} [_{v'} $\text{read}+v$ [_{VP} t_{read} a book]]]]]]

Morphological Merger

(33) predicts correctly that T and v may be affected by a phonological operation, since they are both part of the same phonological domain (PF-domain₂ in (33)). In contrast, the alternative in (32) proves to be too restrictive: According to (32), T and v are part of two different phonological domains (PF-domain₂ and PF-domain₁, respectively), which should lead us to expect that they cannot combine via phonological operations, contrary to facts.

Further empirical support for the extension of phonological domains resulting from a partial overlap of adjacent domains comes from cases where the prosodic phrasing crosses the phonological domains defined in (32). Given the phonological domains in (32), we would perhaps expect that phase boundaries closely match phonological boundaries. While this seems to be largely true of CP, which is commonly assumed to constitute an independent intonational unit, the effects of the supposed close isomorphism between syntactic and phonological cycles are much less clear clause-internally (see e.g. Selkirk and Kratzer 2005 for a critical review of these matters with respect to focus placement). Again, this can be illustrated with the kind of prosodic phrasing exhibited by examples such as (27), repeated here as (35). How can we ensure that the participle may form a prosodic phrase together with the adverb and the finite auxiliary? To see this, consider the more fine-grained syntactic representation of (35) in (36), where the participle *mangiato* is raised to v, with the adverb *già* occupying a position at the left edge of vP:

(35) (Gianni) φ (avrà *già mangiato*) φ (le belle mele) φ

(36) [_{TP} Gianni_i avrà [_{vP} t_i [_{v'} t_i [_{v'} mangiato+v [_{VP} t_V [_{DP} le belle mele]]]]]]].

³² An analysis in terms of affix hopping raises a number of further issues which cannot be addressed here in detail, see Bobaljik (1995), (2002), Lasnik (2000) and Embick & Noyer (2001) for discussion.

According to the definition of phonological domains in (32), the VP (= Σ_1) should form a phonological domain with the right edge of the TP (= Σ_2), that is, spec and head of vP. Crucially, however, the resulting phonological domain is too small to accommodate the fact that the finite auxiliary *avrà* forms a prosodic phrase together with the adverb *già* and the participle *mangiato* as illustrated in (35).³³ Again, the assumption of overlapping domains makes the correct prediction here: T and the left edge of vP are part of the same phonological domain, and thus could be converted into a single prosodic phrase.³⁴

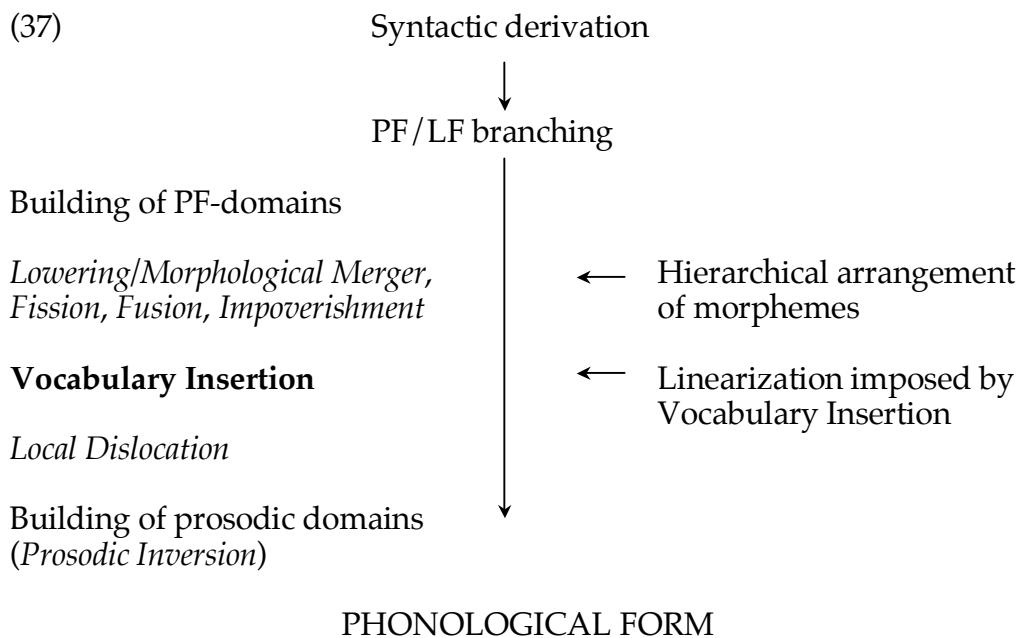
3.4 Section summary

This section has shown that operations of the phonological component may cut across the Spell-Out domains as defined in Chomsky (2000, 2001, 2004, 2005). In particular, I have argued that certain properties of complementizer agreement (adjacency effects, sensitivity to PF deletion processes) suggest that this form of multiple agreement is established by a post-syntactic operation that copies agreement features (valued in the syntax) from T to C under structural adjacency. Accordingly, T and C, which are part of different Spell-Out domains, must be part of a single domain in the phonological component of grammar. I have proposed that the phonological component maps the cyclic output of narrow syntax to phonological domains which are slightly larger than a single Spell-Out domain. More

³³ In Fuß (2007), I propose an independent mechanism enlarging phonological domains in order to account for cases such as (35). More precisely, phonological operations may extend over two complete Spell-Out domains Σ_1 and Σ_2 iff no overt material intervenes between the left edge of Σ_1 and the left edge of Σ_2 . However, apart from being quite ad hoc, this mechanism raises a number of further problems. For example, it cannot account for affix hopping in English and relies crucially on the notion of PF-adjacency ("overt material") which is not compatible with the assumption that at the point of constructing phonological domains no overt material is present yet (Vocabulary Insertion is usually taken to be a late phonological operation, cf. e.g. Embick and Noyer 2001, and section 2 above).

³⁴ Note that the object is part of a separate phonological domain. It is thus correctly predicted that it may form a separate prosodic domain, too. The fact that the subject corresponds to a separate phonological phrase is not a problem as long as we assume that phonological rules may divide a phonological domain into a number of separate prosodic phrases. This assumption is independently needed to ensure that clauses constitute separate intonational units. That is, the only problematic cases would be examples in which phonological operations seem to not respect the phonological domains defined here.

precisely, a phonological domain consists of a Spell-Out domain Σ_n and the right edge of a subsequent Spell-Out domain Σ_{n+1} . In this way, the phonological component can be taken to restore phasal units which have been disrupted by the application of the operation Transfer (affecting TP and VP, but not the phase head and its edge, cf. section 3.1), thereby warranting a close parallelism between the cycles of syntactic and post-syntactic computation. I have then shown that this proposal makes not only correct empirical predictions concerning the realization of complementizer agreement in Germanic; it also serves to maintain the original characterization of phases in terms of interface conditions if we assume that it is not Spell-Out domains, but rather the domains created by the workings of the phonological component which are sent to the SM interface. This leads to the following slightly modified characterization of the PF-branch of grammar, where prior to all other post-syntactic operations, the phonological component creates phonological domains from the cyclic output of the syntactic derivation:³⁵



In addition, I have discussed evidence suggesting that the individual phonological domains assembled in the phonological component do not represent separate discrete units. Rather, phonological domains overlap (i.e., the right edge of a Spell-Out domain Σ_n forms a phonological domain together with a previous Spell-Out Σ_{n-1} , but is also part of the phonological domain created from Σ_n). This extends the scope

³⁵ See Seidl (2001) for a related model of the post-syntactic computation. In contrast to Seidl, I assume that the construction of phonological domains from the output of the syntactic component precedes all other post-syntactic operations.

of phonological operations, which seems to give the right results for processes such as prosodic phrasing, affix-hopping etc. In the next section, I am going to argue that the assumption of overlapping phonological domains plays also an important role when it comes to creating linear order from the hierarchical structures generated by the workings of the syntactic computation.

4 The creation of linear order

In traditional generative approaches to the relationship between phrase structure and linear order, it is generally assumed that phrase markers directly represent the left-to-right order of the constituents (cf. e.g. Chomsky 1965: 123ff.). Language-specific differences in (basic) word order are attributed to different settings of the Head Parameter, which basically states whether a given head precedes or follows its complement(s) in a given language (i.e., the Head Parameter specifies the direction of c-selection either as a default or for particular syntactic heads; cf. e.g. Travis 1989 for some discussion). In other words, linear order is understood as a property encoded in the syntax, with restrictions on possible precedence relations being interpreted as restrictions on the well-formedness of syntactic trees (see Partee et al. 1993: 439ff.; this point of view is also taken up in recent work by Kayne 1994, Uriagereka 1998: 196, Fox and Pesetsky 2003, 2005, Müller 2007, where it is assumed that a violation of principles of linearization lead to a crash of the syntactic derivation). Syntactic linearization is generally taken to be governed by two basic constraints on possible phrase markers, the *Exclusivity Condition*, and the *Nontangling Condition* (Partee et al. 1993: 440):

(38) *The Exclusivity Condition*

In any well-formed constituent structure tree, for any nodes x and y , x and y stand in the precedence relation P , i.e., either $\langle x, y \rangle \in P$ or $\langle y, x \rangle \in P$, if and only if x and y do not stand in the dominance relation D , i.e., neither $\langle x, y \rangle \in D$ nor $\langle y, x \rangle \in D$.

(39) *The Nontangling Condition*

In any well-formed constituent structure tree, for any nodes x and y , if x precedes y , then all nodes dominated by x precede all nodes dominated by y .

The Exclusivity Condition expresses two requirements of linear ordering in syntactic trees. First, it requires that the linear ordering of terminals is total, including ordering statements for each terminal node. Second, it states that the relations of dominance and linear order are mutually exclusive, that is, no ordering statements can be given for nodes that stand in a dominance relation. The Nontangling Condition ensures that a precedence relation between two nodes includes all material dominated by these nodes, ruling out structures with crossing branches, or configurations where a given node is immediately dominated by more than a single node.

One of the most influential and explicit attempts to implement the effects of (38) and (39) is put forward in *Kayne (1994)*, who argues that linear order unambiguously reflects syntactic hierarchical structure. More specifically, *Kayne (1994: 33)* proposes the following principle mapping asymmetric c-command relations into precedence relations (i.e., linear orders of terminal elements):

(40) *Linear Correspondence Axiom (LCA)*

Let X, Y be nonterminals and x, y terminals such that X dominates x and Y dominates y . Then if X asymmetrically c-commands Y , x precedes y .

Kayne further assumes the totality requirement also expressed by the Exclusivity Condition above. That is, for each pair of terminals, it must be possible to state an ordering instruction (in terms of precedence) based on an asymmetric c-command relation between the non-terminals dominating these terminals. Furthermore, the linear ordering of two terminals must be *antisymmetric*, that is, conflicting ordering instructions (resulting from divergent asymmetric c-command relations between non-terminals) are forbidden.³⁶ Interestingly, the LCA serves to derive a number of generalizations/restrictions on properties of phrase structure, such as the fact that cross-linguistically, specifiers are to the left of heads, or the alleged impossibility of rightward movement. However, while the LCA serves to translate hierarchical structures into linear orders in an unambiguous way, it also has a number of quite far-reaching consequences which are not unproblematic. First of all, the LCA leads to an otherwise unmotivated proliferation of structure and a massive increase of

³⁶ It should be noted, though, that *Kayne (1994)* is primarily concerned with deriving basic properties of phrase structure (i.e., X' -theory) that are related to ordering generalizations (position of specifiers relative to heads, status of adjuncts etc.). In other words, while the LCA is commonly understood as a mechanism that translates hierarchical structure into linear order, the exact properties of the mapping from syntax to PF do not figure prominently in *Kayne's* work.

derivational complexity (see e.g. Kremers 2003, and in particular Richards 2004). A direct consequence of the LCA is that there is only a single base order, that is, projections are invariably head-initial, with the specifier universally preceding the head (the so-called Uniform Base Hypothesis, UBH). Under this assumption, there are no basic OV-orders. Instead, OV order must be analyzed as the result of leftward movement of objects (and other complements) up to a higher functional projection. The triggers for these processes often remain unclear or have to be stipulated (case for DPs, related requirements for PPs and oblique complements; see Koizumi 1995, Kural 1997 for early criticism concerning the analysis of strict SOV languages such as Japanese or Turkish). This is especially problematic in accounts that make use of massive leftward (remnant) XP-movement where even SVO orders are often analyzed as the result of raising VP-internal elements to the left of VP, followed by moving the remnant VP to the left of these elements (see e.g. Kayne 1998, Hróarsdóttir 2000, Hinterhölzl 2002, Julien 2002). A related problem comes from the fact that the LCA requires that each projection have only a single specifier position, ruling out additional adjunction to XP (Kayne 1994: 22).³⁷ Again, this leads to a proliferation of empty functional heads to provide landing sites for movement, in particular (i) if it is assumed that adverbs are introduced as specifiers of functional heads (Alexiadou 1997, Cinque 1999), and (ii) if roll-up movement is postulated, i.e., movement of the complement of a (functional) head X to the specifier of XP, followed by moving the whole XP to the specifier of the next higher head etc. (see Julien 2002 for an analysis of SOV grammars in terms of massive roll-up movement; see Kremers 2003 for a critical discussion of roll-up movement focusing on adjectival structures in Arabic).

The next section presents an updated algorithm for linearizing syntactic hierarchical structures, basically following proposals by Epstein et al. (1998) and Richards (2004).

4.1 Symmetric c-command and a phonological Head Parameter

Under current minimalist assumptions, the syntactic component of grammar generates complex hierarchical structures via repeated (recursive) applications of the operation Merge. Merging two elements A and B creates a set {A, B}, with A/B either a primitive element taken from the lexicon, or a complex category created by

³⁷ Due to Kayne's (1994: 16) definition of c-command, multiple specifiers would c-command each other, leading to a violation of the LCA.

previous applications of Merge. Assuming that hierarchical relations are stated in derivational terms, Merge creates a mutual (i.e., *symmetric*) c-command relation between A and B. Moreover, if one of the elements, say B, is a complex category, A is taken to c-command all nodes dominated by B. However, as becomes clear from the set notation, Merge does not create a linear ordering between A and B. Rather, the linear ordering of syntactic terminal nodes is accomplished by operations on the PF branch of grammar that interpret hierarchical syntactic structures (Chomsky 1995: 334; Uriagereka 1998: 217f.; Epstein et al. 1998: 139ff., Nunes 1999: 244, Dobashi 2003, Richards 2004, Embick and Noyer 2007 and many others). Thus, linear order is taken to be a property imposed on the syntax by requirements on the interpretive interfaces: the output of the syntactic computation must be instantiated in real time through articulatory / perceptual mechanisms (be it speech or gesture) that require a strict sequential, one-dimensional concatenation of linguistic signs (the status of this requirement is somewhat less clear with respect to sign languages). However, while linear order is not directly encoded in the syntax, the syntactic hierarchical structures must contain some information that can be put to use by phonological operations to map the output of the syntax into a sequence of words/phonological exponents. As noted by Richards (2007: 215f.), it is a conspicuous property of linearization algorithms based purely on the notion of asymmetric c-command that they cannot linearize the most basic of all syntactic relations, sisterhood, without further assumptions. If precedence relations solely reflect asymmetric c-command relations in the syntax, then the mutual, symmetric c-command relation typically created by Merge raises a problem for the syntax-PF mapping. In particular, “the base pair of every (sub)tree remains unlinearizable, since here only head-level categories (terminals) are combined.” (Richards 2007: 215). Special strategies proposed to circumvent this problem include (obligatory) movement of one sister node (Chomsky 1995: 337), combination of the sister pair into a morphological word (Chomsky 1995: 337, perhaps by cliticization, cf. Uriagereka 1998: 219), or appealing to some version of the traditional (syntactic) head-parameter (Saito and Fukui 1998).

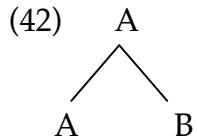
Richards (2004, 2007), basically following Epstein et al. (1998), proposes an alternative linearization algorithm that is not in need of such additional machinery. Epstein et al. (1998: 151) suggest that the mechanism linearizing syntactic structures should exploit the most basic (and most local) structural relation created in the syntax, namely ‘derivational sisterhood’, directly resulting from the application of Merge. In other words, the structural relation relevant for the workings of

linearization is simple c-command rather than asymmetric c-command, contra Kayne (1994). Accordingly, the LCA must be revised as follows:

(41) *Revised LCA* (Epstein et al. 1998: 151)

If X c-commands Y, then the terminals in X precede the terminals in Y.

However, it is immediately clear that the use of the sisterhood relation provides more information than required for linearization. Even worse, symmetric c-command maps into conflicting ordering instructions at PF. Merge of A and B, giving rise to the phrase marker in (42), creates a symmetric c-command relation between A and B.



Accordingly, we derive the following ordering instructions in line with (41): {A>B, B>A}, where each sister node is required to precede the other. In other words, the use of the sisterhood relation actually *overdetermines* linearization (Richards 2007: 216). Since mutual precedence is a logical impossibility, there must be a PF repair strategy that weeds out a subset of the c-command relations for the purposes of linearization. Epstein et al. (1998: 152) propose the following principle that operates at PF and serves to ignore a subset of the symmetric c-command relations created by derivational sisterhood (see also Richards 2004: 24f., 2007: 216):

(43) *The Precedence Resolution Principle*

If two (not necessarily distinct) categories symmetrically c-command each other by virtue of some syntactic operation O, ignore all c-command relations of one of the categories to the terms of the other with respect to establishing precedence via the LCA.

(43) effectively serves to parameterize the LCA (Richards 2004: 24f.), thereby reinstating the Head Parameter as a PF principle that allows a language-specific choice with respect to the set of c-command relations ignored by the linearization algorithm. This is illustrated in (44) for the head-complement relation in VP (cf. Richards 2007: 217; recall that there are no linear relations in the syntax, so the order

of verb and object used in the phrase marker has no direct implications for the actual linear order of elements):

(44) VP	<i>c-command relations</i> $\{V > DP, DP > V\}$	\Rightarrow	<i>ignore</i>	<i>PF-order</i>
			$V > DP$	$DP > V (=OV)$
			$DP > V$	$V > DP (=VO)$

As shown in (44), basic OV order results if the c-command relation $V > DP$ is ignored during the mapping from syntax to PF, while ignoring $DP > V$ leads to basic VO order. Richards (2004, 2007) generalizes this approach to the effect that for any given pair of elements (e.g., V and O) a consistent subset of c-command relations is deleted/ignored in the mapping to PF throughout the whole derivation. The resulting PF version of the Head Parameter, which can be taken to be “an interface strategy for resolving syntactic symmetry” (Richards 2004: 25) is given below:³⁸

(45) *Parametrized LCA* (Richards 2004: 25)

Merge(α, β) $\rightarrow \{\langle \alpha, \beta \rangle, \langle \beta, \alpha \rangle\}$

- a. VO = Ignore all $O > V$ [i.e., $\{\langle \alpha, \beta \rangle, \langle \beta, \alpha \rangle\} \rightarrow \{\langle \alpha, \beta \rangle\}$]
- b. OV = Ignore all $V > O$ [i.e., $\{\langle \alpha, \beta \rangle, \langle \beta, \alpha \rangle\} \rightarrow \{\langle \beta, \alpha \rangle\}$]

The central motivation of this revised version comes from shape conservation effects, where movement operations preserve the original order in the base, the best known being Holmberg’s Generalization on object shift in Scandinavian (Holmberg 1986, 1999). As is well-known, Scandinavian languages such as Swedish, Danish, or Icelandic allow leftward movement of an object only if the verb likewise moves to a position to the left of the shifted object, compare the following examples from Icelandic (Richards 2007: 211):

(46) a. Nemandinn las (bókina) ekki (bókina).
 student-the read book-the not book-the
 ‘The student didn’t read the book.’

b. Nemandinn hefur (*bókina) ekki lesið (bókina).
 student-the has book-the not read book-the
 ‘The student hasn’t read the book.’

³⁸ Note that (45) must be open to parametrization for individual syntactic categories to be able to account for mixed OV/VO languages such as German, which exhibits basic OV with all [+V] categories, but otherwise VO (C, N, D, P etc.).

The fact that object shift preserves the original base order of verb and object immediately follows from the parameter setting (45a), which requires that in a VO language, only ordering instructions of the type V>O are licit. The ungrammatical order with the verb to the right of the object simply cannot be derived/linearized in the relevant languages.

While this account seems to work fine for HG effects, it immediately raises the question of how we can derive movement operations that do *not* preserve the base order of V and O, such as topicalization of the object in V2 clauses in Scandinavian, fronting of a *wh*-object in English, or subject-initial V2 clauses in a basic OV language such as German (with the unpronounced copy marked by strikethrough):

(47) Which book did the student read (~~which book~~)?

(48) Der Student liest das Buch (~~liest~~).
 the student reads the book
 'The student reads the book.'

In the above examples, the base order of V and O is disrupted by subsequent movement operations (moving the *wh*-object to the left of the verb in (47) and the finite verb to the left of the object in (48)). The solution proposed by Richards to account for these apparent differences between movement operations with respect to shape conservation makes use of the notion of cyclic Spell-Out already discussed in section 3 above (see also Fox and Pesetsky 2003, 2005; Müller 2007). More precisely, Richards adopts the idea (Chomsky 2001, 2004) that the syntactic computation proceeds in derivational cycles, called phases. As already noted above, phases are identified as CP and vP. Furthermore, he assumes that the domain (i.e., the complement) of each phase head is transferred to the phonological component when the next phase head is merged (Chomsky's 2001 version of the Phase Impenetrability Condition, PIC). As a result, material which has been subject to Transfer/Spell-Out is gone from the computation and no longer accessible to further operations. This architecture facilitates the following generalizations concerning the behavior of different movement types with respect to shape conservation (Richards 2007: 213):

(49) a. Shape-preserving movement is phase-internal.
 b. Shape-destroying movement is trans-phrasal.

The relevant facts can then be accounted for if it is assumed that the *Parametrized LCA* in (45) “must be relativized to the *phase*” (Richards 2007: 218).³⁹ Thus, Richards claims that c-command relations (giving rise to ordering statements) which have already been transferred to the phonological component are no longer accessible for the computation (recall that it is assumed that the phonological computation proceeds in parallel with the syntactic computation, so that PF operations cannot access information transferred to PF at different cycles):

“The logic of our dynamic linearization system is such that the derivational information that feeds the linearization algorithm in [(45)] is lost (‘forgotten’) beyond the immediate phase level. Once a phasal domain is sent to Spell-Out, there is no memory of the c-command relations and ordering partners (merge-sisters) created within that phase.” (Richards 2007: 218)

As a result, shape conservation effects are confined to short, phase-internal movement such as object shift, which is analyzed as movement to SpecvP, while longer movement steps that cross a phase boundary may give rise to new, and apparently conflicting c-command relations and ordering statements. However, the latter do not create any problems as long as the computation can forget about earlier ordering instructions that are no longer accessible after the relevant phase has been subject to Transfer.⁴⁰

An obvious question not discussed so far concerns the status of copies in this approach. How does PF know that it must not pronounce/linearize a lower copy of an element that has undergone movement? Here, Richards (2007: 220) assumes that the relevant distinction can be made by appealing to the notion of uninterpretable features that render a given syntactic object active for the purposes of the syntactic computation (e.g. Case on nominals, [wh] on *wh*-phrases etc.). More precisely, he suggests that elements with uninterpretable features are invisible for PF operations and therefore do not participate in linearization via (45). This is taken to prevent the pronunciation of (i) the foot of a movement chain, and (ii) intermediate chain links

³⁹ Accordingly, “all” in (45) is confined to c-command relations within a single phase.

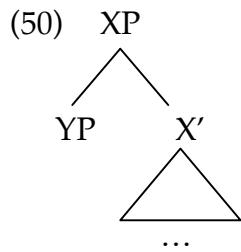
⁴⁰ It should be noted that Richards’ account of shape conservation with object shift includes a number of intricate details which cannot be discussed here in full, e.g., concerning the different behavior of pronouns and full DP objects, or the exact timing of Spell-Out/TRANSFER with respect to the c-command relation V>DP resulting from V-to-T movement (to ensure shape conservation); see Richards (2004), (2007), and Müller (2007) for extensive discussion.

(i.e., elements occupying phase edges). However, this condition on PF-visibility raises a number of further questions, notably concerning the interplay between the operation Agree (which values and eliminates uninterpretable features) and Move/external Merge (which is triggered by EPP or edge features) in an approach making use of cyclic Spell-Out. For example, it is commonly assumed that a chunk of syntactic structure sent to the PF-branch may not contain uninterpretable features anyway. In fact, it is a leading idea of minimalist syntax that syntactic operations are triggered by the need to eliminate uninterpretable features prior to the interfaces. So it is not entirely clear of how we can appeal to the presence of uninterpretable features (in the PF-branch) to render a given element invisible for phonological operations.⁴¹ Further problems are raised by the phenomenon of *wh*-in-situ, which I cannot address here in detail. Instead, I want to focus on another crucial problem for any approach that exploits symmetric c-command relations to create linear orderings. The relevant issue is raised by the generalization (see above) that specifiers uniformly precede rather than follow the head.

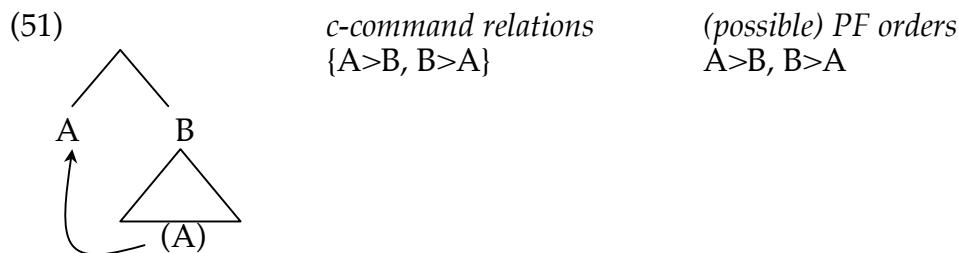
4.1.1 The status of specifiers

Above we have already noted that it seems to be a widely accepted notion that specifiers uniformly precede the head of a projection across languages. This fact follows directly from Kayne's (1994) LCA based on asymmetric c-command: The specifier asymmetrically c-commands and precedes all material dominated by its sister node (note that Kayne assumes that X'-categories may not enter into c-command relations):

⁴¹ This problem appears to be even more obvious in the probe/goal system envisaged in Chomsky (2000) and subsequent work, where Move is not any longer a necessary component of checking operations. In particular, the assumption that the valuation of uninterpretable features is accomplished by an Agree operation that targets elements prior to movement effectively serves to eliminate the possibility of lower copies that contain uninterpretable features (apart from EPP/edge feature driven movement that targets the left edge of phases to circumvent violations of the Phase Impenetrability Condition).



However, if it is assumed that it is *symmetric* c-command relations that map into linear orderings, it appears that the special status of specifiers with respect to precedence cannot be derived in a straightforward manner (this problem is left open in Richards 2004, 2007. Epstein et al. 1998 consider a couple of solutions, which will be discussed shortly). To illustrate this, let us take a look at XP-movement, which is standardly assumed to target specifier positions. Note that conflicting ordering instructions (as a result of symmetric c-command) are not only created by base Merge (i.e., *external Merge*), but also by Merge as part of the more complex operation Move (*internal Merge*). Again, the mutual c-command relation between the moved category A and the node B with which it is merged provides us with contradictory precedence relations:



Of course, it is possible to resolve the contradictory ordering statements by discarding a subset of the c-command relations (along the lines of (45)). However, recall that across languages, specifiers seem to uniformly precede the syntactic object they are merged with. So appealing to some form of the Head Parameter misses a generalization, namely that in fact only one of the options is attested. So how can we ensure the correct outcome? Note that we cannot appeal to some form of shape conservation (recall that trans-phrasal movement is typically shape-destroying; information about the base order is often part of a lower phase and therefore not accessible). Alternatively, one can simply rule out the non-existing options by a stipulation (e.g., by stating that movement must lead to new ordering relations, i.e., 'have an effect on the output', cf. e.g. Epstein et al. 1998: 153; see also Chomsky 2001). To take a concrete example, let's take a look at shape-destroying *wh*-movement in an SVO language such as English:

(52) Which book did the student read (~~which book~~)?

Wh-movement of *which book* creates a specifier that precedes the rest of the clause. So the crucial question is how we can rule out an alternative ordering, with the material dominated by C' preceding *which book* (linearizing the c-command relation C' > *which book*)?

(53) *did the student read (~~which book~~) which book?

In such cases, Epstein et al. (p. 153) attribute the position of specifiers created by Move to the requirement that movement have an effect on the (PF) output. This requirement is not satisfied in (53), where the specifier follows the attracting head, resulting in the same order as without movement (basically an instance of string vacuous movement). However, a ban on string-vacuous movement does not give the correct results for embedded interrogatives in OV-languages like German, where the wrong ordering of C' and the *wh*-phrase has an effect on the output, albeit one that is ungrammatical, namely (55). Moreover, clause-final positioning of *wh*-phrases (due to rightward movement) is an option that is not attested cross-linguistically (cf. e.g. Uriagereka 1998: 215):⁴²

(54) Ich weiß nicht [_{CP} wen [_C C [_{TP} Peter (~~wen~~) gesehen hat]]].
 I know not who.ACC Peter who.ACC seen has
 'I don't know who Peter saw.'

(55) *Ich weiß nicht [_{CP} [_C C [_{TP} Peter (~~wen~~) gesehen hat]] wen].

The next section shows that the issues raised by specifiers can be handled by a modification of Epstein et al.'s and Richards' proposals which is based on a set of assumptions concerning the way ordering information is incrementally built up at the interface to SM from the cyclic output of syntax and the workings of the phonological component, in particular drawing on the notion of overlapping phonological domains developed in section 3.3.

⁴² However, note that Neidle et al. (1997) claim that American Sign Languages exhibits rightward movement of *wh*-phrases (but see Pretonio and Lillo-Martin (1997) for an alternative analysis in terms of leftward *wh*-movement). See also Bokamba (1976) for the claim that rightward *wh*-movement can be observed in a couple of Bantu languages.

4.2 Cyclic linearization and Vocabulary Insertion

This section develops a theory of linearization based on the notion that linear order is created from hierarchical syntactic structures by the workings of Vocabulary Insertion (*The Late Linearization Hypothesis*, see above). It is commonly assumed that Vocabulary Insertion converts hierarchical structures into linear strings of exponents. Furthermore, I assume that the linearization process (henceforth LIN) does not linearize each syntactic node present in the structure. Rather, given that LIN is part of the workings of Vocabulary Insertion, it seems plausible to assume that it affects only phonologically visible material, in the sense that it has to determine the position of each newly inserted phonological exponent relative to the linear string that has been established by previous applications of Vocabulary Insertion. To accomplish this task, LIN may access the hierarchical information made available by the syntactic structure. In addition, the newly inserted exponent must be adjacent to the linear string of words/exponents created so far. In other words, linear order is built up incrementally from successive applications of Vocabulary Insertion.

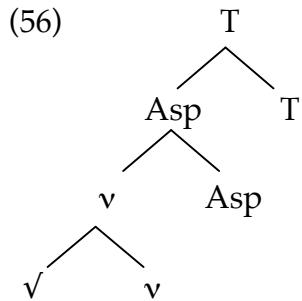
In an approach such as DM, it is natural to assume that the realization of word order (i.e., of exponents that constitute phonological words) proceeds along the same lines as the linearization of word-internal structure. Recall that DM does not recognize a theoretically significant distinction between morphosyntactic features which are realized as (i) bound or (ii) free formatives, that is, words (“syntactic hierarchical structure all the way down”).⁴³ It is expected that this carries over to the workings of LIN. That is, similar to the realization of affix sequences and the prefix/suffix distinction, word order is determined by a combination of lexical properties of certain Vocabulary items and hierarchical structures generated by the syntactic component, resulting in ordering instructions at PF. This is illustrated in the following, beginning with the linearization of X^0 -structures.

4.2.1 LIN and X^0 -structures

Let us first take a closer look at the linearization of head adjunction structures that leads to the linear ordering of word-internal material (i.e., stems and affixes). Compare the following head adjunction structure (assembled by syntactic head

⁴³ Cf. Harley and Noyer (1999: 7): “Features which will eventually be realized as a subpart of a phonological word are treated no differently from features which will eventually be realized as an autonomous word.”

movement) where a verb stem (i.e., the combination of a root plus category defining *v*) combines with functional heads Asp and T:



On the standard assumption that the morphological derivation must mirror the syntactic derivation (the Mirror Principle, Baker 1985, 1988), the phonological exponent of a lower functional head must be closer to the verb stem than the phonological exponents of higher functional heads.⁴⁴ In Fuß (2005), I propose that the effects of the Mirror Principle can be derived if Vocabulary Insertion is taken to proceed in a bottom-up fashion (as widely assumed in Distributed Morphology; see however Legate 1999 and in particular Phillips 1996, 2003 for approaches making use of top-down insertion). In other words, Vocabulary Insertion affects the verbal or nominal root before it affects functional heads the root adjoins to (so-called 'root-out insertion'). For example, in the head adjunction structure (56), Vocabulary Insertion targets first the category-neutral root. Subsequently, the phonological exponent of the closest functional head is inserted (here: *v*), which attaches to the root (creating a verb stem), followed by insertion of the exponent of Asp. Finally, the exponent of T is attached to the existing sequence of phonological exponents, ensuring that the exponent of T is the outermost element in the resulting word. In this fashion, the effects described by the Mirror Principle follow from the way linear order is created at the point of the derivation where phonological exponents are inserted into the terminal elements of the syntactic structure. The ordering restrictions imposed by the hierarchical structure assembled in the syntactic component are supplemented by lexical properties of individual Vocabulary items to create the final ordering instructions which are sent to the articulatory system.⁴⁵ The lexical properties in

⁴⁴ Of course these effects can only be detected if the relevant inflectional markers are located on the same side of the verb stem and if their order is not affected by other processes that take place at MS.

⁴⁵ In addition, morphological operations such as *Local Dislocation* may reorder elements/heads which are adjacent and part of the same constituent at MS. Recall that *Local Dislocation* differs from Morphological Merger in that the latter takes place prior to linearization whereas the former applies to structures that have already been linearized (see Embick and Noyer 2001 for discussion).

question determine the orientation of affixes (i.e., the prefix/suffix distinction). In other words, the linearization of a head adjunction structure depends on selectional properties of bound forms/Vocabulary items: prefixes select a host to their right, whereas suffixes require a host to their left. On these assumptions, a structure such as (56) may give rise to the sequences in (57), but excludes those in (58) (where “V” is used as an abbreviation for the combination of root and category-defining ν):⁴⁶

- (57) a. V+Asp+T (uniformly suffixing)
- b. T+Asp+V (uniformly prefixing)
- c. T+V+Asp (T prefix, Asp suffix)
- d. Asp+V+T (T suffix, Asp prefix)

- (58) a. *V+T+Asp
- b. *Asp+T+V

The prefix/suffix status of the individual Vocabulary items can be recast in terms of a phonological head parameter if we assume that LIN ignores a subset of the symmetric c-command relations established in the syntax (cf. Epstein et al. 1998, Richards 2004, 2007; repeated here for convenience):

- (59) *The Precedence Resolution Principle*

If two (not necessarily distinct) categories symmetrically c-command each other by virtue of some syntactic operation O, ignore all c-command relations of one of the categories to the terms of the other with respect to establishing precedence via the LCA.

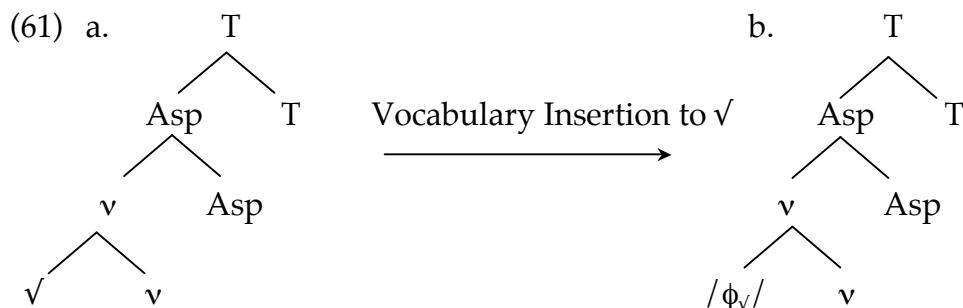
Dependent on the lexical properties of phonological exponents, LIN ignores one c-command relation for each sister pair in a head adjunction structure such as (56) (where symmetric c-command holds between the elements of each sister pair). Thus,

⁴⁶ The sequences in (57c-d) represent possible but typologically marked options which are not very frequent across the world’s languages. This can be attributed to the assumption that there is normally a default orientation for inflectional affixes in a given grammar, that is, languages prefer to be uniformly prefixing or suffixing (cf. e.g. Bybee et al. 1990, Hale 1996). Alternatively, it is conceivable that affixes select not only the direction where their host must appear, but also its category. Then, the marked sequences (57c-d) might follow from the fact that cross-linguistically, T preferably attaches to Asp.

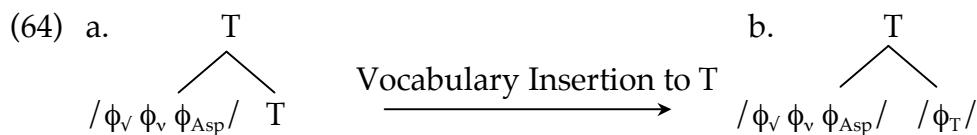
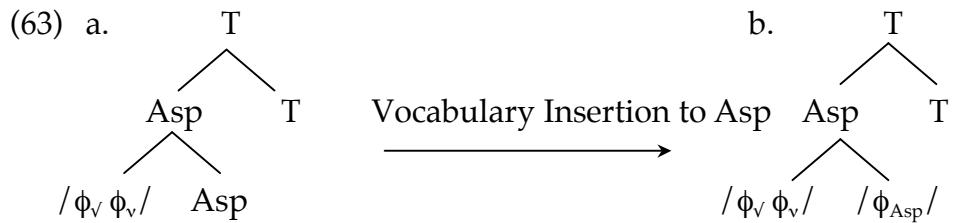
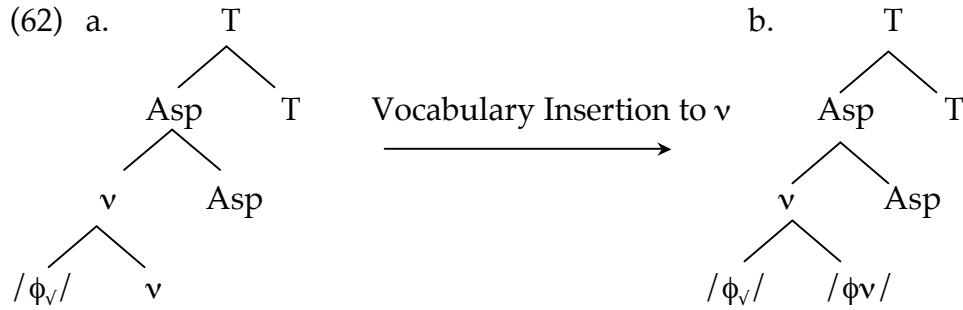
in languages that exhibit the affix orders (57a,c), LIN consistently ignores the c-command relation $\langle \text{Asp}, v \rangle$, ensuring that exponents of Asp are realized as suffixes (i.e., the c-command relation $\langle v, \text{Asp} \rangle$ is mapped onto a linear order where the exponent of v (or, rather \sqrt{v}) precedes the exponent of Asp). Furthermore, it is plausible to assume that the phonological Head Parameter is expressed by individual Vocabulary items (traditionally spoken, lexical properties of individual (functional) heads, cf. the *Lexical Parametrization Hypothesis*, Borer 1984, Wexler and Manzini 1987) at the point of Vocabulary insertion. In other words, the relevant information concerning which c-command relation must be ignored is part of the lexical specification of (individual) Vocabulary items. This can be illustrated with the following Vocabulary item realizing a suffixal past tense T head in English (with “ β ” presumably corresponding to v if no other functional heads intervene):⁴⁷

(60) $[_T \text{ PAST}] \leftrightarrow /d/$
 $\{ \langle T, \beta \rangle, \langle \beta, T \rangle \} \rightarrow \{ \langle \beta, T \rangle \}$

As a result, for each step α of the insertion/linearization procedure, the Head Parameter is linked to the phonological exponent element that is inserted at α . This is illustrated in (61) to (64) for a uniformly suffixing language (phonological exponents are referred to as ϕ and set in slashes):



⁴⁷ Note that in many cases the exponents of functional categories may be null. Still, it seems plausible to assume that null exponents may also carry a specification for the Head Parameter.



As illustrated in (61) to (64), successive applications of Vocabulary Insertion effectively serve to 'flatten' the hierarchical structure created in the syntax, producing a linear string of phonological exponents. At each step of the insertion procedure, the relative order between the newly inserted exponent ϕ and the existing string of elements (resulting from previous applications of Vocabulary Insertion) is determined by lexical properties of ϕ , eventually giving rise to the affix order $/\phi_v \phi_v \phi_{Asp} \phi_T/$ in the example at hand (note that it is always the projecting head that determines affix order). Following work by Ouhalla (1991) and Chomsky (1991, 1995), I tentatively assume that parametric variation is determined by lexical properties of a (closed) class of (exponents of) functional categories. As a result, lexical entries for category-neutral roots are not specified for settings of the Head Parameter. Furthermore, note that the exponent ϕ that is inserted at the top of a head-adjunction structure has to fulfill a double role with respect to LIN. Its lexical properties determine the order of ϕ relative to (i) the exponents of other heads of the head complex, and (ii) the exponent(s) realizing the XP-complement of the relevant head. Thus, the Vocabulary item realizing T in (64) must also contain a linearization instruction that determines the linear order of T and (the exponents of) vP. In the case of English T, this can be stated as follows:⁴⁸

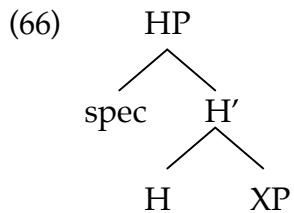
⁴⁸ Of course, it is not very attractive to assume that the relevant linearization instructions are stored separately for each possible phonological exponent of T. In most languages, different realizations of a

(65) $[_T \text{PAST}] \leftrightarrow /d/$
 $\{<T, v>, <v, T>\} \rightarrow \{<v, T>\}$
 $\{<T, vP>, <vP, T>\} \rightarrow \{<T, vP>\}$

In what follows, I am going to examine the notion of a phonological Head Parameter in more detail, addressing the question of how this approach carries over to the linearization of larger structures, that is, the computation of word order.

4.2.2 LIN and XP-structures

It appears that the possible outcome of LIN is more restricted when it comes to the linearization of phrasal structures. In particular, it widely assumed that a structure such as (66) can give rise to (at most) two possible surface strings, namely those listed in (67):



(67) a. $\text{spec} > \text{H} > \text{XP}$
b. $\text{spec} > \text{XP} > \text{H}$

Thus, specifiers are taken to be uniformly to the left, while parametric variation is confined to the ordering relation between the head and its complement (at least in approaches that do not assume Kayne's Universal Base Hypothesis, cf. e.g. Haider 1993, Epstein et al. 1998, Richards 2004). This clearly differs from the situation in head adjunction structures where the hierarchically highest element can be linearized either to right or to the left of the existing string of exponents. Furthermore, work in language typology has revealed that there seem to be further ordering relations that

particular inflectional category (such as present vs. part tense forms of T) do not behave differently with respect to their ordering relative to other inflectional formatives. Thus, I assume that there is a default setting of the Head Parameter in the grammar (possibly in the form of a redundancy rule). Furthermore, the existence of exceptions to the rule (e.g., the small number of postpositions in German) suggests that individual Vocabulary items may be endowed with a fixed lexical value for the Head Parameter which may override the default value (cf. Hale 1996: 120f. for some discussion).

surprisingly do not show up cross-linguistically (cf. e.g. Greenberg 1963, Dryer 1992). Below I will discuss two of these ordering possibilities that are apparently absent from the sample, namely *VO-Aux (Steele 1975, Dryer 1992, Travis 1985, den Besten 1986, Kiparsky 1996, among many others), and [_{CP} Comp ...]-V (i.e., the impossibility of combining a complementizer-initial complement clause with a head-final V, cf. e.g. Bayer 1999, Inaba 2007). So, if we want to maintain the claim (see above) that the workings of LIN are basically the same for X^0 - and XP-structures, the apparent differences concerning the range of possible linear orderings must be attributed to independent (structural) asymmetries between head-adjunction structures and phrasal structures.

First of all, note that the existence of an additional projection level that contains the specifier represents a complication which does not exist for the computation of affix order in head adjunction structures. It seems likely that this phrase-structural difference is responsible for the special status of specifiers with respect to linearization, a hypothesis which we will explore in some more detail shortly.

Another crucial difference between head-adjunction structures and phrasal structures has to do with the assumption of cyclic Spell-Out. More to the point, while head-adjunction structures are always part of the same Spell-Out domain (and the phonological domain created from parts of two consecutive Spell-Out domains), the linearization of phrasal structures has to take into account material that is part of different Spell-Out domains (and, accordingly, phonological domains). Below, I am going to argue that at least some generalizations on non-existing word orders can be derived from the way different phonological domains are ordered relative to each other by the workings of LIN.

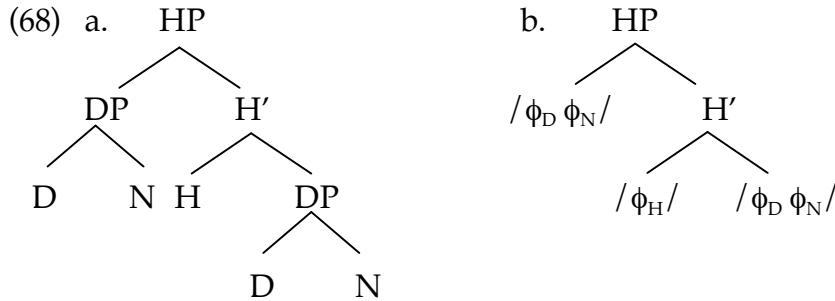
Let us first take a closer look at the way word order is established in the mapping from syntax to phonology and the special role of specifiers in this process. Above, I have already argued that the Head Parameter is confined to lexical properties of individual Vocabulary items (i.e., phonological exponents of functional heads). The relevant lexical properties are imposed on the existing string of exponents when the relevant exponent ϕ of a (functional) head is inserted into the structure as part of the operation of Vocabulary Insertion, establishing an ordering relation between ϕ and its sister (i.e., the exponents corresponding to the complement of a head). Under these assumptions, we can observe that typically, specifiers are not

heads, and therefore may not determine linear order.⁴⁹ Likewise, the sister category of a specifier (traditional referred to as an X' -category) is no head either. In other words, the structure dominated by X' has been subject to previous applications of Vocabulary Insertion, but crucially, the X' -node itself may not be subject to Vocabulary Insertion since it is no terminal node. Thus, we arrive at a point where neither element of the relevant sister pair is capable of determining the linear order of elements. So it appears that in this case, LIN has to resort to other mechanisms in order to establish an ordering relation between a specifier and its X' -sister (i.e., the material dominated by X'). Relevant proposals in the literature (which have already been mentioned above) include the notion of shape conservation (Richards 2004, 2007), or the requirement that movement must have a visible effect on the output (Epstein et al. 1998). However, note that these constraints only warrant leftward orientation of specifiers created by Move/internal Merge (abstracting away from the problems already mentioned). Alternatively, we might attribute the special behavior of specifiers to a property of UG which determines linear order in the absence of other deciding factors (e.g. one might speculate that by default, LIN extends the existing string of phonological exponents only to the left). In what follows, however, I am going to explore another option which capitalizes on the structural differences (i.e., the existence of an intermediate projection level) between head adjunction structures and phrasal units.

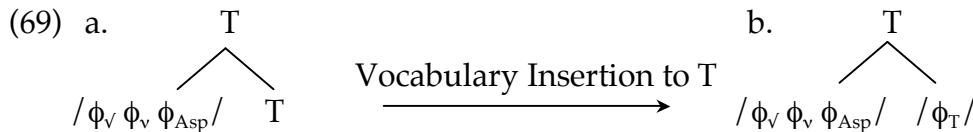
First, note that it seems likely that the insertion of phonological exponents targeting specifiers operates separately from Vocabulary Insertion operating on the main path of embedding, similar to principles governing the constructions of metrical grids (cf. Cinque 1993: 269f.). In other words, the insertion of exponents into terminal nodes that are part of the specifier proceeds in parallel with Vocabulary insertion on the main path of embedding (i.e., head and complement position). Given a syntactic tree such as (68a), Vocabulary insertion leads to the structural

⁴⁹ In particular, the exponents of terminals located in specifiers may impose ordering relations on their sisters specifier-internally (e.g., the exponent of D on the exponent of N), but not on the X' -sister of the specifier position itself. This has to do with the fact that in many cases, the specifier position does not correspond to a terminal node. Accordingly, the node directly corresponding to the specifier is not subject to Vocabulary Insertion and therefore cannot impose an ordering relation on its X' -sister. Moreover, we might suppose that LIN has access to the information of whether a category is minimal or maximal, in the sense that only minimal categories (i.e., projecting terminal nodes) may determine an ordering relation relative to their sister. This would also cover examples where the specifier corresponds to a terminal node as is presumably the case with clitics, which are minimal (since they are non-complex), but also maximal since they do not project.

configuration in (68b) from which the relative order of exponents in the specifier position and within H' must be computed:



If we compare (68b) with the relevant configuration derived by successive application of Vocabulary Insertion to a head adjunction structure (repeated here for convenience), we can perceive a significant asymmetry:⁵⁰

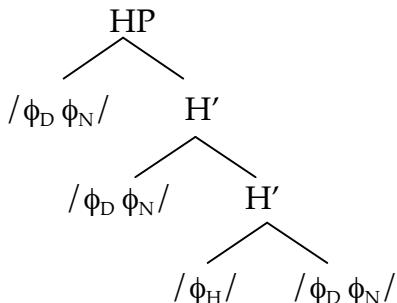


At the point where Vocabulary Insertion applies to T in (69), the insertion site and the existing string of phonological exponents are sisters. In other words, we are dealing with the typical configuration where symmetric c-command holds between two elements, and we thus expect the possibility of parametrization (i.e., any single language may choose to ignore a subset of the corresponding c-command/ordering relations). In contrast, at the point where LIN has to determine the linear ordering of material included in the specifier and material which has been inserted below H' in (68b), the two sets of phonological exponents are not sisters, due to the intervening

⁵⁰ Note that the decisive difference between head complexes and phrasal units depends on the way Vocabulary Insertion removes the hierarchical structure created in the syntax. Thus, I assume that inserting an exponent into a head adjunction structure serves to eliminate the relevant syntactic nodes, including a upper segment of an adjunction structure (e.g. Asp^0 in (69) prior to insertion to T). In contrast, inserting material to H and its complement leaves H' intact up to the point when the linear position of the specifier is determined. This asymmetry can perhaps be attributed to the fact that in the case of head adjunction, the upper segment is treated as being basically identical to the insertion site. In other words, the upper segment has no life on its own for the purposes of the post-syntactic computation. Note that this is required on independent grounds to ensure that the settings of the Head Parameter that determine the linear order of a head and its XP-complement are visible at the top of a head complex (e.g., $\langle T, vP \rangle$ in (69)). In contrast, X' -nodes are syntactic terms on their own (cf. Chomsky 1995: 247), which are left unaffected by Vocabulary Insertion targeting nodes included in X' .

intermediate H' category. As a result, the material in the specifier position asymmetrically c-commands all other elements inserted so far.⁵¹ If c-command relations are taken to map into precedence relations, we may suppose that configurations such as (68b) are uniformly converted into a linear ordering where exponents inserted under the specifier precede all other elements inserted up to this point. So what I am claiming is that in a situation where linear order cannot be determined by lexical properties of individual Vocabulary Items, the relevant ordering information is provided by the structure in terms of asymmetric c-command relations. Note that this also serves to determine the relative order of elements in multiple specifier configurations. Again assuming that Vocabulary Insertion to specifiers proceeds in parallel to Vocabulary Insertion to other nodes, we derive the configuration in (70), in which the inner specifier asymmetrically c-commands the phonological exponents of head and complement while the outer specifier asymmetrically c-commands the exponents of the inner specifier and all other elements. Accordingly, we expect that material in the highest specifier is realized at the leftmost position, as desired.⁵²

(70)



Summing up, it seems that it is possible to derive the above-noted generalizations on possible language specific choices with respect to the Head Parameter. While affix order and the linear order of head and complement are subject to parametrization, material in specifier positions uniformly precedes the remainder of a phrase. More specifically, we have seen that the presence of an intermediate projection level (which is absent in head adjunction structures) creates a structural asymmetry between exponents inserted to the specifier and exponents inserted to nodes

⁵¹ Recall that I assume that H' cannot determine the linear ordering of the relevant sets of exponents (in the specifier and under H' , respectively) since it is not subject to Vocabulary Insertion. Furthermore, note that it is commonly assumed that X' -categories does not c-command since they are neither maximal nor minimal (cf. e.g. Chomsky 1995: 336).

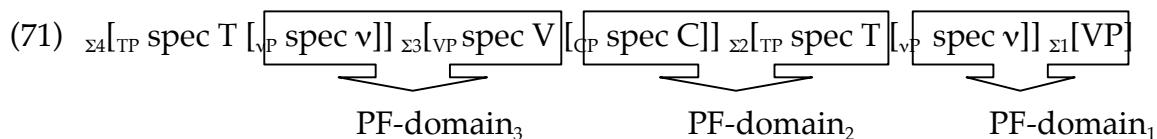
⁵² Note that XP-adjunction possibly raises a couple of further issues, which I cannot address here in detail.

dominated by X' , which is uniformly mapped into a precedence relation by the workings of LIN (as part of the process of Vocabulary Insertion).

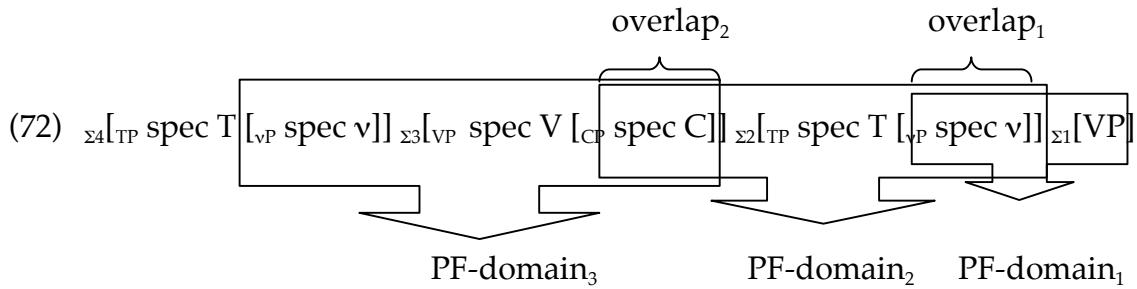
So far, we have only focused on the linearization of material that is part of a single phonological domain. Next, I am going to address the question of how the workings of Vocabulary Insertion create linear orderings between different phonological domains.

4.2.3 LIN and the relative order of phonological domains

The general approach developed so far requires that all post-syntactic operations including Vocabulary Insertion and LIN operate on phonological domains constructed from the cyclic output of the syntactic derivation. This raises the obvious question of how material spelled-out in a previous cycle is linearized relative to material that is spelled-out later (see also Dobashi 2003 for discussion). This becomes immediately clear if we take another look at a structure with non-overlapping PF-domains discussed in section 3.3 above:



If LIN applies cyclically to the output of the syntactic computation, then it appears that after linearization of the first PF-domain, the relevant object is not any longer available to future applications of LIN. That is, the object corresponding to PF-domain₁ cannot be linearized with the object corresponding to PF-domain₂ etc. Of course, this is not the desired result. To solve this problem, Dobashi (2003) assumes that after LIN has applied to a given Spell-Out domain, the leftmost constituent of this domain is not directly mapped to the interfaces, but 'left behind' to be accessible for future applications of LIN. In this way, the relevant constituent serves as a connection between successive Spell-Out domains, ensuring that a linear ordering can be established between material that is part of different Spell-Out domains. Note that this outcome follows more or less directly if we adopt the notion of overlapping phonological domains proposed in section 3.3. The domains for the application of LIN are then as follows:



As illustrated in (72), the left edge of a phonological domain PFD_n overlaps with the right edge of a subsequent domain PFD_{n+1} . The overlap can be exploited to establish a linear ordering between separate phonological domains if we assume that the edge of a lower phonological domain is linearized twice. In (72), material at overlap_1 is first linearized in relation to material that is part of the complement domain of v , and, on a later cycle, in relation to material inserted to T (or higher positions). In this way, the edge of a phonological domain PFD_n provides a connection between two successive phonological domains PFD_n and PFD_{n+1} . In what follows, the workings of this procedure (which by assumption is part of LIN) are illustrated with different settings of the Head Parameter for T and v , which in principle may give rise to the following four logically possible grammars:

(73) a. TP head-initial, vP head-initial (Aux-VO)
 b. TP head-initial, vP head-final (Aux-OV)
 c. TP head-final, vP head-final (OV-Aux)
 d. TP head-final, vP head-initial (VO-Aux)

However, it is a well-known fact that of these four parametric options, only (73a) and (73c) are robustly attested across the world's languages. While (73b) seems to be a possible (albeit rare) option (e.g., Old English, cf. e.g. Pintzuk 1999, 2005, and section 6.2 below), many researchers have claimed that the remaining option, (73d) is absent cross-linguistically (cf. Steele 1975, Travis 1985, Hawkins 1990, Dryer 1992, Holmberg 2000, Fuß and Trips 2002, Pintzuk 2005).⁵³ Recently, this issue has been taken up by

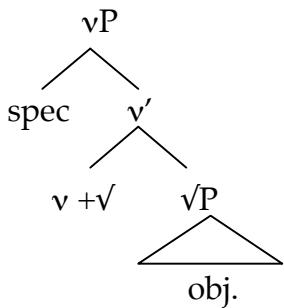
⁵³ There are in fact some languages that apparently exhibit S-V-O-AUX orders. A case in point seems to be Kandoka-Lusi, a dialect of Kaliai-Kove, an Austronesian language spoken on the northern coast of Western New Britain, described by Counts (1969):

(i) $\eta a\text{-} \beta\alpha\text{-} \text{pater} \text{ mu}\gamma\text{a}$.
 I ask priest will
 'I will ask the priest.'
 (Counts 1969: 130)

Biberauer et al. (2007), (2008), who develop an account of the illicitness of VO-Aux orders based on Holmbergs (2000) *Final-Over-Final Constraint* (FOFC) and Kayne's (1994) LCA (see section 6.1.1 for critical discussion). In what follows, I am going to argue that the conspicuous absence of (73d) can be attributed to the way successive phonological domains are put together by the workings of Vocabulary Insertion/LIN. Basically, this is because linearization may not disrupt adjacency between (head-initial) V and Aux when phonological domains are combined by a process in which material at the overlap is replaced by the string of phonological exponents established previously. In other words, I propose that (73d) is absent from the record since it cannot result from the way hierarchical structures are mapped to linear orders in the PF-branch of grammar.

For expository reasons, let's adopt the following (simplifying) assumptions: (i) the verbal root is pronounced together with v , and (ii) the verb and its arguments are pronounced within vP . In (72), LIN applies first to PF-domain₁, establishing a linear ordering between v and material inserted to its complement domain \sqrt{P} (based on lexical properties of the Vocabulary item realizing v). Next, if overtly realized, phonological material is added to the specifier of vP , creating the following orders for simple transitive clauses in SOV and SVO languages, respectively:⁵⁴

(74)

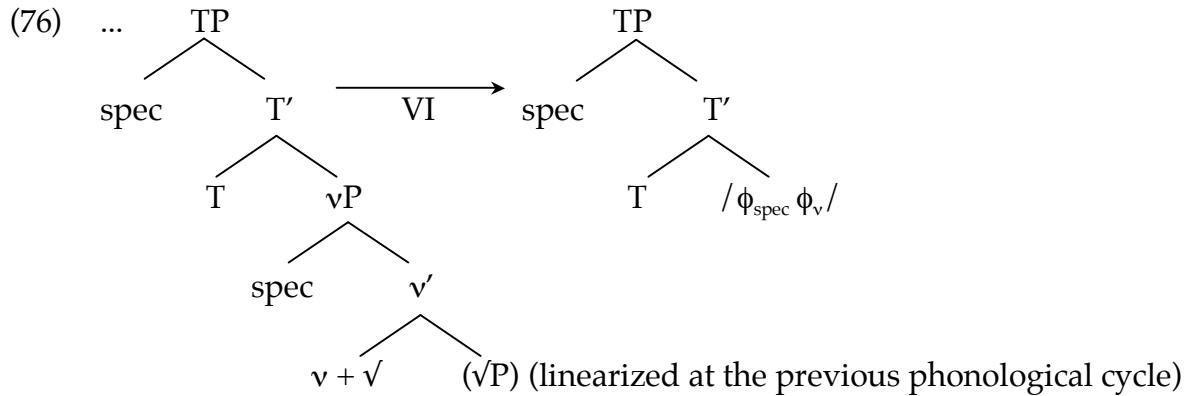


(75) a. / $\phi_{\text{spec}} \phi_v \phi_{\text{object}}$ / SVO
 b. / $\phi_{\text{spec}} \phi_{\text{object}} \phi_v$ / SOV

However, it is not clear at all whether the auxiliary-like elements that appear in clause-final position are really verbal elements; apparently, they do not agree with the subject and show no other signs of finite inflection. Nonetheless, more research is certainly necessary to settle the status of these apparent counterexamples to the generalization in question.

⁵⁴ Note that at first sight, the configuration in (74) seems to lead to us to expect that the $v + \sqrt{}$ complex should always precede the object since the latter is asymmetrically c-commanded by the former. Below, it will be argued that this is only the case in VO grammars, while in OV grammars, the verbal root does not move to v .

On the subsequent cycle, LIN applies to PF-domain₂. Note that under the assumption of overlapping phonological domains, Vocabulary Insertion (proceeding in an bottom-up fashion) again targets *v* and its specifier before *T* is supplied with phonological information (at the moment abstracting away from Vocabulary Insertion to the CP that is also part of PF-domain₂). Due to the fact that the specifier asymmetrically c-commands the head (see above), the exponent(s) linked to SpecvP uniformly precede the exponent of *v*:



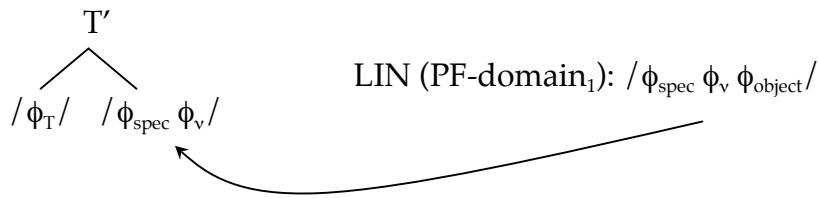
Next, Vocabulary Insertion targets *T* (e.g., modals in English). Dependent on the setting of the Head Parameter for *T*, the exponents of specifier and head of *vP* will then appear either to the left or to the right of the exponent of *T*:



Note that this is the crucial step for establishing an ordering relation between PF-domain₁ and PF-domain₂. As already noted, since SpecvP and *v* are present in both phonological domains, the exponents inserted to these positions effectively provide a connection between two successive phonological domains. More to the point, suppose that at the point when a linear ordering has been established between the elements at the overlap and the phonological exponent of *T*, LIN linearizes PF-domain₁ with PF-domain₂ via replacing the exponents at the right edge of PF-domain₂ with the string of exponents assembled by Vocabulary Insertion to PF-domain₁.⁵⁵ This is illustrated in (78) for a uniformly head-initial grammar:

⁵⁵ Alternatively, we might assume that the mechanism replacing the overlap with the string of exponents linked to PF-domain₁ takes place upon completion of Vocabulary Insertion to PF-domain₂. However, this would require that the replacement procedure be capable of manipulating the string of

(78)



In (78), the relevant operation replaces the overlap ($/\phi_{\text{spec}} \phi_v/$) with the string of exponents inserted to PF-domain₁ ($/\phi_{\text{spec}} \phi_v \phi_{\text{object}}/$). In other words, a linear ordering of PF-domain₁ and PF-domain₂ is established via (i) inspecting the order of the exponents inserted to the overlap relative to the exponent of T (cf. (77)), and (ii) replacing the right edge of PF-domain₂ with the string of exponents assembled earlier:

(79) *Edge Replacement*

After the exponents of the overlapping part of structure have been linearized relative to material in PF-domain_n, they are replaced with the string of exponents assembled at PF-domain_{n-1}.

In case no phonological exponent is inserted to T (as is presumably the case with finite lexical verbs in English), the relevant ordering between PF-domain₁ and PF-domain₂ is determined when Vocabulary Insertion supplies SpecTP with phonological exponents, which will uniformly precede all other exponents inserted so far.

I assume that the replacement mechanism illustrated in (77) is subject to a no-tampering restriction, in the sense that it may not interfere with ordering relations created between the exponent of T and the exponent (at the overlap) directly adjacent to (the exponent of) T. In somewhat more general terms, this can be phrased as follows:

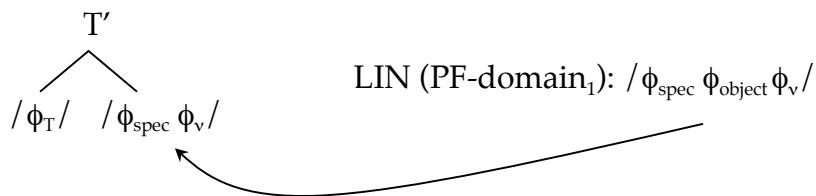
exponents which has been assembled in the course of linearizing PF-domain₂, which raises a number of further issues.

(80) *No-Tampering Condition on Edge Replacement*⁵⁶

Edge Replacement at PF-domain_n may not disrupt ordering relations created between material of PF-domain_n and material at the (overlapping) left edge of PF-domain_{n-1}.

The replacement process illustrated in (78) clearly satisfies this condition, since it leaves the ordering relations established earlier fully intact (i.e., after replacing $/\phi_{\text{spec}} \phi_v/$ with $/\phi_{\text{spec}} \phi_v \phi_{\text{object}}/$, ϕ_T continues to directly left-adjacent to ϕ_{spec}). What happens if we combine a head-final vP with a head-initial TP? As indicated in (81), this operation does not raise any problems either, since it does not disrupt the linear adjacency of ϕ_T and ϕ_{spec} created by application of LIN at PF-domain₂.

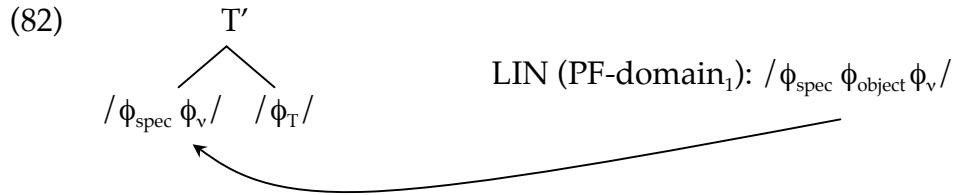
(81)



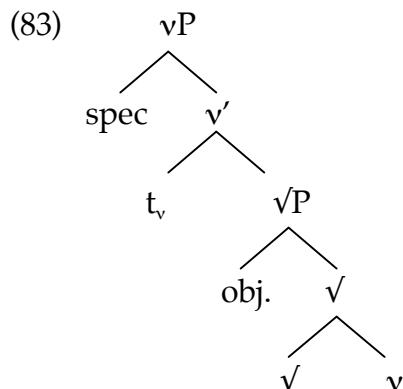
However, note that a problem will arise if the left edge of vP contains only the phase head and no specifier (e.g., after subject movement to SpecTP). See section 6 below for a discussion of relevant orders in Old English (i.e., Aux-OV). Below, I am going to argue that the configuration at hand involves Morphological Merger of v and √ in PF-domain₁, which lowers v to √ prior to Vocabulary Insertion (an operation which by assumption is a characteristic of OV-grammars).

Next, let's address the case of head-final T. If the elements at the overlap are placed to the left of the exponent of T, the element immediately left-adjacent to T is not the exponent of SpecvP, but rather the exponent of v. Accordingly, replacement of the relevant string of elements by the phonological exponents linked to PF-domain₁ must preserve adjacency of ϕ_T and ϕ_v . This can be achieved if the vP is head-final as well, giving rise to the ordering characteristics of uniformly head-final languages as illustrated in (82):

⁵⁶ Not to be confused with the “no tampering condition” proposed in Chomsky (2005: 5) which basically states that syntactic operations such as Merge do not affect the feature content of syntactic objects (i.e., do not add new features etc.).



Again, this operation does not raise an issue with respect to the No-Tampering Condition (80), since ϕ_v continues to be directly left-adjacent to ϕ_T after *Edge Replacement*. This still holds true if we adopt the assumption that it is a general characteristic of OV grammars that they lack verb movement (or, rather, root movement) to inflectional heads (at least in non-V2 contexts, cf. Haider 1993, 2000a; Kiparsky 1996, Vikner 2001, Julien 2002; see e.g. Kuroda 1988, Saito 1992, Fukui and Takano 1998, Saito and Fukui 1998 on Japanese). Obviously, this raises the question of how the category-neutral root can combine with a category defining v-head in a OV grammar. What I want to propose is that in a basic OV grammar, the relevant head adjunction structure is accomplished via Morphological Merger, that is, post-syntactic lowering of v that applies to PF-domain₁ and forms a v+v complex prior to Vocabulary Insertion:



As a result, the original position of v (sister of vP) is not any longer accessible for Vocabulary Insertion in PF-domain₁. Instead, Vocabulary Insertion targets v as part of the v+v complex, which also serves to establish a linear ordering between the v+v complex and material inserted to its complement position (recall that we have tentatively assumed that the Head Parameter is expressed by exponents of functional categories only). The structure in (83) also serves to overcome the problem mentioned in fn. 54, where we have noticed that we would perhaps expect v to asymmetrically c-command the object if the latter remains in a vP-internal position, giving rise to an ordering where the exponent of v uniformly precedes the object (and more generally all elements in vP).

Under these assumptions, the difference between VO and OV grammars boils down to a parameter that determines how roots combine with the category-defining head v : In VO-grammars, the root moves up to v (associated with a head-initial setting of the Head Parameter), ensuring that the verb appears to the left of all its internal arguments and low event-related adverbs (cf. Larson 1988, Stroik 1990, Haider 1993, 2000b, Hinterhölzl 2001, and many others). In OV grammars, however, v and \sqrt{v} are combined via post-syntactic Morphological Merger, which lowers v to the position of \sqrt{v} . As a result, direct objects (and all other $\sqrt{v}P$ -internal elements, including event-related adverbs) precede the $\sqrt{v}+v$ complex (if the exponent of v is associated with an OV-setting of the Head Parameter; see sections 5 and 6 for further discussion of OV-related properties, focusing on German and Old English). I thus propose the following parameter which derives the basic differences between OV and VO grammars:⁵⁷

(84) *The Root Raising Parameter*

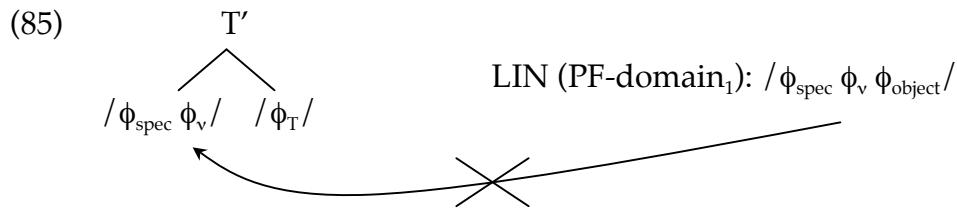
\sqrt{v} raises to category-defining v :

YES (VO-languages: English, Italian, French etc.)

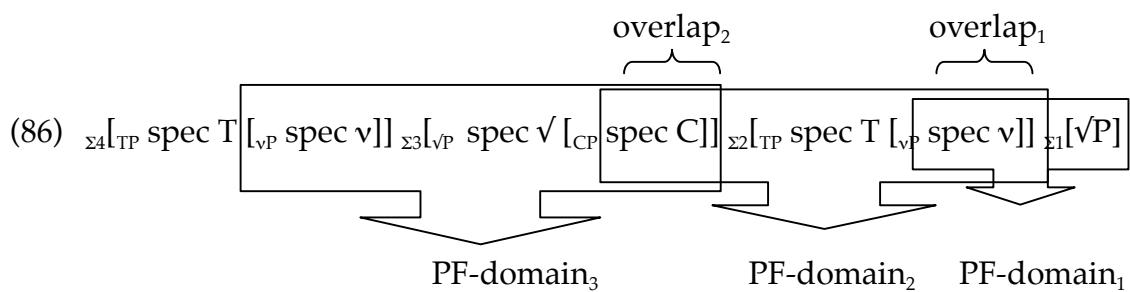
NO (OV-languages: Japanese, German, Hindi etc.)

Finally, let's take a look at the last remaining option, namely a grammar that combines a head-final TP with a head-initial vP . Again, the exponent of v is immediately left-adjacent to the exponent of T after LIN/Vocabulary Insertion has applied to PF-domain₂. By assumption, this ordering relation cannot be tampered with by subsequent workings of LIN. It follows that the previously assembled string of elements that realizes PF-domain₁ ($/\phi_{\text{spec}} \phi_v \phi_{\text{object}}/$) cannot replace the overlapping string $\phi_{\text{spec}} \phi_v/$ in PF-domain₂, since this operation would disrupt adjacency of ϕ_v and ϕ_T :

⁵⁷ Note that the combination of (i) a VO-setting of the Head Parameter for v , and (ii) the lowering option would give rise to a grammar where only direct objects follow the verb, while all other arguments precede the verb. While there are apparently some languages where arguments appear on both sides of the verb, this seems to be a highly marked option. In other words, the vast majority of the world's languages seem to adopt either a positive setting of the *Root Raising Parameter* (giving rise to VO orders), or a combination of a negative setting of this parameter and a OV-setting for the Head-Parameter linked to v (see section 6 for further discussion).



Thus, it seems that we can attribute the apparent non-existence of the combination of a head-final TP with a head-initial vP to properties of the post-syntactic computation, in the sense that a combination of these two parametric choices leads to structures that cannot be linearized by the workings of LIN/Vocabulary Insertion. Note that this effect is restricted to the overlap of the two successive phonological domains, where *Edge Replacement* takes place, that is, clause-internally at the junction of TP and vP, and cross-clausal at the junction of (embedded) CP and (matrix) vP:



Thus, we expect that within a single phonological domain, it is in fact possible that a head-initial projection is embedded under a head-final projection. This seems to give the correct results for a language like German, where NPs/DPs (which are head-initial) may be dominated by VP, which is head-final:⁵⁸

(87) dass Peter [vP [DP die Löschung [der Dateien]] empfohlen] hat
 that Peter the deletion of-the files recommended has
 'that Peter recommended the deletion of the files'

In a similar vein, we predict that head-initial NPs/DPs (or PPs) may be embedded by the small class of pospositions in German:

(88) [PP [DP den höchsten Gipfel [des Gebirges]] hinauf]
 the highest peak of-the mountains up
 'up the highest peak of the mountains'

⁵⁸ Note that this requires that we assume that DPs/NPs do constitute phases on their own.

4.3 Section summary

In this section I have outlined a model for linearizing the cyclic output of the syntactic computation which is based on the following assumptions: (i) The mapping from hierarchic structures to linear orders of phonological exponents is a function of the process of Vocabulary Insertion, which provides syntactic terminal nodes with phonological realizations (the so-called *Late Linearization Hypothesis*, Embick and Noyer 2001); (ii) the linearization procedure applies to individual phonological domains constructed cyclically from the output of the syntactic derivation; (iii) the Head Parameter is phonological in nature, ignoring a subset of symmetric c-command relations created in the syntax. The Head Parameter is confined to exponents of functional categories; (iv) the existence of an intermediate projection level (X') creates an asymmetry between X^0 -structures and XP -structures to the effect that specifiers asymmetrically c-command and precede insertion sites on the main path of embedding (i.e., within X'); (v) the relative order between successive phonological domains can be determined by appealing to the notion of overlapping phonological domains developed in section 3.3. More precisely, I have proposed that material that is part of two successive phonological domains effectively provides a connection between these domains since the relevant exponents are linearized both (a) relative to material in the higher domain and (b) relative to material in the lower domain. The separate phonological domains are then combined by a process called *Edge Replacement* which substitutes the right edge of $PF\text{-domain}_n$ with the string of exponents realizing $PF\text{-domain}_{n-1}$. I have then argued that the process of *Edge Replacement* is subject to a non-tampering condition requiring *Edge Replacement* to preserve adjacency relations established previously between exponents of the higher $PF\text{-domain}_n$ and the exponents at the overlap.⁵⁹ This restriction rules out certain

⁵⁹ Obviously, this rather sketchy outline of the workings of *Edge Replacement* raises a number of further issues which I cannot address here in full (but see sections 5 and 6 for further discussion). Among the most obvious questions are: (i) How does *Edge Replacement* deal with (intermediate/lower) copies of moved elements that fail to be pronounced at the edge? (ii) What happens in cases where the edge does not contain any visible material? Concerning the latter, we may assume that *Edge Replacement* takes place in any case (creating a linear order between two consecutive domains), independent of whether overt material is inserted into the relevant positions at the edge or not. Note that the question raised by copies of moved elements pertains to all approaches that adopt the notion of multiple Spell-Out. Following Richards (2004, 2007), we might assume that elements at the edge of intermediate phases contain uninterpretable features that render them inaccessible for Vocabulary Insertion (but see section 4.1 above for discussions of some problems linked to this assumption). Still, clearly more has to be said here; I leave this issue for future research.

configurations such as the apparently non-attested combination of a head-final TP with a head-initial vP. In addition, I have proposed that a major parametrical difference between OV and VO grammars concerns the availability of raising the (verbal) root to the category-defining v-head, which obligatorily takes place in VO grammars, but is supplanted by post-syntactic v-lowering in OV grammars.

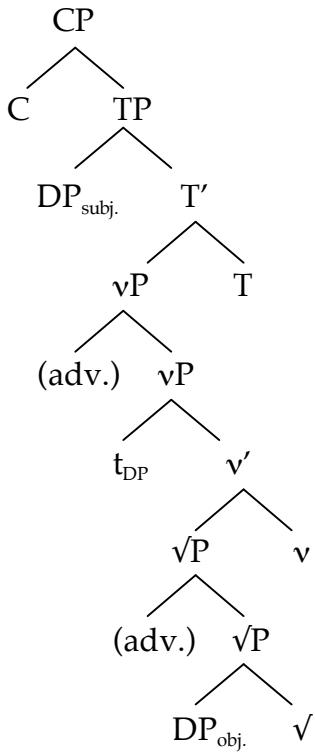
The following sections illustrate the workings of this model of linearization in some more detail, focusing on German and some aspects of the OV-VO change in the history of English. I show that an approach in terms of cyclic linearization makes available new explanations for a set of recalcitrant word order phenomena, which lends further support to this particular account.

5 The distribution of sentential complements in German

As is well-known, German exhibits a major asymmetry between main and embedded clauses. While the former exhibit obligatory verb fronting, giving rise to V2 order in declaratives and *wh*-questions, the latter exhibit a basic SOV order when introduced by a subordinating conjunction, a *wh*-phrase (in indirect questions), or a relative pronoun. In what follows, I will take a closer look at some word order phenomena in embedded clauses from the perspective of cyclic linearization, focusing on the syntactic distribution of finite embedded clauses. First, however, let me add some more general remarks on the analysis of a basic SOV language like German in the framework proposed here. If we adopt the assumption that SOV languages lack verb movement to higher inflectional heads, the structure of an embedded clause looks as follows (with adjunction sites for low adverbs to the left of vP and vP):⁶⁰

⁶⁰ Here and in what follows, the relevant phrase markers for German are depicted with head-final TP, vP, and vP. Note that this is for expository purposes only and should not be taken to imply the existence of a *syntactic* Head Parameter. Recall that by assumption, the syntactic representation contains only hierarchical information.

(89)



Assuming that German lacks verb movement while VO-languages like English must resort to \sqrt -to- v raising (the *Root Raising Parameter*, see above) captures the different placement properties of low event-related adverbs in these languages (as is well-known, time, place and manner adverbs follow the verb in English, but precede the verb in German in the exact mirror image, cf. Haider 2000b, Hinterhölzl 2002, 2004).⁶¹ Accordingly, finite verbs in embedded clauses are the result of a post-syntactic operation that combines the relevant inflectional features and the verbal root (in main clauses, however, the finite verb undergoes syntactic fronting, picking up the relevant inflectional features on its way to C). However, while v can be joined with \sqrt via Morphological Merger, this seems to be impossible for T (containing tense and agreement information) and the $v+\sqrt$ complex, since T and the $v+\sqrt$ complex are part of different PF-domains after v -lowering. A way out of this dilemma is to assume that after Vocabulary Insertion has assembled the string of exponents, the exponents

⁶¹ Note that the analysis of these adverbs raises some further issues which I cannot address here in detail. Apart from the difficulties caused by the fact that the order of the relevant adverbs is different in English and German (German: Time-Place-Manner; English: Manner-Place-Time), it must be assumed that in English, a further movement operation takes place that puts the object to the left of these low adverbs (possibly targeting an outer Spec/P along lines proposed in Chomsky 2005). While this operation is obligatory in English, leftward movement of the object may be triggered in German by information-structural factors which preferably place given information at the beginning of the *Mittelfeld* 'midfield'. See Haider (1993), (2000c), Hinterhölzl (2002), and Grewendorf (2005) for detailed discussion.

of T and the $v+v'$ complex can amalgamate (giving rise to an inflected verb), since they are always string-adjacent in an OV grammar (for similar considerations cf. Julien 2002; see Bobaljik 1995, 2003, Adger 2003: 170 and Wiklund 2005 for discussion and further examples of post-syntactic T-lowering / copying).⁶²

The exceptional behavior of finite sentential complements with respect to the basic OV character of German is a long-standing puzzle of German syntax (cf. e.g. Grewendorf 1988, Webelhuth 1992, Haider 1993, 1995, 1997, Büring 1995, Müller 1995, 1997, Bayer 1996, Büring and Hartmann 1997, and most recently Inaba 2007). While nominal and prepositional complements precede their selecting verb in basic order, finite sentential complements uniformly appear in a postverbal position giving rise to apparent instances of VO order:

(90) a. dass der Peter [_{DP} den Roman] gelesen hat
 that the Peter the novel read has
 'that Peter read the novel'
 b. dass der Peter [_{PP} auf Maria] wartet
 that the Peter on Maria waits
 'that Peter waits for Maria'

(91) a. ?? dass Klaus, [_{CP} dass Peter den Roman gelesen hat], glaubt
 that Klaus that Peter the novel read has thinks
 'that Klaus thinks, that Peter read the novel'
 b. dass Klaus glaubt, [_{CP} dass Peter den Roman gelesen hat]

In generative approaches, the final position of finite complement clauses is often attributed to obligatory rightward movement/extraposition (cf. e.g. Büring 1995, Müller 1995, 1997, Bayer 1996, Büring and Hartmann 1997). However, it has been repeatedly noted in the literature that this approach raises a number of empirical and conceptual problems, which led a number of authors to assume that postverbal finite embedded clauses occupy the verb's complement position (cf. e.g. Webelhuth 1992,

⁶² Alternatively, we may assume that the relevant ϕ -features (i.e., tense and agreement) are not part of T in an OV-grammar, but rather included in the category-defining functional head v (for related proposals cf. Haider 1993, Kiparsky 1996, and Sternefeld 2007). Note that this problem does not arise with periphrastic verb forms, where the auxiliary corresponds to a higher v -head that selects another vP (containing the lexical verb and its arguments) as its complement. See section 6.2.2 for some discussion.

Haider 1993, 1995, 1997; see Inaba 2007 for an overview and detailed discussion). In particular, the obligatory final placement of finite argument clauses comes as a surprise since extraposition is typically an optional process in German (as e.g. in the case of relative clauses or heavy PPs):

(92) a. dass Peter auf den Mann, [den Maria liebt], gewartet hat
 that Peter for the man who-ACC Mary loves waited has
 b. dass Peter auf den Mann __ gewartet hat, [den Maria liebt]
 'that Peter waited for the man who Mary loves'

(93) a. dass Peter sich [auf der Feier im Prinzengarten] gelangweilt hat
 that Peter REFL at the party in-the Prinzengarten bored has
 'that Peter was bored at the party in the Prinzengarten'
 b. dass Peter sich __ gelangweilt hat [auf der Feier im Prinzengarten]

Furthermore, extraposed complement clauses are not islands for extraction, which is somewhat surprising if extraposition is analyzed as adjunction to the right of VP or IP, but expected if the clause occupies its base position as a sister of V (but see Büring and Hartmann 1997, Müller 1997 for the claim that extraction precedes extraposition in the relevant examples):

(94) Wen hat Peter behauptet, [dass die Maria __ liebt]?
 who-acc has Peter claimed that the Mary loves
 'Who did Peter claim that Mary loves?'

The fact that extraposed complement clauses are transparent for extraction contrasts with the behavior of scrambled constituents (which according to the standard analysis occupy an adjunction site as well) and base-generated adjuncts, which are typically islands for *wh*-extraction:

(95) a. [Über wen]_i hat der Peter der Maria [eine Geschichte t]_j
 about whom has the.NOM Peter the.DAT Mary a story
 erzählt?
 told
 'Who did Peter tell Mary a story about?'
 b. *[Über wen]_i hat [eine Geschichte t]_j, der Peter der Maria t, erzählt?

(96) a. *Was_i warst du [nachdem du t_i getrunken hast] krank?
 What were you after you drunk has sick
 'What were you sick after drinking ___?'
 b. *Was_i warst du krank [nachdem du t_i getrunken hast]?
 (Büring and Hartmann 1997: 7)

Another piece of evidence that suggests that complement clauses occupy the verb's complement position comes from topicalization data in which the complement clause is fronted together with its selecting verb (Inaba 2007: 42):

(97) a. [Gesagt, [dass er sie liebt]]_i hat er gestern t_i.
 said that he her loves has he yesterday
 'Yesterday, he said that he loves her.'
 b. *[Dass er sie liebt] gesagt] hat er gestern.

(98) a. [Der Polizei [erzählt [dass er dort war]]] hat er nicht.
 the.DAT police told that he there was has he not
 'He didn't tell the police that he was there.'
 b. [Erzählt, [dass er dort war]], hat er der Polizei nicht.
 c. ?[Der Polizei [erzählt t_i]] hat er nicht [dass er dort war]_i.

As illustrated in (97a), it is possible to front a participle together with its postverbal complement clause. (97b) shows that VP-fronting with the complement clause in preverbal position leads to ungrammaticality. (98a) shows that it is even possible to topicalize a VP that contains both a preverbal nominal (dative) argument and a postverbal finite complement clause, while (98b) shows that VP-fronting can strand the dative argument (which presumably has undergone scrambling prior to VP-topicalization). Furthermore, while the complement clause can be stranded by fronting the participle together with the dative argument (cf. (98c)), the resulting configuration is clearly less acceptable than (98a,b). This contrast suggests that the complement clause occupies a structural position that is configurationally closer to the verb than the position of the dative, similar to the following examples which show that VP-topicalization including the direct object but stranding the dative is more acceptable than pied-piping the dative and stranding the direct object. More

generally, it is a well-known fact that VP-fronting in German preferably pied-pipes arguments that are structurally closer to the verb (cf. e.g. Grewendorf 1988: 297):

(99) a. [Das Buch gegeben]_i hat die Maria dem Peter t_i.
 the.ACC book given hat die.NOM Mary the.DAT Peter
 'Mary gave the book to Peter.'
 b. ?[Dem Peter gegeben] hat die Maria das Buch t_i.

If we accept the notion that postverbal finite clauses are sisters to the verb, this of course raises the question of why they may not show up in preverbal position, in contrast to all other types of arguments. In approaches based on Kayne's (1994) LCA such as Zwart (1997), the different behavior of nominal and clausal complements is attributed to their different licensing requirements: While nominal arguments raise overtly to SpecAgroP in order to receive/check Case, no such requirement exists for clausal complements which accordingly may remain in situ. However, note that this approach cannot explain the distribution of prepositional arguments which pair with nominal arguments, despite the fact that they do not need case (for further conceptual problems raised by LCA-based analyses see section 4 above).

Alternative accounts of the distribution of clausal complements in German have been put forward by Bayer (1996, 1999) and most recently Inaba (2007). Bayer assumes that finite complement clauses are base-generated in preverbal position in German and then undergo short "Argument Shift" which places the clause in an A-position minimally to the right of the verb. According to Bayer, this operation is triggered by a general requirement that the head of a complement clause be adjacent to the selecting matrix verb, which also accounts for the generalization that across languages, head-initial CPs appear in postverbal position (in both SVO and SOV languages), while head-final CPs occupy preverbal position (see also Hawkins 1990, Dryer 1992). While this analysis accounts for the facts in German and cross-linguistically, it should be noted that it is quite ad hoc, being based on a number of non-standard assumptions (e.g., rightward syntactic A-movement).

According to the analysis proposed by Inaba (2007), complement clauses are base-generated in postverbal position in German. The special behavior of sentential complements is then attributed to the way phases are linearized in a model assuming multiple Spell-Out. Inaba assumes that LIN operates in a bottom-up fashion, starting with the most deeply embedded phase. The output of LIN is then successively added to a phonological representation that is incrementally built from right to left. As a

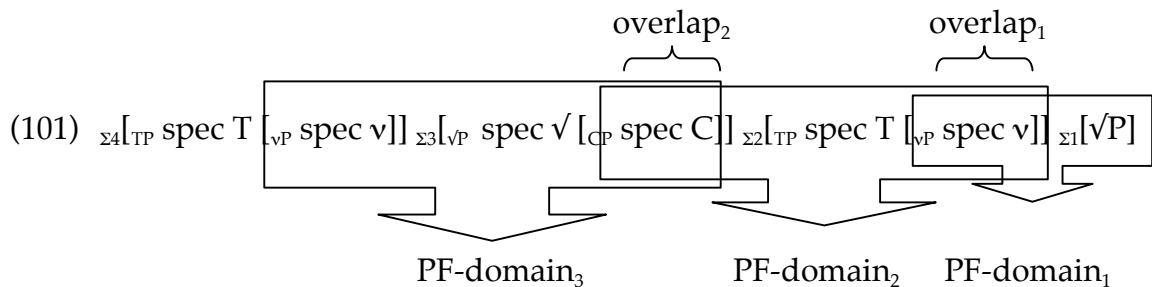
result, phases that are linearized earlier appear more to the right of this representation than phases that are linearized later: $\langle \text{PH}_n \dots \text{PH}_2, \text{PH}_1 \rangle$. This derives the fact that head-initial complement clauses (which are analyzed as separate phases) appear always in final position, since they are the first elements sent to PF. In contrast, nominal complements are linearized together with their selecting verb (since they are part of the same phase), giving rise to OV order (if V is linked to a relevant setting of the Head Parameter). However, while this analysis at first sight provides an elegant explanation of the asymmetry in question, it suffers from a number of shortcomings. First of all, Inaba has to stipulate that complement clauses with head-final complementizers (that show up in preverbal position as e.g. in Japanese) are not CPs (i.e., they are not phases on their own), but rather prepositional or nominal in character (if they were CPs, that is, phases, they would be expected to occur in final position as well). Furthermore, in order to warrant that the left edge of a complement clause is linearized (and sent to PF) together with the embedded clause (and not with the selecting verb), Inaba assumes that the upon completion, the whole phase (and not only the domain of the phase head) is transferred to PF. However, this assumption seems to rule out the possibility of successive-cyclic *wh*-movement via the left edge of a lower phase.

In the following, I am going to develop an account of the placement of finite embedded clauses in German based on the approach to linearization developed in section 4.2. Similar to Inaba (2007), I assume that complement clauses occupy a structural object position in the syntax (that is, they are sisters of $\sqrt{}$), and that their ultimate placement results from the way syntactic structures are mapped to linear orders in the PF-branch of grammar. However, in contrast to Inaba, I am going to argue that the postverbal position of head-initial complement clauses in German (and in general, cross-linguistically) is determined by the way consecutive phonological domains are linked (and linearized) by the process of *Edge Replacement*. More precisely, I assume that the problematic ordering is ruled out by the same properties of LIN that exclude the possibility of VO-Aux orders, that is, the combination of a head-initial vP embedded by a head-final TP. The relevant condition on *Edge Replacement* is repeated here for convenience:

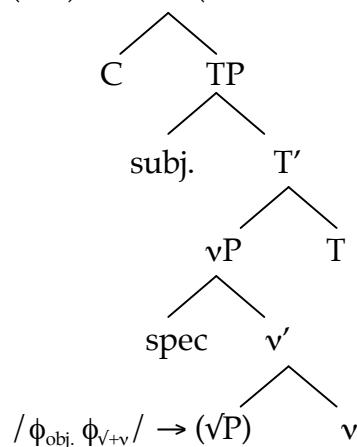
(100) No-Tampering Condition on Edge Replacement

Edge Replacement at PF-domain_n may not disrupt ordering relations created between material of PF-domain_n and material at the (overlapping) left edge of PF-domain_{n-1}.

To see how this works, let us first consider again the relevant phonological domains and the overlap between them that serves to establish the linear order between the selecting root and its complement clause:

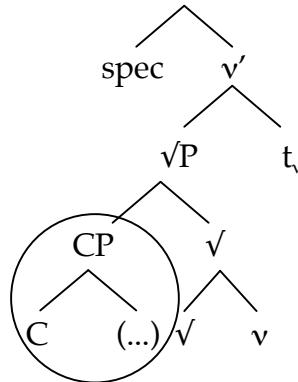


Here, the relevant part of the structure is the area where PF-domain₃ and PF-domain₂ overlap, that is, the left edge of the embedded CP. In the case at hand, the left edge of CP contains only the complementizer, so the relevant chunk of structure that is linearized at PF-domain₂ looks as in (102), giving rise to the string of phonological exponents in (103) if we assume that exponents of C are head-initial in German, while T and v are head-final (with v lowered to \sqrt{v} and therefore part of \sqrt{P} , which has been connected with PF-domain₂ by a previous application of *Edge Replacement*, see above).

(102) CP (=PF-domain₂)(103) /phi_C phi_{subj.} phi_{obj.} phi_{v+v} phi_T/

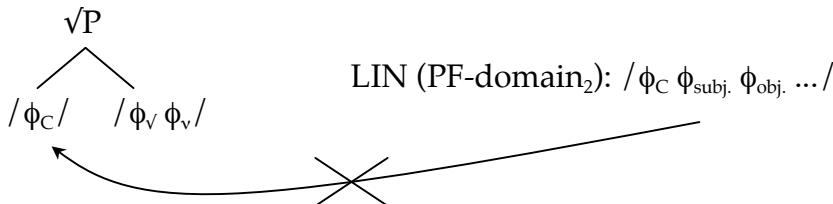
When the next higher PF-domain₃ (compare (104)) is subject to Vocabulary Insertion, the linear ordering of PF-domain₃ and PF-domain₂ is determined via application of *Edge Replacement* that substitutes the overlapping right edge of PF-domain₃ (marked by a circle in (104)) with the string of exponents in (103).

(104) \sqrt{P} (=PF-domain₃)



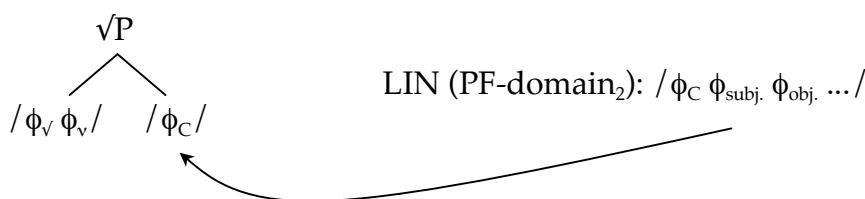
However, it turns out that this procedure violates the *No-Tampering Condition on Edge Replacement* if the relevant string of exponents is inserted into a position to the left of $/\phi_v \phi_v/$. As illustrated in (105), *Edge Replacement* disrupts the linear ordering of $/\phi_C/$ and $/\phi_v \phi_v/$, which has been established previously:

(105)



Thus, embedded head-initial finite CPs can be linearized in an OV grammar only if the Head Parameter of the selecting $\sqrt{v}+v$ complex is ignored, giving rise to postverbal placement of the complement clause:

(106)

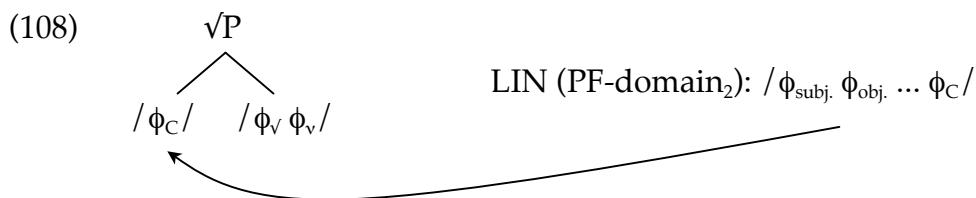


This account in terms of restrictions on the process of *Edge Replacement* does not only account for the placement facts of complement clauses in German, but also captures

the following cross-linguistic generalizations stated by Bayer (1996: 192) (see also Hawkins 1990, Dryer 1992):⁶³

(107) a. CP complements in SVO languages are head-initial.
 b. Those SOV languages which show CP-extrapolation have head-initial CPs.
 c. SOV languages which do not allow CP-extrapolation have head-final CPs.

In particular, we do not have to assume that (preverbal) head-final CPs are of a different syntactic category than (postverbal) head-initial CPs, as has been claimed by Inaba (2007). Rather, we may simply say that the exponent of C is head-final in these languages. As a result, *Edge Replacement* is only possible as long as the exponent of C remains adjacent to the exponent(s) of the $v+v$ complex:



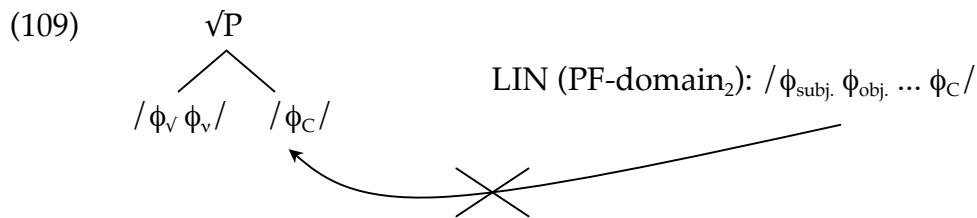
On the other hand, this serves to rule out a configuration where a complement clause introduced by a clause-final complementizer occurs in postverbal position:⁶⁴

⁶³ An apparent exception to this generalization seems to be Lakhota, where complement clauses that contain a final complementizer-like element *ki* may follow the verb (cf. Rood 1973, Dryer 1980, Lehmann 1984: 82):

(i) Tohá slolyáya he [wakpála ektá ohíhpaye ki]
 when you-know Q creek to fall COMP
 'When did you find out that he fell into the creek?
 (Dryer 1980: 132)

However, it is far from clear whether *ki* really is a free-standing complementizer. Rood (1973) glosses it as 'the' and suggests that it functions as a subordinating particle (see also Inaba 2007: 159f. for some discussion). In other words, it might be that *ki* is actually not a free complementizer at the left edge of CP, but rather a subordinating (nominalizing) particle that attaches to the verb. Under this analysis, examples like (i) would not cause a problem for my analysis: in case the left edge of CP is empty, no restrictions are imposed on the linear order of verb and complement clause, and both preverbal and postverbal placement of the embedded clause should in principle be available. As it turns out, this is in line with the facts in Lakhota, where the complement clause can also precede the verb (Rood 1973: 72):

(ii) [Wakpála ektá ohíhpaye ki] slolwáye šni.
 creek to fall COMP I-know NEG
 'I didn't know he fell into the creek.'



A nice minimal pair illustrating the different behavior of embedded clauses introduced by head-final and head-initial complementizers comes from Bengali (Bayer 1996, 1999), where clauses introduced by the clause-initial complementizer *je* uniformly appear in postverbal position, while the use of the clause-final complementizer *bole* forces the complement clause to occupy a preverbal position (CF = 'classifier'): ⁶⁵

(110) a. chele-Ta Sune-che [je [or baba aS-be]]
 boy-CF hear-PAST.3 COMP his father come-FUT.3
 b. *chele-Ta [je [or baba aS-be]] Sune-che
 boy-CF COMP his father come-FUT.3 hear-PAST.3
 'The boy heard that his father would come.'
 (Bengali; Bayer 1999: 259)

(111) a. *chele-Ta Sune-che [[or baba aS-be] bole]
 boy-CF hear-PAST.3 his father come-FUT.3 COMP
 b. chele-Ta [[or baba aS-be] bole] Sune-che
 boy-CF his father come-FUT.3 COMP hear-PAST.3
 'The boy heard that his father would come.'
 (Bengali; Bayer 1996: 255; 1999: 245)

⁶⁴ Note that this analysis predicts another asymmetry between embedded clauses where the C-domain contains merely a final complementizer and embedded clauses where further material is fronted to a specifier position in the C-domain (as e.g. in indirect questions). Since in the latter, the crucial part of the overlap is the specifier position adjacent to $/φ_v φ_v/$, we would actually expect that embedded clauses with a filled SpecCP (and a final complementizer) may (or, rather, must) appear in postverbal position. I leave the assessment of this prediction for future research.

⁶⁵ However, see Dasgupta (2007: 164) for the observation that embedded clauses with final *bole* may occur in postverbal position in Bengali if (i) they are adjuncts, or (ii) if the matrix clause contains scope-taking elements such as emphasis or negation.

Similar data come from other Indic languages such as Gujarati, Marathi, Assamese, or Oriya (cf. Davison 2007), and the Kru languages Vata and Gbadi where tensed complement clauses introduced by the initial complementizer *na* occur in postverbal position, while non-tensed complement clauses that exhibit the final complementizer *ka* precede the matrix verb (Koopman 1984: 108ff.).⁶⁶

Of course, the analysis proposed in this section covers only a small subset of the relevant data, both in German and across languages. For example, more has to be said about the differences between finite complement clauses and other types of embedded clauses. As is well-known, adverbial clauses, attribute clauses and non-finite complement clauses can appear in preverbal position in German.⁶⁷

(112) a. ?dass der Klaus, [weil Peter das Rennen gewonnen hat],
 that the.NOM Klaus since Peter the race won has
 gefeiert hat
 celebrated has
 ‘that Klaus celebrated since Peter won the race’

b. ?dass der Klaus, [nachdem Peter das Rennen gewonnen hat],
 that the.NOM Klaus after Peter the race won has
 gefeiert hat
 celebrated has
 ‘that Klaus celebrated after Peter had won the race’

⁶⁶ Biberauer et al. (2008) propose an analysis of the observation that clause-final complementizers are incompatible with a basic VO-grammar that is based on the assumption that UG rules out a configuration where a head-initial projection is immediately dominated by a head-final projection (*The Final-Over-Final Constraint*, FOFC). In other words, the non-existence of postverbal complement clauses with final complementizers is attributed to an ill-formed structural configuration *inside* the complement clause (e.g. head-final CP dominating head-initial TP). Note that this analysis fails to account for the observation that the position of the complementizer co-varies with the position of the whole complement clause relative to the matrix verb in examples like (110) and (111). See section 6.1.1 for further discussion of Biberauer et al.’s proposals.

⁶⁷ Note that preverbal placement is actually obligatory with non-finite complements that contain ‘pure’ infinitives (i.e., without the infinitival marker *zu*), as illustrated by the ungrammaticality of (i):

(i) *dass Fritz lässt [den Peter das Rennen gewinnen]
 that Fritz lets the.ACC Peter the.ACC race win
 ‘that Fritz lets Peter win the race’

(113) a. dass die Tatsache, [dass Peter das Rennen gewonnen hat],
 that the fact that Peter the race won has
 den Klaus sehr überrascht hat
 the.ACC Klaus very surprised has
 'that the fact that Peter had won the race really surprised Klaus'

b. dass der Klaus den Mann, [der das Rennen gewonnen
 that the.NOM Klaus the-ACC man who the race won
 hat], kennt
 has knows
 'that Klaus knows the man who won the race'

(114) a. dass Peter [das Rennen zu gewinnen] versprochen hat
 that Peter the.ACC race to win promised has
 'that Peter promised to win the race'

b. dass Fritz [den Peter das Rennen gewinnen] lässt
 that Fritz the.ACC Peter the.ACC race win lets
 'that Fritz lets Peter win the race'

How can these data be reconciled with an analysis in terms of *Edge Replacement*? First of all, note that attribute clauses as in (113) are not linearized directly relative to the matrix verb. Rather, LIN first has to determine their position relative to the local N-head that they modify. Subsequently, a linear ordering is established between the matrix verb and the DP containing the attribute clause. So we expect that the attribute clause may appear in preverbal position as long as it is part of another DP. The fact that attribute clauses may also optionally occur in postverbal position is presumably due to an independent stylistic rule of extraposition that is part of the PF-branch of grammar (inserting material at the right linear edge of a cycle, possibly in the sense of Nissenbaum 2000; see also Inaba 2007 for some discussion; basically the same goes for extraposition of relative clauses).

The fact that adjunct clauses are (marginally) acceptable in preverbal position seems to be more problematic. Note that in (112), the relevant embedded clauses are introduced by a clause-initial complementizer-like element. Accordingly, we expect them to occur in postverbal position to warrant adjacency of the exponent of the matrix verb and embedded C. While this is a viable (and actually preferred) possibility (presumably again due to a stylistic rule of extraposition), the acceptability of (112) comes as a surprise. Tentatively, I assume that we can account

for examples like (112) by assuming that adjunct clauses are added to the structure by some late countercyclic operation that may introduce material at the linear edge of a cycle (Nissenbaum 2000, Chomsky 2004). Making use of the phonological domains proposed above, the relevant position can be identified either as left edge of the lowest domain (corresponding to vP), or the right edge of the next higher domain (consisting of CP, TP, and the left edge of vP).

Let's now address the case of the non-finite complement clauses illustrated in (114). First of all, we might say that these complements are not separate phases (cf. e.g. Wurmbrand 2001), that is, they lack a separate CP (and possibly TP) layer. Under this assumption, the relevant clausal complements do not constitute a separate phonological domain and are thus linearized directly together with the matrix verb. We would thus expect that LIN only has to pay attention to the setting of the Head Parameter associated with the matrix verb, giving rise to OV order (again abstracting away from the option of extraposition). Alternatively, we might say that despite appearances, non-finite complement clauses involve more structure, that is, they are CPs on a par with finite clauses (cf. Sabel 1996). In this case, the important thing to note is that the left edge of the relevant non-finite CP does not contain any overt material. As a result, the *No-Tampering Condition on Edge Replacement* does not apply (since no ordering relation has been established between the matrix verb and material at the left edge of the non-finite CP). This gives rise to two possibilities: (i) the non-finite complement clause is linearized in accordance with the Head Parameter expressed by the verb, that is, it is placed in preverbal position in German; (ii) both options (i.e., preverbal or postverbal placement) are in principle available since no explicit linear ordering is established at the overlap. There are some indications that the latter option is more adequate, at least in German. Note that similar to finite complement clauses, non-finite sentential complements can be topicalized together with the matrix verb in cases of VP-fronting. Interestingly, in contrast to finite clauses, both orderings are available here:

(115) a. [[Das Rennen zu gewinnen] versucht] hat der Peter noch nie.
 the race to win tried has the Peter yet not
 'Peter has not yet tried to win the race.'
 b. [Versucht, [das Rennen zu gewinnen]] hat der Peter noch nie.

At this point, let me finally mention two further obvious problems, involving the placement of subject clauses and the fact that 'extraposition' of embedded clauses

targets a position to the right of auxiliary verbs. As illustrated in (116), in cases with periphrastic verb forms (or, more generally, verbal complexes), the complement clause must follow the whole verbal complex; placement directly to the right of the selecting verb, but to the left of the auxiliary, leads to ungrammaticality:

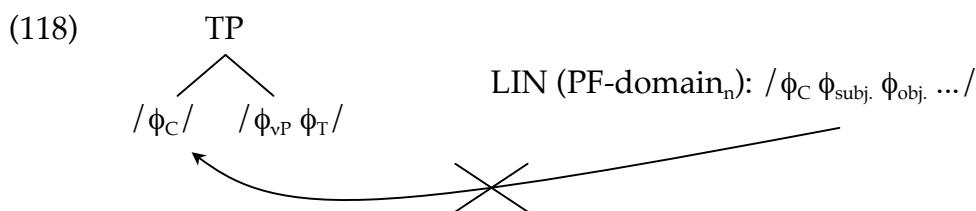
(116) a. dass Klaus geglaubt hat, [dass Peter das Rennen gewonnen hat]
 that Klaus thought has that Peter the race won has
 'that Klaus thought that Peter had won the race'
 b. *dass Klaus geglaubt [dass Peter das Rennen gewonnen hat] hat

At first sight, this seems to raise a serious problem for the analysis developed in this section: The workings of *Edge Replacement* should lead us to expect (116b) to be perfectly grammatical, contrary to facts. In other words, the data in (116) seems to suggest that the postverbal position of the complement clause is due to an extraposition operation that puts the embedded clause to the right of the verbal complex. On the other hand, we have also seen some evidence that finite complement clauses are located in the structural object position (i.e., as a sister to the verb). Recall that postverbal complement clauses are not islands for extraction, as shown in (94), and may undergo fronting together with the verb in cases of VP-topicalization (cf. (97), (98), and (115)). One possibility to resolve this apparent paradox is to assume that there are independent factors that require the parts of the verbal complex to be adjacent. As a result, no material other than verbs may appear inside the verbal complex in an SOV language like German (cf. e.g. Grewendorf 1988, Haider 1993; for discussion see also Truckenbrodt 1995, Büring and Hartmann 1997, Inaba 2007: 51). It seems likely that this adjacency requirement is part of the PF-branch of grammar. More precisely, let us assume that adjacency between the individual parts of a verbal complex is established by post-syntactic operations that apply either prior to Vocabulary Insertion (e.g., via Morphological Merger of verbal heads), or after Vocabulary Insertion (via Local Dislocation that switches the positions of a higher verb and the string of exponents realizing the clausal complement).

Now, let's consider the placement of subject clauses which exhibit a distribution similar to complement clauses, as illustrated by the following examples:

(117) a. ??weil [dass der Peter das Rennen gewonnen hat],
 since that the.NOM Peter the.ACC race won has
 den Klaus überrascht hat
 the-ACC Klaus surprised has
 'since it surprised Klaus that Peter won the race'
 b. weil (es) den Klaus überrascht hat, [dass der Peter das Rennen gewonnen hat]

At first sight, it is not quite clear how the proposed analysis can deal with the distribution of subject clauses since up to now, we have only taken a look at cases where *Edge Replacement* affects material in complement position. However, note that subject clauses presumably also constitute a separate phonological domain. Thus, we should expect that in this case as well, *Edge Replacement* takes place in order to linearize (the exponents of) subject clauses relative to the other exponents inserted to the structure. The mechanics of *Edge Replacement* demand that there is an overlap between the phonological domain of the subject clause and the rest of the structure that allows LIN to connect the strings of exponents associated with the different phonological domains and determine their linear order. Suppose that similar to other clauses, the relevant overlap consists of the ledge edge of the subject clause, that is, the complementizer *dass*. In other words, LIN first determines a linear ordering between the exponent of the complementizer and the exponents of the surrounding structure before *Edge Replacement* replaces the exponent of the complementizer with the string of exponents realizing the whole subject clause. Let's assume (for the sake of the argument) that subject clauses move to SpecTP in the syntax. Recall that due to the special status of specifiers, material in specifiers must precede all other material linearized so far. We then derive the following configuration for the workings of *Edge Replacement* in the case of subject clauses:



Since the exponent of the complementizer (i.e., the element at the overlap between the relevant phonological domains) is contained in the specifier of TP, it must precede and be adjacent to the exponents of vP and T that have been linearized earlier. Thus, in a language with initial complementizers such as German, *Edge*

Replacement invariably leads to a conflict, since it necessarily disrupts adjacency of ϕ_C and the existing set of exponents, as illustrated in (118). What I want to propose is that in this situation, the grammar resorts to obligatory extraposition as a repair mechanism in order to linearize subject clauses. More precisely, we seem to deal with an example of local dislocation, where the exponents that realize the subject clause switch places with the existing string of exponents, giving rise to absolute final placement of the subject clause, to the right of all other elements placed in the *Nachfeld* (i.e., postverbal position). This is illustrated by the following examples for relative clauses and complement clauses, respectively:⁶⁸

(119) a. weil es [den Mann t_{CPrel}] überraschte, [$_{CPrel}$ der die Spiele
 since it the.ACC man surprised who the.ACC games
 eröffnet hatte], [dass Peter das Rennen gewonnen hat]
 opened had that Peter the race won has
 'since it surprised the man who had opened the games that Peter won the
 race.'

b. *weil es [den Mann t_{CPrel}] überraschte, [dass Peter das Rennen gewonnen hat]
 [$_{CPrel}$ der die Spiele eröffnet hatte]

⁶⁸ Note that this approach in terms of repairing a non-linearizable structure via extraposition can possibly also be extended to the placement of complement clauses. An alternative repair strategy consists in moving the argument clause to clause-initial position (i.e., to the *Vorfeld* 'prefield'), which raises similar issues for an account in terms of *Edge Replacement*. Here, one might speculate that the relevant problems can be solved if we assume that A'-moved items behave differently for the purposes of linearization. More to the point, we might suppose that syntactic objects are (internally) linearized before undergoing A'-movement. As a result, material that is moved to a A'-specifier behaves like a single big word (which is related to proposals by Uriagereka 1999), that is, it can be linearized relative to other exponents without *Edge Replacement* (on the syntactic side, this would perhaps explain why syntactic objects in A'-positions are generally islands for extraction). For reasons of time and space, I have to leave these issues for future research.

(120) a. weil es dazu geführt hat [_{CPobj} dass ManU das Finale gewann]
 since it there-to led has that ManU the final won
 [_{CPsubj} dass Anelka den entscheidenden Elfmeter verschoss]
 that Anelka the.ACC decisive penalty missed
 'The fact that Anelka missed with the decisive penalty caused Manchester to win the final.'

b. *weil es dazu geführt hat, [_{CPobj} dass Anelka den entscheidenden Elfmeter verschoss] [_{CPsubj} dass ManU das Finale gewann]

Summing up, in this section I have argued that the postverbal placement of finite complement clauses in German can be analyzed as an effect of restrictions on the mechanism of *Edge Replacement*, which links (and linearizes) material included in two consecutive phonological domains. More specifically, I have claimed that an embedded complement clause introduced by a clause-initial complementizer cannot be realized to the left of the verb in an OV grammar, since *Edge Replacement* would disrupt adjacency of the complementizer and the verb established previously by the workings of LIN as part of Vocabulary Insertion. This analysis also rules out a linear ordering where a complement clause introduced by a clause-final complementizer appears to the right of its selecting verbal head. On the other hand, clausal complements can occupy a preverbal position if the complementizer occupies a clause-final position (as in Japanese or Bengali), which derives the generalization on the interplay between clausal position and complementizer position stated in Bayer (1996, 1999). In addition, I have briefly discussed a set of issues raised by this account for the analysis of German, focusing on (i) cases where clausal complements appear in preverbal position in the so-called *Mittelfeld* 'midfield' of the German clause, (ii) the fact that postverbal complement clauses may not appear inside the verbal complex, and (iii) the placement of subject clauses. To be sure, more has to be said about these issues; in the interest of time and space, however, I want to leave it at that for the moment and turn to another set of phenomena, focusing on typological and diachronic implications of the model developed so far.

6 Possible pathways for word order change

This section aims at exploring some diachronic implications of the model of linearization developed in section 4. In particular, it appears that apart from ruling out certain configurations in synchronic grammars, the *No-Tampering Condition* on

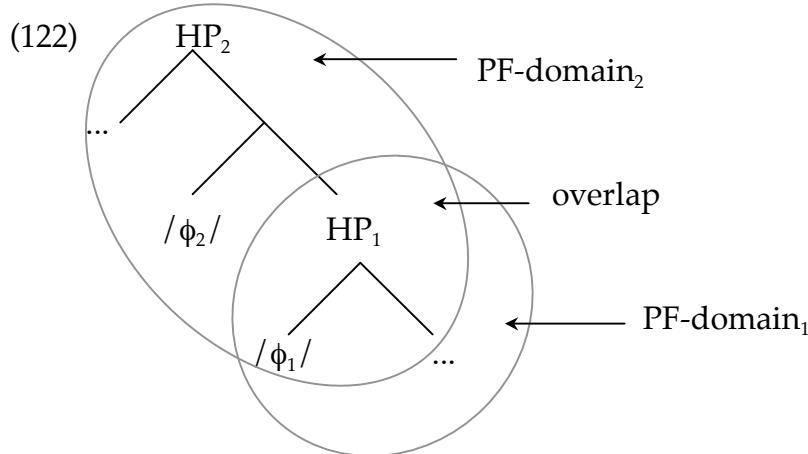
Edge Replacement imposes certain restrictions on the way languages may change over time. Before we review these restrictions, I am going to discuss the grammatical options allowed by the workings of *Edge Replacement* in some more detail, proposing a typology of possible and impossible grammars. In addition, I will compare my analysis of the cross-linguistic absence of VO-Aux orders with a recent alternative analysis developed by Biberauer et al. (2007), (2008), which is based on the *Final-Over-Final Constraint* (FOFC) proposed by Holmberg (2000).

6.1 Possible and impossible grammars

Recall that the no-tampering condition (repeated here for convenience) prohibits outcomes of LIN in which *Edge Replacement* disrupts adjacency between two exponents located at the right edge of phase HP_n and the left edge of phase HP_{n-1} , respectively, as illustrated in (122).

(121) No-Tampering Condition on Edge Replacement

Edge Replacement at $PF\text{-domain}_n$ may not disrupt ordering relations created between material of $PF\text{-domain}_n$ and material at the (overlapping) left edge of $PF\text{-domain}_{n-1}$.



In other words, once LIN has created a linear ordering between $/\phi_1/$ and $/\phi_2/$ (either $</\phi_1/, /\phi_2/>$ or $</\phi_2/, /\phi_1/>$ dependent on the Head Parameter linked to $/\phi_2/$), *Edge Replacement* (i.e., an operation replacing the overlap including $/\phi_1/$ with the whole string of exponents inserted at $PF\text{-domain}_1$), must preserve adjacency of $/\phi_1/$ and $/\phi_2/$. The set of orders ruled out by (121) is then dependent on three factors: (i) the setting of the Head Parameter for $/\phi_2/$, (ii) the make-up of the overlap (i.e., the left

edge of HP_1 / PF-domain₁), and (iii) the make-up of PF-domain₁ inserted via *Edge Replacement* (ER). Let us first take a look at the logical possibilities for orderings that comply with (121) (“...” stands for a string of phonological exponents):

(123) OK:

- a. / ϕ_2 / head-initial, / ϕ_1 / head-initial (result of ER: / ϕ_2 / + / ϕ_1 .../)
- b. / ϕ_2 / head-initial, / ϕ_1 / = SpecHP, $H = \emptyset$ (result of ER: / ϕ_2 / + / ϕ_1 .../)
- c. / ϕ_2 / head-initial, / ϕ_1 / null (i.e., no exponent inserted at overlap)
(result of ER: / ϕ_2 / + / \emptyset .../)
- d. / ϕ_2 / head-initial, / ϕ_1 / the only element in PF-domain₁
(result of ER: / ϕ_2 / + / ϕ_1 /)
- e. / ϕ_2 / head-final, / ϕ_1 / head-final (result of ER: /... ϕ_1 / + / ϕ_2 /)
- f. / ϕ_2 / head-final, / ϕ_1 / null (i.e., no exponent inserted at overlap)
(result of ER: /... \emptyset / + / ϕ_2 /)
- g. / ϕ_2 / head-final, / ϕ_1 / the only element in PF-domain₁
(result of ER: / ϕ_1 / + / ϕ_2 /)

In general, no problems arise in cases where *Edge Replacement* does not affect the ordering relation created between / ϕ_1 / and / ϕ_2 . Thus, ‘harmonic’ configurations such as (123a,g), where the Head Parameter of embedding and embedded category are identical are ruled in since material selected by the category realized by / ϕ_1 / does not intervene between / ϕ_1 / and / ϕ_2 / after *Edge Replacement*. The same holds for the special case (123b) where / ϕ_2 / is head-initial and / ϕ_1 / realizes a specifier at the left edge of phase HP (with no exponent inserted to the head of HP): due to the fact that specifiers at the left-periphery of PF-domain₁ precede all other material in PF-domain₁, *Edge Replacement* does not disrupt adjacency between head-initial / ϕ_2 / and / ϕ_1 / realizing the specifier (see also the discussion of (125c) below). In addition, there are four cases where the no-tampering condition is trivially satisfied, either due to the fact that the overlap does not contain any material (i.e., no ordering has been established between / ϕ_2 / and material inserted to the overlap) as in (123c, f), or because / ϕ_1 / is the only element that receives a Spell-Out in PF-domain₁ (e.g., in cases where HP_1 have been evacuated by syntactic movement) as in (123d, g). Note that in the latter, an explicit ordering relation is established between / ϕ_1 / and / ϕ_2 /, while in the former the relation between successive phonological domains is somewhat less clear since no ordering relation has been established at the overlap.

Above, I have argued that at least in languages like German, this indeterminacy (i.e., cases like (123f)) seems to give rise to some amount of optionality, in the sense that non-finite complement clauses may appear either to the left or to the right of their selecting head:⁶⁹

(124) a. [[Das Rennen zu gewinnen] versucht] hat der Peter noch nie.
the race to win tried has the Peter yet not
'Peter has not yet tried to win the race.'
b. [Versucht, [das Rennen zu gewinnen]] hat der Peter noch nie.

Let's now take a look at the restrictions imposed by (121) on possible orderings. It turns out that there are only three configurations that are ruled out as violations of the *No-Tampering Condition on Edge Replacement*:

(125) RULED OUT:

- a. / ϕ_2 / head-initial, / ϕ_1 / head-final (result of ER: */ ϕ_2 / + /... ϕ_1 /)
- b. / ϕ_2 / head-final, / ϕ_1 / head-initial (result of ER: */ ϕ_1 .../ + / ϕ_2 /)
- c. / ϕ_2 / head-final, / ϕ_1 / = SpecHP, and $H = \emptyset$ (result of ER: */ ϕ_1 .../ + / ϕ_2 /)

All cases listed in (125) involve a configuration in which Edge Replacement affecting an element $/\phi_1/$ at the overlap destroys an ordering relation between $/\phi_1/$ and a phonological exponent $/\phi_2/$ inserted to the head of the projection immediately dominating the overlap. In what follows, I will discuss the problematic cases in some more detail on the basis of relevant empirical phenomena.

First of all, it is important to note that the restrictions apply only to cases of *Edge Replacement*, that is, cases where a linear order has to be established between material that is part of different phonological domains. In all relevant cases, *Edge Replacement*

⁶⁹ But note that in a strict VO-language like English, no such optionality can be observed in clauses introduced by a null complementizer:

- (i) John said [\emptyset he would win the race].
- (ii) *John [\emptyset he would win the race] said.

This might be due to a late PF-rule that deletes the complementizer after linearization. Alternatively, we may assume that in (i) the nominative subject moves to a specifier in the C-domain, eliminating a [uT] feature in C, along lines proposed in Pesetsky and Torrego (2001). As a result, the subject would be the element at the overlap that is linearized with / ϕ_2 / (i.e., the verb), giving rise to obligatory postverbal placement of the complement clause.

disrupts an ordering relation between $/\phi_1/$ and $/\phi_2/$ via placing elements ("..." in (125)) in between $/\phi_1/$ and $/\phi_2/$. This implies that in all problematic cases, the lower phonological domain contains more material than merely the edge element. In the first two cases (125a, b), the relevant problem is caused by different settings of the Head Parameter for the embedding and embedded category, respectively. A relevant example for (125a) comes from the observation that a complement clause with a final complementizer cannot occur to the right of its matrix verb, compare the following pair of examples from Bengali:

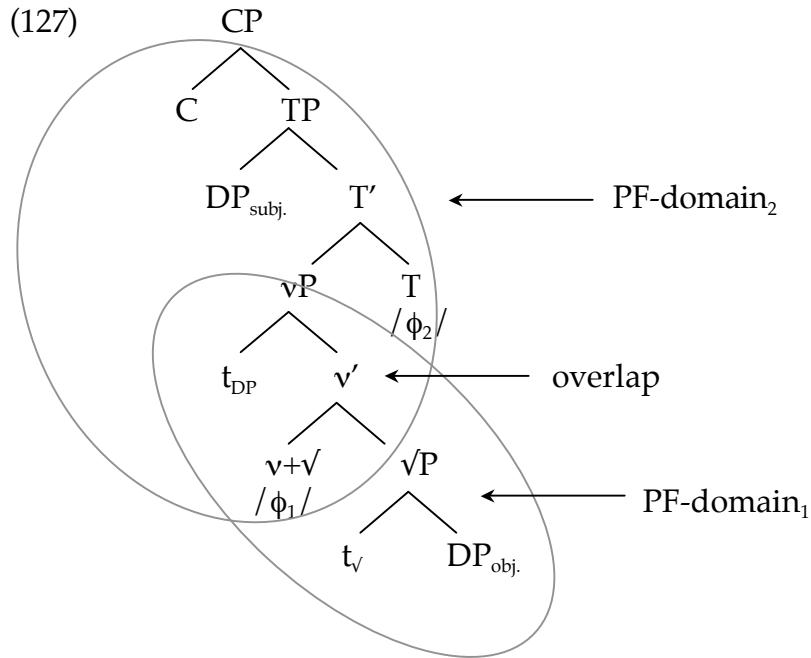
(126) a. *chele-Ta Sune-che [[or baba aS-be] bole]
 boy-CF hear-PAST.3 his father come-FUT.3 COMP
 b. chele-Ta [[or baba aS-be] bole] Sune-che
 boy-CF his father come-FUT.3 COMP hear-PAST.3
 'The boy heard that his father would come.'

(Bengali; Bayer 1996: 255; 1999: 245)

Note that at some point in the post-syntactic computation, LIN has established an ordering relation between *Sune-che* and the complementizer *bole*. *Edge Replacement* of *bole* by the string of exponents linked to the whole embedded clause disrupts adjacency of *Sune-che* and *bole* in the (a)-example and is therefore ruled out (see section 6.2 below for some discussion of apparently problematic Aux-OV orders in Old English).

(125b) excludes the combination of a head-final vP/VP with a head-initial CP, as for example in the case of German. Moreover, it serves to rule out the cross-linguistically non-attested order VO-Aux. However, note that the latter may in principle result from quite a number of different underlying syntactic configurations depending on parametric choices concerning e.g. (i) the base position of the auxiliary/modal (in T/Aux or in v/v?), (ii) the question of whether the auxiliary undergoes movement to a higher functional category (e.g., to T, or Asp), (iii) the question of whether the subject moves to SpecTP etc. It is clearly beyond the scope of this section to explore all of these syntactic options. In what follows, I will demonstrate the issues that arise with merely a couple of typical configurations that (a) involve uniform subject movement to SpecTP and (b) differ with respect to the position of the auxiliary (in T or v/v). Furthermore, recall that by assumption, a head-initial setting for v (which is required to derive vP-internal VO order) is always coupled with v-to-v (the *Root Raising Parameter*, see above).

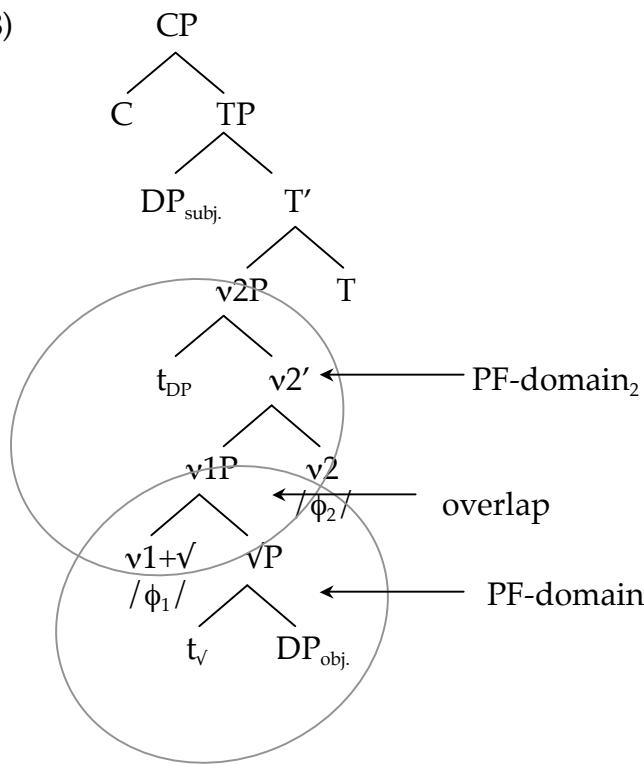
If the auxiliary is directly inserted in (head-final) T, we arrive at the following structure, in which / ϕ_2 / corresponds to head-final T and / ϕ_1 / to head-initial v:



Replacing / ϕ_1 / with the string of exponents assembled at PF-domain₁ (via *Edge Replacement*) would disrupt adjacency of / ϕ_1 / (the exponent of v+v) and / ϕ_2 / (the exponent of Aux/T) at PF-domain₂, which is ruled out by the *No-Tampering Condition on Edge Replacement*. The same goes for cases where the auxiliary is inserted as a head-final v (selecting another head-initial vP that contains the full verb), as long as we assume that the lower v1P constitutes a separate phase.⁷⁰

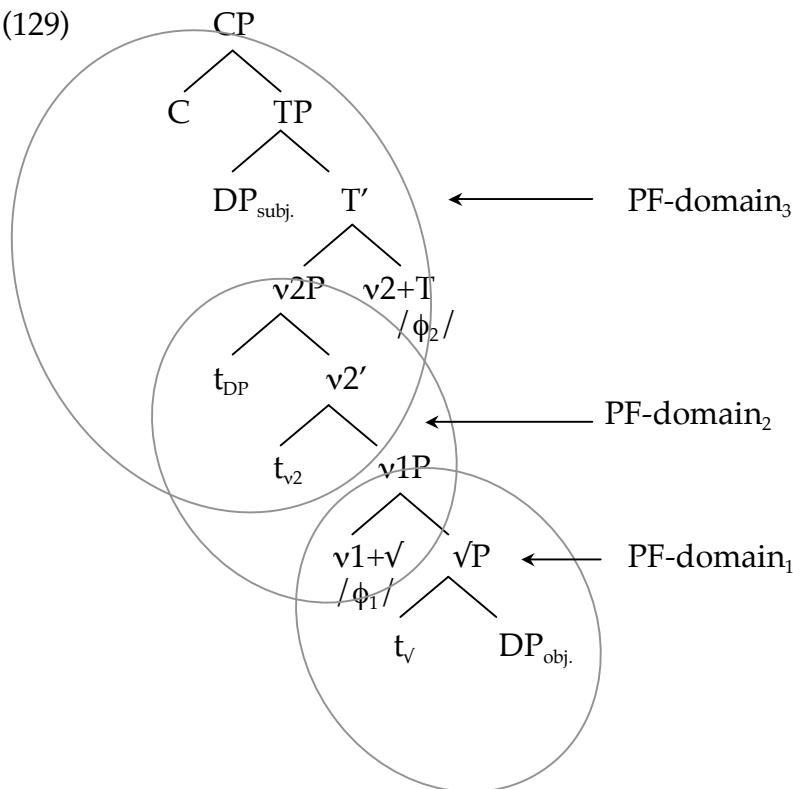
⁷⁰ Note that this structure raises a couple of questions concerning the theta-position of the external argument/subject. It seems plausible that the subject actually receives its theta-role in the lower vP, with the higher v (or, Aux) containing the auxiliary/modal acting as a kind of raising verb. See section 6.2 for some discussion concerning relevant cases in Old English.

(128)



A more problematic case involves a configuration based on structure (128) in which the left edge of v2P (i.e., the overlap between TP and v2P) has been completely evacuated by subject raising and v2-to-T raising:

(129)



It appears that in (129), the no-tampering condition imposes no restrictions on the combination of PF-domain₃ and PF-domain₂, since the overlap between these domain

is empty. Since v2P is not subject to Vocabulary Insertion, the string of exponents assembled at PF-domain₁ (corresponding to v1P, i.e., /φ_{v1} φ_{obj}/) must be linearized relative to the next target of Vocabulary Insertion/LIN, namely the exponents of the v2+T complex. At first sight, this configuration should give rise to the problematic order VO-Aux if the exponent of T is linked to a head-final setting of the Head Parameter. Clearly, this is not the desired result. However, suppose that there is an alternative way of ruling out VO-Aux orders resulting from a structure like (129) that has to do with the fact that the exponents of v2 and T are linked to conflicting settings of the Head Parameter: while (the exponent of) v2 is head-initial, (the exponent of) T is head-final. While this is normally unproblematic (as in the case of V2 in German, where the verb moved to C is head-final, while C itself is head-initial), (129) represents a special case, in that the string of exponents that must be linearized relative to the v2+T complex corresponds to the syntactic complement of v2, that is, v1P. What I want to propose is that (only) in this situation, a conflict arises between the linearization requirements linked to (the exponents of) v2 and T: While v2 requires its complement (i.e., the string of exponents corresponding to v1P) to appear to its right, T requires the existing string of exponents to appear to its left. We may assume that in this situation, the relevant structure either cannot be linearized or is repaired in favor of the more prominent requirement concerning the order of v2 and its complement, giving rise to the order Aux-VO. Crucially, in both cases, the non-existing order VO-Aux is successfully blocked. This state of affairs can be captured by the following condition:⁷¹

(130) In cases where a head α undergoes syntactic movement to a head β , the resulting structure cannot be linearized iff:

- (i) the exponents of α and β differ with respect to the Head Parameter, and
- (ii) the string of exponents linearized relative to the $\alpha+\beta$ complex corresponds to the complement of α .

(130) effectively requires that the setting of the Head Parameter must be identical for exponents of T and an exponent of v moved to T. In section 6.2.3 below, I argue that

⁷¹ Note that (130) raises no problem for successive-cyclic v-to-T-to-C movement in a V2+SOV language such as German. First of all, the constituent linearized relative to C (TP) is neither the complement of T, nor of v. Moreover, (130) clearly applies only in cases where a linear order has to be created between material that is part of different phonological domains, while in the case of V2, the relevant elements/string of exponents are part of the one and the same phonological domain.

this correlation provided a driving force in the change from OV to VO in the history of English. Note that (130) can possibly also be derived as a special case of the no-tampering condition, in the sense that linearization at PF-domain_n conflicts with a ordering relation established at PF-domain_{n-1} (in the case at hand, the linear ordering between v2 and its complement, v1P).

Finally, (125c) states that it is not possible to combine two successive phonological domains where /φ₂/, the exponent of the lowest head of PF-domain_{n+1} is head-final while the only element that receives a pronunciation at the left edge of PF-domain_n is a specifier (=/φ₁/): Application of LIN at PF-domain_{n+1} requires that /φ₁/ precede and be adjacent to /φ₂/, while subsequent *Edge Replacement* would disrupt this ordering relation by inserting the string of exponents assembled at PF-domain_n in between /φ₁/ and /φ₂/. This predicts that in languages where the exponent of v+V is head-final, indirect questions with a fronted *wh*-phrase (or other elements occurring at the left-periphery of the complement clause) must follow the matrix verb if the complementizer position is empty (or realized by a head-initial exponent). A relevant example comes from German, where indirect questions must follow the matrix verb although they exhibit a null complementizer (i.e., no ordering conflict can arise between the head-initial complementizer and the head-final matrix verb):

(131) a. Klaus hat den Peter gefragt [wen der Hans getroffen hat].
 Klaus hat the.ACC Peter asked who.ACC the.NOM Hans met has
 'Klaus asked Peter who Hans met.'
 b. *Klaus hat den Peter [wen der Hans getroffen hat] gefragt.

However, in SOV languages where the indirect question contains a head-final complementizer, we predict that preverbal placement is possible. A case in point seems to be Basque, where indirect questions introduced by a final (clitic) complementizer may precede the matrix verb:⁷²

⁷² For reasons unclear to me, indirect questions may also appear in postverbal position. This option seems to correlate with the presence of negation in the matrix clause:

(i) Ez dakit [noiz etorri d-en herri honetara].
 NEG know when come AUX-COMP this town
 'I don't know when he has come to this town.'
 (Basque; Ortiz de Urbina 1989: 214)

(132) [Noiz etorri d-en] galdetu dut.
 when come AUX-COMP ask AUX
 'I have asked when he has come.'
 (Basque; Ortiz de Urbina 1989: 206)

Another, albeit more indirect piece of evidence supporting the claim that fronting to a left-peripheral specifier of an embedded clause requires extraposition comes from the historical development of subordinators in a number of 'hybrid' South-Asian languages that exhibit both clause-final and clause-initial complementizers (cf. Bayer 1996, 1999). As illustrated in Table 1, clause-final complementizers (such as Bengali *bole*, Oriya *boli*, Assamese *buli*) typically develop from *verba dicendi*, while Table 2 shows that clause-initial complementizers originate from (fronted) operator elements such as *wh*-phrases (in Marathi and Dakkhini-Hindi) or relative pronouns (in Bengali, Oriya, and Assamese).⁷³

Language	Clause-final complementizer	Source
Bengali	<i>bole</i>	past participle of <i>bol-</i> 'say'
Oriya	<i>boli</i>	past participle of <i>bol-</i> 'say'
Assamese	<i>buli</i>	past participle of <i>bol-</i> 'say'
Marathi	<i>mhaṇun</i>	derived from 'say'
Dakkhini-Hindi	<i>bolke</i>	<i>bolkee</i> 'having said'

Table 1: Clause-final complementizers in hybrid languages (Bayer 1999: 237)

Similar facts are reported for Bengali by Dasgupta (2007: 164) who notes that postverbal placement of clauses introduced by final *bole* is improved by "certain scope-taking elements, like emphasis in the matrix [...] or negation":

(ii) amra keu Suni ni [ajke briSTi poRbe bole].
 we anyone have-heard NEG today rain will-fall COMP
 'None of us has heard that it will rain today.'

⁷³ The similarities between the complementizers suggest that at least some of the languages acquired clause-initial subordinators via borrowing. In particular, it seems likely that borrowing is the source of head-initial complementizers in SOV languages that lack (regular) operator fronting. Another relevant example of borrowing comes from Turkish, where the initial subordinator *ki*, which introduces extraposed finite clauses, entered the grammar via language contact with Persian.

Language	Clause-initial complementizer	Source
Bengali	<i>je</i>	relativization operator
Oriya	<i>je</i>	relativization operator
Assamese	<i>je</i>	relativization operator
Marathi	<i>ki</i>	(perhaps) 'what'
Dakhini-Hindi	<i>ki</i>	(perhaps) 'what'

Table 2: Clause-initial complementizers in hybrid languages (Bayer 1999: 237)

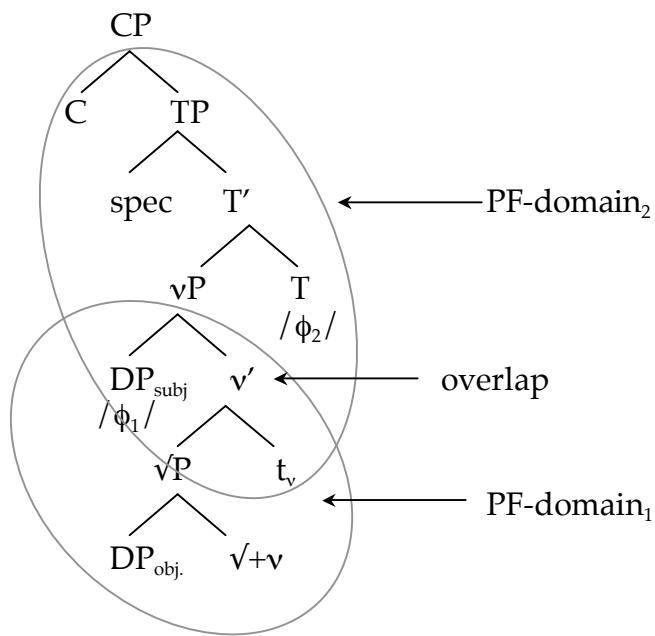
Recall that the position of complement clauses co-varies with the position of the complementizer. In particular, we have seen that clauses introduced by a clause-initial complementizer must occur in postverbal position in these 'hybrid' languages. This implies (if Bayer's generalization proves to be valid) that clause-initial complementizers could not develop in contexts where the embedded clause containing the source element occupied a preverbal position. However, if the relevant embedded clauses had to occupy a postverbal position for the grammaticalization process to take place, then the question arises of why these clauses were extraposed in a basic SOV language in the first place. A possible scenario becomes available under the assumption that the workings of *Edge Replacement* rule out structures like (125c), requiring postverbal placement of embedded clauses in which an operator has been fronted to SpecCP and C is null. In other words, it is not accidental that the set of complementizers introducing postverbal complements (cf. Table 2) developed from syntactic operators, since fronting of the latter (inducing 'extraposition' of embedded clauses) provided the only context where clause-initial complementizers could evolve in a strict SOV grammar. The following example illustrates the possibility of fronting a *wh*-phrase in Assamese (note that in the absence of the complementizer *je*, the *wh*-phrase can also remain in a medial position; see Bayer 1996: 269ff. for discussion and further examples of fronting of relative pronouns/operators in the relevant set of South Asian languages):

(133) a. moi ne-janu [kak_i je bill-e juwal kali t_i dekhisil]
 I NEG-know who COMP Bill-NOM yesterday seen-has
 'I don't know who Bill saw yesterday.'

b. moi ne-janu [bill-e kak_i juwal kali t_i dekhisil]
 I NEG-know Bill-NOM who yesterday seen-has
 'I don't know who Bill saw yesterday.'
 (Assamese; Bayer 1996: 270)

Another issue that is related to the ruled-out pattern (125c) concerns the possibility of leaving the subject in a vP-internal position in the overt syntax. More to the point, (125c) should lead us to expect that the subject must move to SpecTP if (i) T is realized by a head-final exponent, and (ii) $v = \emptyset$ (recall that if v is overtly realized by a head-final exponent, no problem arises for *Edge Replacement* since the latter does not disrupt adjacency of $/\phi_v/$ and $/\phi_T/$). Note that this configuration arises if we assume that in basic SOV languages, v connects with the verbal root via lowering (the *Root Raising Parameter*, see above):

(134)



After *Edge Replacement*, the exponent(s) of the subject would fail to be adjacent to the exponent of T in structure (134), which is therefore ruled out by the non-tampering condition. If valid, (125c) possibly gives us a clue of why verb raising to inflectional heads is generally problematic in SOV languages (and therefore ruled out in many cases, see section 4.2.3 above): Overt spell-out of final T creates a serious complication for the linearization of specifiers at the left edge of vP if v is empty, that is, in cases where it contains only the trace of the moved (finite) verb. This conflict can be circumvented by either (i) raising all relevant elements out of vP (evacuating the edge), or (ii) absence of v-to-T raising. Furthermore, note that several authors have argued that there is a connection between the absence of V-to-I/T movement

and the absence of subject raising in strict SOV languages such as Japanese (cf. Kuroda 1988, Saito 1992; see also Julien 2002). Again, this connection is expected under the present approach to linearization, in the sense that subjects can only be linearized in their base position as long as the verb does not overtly raise to a head-final T/Infl head.

Summing up the discussion so far, we have seen that the *No-Tampering Condition on Edge Replacement* rules out certain combinations of parametric choices due to the fact that they cannot be linearized by the workings of Vocabulary Insertion/LIN. Below, the excluded configurations are listed separately for cross-clausal and clause-internal contexts:

- (135) a. *head-initial v/V embedding a head-final CP, (125a)
- b. *head-final v/V embedding a head-initial CP, (125b)
- c. *head-final v/V embedding CP with filled specifier and C = Ø, (125c)

- (136) a. *head-initial Aux in T embedding a head-final vP (with verb in v), (125a)
- a.' *head-initial Aux in v2 embedding a head-final v1P (with verb in v1), (125a)
- b. *head-final Aux in T embedding a head-initial vP (with verb in v), (125b)
- b.' *head-final Aux in v2 embedding a head-initial v1P (with verb in v1), (125b)
- c. *head-final T embedding vP with filled specifier and v = Ø, (125c)
- c.' *head-final v2 embedding v1P with filled specifier and v = Ø, (125c)⁷⁴

I have proposed that further restrictions on possible grammars/orders are imposed a condition (see (130)) that bans conflicting values of the Head Parameter in a head complex, ruling out movement of head-initial v to head-final T under certain circumstances. Before we turn to the implications of these restrictions for the way languages may change over time, I am going to review an alternative approach to ruling out VO-Aux orders that has been proposed recently in work by Biberauer et al. (2007), (2008).

⁷⁴ Note that the ruled-out configuration (136c') is only listed for the sake of completeness. Under the assumption that in a layered vP, the external argument always occupies a position at least as high as the specifier of the higher vP (that is, the higher v is either a control or a raising verb), the relevant specifier of the lower vP is never realized by overt material and therefore does not cause any problems for linearization.

6.1.1 *Edge Replacement* vs. FOFC

Recently, Biberauer et al. (2007), (2008) have proposed an alternative account of the cross-linguistic absence of VO-Aux orders which takes at its point of departure a generalization proposed in Holmberg (2000: 124):

(137) *Final-Over-Final Constraint* (FOFC)

If α is a head-initial phrase and β is a phrase immediately dominating α , then β must be head-initial. If α is a head-final phrase, and β is a phrase immediately dominating α , then β can be head-initial or head-final.

At first sight, it seems that the empirical coverage of the FOFC is similar to the set of phenomena that are captured by the constraints on *Edge Replacement* proposed in the previous sections. We will see shortly, however, that the two approaches actually make some slightly different predictions concerning the set of possible grammars. Before we turn to these issues, let me first review the theoretical approach of Biberauer et al. in some more detail. Biberauer et al. suggest that the FOFC can be derived if the following set of theoretical assumptions is adopted:

(138) a. PIC in the form proposed in Chomsky (2000): the complement of a (nondefective) phase head H is subject to Transfer when the phase headed by H has been completed.

b. *Radical removal* of material in the Spell-Out domain: material subject to Transfer is linearized immediately and cannot accompany its dominating phase when the latter undergoes EPP-driven movement to a higher specifier (e.g., when vP is moved to SpecTP, VP cannot be spelled-out in the TP domain, but rather is linearized as a clause-final VO string).

c. Kayne's (1994) version of the LCA.

d. Massive roll-up movement (i.e., if a head H hosts an EPP-feature, this triggers movement of the head's complement to the specifier of H).

e. Any head may carry a movement-inducing EPP-feature; the presence of EPP-features is restricted by the following constraint (Biberauer et al. 2008: 101):

(139) If a phase head PH has an EPP-feature, then all heads in its complement domain must have an EPP-feature.

If we limit our attention to the vP domain, this gives rise to the following set of possibilities (Biberauer et al. 2008: 101):

(140) a. $v_{EPP} \ V_{EPP} \rightarrow [[_{VP} O V] v]$ (consistent head-final order)
 b. $v \ V_{EPP} \rightarrow [v [_{VP} O V]]$ (disharmonic non-FOFC-violating order)
 c. $v \ V \rightarrow [v [_{VP} V O]]$ (consistent head-initial order)
 d. $*v_{EPP} \ V \rightarrow [[_{VP} V O] v]$ (FOFC-violating order)

Thus, under the assumption that the finite auxiliary is located in v , the set of assumptions in (138) and (139) seems to successfully rule out the unwanted order *VO-Aux. Note the role of massive roll-up movement in this analysis: for example, in the consistently 'head-final' grammar (140a), the object first moves to the specifier of V before the whole VP is attracted to SpecvP by v 's EPP-feature. The possibility of deriving *VO-Aux via moving a consistently head-initial vP to the specifier of 'head-final' T hosting the auxiliary is then excluded by appealing to assumption (138b), *radical removal*: Recall that by assumption, a head-initial VP must always remain in its complement position where it is spelled-out and linearized upon completion of vP (due to (139), the absence of EPP on V implies the absence of EPP on v). As a result, VP cannot undergo EPP-driven movement of vP to SpecTP and must stay behind in clause-final position. The only elements that can be moved to SpecTP are v and its edge. The problematic combination of a VO grammar with final complementizers is then ruled out in a similar vein: by assumption, the order VO-C can only result from moving a head-initial VP to SpecCP. Similar to *VO-Aux with Aux in T, this is ruled out by (138b), i.e., *radical removal* of the head-initial VP at the previous phase. Thus, independent of the orientation of T/Aux, both SVOAuxC and SAuxVOC languages are excluded.⁷⁵

To account for a couple of apparent counterexamples, including the possibility of head-initial DPs embedded under a head-final VP (as e.g. in German, see above), and clause-final force particles in a basic VO-language such as Mandarin (as illustrated in (142)), the constraint (139) is modified in the following way (Biberauer et al. 2008: 102):

⁷⁵ To open up the possibility of VP-topicalization in VO languages, Biberauer et al. claim that VP-fronting is made available by an A'-related Edge feature which dislocates VP to the left edge of vP from where it can undergo further A'-movement to CP.

(141) If a phase head PH has an EPP-feature, then all heads in its complement domain from which it is non-distinct in categorial features must have an EPP-feature.

(142) a. Xià yǔ le **ma?**
 fall rain PART Q
 'Is it starting to rain?'
 b. Zánmen kuài zǒu **ba!**
 1PL quick go EXCLAM
 'Let's leave immediately!'
 (Mandarin; Biberauer et al. 2008: 100)

As a result, the head-initial character of nominal categories in German is independent of the head-final nature of vP/VP (i.e., EPP on n/N is independent of EPP on v/V). The head-final placement of complementizers in Mandarin is then ruled in by the assumption that the relevant C-elements are nominal in nature, while T and v/V are verbal categories (the same goes for final tense particles in Ma'di and clause-final negation in a number of Central African languages, cf. e.g. Dryer 2007 on the latter). In what follows, I am going to take a closer look at the predictions of the FOFC-based analysis put forward by Biberauer et al. and compare them with the predictions of an account in terms of restrictions on *Edge Replacement*. What I am going to argue is that the LCA-based approach is both too strong (since it rules out existing orders) and too weak (since it predicts ordering possibilities that are not attested cross-linguistically).

First, let's take a look at the word order options of verbs, auxiliaries and objects predicted by Biberauer et al. While at least in the core cases discussed above, their analysis serves to rule out VO-Aux orders, it seems that it may be capable of deriving other orders that are also quite rare among the world's languages. For example, adding the (common) assumption that at least in VO languages, V undergoes systematic movement to v, we derive (S)-V-Aux-O as a possible basic word order for cases with nominal complements (with Aux in T and VP spelled out in situ; the position of the VP in the moved vP is marked by strikethrough):

(143) $T_{EPP} \ v \ V \rightarrow [\cancel{_{TP} [_{vP} V+v [VP]}] T \cancel{t_{vP}} [_{vP} t_V O]] \Rightarrow (S) V \text{ Aux } O$

While this order is attested for certain types of complements, notably complement clauses in Dutch and German, it is very rare (if not absent) as a systematic basic word

order with all types of complements (and in particular nominal objects) cross-linguistically.⁷⁶

Let's now take a closer look at the distribution of complementizers. Recall that an approach in terms of *Edge Replacement* do not impose any ordering restrictions concerning the directionality of heads that are part of the same phonological domain (as long as no other problems arise, e.g. conflicting settings of the Head Parameter after head movement etc.). As a result, it predicts the possibility of a head-initial TP embedded under a head-final CP:

(144) a. S Aux VO Comp
 b. S Aux OV Comp

(144a) is represented by Mandarin (see above); possible examples for option (144b) come from the Kru-languages Vata and Gbadi which exhibit a basic S-Aux-OV syntax. However, non-tensed complement clauses, which appear in preverbal position, are marked by the final complementizer *ka*:

(145) a. mäsłpànyò kà [mò yāmà kā] yí
 healer FUT you healthy-MA COMP come
 'The healer will come to make you healthy.'
 (Vata; Koopman 1984: 46)

b. wà ní yù [zībià plà kà] líb
 they NEG-ADJ child fish-PL buy COMP send
 'They have not sent the child to buy fish.'
 (Vata; Koopman 1984: 57)

⁷⁶ In addition, Biberauer et al. must rule out the possibility of a derivation based on a consistently 'head-final' vP, in which the object moves out of VP (e.g. for Case, labeled CaseP here), followed by movement of the remnant VP to a higher position that is both to the left of the object and the auxiliary, schematically:

(i) ... [_{FP} [_{VP} _{t_{obj}} V] F ... [_{CaseP} obj. ... [_{vP} _{t'_{VP}} v [_{t_{VP}}]]]] ⇒ *VO-Aux

As illustrated in (i), this derivation, if possible, would produce the unwanted option *VO-Aux. Note that these problems do not arise under an analysis in terms of *Edge Replacement*, which is neither capable of deriving systematic V-Aux-O orders nor powerful enough (at least as long as a non-Kaynian approach to syntax is adopted) to derive the unwanted option *VO-Aux.

Both options are in principle excluded by the FOFC and must be ruled in by additional assumptions such as a categorial difference between the clause-final complementizer and T/V. So it seems that at least in its pure form, the LCA-based approach is too restrictive here, while an account in terms of *Edge Replacement* makes correct typological predictions without additional assumptions.

Moreover, it appears that the approach suggested by Biberauer et al. overgenerates since it is capable of deriving at least one grammatical option that is not attested across the world's languages, namely the possibility of clause-medial complementizers. Recall that in a grammar with head-initial VP, the latter must always be spelled out in situ, giving rise to a clause-final VO string (with the position of the TP in the raised vP marked by strikethrough). If this grammatical option is combined with a 'head-final' setting for C (and, accordingly T), triggering roll-up movement of (i) vP to SpecTP and (ii) TP to SpecCP, we derive the following ordering (with Aux either in v or T), where the complementizer appears in the middle of the clause:

$$(146) \quad C_{EPP} \quad T_{EPP} \quad v \quad V \rightarrow [_{CP} [_{TP} [_{vP} V + v [\cancel{vP}]] T t_{vP}] C t_{TP} [_{vP} t_V O]] \Rightarrow (S) \text{ Aux } C \text{ VO}$$

An additional set of problems comes from the position of embedded clauses and the observation that there is a connection between complementizer position and the placement of complement clauses (see above). More precisely, it appears that the approach advocated by Biberauer et al. requires a discontinuous spell-out of the complement clause in languages like German, which are characterized by the following parametric choices (OV+initial complementizers):

$$(147) \quad C \quad T \quad v_{EPP} \quad V_{EPP}$$

Note that the absence of an EPP-feature in C (to derive 'head-initial' complementizers) implies the absence of EPP on T in Biberauer et al.'s system. Furthermore, recall that under the assumption of *radical removal*, the complement of the complementizer, that is, the embedded TP must be spelled out and linearized upon completion of the embedded CP. As a result, the EPP feature linked to matrix V can only attract the head (and the edge) of the embedded CP. After further roll-up movement to the specifier of v (which presumably hosts the auxiliary in German), we derive the following structure and linear ordering for the complementizer, the complement clause, and matrix verbs in German:

(148) ... [_{VP} [_{CP} C [_{TP}]] V t_{CP}] v t_{VP} [_{TP} S O V]] \Rightarrow *S C V Aux [S O V]

As illustrated in (148), the FOFC account of German in terms of roll-up movement and *radical removal* predicts that complementizers should be separated from their clause by the matrix verbal complex, clearly an unwelcome result. Thus, as already noted above in fn. 65 it is not clear how Biberauer et al. can account for the generalization that there is systematic correlation between the orientation of the complementizer (final or initial) and the position of the complement clause relative to the matrix verb. While preverbal placement of complement clauses with final complementizers may be attributed to repeated roll-up movement, the fact that complement clauses with initial complementizers follow the matrix in both OV and VO languages seemingly cannot be accounted for. However, note that both parts of the generalization directly follow from the *No-Tampering Condition on Edge Replacement*.

Summing up, it appears that an approach in terms of *Edge Replacement* is empirically more adequate than the FOFC-based analysis proposed in Biberauer et al. (2007), (2008). In particular, we have seen that the FOFC-based approach requires additional assumptions to rule in attested word order options (such as final complementizers in T-initial languages), which is not necessary under the account developed in this chapter in terms of *Edge Replacement*. Furthermore, the FOFC-based analysis overgenerates, predicting that certain non-existing word orders should in principle be available. In particular, I have demonstrated that the LCA-based account of the FOFC falsely predicts a discontinuous spell-out of complement clauses in SOV languages with initial complementizers.⁷⁷ In the next section, I am going to examine some diachronic implications of the approach to linearization developed in the previous sections.

6.2 Word order variation and change in the history of English

The restrictions on linearization developed above are taken to be 'hard-wired' properties of the syntax-morphology/phonology interface. Accordingly, the relevant

⁷⁷ Moreover, note that the theoretical assumptions of Biberauer et al. are not unproblematic by themselves (see section 4 for some critical discussion of Kayne's version of the LCA). In particular, the assumption of roll-up movement placing the whole TP in SpecCP seems to raise a couple of issues as shown e.g. in Bayer (1999).

constraints ease the process of language acquisition, in the sense that they delimit the set of parametric options that must be taken into consideration when the learner tries to reconstruct the target grammar underlying the utterances he/she is confronted with. For example, if the child encounters regular S-Aux-O-V order in embedded clauses, he/she must assume a negative setting of the *Root Raising Parameter* since this order can only be linearized if the overlap between the domain containing the finite auxiliary and the domain containing the object and the non-finite verb is empty (see (136a) in section 6.1 above). In a similar vein, the ban on conflicting settings of the Head Parameter within a single head adjunction complex (cf. (130)) may impede the acquisition of a OV setting for v in grammars with systematic movement of v to head-initial T (see section 6.2.3 on the loss of OV orders in the ME period).

In this way, the restrictions in question also determine the possibilities for grammar change, that is, the range of parametric choices in which the grammar eventually acquired by the learner may possibly differ from the target grammar. In other words, there can be no changes that give rise to the orderings/grammatical choices ruled by the *No-Tampering Condition on Edge Replacement* (e.g., a head-final TP dominating a head-initial vP).

Interestingly, the restrictions also make certain predictions concerning possible pathways of word order change. In particular, we expect that the often observed change from a strict SOV grammar to a SVO grammar proceeds in a ‘top-down’ fashion, in the sense that a change in the setting of the head parameter must first affect exponents of higher functional heads before it can affect exponents of lower functional heads (see in particular Kiparsky 1996 on the validity of this generalization). Note that “Aux” in the following schemata corresponds to the exponent of T/Infl.⁷⁸

(149) (Comp) S O V Aux > (Comp) S Aux OV > (Comp) S Aux V O

In a similar vein, we expect that the change from a basic VO grammar to OV proceeds in a ‘bottom-up’ fashion, first affecting the setting of the Head Parameter

⁷⁸ See Biberauer et al. (2008: 99) for a related set of conclusions/predictions based on the FOFC. Note that an approach in terms of the FOFC and an account in terms of *Edge Replacement* make different predictions concerning the circumstances under which a head-initial projection may be dominated by a head-final projection. While the FOFC predicts that this should be possible if the relevant heads differ with respect to categorial features, the relevant configuration may arise under *Edge Replacement* in cases where the relevant syntactic objects are properly contained in the same phonological domain.

linked to (exponents of) lower functional heads, before it can reverse the setting for higher functional categories (see also Biberauer et al. 2008: 99):

(150) (Comp) S Aux V O > (Comp) S Aux O V > (Comp) S O V Aux

While (150), that is, the change from a strict VO grammar to a basic OV grammar is only rarely (if at all) attested in the records of the past available to us, we have access to quite some historical data exemplifying the rise of a strict VO grammar along the lines depicted in (149). In what follows, I will therefore focus on the pathway in (149), taking a closer look at the textbook example of the relevant changes, namely the rise of a basic VO grammar in the history of English (cf. e.g. Stockwell 1977, Mitchell 1985, Kemenade 1987, Lightfoot 1991, Kiparsky 1996, Roberts 1997, 2007a, Pintzuk 1999, 2005, Haeberli 1999, Fischer et al. 2000, Kroch and Taylor 2000, Trips 2002, Hinterhözl 2004, Biberauer and Roberts 2006, among many others).

If we take a second look at the schemata in (149) and (150), they seem to imply that word order change proceeds via a set of different discrete stages until the 'destination' of the pathway is eventually reached. However, it is a well-known fact that the different steps on the pathway are actually not discrete historical stages, but rather blend into each other. This observation has given rise to the notion that linguistic change is intimately linked to linguistic variation (cf. e.g. Kroch 1989, 2001, Labov 1994). At first sight, it seems that this kind of variation might be attributed to sociolinguistic factors, that is, incomplete/ongoing diffusion of certain linguistic features in a speech community. In other words, making use of the relevant notions introduced in chapter 1, we might say that the apparent linguistic variation is merely an 'illusion' that results from not distinguishing properly between grammar change affecting the linguistic competence of individuals and the way these changes spread through a speech community. However, work by Anthony Kroch and his collaborators (see in particular Pintzuk 1999, Haeberli 1999 on Old English, OE), has quite convincingly shown that the relevant linguistic variation is also a characteristic of the linguistic output of individuals. Thus, we deal with 'real' linguistic variation that cannot be attributed to sociolinguistic factors (cf. also the work of William Labov on the link between variation and change concerning phonological features of contemporary American English). Interestingly, we can observe that in OE, all three stages of the schema in (149) were valid grammatical options.

The following examples illustrate this fact with embedded clauses of OE (adopting for expository reasons the traditional assumption that the finite auxiliary is

located in T/INFL while the non-finite verb heads a VP; see below for some qualifications).⁷⁹ The examples in (151) display a verbal complex with the finite verb in absolutely final position, which suggests a head-final setting of the Head Parameter for both V and T/INFL (verbs and verbal particles are set in boldface).

(151) a. þe se ealdormon wiþ hiene **gedon** **hæfde**
 that the alderman against him done had
 'that the alderman had done against him'
 (Or 33.13-14; Pintzuk 1991: 107)

 b. þæt man þam halgan were þæt ilce hors eft **bringan** **sceolde**
 that one the holy man that same horse again bring should
 'that one had to bring the holy man the same horse again'
 (GDC, 78.15; Haeberli 1999: 356)

In (152) however, the verbal complex appears in a sentence medial position, followed by a verbal particle and an object in (152a) and all objects including a pronoun in (152b). Since pronouns and verbal particles generally cannot undergo postposition in Germanic, examples like (152) suggest that speakers of OE had in addition access to a grammar where both T/INFL and V were head-initial (cf. Pintzuk 1999).⁸⁰

⁷⁹ If not indicated otherwise, the English examples are taken from the *York-Helsinki-Toronto Parsed Corpus of Old English Prose* (henceforth the York Corpus), and the *Penn-Helsinki Parsed Corpus of Middle English, Second Edition* (henceforth PPCME2).

⁸⁰ Interestingly, this kind of word order variation is not confined to OE, but can also be found in other early Germanic languages such as Old High German (OHG) as illustrated by the following examples (cf. Lenerz 1984, 1985, Tomaselli 1995, Hinterhözl 2004, Axel 2007):

(i) a. bihuuiu man in Judases chunnes fleische Christes bidendi uuas
 why one in Judah-GEN tribe-GEN flesh Christ expecting was
 'why one was expecting Christ in the flesh of the tribe of Judah.'
 (Isidor, 575; Eggers 1964)

 b. dhazs uueroðheoda druhtin sendida mih zi dhir
 that of-hosts lord sent me to you
 'that the Lord of Hosts sent me to you'
 (Isidor, 236; Eggers 1964)

In contrast, the modern Germanic languages do not license a similar freedom with respect to the (basic) serialization of the verb and its complements. The historical development that led to the fixation of linearization is apparently independent of the basic word properties of the languages in question. In other words, there seems to be a general tendency in the Germanic languages that led

(152) a. þæt he **ahof upp** þa earcan
 so-that he lifted up the chest
 'so that he lifted the chest up'
 (GD(C), 42.6-7; Pintzuk 1991: 78)

b. þæt he **wolde geswutelian** swa his digelnyse eow
 that he would reveal so his secrets you
 'that he wanted to reveal his secrets to you in such a way'
 (ÆLS (Thomas) 166; Haeberli 1999: 360)

In addition to 'harmonic' combinations of parametric choices, where (exponents of) T and V exhibit an identical setting of the Head Parameter, there seems to be still another option, namely the 'disharmonic' combination of a head-initial projection hosting the finite auxiliary and a head-final VP:

(153) a. þæt he **warp** þæt sweord **onweg**.
 so-that he **threw** that sword **away**
 'so that he threw away the sword'
 (Bede 38.20; Pintzuk 1999: 57)

b. þæt he **mehte** his feorh **generian**
 that he could his property save
 'that he could save his property'
 (Oros., 48.18; Kemenade 1987: 59)

Under the assumption that verbal particles cannot move rightward, but rather mark the base position of the verb, examples like (153a) suggest that the finite verb has undergone leftward movement to a clause-medial functional head position. However, note that examples like (153b) could alternatively be analyzed as instances of Verb Projection Raising (Haegeman and van Riemsdijk 1986), where the lower VP [*his feorh generian*] moves to the right of the finite verb *mehte*. To show that the relevant disharmonic combination of parametric choices (i.e., head-initial TP + head-final VP) was really an option in OE, Pintzuk (1999) points to examples like the following, in which the element intervening between the parts of the verbal complex

from the existence of both OV and VO orders to the fixation of either OV or VO as a single basic word order (cf. Gerritsen 1984, Weerman 1989).

is a pronoun, which is typically impossible in the modern Germanic languages that exhibit VPR (such as varieties of Swiss German):

(154) þæt he **wolde** hine **læran**

that she would him teach

‘that she would teach him’

(ÆELS 25.173; Pintzuk 1999, p. 73)

(155) a. dass er __ **will** [_{VP} em Peter es Gschänk **gää**]

that he wants the Peter the present give

b. *dass er __ **will** [_{VP} im es Gschänk **gää**]

that he wants him the present give

(Zurich German, Cécile Meier, p.c.)

According to Pintzuk (1999), (2005), the co-occurrence of examples such as (151)–(154) suggests that OE was characterized by a mixed OV/VO character, where in principle both settings of the Head Parameter were available for T/INFL and V. Following Kroch (1989), Pintzuk claims that this kind of word order variation is an indication of change in progress (see also Labov 1994 on the idea that language variation is a necessary ingredient of language change), which can be formally modeled by appealing to the notion of Grammar Competition. Thus, she assumes that speakers of OE had command over more than a single internalized grammar (with the individual grammars differing with respect to the setting of the Head Parameter for V and T/INFL, the so-called *Double Base Hypothesis*, DBH), which produces the variation at the syntactic surface. Schematically, the relevant competing grammars can be illustrated as follows (assuming general V-to-T/INFL movement of the finite verb; cf. Kiparsky 1996: 162, Pintzuk 2005: 119):⁸¹

⁸¹ Note that there are further word order options that are traditionally analyzed as Verb Raising, as illustrated in (i), and postposition of nominal arguments, as shown in (ii). Additional variation is introduced by the placement possibilities for objects in double object constructions, where each object can appear either to the left or to the right of the verbal complex (see Pintzuk 1999, 2005 for detailed discussion).

(i) þæt he Saul ne **dorste ofslean**

that he Saul NEG dared murder

‘that he didn’t dare to murder Saul’

(Oros., 52.33; Kemenade 1987: 59)

(156) a. **head-final IP, head-final VP: S-O-V-Aux**

þæt [IP se biskop [r [VP2 [VP1 þæt cild up aheafan] t_i] wolde_i+INFL]]]
 that the bishop the child up lift wanted

b. **head-initial IP, head-initial VP: S-Aux-V-O**

þæt [IP se biskop [r wolde_i+INFL [VP2 t_i [VP1 aheafan up þæt cild]]]]

c. **head-initial IP, head-final VP: S-Aux-O-V**

þæt [IP se biskop [r wolde_i+INFL [VP2 [VP1 þæt cild up aheafan] t_i]]]

d. ***head-final IP, head-initial VP: S-V-O-Aux**

*þæt [IP se biskop [r [VP2 t_i [VP1 aheafan up þæt cild]] wolde_i+INFL]]

However, as indicated in (156d), there is a gap in the word orders predicted by the DBH. As already noted by Pintzuk (1999) (see also Pintzuk 2005), the combination of a head-final TP/IP and a head-initial VP is apparently not attested in the Old English records.⁸² Pintzuk (1999), (2005) concludes that the problematic combination of parameter settings must be ruled out by a stipulation.⁸³ However, as already pointed

(ii) þæt he **miltisan** **wolde** his agenum slagum
 that he pity would his own executioners
 'that he would pity his own executioners'
 coaelive,+ALS_[Exalt_of:Cross]: 181.5692; Pintzuk 2005: 118)

⁸² According to Pintzuk (2005: 120), there are three examples in the York Corpus where a constituent appears between the verb and a clause-final auxiliary. In all relevant cases, the intervening element is an adverb or a PP:

(i) hu hie **gedon** ymbe þa menn **haefdan**
 how they done about the men had
 'how they had dealt with the men'
 (cobede,Nede_5: 11. 416.25.4189; Pintzuk 2005: 120)

⁸³ Kiparsky (1996) and Fuß and Trips (2002) suggest that this stipulation can be eliminated if it is assumed that the Head Parameter is confined to lexical categories while functional categories are uniformly head-initial. This assumption seems to get the job done, but it is confronted with basically the same problems that are usually raised against Kayne's Uniform Base Hypothesis (clause-final complementizers etc.). Moreover, at a second look, it appears that this proposal requires some further assumptions in order to successfully rule out VO-Aux orders. For example, the claim that functional categories are always head-initial leads to the conclusion that uniformly head-final orders (i.e., (156a)) do not involve V-to-INFL movement. This in turn requires that the combination of a head-initial VP immediately dominated by a head-final VP must be excluded by another stipulation. To rule out the problematic configuration, Fuß and Trips (2002) propose that a head-initial VP is always dominated by a head-initial vP that closes off the series of VP-shells.

out above, there are good reasons to believe that the problematic order is actually excluded on principled grounds. Above we have seen that both an approach in terms of *Edge Replacement* and an LCA-based implementation of Holmberg's (2000) FOFC are capable of ruling out VO-Aux orders. In the next section, I discuss a relevant 'Kaynian' analysis of OE that has recently been proposed by Biberauer and Roberts (2005), (2006).

6.2.1 FOFC again: A pied-piping analysis of OE word order facts

Biberauer and Roberts (2005), (2006) develop a detailed LCA-based analysis of OE and ME word order facts that rules out the illicit order VO-Aux by appealing to the set of assumptions discussed in section 6.1.1.⁸⁴ They assume that OE T and v carry EPP-features which can be satisfied either (i) by attracting the minimal phrase containing the relevant goal (i.e., the subject or the object DP) or (ii) by pied-piping the projection that dominates the nominal goal. The latter option gives rise to roll-up movement, where the complement of the probing head H moves to the specifier of H. An apparently completely head-final configuration (i.e., S-O-V-Aux) is then the result of first moving V to (uniformly head-initial) v, followed by moving the remnant (headless) VP to the inner SpecvP (the outer specifier is occupied by the subject DP). Finally, the whole vP is pied-piped to SpecTP when T (hosting the finite auxiliary) probes the subject DP (this option is linked to the presence of a verbal head with rich inflection):

(157) **V-to-v, EPP in v, pied-piping of VP, pied-piping of vP to SpecTP: S-O-V-Aux**

$[\text{CP} [\text{TP} [\text{vP} \text{ subj.} [\text{v} [\text{VP} \text{ t}_v \text{ obj.}] [\text{v} \text{ V}+\text{v} \text{ t}_{\text{VP}}]]] [\text{T'} \text{ T} \text{ t}_{\text{vP}}]]]$

Assuming further that in OE, v hosts an EPP-feature only if this gives rise to certain information-structural distinctions (EPP-induced movement of the object leads to a defocused interpretation of the latter), the other major word order patterns can be derived as follows:

(158) a. **V-to-v, no EPP in v, subj. to SpecTP: S-Aux-V-O**

$[\text{CP} [\text{TP} \text{ subj.} [\text{T'} \text{ T} [\text{vP} \text{ t}_{\text{subj.}} [\text{v} \text{ V}+\text{v} [\text{VP} \text{ t}_v \text{ obj.}]]]]]]]$

⁸⁴ See Roberts (1997) for an alternative analysis of OE word order options based on Kayne (1994).

b. **V-to-v, EPP in v, pied-piping of VP, subj. to SpecTP: S-Aux-O-V**

$$[_{CP} [_{TP} \text{subj.} [_{_T} T [_{vP} t_{\text{subj.}} [_{_V} [VP t_V \text{obj.}] [_{_V} V+v t_{VP}]]]]]]]$$

Recall that under these approach, the unwanted order VO-Aux can be excluded if we further assume *radical removal* of head-initial VPs (in situ). According to Biberauer and Roberts, VO-order can only be derived when *v* does not carry an EPP-feature, leaving both the object and the VP in situ after V-to-*v* movement. Subsequent operations triggered by T's EPP-feature can then either derive (158a) (via moving the subject to SpecTP), or pied-pipe the whole vP to SpecTP. However, *radical removal* dictates that in the latter case, VP cannot accompany vP-movement to SpecTP and must stay behind in clause-final position. The only elements that can be moved to SpecTP are *v* and its edge. This derives the order S-V-Aux-O (without appealing to an extra mechanism such as extraposition):

(159) **V-to-v, no EPP in v, pied-piping of vP to SpecTP: S-V-Aux-O**

$$[_{CP} [_{TP} [_{vP} \text{subj.} [_{_V} V+v [VP]]]] [_{_T} T t_{vP} [_{VP} t_V \text{obj.}]]]]]$$

In other words, absence of an EPP-feature on *v* implies that the object appears in absolute clause-final position, and VO-Aux orders cannot be derived (see Biberauer and Roberts 2005 and Pintzuk 2005 for discussion of further word order options that can be derived by this analysis).

It is a crucial property of the analysis proposed by Biberauer and Roberts that it does not make use of the assumption of grammar competition. Instead, Biberauer and Roberts assume that the word order variation we observe in OE (and ME) is the result of grammatical options that are used to express information-structural distinctions (such as the basic focus-background structure of a clause) within a single grammar. First of all, note that this hypothesis abandons the insight that there is a link between language change and linguistic variation (with the latter possibly triggered by extensive language contact with Scandinavian and French invaders, cf. Kroch and Taylor 1997, 2000; Trips 2002, Fuß and Trips 2002). Instead, the amount of variation we find in the OE records is treated as an instance of 'principled variation' where each structural option is linked to a certain specific interpretation. In other words, it is assumed that the system we find in OE represents a stable synchronic system, where a large number of different surface patterns was produced by the syntactic computation to express information-structural distinctions. This raises the question of why none of the present-day Germanic languages exhibits a similar

amount of linguistic variation. In particular, as also pointed out by Biberauer and Roberts, it seems that each of the modern Germanic languages is confined to a subset of the parametric choices that were available in OE (e.g., the OV-character of German is analyzed in terms of generalized pied-piping, i.e., (157)). While this fact lends itself naturally to an analysis which assumes that certain competing parametric choices have been lost over time, it basically remains a mystery under the assumption that linguistic variation is always principled variation generated by a single grammar.

Another problem concerns the development of VO-orders. Recall that Biberauer and Roberts assume that objects in postverbal position are typically foci, while preverbal objects are defocused due to movement triggered by an (optional) EPP-feature in *v* (see Hinterhölzl 2004 for a related suggestion, but a slightly different Kaynian-style implementation; see also Roberts 1997 and Nunes 2002). However, this suggestion turns out to be problematic if we take into account quantitative data from OE. In particular, it appears that the frequency of post-verbal objects increases during the OE period (Pintzuk 2002, Pintzuk and Taylor 2006). In this connection, Pintzuk (2005: 122) points out that

“it is of course unlikely that this increase is due to speakers using more and more focused objects as time goes on. Therefore, postverbal position in Old English must have been used for constituents that were not focussed.”

Thus, it is not clear how an approach that assumes a one-to-one link between form/structure and interpretation can account for the changing frequencies of the relevant structural options in the linguistic output of individuals since it is unlikely that speakers of successive generations differ significantly with respect to the number of focused constituents they use. A way out of this dilemma would be to assume that the increase in the frequency of VO orders is not due to an increase in the frequency of focused constituents, but rather has to do with an independent change in which the postverbal position (and therefore the S-Aux-V-O option) gradually lost its status as a designated focus position. A first indication that this is a plausible alternative comes from changes affecting the frequency of examples with phonologically light elements such as pronouns, monosyllabic adverbs, and verbal particles in postverbal position (which Pintzuk 1999 takes to unambiguously reflect a VO-grammar). It is fairly clear that these elements do not constitute foci. According to Pintzuk (1999), the number of relevant examples is quite small in OE (only 18 of 712 relevant clauses in Pintzuk’s sample, i.e., around 2.5%), but increases rapidly

during the ME period. So we might say that the developing possibility of postverbal non-foci reflects a change in which the S-Aux-V-O option lost its link to a certain interpretation (namely, identification of focused constituents).

However, if it is true that it is difficult to detect clear interpretative differences (with respect to information structure) between the different grammatical options devised by Biberauer and Roberts, then of course the question arises of why the grammar produced this amount of variation, in particular if it is assumed that the presence/absence of EPP-features is directly linked to different interpretations.⁸⁵ On the other hand, the existence of word order variation that is not linked to any clear interpretative differences is predicted under the assumption of competing grammars that may generate different base orders which may be linked to identical interpretations.⁸⁶ The increase of the number of VO orders can then be attributed to the fact that over time, one particular grammar (expressing certain parametric choices) wins out over its competitors, both in the speech community and individual speakers.⁸⁷

⁸⁵ This general conceptual concern is supported by another empirical argument raised by Kroch and Taylor (2000). In a quantitative study of Early Middle English (EME), they show that not all SOV orders can be analyzed as the result of leftward movement of the object, and that some of these orders have to be analyzed as OV base orders. Kroch and Taylor demonstrate that only quantified DPs undergo regular leftward movement/scrambling in EME. With non-quantified DPs, however, it is a different matter: if DP-V_{fin}-Pron. is a diagnostic for leftward movement from a VO-base, then only 5% of the preverbal 'scrambled' objects are non-quantified DPs. However, the overall rate of OV-orders with non-quantified object DPs is much higher, namely around 30%. In other words, the few clear cases of leftward scrambling of non-quantified objects are much too rare to account for the high frequency of OV word order found in the EME texts considered by Kroch and Taylor and thus 25% of the preverbal non-quantified DPs have to be analyzed as being in their base positions (i.e., as OV base orders).

⁸⁶ As I understand it, the notion of grammar competition is not *a priori* incompatible with the assumption that the patterns generated by the different grammars are used for different information-structural purposes. See Fuß and Trips (2002) for some discussion. Moreover, additional operations may take place in each individual competing grammar to mark information-structural distinctions (scrambling, extraposition etc.).

⁸⁷ The slow start of the unambiguous VO option is actually expected if we combine the assumption of grammar competition with the generalization that the rise of head-initial projections has to take place in a 'top-down' fashion, first affecting higher functional categories before the change reverses the Head Parameter linked to the verb ((149) above, repeated here for convenience):

(i) (Comp) S O V Aux > (Comp) S Aux OV > (Comp) S Aux V O

If it is true that the development of a medial position for finite auxiliaries is a necessary precondition for a change affecting the Head Parameter of V, then we predict that in a language undergoing a

It should be noted, however, that these conceptual considerations have nothing to say about the specific analyses of OE word order patterns put forward by Biberauer and Roberts. Of course, their approach can easily be reconciled with the assumption of grammar competition if that is considered to be advantageous. One would simply have to assume that the individual grammatical options devised by Biberauer and Roberts constitute the relevant parametric choices in which the set of competing grammars differ. However, we have already noted in section 6.1.1 that there are some general empirical problems that can be raised against a pied-piping analysis (see Pintzuk 2005 for further critical discussion of a pied-piping analysis of OE data). At this point, I just want to add one further critical remark concerning the analysis of S-O-V-Aux orders under Biberauer and Roberts' approach.

Recall that S-O-V-Aux order (i.e., a uniformly head-final structure) is taken to result from repeated pied-piping (i.e., roll-up movement), in which (after V-to-v movement) the VP first moves to SpecvP, followed by vP moving to SpecTP:

(160) **V-to-v, EPP in v, pied-piping of VP, pied-piping of vP to SpecTP: S-O-V-Aux**

$[\text{CP} [\text{TP} [\text{vP} \text{ subj.} [\text{v} [\text{VP} \text{ t}_V \text{ obj.}] [\text{v} \text{ V}+\text{v} \text{ t}_{\text{VP}}]]] [\text{T} \text{ t}_{\text{vP}}]]]$

Crucially, Biberauer and Roberts (2005: 14) assume that this analysis carries over to S-O-V-Aux orders in modern German (which is therefore taken to make use of a subset of the grammatical options that were available in OE, see above). However, it appears that this claim is not unproblematic if we take a closer look at raising constructions in German. Note that according to Biberauer and Roberts, subject movement to SpecTP obligatorily involves pied-piping of the whole vP (containing the lexical verb), giving rise to S-O-V-Aux order. In addition, this analysis is taken to

change from basic OV to basic VO, the competing parametric choice of a head-initial V comes in late, at first manifesting itself only in a small number of cases. This is what we observe in OE: the early beginnings of a basic VO grammar. Subsequently, the change typically proceeds along the S-shaped curve of the logistic function (cf. e.g. Kroch 1989). In the initial stage after introduction/innovation of this option, the growth of its frequency is quite slow (as in the course of OE/EME). Then, there is a stage where growth is approximately exponential (as in ME); finally, the growth slows (in late ME) and stops when the change is eventually completed (i.e., when a certain parametric choice has won out over its competitors). Note that this course of change (its early stages, in particular) is exactly what we expect under the assumption of competing grammars in combination with (i). In contrast, it is much less clear how this course of change can be modeled under the assumption that the amount of variation is generated by a single grammar where different syntactic choices are associated with different interpretations.

explain the absence of subject expletives in German. At first sight, this analysis seems to give the correct results for raising constructions in German where the raising infinitive must always occur in preverbal position:

(161) a. dass [Peter das Rennen zu gewinnen] scheint
 that Peter the race to win seems
 'that Peter seems to win the race'
 b. *dass Peter scheint [das Rennen zu gewinnen]

As indicated by the bracketing in (161a), we might suspect that the whole embedded vP undergoes movement into the matrix clause (first to matrix SpecvP and then together with the matrix vP to matrix SpecTP) in order to check T's EPP-feature (a related analysis is proposed by Biberauer and Roberts to account for verb raising orders, i.e., Aux-V in embedded clauses). In addition, the assumption of obligatory vP-pied-piping seems to explain the ungrammaticality of (161b) where only the subject DP is raised, leaving the raising infinitive in its base position. However, there are some data that suggest that the subject actually may move to the matrix SpecTP on its own, casting some doubts on the assumption that pied-piping is obligatory. Consider the following examples, where the subject occurs to the left of a high speaker-oriented adverb:

(162) a. dass Peter offenkundig [das Rennen zu gewinnen] scheint
 that Peter obviously the race to win seems
 'that Peter obviously seems to win the race'
 b. dass Peter bedauerlicherweise [das Rennen zu gewinnen] scheint
 that Peter regrettably the race to win seems
 'that Peter regrettably seems to win the race'
 c. dass Peter leider [das Rennen zu gewinnen] scheint
 that Peter unfortunately the race to win seems
 'that Peter unfortunately seems to win the race'
 d. dass Peter erstaunlicherweise [das Rennen zu gewinnen] scheint
 that Peter astonishingly the race to win seems
 'that Peter astonishingly seems to win the race'

Crucially, in all the above examples, the high speaker oriented adverb is preferably interpreted as modifying the matrix verb *scheinen* and not the event described by the

raising infinitive.⁸⁸ This clearly shows that the subject DP can undergo movement to SpecTP without necessarily pied-piping the raising infinitive. Moreover, this possibility undermines the account of S-O-V-Aux orders proposed by Biberauer and Roberts: If the subject can move to SpecTP on its own, leaving the vP in situ, then we expect that S-Aux-O-V is a possible word order in embedded clauses with finite auxiliaries in German (Biberauer and Roberts assume that finite auxiliaries are located in T), which is not borne out by the facts:

(163) **V-to-v, EPP in v, pied-piping of VP, subj. to SpecTP: S-Aux-O-V**

[_{CP} [_{TP} subj. [_T [_{VP} _{t_{subj.}} [_v [_{VP} _{t_v} obj.]] [_v V+v _{t_{VP}}]]]]]]]

(164) *dass Peter hat das Rennen gewonnen

that Peter has the race won

‘that Peter has won the race’

To sum up, this section has shown that while being capable of generating the word order facts of OE, the FOFC-based pied-piping analysis proposed by Biberauer and Roberts (2005) raises a number of questions in addition to the general issues discussed in section 6.1.1 above. In particular, I have shown that there is evidence from raising constructions in German that the subject may move to SpecTP on its own, which casts some doubts on the analysis of S-O-V-Aux orders proposed by Biberauer and Roberts. Furthermore, I have argued that an analysis in terms of grammar competition can model the relationship between language change and linguistic variation (and the actual course of the OV-VO change) more adequately than an approach that attributes the amount of variation found in OE to a single grammar that is powerful enough to derive a multitude of surface orders connected to different information-structural distinctions. In the next section, I am going to develop an alternative account of the word order facts of OE which is based on the notion of grammar competition, but differs from the analysis proposed by Pintzuk with respect to the competing parametric choices that generate the linguistic variation found on the syntactic surface. In particular, I show that the problematic order VO-Aux can be ruled by the restrictions on *Edge Replacement* developed in the previous sections.

⁸⁸ Note that the relevant adverbs may also be marginally interpreted as having embedded scope. As pointed out to me by Günther Grewendorf, this can possibly be analyzed as a coherence effect linked to the restructuring properties of raising verbs.

6.2.2 Edge Replacement and OE word order

This section focuses on the (re-) ordering possibilities of the verb and its (nominal) complements in embedded clauses in OE. Following Pintzuk (1999), Haeberli (1999), Kroch and Taylor (1997), Lightfoot (1998) and others, I assume that it is unlikely that the kind of word order variation encountered in the relevant data set is generated by a single grammar. Rather, the amount of linguistic variation we encounter calls for an analysis in terms of several competing grammars that may exist side by side in the mind of an individual speaker (cf. Kroch 1989, 1994, 2001). In addition to grammar competition, the set of basic assumptions I adopt includes the theoretical proposals developed in section 4, in particular the *Root Raising Parameter*, a phonological Head Parameter (i.e., an exponent of a functional category may either be serialized to the left or to the right of the string of exponents with which it is combined), and the restrictions on possible linear orderings imposed by the workings of *Edge Replacement*:

(165) *The Root Raising Parameter*

- ✓ raises to category-defining v :
- YES (VO-languages: English, Italian, French etc.)
- NO (OV-languages: Japanese, German, Hindi etc.)

(166) *Phonological Head Parameter* (based on Richards 2004: 25)

- (i) Merge $(\alpha, \beta) \rightarrow \{\langle \alpha, \beta \rangle, \langle \beta, \alpha \rangle\}$, α a functional category.
- (ii) Upon Vocabulary Insertion,
 - a. 'VO': Ignore all $\beta > \alpha$ (i.e., $\{\langle \alpha, \beta \rangle, \langle \beta, \alpha \rangle\} \rightarrow / \phi_\alpha \phi_\beta /$)
 - b. 'OV': Ignore all $\alpha > \beta$ (i.e., $\{\langle \alpha, \beta \rangle, \langle \beta, \alpha \rangle\} \rightarrow / \phi_\beta \phi_\alpha /$)

(167) *Edge Replacement*

After the exponents of the overlapping part of structure have been linearized relative to material in $PF\text{-domain}_n$, they are replaced with the string of exponents assembled at $PF\text{-domain}_{n-1}$.

(168) *No-Tampering Condition on Edge Replacement*

Edge Replacement at $PF\text{-domain}_n$ may not disrupt ordering relations created between material of $PF\text{-domain}_n$ and material at the (overlapping) left edge of $PF\text{-domain}_{n-1}$.

Recall that by assumption, the *Root Raising Parameter* and the phonological Head Parameter are not completely independent of each other. In section 4.2 I have argued that a positive setting of the *Root Raising Parameter* implies a head-initial setting of the Head Parameter for category-defining *v* (giving rise to a basic VO grammar), while a negative setting implies a head-final setting of *v* (leading to an OV grammar) and requires Morphological Merger of *v* and the root at PF (i.e., ‘*v*-lowering’), see below for some discussion.

There are some indications that the word order variation found in embedded clauses of OE (and presumably in all kinds of clauses) does not involve competing settings of the Head Parameter for (exponents of) of T/INFL (in contrast to the analysis proposed by Pintzuk 1999). This claim is based on evidence from adverb placement which suggests that leftward moved finite verbs occupy different head positions in main and embedded clauses. Fuß and Trips (2002) show there is an asymmetry between main and embedded clauses that can be captured by the following descriptive generalizations (cf. Fuß and Trips 2002: 193f.):⁸⁹

(169) a. In main clauses, adverbs may not intervene between a subject pronoun in second position and a clause-medial finite Verb:

*XP – subject pronoun – (*adverb) – V_{fin} – [...]*

b. In embedded clauses, adverbs may intervene between a subject pronoun and a clause-medial finite Verb:

Comp – subject pronoun – (adverb) – V_{fin} – [...]

The following examples illustrate the relevant differences between main and embedded clauses with respect to adverb placement. In main clauses, a subject pronoun in second clausal position is always adjacent to the finite verb (Fuß and Trips 2002: 192):

⁸⁹ Note that the generalization in (169a) holds only for examples that display the subject pronoun in second position and another fronted XP in sentence initial position; if the pronoun itself is fronted to clause-initial position (i.e., topicalized), adverbs may intervene between the pronoun and the finite verb. Fuß and Trips (p. 192) point out that they found only three apparent counterexamples to this generalization (i.e., orders of the type *XP-pronominal subject-adverb-finite verb*) in the whole Brooklyn-Geneva-Amsterdam-Helsinki Parsed Corpus of Old English. All exceptions contain only a single finite main verb and should probably be analyzed as instances of SOV order in main clauses, presumably an archaic feature reflecting an older Pan-Germanic state. See chapter 3 for some discussion of relevant orders in Old High German.

(170) a. Mid þam haligan ele ge **scylan** þa hæþenan cild
 with the holy oil *you-PL* **should** the heathen child
 mearcian on þam breoste [...]
 mark on the breast
 'With the holy oil you should mark the heathen child on the breast.'
 (AELET3,148.5.317)

b. Nu *þu* **meaht** sweotole ongitan **pæt** **pæt** is good self.
 now *you* **can** openly understand that that is good self
 'Now you can openly understand that that is the good itself.'
 (BOETH,83.6.168)

c. gewislice *ic* **mæg** be him mare secgan.
 certainly *I* **can** about him more tell
 'Certainly, I can tell more about him.'
 (GREGD3,5.20.12.56)

In contrast, adverbs may intervene between the subject pronoun and a clause-medial finite auxiliary/modal in embedded clauses (Fuß and Trips 2002: 193):⁹⁰

(171) a. forðon *þu* nu **scealt** eft to lichoman hweorfan
 because *you* now **shall** again to body turn
 'because you should now turn again to the body'
 (BEDE,13.432.21.566)

⁹⁰ See Koopman (1991) for more examples of elements intervening between a pronominal subject and a clause-medial finite verb in embedded clauses, including cases that display other elements in the position between the subject pronoun and the finite verb: object pronouns as in (i) and a combination of adverbs and other XPs as in (ii).

(i) ðæt *we* hie **sculon** eac milde mode lufian
 that *we* them **must** also mild heart love
 'that we must also love them with mild heart'
 (CP (Cotton) 33.222.5; Koopman 1991: 118)

(ii) ðæt *hie* hiora ða nænne dæl **noldon** on hiora agen **geðiode**
 that *they* of them then no part **not wanted** into their own language translate
 'that they then did not wish to translate any part into their own language'
 (CPLet Wæf 36); Koopman 1991: 118)

According to Fuß and Trips (2002: 192f.), the example in (ii) suggests that the head position hosting the finite verb is to the right of the specifier occupied by nominal elements that have undergone object shift.

b. þa hie ða **hæfdon** Cirinen þa burg ymb seten
 when *they* then **had** C. the stronghold surround
 'when they had surrounded the stronghold C.'
 (OROSIU,66.17.62)

c. þæt hie þonan **mosten** to þæm sawlum becuman
 that *they* thence **must**-PAST to the soul come
 'that they thence had to turn to the soul'
 (OROSIU,102.14.191)

On the plausible assumption that pronominal subjects occupy a fixed position at the left edge of TP/IP in both main and embedded clauses (cf. Pintzuk 1999, Kroch and Taylor 1997, Haeberli 1999 and many others), the asymmetry stated in (169) suggests that the finite verb occupies a lower (functional) head position in embedded clauses. Furthermore, if the position of subject pronouns (at the left edge of TP/IP) is used as a diagnostic for the position of the finite verb, examples like (170) suggest that the finite verb occupies a position in the inflectional domain in main declaratives (cf. Cardinaletti and Roberts 1991, Kiparsky 1995, Eythórsson 1996, Kroch and Taylor 1997, Pintzuk 1999, Haeberli 1999, Fuß and Trips 2002, Fuß 2003; for a different view based on a split CP structure cf. Roberts 1996).⁹¹ Following Fuß and Trips (2002), Fuß (2003), I assume that the relevant functional head occupied by the finite verb in main declaratives is to be identified as T. Under this assumption, examples like (172) suggest that full DP subjects could stay in situ (in their theta-position, SpecvP), while the finite verb is located in (uniformly head-initial) T:

⁹¹ In contrast, inversion is obligatory with all kinds of subjects in the context of fronted operators such as *wh*-phrases or negation (see chapter 3 below for a detailed discussion of word order in main clauses of OE):

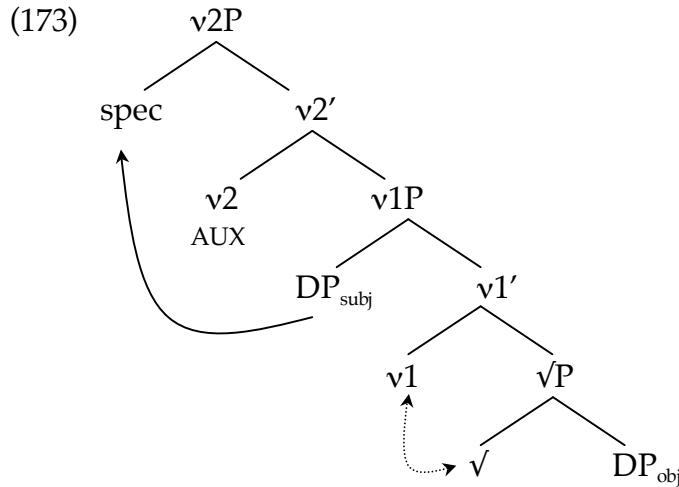
(i) Hu **begæst** þu weorc þin?
 how **go-about** *you* work your
 'how do you go about your work?'
 (Æcoll. 22; Kemenade 1987: 138f.)

(ii) Ne **mæge** we awritan ne mid wordum ascegan ealle þa wundra.
 NEG **can** *we* write NEG with words express all those wonders
 'we can neither write nor express with words all those wonders'
 (ÆLS 21.242; Kiparsky 1995, p. 147)

(172) Fela spella *him* **sædon** þa Beormas.
 many stories him told the Permians
 'the Permians told him many stories'
 (Oros., 14.27; Kemenade 1987:114)

The asymmetries between main and embedded clauses with respect to adverb placement suggest that the finite verb must occupy a lower head position in embedded clauses, which I uniformly take to be the highest *v* of a *vP/vP*-shell structure (see Biberauer and Roberts 2005, 2006 for further arguments against V-to-T movement in OE). Accordingly, it appears that basic patterns of word order variation in embedded clauses of OE should be attributed to a set of competing grammars that differ with respect to parametric choices concerning properties of the (extended) verbal projection, that is, *vP/vP*.⁹² Focusing for the moment on examples with complex verb forms (i.e., a combination of finite auxiliary and a non-finite verb, e.g., a participle), the basic chunk of structure we have to consider for a transitive verb is given in (173), which depicts a 'stacked' *vP*-structure consisting of two light verb projections (see van Gelderen 2004 for a related, but slightly different proposal).

⁹² While I follow Fuß and Trips (2002) in assuming that a good deal of word order variation in embedded clauses of OE involves competing parametric choices that concern the (extended) verbal projection of the clause, it should be noted that my proposal differs significantly from the analysis suggested by Fuß and Trips (2002). In particular, it appears that the latter account is based on a number of stipulations that are not motivated on independent grounds. First of all, to rule out the problematic order *VO-Aux*, Fuß and Trips stipulate that functional categories are uniformly head-initial (i.e., the Head Parameter is confined to lexical categories), raising certain questions about the analysis of clause-final complementizers. Furthermore, it is assumed that a separate, universally head-initial light verb projection is required only if *VP* is head-initial as well. In contrast, a separate *vP* is by assumption merely optional in the case of a head-final *VP*, which is supposed to account for the possibility of *S-Aux-O-V* orders in OE. Note that this weakens the general distinction between *OV* and *VO* grammars, which is much more clearly captured by the approach advocated here in terms of the *Root Raising Parameter* (plus an analysis that rules out *VO-Aux* by appealing to restrictions on *Edge Replacement*). Finally, in order to really rule out *VO-Aux* orders, the analysis put forward by Fuß and Trips requires the additional assumption that the Head Parameter for a given category is uniform for all members of that category; otherwise it should be possible to derive a structure in which a head-final *VP* dominates a head-initial *vP/VP*. This assumption seems to be contradicted by the fact that languages like German exhibit both head-initial and head-final adpositions (e.g., *über den Fluss* 'across the river' vs. *den Fluss entlang* 'along the river').



As illustrated in (173), I assume that the auxiliaries (e.g., *have* and modals) are located in the higher $v2$, while transitive participles enter the derivation as roots merged with their direct object. The resulting \sqrt{P} is then combined with a light verb ($v1$) that introduces the external argument (Larson 1988, Chomsky 1995, Kratzer 1996 and many others) and acts as a category-defining head with which the root must combine, either via syntactic movement or Morphological Merger at PF. Presumably, $v1$ also contains aspectual features that are realized upon Vocabulary Insertion by perfective prefixes such as *ge-* (cf. van Gelderen 2004). Furthermore, I assume that the class of auxiliaries and modals that occupy $v2$ and select $v1P$ as their complement are uniformly raising verbs, which attract the external argument merged in the lower $\text{Spec}v1P$ to $\text{Spec}v2P$ (see Wurmbrand 2001, van Gelderen 2004 on OE, Axel 2001 on Old High German).

Under these assumptions, the relevant set of competing grammars generating the amount of word order variation we find in OE can be identified as follows (note that a positive setting of the *Root Raising Parameter* defines a VO grammar, while a negative setting implies basic OV order):

(174) a. *S-O-V-Aux*

- (i) *Root Raising Parameter: NO*
- (ii) $v2$ head-final (i.e., “ignore all $v2 > \beta$ ” (i.e., $\{\langle v2, \beta \rangle, \langle \beta, v2 \rangle\} \rightarrow / \phi_\beta \phi_{v2} /$)

b. *S-Aux-O-V*

- (i) *Root Raising Parameter: NO*
- (ii) $v2$ head-initial (i.e., “ignore all $\beta > v2$ ” (i.e., $\{\langle v2, \beta \rangle, \langle \beta, v2 \rangle\} \rightarrow / \phi_{v2} \phi_\beta /$)

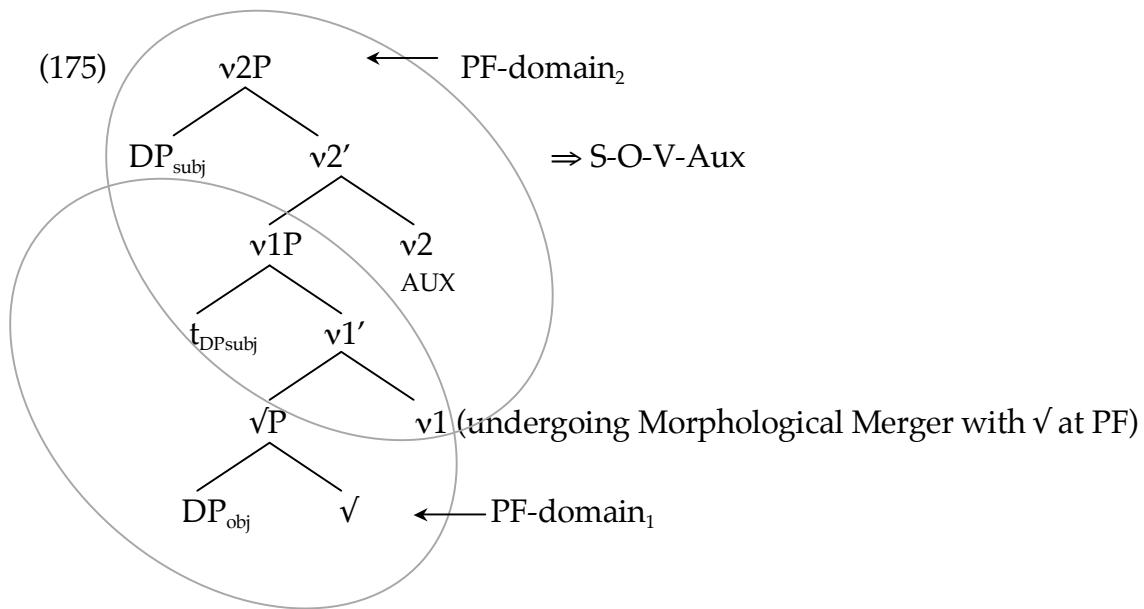
c. *S-Aux-V-O*

- (i) *Root Raising Parameter: YES*
- (ii) $v2$ head-initial (i.e., “ignore all $\beta > v2$ ” (i.e., $\{\langle v2, \beta \rangle, \langle \beta, v2 \rangle\} \rightarrow / \phi_{v2} \phi_\beta /$)

d. *S-V-O-Aux

- (i) *Root Raising Parameter*: YES
- (ii) v_2 head-final (i.e., “ignore all $v_2 > \beta$ ” (i.e., $\{<v_2, \beta>, <\beta, v_2>\} \rightarrow / \phi_\beta \phi_{v_2} /$)

As we will see shortly, the unattested order *S-V-O-Aux can be ruled out on principled grounds since it involves parametric choices which give rise to a structure that cannot be linearized by the workings of Vocabulary Insertion/LIN. But let us first take a closer look at the word order patterns (and parametric choices) attested in OE, beginning with the (presumably most ancient) pattern S-O-V-Aux.⁹³

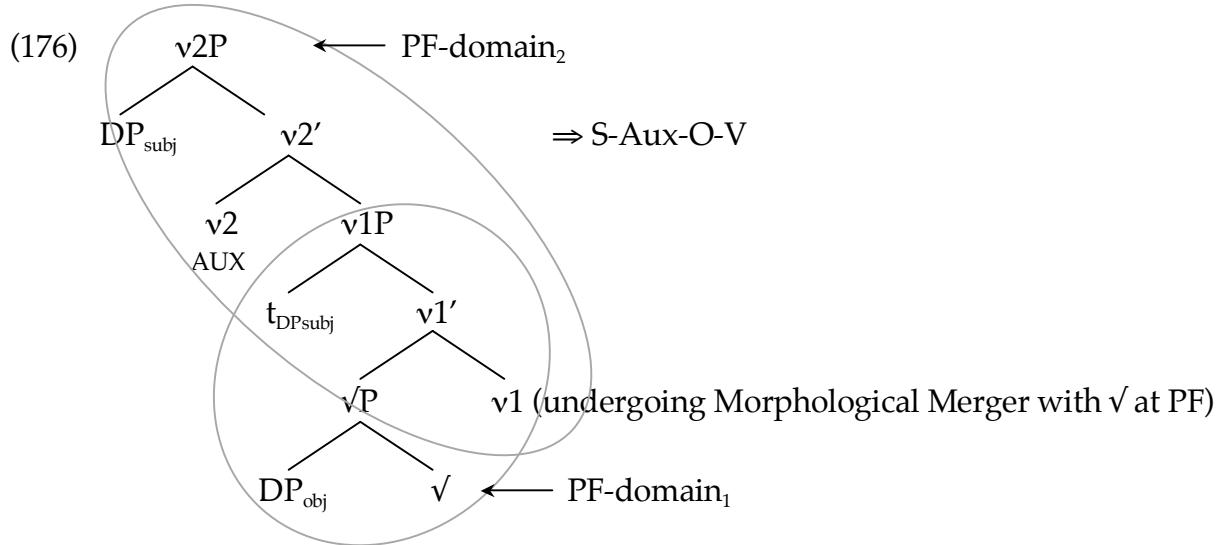


As illustrated in (175), the pattern S-O-V-Aux is generated by the set of parametric choices in (174a), that is, a negative setting of the *Root Raising Parameter* combined with a head-final v_2 . In other words, the root does not raise to category-defining v_1 in the syntax. Rather, v_1 undergoes lowering to \sqrt in the PF-branch of grammar. At Vocabulary Insertion, the $\sqrt+v_1$ complex is then spelled-out by inserting a phonological exponent linked to a head-final setting of the Head Parameter (recall that by assumption, the Head Parameter is expressed by exponents of functional categories, in the case at hand the exponent of v_1). As indicated in (175), I assume that v_1P and v_2P correspond to two different phonological domains (with v_1 and its specifiers constituting the overlap). Due to the fact that the overlap between PF-domain₁ and PF-domain₂ does not contain any overt material at the point of

⁹³ For expository reasons, the following phrase markers differ with respect to the headedness of the relevant heads, although this should not be taken to imply that linear order is directly coded in the syntax. Recall that by assumption, syntactic structures contain only hierarchical information.

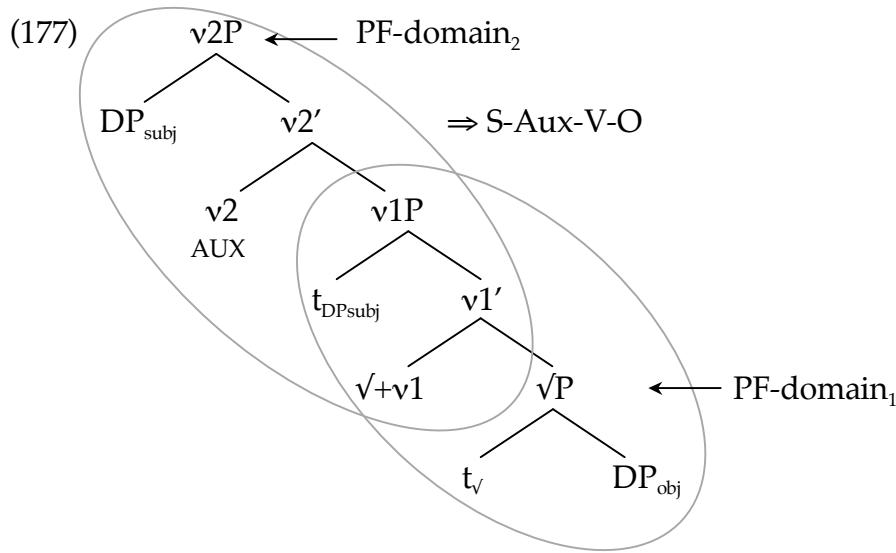
Vocabulary Insertion (i.e., after PF-lowering of v_1 to $\sqrt{ }$), the No-Tampering Condition does not impose any restrictions on *Edge Replacement*, and the string of exponents inserted to the lower PF-domain₁ (i.e., $/\phi_{\text{obj}} \phi_{\sqrt{v+v_1}}/$) can be linearized with the exponent of v_2 , which is head-final as well. Addition of the Vocabulary items realizing Specv2P (which is uniformly to the left) then gives rise to the order S-O-V-Aux.

The same syntactic structure is at the basis of the pattern S-Aux-O-V (derived by grammar (174b)), the only difference being the setting of the Head Parameter for exponents of v_2 (finite auxiliaries, e.g., *have* and modals):

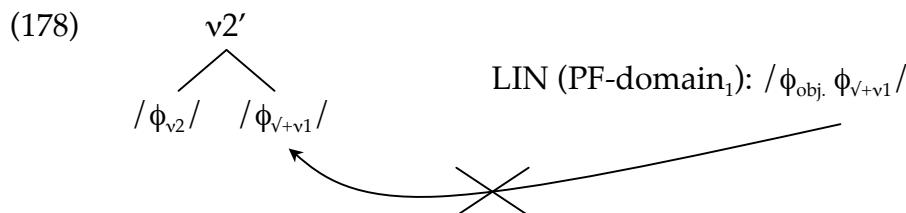


Again, (head-final) v_1 is lowered to $\sqrt{ }$ at PF due to a negative setting of the *Root Raising Parameter*. Similar to (175), this serves to evacuate the overlap between PF-domain₁ and PF-domain₂ (in combination with syntactic movement of the external argument to Specv2P). As a result, *Edge Replacement* can freely apply, inserting the string $/\phi_{\text{obj}} \phi_{\sqrt{v+v_1}}/$ at the overlap. The ordering S-Aux-O-V is then derived by adding the exponents of (head-initial) v_2 and Specv2P.

Grammar (174c), which generates the uniformly head-initial pattern S-Aux-V-O differs minimally from (174b)/(176) in the setting of the *Root Raising Parameter*. As a result, $\sqrt{ }$ raises to category-defining v_1 in the syntax, giving rise to the following structure:

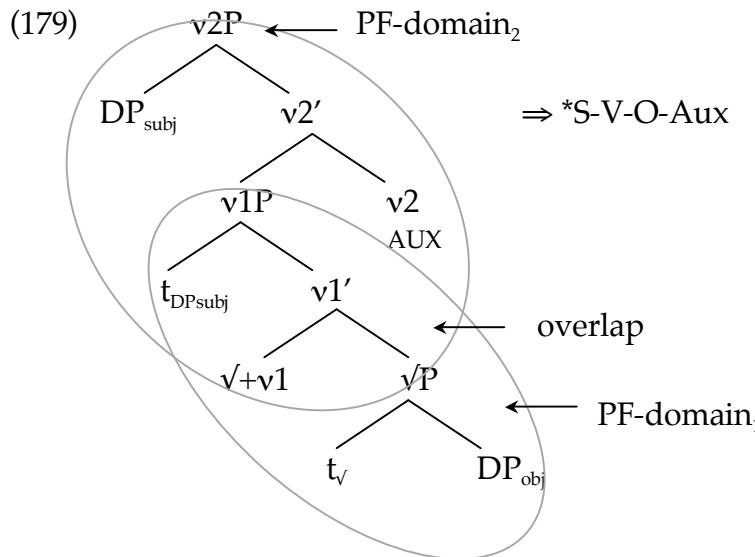


As can be seen from (177), the crucial difference between (176) and (177) comes from the position of the $\sqrt{+v1}$ complex which has been created by syntactic raising of the root. A positive setting of the *Root Raising Parameter* leads to the presence of overt material at the overlap and implies a head-initial setting of the Head Parameter for the exponent of the $\sqrt{+v1}$ complex. Note that the interdependence of Root Raising and headedness follows (at least partially) from the *No-Tampering Condition on Edge Replacement*: If $\sqrt{ }$ raised to $v1$ were realized by a head-final exponent (inserted to the right of the exponent(s) of the object), *Edge Replacement* would be blocked in case $v2$ is realized by a head-initial constituent, since this would disrupt adjacency of the exponents of $v2$ and the $\sqrt{+v1}$ complex:

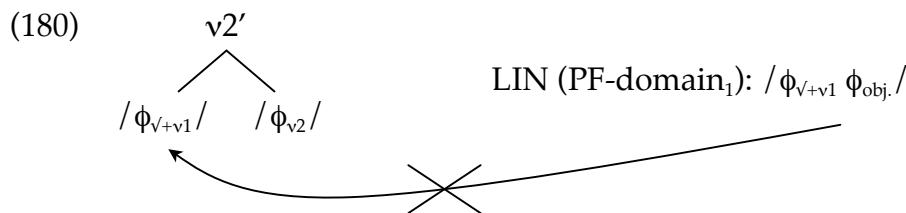


However, no problem arises as long as the exponents of both $v2$ and the $\sqrt{+v1}$ complex attach to the left of the existing string of exponents created so far, giving rise to S-Aux-V-O when Vocabulary Insertion/LIN applies to the structure in (177).

Let's now turn to the question of how the non-attested word order option *S-V-O-Aux can be excluded under the present approach. What I am going to argue is that the non-attested combination of parametric choices (174d) is ruled out since it gives rise to a non-linearizable structure that violates the *No-Tampering Condition on Edge Replacement*. Consider the following structure:



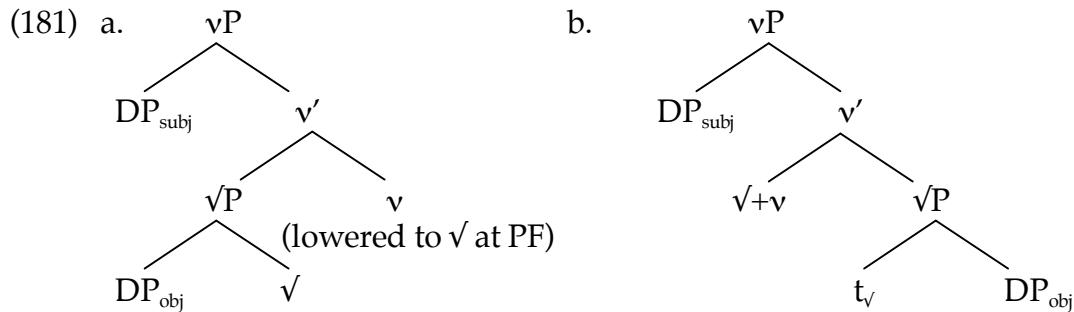
Note that (179) differs minimally from (174d)/(177) in the setting of the Head Parameter linked to exponents of v2 (*have*, modals etc.). As a result, the relevant Vocabulary items must follow the string of exponents created by previous applications of Vocabulary Insertion. The problematic fact about (179) concerns the make-up of the overlap and the way the settings of the Head Parameter for v2 and v1 interact in the course of *Edge Replacement*. Recall that *Edge Replacement* serves to linearize two neighboring PF-domains via replacing the overlap between the relevant PF-domains with the string of exponents assembled at the 'lower' PF-domain. This process is restricted by a no-tampering condition that requires that *Edge Replacement* must preserve ordering relations that have been created between elements at the overlap (in (179), material inserted to the $\sqrt{+}v1$ complex) and exponents inserted to terminal nodes of the higher phonological domain (v2 in (179)). Now, note that according to the set of parametric choices under investigation (i.e., (174d)), v2 is head-final, while v1 is head-initial. At PF-domain₂ this establishes a linear ordering in which ϕ_{v2} immediately follows $\phi_{\sqrt{+}v1}$ (the exponent inserted to the overlap). However, due to the fact that v1 is linked to a head-initial setting of the Head Parameter, application of *Edge Replacement* disrupts adjacency of $\phi_{\sqrt{+}v1}$ and ϕ_{v2} when $\phi_{\sqrt{+}v1}$ is replaced with the string of exponents assembled at PF-domain₁:



Thus, (180) is ruled out by the Non-Tampering Condition on *Edge Replacement*, and no linear ordering can be established between PF-domain₁ and PF-domain₂. In other

words, the specific combination of parametric choices that in principle may give rise to the non-attested ordering S-V-O-Aux (i.e., (174d) is illicit since its output cannot be processed by the workings of Vocabulary Insertion/LIN.⁹⁴

Less variation is expected in examples that contain only a single finite lexical verb since in these cases the structure lacks a separate v2P. Accordingly, the relevant competing grammars differ only in the setting of the *Root Raising Parameter*:



This raises the question of how the present approach can account for examples with finite lexical verbs and verbal particles in which the object appears in between the finite verb and the particle ((153a) above, repeated here for convenience):

(182) þæt he **wærp** þæt sweord **onweg**.
 so-that he **threw** that sword **away**
 'so that he threw away the sword'
 (Bede 38.20; Pintzuk 1999: 57)

According to Pintzuk (1999), in examples like (182) the clause-final particle marks the base-position of the finite verb, which has undergone movement to a medial functional head. At first sight, this seems to be incompatible with both (181a) and (181b). By assumption, (181a) corresponds to a uniformly head-final grammar, where objects always usually appear to the left of the verb (both in the base and after scrambling, if the latter was an option in OE). Moreover, it seems that (182) cannot be analyzed as an instance of Verb Projection Raising (VPR) (operating on (181a)), since rightward movement of particles is generally excluded in the modern Germanic OV-

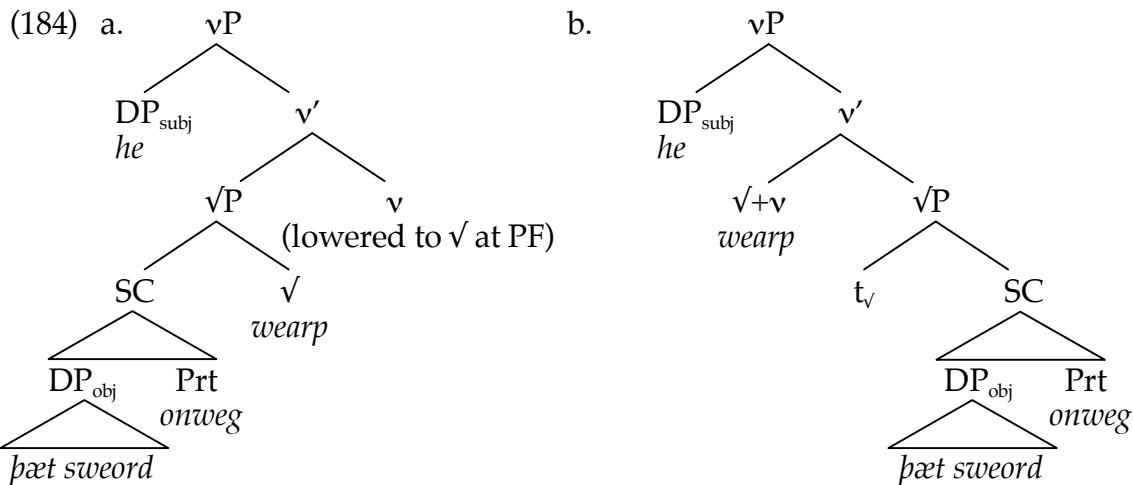
⁹⁴ Under these assumptions, the possibility of DP-extrapolation in examples like (i) (which is highly marked or even ungrammatical in the present-day Germanic OV-languages) can be analyzed as a PF-repair mechanism that relocates the object to the right edge of PF-domain₂ (perhaps as an instance of Local Dislocation), thereby rescuing structures that violate the *No-Tampering Condition on Edge Replacement*:

(i) þæt he **miltsian** **t_i wolde** [his agenum slagum]
 that he pity would his own executioners
 'that he would pity his own executioners'
 coaelive,+ALS_[Exalt_of:Cross]: 181.5692; Pintzuk 2005: 118)

languages. On the other hand, it also seems to be unlikely that (182) can be derived on the basis of (181b) since there is apparently no position to the right of the object that could be occupied by the particle. However, note that orders such as (182) are still possible in present-day English as a result of so-called particle-shift (this has already been pointed out by Kemenade 1987, but see Pintzuk 1999 for some critical discussion):

(183) a. that he threw his sword away
 b. that he threw away his sword

Following Bowers (1993), Svenonius (1994, 1996) and den Dikken (1995), particle-shift can be analyzed in terms of a structure in which a head-initial verb selects a small clause which contains the particle and the object. Thus, (182) and a made-up example corresponding to a uniformly head-final structure can be analyzed by the following pair of structures (where the object occupies the subject position of the small clause):

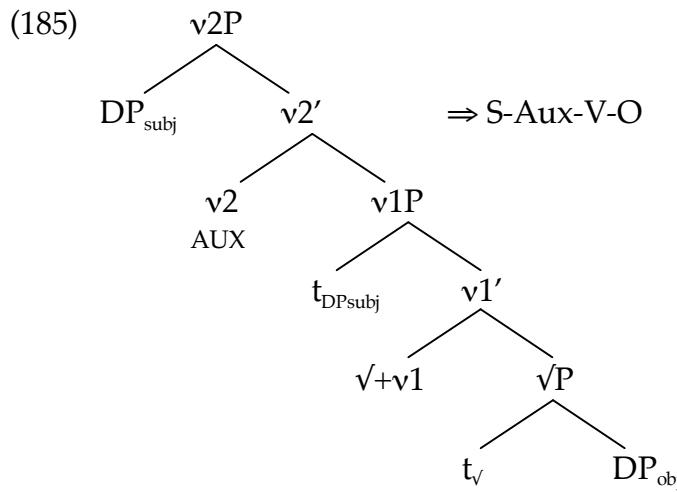


Summing up, in this section I have analyzed word order variation in OE in terms of (i) a 'stacked' v2P-v1P structure for complex verb forms and (ii) competing grammars that differ with respect to the setting of the Head Parameter for the exponent of v2 (the insertion site for auxiliaries and modals) and the *Root Raising Parameter*. The latter determines not only the structural position of the v+v1 complex (formed in the syntax or at PF), but also the basic ordering relation between lexical verbs and their complements. Under these assumptions, the non-attested ordering *S-V-O-Aux can be ruled out as a violation of the No-Tampering Condition of *Edge Replacement* that results from the presence of a head-initial exponent at the overlap in combination

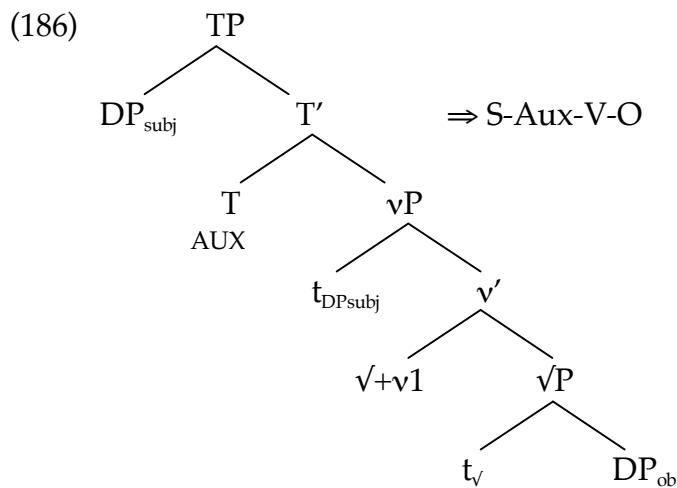
with a head-final setting of the Head Parameter linked to the head of v2P immediately dominating the overlap. In addition, I have taken a brief look at some other word order patterns that can be found in embedded clauses of OE (e.g., with finite lexical verbs), sketching out some preliminary proposals of how these can be subsumed under the approach developed in this section. Of course, this does not begin to represent a comprehensive account of the whole variety of word patterns attested in OE (see e.g. Pintzuk 1999, 2005, Biberauer and Roberts 2005). In particular, more has to be said about the reordering possibilities in double object constructions (see Pintzuk 2005), the role of extraposition and scrambling in deriving additional patterns, and the position of other parts of speech such as adverbs, complement clauses, and negation. These are matters that I hope to address in future work, but see chapter 3 below on some aspects of word order in main clauses of OE. In the next section, I will take a brief look at changes that affected the system described so far in the Middle English period, leading to the development of a strict VO grammar.

6.2.3 Some remarks on the rise of VO in the Middle English period

It is a well-known fact that English underwent a number of major changes in the Middle English (ME) period. One of the most prominent changes concerns the loss of OV structures (concerning both the position of auxiliaries and lexical verbs), which has attracted a lot of attention in the theoretically informed literature on the history of English (cf. e.g. Stockwell 1977, Kemenade 1987, Lightfoot 1991, Roberts 1993, 1997, 2007, Pintzuk 1999, Haeberli 1999, Kroch and Taylor 2000, Fischer et al. 2000, Trips 2002, Biberauer and Roberts 2005, 2006). Under an approach that attributes linguistic variation in OE to competition between different parametric options, the rise of a strict VO-grammar can be modeled in terms of a loss of competing grammars (see also Pintzuk 1999). At first sight, this seems to suggest that the parametric choice that eventually won out over its competitors is represented by the grammar (174d), which combines a positive setting of the *Root Raising Parameter* with a head-initial setting for exponents of v2 hosting auxiliaries and modals, that is, a uniformly head-initial grammar:



However, it is fairly clear that (185) does not correctly represent the structure of Modern English, for which it is generally assumed that auxiliaries are merged directly in an inflectional head (T or Infl). Furthermore, subjects are commonly taken to move obligatorily to SpecTP/IP, giving rise to the following phrase marker (cf. e.g. Chomsky 1995, Radford 2004):



So it seems that in addition to the loss of other competing parametric options, the development of Modern English involved at least two further changes, namely the development of obligatory subject movement to SpecTP, and a change in which auxiliaries were reanalyzed as realizations of an inflectional head (presumably T). In this section, I take a brief look at the historical developments that led to this outcome, focusing on the changes that affected the placement of verbal elements (see chapter 3 for some remarks on the rise of subject movement).

Any account of the relevant changes must address two basic questions, namely (i) What were the specifics of this change, that is, how did it proceed? and (ii) Which were the causes of this change? In other words, in which ways did the triggering

experience change so that OV order failed to be acquired? Let's first take a brief look at (i), that is, the course taken by the relevant changes.

In early generative studies (Kemenade 1987, Lightfoot 1991) it is usually assumed that the change from OV to VO was largely completed already in early ME, i.e., around 1200. However, more recent studies have shown that VO orders began to vastly outnumber OV only after 1300, and that some amount of OV orders continued to show up until about the 16th century (cf. e.g. Kroch and Taylor 2000, van der Wurff and Foster 1997).

As is well-known, the change was accompanied by quite some amount of variation between OV and VO orders, both in OE and ME (cf. e.g. Pintzuk 1999, Kroch and Taylor 2000, Trips 2002). However, it appears that in contrast to OE, the variation is primarily confined to the position of non-finite verbs in ME, whereas finite verbs, and in particular auxiliaries mostly occur in clause-medial position. This is illustrated with the following example from the Early Middle English (EME) *Ormulum*:

(187) Forr þatt I wollde bliþelig þatt all Ennglisse lede wiþþ ære shollde
 For that I would gladly that all English people with ear should
 lisstenn *itt*, wiþþ herte shollde *itt* trowwenn, wiþþ tunge shollde
 listen *it*, with heart should *it* trust, with tongue should
 spellenn *itt*, wiþþ dede sholde *itt* follghenn.
 spell *it*, with deed should *it* follow.
 (CMORM,DED.L113.33; Trips 2002: 112)

Note that OV order continues to be a grammatical option with negative and quantified objects well into the 15th century (cf. van der Wurff 1997, 1999, Kroch and Taylor 2000, Ingham 2002, Pintzuk 2002):

(188) Pei schuld [no meyhir] haue.
 they should no mayor have
 'They should have no mayor.'
 (Capgrave Chronicles 62.23; Fischer et al. 2000: 163)

(189) He hap on vs mercy, for he may [al þynge] do.
 he has on us mercy for he can everything do
 (Barlam 2740; van der Wurff 1999: 8)

This fact suggests that a number of OV orders found in OE were also derived by leftward movement of quantified and negated complements (cf. Roberts 1997, 2007a, Biberauer and Roberts 2005).

A brief glance at the (vast) literature on the transition from OV to VO in the history of English reveals a couple of factors that may have played a role in the change under investigation. Early explanations of the change from OV to VO attributed the reversal of the Head Parameter to a reanalysis of surface word orders derived by (massive) extraposition in the target grammar (Stockwell 1977, Kemenade 1987). This analysis is based on the assumption that non-OV orders in OE were the result of rightward movement (extraposition) that could affect all kinds of elements, including DPs, PPs, and VPs (quite similar to extraposition in the modern Germanic OV languages):

(190) *Extraposition*

þæt ænig mon t_i atellan mæge [ealne þone demm] $_i$
 that any man relate can all the misery
 'that any man can relate all the misery.'

(Oros., 52.6-7; Pintzuk 1993: 14)

(191) *Verb raising*

þæt he Saul ne t_i dorste [ofslean] $_i$
 that he Saul NEG dared murder
 'that he didn't dare to murder Saul'
 (Oros., 52.33; Kemenade 1987: 59)

(192) *Verb Projection Raising*

þæt he t_i mehte [his feorh generian] $_i$.
 that he could his property save
 'that he could save his property.'
 (Oros., 48.18; Kemenade 1987: 59)

It is then assumed that for stylistic reasons, extraposition came to be more and more frequently used in the course of OE and EME. Eventually, it was extended to an ever wider range of phrases, including 'light' DPs, in particular:

(193) Pu hafast gecoren [_{DP} þone wer].
 thou hast chosen the man
 (ApT 34.23; Fischer et al. 2000: 148)

So we might suspect that surface VO orders derived by extraposition weakened the evidence for a basic OV option, which eventually gave rise to a reanalysis in which learners took surface VO orders (derived by rightward movement in the target grammar) to represent the underlying, basic word order.⁹⁵

Another factor that possibly blurred the evidence for a basic OV grammar involves the variable position of verbal particles in OE. As pointed out by Lightfoot (1991), in languages with V2 order in main clauses, underlying OV order can only be detected indirectly by inspecting the position of non-finite verbs and verbal particles that signal the base-position of the verb (relative to objects and elements such as negation and certain adverbs), as illustrated in the following German examples:⁹⁶

(194) a. Peter hat den Professor angerufen.
 Peter has the professor up-called
 b. Peter rief den Professor an.
 Peter called the professor up
 c. Peter rief nicht an.
 Peter called not up

⁹⁵ While intuitively quite plausible, this account can be shown to raise a number of questions that led many researchers to doubt whether the reanalysis of extraposition structures represents a likely scenario for the change under investigation. First of all, it is unclear, why (and how) OE speakers could extend the option of extraposition to light elements that resist this operation in all present-day Germanic languages (including English; note that even extraposition of heavier DPs as in (190) is at least highly marked in present-day OV-languages such as German or Dutch). Furthermore, Pintzuk (1999) shows, based on a thorough quantitative analysis, that there actually was no significant increase in the rate/frequency of extraposition in the relevant period (i.e., OE/EME). Therefore, it is unlikely that the reversal of the Head Parameter can be attributed solely to stylistic changes affecting the frequency of extraposition.

⁹⁶ Note that the scenario developed by Lightfoot is based on the notion of 'Degree-0-Learnability', which states that for the purposes of parameter setting, learners may access only information contained in unembedded contexts.

Lightfoot demonstrates that at least two of these indicators were either absent or unclear in OE. First, clausal negation is a proclitic element attaching to the verb; second, verbal particles could be fronted together with the verb:

(195) Ne meaht þu deman Gallia biscopas buton heora agenre
 NEG might you judge Gaul's bishops but their own
 alderlicnesse, ...
 authority

'You might not judge the Gaul's bishops but their own authority.'

(Bede, *Bede_1:16.74.5.679*)

(196) Stephanus up-astah þurh his blod gewuldorbeagod.
 Stephanus up-rose through his blood glory-crowned
 (Hom. of the Anglo-Saxon Church I, 56; Lightfoot 1991: 61)

Furthermore, recall that more generally, the learner was confronted with a vast variety of different placement options for verbal particles that presumably weakened the status of particles as indicators of basic OV order (either due to grammar competition or different grammatical options linked to different interpretations, as proposed by Biberauer and Roberts 2005).⁹⁷

Furthermore, it has been proposed that the change from OV to VO was influenced by independent changes that led to the loss of case morphology (cf. e.g. Kemenade 1987, Roberts 1997, Kiparsky 1997). Roberts (1997) develops a Kaynian account of the OV-VO change, which is based on the assumption that in OE, OV order was derived by leftward movement of the object NP (or other material). The loss of the relevant movement operation (and therefore the rise of VO) is then linked

⁹⁷ Furthermore, there was a further indicator of OV order in OE, namely the possibility of verb-final main clauses:

(i) he Gode þancode.
 he God thanked
 'He thanked God.'

(Beowulf 625; Lightfoot 1991: 63)

Lightfoot (1991: 65ff.) shows that there was a sharp decline of verb-final main clauses during the OE period; in addition, there was an increase of clauses where a verbal particle is fronted together with finite verb.

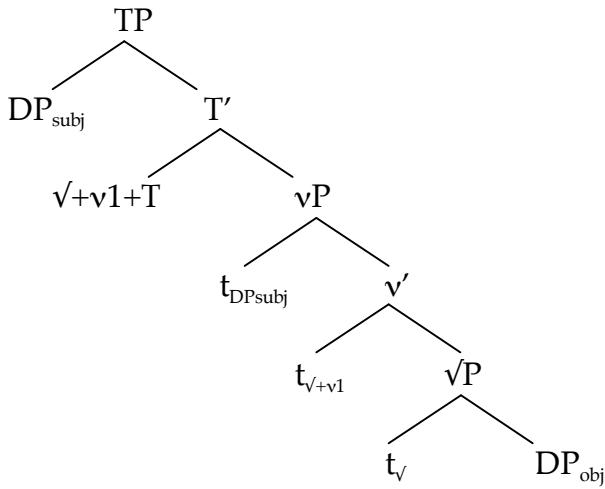
to the independent loss of its morphological trigger (i.e., rich case morphology).⁹⁸ An alternative perspective on the correlation between the rise of VO-order and the loss of case distinctions is put forward by Kiparsky (1997), who suggests that the loss of distinctive case endings required a fixed structural position for licensing the verb's arguments and in order to discriminate grammatical functions.⁹⁹

Suppose that in addition to the above mentioned factors that blurred the evidence for a basic OV grammar, the change from VO to VO was promoted by aspects pertaining to the linearization of syntactic structures in line with the theory developed in the previous sections. What I want to propose is that the loss of parametric options linked to an OV grammar was connected to independent changes that affected properties of T in English, namely (i) the development of obligatory (EPP-driven) subject movement to SpecTP, and (ii) the rise of generalized V-to-T-movement. Under the assumption that T was uniformly head-initial in OE and ME (see section 6.2.2 above, and in particular chapter 3), these changes gave rise to the following structure for clauses with finite lexical (transitive) verbs:

⁹⁸ In addition, Roberts (1997) suggests that derivational economy principles promoted the change from OV to VO. By assumption, the learner chooses the least complex structural option compatible with the input data. In that sense, the non-movement operation – that is, VO – was selected, because it involved less movement operations than the OV option (see also Roberts 2007a). Note that this bears some resemblances to the Transparency Principle of Lightfoot (1979), which may set off a reanalysis if the acquisition task becomes too complex (e.g., when a certain structure or element involves too many exceptional features or rules). A related Kaynian analysis of the change in question has recently been proposed in work by Biberauer and Roberts (2005), (2006) who argue that the loss of O-V-Aux patterns resulted from the combination of (i) the loss of the pied-piping option for satisfying T's and v's EPP feature (see section 6.2.1) and (ii) the complete loss of a EPP-feature associated with v (triggering object/VP movement in OE).

⁹⁹ This proposal seems to be supported by the observation that cross-linguistically, SOV order is linked to the presence of distinctive case endings (cf. e.g. Greenberg's 1963 Universal 41), while languages without case endings typically show SVO order (cf. Venneman 1974, Kiparsky 1996, 1997, Roberts 1997, 2007a for discussion).

(197)



Without going into the details of the latter two developments,¹⁰⁰ there are reasons to believe that they had a crucial impact on the set of possible linear orderings. In particular, consider the following restriction on the process of *Edge Replacement* proposed in section 6.1:

(198) In cases where a head α undergoes syntactic movement to a head β , the resulting structure cannot be linearized iff:

- the exponents of α and β differ with respect to the Head Parameter, and
- the string of exponents linearized relative to the $\alpha+\beta$ complex corresponds to the complement of α .

As noted above, (198) requires that the setting of the Head Parameter must be identical for exponents of T and an exponent of v moved to T. Hence the grammars that resulted from the innovation of syntactic v-to-T movement were incompatible

¹⁰⁰ It seems likely that the rise of v-to-T movement in embedded clauses was promoted by the fact that main clauses exhibited regular v-movement to head-initial T. Fuß and Trips (2002: 212) propose the following change, in which head-initial patterns that were derived vP-internally in the target grammar were reanalyzed as involving movement to SpecTP and T, respectively:

- $[_{CP} \text{comp} [_{TP} \emptyset [_{T'} T [_{vP} \text{subject} [_{v'} v+v [_{vP} t_V \text{object}]]]]]] \Rightarrow$
- $[_{CP} \text{comp} [_{TP} \text{subject} [_{T'} [V+v]+T [_{vP} t_{subj} [_{v'} t_{v+v} [_{vP} t_V \text{object}]]]]]]$

See Kiparsky (1997), Fuß (2003), Biberauer and Roberts (2005), (2006), Trips and Fuß (2008), and chapter 3 below for additional factors triggering the development of EPP-driven subject movement; see Biberauer and Roberts (2005), (2006), and (2008) on the rise of V-to-T movement (due to a reanalysis of subject-initial V2 patterns as involving movement to SpecTP and T, respectively).

with a head-final setting of the Head Parameter for v.¹⁰¹ Over time, when the innovating grammar spread through the speaker community, this further contributed to weakening the status of the competing OV-option, eventually leading to a point where the trigger experience had changed in a way so that underlying OV order could no longer be acquired.

Summing up, this section has suggested that in the ME period, a set of different developments contributed to the loss of OV, leading to a point where the OV setting could no longer be acquired based on the evidence available to the learner. In addition to factors that gradually changed the make-up of the triggering experience and obscured the evidence for a OV grammar (e.g., extraposition, fronting of verbal particles, and loss of case distinctions), the transition from OV to VO was shaped by 'hard-wired' properties of the mapping from syntax to morphology/phonology. More precisely, I have suggested that the independently motivated rise of (i) subject movement to SpecTP and (ii) v-to-T movement blocked the acquisition of an underlying OV option in the innovating grammars, since this combination of parametric choices gives rise to syntactic structures that cannot be linearized by workings of Vocabulary Insertion/LIN.

7 Summary

In this chapter I have examined the interface of syntax and morphology (or, rather, of syntax and the phonological component of grammar), focusing on the way PF deals with the cyclic output of the syntactic computation and the creation of linear orderings in this process. We have seen that an investigation of the way syntactic hierarchical structures are mapped to linear orderings does not only deepen our understanding of certain facts about word order in living languages (e.g., the apparent cross-linguistic absence of VO-Aux orders), but also provides new insights into the ways grammars/languages may change over time. I have proposed that the linearization process is part of the operation of Vocabulary Insertion, which supplies syntactic terminal nodes with phonological material and thereby incrementally builds a linear string of phonological exponents. The decision whether to add a phonological exponent to the left or to the right of the existing string of elements is determined by a phonological Head Parameter which is taken to ignore a subset of

¹⁰¹ Again, this suggests that the apparent connection between V-to-T movement and directionality (i.e., only head-initial languages exhibit verb movement to inflectional heads) can perhaps be attributed to the mechanisms that map hierarchical structures to linear orderings.

the symmetric c-command relations established in the syntax. Assuming a model of cyclic Spell-Out, I have argued that the phonological component recombines the cyclic output of the syntax into larger and partially overlapping phonological domains. The overlap between neighboring phonological domains can then be used to create a linear ordering between the chunks of structure transferred to the phonological component. The central proposal I have put forward is that the linearization of separate phonological domains involves a process called *Edge Replacement* that substitutes the right edge of a phonological domain with the string of exponents created so far. By assumption, this process is subject to a no-tampering condition that requires *Edge Replacement* to preserve ordering relations that have been established between elements at the overlap (between neighboring phonological domains) and higher exponents. I have shown that an approach in terms of *Edge Replacement* is more successful than recent LCA-based analysis (Biberauer et al. 2007, 2008) in deriving a set of generalizations on possible word orders (e.g., the cross-linguistic absence of VO-Aux or the correlation between complementizer position and position of complement clause). Furthermore, I have argued that it is possible to predict a typology of possible and impossible grammars if we combine *Edge Replacement* with certain assumptions about the parametric differences between OV and VO languages (*The Root Raising Parameter*). This model of linearization not only imposes a number of restrictions on possible combinations of parametric choices, but also makes a number of predictions on possible pathways of grammar change (as demonstrated by a discussion of the OV-VO change in the history of English). In what follows, I recapitulate the contents of the individual sections of this chapter in some more detail.

Section 2 has served to establish basic properties of the grammar model adopted in this work (Distributed Morphology), arguing that syntactic terminal nodes are supplied with phonological information in a post-syntactic operation (Vocabulary Insertion) that also serves to establish linear order (the *Late Linearization Hypothesis*). In addition, I have demonstrated that the constituent structure derived in the syntax can be modified by a limited number of post-syntactic operations such as Morphological Merger (giving rise to the impression of syntactic lowering), or Local Dislocation, which reverses the linear order of adjacent phonological exponents.

Section 3 has investigated a set of issues related to the syntax-PF interface that follow from the assumption of cyclic Spell-Out. Focusing on the phenomenon of complementizer agreement in Germanic, I have shown that operations of the phonological component may cut across the Spell-Out domains as defined in

Chomsky (2000) and subsequent work. In particular, we have seen that certain properties of complementizer agreement suggest that this form of multiple agreement is established by a post-syntactic operation that copies agreement features (valued in the syntax) from T to C under structural adjacency. This suggests that T and C, which are part of different Spell-Out domains, must be part of a single domain in the phonological component of grammar. According to the central proposal put forward in this section, this can be ensured under the assumption that the phonological component maps the cyclic output of narrow syntax to phonological domains which are slightly larger than a single Spell-Out domain. More precisely, I have suggested that a phonological domain consists of a Spell-Out domain Σ_n and the right edge of a subsequent Spell-Out domain Σ_{n+1} . In this way, the phonological component can be taken to restore phasal units which have been disrupted by the application of the operation Transfer (affecting TP and VP, but not the phase head and its edge), which warrants a strict isomorphism/parallelism between the cycles of syntactic and post-syntactic computation. In addition, I have discussed evidence suggesting that the individual phonological domains assembled in the phonological component do not represent separate discrete units. Rather, phonological domains overlap, that is, the right edge of a Spell-Out domain Σ_n forms a phonological domain together with a previous Spell-Out Σ_{n-1} , but is also part of the phonological domain created from Σ_n . This extends the scope of phonological operations, giving the right results for processes such as prosodic phrasing, affix-hopping etc.

In section 4, I have argued that the assumption of overlapping phonological domains plays an important role in the linearization of syntactic structures. I have proposed that the mapping from hierarchic structures to linear orders of phonological exponents is a function of the process of Vocabulary Insertion, which applies cyclically to the phonological domains constructed from the output of the syntactic derivation, providing syntactic terminal nodes with phonological realizations. Following Epstein et al. (1998) and Richards (2004), (2007) I assume that the Head Parameter is phonological in nature, which can be formally modeled under the assumption that the insertion procedure ignores a subset of symmetric c-command relations created in the syntax, dependent on lexical properties of individual Vocabulary items. Furthermore, I have suggested that the Head Parameter is confined to exponents of functional categories. Under this approach, the same principles govern the linearization of head-adjunction structures as well as phrasal entities. In the latter, however, the existence of an intermediate projection

level (X') creates an asymmetry between X^0 -structures and XP -structures to the effect that specifiers asymmetrically c-command and precede insertion sites on the main path of embedding (i.e., within X'). I then addressed the question of how individual phonological domains can be linearized relative to each other, arguing that the linearization of successive phonological domains makes use of overlapping phonological domains as defined in section 3.3. Accordingly, material at the overlap between two successive phonological domains (i.e., material which is part of both the left edge of $PF\text{-domain}_n$ and the right edge of $PF\text{-domain}_{n+1}$) provides a connection between these domains since the relevant exponents are linearized both (a) relative to material in the higher domain and (b) relative to material in the lower domain. The separate phonological domains are then combined by a process called *Edge Replacement* which substitutes the right edge of $PF\text{-domain}_{n+1}$ with the string of exponents realizing $PF\text{-domain}_n$. I have argued that this replacement operation is subject to a non-tampering condition requiring *Edge Replacement* to preserve adjacency relations established previously between exponents of the higher $PF\text{-domain}_n$ and the exponents at the overlap. This restriction rules out certain configurations such as the combination of a head-final TP with a head-initial vP . In addition, I have proposed that a major parametrical difference between OV and VO grammars concerns the availability of raising the (verbal) root to the category-defining v -head, which obligatorily takes place in VO grammars, but is supplanted by post-syntactic v -lowering in OV grammars (*The Root Raising Parameter*).

Section 5 has illustrated the workings of this model of linearization in some more detail, focusing on the distribution of finite complement clauses in German, which appear in a postverbal position despite the fact that all other complements are located in preverbal position. I have argued that the peculiar distribution of finite complement clauses can be attributed to the *No-Tampering Condition on Edge Replacement*. More specifically, an embedded complement clause introduced by a clause-initial complementizer cannot be realized to the left of the matrix verb in an OV grammar since *Edge Replacement* destroys the ordering relation between the complementizer and the matrix verb. This analysis also rules out a linear ordering where a complement clause introduced by a clause-final complementizer appears to the right of its selecting verbal head. At the same time, it follows from my analysis that clausal complements can occupy a preverbal position if the complementizer occupies a clause-final position (as in Japanese or Bengali), which derives the correlation between clausal position and complementizer position observed in work by e.g. Bayer (1996, 1999). In addition, I have discussed a set of apparently

problematic cases, focusing on instances of preverbal placement of complement clauses in the so-called *Mittelfeld* 'midfield', the fact that postverbal complement clauses may not appear inside the verbal complex, and the placement of subject clauses.

Section 6 has focused on typological and diachronic implications of the theory of linearization developed in this work. I have demonstrated that the *No-Tampering Condition on Edge Replacement* rules out certain combinations of parametric choices that cannot be linearized by the workings of Vocabulary Insertion and thus gives rise to a typology of possible and impossible grammars. An overview of the ruled-out configurations is given in (199) for cross-clausal contexts and in (200) for clause-internal contexts:

- (199) a. *head-initial v/V embedding a head-final CP
- b. *head-final v/V embedding a head-initial CP
- c. *head-final v/V embedding CP with filled specifier and $C = \emptyset$

- (200) a. *head-initial Aux in T embedding a head-final vP (with verb in v)
- a.' *head-initial Aux in $v2$ embedding a head-final $v1P$ (with verb in $v1$)
- b. *head-final Aux in T embedding a head-initial vP (with verb in v)
- b.' *head-final Aux in $v2$ embedding a head-initial $v1P$ (with verb in $v1$)
- c. *head-final T embedding vP with filled specifier and $v = \emptyset$
- c.' *head-final $v2$ embedding $v1P$ with filled specifier and $v = \emptyset$

Additional restrictions on possible grammars are imposed by a condition that bans conflicting values of the Head Parameter in a head complex (e.g., ruling out movement of head-initial v to head-final T under certain circumstances). I have then argued that an approach in terms of *Edge Replacement* is empirically more adequate than analyses of the absence of *VO-Aux orders that are based on Holmberg's (2000) *Final-Over-Final Constraint* (FOFC) (Biberauer et al. 2007, 2008). In particular, we have seen that on the one hand, the FOFC-based approach requires additional assumptions to rule in attested word order options (such as final complementizers in T-initial languages), while on the other hand it overgenerates, predicting that certain non-existing word orders should in principle be available (e.g., clause-medial complementizers, or discontinuous complement clauses in SOV languages with initial complementizers). I have then explored the diachronic implications of an approach in terms of *Edge Replacement* in some more detail, focusing on the analysis

of word order variation in OE and the transition from OV to VO in the ME period. I have claimed that word order variation in embedded clauses of OE is to be analyzed vP-internally in terms of competing grammars that differ with respect to the setting of the Head Parameter for the exponent of v2 (the insertion site for auxiliaries and modals, assuming a 'stacked' v2P-v1P structure for complex verb forms) and the *Root Raising Parameter*. Under these assumptions, we can rule out the non-attested ordering *S-V-O-Aux as a violation of the No-Tampering Condition of *Edge Replacement* (resulting from the presence of a head-initial exponent at the left edge of v1P in combination with a head-final setting of the Head Parameter linked to the head of v2P). Finally, I have taken a brief look at changes that led to the development of a strict VO grammar in the ME period. I have argued that the transition from OV to VO was shaped by 'hard-wired' properties of the mapping from syntax to morphology/phonology (in addition to factors that gradually changed the make-up of the triggering experience and obscured the evidence for a OV grammar such as extraposition, fronting of verbal particles, and loss of case distinctions). In particular, I have suggested that in ME, the independently motivated rise of (i) subject movement to SpecTP and (ii) v-to-T movement blocked the acquisition of an underlying OV option in the innovating grammars, since this combination of parametric choices gives rise to syntactic structures that cannot be linearized by workings of Vocabulary Insertion/LIN.

Of course, the proposals put forward in this chapter represent only a first step towards developing a theory of how language change is shaped and determined by (hard-wired) properties of the mapping from syntax to PF. While it seems likely that these properties ease the burden on the learner via reducing the number of hypotheses about underlying structures that must be entertained, more has to be said about the range of possible misanalyses and therefore possible deviations from the target grammar. Furthermore, it is fairly clear that the learner's conclusions about word order properties of the target grammar can also be informed by other pieces of information such as (semantic) restrictions on possible argument and event structures, information-structure, stress patterns, or general learning strategies that help the learner to cope with input data that are difficult to analyze. Some of these aspects are investigated in some more detail in the next chapter that deals with the historical development of the V2 property in Germanic.

Chapter 3: V2 in Early Germanic

1 Introduction

In the previous chapter, I have argued that the linearization of syntactic structures is accomplished in the phonological component to satisfy requirements imposed by the serial interface to SM. However, it is fairly clear that the linear ordering of terminal elements is not merely an epiphenomenal property resulting from the necessity that the output of the grammar interact with other cognitive systems. Rather, word order is utilized to encode major structural properties of language that are established in the syntax and determine the interpretation of utterances, in particular grammatical functions and surface-related meaning properties (typically yielded by internal Merge, Chomsky 2002, 2004, 2005) such as scope of operator-like elements, information-structural distinctions (e.g., topic-comment, theme-rheme, and focus-presupposition), and other discourse-related properties such as specificity or anaphoricity.

The central role of word order becomes particularly clear if we adopt the perspective of the language learner, who does not have direct access to the hierarchical structures generated in the syntax, but must reconstruct the relevant structural properties of language (and the relevant featural content of syntactic categories) via inspecting and processing linear strings of words. This non-trivial task is greatly simplified if the learner can apply his hard-wired knowledge of LIN (see chapter 2) during language processing to ‘undo’ the workings of the linearization procedure, reconverting linear ordering into hierarchical structures. Still, there are many factors that may impede detection of properties of the target grammar, ultimately leading to grammar change. For example, it is fairly obvious that the

relation between linear orderings and hierarchical structure is non-unique, that is, a given string of words may be compatible with different underlying structures. In order to decide whether a given order is the result of displacement, or represents the base order/structure of elements, the learner has to take into account additional information. One such type of information involves the (possible) surface interpretation of the utterance.¹ More precisely, the child has to decide whether the pattern in question is systematically linked to certain surface-related meaning properties (scope, information structure etc.), which are typically implemented by displacement/internal Merge (and are often linked to the edge of constructions, e.g., the clausal left periphery). In other words, the proper acquisition of displacement properties requires that the learner can detect the function of the relevant word order pattern in the target grammar. If a particular function is blurred due to independent factors ('noise in the channel', overuse etc., see chapter 1), this may give rise to syntactic change, possible outcomes being: (i) loss of the relevant pattern; (ii) grammaticalization into a 'fossilized' syntactic pattern that fails to be associated with a particular surface meaning (and is triggered by purely morphosyntactic (EPP) features), cf. Simpson (2004); (iii) association of the relevant pattern with a different surface-related meaning/function; or (iv) reanalysis as the base order of elements.²

The present chapter discusses relevant word order changes from the history of the Germanic languages, focusing on the rise and loss of the V2 property in Gothic, Old and Middle English, and Old High German. I am going to show that the historical developments affecting the status of V2 exemplify options (i) and (ii).

It is a well-known fact that earlier stages of Germanic exhibit V2 patterns in non-embedded contexts similar to the Modern Germanic languages (with the notable

¹ Another viable source involves morphological properties, in particular phonological exponents of functional categories that express surface-related meaning properties or syntactic functions, that is, inflections (case, agreement etc.), complementizers, and markers of information-structural distinctions such as topic and focus markers (as in Somali or Japanese). See e.g. Bobaljik (2003) and Haeberli (2004) on the relative significance of syntactic and morphological cues during language acquisition (for similar considerations cf. Anderson 1980). See below for a brief discussion of the claim that the rise of generalized V-to-C movement and V2 in Germanic was connected to the loss of a system of C-related particles which can still be observed in Gothic (Roberts 1996, Ferraresi 1997).

² Note that in recent approaches to word order change based on Kayne's LCA/UBH, possibility (iv) is usually treated as an instance of (ii), cf. Hróarsdóttir (2000) on the rise of VO in Icelandic and Hinterhölzl (2004) on the OV-VO alternation in the history of English and German.

exception of Modern English), compare the following examples from Old English, Old High German, and Old Saxon:

(1) *Old English*

Uneaðe **mæg** *mon* to geleafsuman gesecgan ...
 Hardly may man to faithful speak
 'Hardly may man speak to the faithful ...'
 (Orosius, *Or_3*:9.70.16.1292)

(2) *Old High German*

Dhinera uuomba uwaxsmin **setzu** *ih* ubar min hohsetli
 your womb's fruit place I upon my throne
 'I place your womb's fruit upon my throne.'
 (Isidor, 611; Robinson 1997:9)

(3) *Old Saxon*

Hwat **quiðis** *thu* umbi gōdon?
 what say you about a-good-one
 (Rauch 1992:19)

At a closer look, however, we can identify a variety of types of V2 in the different branches of (Early) Germanic. I am going to argue that the historical facts provide evidence for at least four different underlying structural configurations which may give rise to V2-patterns on the syntactic surface:

(4) a. 'operator V2': V-to-C movement in the context of fronted operators such as *wh*-phrases, negation etc.

b. 'TP-V2': V2 orders resulting from a spec-head configuration in the TP

c. 'pseudo V2': superficial V2 patterns that does not involve a spec-head configuration between fronted element and finite verb

d. 'generalized V2': spec-head configuration of fronted XP and finite verb similar to the Modern Germanic V2 languages

Evidence from Gothic and Old English (OE) suggests that the historical core of the V2 phenomenon reduces to V-to-C movement that is triggered in operator contexts

(cf. Kiparsky 1995, Eythórsson 1995, 1996, Fuß 2003). Therefore, the historical system shares basic properties with Modern English with respect to the limited range of V2/V-to-C movement. As is well-known, Modern English exhibits subject-verb inversion (and V2 effects) with questions, neg-preposing, and in imperatives (and a couple of other contexts; for fuller treatment cf. e.g. Green 1980, McCawley 1998):³

- (5) a. What has Floyd seen?
- b. *What Floyd has seen?
- c. Has Floyd read the book?
- d. *Floyd has read the book? (ok as an echo question)
- e. Never would I do that.
- f. *Never I would do that.
- g. Don't you try again!

In section 2, I discuss a set of apparent deviations from this generalization that can be observed in Gothic. The problematic cases involve orders where a clause-initial *wh*-phrase is not followed by the finite verb, but rather by pronouns or second position particles. Following Fuß (2003), I am going to argue that these exceptions can be attributed to the fact that the Gothic bible is a quite literal translation of a Greek source that often merely imitates the word order patterns of the original text. In addition, I will show that clauses introduced by the discourse connectives *þaruh* 'there' and *þanuh* 'then' constitute another context where systematic verb fronting and inversion take place.

Furthermore, we will see that the apparently more elaborate V2 properties of OE mostly derive from a pattern that does not involve a spec-head relation in the CP, but rather results from merely linear adjacency between a fronted topic and the finite verb which is located in T^0 (dubbed 'pseudo V2' in Fuß 2003). Therefore, V2 in OE reduces to a set of 'core cases' where this pattern is derived by a spec-head configuration between a fronted operator and the finite verb (cf. e.g. Pintzuk 1999).

³ This trait of English syntax is usually referred to as 'residual V2', suggesting that the V2-properties of Modern English represent the residue of a former more elaborate V2-system, which was similar to the 'full V2' grammars of the other present-day Germanic languages. However, research into the historical syntax of English has shown that this term is actually a misnomer; it can be shown that English has never been a V2 language in the sense of the modern Germanic V2 languages (cf. e.g. Kroch and Taylor 1997, Haeberli 1999, Fischer et al. 2000).

Again, this is reminiscent of the limited V2 properties of Modern English, revealing a historical continuity from the oldest records to Modern English (cf. Eyþórsson 1995, 1996, Kiparsky 1995). Similar to Gothic, subject-verb inversion is also obligatorily triggered in the context of a closed class of fronted temporal adverbs. In this work, I focus on the behavior of *þa/þonne* 'then' (cf. Trips and Fuß 2008), arguing that the relevant V2 patterns result from a structural configuration in which the finite verb moves to T while SpecTP is occupied by *þa/þonne* or pronouns.

Section 4 discusses the loss of surface V2 patterns in the Middle English (ME) period. My basic claim is that the changes affecting V2 were linked to the general loss of discourse-configurational properties, which gave rise to a grammar in which word order signals grammatical functions (instead of information-structural distinctions). More specifically, I am going to argue that the loss of pseudo-V2 patterns is to be attributed to the rise of an EPP feature in T.

In section 5, these findings are contrasted with the already much more systematic V2 characteristics of Old High German (OHG). I show that early OHG already exhibits patterns similar to 'generalized V2' in the Modern Germanic languages. Based on an analysis of the left periphery of OHG in terms of a non-split single CP structure, I am going to argue that the historical emergence of generalized V2 can be attributed to a combination of changes that first led to generalized V-to-C movement, followed by the development of semantically/pragmatically vacuous XP-fronting. I show that the latter change proceeded in two steps, involving the rise of an EPP-feature in C and the loss of multiple specifiers in the C-domain (due to the development of C-related expletives) in the course of the OHG period.

Section 6 summarizes the findings of this chapter.

2 Verb fronting and 'operator V2' in Gothic

Traditionally, the V2-phenomenon is considered to be an innovation which is associated with the Germanic branch of the Indo-European language family (the older Indo-European languages such as Sanskrit, Old Greek or Latin do not exhibit systematic V2-patterns), cf. Brugmann and Delbrück (1900), Kiparsky (1995). Except for a few runic inscriptions, the earliest Germanic data handed down to us are of Gothic origin, dating from the fourth century. The bulk of the Gothic data comes from a bible translation by Bishop Wulfila (311-383 AD). The Gothic bible is a fairly literal (i.e. word-by-word) translation of a Greek text. As a result, quite a number of

syntactic properties (word order, in particular) of the Gothic bible can be shown to be influenced by the Greek source text. However, Eythórsson (1995), (1996) and Ferraresi (1997) convincingly argue that there are systematic deviations from the word order found in the original that can be taken to reveal properties of the syntax of Gothic. More specifically, it can be shown that Gothic exhibits systematic verb movement in contexts similar to those that trigger V-to-C movement in present day English (see above). Before we turn to the relevant data, let us first review some basic properties of Gothic. Gothic exhibits basic OV order in both main and embedded clauses (cf. e.g. Eythórsson 1995, Ferraresi 1997):

(6) a. ik in watin izwis **daupja**
 I in water you-PL baptise
 (Matt. 3.11; Roberts 1996: 161)

 b. þaþroh þiudangardi gudis **wailamerjada**
 since-that kingdom of-God is-preached
 jah huzuh in izai **naupjada**
 and everyone into it presses
 (Luk. 16.16; Ferraresi 1997: 277)

However, to rule out influence of the word order of the Greek original, only Gothic examples that show OV order independently of the word order of the Greek text constitute clear evidence in favor of a basic OV syntax. Relevant evidence in favor of an OV base comes from cases where a single Greek verb is represented by a verb together with a (non-pronominal) complement in the Gothic translation. Here, the complement usually precedes the verb (cf. Eythórsson 1995, 1996):

(7) a. dwala **gatawida**
 foolish made
 Gk. emōranen
 'made foolish'
 (1Cor. 1,20; Eythórsson 1996: 109)

b. lofam **slohun**
 palm-PL.DAT smote-3PL
 Gk. errapisan
 '(they) smote (him) with the palm of (their) hands.'
 (Mt. 26,67; Eythórsson 1995: 20)

Starting out from a basic OV order, there are a number of contexts where the finite verb systematically undergoes movement into the left clausal periphery (cf. Longobardi 1994; Eythórsson 1995, 1996; Ferraresi 1997). Interestingly, these instances of regular V-to-C movement seem to be limited to contexts that license the same movement operation in Modern English: *wh*-questions, neg-preposing, and imperatives. Note that in examples (8)-(10), the position of the finite verb in the Gothic sentence has no model in the Greek original and can therefore be taken to reflect genuine properties of the syntax of Gothic (cf. Eythórsson 1995: 22ff.). In (8), the position of the finite verb in the Gothic translation differs from the position occupied by the corresponding verb in the Greek sentence: whereas the finite verb appears in clause final position in the Greek sentence, it undergoes inversion with the subject in the corresponding Gothic *wh*-question. This is reminiscent of the V2 order that characterizes *wh*-questions of present day English, cf. (5a). In the Gothic examples in (9) and (10), a combination of verb + complement renders a single Greek verb. However, in contrast to the examples in (7), the verb precedes its complement in imperatives and examples with clause-initial negation. Again, this suggests that the finite verb has undergone a movement operation that targets a functional head in the left clausal periphery, similar to the relevant cases in Modern English, cf. (5e) and (5g) above:

(8) **h**ua **skuli** þata barn wairþan?
 what shall that child become
 Gk. **ti** ara paidion touto **estai**
 'what shall that child become?'
 (Lk. 1,66; Eythórsson 1996:110)

(9) ni **nimip** arbi
 NEG takes inheritance
 Gk. mē klēronomēsēi
 'shall not be heir'
 (Gal. 4,30; Eythórsson 1996:110)

(10) **wairp** hrains
 become clean
 Gk. katharisthēti
 'become clean!'
 (Mt. 8,3, Mk. 1,42, Lk. 5,13; Eythórsson 1996:110)

Thus, at first glance, it seems that Gothic already exhibited a form of V-to-C movement that shows some similarities to the 'residual' V2 phenomenon of present-day English (at least with respect to the syntactic contexts where this operation is triggered). However, a closer look reveals that at least in *wh*-questions, the facts are actually less clear. The next section discusses a number of cases that conflict with our preliminary conclusion that Gothic exhibited consistent verb fronting to C in operator contexts.

2.1 Apparent counterexamples

This section shows that apart from the finite verb, there are also a couple of other elements – in particular, phonologically 'light' elements such as pronouns and modal particles – that may occupy the second position in *wh*-questions. Following Fuß (2003), I argue that the problematic cases can be attributed to Greek influence and therefore do not represent real counter-examples for the claim that Gothic exhibits systematic V2 effects in *wh*-questions.

2.1.1 Pronoun placement in *wh*-questions

The following examples show that in *wh*-questions, subject pronouns may precede (11) or follow (12) the finite verb in Gothic, conflicting with the claim that *wh*-questions exhibit consistent V2 order:

(11) a. duhe **jus** mitoþ ubila in hairtam izwaraim?
 why you-PL think evil in hearts your
 Gk. hinati **humeis** enthumeisthe ponēra en tais kardiais humōn
 why you-PL think evil in the heart your
 'Why do you think evil in your hearts?'
 (Mt. 9,4; Ferraresi 1997:53)

b. huaiwa **þu** qibis þatei frijai wairþiþ?
 how you-SG say that free become
 Gk. pōs **su** legeis hoti eleutheroi genēsesthe
 how you-SG say that free become
 'How do you say you shall become free?'
 (Jo. 8,33; Ferraresi 1997:53)

(12) a. hua þanamais þaúrbum **weis** weitwode?
 what further need we witness
 Gk. ti eti chreian echomen marturōn
 what further need have-1PL witness
 'What do we need any further witnesses?'
 (Mk. 14,63; Ferraresi 1997:55)

b. hua nuk-kant **þu**, quino?
 what now-know you wife
 Gk. ti gar oidas, gunai
 what therefore know-2SG wife
 'What do you know, wife?'
 (I Cor. 7,16; Ferraresi 1997:55)

As pointed out in Fuß (2003), the word order of the Gothic example in (11) is identical to the order of the Greek original. This suggests that the apparent counter-examples do not tell us much about the syntax of Gothic since they are literal word-by-word translations of the source text. However, a closer look at (12) reveals that the Gothic translations correspond to Greek clauses lacking an overt subject. According to Ferraresi (1997: 58) there are 10 examples where a null subject in the Greek original

is translated by an overt pronoun in the corresponding Gothic main clause.⁴ In 8 of these cases, the newly inserted pronoun is located in clause-initial position. The remaining two cases are the *wh*-questions in (12), in which the added pronoun resides in postverbal position. Therefore, from the evidence available to us we can conclude that examples like (12) can be taken to indicate some real word order properties of Gothic, namely that pronouns originally had to follow the finite verb in *wh*-questions. In other words, the alleged counter-examples actually support the claim made above, namely that *wh*-questions already exhibit systematic V2 in Gothic.

2.1.2 Second position particles

It is a well-known fact that Gothic exhibits an elaborate system of left-peripheral particles, for example the interrogative particle *-u*, the coordinating particle *-uh*, and modal (or emphatic) particles such as *þan*, *nu*, and *auk* (and combinations of these). Traditionally, the placement of these particles is taken to be governed by Wackernagel's law forcing unstressed elements into clausal second-position (Wackernagel 1892). Since it can be shown that most of them are linked to sentential properties such as clause type, focus and the main/embedded distinction, recent generative work on Gothic generally assumes that these particles are generated in the C domain. If the particle in question is a clitic, it attaches to the right of lexical material that has moved into the left clausal periphery (cf. Eythórssón 1995, 1996, and Ferraresi 1997 for comprehensive discussion). In *wh*-questions, the placement of the second-position particles may give rise to violations of V2 in which particles intervene between the finite verb and the fronted *wh*-phrase:

⁴ Furthermore, there are 23 cases where an overt pronoun translates a null subject in an embedded clause. In all these examples, the pronoun directly follows the complementizer. Again, this resembles the placement properties of pronouns in the modern Germanic V2 languages:

(i) witum ei þu **kant** alla
know that you know all

Gk. oidamen hoti oidas panta

'We know that you know all the things.'

(Jo. 16,30; Ferraresi 1997:59)

(13) hjan-uh **þan** þuk sehūm gast jah ga-laþodedum?
 when-PRT PRT you we-saw stranger and PERF-we-invited
 Gk. **pote de** se eidomen xenon kai sunēgagomen
 'And when did we see you as a stranger and invited you?'
 (Mt. 25,38)

(14) a. hua **nu** taujai im frauja þis weinagardis?
 what PRT do them owner of-the vineyard
 Gk. **ti oun** poiēsei autois ho kurios tou ampelōnos
 'What then shall the owner of the vineyard do to them?'
 (Lk. 20,15)

b. hua **auk** boteiþ mannan, jabai gageigaiþ þana fairhū allana
 what PRT profit man if gain-3SG the-DEM world whole
 jah gasleiþeiþ sik saiwalai seinai
 and injure REFL soul his
 Gk. **ti gar** ὅφελει ἄνθροπον κερδεῖσαι τὸν κόσμον καὶ ζημιόθεναι τὸν
 πσυχὴν αὐτοῦ
 'For what does it profit a man, if he gains the whole world, and loses (lit.
 injures) his own soul?'
 (Mk. 8,36)

It is fairly undisputed that this class of second-position particles constitutes a genuine trait of the grammar of Gothic. However, a closer look at the examples in (13) and (14) reveals that the second position particles found in the Gothic examples correspond directly to second position particles in the Greek text (*de* in (13a, b), *oun* in (14a), *gar* in (14b)). The following sections examine the use of second position particles in *wh*-questions in some more detail, arguing that those instances of particle placement that appear to violate the V2 constraint in *wh*-questions can be attributed to word order properties of the Greek original.⁵

⁵ The following discussion does not include the interrogative particle *-u*, which marks main and embedded yes/no questions, but is not used in *wh*-interrogatives (cf. Ferraresi 1997 for details).

2.1.2.1 The coordinating particle *-uh*

The clitic particle *-uh* is primarily used to conjoin main clauses.⁶ In order to conjoin other kinds of phrases or embedded clauses, the tonic conjunction *jah* must be used.⁷ The fact that the particle *-uh* is in complementary distribution with subordinating complementizers such as *ei* 'that' (cf. Eythórsson 1995: 54) suggests that it is sensitive to properties usually associated with C (i.e., the main/embedded distinction). Therefore it is plausible to assume that *-uh* is generated under C (or some functional head of a Split-C system). There are many examples such as (15) in which *-uh* attaches to a fronted finite verb, but as shown in (16) it can also cliticize onto other elements in the left periphery of the clause (cf. Klein and Condon 1993, Eythórsson 1995, 1996; Ferraresi 1997 for details). In (16a), *-uh* cliticizes to a topicalized noun; in (16b), *-uh* does not attach to the right edge of a fronted phrase (here a PP), but rather to its head, splitting the constituent in two parts (similar to the Wackernagel clitics found in Serbo-Croatian, cf. Zec and Inkelaš 1990, Schütze 1994, Ćavar and Wilder 1999).⁸

⁶ Eythórsson (1995) claims that this is the only function of *-uh*. However, it can be shown that the distributional properties of *-uh* are actually more complex (cf. Klein and Condon 1993, Klein 1994, Ferraresi 1997). Apart from its coordinating function, *-uh* apparently also functions as a discourse particle that serves to mark anaphoric relations across clauses. Furthermore, it seems that *-uh* is not always in complementary distribution with the tonic conjunction *jah*. In the following examples, *-uh* appears in the first of two clauses conjoined by *jah*:

- (i) a. **uz-uh-iddja** fram attin jah atiddja in þana fairhū
forth-PRT-came from father and came into the world
'I came forth from the Father and came into the world.'
(Jo. 16,28; Ferraresi 1997:108)
- b. **ip** is **wiss-uh** mitonins ize jah qaþ du þamma mann...
but he knew-PRT thoughts their and said to the man
'But he knew their thoughts and said to the man...'
(Lk. 6,8; Ferraresi 1997:108)

Note that in (ib), *-uh* is placed after the subject pronoun and the finite verb (i.e., it occurs in third position). Eythórsson (1995) shows that this ordering is always observed when a definite subject is topicalized (15 examples).

⁷ Historically, *jah* derives from *ja+uh*, cf. Eythórsson (1995:54, fn. 12).

⁸ This perhaps suggests an analysis where the ultimate position of these particles is determined by post-syntactic processes that may shift a clitic element minimally to the left or right in order to satisfy PF-conditions. One such operation is *Prosodic Inversion* (Halpern 1992), which affects prosodic categories to satisfy PF-conditions such as second position effects. Prosodic Inversion is similar in nature to *Local Dislocation* (Embick and Noyer 2001; see chapter 2 above), although it slightly differs

(15) jah usstigun in skip, iddjedun-**uh** ufar marein
 and they-entered in ship went-PRT over sea
 'and they entered into a ship and crossed the sea'
 (Jo. 6,66; Eythórsson 1996: 120)

(16) a. þuht-**up** þan qipa⁹
 conscience-PRT PRT I-say
 'I speak of conscience'
 (1Cor. 10,29; Eythórsson 1996:119)

b. [_{PP} uz-**uh** þamma mela] managai galipun sponje is
 from-PRT that time many went of-disciples his
 'and from that time many of his disciples went'
 (Jo. 6,66; Eythórsson 1996:120)

According to Ferraresi (1997: 107f.), there are 50 examples where the use of *-uh* has no model in the Greek original, contrasting with 18 cases where *-uh* is used to translate Greek second position particles such as *de*. In addition, there are 122 cases where *-uh* co-occurs with the modal particle *þan*. Again, this combination is often used to translate Greek second position particles such as *de*; in only 7 of these cases the use of *-uh+þan* has no model in the Greek text. Still, these numbers clearly show that the second position effects with *-uh* represent a native trait of Gothic syntax.

Coming back to our initial question concerning the behavior of particles in *wh*-question, it appears that there are only three examples where *-uh* attaches to a clause-initial *wh*-word, resulting in a violation of the V2 constraint (Eythórsson 1995: 102):

from Local Dislocation in that the affected elements are prosodic categories rather than morphological ones. Prosodic Inversion typically does not respect phrase boundaries. Schütze (1994) discusses the well-known case of second position clitics in Serbo-Croatian the placement of which may lead to a split of other phrases, similar to the Gothic examples discussed above. However, see Ćavar and Wilder (1999) and Fanselow and Ćavar (2002) for a critique of phonological accounts of the Serbo-Croatian facts and an alternative analysis in terms of syntactic movement. See Eythórsson (1995), (1996) for an analysis of the splits found in Gothic in terms of syntactic head movement (i.e., incorporation, Baker 1988).

⁹ Here, *uh* changes to *up* due to assimilation rules.

(17) a. **huan-uh þan þuk** sehūm gast jah ga-labodedum?
 when-PRT PRT you saw-1PL stranger and PERF-we-invited
 Gk. **pote de se eidomen xenon kai sunēgagomen**
 'And when did we see you as a stranger and invited you?'
 (Mt. 25,38)

b. **huan-uh þan þuk** sehūm siukana aibþau in karkarai jah
 when-PRT PRT you saw-1PL sick or in prison and
 atiddjedum du þus?
 came-1PL to you
 Gk. **pote de se eidomen asthenounta ē en phulakē kai ēlthomen**
 'And when did we see you sick, or in prison, and came to you?'
 (Mt. 25,39)

c. **hua-uh þan habais þatei ni namt?**
 what-PRT PRT have-2SG that not received
 Gk. **ti de ekheis ho ouk elabes**
 'What do you have that you did not receive?'
 (1Cor. 4,7; Eyþórsson 1995:102)

In (17a-c), *-uh* is directly followed by the modal particle *þan* 'then'. The finite verb then follows either *þan*, as in (17c), or the pronoun *þuk*, as in (17a, b). Note that these patterns correspond exactly to the word orders found in the Greek original. It appears that the combination of *-uh* + *þan* serves to translate the adversative particle *de*, which occupies the second position in the Greek original. In other words, the examples in (17) represent close-to-literal, word-for-word translations of the Greek source text. Accordingly, the distribution of the particle *uh* (+ *þan*) actually does not provide a counterexample to the claim that Gothic *wh*-questions exhibit consistent V2 order.

2.1.2.2 The modal particles *þan*, *nu* and *auk*

The set of second-position modal particles includes: *þan* 'then', *nu* 'now, thus' and *auk* 'because, also, thus'.¹⁰ The placement of these particles is heavily influenced by

¹⁰ According to Ferraresi (1997: 112ff.), *þan* and *nu* can be used either as adverbials or as modal particles. In the latter use, they are confined to second position where they can be preceded by any

properties of the Greek source text. There are 226 examples where *þan* translates Greek second position particles, but only 12 cases where its insertion has no model in the Greek text (Ferraresi 1997: 115f.). Of the 96 examples where *nu* occupies clausal second-position, it translates Greek second position particles 92 times, that is, there are only 4 examples where *nu* is used as an emphatic second position particle without a model in the Greek text (cf. Ferraresi 1997: 118). Moreover, there are no examples where *auk* is inserted without a corresponding Greek particle (Ferraresi 1997: 122).

A search conducted in the online version of the Streitberg edition of the Gothic bible (made available by the TITUS project at the University of Frankfurt; <http://titus.uni-frankfurt.de/>) produced the following numbers for violations of V2 where a clause-initial *wh*-phrase is immediately followed by a modal particle (cf. Fuß 2003).¹¹

other element. No such positional restrictions can be observed if *þan* and *nu* function as adverbials. In contrast, *auk* corresponds exclusively to Greek emphatic particles which appear in second position.

¹¹ No relevant examples were found with the *wh*-words *his* (MASC/NEUT.GEN), *hamma* (MASC/NEUT.DAT), *hizai* (FEM.DAT), *huna* (MASC.ACC), *hurjis* ('which one'), *bihê* ('where of'), *hileiks* ('what ... like'), *helauda* ('how big').

	þan	-uh + þan	nu	auk	þau ¹²	Total
huas (masc.nom)	(1) ¹³	—	—	2	—	3
hva (neut.nom/acc)	—	1	7	5	—	13
huô (fem.nom/acc)	—	—	1 ¹⁴	—	—	1
hê (neut.instr.)	—	—	2	—	—	2
huan ('when')	—	2	—	—	—	2
huaiwa ('how')	—	—	4	—	—	4
huapar ('which of two')	—	—	1	—	—	1
duhuê ('why')	—	—	—	1	2	3

Table 1: *Wh*-words immediately followed by a modal particle in the Gothic NT

Importantly for our purposes, it can be shown that in all instances where the insertion of a modal particle leads to an apparent violation of V2, the position of the Gothic particle imitates the position of a corresponding element in the Greek text. This is shown in Table 2:

¹² According to Wright (1924), *þau* is best translated as 'then, in that case' in this environment.

¹³ There is apparently only a single example where the modal particle/adverb *þan* 'then' (without *-uh*) immediately follows a *wh*-phrase. Notice, however, that (apart from the fact that *þan* is used to translate the Greek second position element *de*) this is actually not a relevant example, since the particle splits up a complex *wh*-phrase (i.e., it appears directly after the initial word), which is in turn directly followed by the finite verb.

(i) [huas **þan** izwara] skalk aigands arjandan aipþau haldandan,
 who PRT of-you.PL servant having ploughing or feeding (cattle)
 saei atgaggandin af haiþai qipai: suns hindarleip anuhkumbei?
 who coming back from field would-say at once come over sit down to eat
 'But which of you, having a servant plowing or feeding cattle, who came in from the field, will say to him: go and sit down to eat at once.'
 (Lk. 17,7; TITUS)

¹⁴ The relevant example involves an embedded clause introduced by *jabai* 'if' (Phil. 2,1).

	Greek 2 nd position particles					no corresponding Greek particle	Tot.
	de	gar	oun	nun	kai		
þan	1	–	–	–		–	1
-uh þan	3	–	–	–		–	3
nu	2	–	12	1		–	15
auk	–	8	–	–		–	8
þau	–	–	–	–	2	–	2

Table 2: Gothic 2nd position particles in *wh*-questions and Greek particles¹⁵

This quantitative study shows there are actually no examples where a second-position particle leading to a violation of V2 in a *wh*-question has no model in the Greek source. Therefore, we can conclude that the placement of Gothic second position particles does not represent a counterexample for the claim that we find systematic V2 in Gothic *wh*-questions. In all problematic cases, the Gothic particle imitates the use and position of a corresponding second position element of the Greek original.

2.2 Further contexts that trigger V2

There is at least one other context that favors V2 order (and inversion) in Gothic. This context involves the particles *þanuh* 'then+uh' and *þaruh* 'there+uh', which Klein (1994: 262) characterizes as "discourse-continuative foregrounding markers, carrying forward the discourse along the time-line of the main story".¹⁶ Both elements can be decomposed into the Proto-Germanic demonstrative root *to- (> *þa-* in early Germanic), an adverbial suffix /-r/, /-n/ (probably signaling local and temporal deixis, respectively; cf. Lockwood 1968: 226, 228, Ramat 1981), and the anaphoric discourse particle *-uh*. As already mentioned above, *-uh* generally signals a sequential or resumptive meaning and is used for clausal conjunction (cf. Klein and Condon

¹⁵ The relevant bible passages (NT) are: Gk. *de* rendered by Gothic *þan*: Lk. 7,17; *de* rendered by *-uh þan*: Mt. 25,38; Mt. 25,39; 1Cor. 4,7; *de* rendered by *nu*: Mt. 11,16; Jo. 9,21; *gar* rendered by *auk*: Mk. 8,36; Rom. 11,34; 1Cor. 4,7; 1Cor. 10,29; 2Cor. 12,13; Phil. 1,18; 1Thess. 2,19; 1Thess. 3,9; *oun* rendered by *nu*: Jo. 9,19; Lk. 7,31; Lk. 7,42; Lk. 20,15; Mk. 15,12; Rom. 7,7; Rom. 9,14; Rom. 9,30; Rom. 10,14; 1Cor. 10,19; 1Cor. 4,26; Phil. 2,1; *nun* rendered by *nu*: Jo. 6,42; *kai* rendered by *þau*: 1Cor. 15,29 and 1Cor. 15,30.

¹⁶ Hirt (1929: 352f.) observes that clause-initial 'then' triggers inversion in other early Indo-European languages as well (including Sanskrit and Old Greek). This might be taken to indicate that the phenomenon in question is actually of greater antiquity.

1993, Klein 1994). Klein (1994: 255) counts 68 cases of *þaruh* and 68 cases of *þanuh* in the Gothic Bible, all of them occurring in sentence-initial position (my own search in the Streitberg edition of the Gothic bible produced somewhat higher numbers, see below).¹⁷ The following examples serve to illustrate both the V2 syntax and the discourse function of *þaruh* and *þanuh* (foregrounding, indication of narrative continuity, and in particular conversational turntaking with *verba dicendi*, often associated with a change of grammatical subject, cf. Klein 1994 for details):¹⁸

(18) Jah þairhgaggands gaumida mann blindamma us gabaurþai. *þaruh* **frehun**
ina siperjos is qipandans [...]
'And passing through, he saw a man blind from birth. *And/then* his disciples
asked him, saying [...]'
(John 9.1-2; Klein 1994: 256)

(19) a. *þaruh* **qaþ** Iesus du þaim twalibim [...]. *þanuh* **andhof** imma Seimon Paitrus
[...]
'Then Jesus said to the twelve [...]. *Then* Simon answered him [...]'
(John 6.67-68; Klein 1994: 260)

b. Qeþun du imma: jai, frauja! *þanuh* **attaitok** augam ize qipands [...]
'They said to him: "Yea, Lord!" *Then* he touched their eyes, saying [...]'
(Mt. 9.28-29; Klein 1994: 260)

In order to verify the status of V2 orders in connection with *þaruh* and *þanuh*, I conducted a search in the online version of the Streitberg edition of the Gothic Bible (New Testament only) made available by the TITUS project. In what follows, I report on the results of that research.

First of all, it can be shown that there some differences between *þaruh* and *þanuh* with respect to verb placement. While *þaruh* triggers V2 (and inversion) in the

¹⁷ According to Klein (1994: 255), there are 31 examples where *þaruh* translates the Greek particle *oun* (all in John), 23 cases where it translates *de* (16 of them in Luke), 4 cases where it translates *kai*, 2x *kai idou*, 1x *idou*, and 1x *ekei*. There are only six instances where the use of *þaruh* does not have a model in the Greek source. The following numbers are given for *þanuh*: 33x *oun* (32 in John), 20x *tote* (16 in Matthew; according to Thompson 1974 the Greek particle *tote* suggests a temporal progression in the relevant Bible passages), 10x *de*, 1x *kai*, and 1x *kai idou*. There are three cases where *þanuh* does not correspond to a Greek particle.

¹⁸ In the following examples, *þaruh/þanuh* are marked by italics, while the finite verb is set in boldface.

majority of clauses (44 of 68, or 64.7%), there are also a significant number of examples where a further element intervenes between *þaruh* and the finite verb (24 of 68, or 35.3%).¹⁹ In particular, there are 15 cases where a subject intervenes between *þaruh* and the finite verb: 10 examples with a subject pronoun, and 5 with a non-pronominal DP subject (of the latter, at least Mk. 10.24 can be discarded since its word order corresponds exactly to the Greek original).²⁰ Interestingly, all violations of V2 caused by a subject pronoun seem to reflect genuine properties of Gothic, since they cannot be attributed to Greek influence. This is illustrated by the following pair of examples:

(20) a. *þaruh is qab du im þatei [...]*
there he said unto them that
 Gk. *ho de eipen pros autos hoti kai [...]*
'And/there he said unto them that [...]'
 (Lk. 4.43)

b. *þaruh is ureisands gastob.*
there he rise-PRES.PART stood forth
 Gk. *kai anastas histē*
'And/there he arose and stood forth.'
 (Lk. 6.8)

In (20a), the Gothic translation exhibits a preverbal subject pronoun, while the subject pronoun follows the verb in the Greek source. In (20b), the Greek source does not contain an overt subject; the newly inserted pronoun shows up in preverbal position (directly after *þaruh*) in the Gothic translation. Thus we can conclude that V2 was the preferred option in main clauses introduced by *þaruh*. Still, there is also a significant number of examples which exhibit divergent word order properties. Next, we will take a closer look at *þanuh* 'then+uh', which is also a foregrounding discourse particle

¹⁹ In three cases (Mt. 9.18, Jo. 6.14, Jo. 6.19) *þaruh* introduces a non-finite clause. Jo. 6.24 exhibits a complex particle *þaruh þan* (+ V2).

²⁰ V2 violations with subject pronouns include: Jo. 6.20, Lk. 4.43, Lk. 5.34, Lk. 6.8, Lk. 6.10, Lk. 7.43, Lk. 8.46, Lk. 15.27, Lk. 15.29, Mk. 14.64. The relevant cases with full DP subjects are: Mt. 9.3, Jo. 13.37, Jo. 18.12, Lk. 9.42, Mk. 10.24. It is noteworthy that 8 out of 10 V2 violations with subject pronouns occur in Luke, which perhaps suggests that the Gothic Bible was not the work of a single translator (cf. Klein 1994: 261, fn. 14 for a related statement).

that signals narrative continuity, marking actions/events that move along the main time line of the narrative. A search conducted in the TITUS online version of the Gothic Bible produced the following results:

V2/inversion	Non-V2				Total
	with Greek model	no direct Greek model	embedded clause	<i>þanuh þan</i>	
52	21 ²¹	1	5	7	86

Table 3: V2 in clauses introduced by *þanuh* in the Gothic Bible

All in all, I found 86 examples that contained the particle *þanuh* (always in clause-initial position). Of the 86 examples, 12 can be discarded immediately, since they involve either the complex particle *þanuh þan* (which is rather to be analyzed as an instance of the backgrounding particle combination *-uh þan*, cf. Klein 1994), or are to be analyzed as embedded clauses). This leaves us with 74 examples, including 22 cases where the finite verb does not occupy second position. Of the latter, 21 can be attributed to Greek influence since the word order of the Gothic translation is more or less identical to the word order of the Greek source text. This is illustrated by the following two examples:

(21) a. þanuh þwairhs sa gardawaldands **qab** du skalka seinamma [...]
 then angry the master-of-house said to servant his
 Gk. tote orgistheis o oikodespotēs **eipen** tō doulō autou
 PRT angry the master-of-house said to servant his
 ‘Then the master of the house, being angry, said to his servant: [...]’
 (Lk. 14.21)

²¹ In Jo. 6.60, Jo. 7.43, Jo. 10.19, Jo. 11.33, Jo. 11.38, Jo. 11.45, and Lk. 8.54, the lower position of the verb corresponds to the Greek original, but the order of the clause-initial elements is reversed in the Gothic translation (subject+PRT > *þanuh* subject). The other examples are: Mt. 27.58, Jo. 6.27, Jo. 7.10, Jo. 7.11, Jo. 9.8, Jo. 12.19, Jo. 16.25, Jo. 19.13, Lk. 14.21, 1Cor. 14.25, 1Cor. 15.18, Phil. 2.23, Col. 3.4, and 1Thess. 5.3.

b. þanuh unweniggo ins **biqimip** fralusts
 then sudden them come-upon destruction
 Gk. tote aiphnidios autois **ephistatai** olethros
 PRT sudden them come-upon destruction
 'Then sudden destruction cometh upon them [...]'
 (1. Thess. 5.3)

This leaves us with a single clear example of a main clause introduced by *þanuh*, given in (22) below, in which the Gothic translation does not place the finite verb in second position. In (22), the translator rendered a single Greek element, *ōrgisthē*, by a combination of adjective and finite auxiliary, with the latter occurring in clause-final position:

(22) þanuh modags **warþ** jah ni wilda inngaggan, [...]
 then angry was and not wanted in-go
 Gk. **ōrgisthē** de kai ouk ēthelen eiselthein
 angry-was PRT PRT not wanted go-in
 'And he was angry, and would not go in.'
 (Lk. 15.28)

Summing up, we can conclude that the particle *þanuh* 'then+uh' regularly triggered V2 in Gothic. Of the relevant 74 cases, only a single example represents a violation of the V2 constraint. All other cases are either clear instances of V2, or deviations from V2 that can be attributed to influence of Greek word order properties.

2.3 Section summary

This section has shown that Gothic exhibits systematic V-to-C movement in a set of syntactic contexts that are quite similar to the contexts that trigger V-to-C movement in present-day English, giving rise to V2 order in *wh*-questions (cf. Eyþórsson 1995, 1996, Ferraresi 1997, Fuß 2003). Furthermore, we have seen that apparent deviations from V2 in *wh*-questions can be attributed to extra-grammatical factors, namely traces of Greek word order that entered the Gothic Bible via literal word-by-word translations of the Greek source text. Thus, the evidence available to us suggests that the earliest stages of Germanic already showed a form of systematic V2 limited to

wh-questions. In addition, I have shown that clauses introduced by the foregrounding particles *þaruh* 'there+uh' and *þanuh* 'then+uh' constitute another context where V2 order occurs regularly. First, we have seen that there is a strong V2 tendency in connection with *þaruh* 'there+uh', where in two-thirds of all examples, the finite verb occupies second position. Interestingly, V2 order is already even more systematic in clauses introduced by *þanuh* 'then+uh', with only a single out of 74 examples exhibiting a clear deviation from V2 that cannot be attributed to Greek influence. The next section examines the apparently more advanced V2 syntax of Old English, arguing that the core V2 properties of this early Germanic language reduce to the very same contexts where V-to-C movement is found in Gothic.

3 Old English: operator V2 + 'pseudo V2'

It is a well-known fact that Old English (OE), exhibits word order patterns reminiscent of the Modern Germanic V2 languages, that is, the finite verb occupies the second position after a fronted XP, leading to subject-verb inversion (examples taken from Trips 2002:231):²²

(23) a. *object–V_{fin}–subject*

[Pæt hus] **hæfdon** *Romane* to ðæm anum tacne geworht ...

that house had Romans to the one sign made

'The Romans had made that house to their sole sign.'

(Orosius, Or_3:5.59.3.1042)

b. *PP–V_{fin}–subject*

[On þysse dune ufanweardre] **bæd** *Sanctus Albanus* fram Gode ...

on this hill higher up bade Saint Alban from God

'On this hill higher up Saint Alban asked from God ...'

(Bede, Bede_1:7.38.30.323)

²² If not indicated otherwise, the OE examples are taken from the *York-Toronto-Helsinki Parsed Corpus of Old English Prose* (henceforth *York Corpus*).

c. *adverb–V_{fin}–subject*

[Uneaðe] **mæg** *mon* to geleafsuman gesecgan ...
 Hardly may man to faithful speak
 'Hardly may man speak to the faithful ...'
 (Orosius, Or_3:9.70.16.1292)

In this section, I take a closer look at the status of V2 and subject-verb inversion in OE, arguing that despite appearances, 'genuine' structural V2 patterns involving a spec-head relationship between a fronted XP and the finite verb were actually restricted to the very same contexts that triggered systematic V-to-C movement in Gothic (which is again reminiscent of the limited V2 properties of Modern English, cf. Eyþórsson 1995, Haeberli 1999, Fuß 2003).

3.1 Systematic deviations from V2 in Old English

Patterns such as those illustrated in (23) have fueled the claim that OE shared basic traits with the OV+V2 syntax of present-day German (cf. e.g. Stockwell 1977, 1984, Kemenade 1987). Upon closer inspection, it has become clear, however, that OE cannot be analyzed on a par with, say, Modern German. First, it can be shown that OE exhibits much more syntactic variation than the present-day Germanic V2 languages, that is, V2 is actually less consistent than suggested by the examples in (23) (see also chapter 2 above on variation in basic word order in OE). Moreover, we can observe systematic deviations from V2 order that do not show up in the modern V2 language. These facts are usually taken to suggest that the kind of V2 exhibited by OE differs significantly from the kind of V2 that is characteristic of the modern Germanic V2 languages (Hulk and van Kemenade 1995, Pintzuk 1999, Haeberli 1999, Fuß 2003, Roberts 1996, 2007 among many others). In what follows, I will give an overview of word order properties in main clauses, focusing on V2 and systematic deviations from it.

3.1.1 V-first and V-final main clauses

In OE, as in all other early Germanic languages (see section 5 below on Old High German), we still find a good deal of V-1 and V-final main declaratives:

(24) *V-first main declarative*

wæs se fruma þus awritten [...]

was the beginning thus written

‘The beginning was written as follows [...]’

(Bede 48.4; Pintzuk 1991: 68)

(25) *V-final main declarative*

se manfulla gast þa martine **gehyrsumode** [...]

the evil spirit then Martin obeyed

‘The evil spirit then obeyed Martin [...]’

(ÆLS 31.1050; Pintzuk 1991: 68)

While matrix SOV order is generally taken to be the residue of an earlier (Pan-Germanic) system that has been replaced by a basic V2 or SVO syntax in the course of time (cf. e.g. Bacquet 1962: 691; Mitchell 1985: 969), V1 order is robustly attested in main clauses of OE. Its exact (discourse) function, however, is less easy to pin down. Suggestions in the literature range from stylistic considerations such as influence of the tendency to place the verb before the first stress in verse texts (originally with non-pronominal DP subjects, later extended to subject pronouns, cf. Campbell 1970) to discourse strategies, compare the following quote taken from Mitchell (1985: 978):²³

“[VS order] is used to contrast with other orders; to link sentences, especially in oral narratives; to introduce new facts or a new train of thought; to change the emphasis; or to narrate a dynamic sequence of events by developing the narrative from stage to stage.”

²³ Note that that there is a conspicuous overlap between the apparent discourse functions of V1 order and the discourse functions of clause-initial *þa, þonne* that trigger obligatory V2 + inversion in OE (see below for details). This perhaps can be taken to suggest that the historical origin of these V2 orders were instances of V1 in which the relevant discourse functions were reinforced by adding a anaphoric/deictic linker in clause-initial position. See below for some discussion.

3.1.2 V3 orders involving non-pronominal DP subjects

OE exhibits word order patterns similar to Modern English, that is, there are examples in which a non-pronominal subject DP fails to undergo inversion with the finite verb. As will be shown later, this pattern is a characteristic feature of OE, which cannot be found in Old High German. According to Kroch and Taylor (1997:304), this type of V3 order often occurs with temporal adverbs that function as 'scene setters':²⁴

(26) a. [Æfter þeossum wordum] [se Hælend] **cwæþ** to his leornerum...
 after these words the Savior spoke to his disciples
 (Blickling 135; Swan 1994:241)

b. [Her] [Oswald se eadiga arceb] **forlet** þis lif.
 in-this-year Oswald the blessed archbishop forsook this life
 (ASC, Laud (992); Kroch and Taylor 1997:304)

Interestingly, it can be shown that this word order pattern is actually quite frequent in the OE data (cf. e.g. Swan 1994, Koopman 1998, and Haeberli 1999, 2000). Based on a quantitative analysis of ten OE text samples, Haeberli (2000: 4) calculates a percentage of 28.7% for cases where fronting of a non-operator does not lead to inversion of a nominal subject and the finite verb. That is, while V2 is the majority pattern (over 70%) in main declaratives with a fronted non-operator, the order XP-DP_{subj}-V_{fin} is a robustly attested characteristic of OE. Below, I am going to argue that the subject occupies SpecTP in these patterns, which anticipates structural properties of Middle and Modern English.

3.1.3 V2 and pronoun placement

A set of well-known violations of V2 involves the placement of pronominal arguments. In clauses with a fronted non-operator, (weak) subject pronouns systematically intervene between the clause-initial XP and the finite verb, giving rise to V3 order:²⁵

²⁴ See Haeberli (2000) for arguments that these violations of V2 cannot be analyzed as instances of verb-final order, which still was a word order option in main clauses of OE (see above).

²⁵ Note that there are very few examples in which the pronoun follows the finite verb and a fronted non-operator, resulting in V2-patterns similar to the modern Germanic languages, cf.

(27) a. [Æfter his gebede] *he ahof þaet cild up.*
 after his prayer he lifted the child up
 ‘After his prayer, he lifted the child up.’
 (AHth, II, 28; Kemenade 1987:110)

b. [þas þing] *we habbaþ be him gewritene.*
 these things we have about him written
 ‘These things we have written about him.’
 (PC, 1087, 143; Kemenade 1987:110)

Furthermore, object pronouns can also intervene between the finite verb and a fronted non-operator XP, as shown in (28). In contrast to the placement of subject pronouns, fronting of pronominal objects is only optional (cf. Mitchell 1985, Haeberli 1999: 337). Moreover, it appears that object pronouns typically occupy a high, preverbal position only if the subject is not a pronoun itself, with the subject DP either fronted, as in (28b) or occupying a lower position, as in (28a) (see Kemenade and Los 2006 on the relative placement of nominal and pronominal arguments in embedded clauses:

(28) a. *Fela spella him sædon þa Beormas.*
 many stories him told the Permians
 ‘the Permians told him many stories’
 (Oros., 14.27; Kemenade 1987:114)

b. *and se halga wer hine betæhte ansundne his fæder*
 and the holy man him.ACC handed-over healthy his father
 ‘and the holy man handed him over to his father healthy’
 (ÆCHom ii. 182.20; Mitchell 1985: 966)

(i) *eadig eart ðu abgar*
 blessed art thou Abgar
 ‘blessed art thou, Abgar...’
 (ÆLS 24.113; Pintzuk 1999:122)

The exceptional status of this pattern is confirmed by a quantitative analysis carried out by Haeberli (2000: 13). Haeberli shows that in seven out of ten OE text samples, inversion with a pronominal subject does not occur at all. In the remaining three texts, there are only six examples (Bede: 1; Chronicle A: 1; Orosius: 4) where the subject pronoun undergoes inversion (leading to V2), contrasting with 89 examples where no inversion takes place.

In other words, it seems to be impossible that an object pronoun is fronted to preverbal position, while the subject pronoun is left behind in postverbal position (see Kemenade and Los 2006 on the relevant facts in embedded clauses). Note, however, that all pronominal elements can form a complex that precedes the finite verb.²⁶ This is illustrated in (29).

(29) and seofon ærendracan *he him hæfde* to asend
 and seven messenger he him had to send
 'and he had to send him seven messengers'
 (Parker, 905; Pintzuk 1999)

While (27) is reminiscent of the word order of Modern English, (28) and (29) suggest that more is at issue here. In particular, it seems that in OE, the placement of pronouns was governed by a special operation that obligatorily moved subject pronouns to a position directly to the left of the finite verb (see below for arguments that non-pronominal subjects normally occupied a lower structural position). Under certain (locality) conditions, the very same operation could also apply to object pronouns, albeit in a merely optional fashion.

However, V3 orders in connection with pronouns do not occur completely freely in OE (cf. Mitchell 1985, Kemenade 1987, Kroch and Taylor 1997, Pintzuk 1999, Haeberli 1999 among many others). As is well-known, strict V2 order (+ inversion) is observed if the fronted element is an operator such as a *wh*-phrase as in (30) or the (clitic) negation *ne* as in (31). Here, the pronoun invariably follows the finite verb:

(30) a. Hwæt **sculon** we þæs nu ma secgan?
 what shall we afterwards now more speak
 'What shall we afterwards speak now more?'
 (Bede,Bede_2:9.132.1.1253)

b. hu **wurð** he elles gelæred?
 how was he otherwise taught
 'How was he taught otherwise?'
 (Bede,BedePref:2.11.153)

²⁶ The order in such clusters appears to be invariably subject–direct object–indirect object, cf. Kemenade and Los (2006: 235).

(31) a. ne **bið** *he* lengra þonne syfan elna lang.
 NEG is he longer than seven ells long
 'He is not taller than seven ells.'
 (Orosius,:1.15.2.149)

b. Ne **meaht** *þu* deman Gallia biscopas buton heora agenre
 NEG might you judge Gaul's bishops but their own
 alderlicnesse, [...]
 authority
 'You might not judge the Gaul's bishops but their own authority [...]'
 (Bede,Bede_1:16.74.5.679)

In addition, subject-verb inversion (with all kinds of subjects, including subject pronouns) is obligatory in clauses introduced by the temporal adverbs *þa*, *þonne* 'then' (cf. Mitchell 1985, Kemenade 1987, Kroch and Taylor 1997, Pintzuk 1999, Kemenade and Los 2006, Trips and Fuß 2008).²⁷

(32) *þa* **for** *he* norþryhte be þæm lande;
 then went *he* northwards to that land
 'Then he went northwards to that land.'
 (Orosius,:1.14.7.128)

(33) *þonne* **ærnað** *hy* ealle toweard þæm feo;
 then run-to they all towards the treasure
 'Then they all ran towards the treasure.'
 (Orosius,:1.17.21.233)

²⁷ V2 order can also be observed with other temporal adverbs such as *nu* 'now', cf.

(i) **Nu** **hæbbe** *we* ymb Affrica landgemæro gesæd.
 now have we about Africa's boundary said
 'Now we have spoken about Africa's boundary.'
 (Orosius,:1.20.25.302)

However, V2 order is much less regular with *nu* than with *þa* and *þonne* (cf. Mitchell and Robinson 1988:69; similar facts hold with respect to *here*, *þær*, *þider*, *þanon*, *swa* and *þeah*, cf. Mitchell 1985 for an overview).

Note that these word order facts are reminiscent of the contexts that trigger systematic subject-verb inversion in Gothic. In particular, it appears that there is a class of adverbial elements that trigger obligatory V2 on a par with syntactic operators such as *wh*-phrases and (fronted) negation. The relevant findings can be summarized as follows:²⁸

(34) *V2 and V3 in non-embedded sentences of OE*

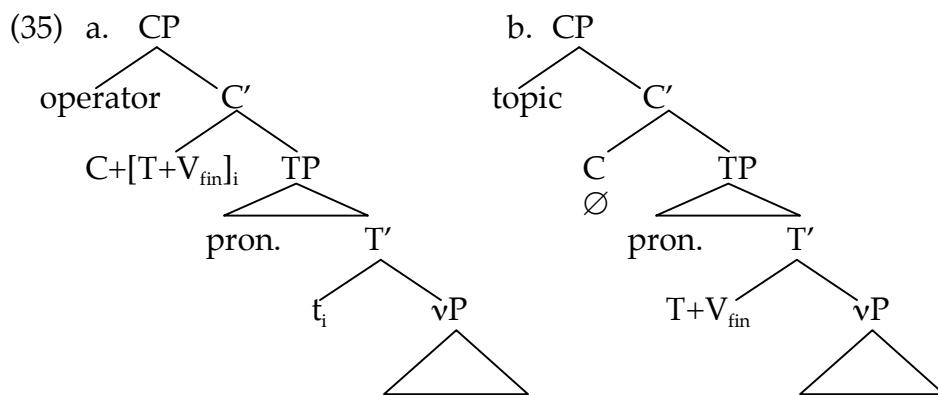
- a. $XP - V_{fin} - DP_{subj.} \dots$
- a'. $XP - DP_{subj.} - V_{fin} \dots$
- b. $XP - (\text{subject}) \text{ pronoun} - V_{fin} \dots$
- b'. $*XP - V_{fin} - \text{subject pronoun} \dots$
- c. $WH/\text{NEG}/\text{þa}/\text{þonne} - V_{fin} - DP \text{ subject/subject pronoun} \dots$
- c'. $*WH/\text{NEG}/\text{þa}/\text{þonne} - DP \text{ subject/subject pronoun} - V_{fin} \dots$

The placement properties of pronominal subjects suggest a diachronic continuity from OE to Modern English: Fronting of non-operators leads to the order *XP – subject pronoun – V_{fin}* , while V2 is forced by fronted operators (abstracting away from *þa* and *þonne*). Thus, a key difference between OE and Modern English seems to consist in the loss of the pattern (34a), that is, subject-verb inversion with full nominal subjects in cases where a non-operator is fronted (cf. Haeberli 1999, 2000, Fuß 2003). In particular, the different distribution of non-pronominal subjects and subject pronouns suggests that in contrast to Modern English, the placement of pronominal elements is governed by a special mechanism in OE. This conclusion is corroborated by the fact that under certain circumstances, object pronouns may occupy a preverbal position as well. Furthermore, another difference concerns the particular behavior of a class of (temporal) adverbs that trigger systematic subject-verb inversion, in contrast to present-day *then* (i.e., (34c)). In the following section, I will outline an analysis of V2 and inversion in OE that pays attention to these considerations, elaborating and modifying proposals put forward in Fuß (2003) and Trips and Fuß (2008).

²⁸ Furthermore, as noted above, OE exhibits a residue of V1 and V-final order in main declaratives, cf. e.g. Mitchell (1985), Pintzuk (1999).

3.2 Toward an analysis

The facts presented in the previous section led many researchers to assume that the clause structure of OE differs considerably from that of the present-day Germanic V2 languages (cf. Cardinaletti and Roberts 1991; Pintzuk 1993, 1999; Hulk and van Kemenade 1995; Kemenade 1997, 1999; Roberts 1996; Kroch and Taylor 1997; Haeberli 1999, 2000, Fischer et al. 2000, Fuß 2003, Fuß and Trips 2002, Trips and Fuß 2008). According to a basic assumption shared by these approaches, weak subject pronouns occupy a fixed position at the left edge of IP (e.g., Cardinaletti and Roberts: Agrs¹⁰; Pintzuk 1999, Kroch and Taylor 1997: adjoined to IP; Roberts 1996: SpecFinP; Haeberli 1999: SpecAgrsP; Fuß 2003, Fuß and Trips 2002, Trips and Fuß 2008: SpecTP). As a result, the relative order of pronoun and finite verb can be used as a diagnostic for the structural position of the verb, which is taken to occupy different positions in V2 and V3 orders (in connection with pronouns): Only in contexts with fronted operators (e.g., neg, wh, imperatives), the finite verb moves into the C-domain, crossing the subject pronoun at the left edge of IP and giving rise to V2 + inversion (possibly due to the presence of criterial features in C which are linked to the licensing of syntactic operators in SpecCP, cf. e.g. Rizzi 1996). In contrast, the finite verb occupies a lower position in the inflectional domain (to the right of subject pronouns) in clauses with fronted non-operators, giving rise to V3 orders.²⁹ Adopting a ‘minimalist’ clause structure (Chomsky 1995, 2000), this can be implemented as follows (cf. Fuß 2003):



²⁹ Pintzuk (1993, 1999) and Kroch and Taylor identify this head position as Infl, Cardinaletti and Roberts (1991) as Agr1, Roberts (1996) as Fin, Kroch and Taylor (1997), Haeberli (1999) as Agrs, Hulk and van Kemenade (1995), Kemenade (1997), Fischer et al. (2000) as a non-specified functional head F, Fuß (2003), Fuß and Trips (2002), Trips and Fuß (2008) as T.

Under the assumption that subject pronouns occupy SpecTP in OE (Fuß 2003), (36) and (37) serve to illustrate the relevant structures for OE clauses with and without inversion in connection with pronominal subjects, respectively:

(36) *wh – V_{fin} – subject pronoun*

[_{CP} hu [_C **wurð**_i+C [_{TP} *he* [_T *t_i*] [_{vP} *elles* gelæred *t_i*]]]]?

 how was he otherwise taught

‘How was he taught otherwise?’

(Bede, BedePref:2.11.153)

(37) *XP – subject pronoun – V_{fin}*

[_{CP} **Æfter** his gebede [_{TP} *he* [_T **ahof**_i [_{vP} **þaet** *cild* up *t_i*]]]

 after his prayer he lifted the child up

‘After his prayer, he lifted the child up.’

(AHth, II, 28; Kemenade 1987:110)

This approach raises the question of how we can account for V2 patterns (+ inversion) in connection with non-pronominal subjects. It is immediately clear that it is not possible to assume that full nominal subjects obligatorily occupy SpecTP on a par with pronominal subjects since this would lead us to expect generalized V3 order. However, as already briefly hinted at above, there are some indications that the syntactic distribution of non-pronominal subjects differed considerably from the distribution of subject pronouns (and pronouns in general) in OE. In particular, it can be shown that full DP subjects may remain in a low, presumably vP-internal position (cf. Pintzuk 1999, Kemenade 1997, Kroch and Taylor 1997, Haeberli 1999, Fischer et al. 2000). The relevant pieces of evidence come from the placement of negation and, in particular, adverbs relative to different kinds of subjects.

First of all, in examples with multiple sentential negation (consisting of the clitic *ne* and the negative adverb *na*) pronominal subjects appear to the left of *na* whereas nominal subjects consistently follow *na* (Haeberli 1999: 340ff., Fischer et al. 2000: 124f.):

(38) a. Ne **het** *he us na leornian heofonas to wyrcenne*
 not ordered he us not learn heavens to make
 'He did not bid us learn to make the heavens.'
 (ÆLS, 127; Fischer et al. 2000:125)

b. **Nis** *na* *se halga gast wuniende on his gecynde*
 not-is not the holy ghost existing in his nature
 swa swa he gesewen wæs
 as he seen was
 'The Holy Ghost is not existing in his nature as he was seen.'
 (ÆCHom I, 22.322.17; Fischer et al. 2000:125)

In a similar vein, other negative adverbs such as *næfre* 'never' may intervene between the finite verb and non-pronominal subject DPs:

(39) ne **abihð** *næfre* *Eadmund Hingware on life [...]*
 not surrenders never Edmund Hingwar on life
 'Edmund never surrenders to Hingwar alive [...]'
 (ÆLS, IV, 322.116; Haeberli 1999: 340)

Under standard assumptions concerning the structural positions of negative adverbs such as *na* – either located in SpecNegP or adjoined to VP – the above examples suggest that nominal subjects can remain in their θ-position. In contrast, it appears that pronominal subjects must occupy a higher position, presumably at the left edge of TP.

In fact, it turns out that all kinds of adverbial elements may intervene between the finite verb and a non-pronominal DP subject (both definite and indefinite) in inversion structures (cf. Haeberli 1999: 341f.). This is illustrated in (40) for short (single-word) adverbs (e.g., *syððan* 'afterwards', *eft* 'again', *semle* 'always', *oft* 'often', *ðær* 'there', *nu* 'now', *eac* 'also'), and in (41) for larger adverbial adjuncts.

(40) a. **gielden** *syððan* *his mægas þone wer*
 pay afterwards his male-kinsmen the man's-legal-value
 'Afterwards, his relatives should pay the man's legal value.'
 (Law2, 120.74.1; Haeberli 1999: 341)

b. þa **cwæð** eft se dry [...]
 then said again the magician
 'Then, the magician said again, [...]'
 (ÆLS, I, 312.71; Haeberli 1999: 341)

c. hu ne **bið** ðonne semle þæt lange yfel wyrse þonne ðæt scorte?
 how not is then always the long evil worse than the short
 'Why isn't then always a long misery worse than a short one?'
 (Bo, 117.31; Haeberli 1999: 341)

(41) a. & ðonne **wyrð** þurh Godes mihte sona deofol swyðe geyrged
 and then gets through God's power soon devil very-much terrified
 'Then, soon, the devil is very much terrified through God's power.'
 (Whom, 176.28; Haeberli 1999: 341)

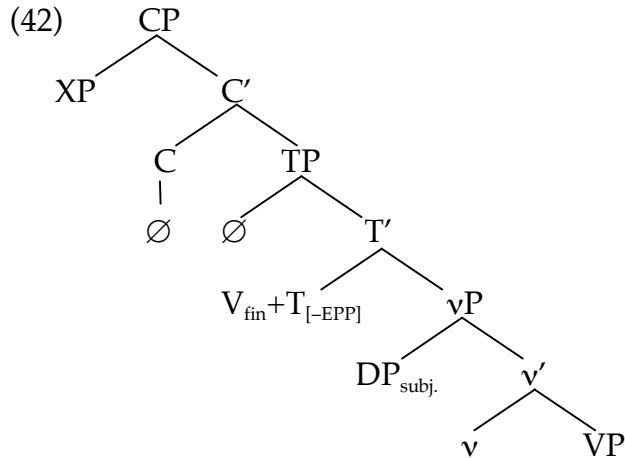
b. & **gearwige** eac to huslgange oft & gelome gehwa
 and prepare also to going-to-Eucharist often and frequently each
 hine sylfne
 him self
 'And everyone should prepare himself often to go to the Eucharist.'
 (Law3, 242.22.1; Haeberli 1999: 341)

c. Ne **bearh** nu foroft gesib gesibban
 not saved now very-often kinsman kinsman
 'At that time, kinsmen didn't help each other very often.'
 (Whom, 269.55; Haeberli 1999: 342)

Interestingly, we can again observe that there is a clear-cut distinction between pronominal vs. non-pronominal subjects with respect to their position relative to adverbial elements: Full subject DPs (both definite and indefinite, compare e.g. (41a) vs. (41c)) may appear to the right of adverbs, while pronouns systematically occur directly adjacent to the finite verb in C, and to the left of all adverbial elements, compare the following statement taken from Haeberli (1999: 342): "The only clear restriction that can be observed with V-XP-SU orders is that the subject is never a pronoun."³⁰

³⁰ According to Haeberli (1999: 342), there is only a single example in the whole OE Corpus he used that exemplifies the order V-XP-subject pronoun:

These facts imply that OE lacked a generalized EPP feature in T that attracts all kinds of subjects.³¹ Accordingly, subject-verb inversion with non-pronominal DP subjects can be attributed to a configuration where the finite verb moves to an (head-initial) inflectional head while the subject stays behind in a lower position (presumably its θ-position SpecvP).



(43) $XP - V_{fin} - DP$ subject

[_{CP} **Pæt** hus [_{TP} Ø [_T **hæfdon** [_{vP} *Romane* to ðæm anum tacne geworht]]]]
 that house had Romans to the one sign made
 'The Romans had made that house to their sole sign.'

(Orosius, Or_3:5.59.3.1042)

In other words, I claim (basically following Fuß 2003) that a vast number of V2 orders found in OE result from a configuration where the fronted XP and the finite

(i) & **ladige** on bam husle *he ana* hine sylfne æt anfealdre spræce
 and exculpate on the Eucharist he alone him self at single charge
 'He alone shall exculpate himself for a single charge at the Eucharist.'

(Bede 252.7; Haeberli 1999: 341)

However, as pointed out by Haeberli, the pronominal subject is quite likely to be stressed in the above example, due to the presence of *ana* 'alone'. Thus, it is actually expected to behave similar to a full DP subject. Importantly, however, there are no clear cases of unstressed pronouns that are non-adjacent to the finite verb.

³¹ Further support for this assumption comes from the fact that OE displays a number of subjectless constructions in which neither a nominative subject nor an expletive element shows up in the subject position (SpecTP), in contrast to Modern English. Relevant examples include weather verbs, experiencer verbs and impersonal passives (cf. the examples in (64) below). Note that the frequent presence of subject pronouns in SpecTP (due to their anaphoric nature) possibly supported the development of [+EPP] T in the ME period (see section 4 below).

verb do not enter into a spec-head configuration in the C-domain. Rather, they are merely linearly adjacent, due to the fact that both C and SpecTP may remain empty. In Fuß (2003), superficial V2 orders resulting from this structural configuration are called 'pseudo V2'.

In contrast, examples exhibiting the pattern $XP - DP_{subj} - V_{fin}$ are probably the result of moving the non-pronominal subject to SpecTP (a structure which is later generalized in the Middle English period, see section 4):

(44) $XP - DP_{subject} - V_{fin}$

[_{CP} *Æfter þeosum wordum* [_{TP} *se Hælend* [_{T'} **cwæp** [_{vP} *to his leornerum*]]]]]
after these words the Savior spoke to his disciples
(Blickling 135; Swan 1994:241)

While this set of assumptions suffices to derive the most frequent word order patterns found in main clauses of OE (further provisos are needed for V1 and V-final orderings, see Fuß 2003), the details of the analyses proposed so far are in need of further discussion (in particular with respect to the triggers of the individual movement processes). In the following, I will focus on two particular questions that have been (and still are) the subject of much debate in the literature, namely (i) the correct analysis of the placement asymmetries between pronominal and non-pronominal subjects,³² and (ii) the status of *þa*, *þonne* 'then' as triggers of obligatory inversion in OE.

Following Kemenade (1987), the fact that the temporal adverbs *þa*, *þonne* force V2 (i.e., V-to-C movement) in OE is commonly accounted for by assuming that these elements should be analyzed as operators on a par with *wh*-phrases and negation:

(45) [_{CP} *þa/þonne* [_{C'} *V_{fin}* [_{TP} *pron.* [_{T'} *t_v* [_{vP} ...]]]]]]]

$\overbrace{}$

³² Proposals concerning the placement of (subject) pronouns of OE can be roughly classified into three major strands: (i) OE subject pronouns are clitics. Their special distribution results from special placement rules (either in the syntax or at PF: left/right adjunction, PF repositioning; Kemenade 1987, Kiparsky 1995, Tomaselli 1995, Kroch and Taylor 1997, Pintzuk 1999; see Koopman 1997 for insightful discussion of the clitic status of OE pronouns); (ii) OE subject pronouns occupy a special structural position associated with the licensing of syntactic clitics/weak pronouns (Cardinaletti and Roberts 1991: Agr1⁰, Roberts 1996: SpecFinP, Hulk and van Kemenade 1995, Fischer et al. 2000: SpecFP); (iii) only pronominal subjects move to the feature checking position of subjects (Haeberli 1999, 2003: SpecAgrsP; Fuß 2003: SpecTP).

However, it is not quite obvious why adverbs like *þa* and *þonne* should have the status of syntactic operators. The only motivation for this analysis seems to come from the fact that they apparently trigger verb movement to C. Moreover, although fronted operators such as *wh*-phrases and negation continue to trigger inversion in Modern English, *then*, the present-day equivalent of OE *þa*, *þonne* fails to do so:

(46) a. *Then **will** Harry/he read that book.
 b. Then Harry/he **will** read that book.

Of course, one might argue that the Modern English reflex of OE *þa*, *þonne* lost its operator status and with it the ability to trigger verb movement. However, this assumption should lead us to expect a dramatic change in the semantics of 'then', which is not borne out by the facts: Mod. English *then* receives an interpretation similar to OE *þa*, *þonne* (cf. e.g. Kroch and Taylor 1997: 303). Despite these problems, Kemenade and Los (2006: 226) revive the operator analysis of *þa*, *þonne* by claiming that clause-initial *þa*, *þonne* is a discourse operator that signals discourse continuity and "triggers movement of the finite verb to C in much the same way as a *wh*-operator or a negative operator". While the purported function of *þa*, *þonne* as markers of discourse continuity seems to be partially correct (see below for detailed discussion), Kemenade and Los fail to be explicit about the notion of 'discourse operator'. In particular, nothing is said about the question of why the finite verb is required to occupy C in the context of *þa*, *þonne*, but not with other elements such as *nu*, *þær* etc. that would equally qualify as 'discourse operators', but fail to consistently trigger inversion in the way *þa*, *þonne* does.

Following basic insights in Trips and Fuß (2008), the next section argues that the special word order patterns linked to subject pronouns and elements such as *þa*, *þonne* are actually not separate phenomena, but two sides of the same coin. I am going to suggest that subject pronouns and *þa*, *þonne* compete for the same structural position, which is identified as SpecTP. In addition, I will claim that the placement properties of pronouns and certain adverbs are connected to the fact that OE was a discourse-configurational language, in which structural positions (and word order), were not primarily associated with grammatical functions, but rather served to implement discourse-related properties (information-structure, anaphoricity etc.).

3.3 The syntax-discourse interface and inversion in OE

To continue a given discourse in a coherent way, a set of conditions concerning the syntax-discourse interface must be met. This is particularly obvious with respect to clausal typing, information-structural distinctions, and the interpretation of anaphoric expressions. For example, sentential mood must be coded in order to distinguish between questions and assertions, information-structural differences such as 'topic' and 'focus' must be properly marked (via word order, by morphological markers, or by assigning certain stress patterns), and anaphoric expressions must be anchored in the discourse to warrant a correct interpretation (e.g., pronouns must receive a referential index). Rizzi (1997) proposes that the first and second of these properties are directly implemented into the structure of the clause. More specifically, clausal typing is associated with properties of Force, which closes off the series of the projections in a split-CP and represents the interface to the discourse context (or a matrix clause), while the encoding of information-structural distinctions such topic and focus is linked to specifier positions made available by the relevant functional heads in the left periphery of the clause (Topic and Focus, respectively). Of course, languages may differ with respect to the extent to which they encode discourse properties by syntactic means such as word order (i.e., via overt movement to positions such as SpecTopP or SpecFocP). Languages where word order does not primarily serve to identify grammatical functions, but rather is used to signal the information-structural status of different elements of the clause are often called 'discourse-configurational languages'.

In current work on OE, it is often pointed out that certain characteristics of OE can be taken to indicate that OE was discourse-configurational as well (cf. e.g. Fischer et al. 2000, Kemenade 2002, Kemenade and Milicev 2005, and Kemenade and Los 2006). The basic observation consists in the fact that topical material referring anaphorically to discourse referents figuring prominently in the discourse contexts (in particular personal pronouns, but sometimes also definite/specific nominal expressions marked e.g. by a demonstrative) typically occupies a position at the left edge of the inflectional domain (i.e., directly to the right of the complementizer in embedded clauses, or adjacent to a fronted finite verb in main clauses). In contrast, non-definite/non-specific noun phrases, which typically represent the focus of the clause (or, more generally, new information), occupy a position further to the right, for example directly to the left of non-finite verbs (quite similar to e.g. Modern German). Kemenade and Los (2006: 237f.) illustrate the syntactic effects of discourse-

configurationality in OE with the following embedded clause in which the object pronoun (representing old information) is situated to the left of *bonne*, while the non-definite subject NP (representing new information) occupies a lower position to the right of *bonne*:

(47) Gif *hine* *bonne* [yfel mon] **hæfð** [...]
 if him then evil man has
 'If an evil man has him...'
 (coboeth,Bo:16.38.26.702; Kemenade and Los 2006: 237f.)

In what follows, I present evidence from OE suggesting that syntactic configurations may also be used to implement the discourse anchoring of anaphoric expressions in a discourse-configurational grammar. Before we turn to the specifics of this proposal, let's have a closer view at the temporal interpretation of 'then', and how these considerations carry over to the analysis of *þa/þonne* in OE.

3.3.1 The temporal interpretation of clause-initial *þa/þonne*

In the literature on the lexical semantics of 'then', the adverb is often analyzed as a temporal anaphor that introduces a temporal relation between the events described by two successive sentences.³³ It has been claimed that the anaphoric character of 'then' requires that it be linked to an anchor time given in the discourse context (cf. Smith 1981, Schiffrin 1992, Glasbey 1993, Thompson 1999).³⁴ More to the point, 'then' is usually taken to express that the event or state described by the 'then'-clause is temporally situated after a time point/interval given in the immediate discourse context.

Furthermore, it has been observed that the temporal interpretation of 'then' is determined by its syntactic position. More precisely, there are systematic differences concerning the interpretation of clause-initial and clause-medial/final 'then' (cf. e.g. Schiffrin 1992, Thompson 1999, Roßdeutscher 2005a,b on Modern English *then*, Roßdeutscher 2005a,b, Roßdeutscher and von Stutterheim 2006 on German *dann*).

³³ See also Partee (1973) for the idea that tenses exhibit anaphoric properties similar to pronominal elements.

³⁴ In somewhat more formal terms, we can say that the interpretation of temporal anaphora requires the assignment of a temporal index given in the discourse.

This is illustrated with examples (48) and (49) from Modern English. When occurring in clause-initial position, *then* is to be analyzed as a sentence adverb (presumably associated with the left edge of the inflectional domain, i.e., IP, or TP) that gives rise to an interpretation where the events described by two successive sentences are understood as temporally ordered – in (48), the speaking event occurs after the visiting event, and there is no temporal overlap between these events (sequential/ordered reading, henceforth *sequential 'then'*):

(48) Mary visited the exhibition. **Then** she spoke to the reporters.

In contrast, clause-final placement of *then* (presumably a VP-adverb, henceforth *cotemporal 'then'*) leads to an interpretation where the event described by the 'then' clause is taken to overlap with the event described by the previous clause – in (49), Mary spoke to the reporters while she was visiting the exhibition (note that clause-final 'then' still seems to indicate that the speaking event began somewhat after the visiting event):

(49) Mary visited the exhibition. She spoke to the reporters **then**.

Let's assume (basically following Thompson 1999, Roßdeutscher and von Stutterheim 2006) that the different readings of 'then' are not to be attributed to different instances of 'then'. Rather, the lexicon contains only a single temporal anaphor 'then', the different interpretations of which are determined by the structural position it occupies in the structure of the clause.

Thompson (1999) argues that the different interpretations of anaphoric 'then' result from linking different times in tense structure with the relevant times given in the immediate discourse context (making use of a Neo-Reichenbachian model of tense structure, cf. Hornstein 1990). According to Thompson, cotemporal 'then' is attached to VP and serves to link the Event time (by assumption associated with VP) of two consecutive clauses, giving rise to an interpretation where the relevant events overlap temporally. In contrast, clause-initial 'then' is taken to be adjoined to IP, linking the Reference time (commonly associated with IP) of its clause with the Reference time of the previous clause, which entails that the relevant events take place one after the other. Leaving aside further technicalities and questions raised by this approach, what's important to keep in mind is that the sequential reading of

'then' becomes available when 'then' is associated with IP, while lower attachment leads to the cotemporal reading illustrated with (49) above.³⁵

Now let's come back to the question of how this carries over to the analysis of fronted temporal anaphora in OE. It has repeatedly been pointed out that in OE, clause-initial *þa*, *þonne* are typically used to mark a sequence of foregrounded successive actions/events that do not overlap temporally (cf. e.g. Foster 1975, Enkvist and Wårvik 1987, Wårvik 1995). Compare the Ohthere interpolation in Alfred's Orosius (reproduction of oral narrative; simple narrative structure):

(50) He sæde þæt he æt sumum cirre wolde fandian hu longe þæt land norþryhte læge, oþþe hwæðer ænig mon be norðan þæm westenne bude. **Þa** for he norþryhte be þæm lande; let him ealne weg þæt weste land on ðæt steorbord & þa widsæ on ðæt bæcbord þrie dagas. **Þa** wæs he swa feor norþ swa þa hwælhuntan firrest faraþ. **Þa** for he þa giet norþryhte swa feor swa he meahte on þæm oþrum þrim dagum gesiglan. **Þa** beag þæt land þær eastryhte, oþþe seo sæ in on ðæt lond, he nysse wæðer buton he wisse ðæt he ðær bad westanwindes & hwon norþan & siglde ða east be lande swa swa he meahte on feower dagum gesiglan. **Þa** sceolde he ðær bidan ryhtnorþanwindes, for ðæm þæt land beag þær suþryhte, oþþe seo sæ in on ðæt land, he nysse hwæþer. **Þa**

³⁵ While the analysis put forward in Thompson (1999) captures the anaphoric character of 'then' in some more or less intuitive way, it can be shown to suffer from a number of shortcomings. In particular, it fails to make explicit how the linking of times in tense structure actually leads to the relevant interpretative differences. For example, while it seems to a certain extent plausible to attribute the cotemporal reading to the linking of Event times, that assumption still fails to account for the fact that even cotemporal 'then' usually leads to an interpretation where the event described by the consequent clause sets in after the event described in the antecedent clause (i.e., the temporal settings overlap, but are not identical). In a similar vein, it is not clear to us how exactly the linking of Reference times results in an ordered reading (Thompson's 1999 article does not give any clear clues here). Of course, one may come up with some additional (pragmatic) machinery (e.g. one might assume that the event described by the first clause is (by default) interpreted as completed when a subsequent clause describes a second action or event that is viewed from the same Reference time (if there is no additional link associating the Event times of the two clauses)), but it would certainly be more desirable if such major aspects followed directly from the central assumptions of the theory. See Trips and Fuß (in prep.) for an account based on a more elaborate (semantic) analysis of 'then' based on DRT-based approaches such as Roßdeutscher (2005a,b) and Roßdeutscher and von Stutterheim (2006).

siglde he þonan suðryhte be lande swa swa he mehte on fif dagum gesiglan. **Ða** læg þær an micel ea up in on þæt land. (Or_1:1.14.5.226-235)

'He said that at one occasion he wanted to find out how far that land extended northwards, or whether any man lived north of the wilderness. **Then** he travelled northwards along the coast; keeping all the way the waste land on the starboard and the open sea on the portside for three days. **Then** he was as far north as the whalehunters go furthest. **Then** he travelled still northwards as far as he could sail in another three days. **Then** the land turned east, or the sea into the land, he didn't know which, but he knew that he there waited for a wind from the west and somewhat from the north and sailed then east along the coast as far as he could sail in four days. **Then** he had to wait for a due north wind, because that land turned there directly to south, or the sea into the land, he didn't know which. **Then** he sailed from there southwards along the coast as far as he could sail in five days. **Then** there was a large river reaching up into the land.' (Enkvist and Wårvik 1987: 234)

A cursory look at the text (50) shows that after a brief backgrounding introduction ('He said that at one occasion...'), the main story line is carried forward by a series of clauses introduced by *þa*. These clauses describe a sequence of actions/events that take place one after the other. Furthermore, note that the discourse referent the subject pronoun refers to remains constant, while the clause describes a new action or a change affecting the state of the discourse referent. In general, it seems that in passages such as (50), the narrative function of fronted *þa* consists in the marking of foregrounded actions/events, while its temporal properties lead to an interpretation where these events are understood as taking place sequentially (i.e., one after the other, without temporal overlap). In other words, it appears that clause-initial *þa*, *þonne* triggering inversion are instances of sequential 'then'. As noted above, this particular temporal interpretation is presumably associated with attaching 'then' to IP/TP, the locus of Reference time. Let's now closer examine the structural position of *þa*, *þonne*, focusing on the question of why clause-initial *þa*, *þonne* obligatorily trigger inversion in OE.

3.3.2 The distribution of sequential *þa/þonne* and subject pronouns in OE

A second look at the syntactic distribution of sequential 'then' in OE reveals that fronted *þa*, *þonne* may be preceded by a topicalized phrase, giving rise to V3 orders that are reminiscent of the kind of V3 typically occurring with subject pronouns (compare (27) above):

(51) a. On þa ilcan tima þa **comon** *hi* to Medeshamstede...
 at the same time then came they to M.

(ChronE_[Plummer]:870.5.1115)

b. Syððan þa **com** *he* to se cyng Eadgar, ...
 afterwards then came he to the king E.

(ChronE_[Plummer]:963.9.1396)

c. Mid þam ða **com** *þæt wif*.
 with that then came that woman

(ACHom_II,_8:67.14.1355)

d. Him þa **andswarode** *se biscop*.
 him then answered the bishop

(GD_1_[C]:4.28.5.293)

(52) a. On ðone sexteoðan dæg ðæs monðes þonne **bið**
 on the sixteenth day of-the month then is
Sancte Marcellus *tid* ðæs *papan*.
 Saint Marcel-GEN feast-day of-the pope-GEN

(Mart_5_[Kotzor]:Ja16,A.1.99)

b. For þi þonne **wacion** *we*, ...
 for that then stay-aware/watch we
 'because then we stay awake/watch...'

(ChrodR_1:14.6.277)

The above examples show that fronting of *þa/þonne* requires subject-verb inversion with both pronominal ((51a,b) and (52b)) and full nominal subjects ((51c) and (52a)). Furthermore, note that while (51a,b) and (52a) initially seem to suggest an analysis in terms of left dislocation, where a fronted adverbial expression co-occurs with an appropriate pronominal form (similar to e.g. Modern German *Am Samstag, da ging er ins Kino* 'On Saturday, he went to the movies.'), examples like (51c,d) and (52b)

clearly show that this analysis cannot be generalized to all cases of V3 with fronted *þa/þonne*.

What should a proper analysis of examples like (51) and (52) look like, then? Recall that in section 3.2 above, I have argued that fronted *þa, þonne* should not be analyzed as syntactic operators. Accordingly, we are led to expect that the finite verb does not move into the C-domain in (51) and (52), but rather occupies Infl/T, as in all other clauses with fronted non-operators. The data in (51) and (52) can then be accounted for under the following set of assumptions:

- (i) The fronted XP occupies SpecCP (or, a relevant spec in a split-C system).
- (ii) *þa, þonne* occupy a specifier in the inflectional domain, presumably SpecTP (the specifier of the head associated with the encoding of Reference time).
- (iii) The finite verb is located in T.
- (iv) All subjects, including pronouns, occupy a lower, vP-internal position.

Accordingly, examples such as (51) and (52) are analyzed as in (53), where *þa, þonne* occupy the specifier of TP (directly adjacent to the finite verb in T), while subjects generally stay behind in their vP-internal theta-position.

(53) *V3 with þa, þonne*

$[\text{CP} \text{ XP} [_{\text{TP}} \text{ } \text{þa/þonne} [_{\text{T}} \text{ V}_{\text{fin}} [_{\text{vP}} \text{ subject (pronoun) ...}]]]]$

If no material is fronted to clause-initial position (i.e., SpecCP), we derive the V2 + inversion order typically triggered by *þa, þonne* in OE ((32) and (33) above, repeated here for convenience):

(54) *þa for he norþryhte be þæm lande;*
 then went *he* northwards to that land
 'Then he went northwards to that land.'
 (Orosius,:1.14.7.128)

(55) *þonne ærnað hy ealle toweard þæm feo;*
 then run-to they all towards the treasure
 'Then they all ran towards the treasure.'
 (Orosius,:1.17.21.233)

Furthermore, if we compare structure (53) with the structure commonly posited for V3 orders with pronominal subjects, it appears that the preverbal position is either filled by the subject pronoun (giving rise to V3 without inversion) or with *þa, bonne* (leading to V2/V3 and obligatory inversion):

(56) *V3 orders with fronted non-operators*

$[_{CP} \text{topic} [_{TP} \text{subject pronoun} [_{T'} V_{\text{fin}} [_{vP} \dots]]]]$

The similarities between the patterns in (53) and (56) suggest that clause-initial *þa, bonne* and subject pronouns compete for the same structural position. Above, it has already been noted that this position is presumably to be identified with SpecIP/TP, since fronted *þa, bonne* specify Reference time, which is commonly associated with IP/TP (Hornstein 1990, Stowell 1995, Thompson 1999). Further evidence for this hypothesis comes from the observation that *þa* seems to require that the finite verb is in the preterite indicative (cf. Mitchell 1985: 308, Wårvik 1995), that is, *þa* has selectional properties linked to the inflectional domain of the clause (note, however, that *bonne* occurs with other tenses/moods as well).

Under these assumptions, the observation that *þa/bonne* seem to force subject pronouns to stay in a lower, post-verbal position can be analyzed as an instance of *Merge over Move* (Chomsky 1995). In other words, sequential *þa* and *bonne* are merged directly in the specifier of TP, thereby blocking the more costly alternative of moving the subject pronoun to this position.³⁶ As a consequence, the pronoun has to stay behind in its theta-position (SpecvP), with Case and agreement checking being accomplished via an Agree-relation initiated by T.³⁷ Next, I take a closer look at the

³⁶ Alternatively, the speaker may choose to merge *þa/bonne* in a lower position to achieve a different communicative effect (cotemporal 'then'). In that case, the (subject) pronoun can freely move to SpecTP, giving rise to a word order option where the pronoun precedes *þa/bonne*. See Roßdeutscher (2005a,b), Roßdeutscher and von Stutterheim (2006) for an analysis that attributes the different readings of 'then' to the relative ordering of subject pronoun and 'then'.

³⁷ To capture the intuition that discourse-relatedness is a property associated with the C-domain, one might assume that *þa, bonne* move to SpecCP after being merged in SpecTP. The shift from SpecTP to a higher position, which took place in the Middle English (ME) period, can then be analyzed as an instance of a change in which a movement dependency is reanalyzed as external Merge in the former target position (while the earlier first Merge position SpecTP is obligatorily occupied by the subject/expletive in ME and ModE, see below for details).

feature content of T and the properties of SpecTP in OE and the question of why both pronouns and certain temporal anaphora seem to be attracted to this position.

3.3.3 Discourse-configurationality and the nature of SpecTP in OE

As already noted above, in current work OE is often portrayed as a discourse-configurational language in which structural positions did not primarily encode grammatical functions, but rather were linked to discourse-related distinctions such as information-structure (cf. Fischer et al. 2000, Kemenade and Los 2006). The basic proposal I want to put forward is that this particular trait of OE was not limited to information-structural categories such as topic or focus, but also included the discourse-anchoring of anaphoric expressions. More precisely, I assume that the interpretation of anaphoric expressions was linked to a certain position in the clause-structure of OE. This hypothesis opens a new perspective on the apparently complementary distribution of referential (subject) pronouns and *þa/ponne* in preverbal position if we further assume that in OE, the relevant functional specifier is to be identified as SpecTP, the position apparently targeted by the elements in question. Let's now take a closer look at the relevant feature specifications shared by (subject) pronouns and the temporal anaphora *þa/ponne* that qualifies them as potential realizations of SpecTP in OE.

A property common to both pronouns and 'then' seems to be that the interpretation of these elements involves a variable that must be bound by (or identified with) a topical element/referent in the given discourse (see e.g. Heim and Kratzer 1998 for (referential) pronouns and Webber et al. 2003 for a unified analysis of (referential) pronouns, 'then', and discourse adverbials such as *otherwise* or *instead*). In what follows, I will use the feature [+anaphoric] to refer to this semantic property.³⁸ Another feature shared by the elements under investigation is morphosyntactic in nature. Historically, *þa* and *ponne* developed from demonstrative

³⁸ Note that this proposal seems to be at odds with Chomsky's (1981, 1986a) classification of pronouns as [+pronominal, -anaphoric] (see also Chomsky and Lasnik 1995: 41). However, recall that Chomsky's characterization of pronouns has originally been formulated for the particular purposes of Binding Theory. In recent theoretic work (following Reinhart and Reuland 1993), there is a tendency to dispense with this featural characterization of pronouns. Instead, their Binding properties are attributed to other feature specifications or the internal structure of pronouns. To avoid terminological confusion, it should be kept in mind that the present use of the feature [+anaphoric] differs from its use in standard GB theory.

pronouns. This aspect of the etymology of these elements is still transparent in OE, where the temporal anaphor *þa* is homophonous with the acc.sg.fem and nom./acc.pl of *se* 'the, that', suggesting that *þa/bonne* are nominal in nature as well. More precisely, they are to be analyzed as elements of the category D, similar to pronouns (see Postal 1969, and much subsequent work). Following Chomsky (2000: 139), I assume that [D] relates to referentiality/specification (i.e., indefinite non-specific noun phrases such as *someone* and bare plurals are merely NPs). Furthermore, anaphoric elements that relate to a (topical) discourse antecedent are commonly assumed to be necessarily specific in nature. In this way, the morphosyntactic character of pronouns and *þa/bonne* can be directly correlated with their interpretative properties, in the sense that [+anaphoric] elements are necessarily specified for a [D] feature.

Let's now address the question of how this relates to the feature specifications associated with T in OE. A different way of phrasing the idea that the interpretation of anaphoric expressions was linked to SpecTP is that the assignment of discourse-related (referential or temporal) indices to variables introduced by pronouns and temporal anaphora was connected to properties of T in OE. However, it seems plausible to assume that the syntactic component is blind to the exact semantic content of these indices, and such semantic processes in general (the only requirement being that the output of the syntactic computation must be legible by the semantic component of grammar, which interprets syntactic structures). Still, in a discourse-configurational language like OE, syntax may make available a structural variant of the relevant semantic assignment procedure that is necessary to interpret anaphoric expressions in a given clause. If we accept the idea that anaphoricity is linked to specificity, which in turn is correlated with the morphosyntactic feature [D], we might say that in OE, T may be endowed with a 'strong' [*D*] feature that requires that a specific/anaphoric element occupies SpecTP in the overt syntax (adopting the notational convention that features assigned a diacritic *_ require overt movement/PF realization, cf. Roberts and Roussou 2003, Sternefeld 2007). We can then ascribe the following feature content to the T-head of OE (recall that T lacked a generalized EPP-feature in OE):³⁹

³⁹ In Trips and Fuß (2008) it is assumed that the relevant feature triggering internal/external Merge of anaphoric elements in SpecTP is [+anaphoric]. However, as pointed out by Halldór Sigurðsson (p.c.), it is far from clear whether the syntactic computation can be driven by features that are purely semantic/pragmatic in nature. Following Chomsky (1995, 2000) and subsequent work, we

(57) T/OE: $[\mathbf{u}\phi, \mathbf{Tns}_{[(\text{PAST, PRESENT etc.})]} (*\mathbf{D}^*)]$

In other words, OE T contained a set of uninterpretable ϕ -features (represented as " $\mathbf{u}\phi$ ", cf. Pesetsky and Torrego 2001) that establish an Agree relation with the subject in SpecVP, an interpretable Tense feature, and, optionally, a (uninterpretable) strong $[*\mathbf{D}^*]$ feature that (if present) requires that specific/anaphoric material occupy SpecTP. The different syntactic distribution of (subject) pronouns, fully referential DPs and *þa*, *þonne* can then be accounted for as follows: subject pronouns are of course characterized by [D] (they are the prototypical [+anaphoric] elements). They are merged in SpecVP where they receive a thematic role and subsequently move to SpecTP (if possible) to eliminate T's $[*\mathbf{D}^*]$ feature. In this position, they are assigned a referential index associated with a topical discourse referent in the semantic component. In contrast, non-specific indefinite full subjects (which are merely NPs) may not move to SpecTP and remain in situ (SpecVP) in OE (but they may move further up into the C-domain to be interpreted as e.g. focus). What about specific full nominal subjects such as *the Bishop* or *his wife*? They are presumably DPs, carrying a [D] feature as well. Accordingly, we should expect them to be able to move to SpecTP as well, and in fact they do, giving rise to the order $\mathbf{XP}\text{-}\mathbf{DP}_{\text{subj}}\text{-}\mathbf{V}_{\text{fin}}$, which is one of the characteristic syntactic patterns of OE (found in 28.7% of the relevant examples, cf. Haeberli 2000). However, if we accept the notion that in OE, SpecTP was a position reserved for anaphoric material that relates to a topical discourse referent, it is fairly clear that not all non-pronominal subject DPs qualify for occupying that position. In particular, if a subject DP is specific, but introduces new information not mentioned in the previous discourse, we do not expect it to move to SpecTP. Rather, we expect it to remain in a position further to the right, which seems to be characteristic of focused elements and constituents introducing new information in OE (cf. Hinterhölzl 2004, Kemenade & Los 2006). In other words, we might say that due to the discourse-configurational status of OE, movement of non-anaphoric material to check T's $[*\mathbf{D}^*]$ feature was ruled out since it would have given rise to a deviant interpretation (although it would have been in principle possible in the syntax).

would rather expect that syntactic operations are triggered by the need to eliminate uninterpretable morphosyntactic features (such as e.g. $[*\mathbf{D}^*]$).

How does this analysis account for the fact observed above (cf. (28), repeated here for convenience) that object pronouns may also occupy SpecTP in cases where the subject is a non-pronominal DP?

(58) a. Fela spella *him* **sædon** þa Beormas.
 many stories him told the Permians
 'the Permians told him many stories'
 (Oros., 14.27; Kemenade 1987:114)

b. and se halga wer *hine* **betæhte** ansundne his fæder
 and the holy man him.ACC handed-over healthy his father
 'and the holy man handed him over to his father healthy'
 (ÆCHom ii. 182.20; Mitchell 1985: 966)

Note that prior to movement to SpecTP, the object pronoun has to move to the left edge of vP (to check Case, and presumably to circumvent the Phase Impenetrability Condition, Chomsky 2000 and subsequent work). After this operation, the object pronoun and the nominal subject are equidistant to T. However, in case the non-pronominal subjects is non-anaphoric, it cannot be attracted by T's [*D*] feature. Moreover, even if topical subjects have anaphoric properties, the pronoun realizes them in a more prototypical way, so that (at least in a discourse-configurational language such as OE) the option of raising the pronoun is preferred. This analysis requires that T's [u ϕ] and [*D*] features act as separate probes, raising a couple of technical questions. What must be ensured is that T agrees with the lower subject, and not with the object pronoun moved to SpecTP. This might be accounted for if we assume that the pronoun may not undergo further A-related operations (Case, agreement) after it has valued and eliminated its Case feature by moving to SpecvP.

However, we have also observed that object pronouns may not cross subject pronouns. This can attributed to the principle *Maximize matching effects* proposed in Chomsky (2001: 15): Due to the fact that the subject pronoun can check both T's [u ϕ] feature and T's [*D*] feature, raising the subject pronoun is more efficient than raising the object pronoun (which would require an additional Agree operation to establish subject-verb agreement). Cases where all pronouns form a complex (cf. (29) above, repeated here for convenience) that precedes the verb can perhaps be analyzed as instances of cluster formation, where prior to movement to a checking

position, lower checkees adjoin to a higher checkee with the same feature specification (cf. Sabel 2001, see Grewendorf 2001 for the formation of *wh*-clusters):

(59) and seofon ærendracan *he him hæfde* to asend
 and seven messenger he him had to send
 'and he had to send him seven messengers'
 (Parker, 905; Pintzuk 1999)

Now, turning to *þa*, *þonne*, recall that similar to pronouns, they are anaphoric elements carrying a [D] feature (i.e., temporal anaphora that specify temporal properties of their clause in relation to the immediate discourse context). If these temporal anaphora are to be interpreted as sentence adverbs (i.e., sequential 'then'), they are directly merged in SpecTP (the locus of Reference time) to receive a temporal index (i.e., they are linked with a Reference time given in the discourse context). Thus, they eliminate T's [*D*] feature via external Merge, blocking movement of subject pronouns to this position (*Merge over Move*, Chomsky 1995).⁴⁰ In cases where *þa*, *þonne* occupy SpecTP, the assignment of a referential index to a lower subject pronoun is accomplished either via a universally available semantic/pragmatic process (similar to non-discourse-configurational languages) or via an AGREE-relation between T and the pronoun (established during the syntactic derivation for independent reasons (Case and agreement)).⁴¹

Of course, this analysis raises a number of further questions, in particular concerning the analysis of embedded clauses, where subject pronouns generally occur to the left of *þa/þonne*, directly adjacent to the complementizer (cf. Mitchell 1985, Koopman 1997, Hauberli 1999):⁴²

⁴⁰ In the absence of subject pronouns, SpecTP may also be targeted by object pronouns, as can be seen from (47) above. Note that in (47), *þonne* most likely receives a cotemporal reading, which suggests that it is adjoined to vP/VP.

⁴¹ Note that this seems to be reminiscent of the relation between *there* and its associate DP in existential constructions. However, there are also major differences between these two constructions. For example, in contrast to *there*, *þa/þonne* are not expletives that lack semantic content. As a result, they may not be deleted (or substituted by the subject pronoun) at LF.

⁴² Furthermore, object pronouns may occur to the left of *þa/þonne*, either alone (if there is no pronominal subject present), or together with the subject pronoun:

(60) *þa hi þa hine geornlice **beheoldon...***
 when they then them carefully beheld...
 'when they then carefully beheld him...'
 (coeust, LS_8_[Eust]:270.286; Kemenade and Los 2006: 236f.)

One possible explanation of this fact might be that the typical rhetoric use of sequential 'then' in OE (marking of foregrounded successive actions/events, cf. Foster 1975, Enkvist and Wårvik 1987) is much less called for in embedded clauses, which are typically associated with backgrounded information. A related problem comes from the observation that contrary to what is expected under the above analysis, a subject pronoun always immediately follows the fronted verb in root *wh*-questions. In other words, the pattern *wh*-*V*_{fin}-*þa/þonne-pron.*_{subj} is apparently not attested in OE. Again, this might be due to independent reasons, for example a morphophonological requirement (or at least strong tendency) that the (weak) subject pronoun must be adjacent to the finite verb in C (similar constraints hold in many present-day Germanic V2-languages). Moreover, a closer look reveals that in many of the relevant examples, *þa/þonne* should rather be interpreted as instances of cotemporal 'then':

(i) *ðætte hie ðonne **gemonnðwærige** sio lufu & sio geferræden hiora niehstena [...]*
 that them then may-humanize the love and the society of-their neighbors
 'that love and the society of their neighbors may humanize them'
 (cocura,CP:47.363.15.2461; Kemenade and Los 2006: 236)

(ii) *gif he hit him ðonne **sellan mæge***
 if he it him then give may
 'if he can give it him then'
 (cocura,CP:44.323.24; Kemenade and Los 2006: 235)

The fact that all pronouns occur to the left of *þa/þonne* in (ii) can again be accounted for under the assumption that pronominal elements may form a cluster prior to movement to SpecTP. In addition, it is of course possible that some of the apparently problematic examples involve instances of cotemporal 'then', which occupies a lower, VP-joined position.

(61) and þonne gyt ne cymð se brydguma; eac swilce þa six ðusend and then still not comes the bridegroom; also the six thousand geara fram Adame beoð geendode and ðonne gyt elcað se brydguma. years from Adam is ended and then still delays the bridegroom.
 [Hu mage we þonne witan hwænne he cymð]?
 How may we then know when he comes?
 (ÆCHom_II,_44:330.117.7427-7430)

Accordingly, one might speculate that the order *wh*-V_{fin}-*þa*/*þonne*-pron._{subj} is not attested in the corpus for the following reasons: First, it would have been quite rare anyway, since it is confined to a very specific context (a *wh*-question concerning a foregrounded sequence of actions/events). In addition, for PF-reasons, subject pronouns are preferably adjacent to a fronted finite verb.⁴³

Summing up, in this section I have argued that in OE, V2 orders resulted from three different underlying structural configurations: First, a spec-head relationship

⁴³ Another more technical question is why T does apparently not allow multiple specifiers in OE, so that both pronouns and *þa*/*þonne* may occupy different specs of T. In this connection, note that the order pronoun-*þa*/*þonne*-V_{fin} is actually possible in OE, as illustrated in (i) (119 tokens in the York Corpus):

(i) *Hig þa forlættan* þone wall & heora burh,
 they then left the wall and their fort
 (Bede_1:9.46.20.406)

These examples are usually analyzed as involving a topicalized pronoun. However, one might also argue that (i) represents an instance where T projects two specifiers, with the temporal adverb in the inner and the pronoun in the outer spec. Furthermore, note that there also examples where the order or pronoun and *þa*/*þonne* is reversed (29 tokens in the York Corpus):

(ii) *Da hig wunedon* on Galilea;
 then they dwelled in Galilea
 (Mt_[WSCp]:17.22.1163)

The order exemplified in (ii) appears to be particularly problematic since it involves instances where fronted *þa*/*þonne* fail to trigger inversion. Again, one might attribute these examples to the possibility of T having multiple specs in OE. But then the question arises of why the subject pronoun follows the finite verb in the vast majority of relevant examples in OE. Presumably, it is more promising to stick to the topicalization analysis of (i) and analyze cases such as (ii) as reflexes of an earlier SOV option in main clauses (with the possibility of extraposing material, as e.g. the PP in (ii)). See Trips and Fuß (in prep.) for a more comprehensive view on these matters, including an in-depth study of the distribution and interpretation of pronouns and *þa*, *þonne* in both main and embedded clauses (see also chapter 2 for some discussion of adverb placement in embedded clauses of OE).

between a fronted operator and the finite verb (in C^0) could give rise to 'residual V2' effects, similar to present-day English. Second, superficial V2 orders could result from linear adjacency between a fronted non-operator XP in SpecCP and the finite verb located in T, with non-pronominal subject DPs remaining in their theta-position SpecvP (due to the absence of a general EPP feature in T). In addition, V2 orders could reflect a spec-head configuration between the temporal adverbs *þa* and *þonne* and the finite verb in T. I have presented evidence suggesting that SpecTP is linked to anaphoricity in OE (due to the discourse-configurational nature of OE), which is established in the syntax via a strong $[*D^*]$ feature that may optionally be added to the content of T. In the absence of *þa* and *þonne*, this feature is checked by pronominal elements that move to SpecTP, giving rise to V3 orders with fronted non-operators. Furthermore, I have speculated that the order $XP\text{-}DP_{\text{subj}}\text{-}V_{\text{fin}}$, which is quite frequent in OE, is actually a precursor of the structure of Modern English, with the full nominal subject moving to SpecTP. By assumption, this option is available only for specific subjects that relate to previously established discourse topic (again in connection with T's $[*D^*]$ feature).

The following section shows that the present approach to V2/V3 in OE receives further support from observations on the loss of V2 patterns in the ME period.

4 The 'loss' of V2 in the Middle English period

It is a well-known fact that English lost much of its V2 character during the Middle English (ME) period, giving rise to the limited V2 properties of Modern English. Upon closer inspection, however, it becomes clear that it is not appropriate to refer to this change in terms of a 'general loss of V2', since English (including OE) has never been a full V2 language in the first place (cf. e.g. Haeberli 1999, 2000 on this point). Rather, what has been lost were structures of the kind $XP\text{-}V\text{-}DP_{\text{subj}}$, that is inversion with non-pronominal DP subjects in clauses with a fronted non-operator. In contrast, the V2 properties of English apparently have not changed much in the context of pronominal subjects. Similar to OE, fronting of an XP leads to non-inversion structures in Modern English, while V2 order is triggered by fronted operators such as *wh*-phrases, or negation:

(62) a. *Last year bought John a house.
 b. Last year, John bought a house.

(63) a. Mary, he really likes.
 b. What did he buy?
 c. Never in his life would he do that.

In other words, what has been lost are instances of 'pseudo V2', where, according to the analysis proposed in Fuß (2003), the finite verb moves to T, with non-pronominal subject DPs remaining in their theta-position, SpecvP. The fact that the change in question affected only orders with non-pronominal subjects can be taken to indicate that there is a connection between the loss of this particular pattern and another change that affected the distribution of subjects in general and led to the development of overt expletives (cf. Hulk and van Kemenade 1995, Kemenade 1997, Haeberli 1999, Fuß 2003).

A first indication that these two changes are interrelated comes from the observation that there are chronological parallels between the loss of V2 patterns of the type XP-V-DP_{subj.} and a set of diachronic developments that led to an obligatorily filled subject position. According to Kemenade (1987), the loss of V2 took place at around 1400. In later work (Kemenade 1997), she somewhat modifies this statement, putting the relevant change in the period roughly from 1350 to 1425. Haeberli (1999) confirms this estimation by taking into account quantitative data from the Penn-Helsinki Parsed Corpus of Middle English.⁴⁴ Interestingly, it can be shown that the loss of XP-V-DP_{subj.} is paralleled by changes affecting the syntactic distribution of subjects.

In OE, we can observe a number of constructions that lack an overt subject. Relevant examples include weather verbs, experiencer verbs and impersonal passives:

(64) a. norþan sniwde
 [from] north snowed
 'it snowed from the north'
 (Seafarer, 31; Kiparsky 1997:471)

⁴⁴ Warner (1997) speculates that the change in question might have begun even over a century earlier as a long-term gradual change that involved competition of different grammars, giving rise to the multitude of word order patterns displayed by Old and Early Middle English.

b. him ofhreow þæs mannes
 him-DAT pitied the man-GEN
 'he pitied the man'
 (AColl, 192.16; Allen 1995:68)

c. þæt eallum folce sy gedemed beforan ðe
 that all people-DAT be judged before thee
 'that all the people be judged before you'
 (Paris Ps. 9.18; Kemenade 1997:335)

In Early Middle English, these constructions began to disappear, a development which is accompanied by the emergence of the expletive *there*. Van Kemenade (1997) shows that English lost subjectless structures at about the same time as it lost the pattern $XP\text{-}V\text{-}DP_{\text{subj}}$, that is 'in the last part of the fourteenth and the early fifteenth century' (p. 349), referring to work by van der Wurff (1990) on the loss of subjectless *easy to please* constructions (cf. also Fischer et al. 2000), Butler (1980) on the rise of *there* in existentials, and Allen (1995) on the loss of subjectless impersonal constructions with experiencer verbs and passives. In a similar vein, Haeberli (1999: 403ff.), argues for systematic parallels between the erosion of the V2 system and the decline of subjectless constructions, citing quantitative data from Breivik (1989), (1990), where it is shown that there is a rapid increase of the use of *there* in existential sentences in the time between 1225 and 1425 (from around 30% to over 80%). Both van Kemenade (1997) and Haeberli (1999) (cf. also Hulk and van Kemenade 1995; Haeberli 2002) then go on to argue that the close connection between the loss of V2 and the loss of subjectless structures motivates an analysis that reduces the apparently independent changes to a single diachronic development, namely the loss of expletive *pro* due to the erosion of verbal inflection. More specifically, Haeberli (1999), (2002) attributes the decline of superficial V2 orders to the loss of an empty expletive *pro* that is inserted as specifier of AgrsP in order to satisfy the EPP, thereby blocking overt movement of non-pronominal subjects to this position in OE. He derives the latter change from an independent morphological change that significantly impoverished the verbal agreement morphology in the ME period.⁴⁵ As a result, expletive *pro* could no longer be licensed by the verbal agreement morphology and dropped out of the grammar. Accordingly, the only remaining

⁴⁵ Due to the loss of the infinitival ending *-n*, the singular verbal agreement endings were no longer distinct from the infinitive.

possibility to satisfy the EPP was overt movement of the subject to SpecAgrsP across the finite verb that is by assumption located in Agrs. Obligatory subject movement to SpecAgrsP disrupts linear adjacency of finite verb and a fronted XP in SpecCP and therefore leads to the loss of V2 orders in connection with non-pronominal DP subjects.⁴⁶

Note, however, that this account raises the question of why structures similar to those found in OE are not possible with an overt expletive in ME prior to the loss of verb movement to Agrs. As pointed out in Fuß (2003: 218), examples of the kind schematically given in (65) are not attested in any historical stage of English.⁴⁷

(65) *[_{CP} The book [_{AgrsP} there [_{Agrs'} read [_{TP} [_{VP} the student]]]]]

In earlier work (Fuß 2003), I proposed an alternative approach of the loss of XP-V-DP_{subj.} patterns. My analysis shares basic insights with the accounts suggested by van Kemenade (1997) and Haeberli (1999), (2002), but crucially differs from Kemenade's and Haeberli's proposals in not assuming a universal EPP (which forces the assumption of empty expletives in Kemenade's and Haeberli's analyses). Rather, it is assumed that the presence of an EPP feature in T represents a language-specific parametric choice (see Roberts 2007 for some discussion of the parametrization of EPP features). Accordingly, I assume that the subjectless constructions in (64) do not involve a zero expletive pronoun. Rather, the absence of an expletive element filling the subject position is taken to follow from the absence of an EPP feature in T.

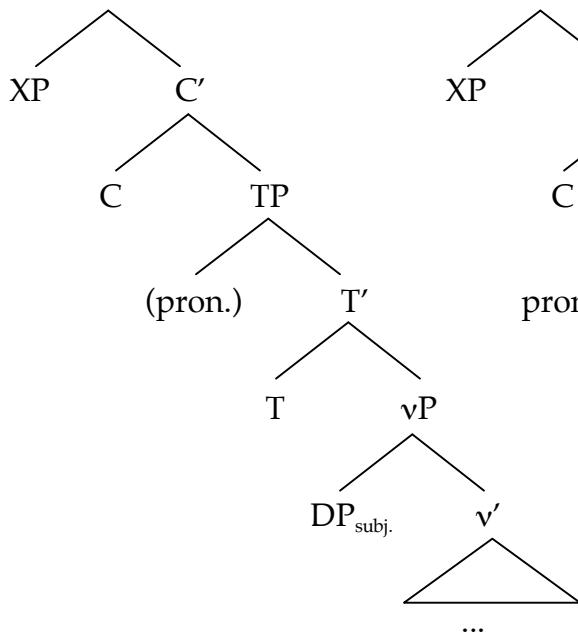
Under this assumption, the change in question can be attributed to the development of an EPP feature in T that requires the subject position (here identified as SpecTP) to be overtly filled – either by a nominal bearing nominative case or a semantically vacuous expletive element such as *there*. Similar to the proposals of van

⁴⁶ Alternative scenarios for the 'loss' of V2 in ME proposed in the literature include Kemenade (1987), Kroch and Taylor (1997), and Lightfoot (1991, 1997). Kemenade (1987) links the obsolescence of (certain) V2 patterns to the loss of clitic subject pronouns (as a result, V3 orders formerly derived by special clitic placement rules were reinterpreted as violations of V2). Kroch and Taylor (1997) attribute the loss of XP-V-DP_{subj.} orders to dialect mixture/contact between northern and southern dialects (cf. chapter 1 for details). According to Lightfoot (1991, 1997), stylistic changes affected the make-up of the triggering experience, so that the ratio of subject-initial clauses crossed a threshold. As a result, the V2 option failed to be triggered, leading to a general reanalysis of subject-initial V2 clauses as IPs.

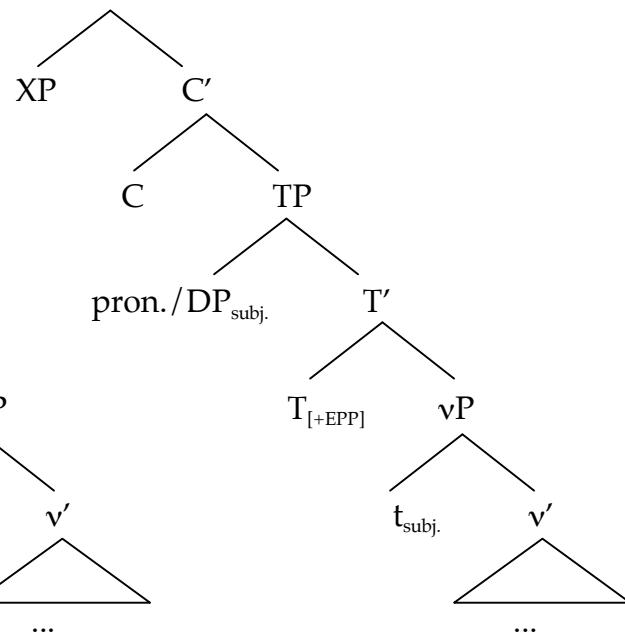
⁴⁷ But see Haeberli (1999:406) for some discussion of this problem.

Kemenade and Haeberli, the resulting obligatory overt realization of SpecTP disrupts the linear adjacency of clause-initial topics and the finite verb in T which led to the loss of 'pseudo V2' configurations and gave rise to the familiar V3 topic-constructions of present day English. The development in question is illustrated by the following pair of phrase markers:⁴⁸

(66) a. OE: CP



b. ME / ModE: CP



According to this view, the partial loss of V2 did not involve changes concerning the structural positions of topics (SpecCP) and the finite verb (T^0). Rather, superficial V2 patterns disappeared because SpecTP came to be obligatorily filled by all kinds of subjects (accompanied by the development of overt expletives). Under these assumptions, the 'loss of V2' in the Middle English period appears to be an epiphenomenon, resulting from the independently motivated development of an EPP-feature.⁴⁹

⁴⁸ Note that phrase marker (66b) does not capture the fact that main verbs occupy different structural positions in ME and ModE. In ME, all finite verbs (auxiliaries as well as main verbs) move to T. In contrast, in ModE this position is only accessible for finite auxiliaries/modals, whereas finite (at least transitive and unergative) main verbs only undergo short movement to v (cf. Larson 1988, Chomsky 1995, Collins 1997, Roberts 1998).

⁴⁹ Sten Vikner pointed out to me that Mainland Scandinavian seems to represent a problem for this proposal: The presence of Infl-related expletives indicates the existence of an EPP-feature in these languages. Nevertheless, Mainland Scandinavian has not lost regular V2 in main clauses. However, this situation might be the result of a different chronological order of the relevant historical processes,

This analysis receives further support from observations concerning another major difference between OE and later stages of English, namely the loss of *þa*, *bonne* + inversion, which took place in the very same period (cf. Fuß and Trips 2003, Trips and Fuß 2008). A survey over *The Penn-Helsinki Parsed Corpus of Middle English* (second edition, PPCME2) shows that 'then' loses its special status as a trigger of V2 in the period from 1340-1475 (Fuß and Trips 2003). The correlation between the loss of 'then'+V2 and development of the expletive *there* becomes particularly clear in the *Ayenbite of Inwit* (1340), a text which exhibits variation between inverted and non-inverted orders after clause-initial *þanne* 'then'.⁵⁰ However, all examples with the expletive *þer* 'there' display V3 order, with the expletive intervening between *þanne* and the finite verb:

(67) a. þanne *þer* **nys** prowesse ariȝt: [...]
 then there not-is prowess properly
 'Then there is no proper prowess [...]'
 (AYENBI,83.1613)

b. þanne *þer* **ne** **is** non noblesse: [...]
 then there not is no nobleness
 'Then there is no nobleness [...]'
 (AYENBI,87.1702)

The systematic absence of V2 orders in clauses in which *þanne* and *þer* co-occur supports the conjecture that there is a close connection between the loss of 'then'+V2 and the rise of an EPP feature in T: in cases where an expletive is inserted in SpecTP to satisfy T's EPP feature, the adverb *þanne* must occupy another position (e.g., in the CP domain, or adjoined to TP).⁵¹ Over time, V2 patterns with 'then' dropped out of the grammar, since SpecTP became a position reserved for subjects/expletives, which could no longer host adverbs:

with the development of full V2 preceding the development of an EPP-feature. This hypothesis is confirmed by the historical facts: Falk (1993) shows that Old Swedish (1225-1526) is a full V2 language that has not yet developed obligatory overt (expletive) subjects.

⁵⁰ The *Ayenbite of Inwit* exhibits 70% inversion with subject pronouns (16 of 23 cases), and 44% inversion with full subject DPs (14 of 32 cases), probably an instance of *Grammar Competition* (Kroch 1989).

⁵¹ In a similar vein, Alexiadou (2000) assumes that SpecTP can host temporal adverbs only if there is no EPP feature in T. However, in languages where such a feature requires subjects to appear in SpecTP, temporal adverbs cannot occur in this position.

(68) $[_{CP} \emptyset [_{TP} 'then' [_{TP} \text{expl. / subj.} [_{T'} V_{\text{fin}} [_{VP} \dots]]]]]$

In this way, both the loss of $XP\text{-}V\text{-}DP_{\text{subj.}}$ patterns and the loss of V2+inversion in the context of clause-initial *þa/bonne* can be attributed to a single independent change – the development of an generalized EPP feature in T. However, as attractive as it may be, this approach ultimately raises another question, namely why and how the latter change came about. This question is addressed in the following section, which shows that there are reasons to believe that the rise of an EPP feature in T was linked to the loss of discourse-configurationality in English.

4.1 The loss of discourse-configurationality and the rise of 'fossilized' movement

This section argues that the loss of certain V2 patterns can be attributed due to the rise of an EPP feature in T. This change was part of a more general development in which English turned into a configurational language in which word order (and structural positions) primarily expresses grammatical functions instead of discourse-related properties such as information-structure and anaphoricity. In what follows, I am going to discuss some details of the overall historical development and the way this change affected word order properties of English.

It is a well-known fact that the grammar of English underwent a major reorganization during the ME period. Major changes affecting syntactic properties of the language included the loss of inversion patterns (i.e., $*XP\text{-}V\text{-}DP_{\text{subj.}}$, see above), the loss of subjectless constructions (accompanied by the rise of expletives), and the development of basic (fixed) SVO order (cf. chapter 2 above). Furthermore, English innovated constructions characteristic of configurational languages such as a structural passive and new ECM constructions. These syntactic changes were accompanied by a wholesale loss of both verbal and nominal inflections (cf. Mitchell 1985, Lightfoot 1979, 1991, Kemenade 1987, Pintzuk 1999, Haeberli 1999, Fischer et al. 2000, among many others). All in all, these changes can be described in terms of a general change from a discourse-configurational language to a configurational language.

Traditionally, the loss of inflections (due to independent processes such as phonological erosion and imperfect language transmission due to intense language

contact with Scandinavian invaders, cf. Taylor 1994, Kroch 1994, 2001, Trips 2002) is seen as the causal factor behind this set of changes (cf. Sapir 1921, Lightfoot 1979, 1991, Kemenade 1987, Kiparsky 1997, among many others). More precisely, it is assumed that the loss of inflectional morphology (in particular, case) required arguments to occur in fixed structural positions ('positional licensing' in the sense of Kiparsky 1997).⁵² According to Kiparsky (1997), the loss of case and verbal agreement morphology led to a situation where subjects could only be licensed in a spec-head relationship with their case-assigning head, T.

While this approach can account for the diachronic facts of English (and a number of other languages, in particular Mainland Scandinavian), it seems to run into some problems if we consider a language such as Icelandic, which exhibits a rich system of verbal and nominal inflections, but crucially lacks free reordering of nominal arguments (in contrast to e.g. Modern German, or OE). In other words, Icelandic apparently has to rely on positional licensing of arguments despite the fact that its case system is rich enough to unambiguously identify the grammatical relations taken up by nominal arguments (cf. Thráinsson 1997). So, in other words, while it is certainly true that the loss of case distinctions is one of the factors that contributed to the loss of free, discourse-driven word order in the history of English, there is apparently no general one-to-one correlation between inflectional morphology and the position of nominal arguments. Next, I explore the rise of the requirement that SpecTP be filled in some more detail, arguing that there is (at least) one other factor that promoted the development of a generalized EPP feature in T.

As already noted above in section 3.1.2, there is one frequently attested word order option in OE that seems to be at odds with the analysis developed in section 3.3., namely V3 orders where the finite verb is preceded by a scene-setting temporal adverb and a full nominal subject, repeated here for convenience:

(69) a. [Æfter þeosum wordum] [se Hælend] **cwæp** to his leornerum...
 after these words the Savior spoke to his disciples
 (Blickling 135; Swan 1994: 241)

⁵² As Sapir (1921: 166) puts it: "As the inflected forms of English became scantier, as the syntactic relations were more and more inadequately expressed by the forms of the words themselves, position in the sentence gradually took over functions originally foreign to it."

b. [Her] [Oswald se eadiga arceb] **forlet** þis lif.
 in-this-year Oswald the blessed archbishop forsook this life
 (ASC, Laud (992); Kroch and Taylor 1997: 304)

As already noted above, surface orders like (69) were quite frequent in OE (28.7% of all clauses with non-pronominal DP subjects according to Haeberli 2000: 4). In other words, they were robustly attested in the input learners of OE/Early Middle English received. Above, I have argued that examples like (69) should be analyzed as involving a configuration in which the DP subject occupies SpecTP, as an early precursor of the Modern English structure. More precisely, I have assumed that topical DP subjects, which relate to a discourse referent established earlier, may move to SpecTP to check T's [*D*] feature. Due to the fact that even if topical, full nominal DPs are not purely anaphoric elements, we might speculate that examples like (69) blurred the original discourse-related function of SpecTP as a structural position linked to anaphoricity. Still, language learners had to account for the very fact that the relevant position is filled in a large portion of the examples they were confronted with (e.g., in all clauses with a pronominal subject), even if they could not detect a clear semantic or pragmatic trigger for that operation. In other words, learners had a choice between (i) developing a grammar that produces an output that differs significantly from the output of the target grammar, or (ii) imitating the patterns they encounter by different syntactic means. Following Simpson (2004), I assume that in this situation, EPP features are available for the language learner as a formal means to cope with dislocation phenomena for which no clear 'substantial' trigger (i.e. semantic, pragmatic or morphological) can be detected, but which are still robustly attested in the input. As a consequence, movement operations are not lost from a structure if the original trigger disappears, but rather are converted into 'fossilized' movement triggered by EPP features.⁵³ As a result, SpecTP ceased to be

⁵³ Simpson (2004) argues convincingly that certain movement operations that apparently have no clear motivation in a certain synchronic stage of the grammar arise historically via a reanalysis of formerly pragmatically/semantically motivated operations (Focus etc.) as EPP-driven movement. This approach provides a diachronic explanation for the rise and existence of EPP-driven movement in general. Furthermore, it leads to two interesting predictions: first, historically, there is no such thing as 'negative reversal-type changes' that is, loss/discontinuation of movement. Second, synchronically, the apparent lack of any understandable motivation for an assumed movement operation "should *not* necessarily cause one to doubt the hypothesis that movement does indeed take place" (Simpson 2004: 186).

associated with anaphoricity in the learner's grammar, turning into a generalized structural position for subjects. This change was possibly promoted by the set of other changes (loss of morphological case etc.) that hindered the acquisition of properties linked to discourse-configurationality.⁵⁴

Summing up, this section has argued that the loss of V2 patterns (i) with inverted non-pronominal subject DPs, and (ii) in the context of clause-initial *þa/bonne* can be attributed to a single underlying change, namely the rise of a generalized EPP feature in T during the ME period. Moreover, I have suggested that learners resorted to positing a semantically vacuous EPP feature in T to mimic word orders attested robustly in the input when the original semantic/pragmatic motivation for movement to SpecTP (licensing of anaphoric elements) became unclear. The development of a structural subject position was part of a large-scale change in

⁵⁴ Fuß (2003) considers the possibility that the rise of an EPP feature in T was linked to the independent development of an elaborate and rigid system of tenses, with a clear functional differentiation of, e.g. past and perfect tense, which is a characteristic of Modern English (in OE, there are only two tenses, past and non-past; cf. Denison 1993 for a comprehensive overview of the diachronic developments in question). Interestingly, there are some hints that the latter development took place at about the same time as the changes discussed above (i.e. the loss of V2 and subjectless constructions). For example, Bauer (1970) claims (in his study of the use of different tenses in the works of Chaucer and Gower) that the differentiation of past and perfect tense was completed by and large by the end of the 14th century. Another indication that these suggestions are perhaps on the right track comes from the fact that in a language such as German, the absence of EPP effects (cf. Haider 1993; Roberts and Roussou 2001) goes hand in hand with a – compared to English – much less rigid tense system, where past and perfect tense are freely interchangeable and the present tense can assume all kinds of temporal functions (cf. Zeller 1994, Grewendorf 1995). Possibly, the rise of a rigid tense system can also be linked to the loss of discourse-configurationality, in the sense that a rigid tense system provides an alternative means to establish discourse/text coherence. While Modern English uses a system of discourse/text coherence based on tense, discourse-configurational languages typically make use of topic continuity to establish coherence (cf. e.g. Hopper 1979, McGinn 1985). This is supported by the observation that discourse-configurational languages such as OE, Hungarian or Russian often exhibit an impoverished system of tense distinctions (typically only a past/non-past distinction), while languages such as English are characterized by a rich (and rigid) system of different tense paradigms. Note that this approach would also capture the behavior of Icelandic, which not only has a fixed subject position, but also an elaborate system of different tenses (present, past, perfect, past perfect, future, and future perfect). To be sure, however, more research will be necessary to strengthen the speculation that there is a systematic link between the presence of an EPP feature in T and properties of the tense system, which should be based on a more detailed study of the diachronic changes affecting the English tense system as well as a larger sample of languages.

which English turned from a discourse-configurational language into a language where word order primarily serves to encode grammatical functions.

The next section examines word order patterns in main clauses of Old High German. It is shown that Old High German differs significantly from the other early Germanic languages. In particular, it will become clear that the V2 properties of Old High German resemble more closely the kind of V2 we encounter in the Modern Germanic languages.

5 V2 in Old High German

5.1 Introduction

This section explores the nature of V2 in early Old High German (OHG), focusing on the *Isidor* and *Tatian* translations in particular (around 800 and 850, respectively). It appears that already at this early stage, we can observe a close-to consistent V2 syntax in main clauses (cf. e.g. Lippert 1974, Robinson 1997 on the OHG *Isidor* translation, Dittmer and Dittmer 1998 on the *Tatian* translation; relevant generative studies include Lenerz 1984, 1985a, Tomaselli 1995; see Axel 2007 for a detailed overview and in-depth discussion).⁵⁵ Similar to OE, a cursory inspection of early OHG reveals a large number of V2 patterns, compare the following examples in which the preverbal position is occupied by a subject (70a), an object (70b), a PP (70c), an adverb (70d), a predicative adjective (70e), and a *zi*-infinitive (70f) (taken from Axel 2007: 4f.):⁵⁶

(70) a. [Druhtin] **suuor** dauite in uuaarnissu
 Lord swore David-DAT in truth
Lt. Iurauit dominus dauid in ueritate
 ‘The Lord swore to David in truth.’
(Isidor, 610)

⁵⁵ This section draws heavily on Axel (2007), the first comprehensive generative study of OHG syntax since Lenerz (1984).

⁵⁶ The line numbers given for the OHG *Isidor* refer to the edition by Eggers (1964). The *Tatian* is cited according to page and line numbers in Masser’s (1994) edition. Numbers in brackets refer to the *Tatian* edition by Sievers (1961).

b. Enti [miin ur teili] **chundit** deotom
 and my judgment declare-3SG nations-DAT
 Lt. et iudicium gentibus nuntiabit
 'And he shall declare my judgment to the nations.'
 (*Monsee Fragments*, V,8; Mt. 12.18)

c. [In dhemu nemin cyres] **ist** *christ* chiuuisso chiforabodot
 in the name Cyres is Christ certainly presaged
 Lt. *In persona enim cyri christus est prophetatus*
 'By the use of the name 'Cyres' Christ is certainly presaged.'
 (*Isidor*, 162)

d. [Chiuuisso] **chioffanodom** *uuir* nu hear dhazs [...]
 certainly revealed we now here that
 Lt. *Probauimus dominum nostrum* [...]
 'Certainly we have now revealed that [...]'
 (*Isidor*, 484)

e. [toot] **ist** *her*
 dead is he
 Lt. *quia mortuus est*
 'He is dead.'
 (*Tatian*, 313,14)

f. [Zi uuizsanne] **ist** nu uns chiuuisso, dhazs fater einemu ist
 to know is now us certainly that father alone is
 dhurahchunt [...] well-known
 Lt. *Scire autem manifestum est solum patrem*
 'We should certainly know now that only the father really knows [...]'
 (*Isidor*, 120)

As is well-known, the study of syntactic properties of OHG is hindered by the fact that, similar to Gothic, but in contrast to OE, the vast majority of early records are translations, mostly of Latin religious texts. As a result, the word order properties of early OHG texts may be influenced by the syntax of the (mostly Latin) original, depending on the quality of the translation.⁵⁷ In order to distinguish between OHG

⁵⁷ Traditional studies have usually praised the quality and free character of the OHG *Isidor* translation, while the *Tatian* translation has been deemed to be of minor quality, in the sense that its word order

patterns affected by properties of the (mostly, Latin) source text and genuine syntactic properties of OHG, we should focus primarily on (i) cases where the word order of the OHG translation (systematically) deviates from the word order of the (presumed) source text, and (ii) evidence involving potential changes that are never carried out in the translation (e.g., a change from V2 in the original to V3 order in the OHG translation), cf. Lippert (1974), Dittmer and Dittmer (1998: 22). For example, the cases of V2 listed in (70) can be considered to reflect a genuine trait of OHG syntax since their word order properties differ from the order found in the Latin source.

In this section, it will become clear that OHG differed significantly from OE in that it already exhibited the early beginnings of a system of generalized V2, that is, (close-to-) regular XP fronting + verb movement into the C-domain similar to the Modern Germanic V2 languages. In addition, I will discuss a set of deviations from V2 that suggest that canonical V2 order had not yet been fully grammaticalized in early OHG.

5.2 Generalized V2?

According to Axel (2007), movement of the finite verb to some head position in the left clausal periphery was already very systematic in early OHG. Similar to OE and Gothic, operator contexts such as interrogatives (both *wh*-question and *yes/no*-questions), imperatives and negated clauses trigger consistent verb fronting, often against the Latin source:

(71) a. Odho **mahti** angil so sama so got mannan chifruman?
 or could angels so same as God man-ACC make
 Lt. *Aut numquid angelus cum deo potuit facere hominem?*
 'Or could an angel make Man the same as God does?'
 (Isidor, 187; Axel 2007: 53)

has been thought to be heavily influenced by properties of the Latin source. As a result, traditional studies of OHG syntax have tended to ignore the *Tatian*. However, in a detailed quantitative study of the *Tatian*, Dittmer and Dittmer (1998) argue convincingly that a study of the *Tatian* can give important insights into the grammar of OHG if we focus on cases where the translators deviate from the Latin source text. In particular, Dittmer and Dittmer show that there are systematic deviations from the word order of the Latin original that can be taken to reflect genuine traits of OHG. See below for some discussion.

b. **bist** thu uuîzago. [...]
 are you prophet
 Lt. *proheta es tú.* [...]
 'Are you the Prophet?'
 (Tatian, 109,14; Axel 2007: 53)

(72) a. [bihuuiu] **uuard** christ in lihi chiboran?
 why became Christ in flesh born
 Lt. *car in carne uenit?*
 'Why was Christ born in the flesh?'
 (Isidor 487; Axel 2007: 55)

b. meistar, [uuanne] **quami** thû héra
 master when came you here
 Lt. *Rabbi, quando hu uenisti.*
 'Master, when did you come here?'
 (Tatian 257, 12; Axel 2007: 55)

(73) a. **tuot** riuua [...]
 do-2PL.IMP repentance
 Lt. *pænitentiam agite* [...]
 'Repent!'
 (Tatian, 103,1; Axel 2007: 56)

b. **Chihori** dhu, israhel [...]
 listen-2SG.IMP you Israel
 Lt. *Audi, israel* [...]
 'Listen, Israel [...]'
 (Isidor, 371; Axel 2007: 57)

(74) a. **nisanta** got sinan sun
 NEG-sent God his son
 Lt. *non enim misit deus filium suum*
 'God did not send his Son.'
 (Tatian, 407, 30; Axel 2007: 61)

b. ni **liugu** *ih* dauid
 neg lie I David
 Lt. si dauid mentiar
 'I will not lie to David.'
 (Isidor, 612; Axel 2007: 62)

The findings on main interrogatives are supported by a quantitative study on verb placement in OHG interrogative clauses carried out by Petrova and Solf (2007). They show that in the *Tatian*, inversion takes place in 207 of 230 *wh*-questions (including five cases where another XP precedes the *wh*-word). In the *Isidor*, inversion takes place in 20 of 23 *wh*-questions (including two cases of XP-*wh*-V_{fin}-subj order).⁵⁸ Interestingly, Petrova and Solf show that inversion is slightly less regular in yes/no-questions (e.g., in 19 of 107 cases in the *Tatian*, the verb fails to undergo inversion with the subject). Furthermore, the *Tatian* exhibits 43 cases where an XP precedes the inverted verb in yes/no-questions. In general, it seems that the small number of examples without inversion in interrogatives is confined to early OHG, while inversion occurs without exceptions in late OHG texts (e.g., Näf 1979: 161 counts 113 examples of *wh*-questions in the first two volumes of Notker's *Consolatio*, all of them exhibiting the finite verb in second position).

Interestingly, it can be shown that early OHG differs significantly from OE and Gothic, since it exhibits systematic verb fronting in contexts where the finite verb regularly occupies a lower position in Gothic and OE. In contrast to OE, subject pronouns often undergo inversion with the finite verb, similar to full nominal subjects (see Axel 2007: ch. 5 and section 5.4 below for a closer view on pronoun placement in OHG):

(75) a. [Dhinera uuomba uuwaxsmin] **setzu** *ih* ubar min hohsetli
 your womb's fruit place I upon my throne
 Lt. *De fructu uentris tui ponam super sedem meam*
 'I will place the fruit of your womb upon my throne.'
 (Isidor, 611; Robinson 1997: 9)

⁵⁸ See section 5.4.1 for some discussion of the deviations from V2 in *wh*-question in the *Isidor*.

b. tho **ersteig** *her úf zi themo itmalen dage*
 then climbed he up to the feast day
 Lt. *tunc & ipse ascendit ad diem festum*
 'then he went also up unto the feast'
 (*Tatian*, 347,12f.)

This suggests that the finite verb moves to a higher functional head position in early OHG, crossing the position of the subject pronoun at the left edge of IP/TP.

Perhaps the most obvious difference between OE and Gothic on the one hand and OHG on the other is that in OHG, verb fronting is not confined to operator contexts, but also applies quite regularly in topicalization contexts (see (70) above; cf. Eyþórsson 1995, 1996 for discussion). Another difference concerns the interaction between verb movement and the insertion of C-related particles. Recall that Gothic exhibits a rich inventory of C-related particles. In the literature, it is often claimed that in Gothic, we can sometimes observe a complementary distribution of verb movement and C-particles and that these particles assumed functions (e.g., lexicalization of heads in the C-domain for reasons of clausal typing) that were later taken up by verb movement when the system of C-particles eroded (cf. Hirt 1934, Hopper 1975, Roberts 1996, Ferraresi 1997).⁵⁹ Interestingly, early OHG had preserved

⁵⁹ Other factors that might have contributed to the development of generalized V-to-C movement include the development of hypotaxis (together with a new class of declarative complementizers, Kiparsky 1995), and the 'syntactification' of verb placement, which was originally determined by prosodic/metrical factors. The latter account dates back to Wackernagel (1892), who argues that the origin of V2 can be traced back to the placement of lightweight, clitic elements in early Indo-European, where we can observe that pronouns and auxiliaries preferably occupy second position in main clauses. (cf. Anderson 1993 for a review of Wackernagel's original insights). More recently, the prosodic explanation of the rise of verb movement has been revived in Dewey (2007). Based on evidence from metre in the Old Icelandic *Edda* and the Old Saxon *Heliand* and intonation marking in the Gothic *Skeireins* (which she takes to reflect properties of spoken language), Dewey argues that verb placement was originally governed by intonational regularities (i.e., Kuhn's Law, 1933) that forced unstressed elements (in particular finite auxiliaries in main clauses) to occur in a position after the first stressed element in main clauses in Germanic alliterative verse (due to the strict intonational pattern s w s w, i.e., alternating strength beginning with a stressed element, cf. Sievers 1893, Árnason 1985, 2002). In contrast, finite lexical verbs were often stressed. As a result, they were placed in clause-final position, which coincided with a prosodic lift in verse/metre. An interesting argument in favor of this hypothesis comes from examples where a fronted finite verb appears after the first word of a fronted XP, giving rise to discontinuous constituency (see also Hopper 1975), compare the

some residues of the earlier Germanic system of C-related particles, namely the interrogative particle *inu/eno* (see Grimm 1890 for the etymology of *inu/eno*), and the affirmative particle *ja/ja*.⁶⁰ However, as shown by Axel (2007: 43ff.) (basically following earlier work by e.g. Erdmann 1874, Behaghel 1928, 1932), in contrast to

following examples from Old Icelandic and Old Saxon (with stress-bearing elements marked by boldface):

- (i) a. **hveriom** ertu **sveini** um **borinn**
which.DAT are-you guy.DAT PARTICLE born
'To what guy have you been born?'
(*Fáfnismál* 1; Dewey 2007: 86)
- b. **atgorðom** kom hann **Gjúka**
to-gardens.DAT came he Gjuki.GEN
'He came to Gjuki's gardens.'
(*Atlaqviða* 1; Dewey 2007: 84)
- (ii) **Maneg** uundrode **Iudeo** **liudio**
many.NOM wondered Jewish.GEN people.GEN
'Many of the Jewish people wondered.'
(*Helian 4109a-4110b*; Dewey 2007: 65)

Note that this behavior in which a finite verb attaches to the first word of the clause, cutting apart the fronted XP, is quite untypical for syntactic movement, but reminiscent of PF-driven placement of clitics (as e.g. in Serbo-Croatian, cf. Schütze 1994). Dewey argues that independent prosodic changes that affected metrical properties of early Germanic led to "the grammaticalization of V2 as a syntactic rather than intonational phenomenon" (p. 56; for related ideas, cf. Stockwell and Minkova 1994). Again, the syntacticization of verb fronting can perhaps be attributed to an acquisition strategy that helps the learner to deal with dislocation phenomena for which he/she cannot identify an independent (in the case at hand, phonological/prosodical) trigger (see section 4.1 above). Thus, we may suppose that at some point, learners could no longer recognize the original metrical trigger of verb fronting. Still, they were under pressure to generate the patterns robustly attested in the input. Following basic insights by Simpson (2004), we may assume that in this situation learners can resort to postulating semantically vacuous syntactic features to generate the input (in the case at hand, a feature attracting the finite verb to C). This strategy was presumably promoted by the fact that there already existed several contexts where systematic verb fronting was triggered by syntactic factors (imperatives, questions etc.).

⁶⁰ The system of C-related particles further eroded during the OHG period. In late OHG prose texts, interrogative particles are already very rare. The erosion of the particle system was accompanied by a development in which V-to-C movement became exceptionless in all main clauses.

Gothic, these C-particles systematically co-occur with verb fronting to the C-domain, compare (76) and (77).⁶¹

(76) a. Inu ni **larut** *ir* huuaz dauid teta
 INU NEG read you.PL what David did
 Lt. *Non legistis quid fecerit dauid*
 ‘Have you not read what David did?’
 (*Monsee Fragments*, IV,6; Mt. 12.3; Axel 2007: 44)

b. eno **habet** *ir* uuaz muoses.
 ENO have you.PL anything food-GEN
 Lt. *numquid pulmentarium habetis*
 (*Tatian* 685, 9; Axel 2007: 43)

(77) ia **ist** sín *muoter* ginemit maria
 IA is his mother named Maria
 Lt. [...] *nonne mater eius dicitur maria* [...]
 ‘His mother is called Mary, right?’
 (*Tatian*, 243,6; Axel 2007: 47)

Again, this suggests that German had already innovated systematic V-to-C movement by the time these early OHG texts were composed. In contrast, it appears that the second ingredient of V2, namely fronting of a single XP to the *Vorfeld* ‘prefield’ had not yet been fully generalized in early OHG. As will be shown in section 5.3, violations of V2 include V1-declaratives and V3 orders with two XPs to the left of the finite verb in main clauses.

Still, Axel (2007) argues that early OHG already exhibits some form of generalized XP-fronting triggered by a semantically vacuous EPP feature in C that attracted the hierarchically highest element in the IP (called “stylistic fronting” in

⁶¹ Note that the negation *ni* was proclitic on the verb and accompanied V-to-C movement in OHG. In addition to yes/no-questions, *inu* could also occur in *wh*-questions in early texts (albeit less frequently):

(i) Inu huu{e}nan **meinit** *ir* daz ih sii
 INU who-ACC think you.PL that I am
 Lt. {Uos autem quem} me esse dicitis
 ‘Who do you think that I am?’
 (*Monsee Fragments* XXXVIII, 1; *St. Augustini sermo*; Axel 2007: 43)

Fanselow 2004 and “formal movement” in Frey 2006). Relevant evidence comes from cases where XP-fronting (+V2) seemingly cannot be attributed to pragmatic/semantic reasons (i.e., topicalization). In the following examples, the clause-initial position is occupied by fronted indefinites and adjuncts. It is usually assumed that these elements cannot act as topics (Axel 2007: 120):⁶²

(78) a. [Neoman] **niuuirdit** fona gote festi [...]
 nobody NEG-becomes by God strengthened
 Lt. *Nemo erit a deo nisi firmus* [...]
 ‘Nobody will become strengthened by God [...]’
 (*Monsee Fragments*, XL,19; *St. Augustini sermo*)

b. [Neo] **nist** zi chilaubanne dhazs fona dhemu salomone
 never NEG-is to believe that of the Salomon
 sii dhiz chiforabodot
 is this prophesied
 Lt. *Numquid de illo salomone creditur prophetatum? minime*
 ‘It can never be believed that this was prophesied by Salomon.’
 (*Isidor*, 638)

c. endi [chiuuiocco] **ist** *christus* in dheru selbun salbidhu chimeinit
 and certainly is Christ in that same salve meant
 Lt. *et utique christus ipsa unctione monstratur*
 ‘And certainly is Christ meant in that same salve.’
 (*Isidor*, 144)

Interestingly, it appears that apart from operator contexts, XP-fronting was confined to referential topics (without verb movement) in earlier stages of Germanic (i.e., Gothic). As pointed out by Axel (2007: 198ff.), this can be taken to indicate that XP-fronting was originally triggered by semantic/pragmatic factors (focus/operator properties and topicalization). In contrast, early OHG already exhibits generalized V2 patterns on a par with Modern German, where the clause-initial position may host all kinds of elements including non-topics. This innovation presumably took place at some point that predates the earliest records of German (see section 5.6.2

⁶² But note that at least (78a,b) can possibly be analyzed as involving fronting of operators. Thus, it is not entirely clear whether it is correct that in these examples fronting is “semantically/ pragmatically vacuous”, as claimed by Axel (2007: 120).

below for some speculations on possible triggers of this change). However, based on the observation that early OHG still allowed a number of violations of V2 (see sections 5.3 and 5.4 below), Axel (2007: 201f.) suggests the following structure for the left periphery for early OHG (with C possibly split up into more than a single head):

(79) $[_{CP} XP_i [_C' C [_{FinP} \dots [_{Fin'} V_j + Fin [\dots t_i \dots t_j]]]]$

Crucially, the fronted XP and the finite verb do not enter into a spec-head relation in (79). That opens up the possibility that further material intervenes between fronted XP and finite verb, giving rise to V3 orders (see section 3.3 above for a related analysis of OE). Axel then claims that in the course of OHG, the split-CP was conflated into a single projection, leading to the loss of V3 orders in late OHG:

(80) $[_{FinP} XP_i [_{Fin'} V_j + Fin [\dots t_i \dots t_j]]]$

Further evidence that some form of the V2 constraint was already at work in early OHG comes from a detailed investigation of the OHG *Tatian* carried out by Dittmer and Dittmer (1998). Dittmer and Dittmer demonstrate that in almost all cases in which the OHG translation systematically deviates from the word order of the Latin original (so-called version “G”), the relevant reorderings or additions led to patterns that are reminiscent of the V2 syntax of present-day German. This suggests that early stages of OHG were already characterized by the beginnings of a generalized V2 syntax, compare the following quote taken from Dittmer and Dittmer (1998: 18):⁶³

“1. Jede topologische Abweichung des Tatian von G erreicht eine genuine deutsche Wortstellung oder nähert sich einer solchen. Der Tatian weicht nie in nichtdeutscher Richtung von G ab. Die Abweichungen in der Verbstellung bestehen darin, dass im Hauptsatz das finite Verb nach links, an die zweite Stelle, und im Nebensatz nach rechts, hinter die zweite Stelle rückt.

⁶³ In a similar vein, Lippert (1974: 83) shows that verb second is the most frequent word order pattern in main declaratives in the OHG *Isidor*. Lippert notes that of 380 main declaratives, 280 exhibit V2 order. Of the latter, only 75 render a Latin V2 pattern, while in 205 cases, V2 order is established against the Latin source. The most frequent autochthonous deviations from V2 are V3/V4 orders after light elements such as pronouns and adverbs (20 cases), and 17 instances where more than one phrasal constituent precedes the finite verb. See sections 5.3 and 5.4 for some discussion.

2. Einer nichtdeutschen Wortstellung im Tatian entspricht immer eine nichtdeutsche Wortstellung in G. Das heißt: Jedesmal, wenn der Tatian eine nichtdeutsche Wortstellung hat, steht in G die gleiche Wortstellung."

'1. Each instance where the Tatian deviates from G topologically leads to a genuinely German word order or comes close to it. The Tatian never deviates from G in the direction of a non-German word order. The deviations concerning verb placement can be described as follows: In main clauses, the finite verb moves to the left, to the second position. In subclauses, the finite verb moves to the right of the second position.

2. In the Tatian, a non-German word order always corresponds to a non-German word order in G. That is, each time the Tatian exhibits a non-German word order, G exhibits the same word order.' (translation: EF)

The following table lists their observations for chapters 106-109 and 140-150 of the OHG *Tatian*:

	Latin version "G" – preverbal area	OHG – <i>Vorfeld</i> 'prefield'	Total
Word order of translation corresponds to original	empty single XP two XPs (or more)	empty single XP two XPs (or more)	40 142 29
Word order of translation deviates from original	two XPs (or more) two XPs (or more) single XP single XP	less elements, but > 1 XP single XP empty two XPs (or more)	2 16 3 –
Insertion of elements absent in the original	empty single XP (or more)	single XP single XP (or more) + 1	46 –
Insertion of finite verb leads to V2 order			2

Table 4: Prefield elements in the OHG Tatian (ch. 106-109 & 140-150), Dittmer and Dittmer (1998)

Even a brief glance at Table 4 reveals that the vast majority of changes result in V2 orders where the prefield is occupied by a single XP only. This is achieved either (i) by reducing the number of preverbal elements found in the Latin original (via shifting elements to a postverbal position, mostly to the *Mittelfeld* 'midfield'), or (ii) by inserting or shifting elements to the prefield in cases where the Latin original

exhibits V1 order. In the following example, the number of elements in the prefield is reduced to one by shifting the pronoun to postverbal position (i.e., to the midfield):

(81) *unum tibi deest* → *ein ist thir uuān*.
 one thing you lack one thing is you-DAT lacking
 'thou lackest one thing'
 (*Tatian*, 357,15 [106,3]; Dittmer and Dittmer 1998: 92)

As will become clear later on, strategy (ii) is of particular interest for the purposes of the present study. According to Dittmer and Dittmer (1998: 95), the chapters 106-109 of the OHG *Tatian* contain 12 cases where an empty preverbal position in Latin is rendered by an OHG clause with a single element in front of the finite verb (in addition, there are 34 relevant examples in chapters 140-150). Dittmer and Dittmer further note that these elements are always light elements such as subject pronouns (e.g., *ih* 'I', 3 examples), *thō* ('then', 8 examples), and *thanne* ('then', 1 example):

(82) **dixit** illi. → *thō quad her imo*.
 said him then said he him
 'then he said to him'
 (*Tatian*, 357,1 [106,2]; Dittmer and Dittmer 1998: 92)

(83) **rogo ergo te pater** → *ih bītiu thih fater*
 pray-1SG therefore you father I pray you father
 'I pray thee therefore father'
 (*Tatian*, 365,5 [107,3]; Dittmer and Dittmer 1998: 95)

Note that these are the very same elements that I have taken to occupy the preverbal position SpecTP in main clauses of OE. In section 5.6.2, I am going to argue that these orders played a special role in the rise of generalized V2 in German and presumably other Germanic languages. However, before we discuss the diachronic development of V2, let us first take a closer look at the syntax of main clauses in OHG, focusing on apparent deviations from V2.

5.3 Deviations from V2 in main declaratives

5.3.1 V-first order in main clauses

In the traditional literature, it has repeatedly been noted that V1 order constituted a well-established grammatical option in main declaratives of OHG (cf. Reis 1901, Maurer 1924, Behaghel 1932, Fourquet 1938, Lockwood 1968, Lippert 1974; for overviews and more complete references see Axel 2007, and in particular Önnerfors 1997). Evidence suggesting that V1 order in main declaratives was an autochthonous trait of OHG syntax comes from early cases in which V1 order is established independent of the Latin source:

(84) a. **uuarun** thô *hirta* In therō lantskeffi uuahante [...]
 were then/there shepherds in that country abiding
 Lt. *Et pastores erant In regione eadem. uigilantes [...]*
 'And there were shepherds in that country abiding [...]'
 (Tatian, 85,29; Axel 2007: 113)

b. {See **qui}mit** *der bruti gomo* gaat uz ingegin imo
 behold comes the bridegroom go out toward him
 Lt. *Ecce sponsus uenit, exite obuiam ei*
 'See, the bridegroom is coming! Go out and meet him!'
 (Monsee Fragments, XX, 8; Mt. 25,6; Axel 2007: 114)

c. Endi **uuirdit** *siin herduom* oba sinem sculdrom,
 and become his power over his shoulders
 endi **uuirdit** *siin namo* chinemnit uundarliih
 and become his name called admirable
 Lt. *et factus est principatus eius super humerum eius, et uocabitur nomen eius admirabilis.*
 'And his power will come over his shoulders, and his name will be called admirable.'
 (Isidor, 387; Robinson 1997: 24)

Further evidence that can be taken to indicate that relevant V1 patterns are a native phenomenon comes from later OHG texts such as Otfried's Gospel Harmony, in which this order is particularly frequent (cf. McKnight 1897), and autochthonous OHG texts such as the *Muspili*. Moreover, it is a well-known fact that V1 orders can

also be found in other early Germanic languages, which again suggests that V1 is a common trait of Germanic syntax (cf. Önnerfors 1997). The pragmatic function of V1 in OHG is not entirely clear. For example, Lenerz (1984, 1985a) assumes that V1 order signals that the complete proposition is rhematic, and that the sentence does not contain a thematic element (similar to narrative V1 declaratives in present-day German). However, Axel (2007) shows that V1 declaratives may contain thematic material such as pronouns, as in the following example:

(85) [...] **árstuont** *siu* *tho* *uf*
 arose she then/there up
 Lt. [...] & *surrexit*
 'She arose.'
 (*Tatian*, 183,14; Axel 2007: 124)

In general, it seems that V1 declaratives are linked to a oral, lively narrative style (cf. e.g. Önnerfors 1997). Further discourse functions mentioned in the literature seem to be comparable to those noted above for V1 orders in OE (see section 3.1.1 above): V1 orders are often used to introduce new facts, or the beginning of a new passage (often accompanied by a change of discourse topic); in particular, they seem to link clauses describing sequences of foregrounded actions/events along the main story line (cf. Petrova 2006).⁶⁴ Axel (2007: 167f.) notes that in early OHG, V1 order is often accompanied by the element *thô* 'then, there', which she analyzes as a discourse particle that reinforces the discourse function and narrative character of V1 order (see also Betten 1987 and section 5.5 below).

It can be shown that the cases of V1 found in early OHG can be subdivided into a couple of major patterns. According to Axel (2007: ch. 3), a major class of V1 patterns can be attributed to the fact that OHG had not yet developed an expletive *es* filling the prefield in existential constructions such as (84a), or presentational clauses such as (84b). Thus, V1 order resulted in cases where the prefield was not filled by a constituent for pragmatic/information structural reasons (as is typical of existential and thetic constructions).

⁶⁴ Robinson (1994, 1997: 24f.) claims that in the *Isidor*, V1 orders served as a marker of Bible quotations in the text. Furthermore, Robinson (1994) notes that V1 order is less frequent in the *Isidor*. This might have to do with the fact that in contrast to the *Tatian*, the *Isidor* is not a narrative, but rather a scholarly text, which does not call for a form type signaling oral narrative style (cf. Axel 2007: 166).

Another context that seems to favor V1 order involves unaccusative predicates such as mutative verbs (in particular verbs of movement) and passivized predicates (cf. Axel 2007: 124ff.):⁶⁵

(86) a. **arstarp** ouh *ther otago* Inti uuard bigraban [...]
dies also the rich and became buried

Lt. *mortuus est autem & diues & sepultus est* [...]

'The rich man also died and was buried.'

(*Tatian*, 363,11; Axel 2007: 125)

b. **Argengun** dhuo uz *pharisara* [...]
went PRT out Pharisees

Lt. *Exeuntes autem pharisaei* [...]

(*Monsee Fragments* IV, 30; Mt. 12.14; Axel 2007: 126)

c. **gieng** thô zuo *ther costari*
went then/there to the tempter

Lt. *& accedens temptator*

(*Tatian*, 113,28; Axel 2007: 126)

(87) a. **uuard** tho giheilit *ther kneht* in thero ziti.
became then/there healed the servant in that hour

Lt. *& sanatus est puer in illa hora*

'The servant was healed at that very hour.'

(*Tatian* 183,7; Axel 2007: 127)

b. **sint** thir furlazano *sunta*
are you.DAT forgiven sins.NOM

Lt. *dimituntur tibi peccata*

'Your sins have been forgiven you.'

(*Tatian* 193,24; Axel 2007: 127)

⁶⁵ Axel (p. 133f.) notes that in these constructions, V1 order is often accompanied by extraposition of the subject (cf. e.g. (86b,c) and (87a,b)) and speculates that extraposition renders the subject less accessible for fronting. In derivational terms, this seems to imply that extraposition precedes fronting. However, without an explicit theory of extraposition, it is not clear how this alleged correlation can be phrased in more formal terms.

Again, a subset of the relevant examples are thetic constructions which would be rendered in present-day German by a construction in which the prefield is occupied by the expletive *es* or a light adverbial element such as *da* 'then, there'. The latter is also true of other contexts where V1 is regularly attested such as *verba dicendi* and certain impersonal constructions in connection with nominal, adjectival or adverbial predicates, and 'psych-verbs' with accusative or dative experiencers:⁶⁶

(88) **quad** tho *maria* zi themo engile.

 said then/there Mary to the angel

 Lt. *Dixit autem maria ad angelem*

 'Mary said to the angel [...]'

 (*Tatian*, 71,24; Axel 2007: 150)

(89) **uuas** tho zit nah sehsta

 was then/there hour after six

 Lt. *hora erat quasi sexta*

 'It was about the sixth hour.'

 (*Tatian*, 275,29; Axel 2007: 142)

(90) **lustida** sie [christinheidi chilaupnissa chihoran]

 desired them.ACC Christianity belief hear

 Lt. *christi fidem delectantur audire*

 'They wanted to hear the belief of Christianity.'

 (*Isidor*, 694; Axel 2007: 142)

Furthermore, V1 order is regularly triggered in negated clauses, where the finite verb is fronted to initial position together with the proclitic negation *ni*:⁶⁷

⁶⁶ Interestingly, V1 with psych-verbs seems to be restricted to 3rd person experiencers, while first and second person experiencers regularly trigger V2 order:

(i) **íu** **gilimphit** [abur giboran uuerdan]

 you-DAT behoves again born become

 Lt. *oportet uos nasci denuo*

 'You must be born again.'

 ((*Tatian*, 405,26; Axel 2007: 146)

⁶⁷ However, we might assume that similar to OE, these examples involve an empty operator in SpecCP and thus may be subsumed under verb fronting in operator contexts.

(91) nisanta got sinan sun
 NEG-sent God his son
 Lt. non enim misit deus filium suum
 'God did not send his Son.'
 (*Tatian*, 407, 30; Axel 2007: 61)

Thus, the general picture that emerges is that V1 order is often triggered in declaratives when there is no pragmatic/semantic reason for XP-fronting (as e.g. in existential or presentational constructions) or when there is no adequate candidate available for fronting (as e.g. in impersonal constructions). In particular, OHG lacked an expletive *es* that could be inserted to SpecCP in these contexts. Under the assumption that the existence of expletives is a diagnostic for the existence of a structural specifier position that must be obligatorily filled (Haider 1993), the latter fact and the general availability of V1 orders suggest that OHG had not yet fully developed a generalized EPP feature in C (see Abraham 1993 for the hypothesis that the emergence of Vorfeld-*es* is connected to the rise of a structural topic position, SpecCP).⁶⁸ In the next section, I show that XP-fronting in OHG was still often linked to information-structural distinctions such as topic or focus, and that it was possible to front more than a single constituent if that was called for by pragmatic/discourse reasons. Again, this can be taken to indicate that OHG had not yet fully generalized semantically vacuous, EPP-driven XP-movement to clause-initial position.

5.3.2 V3 orders

Another set of data that suggests that the V2 constraint had not yet fully developed (at least in early OHG) involves cases where apparently more than a single constituent appears to the left of the finite verb in main clauses. According to Axel (2007: 202), we must distinguish at least six different types where the finite verb occupies a position further to the right, mostly giving rise to V3 order:

⁶⁸ Axel (2007) argues convincingly that it is not likely that the Vorfeld is occupied by a null expletive in V1 declaratives of OHG. So we may conclude that the rise of generalized V2 involved the development of a semantically vacuous EPP feature in C that requires XP-fronting in all main declaratives (see section 5.6 for detailed discussion). Furthermore, note that this insight suggests that at least from a diachronic point of view, an analysis of V2 in terms of vP-fronting (Müller 2004) is not plausible for German.

(92) a. V3 order after dislocated topics
 b. V3 order after two adverbial expressions (that in most cases can be taken to form some kind of unit, e.g. $[_{XP} [XP] [XP]] V_{fin} \dots$)
 c. V3 order induced by intervening sentence adverbs
 d. V3 order induced by intervening personal pronouns
 e. V3 induced by intervening short adverbs (mostly *thô* 'then')
 f. V3 after preposed adverbial clauses

However, Axel argues convincingly that a large portion of these apparent violations of V2 represent types of 'V3' that are still possible in present-day German. For example, there are quite a number of surface V3 orders which are reminiscent of left dislocation in present-day German.⁶⁹

(93) [thie morganlihho tág] [ther] **bisuorg&** sih selbo
 the tomorrow day-NOM that-NOM worries REFL self
Lt. crastinus enim dies. sollicitus erit sibi ipse.
 'For tomorrow will worry about itself.'

(*Tatian*, 157,14; Axel 2007: 204)

Similarly, many cases where apparently two adverbial constituents appear to the left of the finite verb are quite similar to instances of apparent multiple fronting of adverbials in present-day German. The latter are often taken to involve a structure in which the two adverbial expressions combine to form a single adverbial complex (cf. Haider 1982, Wunderlich 1984, Müller 2005).

⁶⁹ Note, however, that OHG left dislocation differs from its modern equivalent in a number of properties. For example, it appears that left dislocation with a resumptive *d*-pronoun (as in (93)) is confined to nominative elements in early OHG. Furthermore, we find other types of resumptive elements such as personal pronouns (in pre- as well as postfinite position), or null resumptives. See Axel (2007) for details.

(94) [In haubide dhes libelles] [azs erist] **ist** chisriban umbi mih, dhazs [...] in head of-the book at first is written about me that
 Lt. *In capite libri scriptum est de me, ut [...]*
 'In the beginning of the book it is written about me that [...]'
 (Isidor, 294; Axel 2007: 212)

(95) [[_{AdvP} Heute] [_{PP} auf dem Weg zur Arbeit]] **ist** der Max eingeschlafen.
 today on the to-the work is the Max fell-asleep
 'Today, on his way to work, Max fell asleep.'

Likewise, a set of V3 orders induced by sentence adverbs such as *giwisso* 'certainly' or *wārlīh(h)o* 'truly, really, indeed' can be likened to relevant constructions in Modern German, where sentence adverbs may intervene between a fronted XP and the finite verb if they function as discourse connectives (cf. Axel 2007: 217; for Modern German cf. Thim-Mabrey 1988, Dürscheid 1989):

(96) [Dhiu] [chiuuisso] **ist** bighin gotes sunes
 that-fem.nom certainly is origin of-God's Son
 Lt. *origo scilicet filii dei*
 'That certainly is the origin of the Son of God.'
 (Isidor, 116)

(97) Bei der Abschlussfeier aber/hingegen/indessen/allerdings/
 by the leaving party however/in contrast/on the other hand/however/
 freilich/schließlich können die Lehrer nur zusehen.
 sure enough/finally can the teachers only watch
 'The teachers ADV can only watch the leaving party.'

In what follows, I will focus on a couple of cases of V3 order that are no longer available in Modern Standard German and thus can be taken to represent clear deviations from a generalized V2 constraint (see Axel 2007 for a comprehensive overview). Furthermore, note that examples involving personal pronouns and short adverbs (in particular *thō*) are discussed separately (in sections 5.4 and 5.5, respectively).

However, before we can turn to genuine deviations from V2 in OHG, another caveat is in order. With respect to apparent deviations from V2, we have to bear in mind that OHG translations often respect the line breaks of the original. That is, if a sentence runs over two lines in the original, material which is part of the first line must not be shifted to the second line and vice versa in the OHG translation (see e.g. Masser 1997 and Dittmer and Dittmer 1998: 23 on the OHG *Tatian* translation). In many cases, this gives rise to word orders in which the finite verb occurs further to the right in main clauses (V3 in the following example):

(98) Auditis autem his discipuli → then gihorten thi iungoron
mirabantur ualde dicentes **uuntrotun** thrato sus quedenti:
 ‘And when they had heard this, the disciples **wondered** much, saying:
 (Tatian 359,6 [106,4]; Dittmer and Dittmer 1998: 98)

Of course, these orders do not reflect genuine properties of OHG. However, in what follows I will discuss genuine instances of V3 including: (i) topicalization/left dislocation of DPs in interrogatives; (ii) cases of XP-XP-V_{fin} where the fronted XPs cannot be analyzed as forming a unit; (iii) XP-adv-V_{fin} orders, where the sentence adverb cannot be analyzed as a discourse connective, and (iv) V3 orders with fronted adverbial clauses.

First of all, topicalized elements may occur in interrogatives introduced by the interrogative particle *inu/eno*. In these cases, the topic uniformly appears to the right of the particle, similar to *wh*-phrases (cf. fn. 61 above):⁷⁰

(99) Inu ni [angil] **nist** anaebanchiliih gote?
 INU NEG angel-NOM NEG-is identical God-DAT
 Lt. *Num angelus equalem cum deo habet imaginem?*
 ‘Is an angel not identical to God?’
 (Isidor 184; Axel 2007: 206)

According to Axel (2007: 207), (99) shows that the particle and the finite verb are not in a specifier-head relation in OHG. Moreover, she claims that examples like (99) suggest that CP splits up into several projections in OHG (Rizzi 1997), with *inu/eno*

⁷⁰ Axel (2007: 207) assumes that cases such as (99) contain a null resumptive pronoun, similar to other cases of left dislocation in OHG.

located in the ForceP (where sentence type is encoded) and the finite verb in a lower head position, presumably Fin⁰. Note that this conclusion is not by all means necessary; alternatively we might posit a structure with a single head and multiple specifiers, along the lines proposed in Chomsky (1995) and subsequent work. See section 5.6 for more discussion.

Furthermore, topics may occur to the left of fronted *wh*-phrases, which is reminiscent of the structure of the left periphery proposed by Kiparsky (1995) for Indo-European and Proto-Germanic:⁷¹

(100) a. [ir uuarlichō]/ uuēn mih **quedet** uuesen/

 you-NOM in-fact who me-ACC say be

 Lt. / [...] *uos autem / quem me esse dicitis*/

 ‘But you in fact, who do you say who I am?’

 (*Tatian*, 299,32; Axel 2007: 209)

 b. [Uuexsal dhes nemin]/ huuazs **bauhnida**?

 changing-NOM of-the name what meant

 Lt. *Mutatio nominis quid significabat*?

 ‘The changing of the name, what did it mean?’

 (*Isidor*, 532; Axel 2007: 209)

Robinson (1997: 26ff.) points out that the OHG *Isidor* exhibits another set of deviations from V2. In the relevant examples, we find V3 orders of the type XP-XP-V_{fin}. However, in contrast to (94), it is unlikely that the two fronted XPs form a single unit:

(101) a. [Dhea uuehhun] [auur] [in heilegim quhidim] **arfullant** sibun iaar.

 the weeks however in sacred language fulfil seven years

 Lt. *Ebdomada namque in sacris eloquiis septem annis terminatur*.

 ‘The weeks, however, take seven years in sacred language.’

 (*Isidor*, 457; Robinson 1997: 26)

⁷¹ Note, however, that both examples in (100) are very similar to the Latin source. In (100b), the example taken from the *Isidor*, the word order of the OHG translation is fully identical to the Latin text, while in (100a), the only difference between the OHG text and the Latin source concerns the relative order of the clause-final verbs. In other words, it is not entirely clear whether these example really can be taken to reflect genuine properties of OHG.

b. [So] [auh in andreru stedi] [dhurah dhen selbun heilegum forasagun]
 so also in other places through the same holy prophet
uuard dhera dhrinissa bauhnunc sus **araughit:** [...]
 became the-GEN Trinity-GEN meaning in this way demonstrated

Lt. *Item alibi per eundem prophetam trinitatis sic demonstratur significantia: [...]*

'In this way, also elsewhere the meaning of the Trinity was demonstrated by
 the same holy prophet: [...]'

(*Isidor*, 328; Robinson 1997: 27)

Another potential violation of V2 involves clauses with fronted sentence adverbs where it is unlikely that they are used as sentence connectives, in contrast to (96) above. Rather, it seems that in the following examples, the sentence adverbs *uuarlichō* and *chiuuisso* are used to express an epistemic meaning (cf. Axel 2007):

(102) / min fleisg *uuarlichō* **ist** muos/ inti min bluot *uuarlichō* **ist** trang/
 my flesh truly is food and my blood truly is drink

Lt. */Caro enim mea. uere est cibus./& sanguis meus uere est potus; /*

'My flesh is truly food and my blood is truly drink.'

(*Tatian*, 263,11; Axel 2007: 221)

(103) endi bidhiu iu *chiuuisso* **quham** *christ*
 and therefore already certainly came Christ

Lt. *Ideoque iam aduenit christus*

'And that's why Christ has certainly already come.'

(*Isidor* 464; Axel 2007: 221)

Finally, it appears that fronted adverbial clauses trigger systematic violations of V2 in early OHG. The special placement properties of adverbial clauses lead Axel (2007) to assume that they are not properly embedded in their matrix clause. Instead, she claims that (fronted) adverbial clauses are left-adjoined to the root node (i.e., ForceP in her analysis). This analysis is supported by the fact that adverbial clauses always occur at the outermost left edge of the clause. In main clauses, they appear to the left of the interrogative particle *inu/eno* (the only other elements that may occur in this position are vocatives) as illustrated in (104). In embedded clauses, they may precede

the complementizer, cf. (105). Moreover, they may occupy a position to the left of coordinating conjunctions, as shown in (106):

(104) /[thanne ih iuuuih santa/ uzzan seckil] / [...] / eno **uuas**
 when I you sent without bag PRT was
 iu iouuiht thes uuauan
 you anything of-that need
 Lt. */quando misi uos / sine sacco / [...] / numquid aliquid defuit uobis*
 'When I sent you without a bag [...], did you lack anything?'
 (Tatian, 575,1; Axel 2007: 210)

(105) [...] neist tés nîomannen vuúnder. [[sô der uuínt uuáhet]
 neg-is that-gen noboday-acc wonder when the wind blows
táz tiu uuélla án den stád sláhet]
 that the wave at the shore crashes
 Lt. *Nemo miratur flamina chori. túndere litus frementi flutv*
 'Nobody is surprised that the wave crashes at the shore when the wind is
 blowing.'

(Notker BCon IV 211,2; Axel 2007: 230)

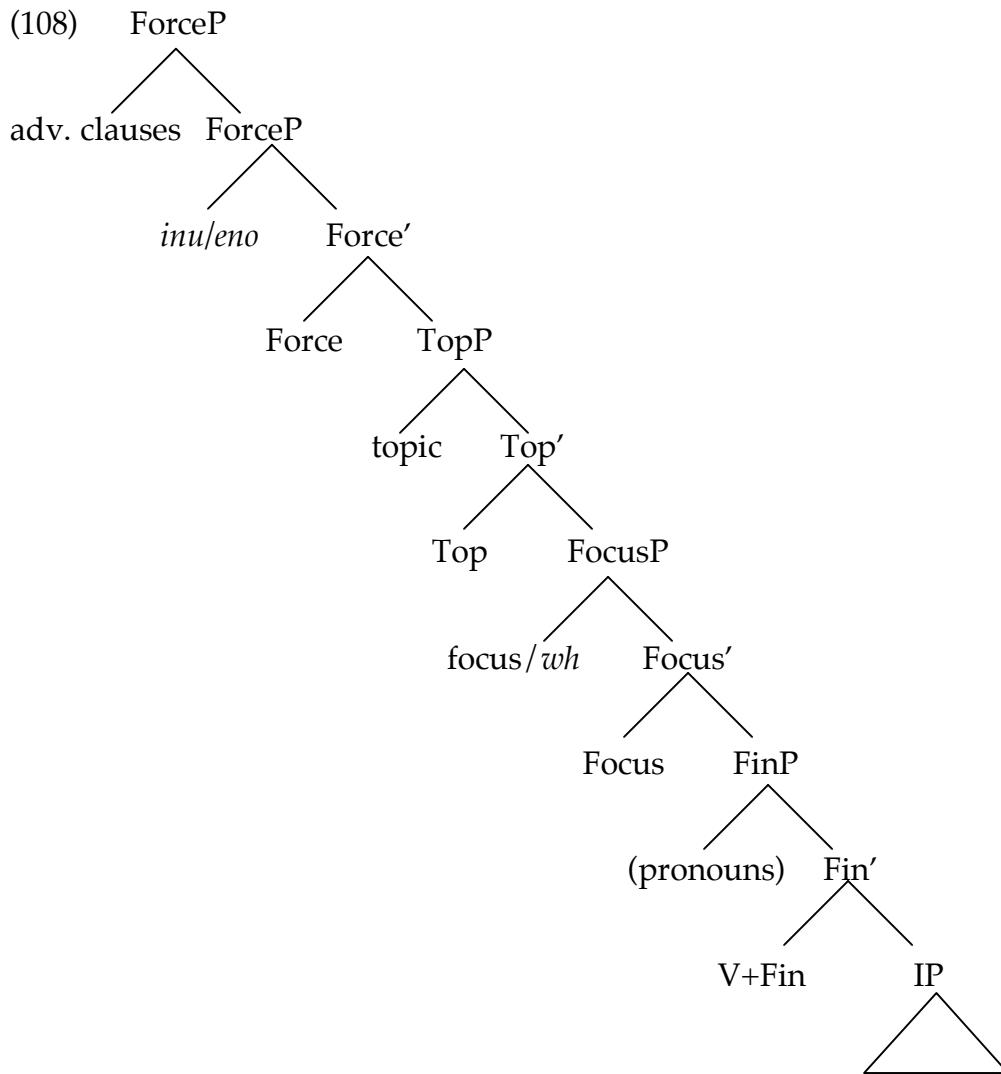
(106) Enti [ibu daz {hus sii} uuir dich] enti *iuuuer fridu* **quuimit**
 and if that house be worthy and your peace comes
 ubar daz hūs
 upon the house
 Lt. *et sie quidem fuerit domus digna, ueniat pax uestra super eam*
 'And if the house is worthy, let your peace come upon it.'
 (Monsee Fragments, II,2; Mt. 10.13; Axel 2007: 229)

Based on these observations, Axel (2007: 210) proposes the following generalizations on the relative orderings of elements in the left periphery of interrogatives and declaratives:

(107) a. adverbial clause > *inu/eno* > disloc. topic > *wh* > (pron.) > V_{fin} ... (interrogatives)
 b. adverbial clause > disloc. topic > XP > (pron.) > V_{fin} ... (declaratives)

Note that these generalizations highlight a set of significant differences between OHG and OE. In particular, OHG does not exhibit the pattern $XP\text{-}DP_{\text{subj}}\text{-}V_{\text{fin}}\text{-}...$ which was quite productive in OE (in around 28% of all main clauses according to Haeberli 2000). Likewise, we do not find systematic V3 orders in the context of (subject) pronouns, as already briefly noted above. Still, Axel takes these generalizations to imply that V2 orders did not reflect an underlying spec-head configuration in early OHG. Rather, verb movement targeted a low head in the C-domain (Fin), while XPs could be fronted to a number of specifiers in a split CP, to the right of the interrogative particle *inu/eno*, which by assumption is merged as SpecForceP. Some of these operations are triggered by semantic/pragmatic factors (topic, focus, *wh*), but recall that early OHG had already developed the option of XP-fronting driven by a semantically vacuous EPP-feature. Still another option is reflected by the placement of adverbial clauses, which Axel analyzes as adjuncts to ForceP (which is by hypothesis an archaic trait of Proto-Germanic/Indo-European):⁷²

⁷² Axel (2007: 234) notes that “In the OHG sources there is no evidence that topics and *wh*-phrases occupied different positions.” However, this claim is at odds with the word order generalizations she proposes on page 210 (my (107)), where (dislocated) topics appear to the right of *inu/eno* and to the left of *wh*-phrases. In particular, in contrast to what seems to be implied by Axel on page 234, dislocated topics cannot be analyzed in terms of adjunction to the root node, since they should otherwise appear to the left of the interrogative particle. Accordingly, I opted for the structure in (108) with different positions for topics and foci, even if that slightly misrepresents Axel’s original proposals.



Axel assumes that the various possibilities of filling (different) positions in the left clausal periphery were lost in the course of the OHG period, giving rise to the clause structure that still characterizes Modern German. Thus, she claims that at some point, the formerly split CP was conflated into a structure with only two positions in the CP (see (80 above, repeated here for convenience):

(109) [_{FinP} XP_i [_{Fin'} V_j+Fin [... t_i ... t_j]]]

Note that this analysis raises a number of questions, for example concerning the possibility of adjunction in a split-CP approach, or the reasons that motivated the loss of the split-CP structure. Some of these are discussed in more detail in section 5.6. Concerning the conflation of CP, Axel claims that an important part of this change was the loss of a special position for pronominal elements in the left clausal periphery, which she identifies as SpecFinP (see also Roberts 1996). This claim is

examined in the next section, which focuses on the distribution of subject pronouns in OHG.

5.4 V2 and pronoun placement in Old High German

Recall that in OE main declaratives, subjects pronouns consistently fail to undergo inversion with the finite verb, giving rise to systematic violations of V2. This fact is considered by many researchers to be the crucial piece of evidence suggesting that OE differs significantly from the Modern Germanic V2 languages. Above, we have already noted that OHG differs from OE in that subject pronouns more often follow than precede the finite verb in non-embedded declaratives, giving rise to V2 patterns similar to the modern Germanic languages. In this section, I will take a closer look at the placement of pronouns in early OHG, arguing that deviations from V2 are due to an archaic SOV option in main clauses (cf. Lenerz 1984), *pace* Axel (2007) who proposes an analysis in terms of pronoun movement into the left clausal periphery (SpecFinP).

5.4.1 V2 and pronoun placement in the OHG *Isidor*

In the following examples taken from the OHG *Isidor*, a non-V2 order lacking an overt subject pronoun (in the Latin original) is rendered by a V2 declarative in which the newly inserted subject pronoun follows the finite verb. This deviation from Latin can be taken to reflect genuine properties of OHG, suggesting that in contrast to OE, pronouns underwent subject-verb inversion as early as in the OHG *Isidor* translation:

(110) a. [Dhinera uuomba uuwaxsmin] **setzu** *ih* ubar min hohsetli
 your womb's fruit place I upon my throne
 Lt. *De fructu uentris tui ponam super sedem meam*
 'I will place the fruit of your womb upon my throne.'
 (*Isidor*, 611; Robinson 1997: 9)

b. [In dhemu uuorde] **chundida** *ir* bifora umbi christan
 in those words prophesied he before about Christ
 himilischen druhtin, dhazs *ir* in sines edhiles fleische
 heavenly Lord that he in his noble flesh
 quhoman scolda uuerdan.
 come should will

Lt. *Quo uerbo Christum deum celi de genere suo testabatur in carne esse uenturum*

'In those words he prophesied about Christ the Lord of heaven that He
 would come in the flesh of his noble line.'

(*Isidor*, 559; Robinson 1997: 72)

c. endi [bidhiu] **uuard** *ir* uns chiboran
 and therefore was he us born

Lt. *et ideo nobis natus est*

'And therefore he was born to us.'

(*Isidor*, 394)

However, as noted by Fourquet (1938) and Lippert (1974), the OHG *Isidor* translation also exhibits patterns in which the subject pronoun fails to undergo inversion with the finite verb, similar to the regular V3 patterns of OE (see Eythórsson 1995, Tomaselli 1995, Robinson 1997, and in particular Axel 2007 for discussion). According to Eythórsson (1995) and Axel (2007), the order XP-pron.-V_{fin} is actually quite frequent in the *Monsee Fragments* and the *Isidor*. Eythórsson (p. 327) counts 26 matrix declaratives in which a pronominal element (which he analyzes as a clitic) intervenes between a fronted XP and the finite verb (according to Axel, p. 242, the number of relevant examples amounts to 27 cases). Both authors note that there are roughly thirty examples in which the relevant pronominal elements (and certain adverbs, see below) follow the finite verb in main clauses. Thus, the ratio of inversion with (light) pronouns appears to be about 50% in main declaratives in the OHG *Isidor*. The examples in (111) serve to illustrate non-inversion with pronouns. In the first conjunct clause of (111a), the subject pronoun *ih* intervenes between the fronted object and the finite verb. In a similar vein, the subject pronoun *uuir* leads to V3 order in (111b). In the second conjunct of (111a), an object pronoun appears in prefinite position, in (111c) a reflexive pronoun, and (111d) illustrates that both subject and object pronoun can intervene between fronted elements and the finite verb. Note that in all the examples given below, the position of the pronominal element does not

correspond to the Latin source, mostly because there is no pronoun present in the original.

(111) a. [Erino portun] *ih firchnissu*, *iinsnne grindila firbrihu*
 bronze portals I destroy-1SG iron locks break-1SG
endi [dhiu chiborgonun hort] dhir ghibu
 and the hidden treasures you give-1SG

Lt. *Portas aereas conteram et uestes ferreos confringam et dabo tibi thesauros absconditos*
 'I destroy bronze portals, break iron locks and give you the hidden
 treasures.'

(Isidor, 157; Robinson 1997: 17)

b. [Dhes martyrunga endi dodh] *uuir findemes* mit urchundin
 of-his martyrdom and death we prove with testimony
dhes heilegin chisribes
 of-the holy scripture

Lt. *Cuius passionem et mortem in suo loco scripturarum testmoniis adprobabimus*

(Isidor, 516; Robinson 1997: 17)

c. enti [diu himilo megin] *sih hruorent*
 and the heavens' powers REFL agitate

Lt. *et uitutes caelorum commouebuntur*

'And the powers of the heavens will be shaken.'

(Monsee Fragments, XIX,4, Mt. 24.29; Axel 2007: 243)

d. [Fona hreue] [aer lucifere] *ih dhih chibar*
 from womb before Lucifer I you-ACC bore

Lt. *Ex utero ante luciferum genui te*

'I bore you out of the womb before Lucifer.'

(Isidor, 409; Robinson 1997: 17)

However, note that in the vast majority of these cases, the finite verb appears in absolute clause-final position (17 examples, according to Eythórsson 1995: 327), as in (111a, c, d). This led Lenerz (1984) to conclude that the order XP-pron.-V_{fin} should actually be analyzed as an instance of matrix SOV order (in combination with fronting of topic XPs), which he treats as an archaic vestige of an earlier (Pan-Germanic) grammatical system (see the brief discussion of matrix SOV order in OE in

section 3.1.1 above). This view is shared by Eyþórsson (1995), who follows Behaghel (1932: 14) in assuming that the pattern $XP\text{-pron}\text{-}V_{fin}$ was triggered for (archaic) metrical reasons (see Axel 2007: 252 for critical discussion):

“The fact that the pattern clitic-verb is almost entirely restricted to environments where the clitic would have been in final position had it followed the verb indicates that the pattern was preserved, for general metrical considerations, to avoid an unstressed element in absolute clause-final position.” (Eyþórsson 1995: 327f.)

In a more formal way, the presence of archaic word order patterns can be analyzed as an instance of grammar competition, in which speakers have access to more than one internalized grammar (cf. e.g. Kroch 1989). Under these assumptions, orders such as (111b), in which another XP appears to the right of the finite verb, can be attributed to extraposition/exbraciation of material to the right edge of the clause (note that this process is still available in Modern German for embedded clauses, PPs, and other heavy constituents).

The hypothesis that the order $XP\text{-pron}\text{-}V_{fin}$ in fact represents the residue of an archaic grammatical option which was not any longer productive in (early) OHG is further supported by the fact that violations of V2 in connection with pronouns soon became obsolescent in later stages of OHG. In the *Tatian* (around 850), the order $XP\text{-pron}\text{-}V_{fin}$ is already much less frequent (although we can observe a number of V2 violations with light adverbs; see section 5.4.2 for details) while the relevant pattern is very rare in late OHG records (ca. 950-1050), compare the following quote taken from Axel (2007: 240): “In late OHG texts [...] the pronouns almost always occur postfinitely [...] The prefinite placement is only attested very sporadically”.

As already briefly noted above, however, Axel (2007) argues against an analysis of examples such as (111) in terms of non-embedded SOV order. Instead, she claims that the relevant OHG examples should be analyzed in terms of a split-CP structure (Rizzi 1997), in which the finite verb occupies Fin^0 , while the pronoun is located in $SpecFinP$. Violations of V2 in connection with pronouns are then attributed to movement of fronted XPs to other topic- or focus-related specifier positions in the left clausal periphery (see Roberts 1996 for a similar proposal; see section 5.6 for further discussion). Axel’s case against the SOV analysis originally put forward by Lenerz (1984) is based on two arguments. First, she observes that an SOV analysis is not

possible in cases of XP-pron.-V_{fin} where further pronouns appear in postverbal position (due to the fact that extraposition/exbraciation of light elements such as pronouns is generally ruled out across Germanic):

(112) Vnde [dô] íu **hâbeta** si lêid in-fágén in íro hérzen.
 and then you.DAT.PL had she sorrow received in her heart
 'And then her heart was filled with sorrow for you.'

(*Notker Ps VII 23,26*; Lenerz 1985a: 121)

Axel's second argument is based on her observation (Axel 2007: ch. 6) that in OHG, null subjects are confined to postfinite position (i.e., to a position to the right of the finite verb, which occupies a functional head in the C-domain; note that similar restrictions on the distribution of null subjects have been reported for other V2 languages such as Old French, cf. e.g. Kuen 1957, Roberts 1993a). According to Axel, this observation forces us to assume that the finite verb has undergone movement to C/Fin in examples such as (113) which (i) exhibit a null subject (which by assumption can only be placed in postfinite position) and (ii) where the verb occupies the clause-final position on the syntactic surface. However, if we have to assume that verb movement has taken place in examples such as (113), then, so the argument goes, nothing prevents us from analyzing other examples with clause-final verbs along the same lines. At first sight, this appears to be the most parsimonious option, since it eliminates the need for positing an additional grammatical option of matrix SOV structures for OHG (cf. Axel 2007: 279).

(113) [Auuar] *iu* **sagem** [...]
 again you.DAT.PL say-1SG
 Lt. *Iterum dio uobis* [...]
 'Again, (I) say to you [...]'
 (*Monsee Fragments XI,18*; Mt. 18.19; Axel 2007: 241)

However, there are some facts that cast some doubts on the validity of Axel's arguments. Starting with the second argument based on the alleged correlation between verb placement and the licensing of null subjects, it is not clear at all whether examples such as (113) really force us to assume that the null subject occupies a postverbal position. As already briefly noted, the restriction that null

subjects are confined to postfinite position (after verb movement to C) is apparently a characteristic of V2 languages only (such as OHG or Old French). No such restriction holds for other pro-drop languages such as Italian, for example. Furthermore, there are many SOV languages that exhibit null subjects in preverbal position (Latin, Hindi, Korean, Japanese, to name only a few). So, if the requirement that null subjects can only be licensed in postfinite position is systematically connected to the V2 property, we should perhaps doubt that this restriction is also at work in examples like (113) where the V2 constraint is violated. On the other hand, if examples such as (113) are attributed to the availability of an (archaic) non-V2 matrix SOV grammar in early OHG (a residue of an earlier stage, presumably an instance of grammar competition), then we actually do not expect that null subjects are located in postverbal position in clauses like (113). The latter analysis in terms of an archaic grammar option receives further support from the fact that exactly this system (matrix SOV + pro-drop) can be observed in Gothic, the earliest attested stage of Germanic (cf. Eythórsson 1995, Ferraresi 1997). Thus, we can conclude that examples such as (113) do not provide clear evidence against an analysis of XP-pron.-V_{fin} orders in terms of an matrix SOV option (as proposed by e.g. Lenerz 1984).⁷³

⁷³ Furthermore, note that quite a number of the relevant violations of V2 (cf. e.g. (111a) and (111c) above) occur in clauses introduced by the conjunction *endi* 'and'. In many early Germanic languages, we can observe asymmetries with respect to verb placement in conjoined main clauses. While the first conjunct regularly exhibits V2 order/inversion, second conjuncts often exhibit basic SOV order, which led some researchers to speculate that 'and' was originally ambiguous between a coordinating and a subordinating conjunction (cf. Mitchell 1985, Kiparsky 1995 for OE, Behaghel 1932: 25f. for OHG). So, perhaps, some of the apparent violations of V2 in connection with pronouns can be attributed to the influence of the conjunction 'and' on word order in second conjuncts. Interestingly, Behaghel (1932) shows that the phenomenon in question (OV order in second main conjunct clauses) is not confined to OHG, but can be observed in all historical stages of German:

- (i) getaten sie mih pogen unde
made they me bow and
iro werch ieo ze erdo **sahen**
their works/deeds ever/always to earth saw-PL
(OHG, Np, 216,14)
- (ii) a. nu haben wir niwan dri wochen dahin, daz [...]
now have we merely three weeks till then that
unde sich Simeon so lange darzuo **bereite**
and REFL Simon that long prepare
(MHG, Berth. I, 567,3)

Now, turning to the first argument raised by Axel, note that examples such as (112), where the finite verb is sandwiched between two pronouns, are very rare in all stages of OHG. In fact, the example given in (112) is the only one I am aware of, at least in a prose text. In particular, no relevant examples of the type $XP\text{-pron-}V_{\text{fin}}\text{-pron}$ can be found in the early stages of OHG where violations of V2 in connection with pronouns are more frequent (i.e., in the *Isidor* and the *Monsee Fragments*). Instead, (112) comes from a late OHG text (Notker's Psalter, i.e., a translation of a book of Latin psalms), that is from a stage of OHG for which it is generally assumed that the V2 constraint had already been generalized to all main clause types (cf. e.g. Axel 2007: 200: "[...] late OHG grammar seems to be as restrictive towards verb-third orders as the modern German grammar. There are hardly any verb-third orders attested in Notker's text."). Thus, even if the status of (112) is unclear – whether its particular word order was used to mimic an archaic system, or the word order of the Latin original – it presumably cannot be used as an indication of syntactic properties of early OHG since it was generated by a grammar with different properties (i.e., a grammar with generalized V2).

Another interesting observation made by Axel (2007: 244f.) and Petrova and Solf (2007: 17) is that there are three examples where pronouns intervene between a fronted *wh*-word and the finite verb (without a relevant model in the Latin text). This contrasts with the distribution of pronouns in OE, where they consistently follow the finite verb:

b. daz ist ein lauter bloz ancapfen des obersten gutes,
 that is a mere tapping of.the greatest good
 und daz oberst gut sich reichlich **ergenzet** in den Geist
 and that geatest good REFL amply adds into the spirit
 (MHG, Jostes, Eckhart, 47,14)

(iii) a. denn Gott sind alle ding möglich und wyr niht **wissen** [...]
 since God are all things possible and we not know
 (NHG, Luth. III, 73,17)

b. so wird der Stockfisch alle gefangen, und große Scheunen voll **sind**
 so is the stockfish all caught and huge barns full are
 (NHG, Diez, 168)

(114) a. [christes chiburt] [uuwer] *sia* **chirahhoda?**
 Christ's birth who it[SG.FEM.ACC] reported
 Lt. *Generationem eius quis enarrauit?*
 '[The birth of Christ], who made it, fully known?'
 (Isidor, 106)

b. [Dhiu uurza dhera spaida] [huuemu] *siu* **uuard**
 the root of-the wisdom whom it [SG.FEM.NOM] became
 antdhechidiu
 revealed
 Lt. *Radix sapientę cui reuelata est*
 'The root of wisdom, to whom has it been revealed?'
 (Isidor, 115)

c. [uuwer] *sih* [dhes] **biheizssit** *sia zi archennenne* [...]
 who refl that-gen insists her to recognize
 Lt. *quis confitebitur nosse* [...]
 'Who claims to know her?'
 (Isidor, 108)

Note that all three of these examples involve some form of resumption, or a correlative construction. In (114a) and (114b), the intervening pronoun is a resumptive element that takes up a left-dislocated DP, while in (114c), the prefinite position is occupied by a reflexive pronoun and the demonstrative *dhes*, which is a correlative to the (extraposed) infinitival clause *sia zi archennenne*. So even if these examples cannot be attributed to Latin influence, we might speculate that their deviating behavior has something to do with special properties of resumption/correlatives in OHG that override the V2 constraint in the above examples. Moreover, Petrova and Solf (2007) demonstrate that across OHG, violations of V2 in *wh*-questions are very rare. According to Petrova and Solf, only five of 417 *wh*-questions in their corpus (the examples in (114) are three of them) can be considered to be clear violations of V2 that cannot be attributed to Latin influence. In the *Isidor*, they count 18 other cases where strict V2 order is observed in *wh*-questions. Thus, we may conclude that despite the questions raised by the examples in (114), the OHG *Isidor* exhibits systematic V-to-C movement in *wh*-questions as well.

Summing up, while early OHG exhibits a number of word order patterns that deviate from the kind of generalized V2 we find in Modern German (in particular, matrix V1 and SOV orders, V3 orders of the type XP-XP-V_{fin}), we do not observe systematic deviations from V2 in connection with pronouns, in contrast to OE. Rather, the relevant cases of XP-pron.-V_{fin} order in the OHG *Isidor* and the *Monsee Fragments* are more likely to be the result of an archaic SOV grammar option in main clauses, which was already in decline in early OHG. The next section shows that this view is supported by the properties of pronoun placement that we can observe in the OHG *Tatian*.

5.4.2 V2 and pronoun placement in the *Tatian*

As already briefly noted above, V3 orders with pronouns intervening between fronted elements and the finite verb are significantly less frequent in the *Tatian*. Axel (2007: 242) gives the following example (note that while the finite verb has been shifted further to the left, the relative order of adverb and subject pronoun still corresponds to the order of the Latin source):

(115) [ziuuare] **thu** **bist** fon then
 surely you are from them
 Lt. *uere & tu ex illis es*
 ‘Surely you are one of them.’
 (*Tatian*, 613,12)

However, violations of V2 such as (115) are vastly outnumbered by cases in which subject pronouns are moved (or newly inserted) to postfinite position (i.e., to the *Mittelfeld*). Dittmer and Dittmer (1998: 79) show that there are 21 examples in which the subject pronoun is shifted to a postverbal position, giving rise to V2 order (in nine of the relevant examples, shifting the subject to postfinite position leaves a single element in the *Vorfeld*):

(116) tunc & **ipse ascendit** → tho **ersteig** her úf.
 ad diem festum zi themo itmalen dage
 ‘then *he went* also up unto the feast’
 (*Tatian*, 347,12f. [104,3]; Dittmer and Dittmer 1998: 79)

In addition, there are 76 examples in which the Latin source does not exhibit an overt subject while the OHG translation exhibits a light subject that has been inserted into the midfield (Dittmer and Dittmer 1998: 88):

(117) **eduxit** autem eos foras → **tho leita** *hér sie úz*
 In bethaniam In bethaniam
 'And *he led* them out as far as to Bethany'
 (Tatian, 695,1f. [244,2]; Dittmer and Dittmer 1998: 64)

Furthermore, a number of examples where the insertion of a pronouns leads to V3 order can be analyzed as SOV main clauses, similar to the relevant examples in the OHG *Isidor*:

(118) Nemo potest duobus dominis seruire > Nioman nimag zuuein herron thionon
 aut enim unum **odio habebit** odo *her einan hazzot*
 et alterum diligit. inti anderan minnot.
 aut unum sustinebit. odo einan gitregit
 et alterum contemnet inti anderan ubarhugit.
 'No man can serve two masters: for either he **will hate** the one, and love the
 other; or else he will hold to the one, and despise the other.'
 (Tatian 85,29 [6,1]; Dittmer and Dittmer 1998: 100)

A related example exhibiting the finite verb in absolute clause-final position is given in (119). Again, a subject pronoun has been inserted, which enlarges the preverbal field. Similar to the previous example, however, the result can be analyzed as a main SOV clause. Note that due to the position of the line breaks, the verb could not be shifted into a position to the left of its object in this example:

(119) nam digna → *uuir uuirdigen*
 factis **recepimus** tatin **intfahemes**.
 'for we **receive** the due reward of our deeds'
 (Tatian 85,29 [6,1]; Dittmer and Dittmer 1998: 100)

Similar to the *Monsee Fragments/Isidor*, there are a couple of *wh*-questions where a (object) pronoun intervenes between the *wh*-word and the finite verb (Axel 2007: 244f.):

(120) a. [uuanan] **uns** **sint** in uuostino so manigu brot
 whence us-DAT are in desert so many breads
 Lt. [...] *unde ergo nobis in deserto panos tantos?*
 'Where are we to get so many loaves of bread in the wilderness?'
 (*Tatian*, 295,23)

 b. [uuaz] **mih** **frages** fon guote?
 what me-ACC ask of good
 Lt. *quid me Interrogas de bono*
 'Why do you ask me about what is good?'
 (*Tatian*, 355,29)

 c. [uuer] **mih** **sazta** zi duomen [...]
 who me-ACC installed to judge
 Lt. *quis me constituit iudicem*
 (*Tatian*, 353,22)

But note that in the above examples, the word order of the OHG text is either very similar or fully identical to the order found in the Latin source. Thus, these examples are not reliable indications of properties of OHG and should perhaps be discarded (see also Petrova and Solf 2007 for some discussion).

Summing up, it appears that at least with respect to V2, the distribution of pronouns in the *Tatian* is very similar to the system of Modern German, in the sense that there is a strong tendency to place pronouns directly to the right of the finite verb in main clauses.

5.4.3 Section summary

This section has shown that in early OHG, pronouns already regularly followed the finite verb in main clauses. Thus, in contrast to OE, both pronominal and non-pronominal subjects undergo regular inversion with the verb in main clauses of OHG. Furthermore, I have argued that the evidence put forward by Axel (2007) in favor of an optional (archaic) mechanism placing pronouns in SpecFinP does not

withstand closer scrutiny. Instead, the fact that the finite verb appears in absolute clause-final position in most of the relevant examples suggests that these deviations from V2 lend themselves to an analysis in terms of main SOV clauses, representing an archaic grammatical option that was already in decline in early OHG.⁷⁴ This can be taken to suggest that already prior to early OHG, the relevant specifier had turned into a position that could be targeted by all kinds of XP-fronting (including *wh*-movement and movement triggered by a semantically vacuous EPP feature). In addition, we have seen before that in early OHG, syntactic operations could target positions further to the left (cf. the distribution of dislocated topics, adverbial clauses, and the interrogative particle *inu/eno*). In section 5.6, I am going to argue that the relevant data can also be captured by an analysis that assumes only a single head position in the left periphery that may project multiple specifiers in early OHG (an option which is lost in the course of the OHG period). Before we can turn to the relevant proposal, let us first take a look at another element that frequently triggers V2 in OHG, namely the temporal adverb *thô* 'then' (which is cognate with Gothic *þan(uh)* and OE *þa*).

5.5 V2 and the placement of the adverb *thô*

Similar to its Gothic and OE cognates *þanuh* and *þa*, the adverb *thô* 'then' is frequently used as a clause-initial sentence connective in OHG (Betten 1987: 397 counts 212 instances of the sentence connective *thô* in the first quarter of the *Tatian*). In the following passage from the OHG *Tatian*, we find five instances of *thô*, four of them in clause-initial position. Furthermore, it appears that similar to the corresponding Gothic and OE forms, clause-initial *thô* obligatorily triggers inversion in OHG:

(121) Thô gihortun inan thiie iungiron sprechantan inti folgetun themo heilante. Thô hiuuanta sih ther heilant inti gisah sie imo folgente, quad in: uuaz suochet ir? Sie quadun imo: rabbi (thaz ist arrekit meistar) uuâr artos? Thô quad her in: quemet inti gisehet. **Quamun** sie thô inti gisahun uuâr her uuoneta, inti uuonetun mit imo then tag; thô uuas thiu zehenta zît thes tages. (*Tatian* [16.2])

⁷⁴ Note that Axel (2007: 51f.) also considers the possibility of an archaic SOV option in the context of verb-final main clauses introduced by the affirmative particle *ja*.

Lt. John 1.37. *Et audierunt eum discipuli loquentem et secuti sunt Ihesum.* 38. *Conversus autem Ihesus et videns eos sequentes se, dicit eis: quid queritis? Qui dixerunt ei: rabbi (quuod dicitur interpretatum magister) ubi habitas?* 39. *Dicit eis: venite et videte. Venerunt et viderunt ubi maneret, et apud eum manserunt de illo; hora autem erat quasi decima.*

'The two disciples heard him speak, and they followed Jesus. Then Jesus turned, and seeing them following, said to them, "What do you seek?" They said to Him, "Rabbi" (which is to say, when translated, Teacher), "where are You staying?" He said to them, "Come and see." They came and saw where He was staying, and remained with Him that day (now it was about the tenth hour).'

If we compare the OHG passage with the Latin text, we can see that there are no clear correspondences between the use of *thô* and Latin sentence connectives. In two cases, the OHG translation exhibits *thô* where the Latin clause uses the connective *autem* (each time postverbally); in one case, *thô* corresponds to clause-initial *et*. The two other instances of *thô* are inserted without any Latin correspondence. Betten (1987) shows that it is not possible to link the use of *thô* in the OHG *Tatian* to any particular (discourse) particle used in the Latin source. In the first quarter of the *Tatian* (chapters 1-61) examined by Betten, there are 40 cases, where *thô* corresponds to Lt. *autem*, while there are more than 80 instances of *autem* that fail to be translated in the OHG text. In 48 cases, *thô* is inserted without any Latin model, mostly in clause-initial position. Other Latin discourse particles such as *tunc*, *enim*, or *ergo* fail to be translated by *thô* on any regular basis. Betten (1987: 401) concludes:

"Es zeigt sich somit, dass *thô* zur Wiedergabe argumentativer Partikeln, die im Lateinischen bevorzugt in der wörtlichen Rede eingesetzt werden, keine bedeutende Rolle spielt".

'It becomes clear that *thô* does not play an important role in the rendering of argumentative particles that are used primarily in direct speech in Latin.' (translation: EF)

Moreover, the fact that *thô* is used to translate a number of different Latin elements (or is often used without a Latin model) can be taken to indicate that it was

semantically underspecified. Further support for this assumption comes from cases where *thô* is used in combination with additional temporal or locative adverbials (Axel 2007: 156):⁷⁵

(122) a. inti **uuas** *tho* giheilit ira tohther fon dero ziti
and was then/there healed her daughter from that hour

Lt. [...] & *sanata est filia illius ex illa hora.*

'And her daughter was healed from that hour.'

(*Tatian*, 273,31)

b. senonu *tho* **uuas** man In hierusalem thes namo uuas
INTERJ then/there was man in Jerusalem the.one's name was
gihezzan simeon
called Simeon

Lt. & *ecce homo erat In hierusalem. cui nomen simeon*

'And, lo, there was a man in Jerusalem, whose name was Simeon.'

(*Tatian*, 89,23)

In section 5.2 above, we have already noted that *thô* is often used to fill the clause-initial position in cases where the Latin text exhibits V1 order. This has been taken to suggest that there is a general tendency toward (strict) V2 order in main clauses of early OHG (Dittmer and Dittmer 1998):

(123) **dixit** illi. → *thô* **quad** her imo.
said him then said he him
'then he said to him'
(*Tatian*, 357,1 [106,2]; Dittmer and Dittmer 1998: 92)

Together with its weak semantics, the fact that *thô* is often used as a prefield filler element can perhaps be taken to suggest that *thô* could also function as a precursor of the expletive 'Vorfeld-*es*' that developed later in the MHG period (cf. Brugmann 1917: 37 for a related suggestion; see also Axel 2007: 156). This hypothesis is supported by the fact that there are present-day German dialects which use *do* instead of *es* as a prefield expletive (cf. Weiß 1998: 102):

⁷⁵ Despite its apparently weak semantics, *thô* is sensitive to the temporal setting in which it appears in that it can only appear with preterite verbs, similar to OE *þa* (cf. Lawson 1980).

(124) Do is a *Ungligg* bassierd.
 there is an accident happened
 'An accident has happened.'

Turning now to the discourse function of *thô*, it appears that *thô* is particularly frequent in narrative passages, while it never occurs in direct speech (in contrast to *thanne*, which all in all is much less frequent than *thô*, but can be used to translate Latin discourse particles in direct speech), cf. Betten (1987: 403). Betten characterizes *thô* as a marker of oral style in narratives which is used to mark discourse continuity and achieve text coherence, quite similar to Gothic *þanuh* and OE *þa* (see above). Furthermore, OHG *thô* is often used as a marker that introduces new, foregrounded actions/events along the main story line. Betten stresses that *thô* serves to attract the reader's/listener's attention, and that its discourse function is therefore similar to the function of V1 order:

"[Die Spitzenstellung des Verbs im Aussagesatz] verstkt noch die durch *thô* bewirkte proncierte Hervorhebung des Folgenden. Alles konzentriert sich auf den neu ansetzenden Text. Als das eigentliche Aufmerksamkeitssignal ist jedoch *thô* zu betrachten." (Betten 1987: 404)

'In addition to *thô*, [initial placement of the verb in main declaratives] reinforces the emphasis on what follows. Everything focuses on the newly coming text. However, the actual attention attracting signal is *thô*.' (translation: EF)

Thus, it appears that *thô* was placed in the left periphery to reinforce the original discourse function of V1 order in main declaratives.⁷⁶ Note that *thô* also often occurs in postverbal position in V1 clauses, as for example in (122a) above (see also section 5.3.1 for more relevant examples). This suggests that this word order option reflects the original position of *thô* as an element merged in SpecTP (see section 3.3 on OE),

⁷⁶ Due to its weak semantics and its discourse-related functions, Axel (2007: 169) prefers to analyze *thô* as a sentence particle when occurring in the left periphery of the clause. Furthermore, Axel speculates that the presence of *thô* as a "narrative-declarative" sentence particle blocked EPP-driven XP-fronting in these cases and puts forward the hypothesis that the overall loss of sentence-particles contributed to the rise of generalized XP-fronting.

with V1 order in early OHG being the result of generalized V-to-C movement. V2 orders with *thô* in clause-initial position are then derived by moving the discourse connective to SpecCP.

Furthermore, preverbal placement of *thô* may give rise to V3 order, with a fronted XP occupying the clause-initial position (cf. Axel 2007: 223ff.). Again, it appears that *thô* is not used in its original meaning as a temporal adverb in these examples. Rather, it is to be interpreted as a discourse particle, the interpretation of which seems to be similar to sentence connectives such as *wārlīhō* discussed above. Axel (2007: 225) claims that in the relevant examples, *thô* is used as a contrastive particle that “signals a change in the discourse topic”. We may thus assume that in these examples, *thô* in fact does not occupy a separate position in the left periphery but rather attaches directly to the fronted XP to mark it as a new discourse topic or contrastive focus (similar to present-day German sentence connectives such as ‘aber’, ‘hingegen’ etc.):⁷⁷

(125) a. / [her] tho **antuurtita** inti quad in/
he then answered and said them

Lt. *qui respondens ait eis*;

‘But he answered and said to them [...]’

(*Tatian*, 335,18; Axel 2007: 224)

b. / [...] [siu] tho **giuuanta** sih/
she then turned herself

Lt. / [...] *conuersa illa* /

‘She then turned around.’

(*Tatian*, 665,19; Axel 2007: 224)

Summing up, it appears that the discourse function of fronted *thô* seems to be quite similar to the function of its cognates OE *þa* and Gothic *þanuh*, introducing new foregrounded actions/events or discourse topics along the main story line. Similar to its cognates, *thô* often precedes the finite verb in inversion contexts, although this characteristic is less salient in OHG due to the existence of generalized V-to-C movement. Furthermore, we have seen that *thô* also quite often appears in postverbal

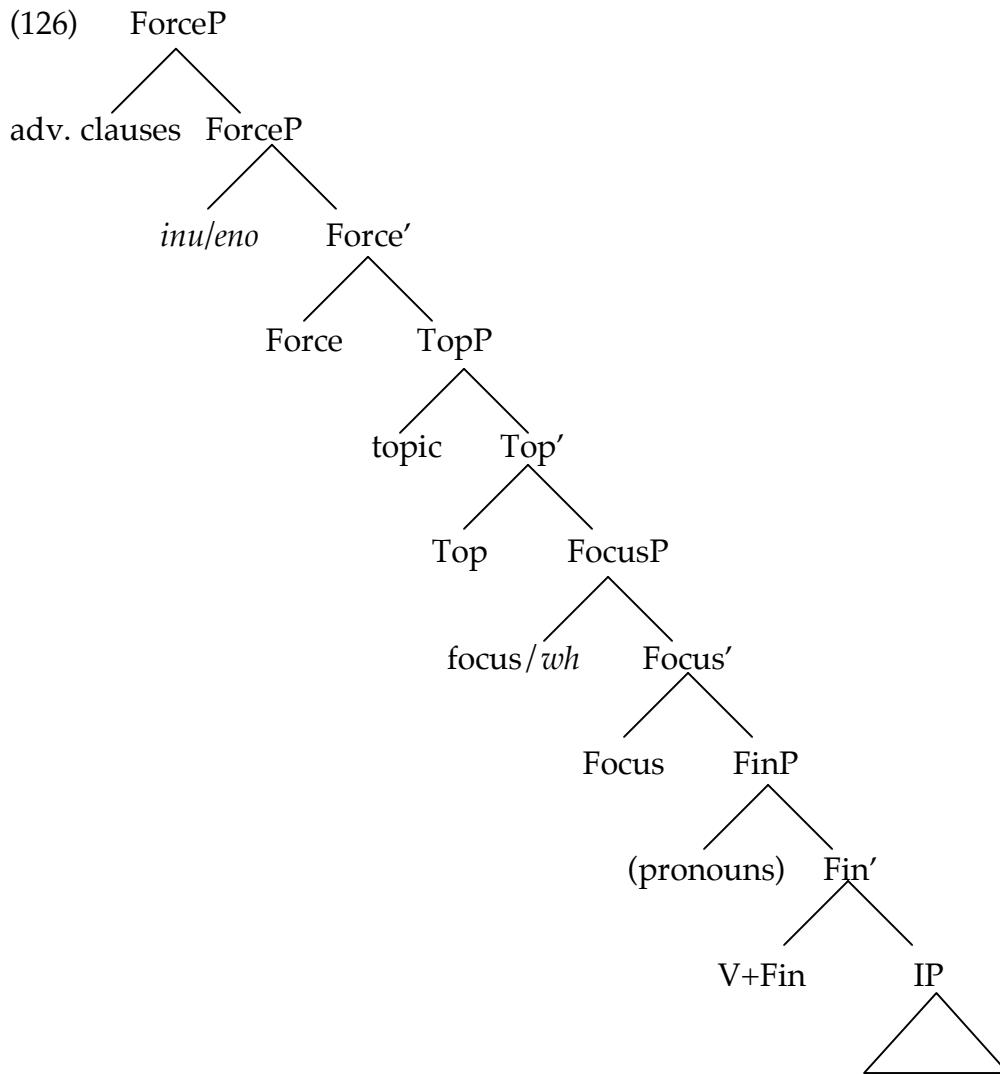
⁷⁷ Other light adverbs that may intervene between a fronted XP and the finite verb include *iu* ‘already’, *thanne* ‘then’, *dhar* ‘there’, *sō* ‘so, thus’, *nū* ‘now’. Since those are much less frequent than *thô*, I chose to focus on the latter here.

position in V1 declaratives, which may be taken to reflect the original position of *thô* in SpecTP, which is crossed by (generalized) V-to-C movement in these examples. See section 5.6.2 below for more discussion of how the frequent (over)use of *thô* (and the concomitant loss of pragmatic/semantic content) may have contributed to the development of generalized V2 in the history of German.

In the next section, I examine the structure of the left clausal periphery of OHG in some more detail, arguing that an analysis in terms of a single head projecting multiple specifiers accounts for the facts more adequately than an analysis in terms of a split CP, *pace* Axel (2007).

5.6 Some remarks on the analysis of V2 in Old High German

The previous sections have shown that while a basic V2 grammar was already quite firmly established in early OHG, there are also some indications that other word order options were still available. In particular, we have seen that there are still a number of examples in which more than a single constituent appears to the left of the finite verb in main clauses. To account for these facts, Axel (2007) proposes an analysis of OHG in terms of a split CP, repeated here for convenience:



According to Axel, this structure collapses into a non-split CP in the course of the OHG, giving rise to the following structure of the left periphery which by assumption still characterizes present-day German:

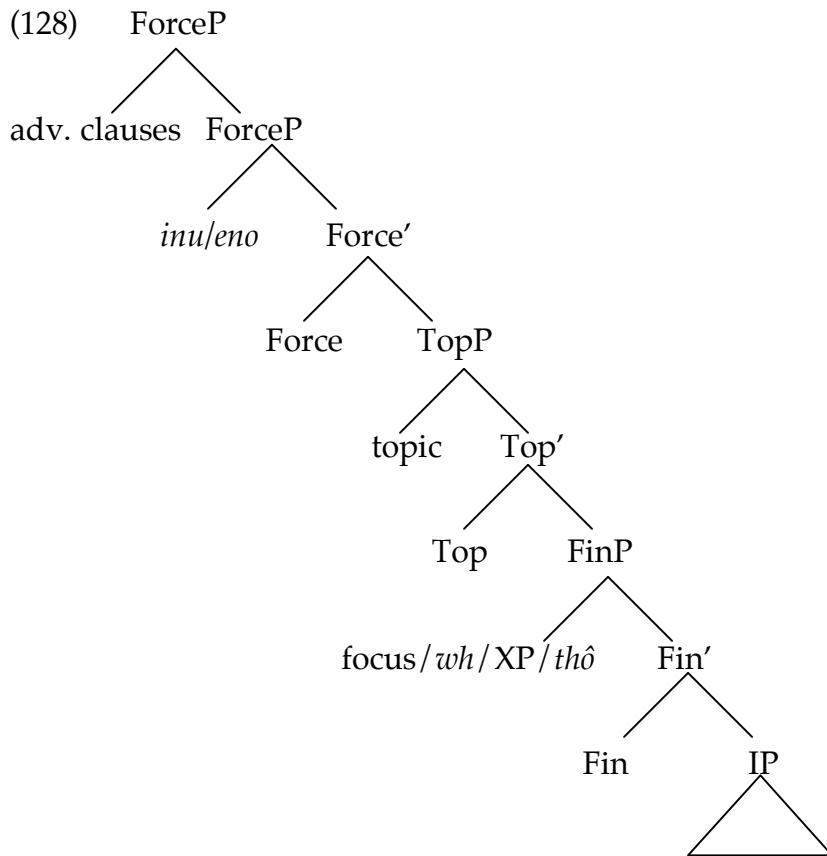
(127) $[_{\text{FinP}} \text{XP}_i [_{\text{Fin}'_j} \text{V}_j + \text{Fin} [\dots \text{t}_i \dots \text{t}_j]]]$

Based on a critical discussion of Axel's analysis, this section proposes an alternative analysis of the left periphery of OHG. I am going to show that the evidence available to us suggests that the C-domain of OHG in fact consists only of a single head that may project multiple specifiers in early OHG. Before we turn to the specifics of this proposal, let me review the account developed by Axel (2007) in some more detail.

First of all, the discussion of pronoun placement in section 5.4 suggests that the structure in (126) must be slightly modified. Recall that I have argued that OHG does not exhibit a designated position for (subject) pronouns in the left clausal periphery. Rather, the relevant V3 orders are more likely the result of an archaic SOV option in

main clauses (presumably an instance of grammar competition in the sense of Kroch 1989). Accordingly, it seems likely that the change which gave rise to (127) (i.e., a multi-purpose position for *wh*-/ focused phrases as well as elements attracted by an EPP feature in Fin) had already been completed by and large in early OHG. Under a split-CP analysis, this means that SpecFinP is now also the position of *thô* (which is presumably directly merged in this position when it is used as a sentence particle). To account for the placement of material occurring to the left of this position, we might posit a higher TopP and ForceP. Furthermore, let's assume that a separate FocusP is only projected if the PLD contains clear evidence in favor of its presence. Note that OHG neither exhibits a set of focus markers (located in Foc) nor other indications that *wh*-/ focused phrases occupy a special separate (specifier) position in the left periphery. Accordingly, the most economical analysis is one where FocP is simply absent from the structure. If we want to stick to the traditional idea that the fronting of *wh*-/ focused phrases is triggered by a relevant set of semantic/pragmatic features, this can be modeled by assuming that these features are optionally hosted by Fin (cf. Fanselow 2002).⁷⁸

⁷⁸ But note that in recent work, Chomsky (2005) denies that such features play a particular role in the syntactic derivation. Instead, it is assumed that the fronting of *wh*-/ focused phrases is simply triggered by EPP-features, with the relevant interpretative effects relegated to the interfaces.



In what follows, it is shown that even this 'slimmer' revised split CP analysis raises a number of problems. My argument is threefold. First, I argue that the split CP structure is incompatible with the assumption that adverbial clauses are adjoined to the root node. Second, I show that there is no clear evidence for more than a single head position in the C-domain of OHG. Third, I will point out a number of conceptual problems that follow from the assumption that the split CP is conflated into a simplex CP in the course of OHG.

First of all, note that a split-CP analysis seems to be incompatible with the assumption that material may be adjoined to the root node, which is proposed in Axel (2007) to account for the position of adverbial clauses and (albeit somewhat less explicitly) dislocated topics. Rather, cartographic approaches usually allow only a single specifier per projection, ruling out adjunction to phrasal categories (basically following Kayne 1994).⁷⁹

⁷⁹ Note that if multiple specifiers or XP-adjunction were allowed in the cartographic approach, most arguments put forward by Rizzi (1997) in support of a split CP would not go through. Furthermore, assuming both multiple projections and multiple specifiers/adjunction positions would give rise to an inflated theory of phrase structure that combines otherwise complementary theoretical assumptions, see e.g. Lahne (2007) for some discussion.

Second, note that apart from complementizers, which typically appear in the lowest head position of the C-domain in OHG (directly above IP/TP, i.e., in the position occupied by the finite verb in main clauses), no further projecting head can be observed in the left periphery of the clause. In particular, while there are still some residues of an older (presumably Pan-Germanic) system of sentence particles (including the affirmative particle *jā*, the interrogative particle *inu/eno*, and possibly the narrative marker *thō*), it can be shown that these elements do not spell-out head positions in the C-domain. Rather, the fact that they seem to interact with XP-movement rather than with verb fronting (cf. Axel 2007) suggests that they occupy specifier positions in the C-domain (a similar conclusion is reached by Axel 2007). Moreover, there are examples with more than a single sentence particle in the left periphery:

(129) a. eno nu ia sint zuelif citi thes tages?

PRT PRT PRT are twelve hours of-the day

Lt. *nonne XII hore sunt diei?*

'Are there not twelve hours of daylight?'

(Tatian, [135,5]; TITUS)

b. / [...] eno ia uurdun zéheni gihéilte/

prt prt were ten healed

Lt. [...] *nonne decem mundati sunt*

'Were the ten healed?'

(Tatian 379,10; Axel 2007: 44)

Axel (2007: 209ff.) assumes that *inu/eno* occupies SpecForceP, while *ia* is located in SpecFinP. According to Axel, the latter assumption is supported by the fact that *ia* usually occurs directly to the left of the (inverted) finite verb. However, this implies that the particle *nu* occupies a further specifier position in examples like (129a), possibly SpecTopP if structure (128) is adopted for OHG. Note that this outcome is not motivated by any independent considerations (why should a discourse particle such as *nu* 'now' be placed in a position normally reserved for topical material?), but rather follows automatically from the logic of an approach in terms of multiple projections.⁸⁰

⁸⁰ Of course, it is always possible to posit a separate projection that hosts *nu*. But such an assumption is clearly quite ad hoc as long as it is not based on any further evidence motivating the existence of such

Third, recall that an analysis in terms of multiple projections must assume that either the previously split CP has been conflated into a simplex CP (as is proposed in Axel 2007), or that modern Germanic V2 languages still have a split CP, but have developed some special restrictions that rule out multiple XP fronting (for concrete proposals cf. e.g. Grewendorf 2002, Frey 2004, 2006; see Fanselow 2002, 2004, 2006 for some discussion). Neither possibility seems to be particularly attractive. As to the conflation hypothesis, Axel (2007: 235) concedes that “it is unclear how this process was triggered.” The same goes for the hypothetical historical development of restrictions ruling out V3 orders in a split CP. Most of the relevant proposals to present-day German involve some sort of hard-wired locality restriction in the spirit of Relativized Minimality (Rizzi 1990), or the Minimal Link Condition (Chomsky 1995). For example, Grewendorf (2002) proposes that the presence of an EPP feature in Fin blocks further (overt) movement to higher specifiers of TopP or FocP (since the XP in SpecFinP creates an intervention effect).⁸¹ Alternatively, one might invoke the Minimal Link Condition to delimit EPP-driven movement into the C-domain to the highest constituent in the midfield (cf. Frey 2006, Fanselow 2002, 2004). While these approaches serve to model the facts in present-day German, it is hard to see how such restrictions can develop historically, in particular if a multiple projections model is adopted that licensed multiple fronting in earlier stages of Germanic.

In what follows, I will develop an alternative analysis of the left periphery in early OHG and the changes that took place in the course of the OHG period based on the assumption of multiple specifiers instead of multiple projections (cf. Chomsky 1995, 2001 and subsequent work).

5.6.1 A multiple specifiers approach to V2 in Old High German

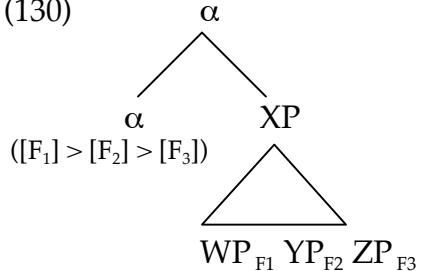
An alternative to an analysis of multiple XP-fronting in terms of multiple projections is to assume that the C-domain is made up by only a single functional head (C), which may project multiple specifiers hosting fronted XPs, or particles directly

a projection (e.g., further word order facts concerning the placement of *nu* relative to other fronted elements, or considerations concerning the semantic impact of the relevant functional projection).

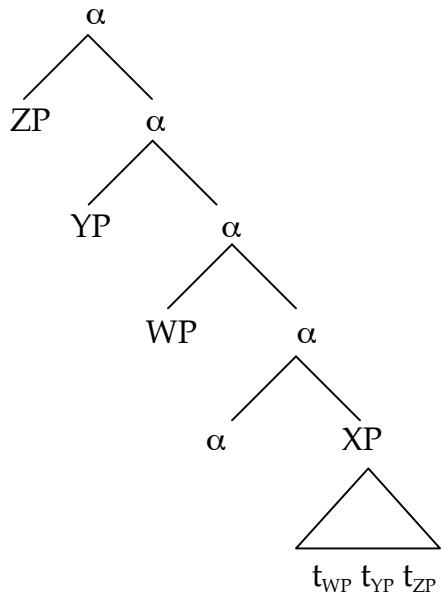
⁸¹ Note that this proposal raises a couple of questions, since Grewendorf assumes that movement from SpecFinP to SpecTopP is in principle possible (e.g., in cases of left dislocation in German).

merged in the left clausal periphery.⁸² This approach implies that semantic/pragmatic features triggering Merge operations are hierarchically ordered (cf. e.g. Grewendorf and Sabel 1999 on scrambling, and Lahne 2007 for the structure of the left periphery), ensuring that the relevant features must be checked off in a certain order.⁸³ Furthermore, the outcome of the movement (i.e., internal Merge) operations triggered by these features must reflect the feature hierarchy. More precisely, let us assume that given a functional head α with the feature hierarchy $[F_1] > [F_2] > \dots > [F_n]$, $[F_1]$ first triggers second Merge creating the closest specifier of α . Subsequently, $[F_2]$ triggers third Merge creating an outer specifier and so on. In other words, higher specifiers correspond to features lower in the hierarchy (abstracting away from the possibility of 'tucking in', Richards 2001):

(130)



(131)



⁸² See Lahne (2007) for a number of conceptual advantages of an approach to the left clausal periphery in terms of multiple specifiers.

⁸³ The relevant feature hierarchy for a given functional head is presumably determined by (semantic) conditions holding at the interfaces, in the sense that a 'wrong' hierarchy of specifiers hosting the relevant elements could not be interpreted at the interface to C-I.

In more formal terms, this can be expressed by the following condition (Lahne 2007: 10):

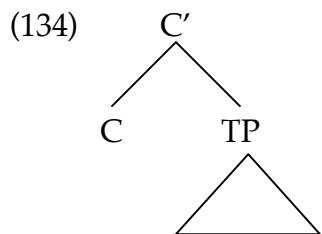
(132) *Condition on hierarchy-driven derivation*

- a. A feature $[F]$ of a head α is to be satisfied at a point P of the derivation iff (i) and (ii):
 - (i) α is the active head.
 - (ii) $[F]$ is the active feature.
- b. **Active head**
A head is active at a point P of the derivation iff it is a probe at P .
- c. **Active feature**
A feature is active at a point P of the derivation iff it is the highest unsatisfied (unchecked/unvalued) feature in the feature hierarchy of an active head at P .

Thus, at any point during the syntactic derivation, syntactic operations may only be triggered by the active feature of an active head. Under these assumptions, the ordering restrictions that we have observed in the left periphery of OHG (i.e., Topic-Focus/*wh*- V_{fin} -...) can be derived as follows. Let us take a look at the derivation of (100), repeated here for convenience:

(133) [Uuexsal dhes nemin] huuazs **bauhnida?**
changing-NOM of-the name what meant
Lt. *Mutatio nominis quid significabat?*
'The changing of the name, what did it mean?'
(*Isidor*, 532; Axel 2007: 209)

Focusing on the left clausal periphery, first Merge of C with TP creates the structure in (134) (using traditional X'-notation for convenience):

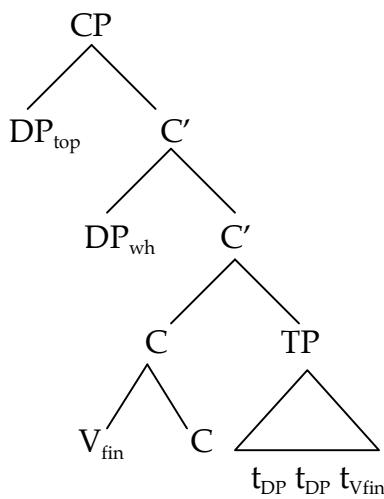


Adopting the notational convention that features assigned a diacritic *_* require overt movement/PF realization (cf. Roberts and Roussou 2003, Sternefeld 2007), (133) can be derived by assuming that C hosts the following features: $[\text{*fin}/\text{_V}^*]$, which requires attraction of a finite element of the category V (cf. Lahne 2007),⁸⁴ $[\text{*wh}^*]$, and $[\text{*top}^*]$, ranked according to the following hierarchy:⁸⁵

(135) $[\text{*fin}/\text{_V}^*] > [\text{*wh}^*] > [\text{*top}^*]$

Thus, C must first attract the finite verb. Subsequently, a *wh*-specifier and a topic specifier are added by recursive applications of Merge:

(136)



However, while this analysis successfully derives the order found in (133), it fails to capture the fact that already early OHG exhibited generalized V2 effects, which we have taken as an indication that C was already endowed with a semantically vacuous EPP feature that could attract all kinds of phrases. In particular, we have noted that there already existed quite a strong tendency to place only a single constituent in the prefield. Both these facts, that is, fronting independent of semantic/pragmatic

⁸⁴ Recall that OHG already exhibited generalized V-to-C movement.

⁸⁵ As already briefly noted above, the ranking of semantic/pragmatic features is presumably determined by properties of the interfaces. Furthermore, note that the ranking of semantic/pragmatic features in functional heads represent the reverse of what presumably holds at the interface to C-I, with 'lower' functional features giving rise to higher specifiers (see also Müller 2007). In addition, morphosyntactic features seem to have primacy over 'peripheral' semantic/pragmatic features (i.e., must be satisfied first). This might have to do with the fact that unvalued/unchecked morphosyntactic features lead to a crashing derivation, while unvalued/unchecked semantic/pragmatic features probably merely give rise to deviant interpretations.

factors, and the strong tendency to have only a single prefinite constituent, are kind of unexpected if movement is triggered directly by 'strong' (i.e., starred) semantic/pragmatic features. Rather, the latter is presumably a characteristic of discourse-configurational languages, and we would predict semantically/pragmatically driven V3 orders to be much more common if XP-fronting in OHG were discourse-configurational along these lines. In other words, the basic question we have to address is how we can formally model the 'early beginnings' of generalized V2 in OHG.

Of course, an important ingredient of a relevant approach is the presence of an semantically vacuous EPP feature in C, which is not linked to a particular semantic/pragmatic feature. Following ideas put forward in Simpson (2004) on the historical origin of EPP features, we might speculate that the EPP feature in C developed historically when the original motivation for fronting became non-transparent to the learner, who in turn posited a semantically vacuous feature in C to mimic the relevant orders that are part of the PLD. This possibility can be attributed to a mechanism devised by Chomsky (2000) to the effect that semantically vacuous, structure-building EPP features (or, edge features, Chomsky 2005) may be optionally added to phase heads, possibly in the course of the syntactic derivation (Chomsky 2000, 2001, 2005; cf. Müller 2007 for discussion):

(137) The head H of phase Ph may be assigned an EPP-feature (Chomsky 2000: 109)

Thus, generalized V2 effects due to XP-fronting can be analyzed in terms of assigning an EPP feature to C. In contrast to Modern German, however, this procedure has not yet been fully generalized in OHG (in contrast to V-to-C movement), as evidenced by the (frequent) possibility of V1-declaratives.⁸⁶ Moreover, independent of whether C hosted an EPP feature or not, the feature [*wh*] required overt movement of *wh*-phrases, which is a property found in all early Germanic languages, including Gothic ('operator V2', see section 2 above). This suggests that in OHG, XP-fronting could be triggered by either 'strong' (i.e., starred) semantic/pragmatic features or by an EPP-

⁸⁶ Note that the assignment of EPP-features is probably governed by lexical features of C. For example, an EPP-feature is usually not assigned in yes-no questions and imperatives. In a similar vein, one might suspect that the special narrative function of V1 order was encoded by a lexical feature of C (e.g., [+narrative]) that prevented assignment of an EPP-feature (alternatively, one might assume that all these constructions involve an empty operator in SpecCP that checks the relevant EPP-feature).

feature optionally added to C. Given that both options are in principle available, this of course raises the question of how EPP-features and starred features interact in cases of multiple XP-fronting. In particular, we must ensure that certain elements appear closer to the finite verb, while others occupy outer specifiers of CP. To these ends, let's take another look at the generalizations stated in Axel (2007) on the relative order of elements in the left periphery of OHG. Note that the following statements are slightly modified, that is, they do not include a special position for pronouns (cf. section 5.4 above) but explicitly mention the narrative discourse marker *thô*:

(138) a. adverbial clause > *inu/eno* > disloc. topic > *wh* > V_{fin} ... (interrogatives)
 b. adverbial clause > disloc. topic > $XP / thô$ > V_{fin} ... (declaratives)

Let's start with the first specifier of CP created by second Merge. In cases where C has been endowed with an EPP feature, it attracts the highest XP in the midfield (as a result of the MLC/Attract Closest Principle, cf. Chomsky 1995, 2000; cf. Fanselow 2002, 2004, Frey 2004, 2006 for details of this particular analysis of V2 effects). In many cases, this element is the discourse-continuative marker *thô*, which often appears directly to the right of the finite verb in V1-declaratives and moves to prefinite position if declarative C is endowed with an EPP feature (as argued in section 3.3 above, discourse-continuative 'then' (*thô* in OHG) was originally inserted in SpecTP in early Germanic):

(139) a. $[V + C_{[-EPP]} [_{TP} thô \dots t_V]] \rightarrow V1 \text{ declarative}$
 b. $[_{CP} thô [V + C_{[+EPP]} [_{TP} t_{thô} \dots t_V]] \rightarrow V2 \text{ declarative}$

The fact that left-peripheral *thô* seems to have a similar interpretation in V1 and V2 clauses suggests that we deal with a single element that may occupy different positions depending on the feature content of C. In particular, I assume that in early OHG, the discourse-continuative, sequential meaning of *thô* was still associated with SpecTP, similar to OE (and presumably Gothic). V1 order with postfinite *thô* is then the result of generalized V-to-C movement innovated by the (Western) Continental branch of Germanic (a development which did not take place in OE). Furthermore, the additional innovation of a semantically vacuous EPP feature in C led to V2 orders with prefinite *thô*. Due to the fact that C's EPP feature is not linked to any peripheral

semantic/pragmatic feature (such as [top] or [foc]), the original meaning of *thô* linked to SpecTP is preserved after movement to SpecCP. The variation concerning the presence vs. absence of an EPP feature in C can be analyzed as an instance of grammar competition (Kroch 1989), with the parametric option [-EPP] disappearing in the course of OHG. In case the numeration does not contain a discourse-continuative element such as *thô*, another XP occupying the highest position in the midfield (including indefinites) may be fronted to satisfy C's EPP feature.⁸⁷ In contrast, in *wh*-questions the first specifier is systematically occupied by a fronted *wh*-phrase. While this directly follows from the feature hierarchy proposed above, we still might ask how EPP-driven movement and movement triggered by peripheral features such as [*wh*] interact since it is at least at first sight not clear which operation should have preference over the other. As already pointed out in fn. 85, one might argue that the checking/valuation of morphosyntactic features is imperative, since if unvalued, these constitute genuine uninterpretable features that cause a derivation to crash at both interfaces. So for the time being, let's adopt the preliminary assumption that the semantically vacuous EPP feature must be checked first (presumably after verb movement has taken place, although nothing hinges on that particular choice):

(140) [*fin/_V*] > [EPP] > [*wh*] > [*top*]

Now, this raises the question of how we can account for the fact that *wh*-phrases obligatorily occupy the closest specifier of C in OHG. More to the point, one might perhaps expect that EPP-checking by an element such as *thô* preempts *wh*-movement to the closest specifier of C due the fact that [EPP] is ranked higher than [*wh*]. It is a fact, however, that this apparently never happens in OHG, that is, fronted *wh*-phrases systematically appear directly to the left of the finite verb in main clauses (Petrova and Solf 2007; see *ibid.* and section 5.4 above for discussion of apparent counter-examples).⁸⁸ I suppose that the distribution of *wh*-phrases can be attributed to the principle *Maximize matching effects* (Chomsky 2001: 15). Note that fronting of

⁸⁷ Note that there also cases where *thô* occupies a lower position in the midfield. In these cases, *thô* is used in its original temporal meaning (i.e., mostly cotemporal 'then').

⁸⁸ As pointed out in section 3.3.3 above for OE, the fact that discourse-continuative, sequential *thô* does not co-occur with *wh*-phrases (in the left periphery) might also be due to the fact that the discourse function of *thô* is not compatible with the pragmatic status/discourse function of *wh*-questions.

wh-phrases serves to check/value both C's EPP feature and its [*wh*] feature, while purely EPP-driven fronting merely checks a single feature. In other words, the set of features checked/valued by fronting of *wh*-phrases is a superset of the features checked by semantically vacuous movement solely triggered by C's EPP. Now, if there are two goals that are in principle accessible to probe C, C will attract only that element that serves to check/values the greatest subset of features contained in C. In the case at hand, both goals are part of the same phase and therefore equidistant to C (after previous movement of *wh* to the left edge of vP). Moreover, elements such as *thô* presumably do not create an intervention effect for *wh*-movement. Accordingly, C will attract the *wh*-phrase, leaving *thô* in a lower, postfinite position. Furthermore, it is not possible to raise *thô* to a second/outer specifier, since *thô* can only be attracted by a semantically vacuous EPP feature (which has already been checked by *wh*-movement in the case at hand): *thô* does not constitute a possible topic (the same goes for other indefinites) and therefore cannot be attracted by a feature such as [*top*]. This prediction is borne out by the facts: There are apparently no cases where a fronted *wh*-phrase is preceded by a non-topic such as *thô* (cf. Petrova and Solf 2007).

Next, let's take a look at the apparent possibility of fronting a topic to the left of *wh*, as in (133) above. This can be attributed to the presence of a strong [*top*] feature in C. If the C head hosts the features [EPP], [*wh*], and [*top*], *wh*-fronting checks both [EPP] and [*wh*], and [*top*] may be checked by a later operation raising a topical element to an outer specifier of C. Note that the feature hierarchy in (140) correctly rules out a derivation where the order of operations is reversed, with the topic raising to the closest specifier of C (to check [EPP] and [*top*]), and *wh*-fronting targeting an outer specifier to check [*wh*].

Another set of examples with multiple fronting includes cases like the following from the OHG *Isidor*, where V3 orders of the type XP-XP-V_{fin} were (still) more productive than in other OHG texts (cf. Robinson 1997):

(141) a. [Dhea uehhun] [auur] [in heilegim quhidim] **arfullant** sibun iaar.
the weeks however in sacred language fulfil seven years

Lt. *Ebdomada namque in sacris eloquiis septem annis terminatur.*

'The weeks, however, take seven years in sacred language.'

(*Isidor*, 457; Robinson 1997: 26)

b. [So] [auh in andreru stedi] [dhurah dhen selbun heilegum forasagun]
 so also in other places through the same holy prophet
uuard dhera dhrinissa bauhnunc sus araughit: [...]
 became the-GEN Trinity-GEN meaning in this way demonstrated

Lt. *Item alibi per eundem prophetam trinitatis sic demonstratur significantia: [...]*

'In this way, also elsewhere the meaning of the Trinity was demonstrated by the same holy prophet: [...]'

(*Isidor*, 328; Robinson 1997: 27)

In (141a), it can be shown that the two fronted XPs have a different information-structural status. While *dhea uehhun* refers back to a known information previously mentioned (in the previous clause),⁸⁹ *in heilegim quhidim* clearly is a contrastive focus. This particular order is reminiscent of the clause structure of Indo-European/Proto-Germanic proposed in Kiparsky (1995), where the left periphery is taken to include a focus position (hosting operators and in particular *wh*-phrases) and a higher topic position. Examples such as (141a) can then be accounted for if we slightly modify the feature hierarchy in (140), replacing the feature [*wh*] by the more general feature [*foc*] (see e.g. Sabel 1998 for the idea that *wh*-movement is triggered by a focus feature):

(142) [*fin/_V*] > [EPP] > [*foc*] > [*top*]

As a result, *in heilegim quhidim* moves to the first specifier of C to check both the EPP and the [*foc*] feature. In a subsequent operation, *dhea uehhun* is merged in the outer SpecCP to check [*top*]. The most likely interpretation of (141b), on the other hand, is that it involves multiple fronting of topics. This can be modeled by the

⁸⁹ Compare the relevant passage from the OHG *Isidor*:

Chiuuisso nu, ibu dhea sibunzo uehhono fona daniheles zide uuerdhant chizelido, buuzzsan einigan zuuiuun ist dhanne archennit, dhazs dher allero heilegono heilego druhtin nerrendeo christ iu ist langhe quhoman. Dhea uehhun auur in heilegim quhidim arfullant sibun iaar.

Lt. *Quę scilicet LXX ebdomadę, si a tempore danielis numerentur, procul dubio sanctus sanctorum dominus iesus christus olim uenisse cognoscitur. Ebdomada namque in sacris eloquiis septem annis terminatur.*

'Certainly now, if the 70 weeks are counted from Daniel's time on, it is without doubt that the holiest of the holy, Christ the Lord has already come. The weeks, however, take seven years in sacred language.'

(*Isidor*, 453-457; TITUS)

assumption that we can have more than a single [*top*] feature each instance of which must be checked by a separate move operation (see e.g. Rizzi 1997 for the observation that topics can be iterated).⁹⁰

It should be noted, however, that V3 orders of the type exemplified in (141) were not very frequent in early OHG; they soon disappeared in the course of OHG. In late OHG, we can already observe strict V2 order in main declaratives (cf. Axel 2007). In fact, relevant V3 orders are attested primarily in the OHG *Isidor* (cf. Robinson 1997, Axel 2007). In other words, the presence of a [*top*] feature in C seems to be an archaic trait which is already in decline in early OHG. Thus, the loss of V3 orders can be taken to be part of a more general development in which fronting to clause-initial position lost its original discourse-related, semantic/pragmatic motivation in the history of German. The fact that fronted XPs may still be interpreted as topics can be modeled by assuming that C may still contain a [top] feature that is no longer able to trigger movement on its own. As a result, fronting of XPs interpreted as topics is accomplished by C's EPP feature in later stages of German, leading to the loss of V3 effects in connection with fronted topics (but see Chomsky 2005 and Fanselow 2006 for critical discussion of the assumption that semantic/pragmatic features such as [top] are directly encoded in the syntax).

A further question concerning the relation between topicalization and EPP-driven fronting is raised by the following V3 examples, in which a fronted pronoun appears to the left of *thô*, which by assumption could only have been attracted by C's EPP-feature:

(143) a. / [her] tho **antuurtita** inti quad in/
he then answered and said them

Lt. *qui respondens ait eis;*

'But he answered and said to them [...]'

(*Tatian*, 335,18; Axel 2007: 224)

⁹⁰ Thus, it must be ruled out that *Maximize matching effects* leads to multiple checking of both [*top*] features. This can be achieved either by a uniqueness principle ensuring that a single element can maximally satisfy a single substantial semantic/pragmatic feature, or by a more fine-grained distinction between different kinds of [*top*] features linked to the well-known distinction between aboutness topics, familiarity topics etc.

b. / [...] [siu] tho **giuuanta** sih/
she then turned herself

Lt. / [...] *conuersa illa*/
‘She then turned around.’

(*Tatian*, 665,19; Axel 2007: 224)

At first sight, it seems that these examples should be ruled out by *Maximize matching effects*: Since fronting of the pronoun serves to check both C’s EPP feature and its [top] feature (independent of whether [top] is starred or not), we should expect that the pronoun occupies the first specifier of C, with *tho* staying behind in postfinite position. However, recall that in examples like (143), *tho* is in fact to be interpreted as a sentence connective which signals a change in discourse topic and is attached to the right of a fronted XP (cf. section 5.5 above). Thus, similar to examples where sentence adverbs such as *giwisso* ‘certainly’ or *wārlīh(h)o* ‘truly, really, indeed’ intervene between a fronted XP and the finite verb, cases like (143) do not constitute genuine violations of the V2 constraint (similar to apparent V3 orders of the type XP-*aber/hingegen/etc.-V_{fin}* in present-day German).

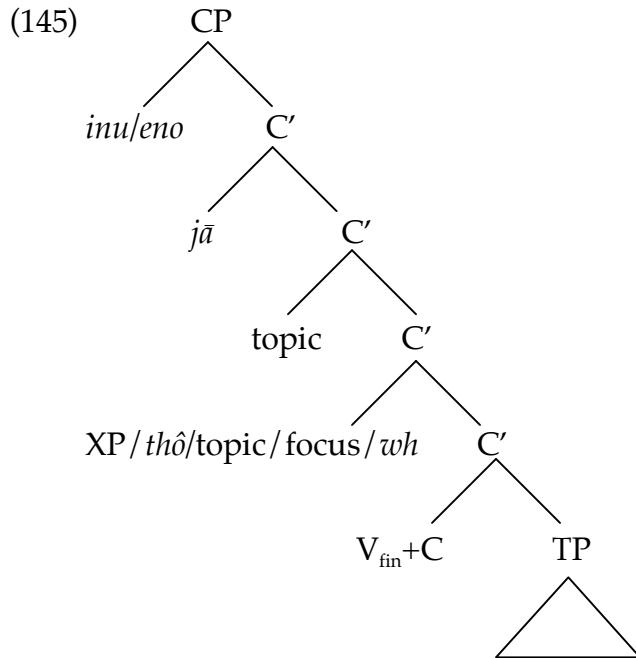
At this point, there are still two of Axel’s generalizations left that require an alternative explanation in terms of the multiple specifiers approach devised here. First, we have to account for the high position of sentence particles such as interrogative *inu/eno* and affirmative *ja*. Second, we have to provide an explanation for the peculiar placement properties of fronted adverbial clauses. Concerning the latter, recall that fronted adverbial clauses occur at the leftmost position in both main and embedded clauses, preceding all other material, including complementizers, and sometimes even coordinating conjunctions. Axel (2007) analyzes the distribution of fronted adverbial clauses in terms of adjunction to the root node (i.e., ForceP in her model). However, it has already been noted that this account is not very attractive, in particular under an approach to the left periphery in terms of multiple projections. Alternatively, we may adopt proposals by Nissenbaum (2000) and Chomsky (2004) to account for the placement of adverbial clauses in OHG. Nissenbaum and Chomsky argue that adjuncts are introduced by countercyclic late Merge which applies (optionally) to a given syntactic object at the point of Spell-Out/Transfer. As a result, adjuncts can only appear at the linear edge of a Spell-Out Domain. This appears to give the correct results for the placement of adverbial clauses in OHG and serves to model the fact that these clauses are apparently not fully integrated in their

matrix clause (although a couple of questions remain, concerning e.g. the exact size of Spell-Out domains, and the left/right orientation of elements introduced by late Merge; see chapter 2 for some discussion).

Concerning the placement of particles such as *inu/eno*, recall that OHG exhibits the residue of a formerly more refined system of left-peripheral particles linked to discourse properties and clausal typing. This can be modeled by expanding our feature hierarchy by adding two more features linked to the coding of affirmativity and the typing of all kinds of interrogatives. This gives us the following (final) feature hierarchy for the content of C in early OHG (note that I also included the variation between strong and weak [top] that can be observed in OHG, again presumably an instance of grammar competition):

(144) [*fin/_V*] > [EPP] > [*foc*] > [*top*]/[top] > [*affirm*] > [*interrog*]

Schematically, this hierarchy translates into the following phrase marker (however, note that it is in fact quite unlikely that all positions can be filled in a single sentence):



As becomes clear from (145), apart from the absence/presence of multiple head positions in the left periphery, the main difference between an approach in terms of multiple projections and an approach making use of multiple specifiers concerns the specifier which is closest to the position of the finite verb. While in a multiple projections analysis, this position is typically identified as SpecFinP, possibly

reserved for pronouns and non-topic/non-focus elements attracted by C's EPP feature (with other elements targeting higher specifiers in early OHG, cf. Axel 2007), it turns into a multi-purpose position in a structure like (145). The actual content of the specifier closest to C is then determined by (i) the feature hierarchy proposed in (144), (ii) the actual feature content of C in each individual sentence, and (iii) the interaction between C's EPP feature and other substantial semantic/pragmatic features in terms of *Maximize matching effects*.⁹¹ Note that the multi-purpose character of this position carries over to present-day German (cf. e.g. Fanselow 2004, 2006), the only difference being that in present-day German, C has apparently lost its ability to project more than a single specifier. In the next section, I argue that the loss of multiple positions in the *Vorfeld* can be analyzed in terms of a simple parametric change, which further supports the analysis of the left periphery of OHG developed in this section.

5.6.2 The loss of V3 orders in OHG

This section focuses on the emergence of a strict V2 grammar in the course of the OHG period. I show that the two approaches discussed so far (i.e., multiple projections vs. multiple specifiers) differ significantly when it comes to modeling the

⁹¹ Above, we have noted that there is a crucial difference between OE and OHG concerning the availability of $XP-DP_{subj}-V_{fin}$ order in main clauses. While this type of V3 order is quite frequent in OE (around 28% of all relevant clauses according to Haeberli 2000), it is practically absent from the OHG record. In section 3.3.3, I have argued that the relevant OE examples presumably involved movement of topical DP subjects to SpecTP (with the fronted XP occupying SpecCP), which anticipated the structure of later stages of English (furthermore, note that van Gelderen 2004 claims that OE did not exhibit multiple positions in the left periphery. If she is right, then the relevant OE examples cannot be attributed to multiple fronting to the C-domain). In contrast, the derivation of $XP-DP_{subj}-V_{fin}$ orders was significantly obstructed by changes that had taken place by the time the earliest OHG records were written. After the rise of generalized V-to-C movement and the development of an (optional) EPP-feature in C, the order $XP-DP_{subj}-V_{fin}$ could only be derived by the presence of a [*top*] feature in C that attracts another XP in addition to the subject (which for reasons of locality is raised to the first SpecCP to check C's EPP and perhaps some other substantial feature). A further complication is created by the fact that subjects are usually highly topical (cf. e.g. Givón 1976). As a result, the availability of the order $XP-Subj.-V_{fin}$ reduces to cases where C contains either a [*top*] and a [*foc*] feature (with the subject being a focus) or two [*top*] features (with the subject being a topic). Recall, however, that the possibility of endowing C with a strong [*top*] feature was apparently already in decline in early OHG. As a result, we do not expect to find $XP-DP_{subj}-V_{fin}$ orders in significant numbers in early OHG, which is borne out by the facts.

diachrony of V2 in OHG. In particular, it will become clear that an approach that assumes only a single head in the C-domain can capture the relevant historical developments in terms of a simple parametric change, namely the loss of multiple specifiers.

We already know that V2 order was generalized in the course of OHG. In late OHG texts, we typically find consistent V2 and inversion in case a non-subject is fronted to clause-initial position in all kinds of main clauses (apart from imperatives and yes/no questions), cf. Behaghel (1932), Näf (1979), Lenerz (1984), and more recently Schrodt (2004: 204f.) and Axel (2007: 200f.) (see Petrova and Solf 2007 for V2 in interrogatives). As already mentioned, Axel (2007) assumes that generalized V2 evolved when the previously split CP turned into a simplex CP in the course of OHG, although “it is unclear how this process was triggered” (Axel 2007: 235).

In contrast, an approach in terms of multiple specifiers can attribute the rise of generalized V2 to a simple parametric change, namely the loss of the availability of multiple specifiers in the CP. In what follows, I am going to argue that this development was triggered by two factors that are both related to the presence of a generalized EPP-feature in C: First, at some point, the original semantic/pragmatic function of XP-fronting (apart from operator contexts) became opaque. As a result, learners posited a semantically vacuous EPP-feature in C to mimic the relevant orders (cf. Simpson 2004). In addition, the rise of an EPP-feature was promoted by examples with clause-initial *thô* (see below for details) Second, when the original discourse function of clause-initial *thô* became unclear, the discourse connective was reanalyzed as a semantically light expletive-like element that is directly merged in SpecCP. There are reasons to believe that the latter change was decisive for the development of generalized V2 in German. In particular, I am going to argue that the presence of expletives signals to the learner that a functional head may project only a single specifier, which led to the loss of V3 orders in late OHG. However, before I turn to the relevant developments that took place in the recorded span of OHG, let me first add some (admittedly speculative) remarks on the pre-OHG rise of an EPP feature in C.

Axel (2007: 169, 235) conjectures that both the rise of an EPP feature (giving rise to generalized XP-fronting) and the development of a simplex CP were linked to the loss of an older system of left peripheral sentence particles (including *inu/eno*, *ja*, and *thô*) that were used for various discourse-related purposes including clausal typing. According to this scenario, C’s EPP feature developed when different sentence types

(including the main/embedded distinction) could no longer be identified by lexical means (via Merge of sentence particles): “As a consequence, the distinction of sentence types was primarily signalled by means of word order properties (verb-first, verb-second), instead of by sentence-typing particles.” (Axel 2007: 235). In addition, the loss of sentence particles is taken to have contributed to the conflation of the formerly split CP via blurring the evidence for the presence of different structural positions in the C-domain. However, while it seems likely that the loss of sentence particles did play a role in the diachronic development of Germanic syntax (note that the loss of sentence particles is also commonly taken to have instigated the rise of generalized verb movement in Germanic, cf. Eyþórsson 1995, Roberts 1996, Ferraresi 1997), it is not clear whether the changes in question can be attributed to the loss of particles alone. First of all, the hypothesis that XP-fronting developed as an alternative strategy of sentence typing has a somewhat functionalist flavor to it; it is not clear whether language learners really have the ability to carry out such changes to ‘repair’ functional deficits of the target grammar. From a theoretically informed perspective in terms of grammar change (see chapter 1), it would certainly be more desirable if we could identify alternative (formal) factors that contributed to this change in the course of language acquisition. For example, we might assume that after the loss of sentence particles, systematic word order differences between main and embedded clauses were mistakenly associated with clause type features in main C that required PF realization via XP-fronting (perhaps on the analogy of clausal typing by complementizers in embedded clauses, cf. Kiparsky 1995). However, note that in early Germanic, the differences between main and embedded clauses in terms of word order were less clear-cut than they are in later stages of German (apart from the presence/absence of complementizers). In other words, it is not clear whether these differences could really be exploited to ascribe a clause-typing function to XP-fronting. Moreover, it seems that this explanation does not carry over straightforwardly to the rise of V2 in Northern Germanic, where we apparently find symmetric V2 in both main and embedded clauses from quite early on (cf. e.g. Eyþórsson 1995, Rögnvaldsson 1996, and Faarlund 2004 on Old Norse).⁹² In a similar vein, it is doubtful whether the loss of sentence particles can be linked in any systematic way to the loss of a split CP. Note that there are many languages such as

⁹² Recently, however, Dewey (2007: 88f.) has argued that at least in early Northern Germanic poetic texts (i.e., the Eddic corpus), the finite verb occupies clause-final position in “the vast majority of bound clauses”.

French, Italian or English which are commonly analyzed as exhibiting a split-CP despite the fact that they lack a refined system of sentence particles in the left clausal periphery. In what follows, I will therefore explore an alternative explanation of the loss of V3 orders which is based on the analysis in terms of multiple specifiers proposed in the previous section. First, let's take a look at factors that might have contributed to the development of a semantically vacuous EPP feature in C (in pre-OHG times).

Note that in contrast to a split-CP model of the left periphery, there is no one-to-one relationship between semantic/pragmatic functions and (separate) structural positions in a model making use of multiple specifiers. In particular, the closest specifier of C may host elements linked to different semantic/pragmatic functions (topics, foci, *wh*-phrases), with the effects of the feature hierarchy proposed in (144) only becoming apparent when C contains more than a single substantial feature marked for PF-realization (giving rise to the projection of multiple specifiers). Thus, we might suppose that even in pre-OHG times, when C had not yet developed a semantically vacuous EPP feature and fronting was exclusively triggered by semantic or discourse-related factors, the multi-purpose character of this position blurred the original semantic/pragmatic motivation in cases where only a single XP was fronted to clause-initial position. At some point, possibly fueled by other developments such as the ongoing loss of sentence particles, this probably led to a situation where learners could not unambiguously identify the semantic/pragmatic triggers associated with fronting (to C's first specifier) in the target grammar (see also Hinterhölzl et al. 2005 for a related conclusion). Still, they somehow had to account for the word order facts in the relevant clauses. Following Simpson (2004), this is a situation where learners can resort to the insertion of semantically vacuous EPP-features to replicate the patterns encountered in the input data. Together with the independent (and earlier) development of generalized V-to-C movement, this eventually led to the grammar of early OHG, where C may be endowed with an EPP-feature, which has not yet been generalized to all main clauses, as evidenced by the possibility of V1 order in main declaratives.

Another factor that possibly promoted the development of an EPP feature in C involves the distribution of discourse-connectives such as OHG *thô*. Recall that in section 3.3 above, we analyzed obligatory inversion triggered by OE *þa/þonne* 'then' in terms of a structure where sequential, discourse-continuative *þa/þonne* is merged in SpecTP, forcing the subject to stay behind in its theta-position, to the right of the

finite verb raised to T (the same analysis presumably carries over to the Gothic data in section 2.2). If we take this structure to be common to all early Germanic languages, then it might also have contributed to the rise of C's EPP feature after Continental Germanic had developed generalized V-to-C movement. Note that at this point, learners were confronted with a frequent pattern in which the equivalent of OE *þa/ponne* and OHG *thô* triggered subject-verb inversion. As already noted, this pattern was derived TP-internally in the (older) target grammar. However, after the innovation of V-to-C movement, such patterns had to be reanalyzed with the verb located in C. As a result, prefinite *thô* 'then' could no longer be placed in SpecTP. Rather, the learner had to assume that *thô* occupies the closest specifier of C.⁹³

(146) a. $[_{CP} \dots C_{[-EPP]} [_{TP} thô [_{T'} T + V_{fin} [_{vP} \dots]]]] \rightarrow$
 b. $[_{CP} \dots thô [_{C'} V_{fin} + C_{[+EPP]} [_{TP} t_{thô} T [_{vP} \dots]]]]$

Note that the learner had to posit a trigger for placing *thô* in SpecCP. Assuming that its original discourse function was still linked to SpecTP at this stage (i.e., the relevant grammar was still discourse-configurational), the learner had to assume a movement process linking fronted *thô* to its original position. However, since *thô* could neither be analyzed as a topic nor as a focus, again the only possibility the learner could resort to was assuming that fronting of *thô* resulted from the presence of a semantically vacuous EPP feature in C. Let's now address the question of how the resulting grammar, which still had (limited) access to multiple specifiers in the C-domain, turned into a generalized V2 grammar in the course of the OHG period. Again, we will see that the status of *thô* was instrumental in that later change.

It appears that many facets of OHG word order were determined by information-structural properties (cf. Hinterhölzl 2004, Schlachter 2004, Hinterhölzl et al. 2005). For example, it has been observed that new/focused and heavy constituents preferably occupy a postverbal position (i.e., the *Nachfeld*), while preverbal position was apparently reserved for light and topical elements. This suggests that the OV-VO alternation could be used to signal differences between given and new information, expressing the focus-background structure of utterances (Hinterhölzl 2004, Schlachter 2004). In a similar vein, Hinterhölzl et al. (2005) show that in early OHG main clauses, V2 order was primarily used in clauses containing

⁹³ Note that this change could not take place in OE since English failed to innovate generalized V-to-C movement.

given discourse referents, signaling discourse continuity, while V1 order signaled a change in discourse topic, often introducing new discourse referents (see also section 5.3.1 above). They further argue that in early OHG, the position of the finite verb served to separate topic and focus domains in the clause. Thus, it seems that (early) OHG can be classified as a discourse-configurational language. However, what we can observe is that already in the OHG period, verb placement ceased to be controlled by considerations of topicality and focus (cf. Hinterhölzl 2004, Hinterhölzl et al. 2005, Axel 2007). As a result, the word order became more and more fixed, eventually leading to the V2 + SOV grammar characteristic of present-day German, where the prefinite position in main clauses is not linked to any particular function in terms of information-structural distinctions. As already noted, this development gave rise to a semantically vacuous EPP-feature in C (in order to mimic V2 orders in the input after the original semantic/pragmatic motivation for XP-fronting was no longer transparent). This presumably led to a general decrease in multiple XP-fronting in the output of learners that failed to recognize the original semantic/pragmatic function of XP-fronting (why should the learner carry out multiple operations the motivation and information-structural function of which are unclear to him?). Still, given the above analysis of the left clausal periphery of OHG in terms of multiple specifiers, the question is why at some point, the option of projecting more than a single specifier of C became unavailable (giving rise to strict generalized V2 order in main declaratives). In what follows, I am going to argue that the change that eventually tipped the scales in favor of a strict V2 grammar concerned the status of the frequently-used discourse connective *thô* 'then'.

Recall that similar to OE, we expect that the loss of discourse-configurationality motivated a change in which elements such as *thô* could no longer be externally merged as a specifier of T (undergoing subsequent movement to SpecCP), which turned into a licensing position for subjects, exclusively.⁹⁴ As a result, clauses with prefinite *thô* were subject to the reanalysis in (147), in which learners took clause-initial *thô* to be base-generated in SpecCP, discarding the movement operation that

⁹⁴ Still, it is fairly clear that the properties of T in German differ from the properties of T in English. For example, German did not develop an obligatory subject expletive such as English *there*, and there are a number of constructions where the subject position is not filled by an overt element (cf. e.g. Grewendorf 1989, Haider 1993 for discussion). Note that under the assumption that the availability of multiple specifiers is dependent on the absence of expletives (see main text), the lack of a subject expletive may be linked to the possibility of scrambling to pre-subject position (if the latter is analyzed in terms of multiple specifiers, see Grewendorf and Sabel 1999 for relevant discussion).

accomplished *thô*-fronting in the target grammar. This reanalysis was presumably further promoted by least effort strategies that favor the least costly derivation in case the input is ambiguous (cf. Roberts and Roussou 2003).

(147) a. $[_{CP} \dots thô [C' V_{fin} + C_{[+EPP]} [_{TP} t_{thô} T [_{vP} \dots]]]] \rightarrow$
 b. $[_{CP} thô [C' V_{fin} + C_{[+EPP]} [_{TP} \dots T [_{vP} \dots]]]]$

The reanalysis in (147) effectively turned *thô* into an expletive element inserted in SpecCP to satisfy C's EPP-feature. Thus, *thô* can be described as a precursor of the expletive *es* 'it' (so-called *Vorfeld-es*) that developed in the MHG period (Brugmann 1917, Behaghel 1928, Lenerz 1985b, Abraham 1993, Axel 2007; see also section 5.5 above). Recall that the expletive use of clause-initial *da* can still be observed in present-day German dialects, which often make only sparse use of the *Vorfeld-es* (cf. e.g. Weiß 1998 on Bavarian):

(148) Do is a *Ungligg* bassierd.
 there is an accident happened
 'An accident has happened.'
 (Weiß 1998: 102)

Let us now turn to the question of how the reanalysis in (147) might have affected C's capability of projecting multiple specifiers. What I want to argue is that the robust presence of expletive *thô* (and later *es*) in the input signaled to the learner that C can project only a single specifier.

A first indication that there is a connection between the presence of expletive elements in the specifier of a functional category F and F's capability of projecting more than a single specifier comes from the syntactic distribution of expletives across Germanic. It seems that expletives are either of the subject-type (occurring in SpecTP, as e.g. in English, Scandinavian, and, to some extent, Dutch) or of the V2-type (occupying SpecCP in all Germanic V2 languages), cf. Chomsky (1995: 362f.) for some discussion. In both cases, it appears that the availability of an expletive not only signals that the relevant syntactic position must be obligatorily filled in a certain context (e.g., the subject position of tensed clauses in English or the *Vorfeld* of main declaratives in the Germanic V2 languages), but also that the relevant syntactic position is unique, in the sense that the relevant functional head (T and C,

respectively) may not project more than a single specifier (cf. the non-availability of scrambling to pre-subject position in English and Scandinavian, and the absence of V3 effects in V2 languages).^{95,96} In contrast, the absence of subject expletives like *there* seems to be correlated with the possibility of scrambling elements to the left of the subject, which can be taken to indicate the availability of multiple specifiers (or adjunction sites) in the relevant functional projection (as for example in German, cf. Grewendorf and Sabel 1999 for discussion). Provisionally, these observations can be phrased as in the following descriptive generalization:

(149) *Generalization: Expletives and multiple specifiers*

A functional head F can project multiple specifiers only if the grammar does not contain an expletive related to F.

In other words, it seems that the requirement that a certain structural position be overtly filled (either by movement or by externally merging a semantically vacuous element) blocks the availability of multiple specifiers. Of course, this raises the question of whether it is possible to derive (149) from deeper properties of grammar (i.e., the theory of phrase structure). What I am going to argue is that the special behavior of expletives, which seem to prevent multiple specifiers, thereby 'closing off' the projection of a functional head, follows from strict cyclicity in combination with the assumption that the expletive itself acts as a probe, initiating an Agree relation with a functional head F after the expletive has been merged as specifier of F (Chomsky 2000, 2004).

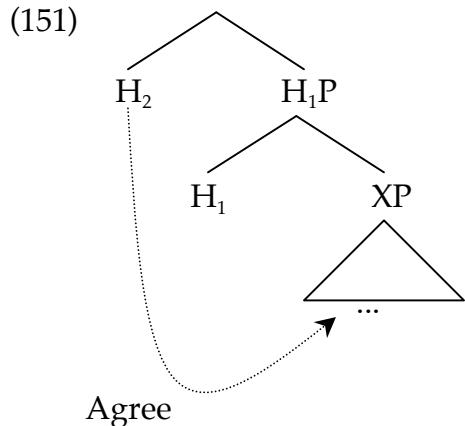
First of all, let's adopt a version of strict cyclicity in which a lower head H_1 may not any longer trigger syntactic operations (Agree, Move/internal Merge) after a higher head H_2 has been merged, acting as a probe (Chomsky 2000: 132):

⁹⁵ The (restricted) possibility of placing object pronouns to the left of the subject in Dutch differs significantly from free constituent reordering in the German midfield. This suggests that pronoun fronting (in Dutch, and possibly in general) should not be analyzed on a par with scrambling, cf. Haider (2000c).

⁹⁶ Abstracting away from Chomsky's (1995) analysis of transitive expletive constructions in Icelandic, where it is assumed that expletive and associate occupy multiple specifiers of TP. However, note that this analysis raises a number of questions. For example, it falsely predicts the availability of V3 orders in the relevant contexts, which can only be circumvented by assuming that PF-operations place the finite verb between expletive and associate (cf. Chomsky 1995: 368).

(150) Properties of the probe/selector α must be satisfied before new elements of the lexical subarray are accessed to drive further operations.

Thus, in a structure like (151), H_1 is inert after H_2 (which has been subsequently added to the structure) has initiated an Agree operation:



This assumption seems to be implicit in most work on the strict cycle (for related discussion cf. e.g. Chomsky 1995: 234f., Collins 1997: 81ff., and in particular Chomsky 2000: 132f.); it follows more or less directly if phases are equated with phrases as for example in Müller (2007).⁹⁷

The second ingredient of our explanation of (149) involves the status of expletives in the probe/goal system envisaged in Chomsky (2000) and subsequent work. Recall that originally (cf. e.g. Chomsky 1995) it has been assumed that expletives are inserted into SpecTP to check T's EPP feature and an uninterpretable categorial [D] feature of the expletive. In Chomsky (2000), this analysis is slightly revised by assuming that the relevant uninterpretable feature is [person]. In the same work, Chomsky proposes a major modification of the syntactic computational system, arguing that feature checking/valuation does not proceed via specifier-head configurations. Rather, he assumes that a feature/head P capable of triggering a syntactic Agree operation (a *probe*) seeks a matching *goal* G in its c-command domain (i.e., its sister), leading to valuation (and deletion) of uninterpretable features in both

⁹⁷ The status of (150) is somewhat less clear under the assumption that T may initiate syntactic operations only after it has inherited the relevant uninterpretable features from C (Chomsky 2004, 2005). One might argue, however, that this particular situation does not conflict with (150), since T in fact has no probe properties prior to Merge of C. After C has been added, the relevant features (e.g., $u\phi$, EPP) are handed over to T, giving rise to cases of 'parallel probing' where the operations triggered by C and T apply in parallel.

P and G. Furthermore, it is assumed that checking/elimination of EPP features (now called *edge features* EF) is parasitic on a previous Agree relation. However, this mechanism raises a problem for the analysis of expletives: since SpecTP is no longer part of the checking domain of T, T cannot enter into a checking relation with the expletive if the latter is directly merged as SpecTP. There are two possible solutions to this problem (see Radford 2004: 300ff. for discussion).

First, one might assume that the expletive is merged in a lower, vP-internal position (cf. Bowers 2002). It is then part of the checking domain of T (i.e., T's sister) and may be accessed by an Agree operation initiated by T, subsequently undergoing raising to SpecTP to eliminate T's EPP feature. However, this approach faces two conceptual shortcomings: (i) The semantically vacuous character of expletives is at odds with the assumption that the expletive is merged within the theta-domain (in a potential theta-position); (ii) this analysis does not capture the intuition that expletives are inserted as a kind of last resort to fill a structural position that requires overt realization (SpecTP in English, SpecCP in the Germanic V2 languages).

Alternatively, we may maintain the original assumption that expletives are externally merged directly in the position where they are needed (i.e., SpecTP, SpecCP). Assuming the probe/goal system, this requires that it is the expletive itself that acts as a probe (cf. Chomsky 2000: 128, 2004: 114):

“The results are expected if Expl is an X^0 head and its [person] feature is uninterpretable, therefore able to probe its domain T' ($= D(Expl)$), locating the ϕ -set of T as the closest goal. The uninterpretable probe deletes, and the ϕ -set of T as well if Expl has a full ϕ -set.” (Chomsky 2000: 128)

Prior to merging the expletive in SpecTP, T accesses the vP-internal associate, establishing agreement with it. To ensure that T's ϕ -set can nonetheless serve as an active goal for the expletive probe, one has to assume that the relevant valued features do not delete immediately, but remain accessible for the syntactic computation until the next phase head (C) is merged (perhaps along the lines proposed in Pesetsky and Torrego 2001, who suggest that the relevant features are only *marked for deletion* on the TP cycle). In case T is defective, that is, it does not

contain a full set of ϕ -features, deletion of the expletive's [person] feature must wait until a higher ϕ -complete probe is merged.⁹⁸

Together with the assumption that a head H_1 is inert after a higher head H_2 has acted as a probe (i.e., strict cyclicity in the sense of (150)), we are now in a position to derive generalization (149) from independent properties of the computational system: After an expletive probe has been added to the structure, establishing an Agree relation with C or T as the closest goal, C/T become inert and may not trigger further operations. As a result, they can neither attract further elements nor project additional specifiers, which rules out the possibility of multiple specifiers after the expletive has been merged. In this way, we derive that Merge of an expletive effectively 'closes off' the projection of the relevant functional head.

Now, coming back to our initial question, let's see how these considerations bear on the loss of V3 orders (i.e., the option of multiple specifiers in the CP) in the course of OHG. Recall that at some point in the OHG period, *thô* was reanalyzed as an expletive-like element directly merged in SpecCP (cf. (147) above, repeated here for convenience):

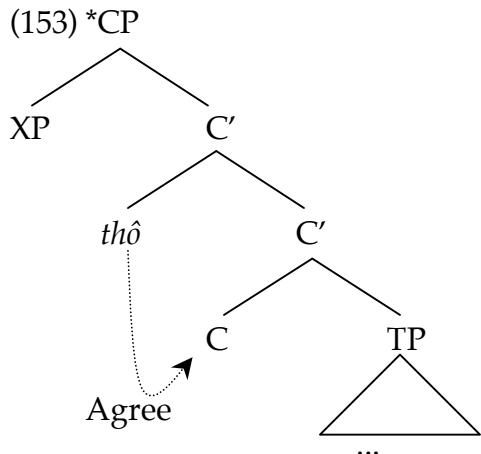
(152) a. $[_{CP} \dots thô [_{C'} V_{fin} + C_{[+EPP]} [_{TP} t_{thô} T [_{vP} \dots]]]] \rightarrow$
 b. $[_{CP} thô [_{C'} V_{fin} + C_{[+EPP]} [_{TP} \dots T [_{vP} \dots]]]]$

In the resulting structure, external Merge of *thô* in SpecCP eliminates C's EPP-feature. Furthermore, *thô* carries an uninterpretable feature [uF] that renders it active and must be eliminated as well.⁹⁹ Adopting the analysis of expletives devised in

⁹⁸ Under these assumptions, expletives may act as both probes and goals, depending on the syntactic context in which they occur. In this connection, the question arises of whether we would predict that it is in fact the expletive probe that determines the label of the relevant term, giving rise to an *ExplP* (instead of *TP*) in conventional notation. This may be ruled out by the assumption that the wrong choice cancels the derivation at a later stage, since *ExplP* cannot be selected by C. Alternatively, we might speculate that defective categories may not project, giving rise to some perhaps quite interesting predictions for non-finite clauses, which would be *vPs* instead of *TPs* after Merge of defective T with *vP*. However, for reasons of time and space, I cannot pursue these matters here. Note that further questions arise for expletives that are ϕ -complete such as French *il*. Here, Chomsky (2004: 114) assumes that the agreement relation established between the expletive and T overwrites the values that has been established previously by the Agree relation between T and the associate.

⁹⁹ As to the nature of [uF], we might speculate that it relates either to C's clause type features (i.e., [+declarative] in the case at hand) or to the fact that C in V2 languages is typically linked to

Chomsky (2000, 2004), *thô* acts as a probe that accesses C as the closest goal. As a result, *thô*'s [uF] deletes. Moreover (and this what is crucial for our purposes), C is inert and cannot trigger any further operations after it has been accessed by the expletive probe. Thus, C may not project further specifiers, ruling out a structure as in (153).



In this way, examples with clause-initial *thô* provided positive evidence to the learner that at least in a certain context, C could not project more than a single specifier. Furthermore, recall that relevant examples were particularly frequent in OHG, which suggests that they played an important role in the PLD constructed from the input.¹⁰⁰ Together with the fact that the original semantic/pragmatic motivation for XP-fronting was becoming more and more opaque, the reanalysis (147) can thus be taken to have tipped the scales in favor of a strict V2 grammar that lacks the possibility of multiple specifiers in the C-domain.¹⁰¹

finiteness. The latter might be taken to indicate that both C and the expletive *thô* carry an uninterpretable tense specification [uTns]. This seems to make the correct typological prediction that cross-linguistically, C-related expletives are confined to V2 languages. I leave this point open for future research.

¹⁰⁰ There is an interesting difference between present-day German and Scandinavian languages such as Swedish in that the latter exhibit a higher frequency of lightweight, expletive elements such as *det* or *så* in clause-initial position, while German has apparently developed a tendency to place adverbs, rhematic information and phonologically heavy elements in the prefield (cf. Bohnacker 2006, Bohnacker and Rosén 2007). This suggests that the Scandinavian languages have preserved more faithfully the original origins of a strict V2 grammar.

¹⁰¹ Under a cue-based model of language acquisition (Lightfoot 1999), one might possibly say that the relevant cue is not exactly the presence of expletives, but rather the observation that in a given well-defined context (SpecTP of tensed clauses in English, SpecCP of matrix declaratives in German), a certain position cannot be left unfilled.

5.7 Section summary: V2 in Old High German

This section has shown that the basic ingredients of V2 – generalized V-to-C movement and fronting of a single XP into prefinite position – were already very much in place at the time when the earliest OHG texts were composed. If we compare OHG with other early Germanic languages, we can observe that in contrast to Gothic and OE, V-to-C movement is not confined to operator contexts. Rather, verb fronting to C has been generalized to all main clause contexts at some pre-OHG stage. This also explains why OHG does not exhibit systematic deviations from V2 in the context of (subject) pronouns (in contrast to OE). I have argued that apparent cases of XP-pron-V_{fin} order can be attributed to the fact that early OHG had access to another grammatical option, namely basic SOV order (+fronting or extraposition) in main clauses, which presumably represents an archaic trait that was already in decline in the earliest attested stages of OHG (contra Axel 2007, but in line with e.g. Lenerz 1984). Other patterns that in which OHG differs from the present-day V2 languages include (i) the frequent use of V1 order in main declaratives to express certain discourse-related distinctions (introduction of new discourse topic, foregrounding), and (ii) residues of multiple XP-fronting, which suggests that XP-fronting could still be used to express (and was partially motivated by) information-structural distinctions such as topic or focus. The nature of these deviations from V2 shows that in contrast to V-to-C movement, the second ingredient of V2 (XP-fronting) has not yet been fully generalized in early OHG. Still, the fact that XP-fronting also affects non-topical material such as indefinites suggests that already at this stage, C (optionally) carried a semantically vacuous EPP-feature. Based on a critical discussion of a split CP analysis of the left periphery of OHG (cf. Axel 2007), I have argued that the relevant empirical facts can be more adequately captured if we assume that the C-domain of early OHG contained only a single functional head (C) which could project multiple specifiers. I have then proposed that multiple XP-fronting was triggered by a set of hierarchically ordered, discourse-related features in C (topic, focus etc.). Under these assumptions, I have claimed that the rise of a strict V2 syntax proceeded via two stages. First, when learners could no longer detect the original semantic/pragmatic trigger of XP-fronting, they posited an EPP-feature in C to mimic dislocation phenomena the original motivation of which was no longer transparent. The rise of an EPP-feature was promoted by a highly

frequent V2 pattern in which the clause-initial position was occupied by *thô* 'then', the OHG equivalent to Gothic *þanuh* and OE *þa*: After the development of generalized V-to-C movement, such examples could only be parsed as involving fronting of *thô* to SpecCP. This instigated the rise of an EPP-feature since *thô*-fronting could not be attributed to 'strong' topic or focus features in C. In later stages of OHG, *thô* underwent a second reanalysis in which it turned into a semantically light expletive-like element directly merged in SpecCP. I have claimed that the latter change was decisive for the development of generalized V2 in German since the presence of expletives signaled to the learner that a functional head may project only a single specifier, which led to the loss of V3 orders. The latter change, which was shaped by a hard-wired property of grammar concerning the syntax of expletives, can be elegantly captured as a simple parametric change (availability of multiple specifiers) in a model making use of a single left-peripheral functional head, while it raises a number of questions for analyses of OHG which are based on a split CP.

6 Summary

In this chapter, I have taken a closer look at the properties and historical development of the V2 phenomenon in early Germanic, arguing that we can identify a couple of different structural configurations that may give rise to surface V2 orders in early Germanic. First, we have seen that contexts such as questions, imperatives, and examples with fronted negation, in which the finite verb accompanies operator movement to the C-domain, presumably constitutes the historical origin of the V2 phenomenon in Germanic (see also Eyþórsson 1995, 1996; Kiparsky 1995). Another context that triggers systematic verb fronting in all early Germanic languages are clauses introduced by certain temporal adverbs such as Gothic *þanuh*, OE *þa/þonne* and OHG *thô*, all roughly meaning 'then'. It appears that these elements share a common discourse function across early Germanic, in that they typically introduce new actions/events along the main time line of the story, often in connection with the introduction of new discourse topics. I have presented evidence suggesting that the anaphoric character of 'then' was linked to the specifier of TP in early Germanic, which by assumption served to express the discourse-related property of anaphoricity in the early Germanic languages. The fact that the further discourse functions of these elements seems to quite similar to the discourse function of V1-order might be taken to suggest that the sentence connecting adverbs were originally

added to reinforce the functions expressed by V1 structures (cf. Betten 1987 on OHG). Another source of surface V2 orders is so-called 'pseudo V2' in OE, which does not result from a Spec-head relation between the fronted XP and the finite verb, but rather involves a configuration in which the two elements are merely linearly adjacent (with the fronted XP in SpecCP and the finite verb in T). Due to the fact that T did not carry an EPP-feature in OE, pseudo-V2 could also give rise to inversion with full DP subjects, which stay behind in their vP-internal theta-position. Finally, there is the parametric option of 'generalized V2', in which the finite verb occupies C in all main clauses, accompanied by EPP-triggered XP-fronting in declaratives. This option is found already in early OHG, gaining a wider distribution in the course of the OHG period.

Due to these parametric differences, we do not find a uniform behavior with respect to V2 in early Germanic (apart from inversion/verb fronting which is systematically triggered in operator contexts and clauses introduced by 'then'). That is, while all early Germanic languages exhibit the option of V1 or V-final order in main declaratives, we still find different patterns of violations of V2, dependent on which structural configurations are available in the individual languages. In Gothic, systematic V2 order is confined to operator contexts and clauses introduced by *þanuh* (and, to a lesser extent, *þaruh*), while a wide variety of other orders seems to be possible in main clauses. In contrast, OE exhibits additional V2 patterns due to the availability of 'pseudo-V2' which presumably emerged after the innovation of systematic V-to-T movement in pre-OE times. Still, it is fairly clear that OE differed significantly from the modern Germanic V2 languages (and OHG) since it exhibits a considerable number of systematic V3 orders in main clauses, in particular the pattern XP-subj.-V_{fin}, which I have accounted for by assuming that the finite verb is located in T in main declaratives of OE while topical/anaphoric subjects may move to SpecTP in order to check a feature in T linked to anaphoricity ([*D*]).

A major force in the historical development of the V2 phenomenon are acquisition scenarios in which learners (or, a generation of learners) encounter a word order pattern (involving dislocation) for which they cannot identify a clear semantic/pragmatic trigger. Following proposals by Simpson (2004), I have taken the view that in such a situation, learners may posit an EPP-feature in the relevant functional head to imitate the word orders they are confronted with in the input, giving rise to generalized XP-fronting in the history of German (via the development of a generalized EPP-feature in C), and the loss of surface 'pseudo V2' patterns in

English (via the development of a generalized EPP-feature in T). In what follows, I briefly summarize the contents of the individual sections of this chapter.

In section 1, I have argued that Gothic exhibits systematic V-to-C movement in a set of syntactic contexts that are quite similar to the contexts that trigger V-to-C movement in present-day English (imperatives, interrogatives, examples with fronted negation), giving rise to V2 order in *wh*-questions. Furthermore, it has become clear that apparent deviations from V2 in *wh*-questions can be attributed to extra-grammatical factors, namely influence of the word order of the Greek source text. In addition, we have seen that clauses introduced by the foregrounding particles *þaruh* 'there+uh' and in particular *þanuh* 'then+uh' constitute another context where V2 order occurs regularly.

In section 3, I have examined the apparently more advanced V2 syntax of Old English, arguing that the core V2 properties of this early Germanic language reduce to the very same contexts where V-to-C movement is found in Gothic. More to the point, I have proposed that in OE, V2 orders resulted from three different underlying structural configurations: First, a spec-head relationship between a fronted operator and the finite verb (in C⁰) could give rise to 'residual V2' effects, similar to present-day English. Second, superficial V2 orders could result from linear adjacency between a fronted non-operator XP in SpecCP and the finite verb located in T, with non-pronominal subject DPs remaining in their theta-position SpecvP (due to the absence of a general EPP feature in T). In addition, V2 orders could reflect a spec-head configuration between the temporal adverbs *þa* and *þonne* and the finite verb in T. I have suggested that this option can be attributed to the discourse-configurational nature of OE, where SpecTP was linked to anaphoricity. By assumption, this property is established in the syntax via a strong [*D*] feature that may optionally be added to the content of T in OE. In the absence of *þa* and *þonne*, this feature is checked by pronominal elements that move to SpecTP, giving rise to V3 orders with fronted non-operators. Furthermore, I have analyzed the robustly attested order XP-DP_{subj}-V_{finv} as a precursor of the structure of Modern English, with the full nominal subject moved to SpecTP. By assumption, this option is available only for specific subjects that relate to a previously established discourse topic (again in connection with T's [*D*] feature).

Section 4 has shown that this approach to V2/V3 in OE receives further support from observations on the loss of V2 patterns in the ME period. I have argued that the loss of V2 patterns (i) with inverted non-pronominal subject DPs, and (ii) in the

context of clause-initial *þa/þonne* can be attributed to a single underlying change, namely the rise of a generalized EPP feature in T during the ME period. Following proposals by Simpson (2004), I have assumed that learners resorted to positing a semantically vacuous EPP feature in T to mimic word orders attested robustly in the input when the original semantic/pragmatic motivation for subject movement to SpecTP (licensing of topical/anaphoric elements) became unclear. The development of a structural subject position was part of a large-scale change in which English turned from a discourse-configurational language into a language where word order primarily serves to encode grammatical functions.

Finally, section 5 has shown that OHG represents a more advanced stage of V2, which already closely resembles the modern Germanic V2 languages. In particular, verb fronting is neither confined to operator contexts nor does OHG exhibit systematic deviations from V2 in the context of (subject) pronouns (in contrast to OE). This clearly indicates that the finite verb undergoes systematic fronting to C in early OHG, crossing the position of weak pronouns at the left edge of IP/TP. However, it has also become clear that early OHG exhibited a number of patterns that cannot be found in the present-day V2 languages. The deviations from V2 show that in contrast to V-to-C movement, the second ingredient of V2 – fronting of single XP – has not yet been fully generalized in early OHG. In particular, we can frequently observe V1 order, and to some extent multiple XP-fronting. The latter indicates that XP-fronting could still be used to express information-structural distinctions such as topic or focus (similar to OE). On the other hand, XP-fronting can also affect non-topical material (i.e., indefinites), which suggests that already at this stage, C (optionally) carried a semantically vacuous EPP-feature. I have argued that the relevant empirical facts, including the changes that took place in the course of the OHG period, should be accounted for by a multiple-specifier analysis that posits only a single functional head in the left clausal periphery (C). Under this assumption, multiple XP-fronting in early OHG has been analyzed as being triggered by a set of discourse-related features in C (topic, focus etc.), which are hierarchically ordered. Generalized V2 began to emerge when learners could no longer detect the original semantic/pragmatic trigger of XP-fronting: To account for the relevant orders in the input, learners posited an EPP-feature in C. Furthermore, I have argued that examples with clause-initial *thô* played a particular role in the development of a strict V2 grammar. Similar to Gothic *þanuh* and OE *þa*, clause-initial OHG *thô* triggers systematic inversion, introducing new foregrounded actions/events and/or

discourse topics. After the development of generalized V-to-C movement, such examples could only be parsed as involving fronting of *thô* to SpecCP. This further promoted the rise of an EPP-feature, since *thô*-fronting could not be attributed to topic or focus features in C. When its original discourse function became unclear, *thô* was reanalyzed as a semantically light expletive-like element directly merged in SpecCP. I have suggested that the latter change was decisive for the development of generalized V2 in German since the presence of expletives signaled to the learner that a functional head may project only a single specifier, which led to the loss of V3 orders.

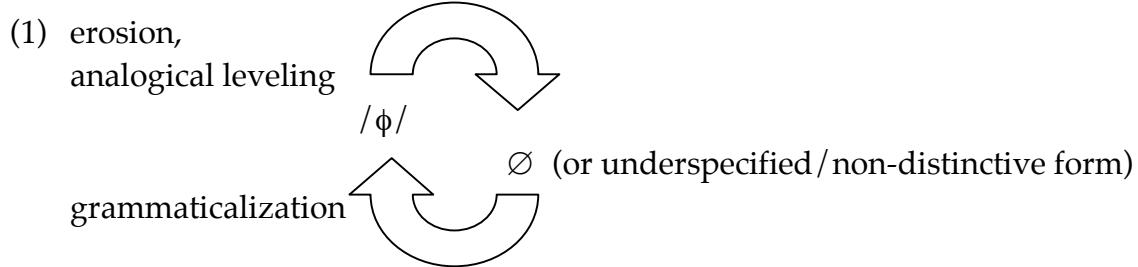
In this chapter, we have seen again that grammar change is shaped by different factors during language acquisition. On the one hand, change may be triggered (or restricted) by hard-wired properties of grammar, such as the principles that governed the loss of multiple specifiers discussed in section 5.6.2. On the other hand, I have argued that the task of the learner is eased by an acquisition strategy that enables the child to accommodate dislocation phenomena for which no substantial trigger can be detected in PLD. More precisely, I have claimed that learners may make use of EPP features to imitate relevant word order patterns encountered in the input. The next chapter examines another set of such 'learning' strategies, focusing on the acquisition of inflectional morphology and the observation that morphological change often proceeds in a cyclic fashion.

Chapter 4: On the cyclic nature of language change

1 Introduction

In the previous chapters, we have seen that the course of language change may be shaped by (i) hard-wired properties of the syntax-morphology mapping, and (ii) acquisition strategies that the learner applies to the input data in case the evidence for a certain structural property is ambiguous (e.g., via positing EPP-features to imitate dislocation phenomena the original semantic/pragmatic trigger of which has become unclear). This chapter examines another facet of the interaction between language acquisition and language change, focusing on the way learners acquire phonological exponents of inflectional categories. For example, when learners acquire the inventory of agreement markers in a given language, they must be able to identify the phonological shape, the distribution, and the featural specification of the individual inflectional formatives. The latter in particular is not a trivial task since Vocabulary items are typically underspecified for the features present in the syntactic representation (e.g., it is usually assumed that German 3pl /-n/ is not fully specified for person and number features. Rather, the fact that the very same formative appears in 1pl contexts is taken to suggest that it is underspecified for person features). Following proposals by Noyer (1997) and Halle (1997), I will assume that the acquisition of inflectional formatives involves learning strategies that are applied to the input data in order to identify the set of relevant Vocabulary items and their featural specifications (see also Fuß 2005 for some discussion).

It will become clear that taking a closer look at the workings of the relevant acquisition strategies offers a new perspective on the observation that morphological change often proceeds in a cyclic fashion, as illustrated in (1):



In particular, we can observe that the phonological realization of inflectional categories is affected by changes with reverse outcomes in the historical development of languages. On the one hand, distinctive morphology is lost via phonological erosion and analogical leveling. On the other hand, the loss of distinctions is compensated for by grammaticalization processes that provide new and more expressive (i.e., more specified) phonological exponents of inflectional categories.¹ Traditional explanations for the ups and downs on this cycle often appeal to functionalist notions such as speaker- vs. hearer-oriented economy, or the need to coin new forms to pursue certain communicative goals. In this chapter, I explore whether it is possible to account for the interplay between paradigm leveling and grammaticalization in more formal terms, focusing on the historical development of verbal agreement marking in a set of German dialects (notably Bavarian and Alemannic). I am going to argue that we can gain a deeper understanding of the cycle in (1) if we take a closer look at the feature specifications of individual Vocabulary items (alone, and in relation to other formatives in a paradigm) and the way these specifications are learned in the process of first language acquisition. The central proposal put forward in this chapter is that the cyclic nature of morphological change is guided by (apparently) conflicting strategies that shape the acquisition of inflectional morphology.

First, I assume that there is a learning strategy based on morphological blocking that selects the most specified variant in case the input contains more than a single potential realization of a given inflectional category (cf. Fuß 2005). The workings of this learning strategy typically promote grammaticalization processes that lead to new and more distinctive phonological realizations of underlying inflectional categories. Of course, this is reminiscent of synchronic blocking effects driven by

¹ Cf. Paul (1880) or Gabelentz (1891). The idea that grammaticalization processes are motivated by the need to compensate for the loss of distinctions due to phonological erosion is widely held in typological/functional approaches to grammaticalization, cf. Lüdtke (1980), Hopper and Traugott (1993), Siewierska (1999), (2004), Ariel (2000), and Lehmann (2002), among others.

some form of the *Elsewhere Condition*, which requires that the availability of a more specific form or rule blocks the use of a less specific form or rule (cf. Kiparsky 1982 and in particular Halle's 1997 *Subset Principle* discussed in chapter 2 above). A well-known example for this kind of inflectional blocking comes from English, where the existence of a more specified exponent /z/ blocks the use of the completely underspecified null suffix in the context of 3SG.PRES.INDIC.ACTIVE:²

(2) he/she/it run-s vs. *he/she/it run-Ø

The effects of blocking-induced change are balanced by a second acquisition strategy that aims at minimizing the number of elements/features stored in the lexicon, which I call *Minimize Feature Content* (cf. Halle 1997). I am going to argue that this strategy may give rise to effects that are traditionally subsumed under the notion of analogical change. In this way, more regular/less specified variants may be introduced into the grammar. Furthermore, it will become clear that the workings of this acquisition strategy may lead to a more transparent relation between form and function/meaning. In contrast to functionalist approaches, however, I claim that this particular outcome (sometimes referred as the "agglutinative ideal") does not drive language change, but is rather to be analyzed as a side-effect of the workings of *Minimize Feature Content*.

This chapter is organized as follows. In section 2, I explore the role of morphological blocking in grammaticalization processes leading to more specific exponents of inflectional categories. The discussion focuses on the development of new plural agreement suffixes in the history of Bavarian, basically following Fuß (2005). Section 3 examines cases of analogical leveling in a set of Alemannic varieties, focusing on the development of *Einheitsplural* 'common plural' (e.g., Zurich German 1pl/2pl/3pl *träffed* 'meet'). It will become clear that at least core cases of analogical changes can be attributed to a preference for minimizing the number of features (or, lexical items) mentioned in the lexicon. Section 4 deals with the role of morphological blocking in another well-known cyclic change, namely the rise and fall of null subjects. Focusing on relevant developments in Bavarian, French, and Finnish, I am going to argue that the availability of a null realization of a weak pronominal D-head

² For a general discussion of blocking/elsewhere effects cf. e.g. Kiparsky (1973), (1982); Aronoff (1976); Anderson (1986), (1992); Halle and Marantz (1993), Halle (1997); Noyer (1997); Giegerich (2001); Embick and Marantz (2006).

may be dependent on properties of the inventory of overt pronominal forms. More to the point, I propose that a null realization may become available if a competing overt form is lost due to independent changes (so-called *deblocking*). This analysis predicts that the development of new weak pronouns can block a null realization of pronominal D that was formerly available in the grammar. We will see that a possible case in point is Colloquial Finnish, where the loss of null subjects is accompanied by the development of a new series of unstressed pronouns (at the end of this chapter you will find an appendix in which I take a cursory look at alternative pathways to pro-drop, focusing on the emergence of discourse-oriented pro-drop phenomena in a set of Creoles). Section 5 provides a concluding summary of the findings reached in this chapter.

2 Blocking and the grammaticalization of verbal inflection

Across languages, we can observe that the grammaticalization of inflectional markers does not replace existing formatives in a random fashion. Focusing on the rise of verbal agreement marking, it seems that the creation of new inflectional material complies with the following generalization (cf. Fuß 2005: 1):³

- (3) New verbal agreement formatives arise only for those slots of the agreement paradigm where the existing inflections are non-distinctive.

In previous work, I have argued that this generalization can be formally accounted for if we assume that the acquisition of phonological exponents of inflectional categories is shaped by an economy constraint that favors the use of more specified exponents over less specified exponents (cf. Fuß 2005: 233). The relevant acquisition strategy, dubbed the *Blocking Principle*, is stated in (4):

³ Relevant examples include: various Italian and Rhaeto-Romance dialects (cf. e.g. Spiess 1956, Kuen 1957, Renzi and Vanelli 1983, Linder 1987, Haiman and Benincà 1992, Poletto 1997, Gerlach 2002), Non-Standard French (Wartburg 1970, Ashby 1977, Harris 1978, Auger 1993, Gerlach 2002), a number of German varieties (Wiesinger 1989, Dal Negro 2004, Fuß 2005, Kolmer 2005), Khinalug (North-East Caucasian, Corbett 1991: 123), Kisar (Austronesian, Blood 1992, Siewierska 1999, 2004), Maricopa (Yuman, Siewierska 1999, 2004), and the Papuan language Skou (Donohue 2002).

(4) *Blocking Principle (BP)*

If several appropriate phonological realizations of a given morpheme are attested in the Primary Linguistic Data (PLD), the candidate matching the greatest subset of the morphosyntactic features included in the morpheme must be chosen for storage in the lexicon.

Thus, I assume that child learners scan the input they receive for the most specific phonological realization of a given underlying inflectional category. Similar to structural economy principles (cf. e.g. Clark and Roberts 1993a, Roberts and Roussou 2003), the BP is called into service only if the cues provided by the input data are for some reason ambiguous and not sufficient for identifying the exponent of an underlying morpheme on independent grounds. Relevant examples discussed in Fuß (2005) involve cases where independent, gradual changes (e.g., phonological erosion of a pronominal clitic) lead to a situation where the trigger experience can be taken to contain more than one potential exponent for a given agreement head/morpheme.⁴ The BP is then invoked to decide which of the candidates is eventually stored in the lexicon. In this way, the BP ensures that the development of new inflections can affect only underspecified slots of the paradigm, replacing non-distinctive markers.⁵ An instructive example for this kind of blocking-induced change comes from the historical development of the verbal agreement paradigm of Bavarian.

2.1 The grammaticalization of agreement markers in German/Bavarian

In a number of varieties of German, new agreement suffixes developed via a reanalysis of subject enclitics in inversion contexts (cf. e.g. Pfalz 1918, Bayer 1984, Wiesinger 1989, Weiß 2002, Fuß 2005 on Bavarian). The former clitics mostly turned into enlargements of the existing inherited agreement endings. The most widespread of these changes led to the 2sg suffix *-st*, which is commonly analyzed as a combination of the inherited ending 2sg *-s* and the onset of the 2sg nominative clitic

⁴ As argued for in Fuß (2005), the reanalysis of a pronoun as agreement marker is only possible if a set of independent necessary conditions is met (e.g., the pronominal element must be a phonologically reduced clitic that cannot receive stress, the reanalysis has to satisfy a set of syntactic conditions such as adjacency to the host, etc.).

⁵ Note that it is presumably more adequate to characterize the Blocking Principle as a restriction and not as a driving force in grammar change (in contrast to what is suggested in Fuß 2005). See below for some discussion of the interaction between blocking and analogical change.

thu (cf. e.g. Brinkmann 1931, Paul 1952: 192, Braune and Reiffenstein¹⁵ 2004: 261). In this section, I will focus on the changes that affected the verbal agreement paradigm of Bavarian.

Interestingly, it appears that the reanalysis of clitics as realizations of verbal agreement morphemes did not take place in a wholesale fashion, despite the fact that the relevant varieties of Bavarian exhibited a full paradigm of subject clitics. Rather, the change in question is confined to the following contexts:

(5) a. 2sg *-s* ⇒ *-st* (early OHG; *-st* in most modern varieties of German)
 b. 2pl *-t* ⇒ *-ts* (13th century Bavarian; attested in all modern Bav. varieties)
 c. 1pl *-an* ⇒ *-ma* (18th century; extension to verbs in clause-final position
 in e.g. some Lower Bavarian dialects, cf. Fuß 2005 for details)

Of course, the limited scope of this grammaticalization process raises the question of whether there is a principled explanation of why the reanalysis of clitics took place in some contexts but not in others. In what follows, I show that the facts in (5) can be directly related to the workings of the BP, focusing on the changes that affected 2pl and 1pl forms.⁶

If we take a closer look at the changes that took place in the history of Bavarian, it becomes apparent that the development of the new endings 2pl *-ts*, 1pl *-ma* served to eliminate syncretism in the verbal agreement paradigm. Table 1 illustrates the effects of the rise of 2pl *-ts* (orig. 2pl ending *-t* + clit. 2pl *-(ē)s*), which is first attested in 13th century texts (in Northern and Middle Bavarian, cf. Wiesinger 1989: 72f.):

⁶ For reasons of time and space, I do not go into the details of the earlier development giving rise to 2sg *-st*. In Fuß (2005: 235ff.), it is argued that the change in question was promoted by the fact that the resulting form was unambiguously specified for verbal mood (indicative) and therefore proceeded in line with the BP. Another causal factor involved in this change was presumably the fact that other verbs already showed *-st* for the 2sg present indicative (notably, the class of preterite-presents, e.g. *kanst* 'can-2sg', *tarst* 'dare-2sg', *muost* 'must-2sg', *weist* 'know-2sg', and the 2sg of 'be' *bist*, which resulted from an independent and earlier development, cf. Paul 1879, Lühr 1984).

	Old paradigm	New paradigm
1sg	-Ø	-Ø
2sg	-st	-st
3sg	-t	-t
1pl	-an	-an
2pl	-t	-ts
3pl	-ant	-ant

Table 1: Verbal agreement paradigms (pres. indic.), 13th century Bavarian

A look at the shaded lines reveals that prior to the reanalysis, the agreement suffixes for 3sg and 2pl were identical. The reanalysis of the 2pl clitic *-s* as an enlargement of the existing agreement formative 2pl *-t* removed this syncretism from the paradigm, giving rise to fully distinctive 2pl and 3sg markers.

By the 18th century, 3pl and 1pl forms had fallen together in many Bavarian dialects (due to erosion of final *-t* in 3pl forms). In some varieties, the resulting syncretism was eliminated by the development of a new agreement ending 1pl *-ma*, compare the shaded lines in Table 2:⁷

	Old paradigm	New paradigm
1sg	-Ø	-Ø
2sg	-st	-st
3sg	-t	-t
1pl	-an	-ma
2pl	-ts	-ts
3pl	-an(t)	-an(t)

Table 2: Verbal agreement paradigms (pres. indic.), late 18th century Bavarian

These observations suggest that the reanalysis of clitics as agreement markers is connected to the elimination of syncretisms in the paradigm. This is exactly what we expect under the assumption that the acquisition (and grammaticalization) of inflectional morphology is governed by blocking constraints which operate during language acquisition and scan the input for the most specific realization of a given agreement morpheme. In the following, I show that the new agreement suffixes 2pl *-ts*/, 1pl *-ma*/ satisfy the Blocking Principle since they realize a greater subset of agreement features than their respective predecessors (cf. Fuß 2005 for details).

⁷ Some of these dialects are spoken in the Bavarian Forest, in an area the boundaries of which are (roughly) marked by Cham in the west, Lam in the east, Furth i.W. in the north and Kötzting in the south (cf. Pfalz 1918, Kollmer 1987; Wiesinger 1989, Weiß 1998).

[+speaker, +hearer]	1st person inclusive
[+speaker, -hearer]	1st person exclusive
[-speaker, +hearer]	2nd person
[-speaker, -hearer]	3rd person

Table 3: Binary system of person features

Adopting the binary system of person features illustrated in Table 3, (cf. Benveniste 1950, 1966, Halle 1997, Noyer 1997, Harley and Ritter 2002, Cysouw 2003, and many others), the relevant changes can be accounted for in terms of blocking effects in the following way. First, the new formative */-ts/* is unambiguously specified for both 2nd person (i.e., *[-speaker], [+hearer]*) and number (*[+pl]*), while the former exponent */-t/* is clearly underspecified for number since it occurs in both 3sg and 2pl contexts. */-t/* may be linked to a person specification *[-speaker]*, though, since this is the feature common to both 1st and 3rd person contexts (see section 3.1.1 below for some discussion):

(6) *New 2pl /-ts/ vs. old 2pl /-t/*

- a. *[-speaker, +hearer, +pl]* \leftrightarrow */-ts/*
- b. *[-speaker]* \leftrightarrow */-t/*

The later change affecting 1pl also proceeded in line with the BP. Note that */-ma/* signals first person (i.e., *[+speaker], [-hearer]*) and number (*[+pl]*), while *-an* is presumably only specified for number since it occurs in both 1pl and 3pl contexts:⁸

(7) *New 1pl /-ma/ vs. old 1pl /-an/*

- a. *[+speaker, -hearer, +pl]* \leftrightarrow */-ma/*
- b. *[+pl]* \leftrightarrow */-an/*

Thus, the BP makes available an explanation of why the rise of new agreement formatives took place in some contexts, but not in others: The relevant

⁸ It appears that the form */-an/* may also be specified for *[-hearer]*, which characterizes both first and third person forms. However, this additional specification is not necessary to predict the distribution of */-an/* and should therefore be avoided (under the assumption that the lexicon contains the least numbers of features/elements necessary to derive the forms of a paradigm, cf. Halle 1997 and section 3.1 below). Furthermore, note that even if we took */-an/* to be specified for person, the new formative */-ma/* would still be more specific than */-an/* since it is specified for *[+speaker]* and *[-hearer]*, that is, it unambiguously identifies 1st person.

grammaticalization processes could affect only contexts where the potential new agreement formative was more distinctive than the existing marker.

2.1.1 Properties of blocking-induced change

From the above discussion it is clear that blocking-induced changes select the most specific marker of a set of candidates (robustly) attested in the PLD, dismissing other potential (less specified) realizations of the same inflectional category. Thus, it is a characteristic property of blocking-induced change that the resulting grammar produces less linguistic variation than the target grammar. In somewhat more formal terms, this can be stated as in (8).

(8) The PLD contains more than a single potential phonological realization of an inflectional category X with features $\{F_1, F_2 \dots F_n\}$:

$$\left. \begin{array}{l} / \alpha / \leftrightarrow [x F_1 \dots F_i] \\ / \beta / \leftrightarrow [x F'_1 \dots F'_j] \end{array} \right\} \quad \begin{array}{l} / \beta / \leftrightarrow [x F'_1 \dots F'_j] \\ (\text{selected by the BP if } | \{F'_1 \dots F'_j\} | > | \{F_1 \dots F_i\} |) \end{array}$$

Suppose the learner is confronted with two potential realizations ($/ \alpha /$ and $/ \beta /$) of a given underlying abstract morpheme X that contains a set of morphosyntactic features $\{F_1, F_2 \dots F_n\}$. All other things being equal, the BP will ensure that $/ \beta /$ is stored as the Vocabulary item realizing X if the cardinality of the set of features realized by exponent $/ \beta /$ is greater than the cardinality of the set of features realized by $/ \alpha /$. Note that it is likely that the application of the BP may be preempted by other factors such as relative frequency of the competing formatives. As pointed out in Fuß (2005: 287), the more specified exponent can only be acquired if it is robustly attested in the PLD. Accordingly, if a less specified form is much more frequent than the more specified form, then the learner will probably acquire the less specified form, despite the workings of the BP. Furthermore, as will be discussed in the following section, analogical change is another factor that might work against the effects of the BP: In the absence of robust evidence for a irregular/more specified form, the learner may acquire a regular/underspecified form for a given verb as a default. Evidence from language acquisition (Prasada and Pinker 1993, Clark 1998) and language change (Taylor 1994) shows that this process affects primarily verbs that are less frequent in the PLD, a fact which seems to be in line with the assumption that the BP selects between forms that are robustly attested in the input.

However, it can be shown that the less specified alternates do not disappear at once. In many cases, the new and more specified variants are first confined to certain contexts before they replace the older formatives entirely. Thus, blocking-induced grammaticalization processes may introduce systematic variation between old and innovated forms. Again, this facet of blocking-induced change can be illustrated with the rise of new agreement formatives in Bavarian (cf. Fuß 2005). In particular, we can observe that in a set of North-eastern varieties, the old and the new agreement markers seem to occur in complementary distribution. Finite verbs fronted to second position (and complementizers) systematically carry the new ending 2pl /-ts/, while verbs in clause-final position still exhibit the older ending /-t/:

(9) [Wei-ts iwə t'pruk khum-t] sea-ts s'wiətshaus.
 when-2PL over the-bridge come-2PL see-2PL the-tavern
 'When you cross the bridge, you'll see the tavern.'
 (variety spoken in Lauterbach/Sangerberg; Pfalz 1918: 232)

The different distribution of the old and new variant of the 2pl formative suggest that the change in question has reached a stage where the relevant inflectional formatives are not in free variation. Rather, we deal with an instance of contextual allomorphy, where the new ending is initially confined to a certain structural position (i.e., agreement morphemes located in the C-head), before it turns into a general realization of 2sg verbal agreement, spreading to verbs in clause-final position (see Fuß 2005 for details).⁹

Summing up, this section has argued that grammaticalization processes are shaped by an acquisition strategy (the so-called Blocking Principle) that favors the acquisition of more specified over less specified inflectional formatives. In case the linguistic input contains more than a single potential phonological realization of a abstract inflectional morpheme, the BP will ensure that the most specified candidate is stored as a Vocabulary items realizing that morpheme, all other things being equal (relative frequency of the candidates etc.). I have further argued that blocking-induced change typically reduces linguistic variation attested in the input, giving rise to either the loss of underspecified, less distinctive forms, or leading to contextual

⁹ Alternatively, the new endings may be confined to a certain morphological context before they gain a wider distribution. For example, the new ending 2sg /-st/, which developed in early OHG, first showed up systematically in the present indicative (cf. Brinkmann 1931).

allomorphy, where the competing forms are associated with different insertion contexts. Of course, this raises the question of how morphological variants enter the grammar in the first place (apart from phonological erosion that blurs categorial differences between e.g. agreement formatives and former clitic pronouns). In the next section, I am going to argue that one particular source of linguistic variation are analogical changes that expand the domain of less distinctive formatives in a paradigm.

3 Analogical change

It is a well-known fact that analogical change may create regular variants of originally irregular forms, as illustrated in (10) and (11) with alternating regular and irregular past tense forms and past participles in German:

- (10) a. buk (irreg.) vs. backte (regular) 'I/he/she/it baked'
- b. glomm (irreg.) vs. glimmte (regular) 'I/he/she/it glowed'

- (11) a. gegoren (irreg.) vs. gegärt (regular) 'fermented (participle)'
- b. geblichen (irreg.) vs. gebleicht (regular) 'bleached (participle)'

As already briefly noted above, innovations giving rise to more regular formatives usually affect less frequent forms more readily (e.g. due to overgeneralizations that fail to be corrected). Furthermore, analogical changes typically lead to more uniformity among forms organized in a paradigm. In (10) and (11), for example, we can observe reduction to a single stem form via the elimination of stem vowel alternations (see Albright 2002, Fuß 2005). In frameworks such as Natural Morphology, the apparent drift towards more uniformity in a paradigm is often analyzed as a natural development toward a one-to-one correspondence between form and meaning/function (cf. e.g. Mayerthaler 1980).

This section seeks to develop a formal account of the phenomenon of analogical change. I am going to propose that at least a subset of apparent analogical changes is triggered by an acquisition strategy that aims at minimizing the number of features/elements stored in the lexicon. Moreover, it will become clear that the drift toward a one-to-one correspondence between form and meaning/function is in fact an epiphenomenon resulting from the workings of this acquisition strategy.

Another issue I am going to address concerns the relationship between blocking-induced change and analogical change. It is immediately clear that the phenomenon of analogical change raises an issue for the claims put forward in the previous section since it is usual the regular, less specified form that wins out over the irregular form, and not vice versa. We will see that this apparent tension can be solved if we take a closer look at the contexts in which the different types of changes are set off. The next section deals with one type of analogical change which seems to be particularly problematic from the perspective of Blocking Principle, namely instances of paradigm leveling where less distinctive formatives gain a wider distribution in a paradigm, replacing forms that are apparently more distinctive.¹⁰

3.1 The development of *Einheitsplural* in Alemannic

The development of *Einheitsplural* (henceforth ‘common plural’) in Alemannic varieties of German constitutes a particularly instructive example of analogical leveling expanding the domain of less distinctive formatives in paradigm. As is well-known, most Alemannic dialects spoken in Switzerland and Southwest Germany exhibit only a single plural agreement formative /-ə(n)t/, which originated from the 3pl *-ent* (via vowel reduction and, in some varieties, elision of /n/):

	Present indicative
1sg	-ə
2sg	-ʃ
3sg	-t
1pl	-ə(n)t
2pl	-ə(n)t
3pl	-ə(n)t

Table 4: *Einheitsplural* in Alemannic

Table 5 gives a rough overview of the different historical stages that eventually led to the paradigm in Table 4. The rise of the common plural began already in the Old High German (OHG) period after 3sg (previously /-it/ with strong verbs and weak verbs of class I) and 2pl had fallen together in *-et*, due to a general reduction of vowels in non-stressed (final) syllables, consider the second column in Table 5 (attested in the works of Otfried, mid-8th century). Interestingly, it appears that this

¹⁰ Note that leveling via phonological erosion is not a real issue here: When reduction processes lead to the erosion of inflectional distinctions, the relevant forms simply disappear from the input and fail to be acquired.

change, which was driven by phonological erosion, led to the very same set of distinctions that marked the outset of the changes that took place in Bavarian (see section 2.1). However, in contrast to Bavarian, Alemannic did not choose to eliminate the syncretism of 3sg with 2pl via grammaticalizing a new 2pl formative.¹¹ Instead, it gradually extended the original 3pl form to all plural contexts.

	Original paradigm (Early OHG, ca. 800)	1pl <i>-mēs</i> → <i>-en</i> (Otfrid, ca. 865)	2pl <i>-et</i> → <i>-ent</i> (Notker, OHG/ Alem., ca. 1000)	1pl <i>-ēn</i> → <i>-ent</i> (MHG/Alem., 13th- 15th cent.)
1sg	-u	-u	-o	-e(n)
2sg	-is	-ist	-est	-eſ(t)
3sg	-it	-it (→ -et)	-et	-(e)t
1pl	-mēs	-ēn ¹²	-ēn	-ent
2pl	-et	-et	-ent	-ent
3pl	-ent	-ent	-ent	-ent

Table 5: The development of *Einheitsplural* in OHG/Alemannic, present indicative (inflections of strong verbs and weak verbs of class I, including theme vowels)

Upon closer inspection, it becomes clear that the development of a ‘common plural’ proceeded via two major stages (cf. Braune and Reiffenstein¹⁵ 2004: 263, Paul²⁴ 1998: 240, Paul 1952: 194, and in particular Weinhold 1863: 332ff., Schirmunski 1962: 521ff., Besch 1967: 310ff.): In a first change, the 3pl ending *-nt* replaced the former 2pl *-t*. The earliest instances of 2pl *-nt* are attested in 8th and 9th century OHG (in the Paris and St. Gallen manuscripts of the glossary of Abrogans; all of the earliest forms are imperatives: *haffent*, *dannent* (Paris), *firnēmant* (St. Gallen)). The innovation is mostly confined to Alemannic varieties of OHG (although there are also some relevant examples in the OHG *Tatian*, cf. Sievers 1961). In the work of Notker (950-1022), the new 2pl formative is found consistently in all tenses and moods. Then, in the Middle High German (MHG) period, *-nt* spread to 1pl (formerly *-ēn*), leading to the complete loss of person distinctions in the plural part of the verbal agreement paradigm.

¹¹ Presumably, Alemannic failed to grammaticalize a new, more distinctive 2pl formative since it lacked an appropriate pronominal source. First of all, the relevant 2pl clitic *er* (full pronoun: *ir*) was very similar to 3sg.masc and therefore perhaps not distinctive enough for the purposes of the Blocking Principle. Moreover, the reanalysis of the clitic as part of the verbal agreement ending was perhaps hindered by the fact that the relevant reanalysis (giving rise to a new 2pl formative */-tir/) would have changed both the syllable structure and the accent structure of the verbs affected by that change (in contrast to Bavarian, where the relevant properties were largely unaffected by the reanalysis of the consonantal onsets of the subject clitics).

¹² It is commonly assumed that the 1pl *-en*, which replaced *-mēs*, originated in the subjunctive 1pl *-(e)m*.

Traditionally, the rise of the common plural is analyzed as an analogical change on the model of the 3pl (cf. e.g. Weinhold 1863, Braune and Reiffenstein¹⁵ 2004: 263). However, even if we accept an explanation in terms of analogy, certain open questions remain. For example, we might ask why Alemannic chose to innovate 2pl on the model of 3pl (and not vice versa). In what follows, I am going to argue that we can gain a deeper understanding of the historical developments that led to the rise of the common plural if we take a closer look at the feature specifications of the individual Vocabulary items that are part of the verbal agreement paradigm, and the way these Vocabulary items (and their feature specifications) are acquired by the learner. In particular, I claim that the relevant 'analogical' changes were triggered by an acquisition strategy that aims at minimizing the number of elements/features mentioned in the lexicon (which may lead to the impression of a more transparent relation between form and function/meaning).

3.1.1 The extension of 3pl *-nt* to 2pl

This section focuses on the first stage of the development of the common plural in Alemannic, that is, the early (OHG) change in which the 3pl marker /-nt/ was extended to 2pl contexts, replacing the original 2pl marker /-t/. Table 6 lists the forms of the verbal agreement paradigm of OHG before and after the extension of 3pl to 2pl (note that /e/ is merely a theme vowel):

	Paradigm after conflation of 3sg, 2pl	Paradigm after 3pl → 2pl
1sg	-o	-o
2sg	-est	-est
3sg	-et	-et
1pl	-ēn	-ēn
2pl	-et	-ent
3pl	-ent	-ent

Table 6: 2pl /-t/ → /-nt/ (pres.indic) in OHG/Early Alemannic¹³

As already noted, traditional approaches treat this change as a typical case of analogical leveling. Thus, it is usually assumed that the exponent of 2pl has been reshaped on the model of the formative realizing 3pl. While this is a possible account, it leaves many aspects of the change in the dark. For example, it says nothing about possible motivations that might lead speakers to favor /-nt/ over /-t/ as a

¹³ Note that at this stage, the theme vowel still served to distinguish conjugation classes (cf. class II: 3pl *-ont*, class III: 3pl *-ēnt*).

realization of 2pl. In particular, the fact that the change proceeded in the way it did, replacing the 2pl formative with the 3pl formative, comes out as completely accidental. Under the analogy approach, it could also have been the other way around. Thus, it appears that an analysis appealing to the notion of analogical change lacks explanatory force. Of course, it is not always possible to give principled explanations for all aspects of language change (cf. Lightfoot 1999 on this point), but at least we should try and see if we can do better than simply invoking analogy. In what follows, I am going to argue that it is in fact possible to isolate a set of factors that possibly acted as a driving force in the emergence of the common plural if we examine the individual Vocabulary items listed in Table 6 and their feature specifications in some more detail. What I am going to claim is that the extension of 3pl /nt/ to 2pl was part of a major reanalysis that affected the form-function pairings in the verbal agreement paradigm of early Alemannic.

My proposal is based on two assumptions. First, I adopt the system of person features proposed above, that is, I assume that the traditional 3-way distinction for person features must be decomposed by using the binary features [±speaker] and [±hearer], repeated here for convenience:

[+speaker, +hearer]	1 st person inclusive
[+speaker, -hearer]	1 st person exclusive
[-speaker, +hearer]	2 nd person
[-speaker, -hearer]	3 rd person

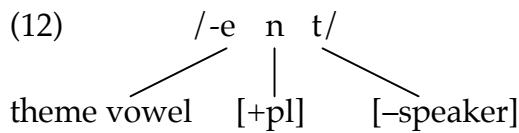
Table 7: Binary system of person features

In addition, I will assume a lexical decomposition analysis in which the traditional inflectional markers of the agreement paradigm are split up into smaller units of exponence (cf. e.g. Wiese 1994, Müller 2006a, 2006b on German). Under these assumptions, the extension of 3pl /nt/ to 2pl can be analyzed as the result of two separate changes.

First, let us suppose that the earlier innovation of 2sg -st (inherited ending /-s/ + onset of subject clitic *thu*) made available a reanalysis of the segment /t/ as a realization of the feature [-speaker] since final /t/ appears in all 2nd and 3rd person forms (cf. Table 5 and Table 6).¹⁴ An additional change led to nasalization of 2pl

¹⁴ The absence of /-t/ in 3sg person preterite forms can be attributed to an Impoverishment rule that deletes the feature [-speaker] in the relevant context ([−hearer, −pl, +past]), giving rise to identical 1sg and 3sg preterite forms in these dialects (cf. Müller 2006b: 104 for a related analysis of the conflation of 1st and 3rd person preterite forms in Standard German). Note that the present-day Alemannic dialects generally lack preterite forms (preterite forms began to disappear in the 16th century).

forms ($/-t/ \rightarrow /-nt/$), which seems to be a common strategy across Alemannic to reinforce/strengthen inflectional formatives (cf. Weinhold 1863). Note that after nasalization of 2pl, the segment $/n/$ could be analyzed as being uniquely paired with the inflectional feature $[+pl]$, since it occurs in all plural forms and nowhere else. Thus, at some point, the distribution of $/t/$ and $/n/$ in the verbal agreement paradigm led to a reanalysis that affected the feature specifications of these segments. The result of this change is illustrated in (12), where the inflectional marker $/-ent/$ (including a theme vowel) is decomposed as a combination of smaller phonological exponents:¹⁵



Interestingly, there are reasons to believe that the nasalization of 2pl was promoted by additional factors apart from merely phonetic reinforcement. If we apply lexical decomposition to all inflectional formatives of the verbal agreement paradigm, then it appears that the innovation of 2pl $/-nt/$ gave rise to a more transparent relationship between form and function/meaning via creating a phonological exponent which was uniquely paired with the feature $[+pl]$ (compare the shaded line in Table 8).

Old feature specification	Exponent	New feature specification
$[+speaker, -pl]$	$/-o/$ ¹⁶	$[+speaker, -pl]$
$[+hearer, -pl]$	$/-s/$	$[+hearer, -pl]$
$[-hearer, +pl]$	$/-n/$	$[+pl]$
$[-speaker]$	$/-t/$	$[-speaker]$

Table 8: Reanalysis giving rise to 2pl $/-nt/$, OHG/early Alemannic

From a functionalist perspective, the change in question certainly led to a welcome result (cf. conditions on analogical change proposed in the framework of Natural

¹⁵ Decomposing the relevant agreement markers requires that the relevant inflectional head may split up into several insertion sites prior to the insertion of phonological exponents (so-called 'Fission', see e.g. Noyer 1997 and chapter 2 above). Under the assumption that Vocabulary Insertion discharges morphosyntactic features of the underlying morpheme, exponents compatible with the (remaining) feature set may be inserted as long as there are features left that can be discharged. See Müller (2006a), (2006b) for a related analysis of the verbal inflection of Standard German.

¹⁶ If $/-o/$ is analyzed as a theme vowel, the combination of features underlying 1sg is realized by \emptyset , the zero exponent.

Morphology, e.g., Mayerthaler 1980). However, it appears that we do not need to appeal to functionalist notions in order to explain this effect. As it turns out, it can also be modeled in purely formal terms if we assume that the acquisition of phonological exponents and their feature specifications is subject to the following constraint (cf. Halle 1997: 430):

(13) *Minimize Feature Content*

The number of features mentioned in the Vocabulary [i.e., the lexicon, EF] must be minimized.

According to (13), child learners acquire the most economical lexical inventory compatible with the input they are exposed to.¹⁷ This has the following two consequences for the acquisition of phonological exponents and their featural properties. First, the set of lexical entries/Vocabulary items stored in the lexicon consists of the minimal number of formatives required for generating the input. Second, each inflectional marker is associated with the most economical feature specification compatible with the input data. In other words, the learner acquires the minimal set of feature specifications that is necessary for deriving the distribution of a given phonological exponent/Vocabulary item.

Interestingly, upon closer inspection it turns out that the workings of (13) may also lead to a more transparent relation between form and function/meaning, in particular if inflectional markers are decomposed into smaller units of exponence, as in (12): The smaller the individual units of exponence are, the more likely it is that (13) leads to a one-to-one relation between form and meaning. The development of a unique plural formative (without an additional [person] specification, compare Table 8) clearly is a case in point. In other words, the fact that the change in question led to a more transparent relationship between form and function/meaning was not a driving force, but rather merely an epiphenomenal outcome of the reanalyses giving rise to (12). We might suspect that the in-built tendency to posit an economical system of featural distinctions may lead learners to coin new variants that are not part of the input (or associated with a different feature specification in the target grammar), but comply with (13):

¹⁷ According to Halle (1997: 430), independent motivation for this constraint comes from considerations of memory load: "Such an economy constraint is entirely plausible, because the Vocabulary entries represent items that speakers must memorize, and since our memories are finite, the load on memory must be minimized."

(14) The learner innovates a regular/less specified phonological exponent /β/ of an inflectional category X with features {F₁, F₂ ... F_n} (an irregular/more specified form /α/ may be part of the input):

$$\left. \left(/ \alpha / \leftrightarrow [_{\mathbf{x}} F_1 \dots F_i] \right) \right\} \begin{array}{l} / \beta / \leftrightarrow [_{\mathbf{x}} F'_1 \dots F'_j] \\ \text{(innovated form, with } | \{F'_1 \dots F'_j\} | \leq | \{F_1 \dots F_i\} | \text{)} \\ / \alpha / \leftrightarrow [_{\mathbf{x}} F_1 \dots F_i] \end{array}$$

(14) states that 'analogical' change may introduce new variants formerly absent in the grammar and not attested in the linguistic input the learner receives. Crucially, the innovations are typically more regular/less distinctive than the existing forms, which is captured by the statement that the cardinality of the set of features realized by the innovating form /β/ is smaller than (or equal to) the cardinality of the set of features realized by the existing Vocabulary item /α/. The innovating form may result from overgeneralization (after the learner has mastered the relevant inflectional rule) or from the workings of the acquisition strategy (13), which compels the learner to minimize the number of features mentioned in the Vocabulary. In this way, (13) may promote the introduction of new, more economical variants that enter into a competition with older formatives that are more distinctive (i.e., specified for more features), but 'harder' to acquire, since the relation between exponent and features is one-to-many. Over time, the new variant may win out over the original Vocabulary item (possibly in a word-by-word fashion), giving rise to the effect of paradigm leveling when the change has been completed (see below for more discussion and the interaction between (13) and morphological blocking).

In the next section, I am going to examine the second historical stage of the development of the *Einheitsplural* in Alemannic, in which the formative /nt/ was extended to 1pl contexts during the MHG period.

3.1.2 The rise of a general plural marker: Extension to 1pl

Between the 13th and 15th century, /-(e)nt/ evolved into the general plural marker for all persons, replacing 1pl /-(e)n/ (cf. Weinhold 1863: 366, Schirmunski 1962: 521ff., Besch 1967: 310ff., Paul ²⁴1998: 240). This change is represented in the following table, which lists the relevant forms (with theme vowels):

	Old paradigm	New paradigm
1sg	-e	-e(n)
2sg	-st	-ſ(t)
3sg	-et	-et
1pl	-en	-ent
2pl	-ent	-ent
3pl	-ent	-ent

Table 9: 1pl /-en/ → /-ent/ (pres.indic) in MHG/Alemannic (~13th-15th century)

Again, we may ask whether the extension to 1pl contexts was merely an analogical change in the traditional sense, or whether it is possible to identify some deeper motivation for rise of the *Einheitsplural*. More to the point, the extension of /nt/ to 1pl can be connected to two other changes that altered the make-up of the verbal agreement paradigm of Alemannic. In particular, there are reasons to believe that the changes affecting the shape of the exponents of 1sg and 2sg required a major reorganization of the form-function pairings in other parts of the paradigm.

More precisely, what we can observe is that the extension of /-nt/ to 1pl was accompanied by the loss of final /-t/ in 2sg contexts (presumably due to phonological erosion) and nasalization of 1sg giving rise to a new 1sg exponent /-(e)n/.¹⁸ Interestingly, it seems that there is a systematic connection between the three changes highlighted by shading in Table 9. First of all, we can observe that the phenomenon of ‘common plural’ is also a characteristic of Low German dialects: Western Low German dialects exhibit the form /-(ə)t/, while /-ən/ is the typical ending found in Eastern Low German dialects (cf. Schirmunski 1962: 543f. for details). Interestingly, many of these dialects also exhibit loss of final /-t/ in 2sg forms, similar to Alemannic (Schirmunski 1962: 544).¹⁹ Second, Besch (1967: 301) observes that there is a geographic connection between the extension of the *Einheitsplural* to 1pl and the presence of the 1sg form /-(e)n/, in the sense that in the 15th century, 1sg /-n/ is found in particular in those dialectal areas that also participated in the development of the *Einheitsplural*.²⁰ Thus, we may conclude that

¹⁸ 2sg /-st/ → /-ſ/ after /-st/ → /-ſt/ in most varieties; cf. Weinhold (1863: 365), Schirmunski (1962: 520f.), Weber (1987: 174). The nasalization of 1sg forms is traditionally analyzed as an extension of the relevant 1sg ending of the weak verbs of classes II and III, cf. Schirmunski (1962: 519).

¹⁹ The possible connection between the loss of 2sg /-t/ and the rise of the common plural /-nt/ was pointed out to me by Helmut Weiß.

²⁰ Compare the following quote taken from Besch (1967: 301):

the joint appearance of (i) the changes affecting the 1sg/2sg forms and (ii) the rise of the *Einheitsplural* is not coincidental.

Let's now address the question of how the apparent link between these changes can be modeled in a more formal way. First of all, note that due to the loss of final /-t/ in 2sg, /-t/ could no longer be analyzed as a marker realizing the feature [-speaker] (otherwise we would expect /-t/ to occur in all 2nd and 3rd person contexts). Thus it appears that this change not only affected the shape of 2sg forms, but also had a considerable impact on the whole system of feature distinctions that underlies the verbal agreement paradigm. In a similar vein, nasalization of 1sg leading to 1sg /-en/ did not only alter the shape of 1sg forms, but also affected the system of form-function pairings in the plural part of the paradigm since /-n/ could no longer be analyzed as the realization of [+pl].

In other words, it appears that the changes that altered the shape of the 1sg and 2sg formatives led to a major reorganization of the verbal agreement paradigm in the relevant varieties of Alemannic. After /-t/ and /-n/ could no longer be paired with a unique feature value, the 'analogical' extension of /-nt/ to 1pl facilitated a reanalysis of the combination /-nt/ as a pure plural marker, with /-t/ turning into the elsewhere marker.²¹ Furthermore, the systematic absence of person distinctions in the plural suggests an analysis in terms of an Impoverishment rule that deletes person features in the context [+pl] (see Halle 1997, Noyer 1997, Embick & Noyer 2007 and chapter 2 above on the notion of Impoverishment):

(15) *Impoverishment in Alemannic (Einheitsplural)*

$$[\pm\text{speaker}], [\pm\text{hearer}] \rightarrow \emptyset / [+pl]$$

It seems likely that the development of the Impoverishment rule in (15) was promoted by *Minimize Feature Content* since Impoverishment typically serves to expand the domain of less specified (and therefore less costly) exponents (see e.g.

"[1sg] -n hält sich im 15. Jh. besonders dort, wo sich in den Quellen der Einheitsplural auf -nt durchsetzt, es fehlt ganz in den Schreiblandschaften, die stark zu einem -n Plural (beim Verbum) neigen [...]"

'In the 15th century, [1sg] -n is preserved in particular in those areas where the common plural ending -nt gains acceptance in the historical records. It is completely absent in areas which strongly prefer the -n plural with verbs [...]' (translation: EF)

²¹ The extension of /-nt/ to 1pl was probably facilitated by the fact that the endings /-n/ and /-nt/ are very similar phonetically.

Noyer 1997 on the role of Impoverishment in language change). As illustrated in (16), this set of changes eventually led to a highly economical agreement paradigm, where each phonological exponent is uniquely paired with a single syntactico-semantic feature. Again, this result is clearly in line with the constraints on language acquisition imposed by *Minimize Feature Content*.²²

(16) a. [+speaker] \leftrightarrow /-n/
 b. [+hearer] \leftrightarrow /-ʃ/
 c. [+pl] \leftrightarrow /-nt/
 d. elsewhere \leftrightarrow /-t/

The *Subset Principle* (see chapter 2 above, repeated here for convenience) guarantees that the phonological exponents realizing person features (1sg /-n/, 2sg /-ʃ/) cannot be inserted in plural contexts, since they contain features not present in the agreement morpheme after Impoverishment has taken place.

(17) *The Subset Principle* (Halle 1997: 428)

The phonological exponent of a Vocabulary item is inserted into a morpheme in the terminal string if the item matches all of a subset of the grammatical features specified in the terminal morpheme. Insertion does not take place if the Vocabulary item contains features not present in the morpheme. Where several Vocabulary items meet the conditions for insertion, the item matching the greatest number of features specified in the terminal morpheme must be chosen.

As a result, the relevant Vocabulary items need not be specified for number, giving rise to an one-to-one relation between form and function. Crucially, this 'optimal' outcome can be analyzed in purely formal terms via attributing the relevant changes to the interaction between learning strategies (*Minimize Feature Content*) and operations of the phonological component (Impoverishment), without appealing to any functionalist considerations.

In sum, it seems that learners can slightly rearrange the form/function pairings they encounter in the input if the outcome complies with *Minimize Feature Content*. In

²² Note that this analysis raises a number of issues which I cannot address in detail here. For example, more has to be said about (i) the status of the elsewhere marker /-t/ if Agr is still subject to Fission at this stage, and (ii) the status of Impoverishment in the relevant preterite forms, cf. fn. 14 (but note that the preterite began to diminish shortly after the rise of the common plural).

the case at hand, this can be taken to have motivated the enlargement of 1pl via adding final /t/, giving rise to a common plural /-nt/ and the paradigm in (16), which represents the most economical *and* transparent linking of form and function/meaning compatible with the input after /-t/ and /-n/ could no longer be analyzed as marking [–speaker] and [+pl], respectively.²³ Again, it turns out that an ‘optimal’ system that involves a one-to-one relation between form and function may evolve as a side-effect of other, purely formal mechanisms that govern the acquisition of Vocabulary items and their feature specifications.²⁴

3.2 Analogical change vs. blocking-induced change

This section focuses on the question of how the apparently conflicting properties of blocking-induced change and (‘analogical’) change driven by *Minimize Feature Content* can be reconciled. As noted above, we can characterize the latter as follows:

²³ Note that this seems to be reminiscent of the Transparency Principle, which according to Lightfoot (1979) may set off a reanalysis in case the acquisition task becomes too complicated (due to the presence of too many exceptional features or rules). In fact, we can observe that an alternative paradigm that lacks a common plural and incorporates /n/ as a marker for both 1sg and 1pl turns out to be much more complex than (16) (due to the fact that it prevents the application of Impoverishment). In particular, /-ʃ/ would turn out as being specified for [+hearer, -pl] while /nt/ would correspond to [–speaker, +pl]:

- (i) a. [+speaker] \leftrightarrow /-n/
- b. [+hearer, -pl] \leftrightarrow /-ʃ/
- c. [–speaker, +pl] \leftrightarrow /-nt/
- d. elsewhere \leftrightarrow /-t/

Furthermore, note that the hypothetical system in (i) introduces a number syncretism (1sg/1pl), while the actually attested system in (16) is characterized by a person syncretism (i.e., all plural forms are identical). There are additional reasons to believe that the latter system is preferred over (i). Corbett (2000: 277) observes that there is a cross-linguistic tendency to signal number distinctions in the 1st person (“if number is not available for all persons, then it will be found first of all in the first person, then in the second, and in the third only if in both first and second also [...]”). The preference for having number distinctions in 1st person forms can possibly be modeled as a blocking effect that prefers the acquisition of distinct 1sg and 1pl forms, even if this leads to syncretism in another part of the paradigm (in the case at hand, person syncretism in the plural, i.e., *Einheitsplural*).

²⁴ Later changes, which are most probably due to purely phonological factors, then led to the present-day paradigm (in some varieties, cf. Table 4): (i) cluster reduction of /nt/ via elision of /n/ (cross-linguistically a common change, which is usually attributed to a tendency to preserve the least sonorous element of the target cluster, cf. Ohala 1996, 1999, Pater and Barlow 2003); (ii) loss of final /-n/ in 1sg forms (in most dialects).

(18) The learner innovates a regular/less specified phonological exponent /β/ of an inflectional category X with features {F₁, F₂ ... F_n} (an irregular/more specified form /α/ may be part of the input):

$$(/α/ \leftrightarrow [x F_1 \dots F_i]) \quad \left. \begin{array}{l} /β/ \leftrightarrow [x F'_1 \dots F'_j] \\ \text{(innovated form, with } |F'_1 \dots F'_j| \leq |F_1 \dots F_i| \text{)} \\ (/α/ \leftrightarrow [x F_1 \dots F_i]) \end{array} \right\}$$

This type of ‘analogical’ change introducing new (more regular/less distinctive) variants that were formerly absent in the grammar typically originates at an early point during language acquisition when children begin to master inflectional rules. It is triggered either by overgeneralization (after the learner has mastered the relevant rule) or by an acquisition strategy (*Minimize Feature Content*) that aims at minimizing the number of elements/features stored in the lexicon (and may bring about a more transparent relation between form and function/meaning as a side-effect). Now, the question arises of how the learner reacts when he/she becomes aware of the fact that the input contains a competing Vocabulary item that is apparently more distinctive than the innovating variant. Several scenarios are possible depending on the overall frequency of the older, more marked variant: First, if a more distinctive/irregular formative is robustly attested in the input, it will replace the innovated variant due to blocking effects, and no change will occur. The interesting case involves a scenario where the older form occurs less frequently. If the older form is not frequent enough, the child may fail to acquire it and the innovated form will replace the older form.²⁵ Alternatively the learner may acquire the older form in addition to the innovated variant, giving rise to morphological doublets and linguistic variation. As is well-known, the distribution of doublets introduced by analogical changes is typically determined by extragrammatical factors such as style or register.

In contrast, blocking-induced changes select an inflectional marker out of a set of candidates robustly attested in the input:

²⁵ Recall that analogical change/overgeneralization typically affects less frequent elements, cf. e.g. Prasada and Pinker (1993), Kroch (1994), Clark (1998).

(19) The PLD contains more than a single potential phonological realization of an inflectional category X with features $\{F_1, F_2 \dots F_n\}$:

$$\left. \begin{array}{l} / \alpha / \leftrightarrow [x F_1 \dots F_i] \\ / \beta / \leftrightarrow [x F'_1 \dots F'_j] \end{array} \right\} \quad \begin{array}{l} / \beta / \leftrightarrow [x F'_1 \dots F'_j] \\ \text{(selected by the BP if } | \{F'_1 \dots F'_j\} | > | \{F_1 \dots F_i\} | \text{)} \end{array}$$

Thus, blocking-induced changes originate at a later stage during language acquisition, selecting between candidates robustly attested in the input, effectively reducing linguistic variation (or introducing systematic variation in the form of contextual allomorphy, see above). So it turns out that there are significant differences concerning the contexts where the apparently conflicting acquisition strategies (*Blocking Principle* vs. *Minimize Feature Content*) apply. First of all, while analogical changes typically affect forms that are less frequent and therefore less robustly attested, blocking applies in cases where the input contains robust evidence for more than a single potential candidate realizing a certain abstract morpheme. Crucially, regular/less distinctive forms are always potentially available (due to overgeneralization and acquisition strategies such as *Minimize Feature Content*), while the acquisition of irregular/more marked forms is only possible via the input the child receives. This perhaps explains the apparent predominance of analogical leveling across languages and times.

Furthermore, as pointed out in Fuß (2005: 289), note that *Minimize Feature Content* does not require the learner to select the least marked/specified formative for storage in the lexicon. Rather, it ensures that the child acquires the most economical lexical inventory compatible with the input he/she is exposed to. For example, if a feature specification [+speaker] is sufficient to guarantee that a 1sg exponent is inserted in the contexts where it appears in the input, the child will not acquire a redundant feature specification [+speaker, -hearer] for this exponent (compare (16) above). This function of *Minimize Feature Content* does not interfere with the claim that the learner scans the input for the most marked (and therefore salient) realization of a given inflectional head. In other words, the Blocking Principle ensures that the most specified candidate is selected while *Minimize Feature Content* warrants that this candidate is assigned a non-redundant feature specification. In addition, it seems that blocking operates in a local fashion, comparing two possible candidates for realizing a certain terminal node. In contrast, the discussion of the rise of *Einheitsplural* in Alemannic seems to suggest that the scope of *Minimize Feature Content* is wider, including the featural make-up of whole paradigms.

So we may conclude that the conflict between the different learning strategies is merely apparent. Rather, blocking and *Minimize Feature Content* differ both with respect to their scope and the contexts where they apply. In this way, the two principles may actually be taken to work hand in hand during language acquisition, warranting that the learner selects an optimal paradigm and lexicon structure based on the evidence available to him/her.

3.3 Section summary

This section has argued that at least in the domain of inflectional morphology, the cyclic nature of language change can be attributed to the workings of two apparently conflicting acquisition strategies that help the learner to identify phonological exponents (of inflectional categories) and their feature specifications on the basis of the linguistic input he/she receives.

First, we have seen that grammaticalization processes giving rise to new and more distinctive inflectional formatives are shaped by a learning strategy based on morphological blocking that selects the most specified variant in case the input contains more than a single potential realization of a given inflectional category (the *Blocking Principle*). In this way, blocking-induced change typically leads to a grammar that generates less linguistic variation than the target grammar. I have illustrated the workings of the Blocking Principle with the historical development of verbal agreement marking in Bavarian (cf. Fuß 2005). It has become clear that the relevant grammaticalization processes affected only those slots of the agreement paradigm where the existing markers were less distinctive than the newly coined formatives.

Second, I have presented evidence suggesting that the acquisition of phonological exponents and their feature specifications is influenced by another strategy that aims at minimizing the number of elements/features stored in the lexicon (*Minimize Feature Content*, cf. Halle 1997) and ensures that learners acquire the most economical lexical inventory compatible with the input they are exposed to. The tendency to posit an economical system of featural distinctions may lead learners to innovate 'less costly' variants that are not part of the input. The innovating forms may then replace older, more distinctive forms if the latter fail to be robustly attested in the input the learner receives. The workings of *Minimize Feature Content* have been illustrated with the rise of the so-called *Einheitsplural* in Alemannic, in which the 3pl /nt/ replaced all other plural formatives. Upon closer inspection of the relevant Vocabulary items (applying a lexical decomposition analysis), each of the individual

stages of this change complied with the notion of *Minimize Feature Content*, giving rise to highly economical inventory of lexical entries, in which each phonological exponent is uniquely paired with a single syntactico-semantic feature. I have argued that in contrast to claims in the functionalist literature, this particular outcome (a one-to-one relationship of form and function) was not a driving force behind the relevant set of changes, but rather merely an epiphenomenal result of the workings of *Minimize Feature Content*.

I have then taken a closer look at the interaction of the two apparently conflicting acquisition strategies, arguing that blocking and *Minimize Feature Content* apply in different contexts during language acquisition and that it is in fact possible to reconcile the two strategies as separate devices employed by the learner to select an optimal inventory of Vocabulary items based on the evidence available to him/her.

The next section shows that blocking effects also play a major role in another cyclic change, namely the rise and fall of null subjects.

4 The rise and fall of null subjects

In a couple of languages, we can observe a cluster of diachronic developments involving pronouns, verbal agreement and the pro-drop property. Perhaps the best-studied example of the relevant changes comes from the history of French, where we can observe that the loss and rise of null subjects (and the accompanying changes in the verbal agreement paradigm) seem to take place in a cyclic fashion (cf. e.g. Bally 1965, Guiraud 1968, Wartburg 1970, Ashby 1977, Harris 1978, Lambrecht 1981, Adams 1987, Roberge 1990, Roberts 1993a, 2007b, Zribi-Hertz 1994; see e.g. Poletto 1995 on related changes in Northern Italian dialects). The different historical stages of the pro-drop cycle in French can be characterized as follows:

(20) *Pro-drop cycle in the history of French*

- (i) distinctive verbal Agr/pro-drop (OFr.)
- (ii) loss of Agr \Rightarrow loss of pro-drop (Middle Fr., 14th-16th century)²⁶
- (iii) subject pronouns lose emphatic force and become clitics (15th-18th century)
- (iv) clitics are reanalyzed as agreement markers \Rightarrow rise of pro-drop²⁷

(ongoing change/Colloquial French)

This section argues that the workings of blocking effects can also be detected in the cyclic changes affecting the availability of null subjects. More precisely, it will be shown that the relevant changes not only involve the notion of blocking as a learning strategy, but also as a principle of grammar that blocks the insertion of a less specified Vocabulary item if a more specified candidate is listed in the lexicon (Kiparsky's 1973, 1982 Elsewhere Condition or Halle's 1997 *Subset Principle*). The basic proposal I want to explore is that the rise of agreement-related null arguments may be guided by the following two factors, which both involve the notion of morphological blocking:

- (i) The *Blocking Principle* may trigger the reanalysis of a particular pronominal clitic if the resulting agreement marker is more distinctive than the existing verbal inflection (cf. section 2 above);
- (ii) Changes leading to a gap in the paradigm of weak pronouns trigger *deblocking* of a (universally available) null realization of weak/clitic pronominal forms (assuming that null subjects are to be analyzed as the zero realization of regular pronominal forms, Holmberg 2005, Roberts 2007b).

(ii) predicts that if for independent reasons a more distinctive candidate is lost, a less specified phonological formative (formerly blocked by the presence of a more

²⁶ Note that according to Wartburg (1970: 72) and Harris (1978: 113), the rise of overt pronouns (in Middle French) is not directly related to the loss of agreement morphology; rather it is linked to word order properties and prosodic factors (in fact, Harris claims that subject pronouns became obligatory prior to the erosion of the agreement system).

²⁷ Givón (1976) claims that in Colloquial French, verbal agreement markers have developed via a reanalysis of a former topic left dislocation structure. However, there are at least some indications that the relevant syntactic environment was *not* topic left dislocation, but rather a structure where a reinforcing full form (e.g. the oblique 1sg form *moi*) has been added to the non-stressable clitic for reasons of emphasis/focus (cf. Guiraud 1968, Wartburg 1970, and Ashby 1977 for details).

distinctive overt spell-out) becomes available for realizing a certain functional head. Likewise, the development of a more distinctive spell-out is predicted to drive a less specific realization out of the grammar. These expectations seem to be borne out by data from the historical development of weak pronominal forms in Bavarian, French, and Finnish. More precisely, I am going to argue that historically, a null-spell out of (weak) pronominal D-heads may emerge due to the loss of overt weak pronouns, while the grammaticalization of new overt weak forms may lead to the loss of pro-drop properties formerly present in the grammar.

The possibility that the availability of null arguments is connected to the inventory of overt pronominal forms has been largely neglected by generative approaches to pro-drop phenomena, where it is standardly assumed that there is a correlation between the presence of rich verbal inflection and certain syntactic phenomena, including the availability of null subjects (cf. Rizzi 1982, 1986a, Jaeggli and Safir 1989, Roberts 1993a, Vikner 1997, Rohrbacher 1999, and more recently Müller 2006b, among many others; but see e.g. Gilligan 1987, Haider 1994 for a critical review).²⁸ Note that this view, sometimes referred to as the Rich Agreement Hypothesis (RAH), has strong implications for the historical development of null arguments which are largely independent of the exact theoretical implementation of the RAH: On the one hand, it is expected that the loss of rich verbal agreement leads to the loss of pro-drop. On the other hand, it is expected that historically, pro-drop

²⁸ See however Sprouse and Vance (1999) who link the loss of null subjects in Swiss Rhaeto-Romance varieties (e.g., Surselvan) and in the history of French to the presence of competing overt weak forms in the input the child receives. Sprouse and Vance argue that overt forms are easier to parse than null arguments. In the course of time, this may lead to a loss of the null realization (*Differential Parsing Model of Change Through Competition*, DPM). Note that this approach raises the question of how null realizations can develop in the first place given that they are inherently harder to parse than overt pronouns. Furthermore, this approach does not address the question of how the innovation of new overt forms affects the possibility of pro-drop (i.e., the actual grammar change leading to the competition of null and overt forms). In the approach developed here, the observation that a language may have both null and overt forms can be analyzed as an instance of grammar competition (Kroch 1989), where learners may have access to more than one internalized grammar. The effect observed by Sprouse and Vance may then be modeled by assuming that over time, the grammar with overt pronouns wins out over the competing variant with null forms. Crucially, in an approach making use of grammar competition, the gradual loss of the null variants can be attributed to a parsing preference (governing the choice of grammars) without giving rise to the problem mentioned above (how can null forms develop in the first place?), since a proper distinction is made between language change (loss of competing grammars) and grammar change (innovation giving rise to new overt forms).

emerges when the richness of verbal agreement marking crosses a certain threshold. Moreover, since the 'pro-drop parameter' is generally conceived of as being binary in nature (i.e., referential pro-drop is either generally available or completely absent), the rise of pro-drop is predicted to proceed in an across-the-board fashion, affecting all persons and numbers at once.

While the first prediction seems to be borne out by the facts (at least to some extent, in the sense that there are cases where the loss of null subjects has been preceded by the loss of agreement distinctions, cf. Falk 1993 on the history of Swedish, Adams 1987, Roberts 1993a, among many others, on French; but see e.g. Sprouse and Vance 1999 for a critical evaluation), the second implication has received less attention in the literature (but see Haider 1994 and Wratil 2008 for some discussion).

In the following, I show that the historical development of null subjects in Bavarian, French, and Finnish is at odds with the predictions of the RAH,²⁹ but can be successfully captured by an approach that analyzes the ups and downs on the pro-drop cycle in terms of blocking and deblocking of a universally available null spell-out of arguments.

4.1 The rise and loss of partial pro-drop

It has variously been suggested that null subjects may develop as a by-product of the reanalysis of pronominal clitics as verbal agreement markers (see e.g. Rizzi 1986b, Brandi and Cordin 1989, Haider 1994, Poletto 1995, Roberts and Roussou 2003 on Northern Italian dialects; Roberge 1990, Auger 1993, 1994a, Zribi-Hertz 1994, Fonseca-Greber 2000, Fonseca-Greber and Waugh 2003, Gerlach 2002 on Non-Standard French). However, the theoretical implications of these observations for the theory of pro-drop has largely been left unaddressed (but see Haider 1994, Roberts and Roussou 2003: 185f., and Wratil 2008). In this section, I demonstrate that the development of partial pro-drop in Bavarian and non-standard varieties of French raises a couple of questions for the traditional notion that there is a one-to-one relation between rich verbal agreement and the presence of null-subjects (see the appendix for another set of problematic data coming from the development of pro-drop in Creoles). In particular, it will become clear that the rise of agreement-related pro-drop typically takes place in a piecemeal fashion, that is, referential pro-drop is

²⁹ See appendix I for an alternative pathway to (non-agreement-related) null arguments that can be observed in a number of Creole languages.

initially restricted to certain slots of the paradigm, before it eventually extends to all persons and numbers. This is illustrated in (21) and (22) with examples from present-day Bavarian, where referential pro-drop is restricted to second person (plus 1pl in some dialects, see below) (cf. e.g. Bayer 1984, Weiß 1998, 2002, 2006):

(21) a. Kummst noch Minga, dann muaßt me b'suacha.
 come-2SG to Munich then must-2SG me visit
 'If (you) come to Munich (you) must visit me.'
 b. Kummts noch Minga, dann müaßts me b'suacha.
 come-2PL to Munich then must-2PL me visit
 'If (you.PL) come to Munich (you.PL) must visit me.'

(22) a. *Kumm noch Minga?
 come-1SG to Munich
 'Will (I) come to Munich?'
 b. *Kumm-t noch Minga?
 come-3SG to Munich
 'Will (he/she/it) come to Munich?'

Interestingly, these are the very same contexts in which Bavarian exhibits the phenomenon of complementizer agreement:³⁰

(23) a. ob-st (du) noch Minga kumm-st
 whether-2SG you.SG to Munich come-2SG
 'whether you come to Munich'

³⁰ Further languages that exhibit the phenomenon of partial pro-drop include Finnish (Vainikka and Levy 1999, Holmberg 2005, Koeneman 2007; see also below) and Frisian, where pro-drop is also limited to the contexts where complementizer agreement is available (2sg):

(i) a. Kom-st (do) jün?
 come-2SG you tonight
 'Do you come tonight?'
 b. dat-st (do) jün kom-st
 that-2SG you tonight come-2SG
 'that you come tonight'
 (Zwart 1993: 256)

b. ob-**ts** (ees/ihr) noch Minga kumm-**ts**
 whether-2PL you.PL to Munich come-2PL
 'whether you(PL) come to Munich'

The formatives *-st*, *-ts* that attach to the complementizer in (23) are obligatorily present and cannot be replaced by the relevant tonic subject pronouns. The latter are only acceptable if they co-occur with *-st/-ts*, cf. (24). This contrasts with the behavior of genuine subject clitics, cf. (25) and can be taken to indicate that the *-st* and *-ts* are not pronominal clitics, but rather inflections. Furthermore, the fact that it is not possible to attach the alleged 2nd person 'clitics' *-st/-ts* to the inflected verb (forms such as 2sg **kummst-st* or 2pl *kummts-ts* are not well-formed) can be taken to indicate that Bavarian lacks 2nd person subject clitics altogether (that is, there are gaps in the paradigm of clitic pronouns; see Altmann 1984, Bayer 1984, Fuß 2005).³¹ Accordingly, the sentences in (21) must be analyzed as instances of referential pro-drop.

(24) a. *ob du noch Minga kumm-*st*
 whether you.SG to Munich come-2SG
 'whether you come to Munich'

³¹ Of course, one might argue that forms with 'double attachment' of agreement ending plus clitic such as **kumm-st-st* 'come-2sg-2sg.clitic' are excluded for phonological reasons. However, there is evidence from the distribution of complementizer agreement in comparatives that clearly shows that Bavarian lacks 2nd person clitics. In comparatives, complementizer agreement is only possible if the finite verb is overtly realized (cf. chapter 2 above). As shown in (ib) and (ic), deletion of the verb renders complementizer agreement ungrammatical (Bayer 1984: 269):

(i) a. D'Resl is gresser [als wia-*st* du *bist*].
 the-Resl is taller than as-2SG you are
 'Resl is taller than you are.'
 b. *D'Resl is gresser [als wia-*st* du].
 the-Resl is taller than as-2SG you
 c. D'Resl is gresser [als wia du].
 the-Resl is taller than as you

Now, under the assumption that 'double attachment' is excluded for purely phonological reasons, we would actually expect the putative clitic 2sg *-st* to show up in cases like (ic) where complementizer agreement is ruled out. This prediction is not borne out by the facts:

(ii) *D'Resl is gresser [als wia=*st*].
 the-Resl is taller than as='CLIT.2SG'

We can therefore conclude that there are no second person clitics in Bavarian, and that the paradigm of subject clitics exhibits a gap in these contexts.

b. *ob ees/ihr noch Minga kumm-ts
 whether you.PL to Munich come-2PL
 'whether you come to Munich'

(25) a. ob'e (*i) noch Minga kumm
 whether-CLIT.1SG I to Munich come.1SG
 b. ob i noch Minga kumm
 whether I to Munich come.1SG
 'whether I come to Munich'

An additional instance of complementizer agreement can be observed in some Carinthian and Lower Bavarian varieties, where the 1pl subject enclitic *-ma* turned into a C-related inflection (cf. Pfalz 1918, Schirmunski 1962: 525, Bayer 1984, Altmann 1984, Kollmer 1987, Wiesinger 1989, Abraham 1995, Weiß 1998, 2002, Fuß 2005). As a result of that change, the formative 1pl *-ma* exhibits a similar behavior as the 2nd person inflections: it is obligatory in all contexts and it can be doubled by full pronoun, as illustrated in (26) and (27). Moreover, (28) shows that in a subset of the relevant varieties, the new agreement formative has fully replaced the older ending /-an/ (albeit this extension is restricted to auxiliaries and a couple of modals):

(26) a. wem-**ma** aaf Minga fon
 when-1PL to Munich drive
 b. wem-**ma** mia aaf Minga fon
 when-1PL we to Munich drive
 c. *wem mia aaf Minga fon
 when we to Munich drive
 'when we drive to Munich'
 (Weiß 2002:9)

(27) a. Mia fom-**ma** hoam.

we drive-1PL home

‘We go home.’

(Weiß 2002:9)

b. *Mia fon hoam

we drive home

‘We go home.’

(Helmut Weiß, p.c.)

(28) a. dass-**ma** (mia) koā geid ned hā-**ma** [instead of 1pl *hā-n*]

that-1PL we no money not have-1PL

‘that we have no money’

(Kollmer 1987: I, 362)

b. we-**ma** (mia) des ned dou-**ma**... [instead of 1pl *dou-n*]

if-1PL we that not do-1PL

‘if we don’t do that...’

(Kollmer 1987: I, 358)

Accordingly, sentences like (29) must presumably be analyzed as involving a null subject, similar to related examples with 2nd person forms (cf. Bayer 1984: 252):

(29) Fahr-**ma** Ø noch Minga?

drive-1PL to Munich

‘Will (we) go to Munich?’

The fact that pro-drop is limited to the same contexts where inflected complementizers appear has led some researchers to propose that the overt manifestation of agreement in C serves to license referential null arguments in the subject position (cf. e.g. Bayer 1984, Weiß 2002, 2006, Axel and Weiß 2007). While this proposal raises a number of questions from a purely synchronic point of view,³² it is fairly clear that the two phenomena are correlated historically.

³² For example, it is unclear why complementizer agreement and pro-drop are limited to 2nd person contexts in Bavarian. From a purely synchronic point of view, this restriction appears to be completely arbitrary, as it does not relate to any other properties of the grammar (e.g., special properties of 2nd person inflection). Furthermore, we may ask why pro-drop is also licensed in inversion contexts, where complementizer agreement is not visible (cf. (21)). Note that in these

As already noted in section 2.1, Bavarian (and other Germanic varieties) was subject to a historical development in which new verbal agreement suffixes evolved via a reanalysis of subject enclitics (Brinkmann 1931, Sommer 1994, Braune and Reiffenstein 2004: 261; cf. Bayer 1984, Weiß 2002, Fuß 2005 for details of the changes affecting Bavarian). In Fuß (2004, 2005), it is argued that the transition of pronouns to agreement markers forced the learner to assume the presence of a referential null subject (*pro*) receiving the thematic role of the external argument, which had formerly been assigned to the clitic pronoun.³³ By assumption, this gave rise to the limited pro-drop properties of the present-day language (cf. Weiß 2002 for a related proposal):

(30) $[\text{CP} \text{XP} [\text{C}' \text{V}_{\text{fin}} [\text{IP} \text{clitic}_{\text{subj}} \dots]]] \rightarrow [\text{CP} \text{XP} [\text{C}' \text{V}_{\text{fin}} + \text{AGR} [\text{IP} \text{pro} \dots]]]$

- a. 2sg: /-s/ + /t/ (<<< clit. 2sg *t(hu)*)
- b. 2pl: /-t/ + /s/ (<<< clit. 2pl *(ee)s*)
- c. 1pl: /an/ \rightarrow /ma/ (<<< clit. 1pl *ma*)

In section 2.1, we have already developed an explanation of why the reanalysis did not affect all existing agreement endings, giving rise to full-fledged pro-drop, but rather was confined to 2nd person forms (plus 1pl in some varieties): due to the workings of the *Blocking Principle*, new agreement markers (and therefore null subjects) could only develop for those slots of the agreement paradigm where the new agreement formatives were more distinctive than the existing markers. Thus, the distribution of null-subjects in the present-day language can be explained in terms of morphological conditions on the reanalysis in (30). Thus, it seems that the replacement of clitic pronouns by null subjects is not directly related to properties of

contexts, the restriction to 2nd person cannot be attributed to some special morphological property of the 2nd person verbal agreement suffixes, in the sense that 2nd person forms are more distinctive than 1sg or 3sg. Finally, note that in other varieties such as West Flemish and dialects spoken in the east and south of the Netherlands, the presence of complementizer agreement does not license pro-drop (cf. e.g. Zwart 1993b).

³³ The evidence available to us suggests that the change in question proceeded as follows (cf. Fuß 2005, ch. 5 and 6 for details):

- (i) a. V + enclitic (inversion contexts) \rightarrow V+Agr + *pro*
- b. Bavarian: extension to other C-related elements such as complementizers, relative pronouns etc.
- c. Extension of the new ending to verbs in clause-final positions

the agreement paradigm as a whole, but rather depends on the reanalysis of individual clitics as verbal agreement morphology.

Still, we may ask why the absence of an alternative overt carrier of the relevant thematic role (e.g., a left-dislocated subject double) did not simply block the reanalysis of the pronominal clitics (instead of giving rise to the marked parametric option of partial pro-drop). This question is dealt with in the next section.

4.1.1 Deblocking the zero spell-out of weak pronouns

In this section I take a closer look at the circumstances under which the transition from pronominal clitics to inflectional markers may give rise to null subjects. In particular, I want to explore the question of how agreement-related null arguments can develop in a language such as Bavarian that otherwise lacks the preconditions (i.e., rich verbal inflection) for full-fledged 'Italian-style' referential pro-drop.

The proposal I want to put forward is based on the idea that the type of partial pro-drop found in Bavarian is systematically linked to gaps in the paradigm of weak (or clitic) pronominal forms. Recall that the evidence available to us suggests that the reanalysis of clitic pronouns did not only lead to new agreement suffixes, but also produced gaps in the paradigm of weak/clitic pronouns (cf. e.g. Altmann 1984: 200):

	Verbal agreement	Subject clitics
1sg	-Ø	e
2sg	-st	—
3sg	-t	a/s
1pl	-an(d) -ma (in some varieties)	-ma —
2pl	-ts	—
3pl	-an(d)	s

Table 10: Agr suffixes (pres.indic.) and subject clitics in present-day Bavarian

A brief look at Table 10 suggests that pro-drop becomes available in those contexts where the clitic paradigm exhibits a gap. So the relevant empirical generalization seems to be that null subjects may be licensed in case there is no visible weak form stored in the lexicon. Now, of course this raises the question of how we can formally explain the apparent correlation between the availability of empty subjects and apparent gaps in the paradigm of weak forms. The basic idea I want to pursue is that the availability of null subjects is systematically linked to properties of the paradigm of overt (weak) pronouns (see also Neeleman and Szendrői 2007). More precisely, I assume that a null spell-out of weak pronouns becomes available if there is no

competing overt candidate available that realizes a greater subset of the underlying morphosyntactic features than the null spell-out. In other words, partial pro-drop is analyzed as an instance of deblocking, where a less specified phonological realization, namely zero, becomes available in the absence of a more specified competing form. Let's now take a closer look at the specifics of this proposal.

Following Holmberg (2005), Roberts (2007b), and Neeleman and Szendrői (2007), I assume that the phenomenon of pro-drop does not involve a special empty category like *pro*. Adopting a Late Insertion approach where syntactic nodes are associated with phonological features post-syntactically (cf. e.g. Halle and Marantz 1993), null arguments are analyzed as regular pronominal forms that fail to be associated with a phonological matrix at the point of Vocabulary Insertion. More precisely, I assume that syntactically, null pronouns are a particular variant of weak forms (cf. Cardinaletti and Starke 1999). Adopting a bare phrase structure approach (Chomsky 1995),³⁴ let's assume that the syntactic structure corresponding to a weak pronoun is a category $D^{\min/\max}$ (D^0 in traditional X-bar notation) that is both minimal (since it is non-complex) and maximal (since it is merged in a thematic specifier position and does not project) at the same time (Chomsky 1995: 249, Roberts 2007b). In contrast, full tonic pronouns are DPs (cf. e.g. Chomsky 1995: 249). A pronominal $D^{\min/\max}$ is characterized by the binary features $[\pm\text{pronominal}]$, $[\pm\text{anaphoric}]$ (cf. Chomsky 1982), a definiteness feature $([+\text{definite}])$, and a set of ϕ -features (at least person, number, and case). A relevant pronunciation rule (or, Vocabulary item) that gives rise to a null weak subject pronoun would look like (31) (see also Neeleman and Szendrői 2007: 682):³⁵

³⁴ See e.g. Uriagereka (1995), Cardinaletti and Starke (1999), Déchaine and Wiltschko (2002), Neeleman and Szendrői (2007), and Holmberg (2005) for more elaborate theories of the internal structure of pronominal elements.

³⁵ As pointed out by Denis Delfitto to me, it is somewhat unexpected that a supposedly universal spell-out rule such as (31) makes reference to a language-specific feature such as $[\pm\text{NOM}]$. This shortcoming could perhaps be repaired by making use of more basic features, adopting an analysis in which traditional case features are decomposed into a set of semantic primitives (Jakobson 1936 [1971], Bierwisch 1967). Alternatively, we may assume that the relevant case specification is in fact $[+\text{DEFAULT CASE}]$ (Denis Delfitto, p.c.), giving rise to null subjects in languages where the relevant default case is nominative, as for example in Bavarian. Furthermore, in languages that lack the category of case altogether (such as Chinese), we would then perhaps expect that (31) becomes available for all kinds of arguments, giving rise to 'radical' pro-drop (Neeleman and Szendrői 2007). At first sight, this seems to be borne out by the facts, but obviously, more research is necessary to

$$(31) [_{D_{\min/\max}} +\text{pron.}, -\text{anaph.}, +\text{NOM}] \leftrightarrow \emptyset$$

(31) states that a $D^{\min/\max}$ with a feature combination that is characteristic of a pronominal subject can be realized as zero when Vocabulary Insertion applies. I follow Neeleman and Szendrői in assuming that (31) is universally available as the unmarked realization of weak/clitic pronominal forms. In other words, the setting [+pro-drop] is conceived of as the default parameter option.³⁶ Furthermore, I assume that (31) is restricted by an identification requirement, that is, a null spell-out is only possible if the morphosyntactic content of the empty pronoun can be recovered. Relevant pieces of information that serve to identify the missing argument may come from (i) the presence of rich agreement morphology, (ii) the immediate discourse context, or, as will become clear shortly, (iii) gaps in the paradigm of weak pronominal forms.

Of course, (31) must be complemented by insertion rules that determine the realization of overt pronouns. Note that the Vocabulary items that are associated with individual overt forms are usually much more specific than the very general rule (31), compare the following insertion rule that realizes the 3sg.masc subject clitic *a* in Bavarian:

$$(32) [_{D_{\min/\max}} +\text{pron.}, -\text{anaph.}, +\text{NOM}, -\text{speaker}, -\text{hearer}, -\text{pl}, +\text{masc}] \leftrightarrow /a/$$

Under the assumption that the insertion of phonological material is governed by conditions that favor more specified over less specified Vocabulary items (cf. the Elsewhere Condition of Kiparsky 1973, 1982, or Halle's 1997 *Subset Principle*), the availability of 'overt' forms such as (32) – all things being equal – blocks the null spell-out of weak pronouns (i.e., the relevant $D^{\min/\max}$), since the overt forms clearly realize more morphosyntactic features than the Vocabulary item in (31).³⁷

substantiate this claim. For expository reasons, however, I will stick to the case specification [$\pm\text{NOM}$] in what follows.

³⁶ This assumption is supported by the observation that cross-linguistically – even in non-null subject languages like English – children produce null subjects at an early stage of the acquisition process (cf. Hyams 1986 and the following quote taken from O'Grady 1997: 83, "[...] subject drop seems to be a universal feature of syntactic development [...]").

³⁷ See Neeleman and Szendrői (2007) for an analysis of 'radical' pro-drop in languages such as Chinese based on the assumption that both insertion rules may optionally apply in case they satisfy different

Note that the availability of overt tonic pronouns does not prevent the application of (31), since they correspond to a different syntactic structure (by assumption, DP) and therefore do not compete with the null form for realizing pronominal $D^{\min/\max}$ (note that this implies that Vocabulary Insertion may target not only terminal nodes, but also larger pieces of phrase structure, in the case at hand a whole DP node, cf. Weerman and Evers-Vermeul 2002, Neeleman and Szendrői 2007). That is, the availability of the full pronoun 2sg /du:/ does not interfere with the null realization of a pronominal 2sg $D^{\min/\max}$ (as a result of (31)) in Bavarian, since the relevant insertion rules target different nodes in the syntactic structure:

(33) $[\substack{+pron., -anaph., +NOM, -speaker, +hearer, -pl}] \leftrightarrow /du:/$

Under these assumptions, the development of partial referential pro-drop in the history of Bavarian can be accounted for in terms of deblocking: At the point where the continuing phonological erosion of subject clitics made available a reanalysis of these forms as inflectional formatives, the clitics affected by this process dropped out of the grammar, giving rise to gaps in the paradigm of weak pronominal forms. The disappearance of clitic forms caused the emergence of a previously blocked option, namely the null spell-out of pronominal $D^{\min/\max}$ due to the application of the (universally available) insertion rule (31). I assume that the content of the phonologically empty pronoun can be recovered via the particular agreement morphology associated with C in Bavarian, which unambiguously signals person and number of the subject (see Weiß 2002 for a related proposal).³⁸

This approach makes an interesting prediction: the development of new clitic forms that fill the relevant gaps in the paradigm as new phonological realizations of pronominal $D^{\min/\max}$ is expected to lead to the loss of (partial) pro-drop in the relevant

parts of the Elsewhere Principle ('realize more features' vs. 'realize bigger chunks of structure'). However, note that this analysis is based on the assumption that null forms correspond to phrasal categories (DP), contrasting with the view adopted here that clitic pronouns are significantly smaller structural units (i.e., D^0 's).

³⁸ Note that the presence of complementizer agreement does not automatically give rise to referential pro-drop. A case in point is West Flemish, which has complementizer agreement in all persons and numbers (1sg, 1pl, 3pl: /-n/, 2sg, 2pl, 3sg: /-t/), but fails to exhibit pro-drop (cf. Haegeman 1992). Under the above assumptions, the absence of referential null subjects can be attributed to the fact that West Flemish has a complete series of clitic subject pronouns, which blocks a null realization of pronominal $D^{\min/\max}$ (moreover, note that the inflection associated with C is highly syncretic and therefore fails to unambiguously identify a null subject).

contexts. Some evidence that this prediction is on the right track comes from recent developments that affected the grammar of Colloquial Finnish (cf. Vainikka and Levy 1999).

4.1.2 Blocking the zero spell-out of weak pronouns: Colloquial Finnish

It has variously been noted that Finnish raises a problem for the assumption that there is a correlation between rich morphology and the availability of full-fledged pro-drop (cf. e.g. Vainikka and Levy 1999, Holmberg 2005, Koeneman 2007). Despite the fact that Standard Finnish exhibits a fully distinctive verbal agreement paradigm (similar to Italian), null subjects are limited to first and second person (examples taken from Holmberg 2005: 539):³⁹

(34) a. (Minä) puhun englantia.
 I speak-1SG English
 b. (Sinä) puhut englantia.
 you speak-2SG English
 c. *(Hän) puhuu englantia.
 he/she speak-3SG English

³⁹ Vainikka and Levy (1999: 661f.) argue that the licensing of null subjects in Standard Finnish is linked to similarities between verbal agreement morphology and the inventory of subject pronouns: Historically, the 1st and 2nd person verbal agreement markers developed from pronouns. This relationship is particularly clear in the case of 1pl and 2pl. In the singular, the link is less transparent, but can be easily reconstructed historically (in the case of 2sg, the original pronoun was *tinä*, which later changed into *sinä* due to a general phonological rule /ti/ >>> /si/ which is still at work in present-day Finnish). The 1sg suffix /-n/ developed from former /-m/). No such relation can be constructed for the 3rd person endings, which developed from an active present participle suffix. Vainikka and Levy (1999) suggest that the systematic differences between 1st and 2nd person agreement formatives on the one hand and 3rd person forms on the other can be detected by the learner. More precisely, it is assumed that the morphological similarities between 1st and 2nd person agreement markers and pronouns signal to the learner that the relevant agreement endings are [+pronominal], giving rise to (partial) pro-drop in these contexts. While this proposal seems to account for the Finnish data, it raises a number of questions (see also Koeneman 2007). In particular, it is not clear whether learners of Finnish are really capable of identifying the singular endings as [+pronominal]. While the diachronic relation between *minä* and *-n* and *sinä* and *-t* is fairly clear from a (historical) linguist's viewpoint, it is not very likely that all speakers of contemporary Finnish have access to that piece of information.

- d. (Me) puhumme englantia.
we speak-1PL English
- e. (Te) puhutte englantia.
you speak-2PL English
- f. *(He) puhuvat englantia.
they speak-3PL English

As noted by Vainikka and Levy (1999), Colloquial Finnish differs from the standard variety in that it requires the presence of overt pronouns (i.e., pro-drop has been completely lost). Interestingly, this change is accompanied by a set of further changes that affected the shape of pronouns (and the inventory of agreement markers):

	Pronouns	Agreement
1sg	minä → mä	-n
2sg	sinä → sä	-t
3sg	hän → se	-V
1pl	me	-tAAn
2pl	te	-tte
3pl	he → ne	-V

Table 11: Pronouns and subject agreement in Colloquial Finnish⁴⁰

As can be gathered from Table 11, Colloquial Finnish has developed new reduced forms for 1sg and 2sg (in addition, the 3rd person pronouns 3sg *hän* and 3pl *he* have been replaced by the relevant demonstrative forms, *se* and *ne*, respectively). The new shortened forms are generally unstressed (cf. e.g. Holmberg and Nikanne 2006: 5). Furthermore, the 1pl verbal agreement suffix is replaced by -tAAn, originally an impersonal passive affix, and the 3rd person endings have fallen together. Vainikka and Levy suggest that these changes disrupted the systematic similarities between 1st and 2nd person pronouns and agreement endings. As a consequence, the latter lose their argumental status, leading to the loss of (partial) pro-drop in Colloquial Finnish (see Koeneman 2007 for an alternative analysis that attributes the loss of pro-drop to the loss of a fully distinctive agreement paradigm).

However, it seems that the facts observed in Colloquial Finnish can possibly also be subsumed under the account of (partial) pro-drop developed above. More specifically, at least in the case of 1sg and 2sg, the loss of pro-drop can be directly related to the development of new weak forms that are more specified than a null

⁴⁰ “-V” represents an empty vowel that is similar to the preceding vowel and results in vowel lengthening. Capital “A” represents a vowel undergoing vowel harmony.

spell-out of pronominal $D^{\min/\max}$ and therefore block the latter. Of course, further research is necessary to substantiate this conjecture, in particular concerning the status of the 1st and 2nd person plural forms, which at first sight seem to be identical to the relevant forms in the standard language.

Interestingly, there are some observations concerning the system of pronouns in spoken Finnish that suggest that the inventory of forms is in fact larger than it appears at first sight. More to the point, it appears that in spoken Finnish, there are differences in vowel length that discriminate between stressed and unstressed forms (cf. e.g. the description of the vernacular spoken in Jyväskylä on <http://www.cc.jyu.fi/~tojan/rlang/finn2.htm>). Anne Vainikka (p.c.) informed me that in her dialect (the variety spoken in Tampere), there are three kinds of pronouns: (i) an unstressed variant with a short vowel (the default case), (ii) a stressed form with a long vowel, and (iii) an unstressed form with a long vowel. Thus, it seems that in the Tampere variety, weak pronominal forms differ systematically in vowel length from tonic pronouns and can therefore be analyzed as an overt spell-out of weak pronominal $D^{\min/\max}$. We thus deal with at least two series of pronouns: strong forms that may bear stress and weak forms that are characterized by a short vowel. Under the assumptions put forward in this section, the latter can be taken to block the competing null realization of $D^{\min/\max}$. These observations concerning the emergence of a separate series of clearly identifiable weak pronominal forms in spoken Finnish are suggestive and can be taken to indicate that the particular approach to null subjects developed here might be on the right track.

In this section, I have proposed that in the history of Bavarian, partial pro-drop developed as a side-effect of the reanalysis of clitic forms that turned into agreement markers. In particular, we have seen that this change led to gaps in the paradigm of clitic pronouns, which made available a null realization of pronominal $D^{\min/\max}$ in exactly those contexts where the reanalysis took place. In more formal terms, the emergence of null subjects has been analyzed in terms of deblocking of an underspecified Vocabulary item (i.e., $/Ø/$), the insertion of which was formerly blocked by the availability of more specified overt candidates.⁴¹ This analysis is

⁴¹ Of course, this section left many important questions unaddressed. For example, a couple of open questions remain concerning the analysis of partial pro-drop in Standard Finnish. While one might argue that Standard Finnish lacks a separate series of weak pronouns, leading to deblocking of the null spell-out, this still would raise the question of why pro-drop is limited to first and second person. In addition, more has to be said about possible implications for the analysis of agreement-related pro-drop in languages like Italian, or the question of how we can account for the general non-

supported by the observation that spoken Colloquial Finnish is characterized by a loss of null subjects which can be traced back to the development of a new series of overt weak pronouns blocking a null-spell out of pronominal $D^{\min/\max}$ (still available in the standard language).

What is the significance of these findings for the theory of pro-drop? First of all, the changes affecting Bavarian suggest that pro-drop does not necessarily develop in an across-the-board fashion for all persons and numbers (when the richness of verbal inflection crosses a certain threshold), contrary to what is expected under standard assumptions. Instead, it appears that the development of null subjects is confined to contexts where the paradigm of clitic forms exhibits gaps due to the reanalysis of pronominal elements as agreement markers. More precisely, it appears that the emergence of (partial) pro-drop involves an intricate interaction between morphological blocking and deblocking: First, the development of null subjects is sensitive to properties of the series of weak/clitic pronouns (i.e., deblocking of the null spell-out via gaps in the paradigm), and second, it is indirectly related to properties of verbal agreement, namely via morphological mechanisms that promote a reanalysis of subject clitics if this results in more distinctive agreement markers (due to blocking effects that favor more specified Vocabulary items).

Note that these findings do not falsify agreement-related theories of pro-drop entirely. They merely suggest that these approaches cannot be maintained in a strong form (e.g., positing that referential pro-drop is available *only* in the presence of a fully distinctive agreement paradigm). Rather, it seems that languages may develop restricted pro-drop properties linked to contexts where pronominal elements turn into agreement markers. In other words, an empirically more adequate generalization would be that full referential pro-drop in all persons and numbers requires rich verbal agreement, while partial pro-drop does not. Accordingly, it seems that the assumption that pro-drop operates in an all-or-nothing fashion must be abandoned. Still, we may wonder whether a typological change from [-pro-drop] to full fledged referential pro-drop can be the result of a wholesale reanalysis of pronouns affecting larger parts of the agreement paradigm (either in one fell swoop

availability of pro-drop in languages such as English. However, note that full pro-drop languages such as Italian or Spanish typically lack a series of clitic subject pronouns, which is expected under the approach advocated here. Furthermore, the behavior of non-null subject languages may possibly be accounted for under the assumption that languages like English lacks the syntactic category of weak pronominal $D^{\min/\max}$, that is, the output of the syntax does not contain a structure that can be realized by the insertion rule (31). I leave these matters for future research.

or one item after the other). A possible case in point are recent developments that have been affecting the grammar of Colloquial French.

4.2 The rise of full-fledged pro-drop: Colloquial French

It is a well-known fact that the grammar of Colloquial French exhibits a number of properties that sets it apart from the standard language. These differences also concern the realization of subject pronouns. Authors such as Roberge (1990), Friedemann (1997), or Fonseca-Greber (2000) (among others) argue that Colloquial French exhibits an ongoing transition from a grammar without null subjects to a +pro-drop grammar.⁴² Similar to Bavarian, this development involves a change in which subject clitics turn into (prefixal) agreement markers. This transition is manifested by a set of properties in which the subject 'clitics' of Colloquial French differ from those of the standard language (cf. Wartburg 1970, Ashby 1977, Harris 1978, Lambrecht 1981, Roberge 1990, Auger 1993, 1994a, Fonseca-Greber 2000, Fonseca-Greber and Waugh 2003, Gerlach 2002; however see de Cat 2005 for an opposing view).

First of all, the subject clitics are obligatory and cannot be replaced by full tonic pronouns (historically an oblique form).⁴³ Furthermore, sentences with apparent clitic doubling generally favor a basic, non-dislocated interpretation:

Colloquial French

(35) a. (Moi) je porte la table.
 me 1SG carry the table
 'I carry the table.'

⁴² Apparently, a similar development has taken place in a number of North Italian dialects, cf. Vanelli (1987), Renzi (1992). See also Rizzi (1986b), Brandi and Cordin (1989) and Poletto (1995) for discussion.

⁴³ Friedemann (1997) claims that doubling is merely optional in all non-standard varieties of French. However, Fonseca-Greber and Waugh (2003), examining a corpus of contemporary spoken French, observe that there are no cases where a tonic 1st or 2nd pronoun occurs without a clitic (i.e., doubling seems to obligatory). With 3rd person forms, doubling is slightly less frequent (3sg clitics are present in 91.5% of the relevant cases, 3pl forms in 93.6%). A similar finding is reached by Gerlach (2002).

b. Moi *(je) porte la table.
me 1SG carry the table
'I carry the table.'
(Gerlach 2002:224)

In conjoined clauses, subject clitics must be repeated before each finite verb (cf. Lambrecht 1981, Fonseca-Greber and Waugh 2003), while standard French exhibits the typical elision of pronominal forms under identity with the subject of the first conjunct clause:

Standard French

(36) Il mange et boit comme un cochon.
he eats and drinks like a pig

Colloquial French

(37) I mange et *(i) boit comme un cochon.
he eats and he drinks like a pig

The preverbal 'clitics' occur in a fixed position relative to the verb stem. For example, they fail to undergo subject-verb inversion in matrix interrogatives, in contrast to the subject clitics of the standard language. This is shown in (38) and (39) (Friedemann 1997: 3f.):

Standard French

(38) Où est-il parti?
where is=he gone
'Where did he go to?'

Colloquial French

(39) Où il-est parti?
where he-is gone
'Where did he go to?'

These properties are commonly taken to suggest that the 'clitics' are in fact better analyzed as instances of preverbal agreement markers. As a consequence, clauses without a subject double must be analyzed as instances of pro-drop.

However, there are some indications that the transition from pronouns to agreement markers is not yet fully completed. Several authors have noted that quantified expressions, indefinite NPs, and *wh*-phrases cannot be doubled by subject clitics (cf. e.g. Roberge 1990: 95, Friedemann 1997: 125):

Colloquial French

(40) *Personne il a parlé.
 nobody he has spoken
 'Nobody spoke.'

Colloquial French

(41) *Un ami il est toujours là.
 a friend he is always there
 'A friend is always there.'

Colloquial French

(42) *Qui il aime la tarte?
 who he likes the pie
 'Who likes the pie?'

However, examples similar to (40) and (41) are well-formed in other non-standard varieties of French such as Picard, or Pied-Noir, which suggests that in the latter, the grammaticalization of prefixal agreement is more advanced (cf. Roberge 1990, Friedemann 1997, Auger 1994b, 2003):

(43) Personne i(l) sait qui c'est leur mère.
 nobody he knows who that-is their mother
 'Nobody knows who is their mother.'
 (Pied-Noir, Friedemann 1997: 125)

(44) Un homme il vient.
 a man he comes
 (Pied-Noir, Roberge 1990: 97)

(45) Chacun il a sa chimère.
 everybody he has his spleen
 'Everybody has a spleen.'
 (Picard, Friedemann 1997: 125)

While doubling of *wh*-phrases is ruled in Pied-Noir (Roberge 1990: 120), a default 3sg.masc clitic is present in *wh*-questions in Picard; furthermore, subject-relatives exhibit resumptive subject clitics (Auger 1994b, 2003):

(46) tchèche qu' il a dit qu' i folloait nin finir?
 who that he has said that it had-to of-it to-finish
 'Who said we had to put an end to it?'
 (Picard, Auger 2003: 5)

(47) inne grosse féme éd Tours qu' al étoait rouge...
 a fat woman from Tours that she was red
 'a fat woman from Tours who was red...'
 (Picard, Auger 2003: 5)

Moreover, corpus studies carried out by Fonseca-Greber (2000) and Fonseca-Greber and Waugh (2003) show that doubling is extending to contexts with quantified NPs in 'normal' Colloquial French as well.

Summing up, it appears that different non-standard varieties exhibit different stages of a development in which clitics turn into prefixal agreement markers, eventually giving rise to a grammar with null subjects. To the extent that the reanalysis is completed, the evidence available to us suggests that the emergence of pro-drop in Colloquial French can possibly be treated on a par with the developments in Bavarian, that is, in terms of the deblocking of a null spell-out in those contexts where the reanalysis of clitic forms has given rise to gaps in the paradigm of weak/clitic forms (note that the resulting system of agreement marking is sufficiently distinctive to recover the content of the missing argument).

Again, the changes affecting the status of the subject clitics can be related to properties of the existing agreement paradigm (cf. Gerlach 2002, Fuß 2005). It appears that there are some differences between the individual subject clitics. Most importantly, it seems that not all subject clitics are obligatorily present:

1sg	obligatory
2sg	obligatory
3sg	almost obligatory (cf. fn. 43)
1pl	obligatory (<i>on</i>)
2pl	almost obligatory
3pl	almost obligatory (cf. fn. 43)

Table 12: Presence of subject clitics in Colloquial French

Doubling of full forms is compulsory in the contexts of 1sg, 2sg, and 1pl (where *on* has replaced *nous* in the spoken language), while it is slightly less frequent in the other contexts (see below for some remarks on the status of 2pl). Thus, the grammaticalization process lags behind for 3rd person forms and 2pl. Interestingly, the differences in the behavior of the 'clitic' forms correlates with properties of the existing suffixal agreement morphology. Apparently, at least in non-3rd person contexts, clitics are obligatory in case the verbal inflection is underspecified for subject agreement features (Gerlach 2002: 225f.):

	Written language	Phonetic form
1sg	porte	[pɔRT]
2sg	portes	[pɔRT]
3sg	porte	[pɔRT]
1pl	(on) porte	[pɔRT]
	(nous) portons	not used in Colloquial French
2pl	portez	[pɔR'te:]
3pl	portent	[pɔRT]

Table 13: Subject agreement in written/spoken French

Note that only the 2pl ending /-e:/ signals person and number of the subject; elsewhere we find the completely underspecified zero ending. This can be linked to the distribution of clitics in the following way (Gerlach 2002):

(48) *Verbal agreement and the distribution of subject clitics*

In Colloquial French, subject clitics are obligatory only

- (a) in non-third person contexts and
- (b) if they serve to express ϕ -features not marked by the existing suffixal agreement morphology.

Similar to Bavarian, the distribution of the obligatory agreement marking forms can be attributed to the workings of the Blocking Principle: The grammaticalization of new agreement markers (and the rise of the null subjects) is triggered only in contexts where the new inflections are clearly more specified than the existing elsewhere marker /Ø/:

(49) a. [+speaker, +hearer, -pl] ↔ /ʒə/ (1sg)
 b. [-speaker, +hearer, -pl] ↔ /tY/ (2sg)
 c. [+speaker, -hearer, +PL] ↔ /ʒ/ (1pl)

In those varieties where the 2pl clitic is still merely optional, its different behavior can be attributed to the fact that the existing agreement morphology is still distinctive, which hinders a reanalysis of the subject clitic. However, note that based on a study of a corpus of spoken French, Fonseca-Greber and Waugh (2003) claim that the 2pl subject clitic *vous* has also developed into a fully morphologized agreement marker. This can possibly be explained as the result of analogical extension on the model of the other former clitics (possibly promoted by factors such as paradigm uniformity, that is, a general preference for a uniformly prefixing or suffixing set of agreement markers).

A more serious question concerns the status of 3rd person forms (3sg.fem *elle*, 3sg.masc *il*, 3pl.fem *elles*, 3pl.masc *ils*). At first sight, they should qualify for a reanalysis as agreement markers from the viewpoint of the Blocking Principle. Obviously, they are specified for person, number and gender, so they should meet the condition that they be more distinctive than the existing zero marker. However, upon closer inspection, the 3rd person forms turn out to be less specified than they appear to be.

Let's begin by taking a second look at the putative person specification of the relevant pronouns. It has repeatedly been pointed out in the literature that '3rd person' should actually be treated as the 'non-person', that is, 3rd person forms are better analyzed as being underspecified for person features (cf. Benveniste 1950, 1966, Halle 1997, Noyer 1997, Harley and Ritter 2002, Cysouw 2003, among many others; however, see Trommer 2006 for an opposing view). If this view turns out to be correct, the 3rd person forms fail to be more specified than the existing zero marker with respect to the category of person. Accordingly, at least in this respect,

they do not qualify as more distinctive forms that may outrank the existing markers due to blocking effects.⁴⁴

What about the number specification? On the face of it, the apparent contrast between 3sg *il/elle* and 3pl *ils/elles* should suffice to mark the clitics as more specified than the existing zero marker. Note, however, that the number marking of the plural forms is only perceivable if the verb following the clitic begins with a vowel. Hence, the number marking of the 3rd person forms is actually less salient than it appears at first sight. In some varieties, the visibility of number marking is further weakened by the tendency to use a reduced form *i(l)* for all 3rd person contexts (sometimes accompanied by 3pl *eux*, cf. Ashby 1977, Fonseca-Greber and Waugh 2003: 102):⁴⁵

(50) Mes petites cousins eux i-savaient...
 my little cousins.FEM 3PL.MASC 3-knew
 'My little cousins knew...'
 (Fonseca-Greber and Waugh 2003: 102)

So it appears that the set of contexts where the number marking on the 3rd person pronouns is really visible is actually quite small, presumably too small to count as robust evidence for the purposes of the Blocking Principle.

Finally, let's turn to gender. Interestingly, we can observe that there is a tendency in Colloquial French to use *i(l)* as a general marker of 3rd person that can also be used in 3sg.fem contexts. This is illustrated by the following examples taken from Wartburg (1970: 74) and Ashby (1977: 68), respectively. This can be taken to blur the gender distinctions originally signaled by the subject clitics.⁴⁶

⁴⁴ Note that cross-linguistically, 3rd person verbal agreement is much less common than 1st and 2nd person agreement. See Fuß (2005) for an explanation of this fact in terms of the Blocking Principle, making use of the assumption that 3rd person forms are inherently underspecified for [person].

⁴⁵ Similarly, the forms for 3sg and 3pl clitics have merged in Picard and Pied-Noir French (e.g. Pied-Noir 3sg.masc.sg, 3sg.masc.pl. /i/, 3sg.fem.sg, 3sg.fem.pl /ɛl/), cf. Roberge (1990: 191) on Pied Noir and Auger (2003: 5) on Picard.

⁴⁶ Cf. Fuß (2005: 255f.) for an alternative explanation (making use of a feature geometry) which is based on the assumption that the grammaticalization of gender agreement requires the presence of number marking for all persons.

(51) a. Ma femme il est venu.
 my wife he is come
 'My wife came.'

 b. Ma soeur i'chante.
 my sister 3-sing
 'My sister is singing.'

Thus, we may conclude that the 3rd person clitics are actually less distinctive than it appears at first sight. We might suspect that this obstructs the reanalysis of 3rd person forms as agreement markers (due to their reduced visibility to the workings of the Blocking Principle). Furthermore, the fact that the 3rd person forms have not yet fully grammaticalized into agreement markers in Colloquial French is presumably also the reason why doubling of quantified expressions (which are usually 3rd person NPs) is still ruled out – in contrast to other non-standard varieties such as Pied-Noir or Picard.

Summing up, it appears that Colloquial French (and other non-standard varieties of French) exhibits an ongoing change in which preverbal subject clitics turn into prefixal agreement markers, giving rise to pro-drop properties formerly absent in the grammar. Again, we witness a development in which pro-drop does not evolve at once for all persons and numbers. Rather, the rise of null subjects is intimately related to the reanalysis of individual subject clitics as agreement markers, which can be analyzed in terms of deblocking of the null spell-out due to the ongoing erosion and eventual loss of clitic forms. Similar to Bavarian, the transition of clitics into agreement markers is linked to properties of the existing agreement paradigm (new markers are more specified). When the change is eventually completed for all persons and numbers, this may give rise to full agreement-related referential pro-drop in (future) Colloquial French.

5 Summary

In this chapter I have taken a brief look at another set of examples where the specific course taken by language change can be attributed to the workings of acquisition strategies that enable the learner to determine structural properties of the grammar in case the input is ambiguous or defective. In particular, I have argued that at least in the domain of inflectional morphology, the cyclic nature of language change can be attributed to the workings of two apparently conflicting acquisition strategies that

help the learner to identify phonological exponents (of inflectional categories) and their feature specifications on the basis of the linguistic input he / she receives.

Section 2 has been shown that the course of grammaticalization processes giving rise to new and more distinctive inflectional formatives is shaped by a learning strategy (dubbed the *Blocking Principle*) that selects the most specified variant in case the input contains more than a single potential realization of a given inflectional category. This has been illustrated with the historical development of verbal agreement marking in Bavarian, where new formatives developed only for those slots of the paradigm where the existing markers were less distinctive than the newly coined ones. Due to the fact that less distinctive variants are dismissed in the course of language acquisition, blocking-induced change typically leads to a grammar that generates less linguistic variation than the target grammar.

Section 3 has been concerned with historical developments that are traditionally subsumed under the notion of analogical change. I have suggested that we can gain a deeper understanding of the relevant diachronic phenomena if we assume that the acquisition of phonological exponents and their feature specifications is influenced by a strategy that aims at minimizing the number of elements/features stored in the lexicon (*Minimize Feature Content*, Halle 1997). If true, this predicts that learners acquire the most economical lexical inventory compatible with the input they are exposed to. The tendency to posit an economical system of featural distinctions may lead learners to innovate 'less costly' variants that are not part of the input. To illustrate the workings of *Minimize Feature Content*, I have taken a closer look at a prototypical instance of analogical change, the development of so-called *Einheitsplural* in Alemannic. After a subanalysis of the relevant inflectional markers, it has become clear that each of the individual historical stages of the extension of 3pl /nt/ to other plural contexts complied with the notion of *Minimize Feature Content*. Furthermore, I have argued that the relevant changes were not motivated by their outcome, that is, the fact that they led to a highly economical inventory of lexical entries, in which each phonological exponent is uniquely paired with a single syntactico-semantic feature (in contrast to ideas widely held in the functionalist literature on language change). Rather, it has become clear that this particular result was merely a side-effect of the workings of *Minimize Feature Content*. I have then discussed the interaction of the blocking-induced change and the workings of *Minimize Feature Content*. Upon closer inspection, it appears that it is in fact possible to reconcile the apparently conflicting strategies as separate devices employed by the learner to select an optimal inventory of Vocabulary items based on the evidence

available to him/her. In particular, I have argued (i) that the Blocking Principle is invoked only in cases where there is more than a single candidate robustly attested in the input, and (ii) that it pays only attention to feature specifications of individual Vocabulary items. In contrast, 'analogical' changes typically coin new variants of less frequent forms, which may replace the older, possibly more distinctive forms if the latter fail to be robustly attested in the input the learner receives. Moreover, I have suggested that the scope of *Minimize Feature Content* is wider than the scope of blocking, in the sense that *Minimize Feature Content* not only considers the feature specifications of individual exponents, but also the featural make-up of whole paradigms.

In section 4, I have taken a brief look at the workings of blocking effects in another major cyclic change, namely the rise and fall of null subjects, discussing a small selection of different historical paths along which languages can develop and lose (partial) pro-drop. Drawing on data from Bavarian and Non-Standard French, it has been shown that referential null subjects may develop as a side-effect of the transition from pronouns to agreement markers. I have argued that the specifics of this change can be analyzed in terms of an intricate interplay between blocking and deblocking phenomena in morphology. Under this approach, the reanalysis of clitic pronouns as agreement affixes is taken to be governed by blocking effects which require new agreement markers to be more specified than existing inflectional markers (which accounts for the restrictions on the contexts where this change applies). The emergence of (partial) pro-drop can then be analyzed as an instance of deblocking, where a (by assumption universally available) null realization of weak pronouns becomes available due the loss of a more specific spell-out (i.e., the former clitic forms). The relevant observations suggest that agreement-related null subjects do not develop in an across-the-board fashion, but are initially restricted to those contexts where pronouns turn into agreement markers, contradicting the relevant diachronic predictions of standard generative theories of (agreement-related) pro-drop. When the development of new inflections is completed for all persons and numbers, which is presumably the case in certain non-standard varieties of French, this may give the impression that full referential pro-drop is linked to rich verbal agreement. In addition, we have seen that the loss of null subjects may be linked to the development of new overt weak pronominal forms. A relevant case in point seems to be the Colloquial Finnish, where the loss of partial pro-drop is accompanied by the development of new series of weak subject pronouns. I have argued that the apparent correlation between changes affecting the availability of null subjects and

changes affecting the paradigm of weak pronouns can be analyzed as an instance of blocking, where the availability of a more specified phonological exponent (the new overt weak pronouns) blocks the less distinctive null realization of weak pronominal $D^{\min/\max}$.

Appendix: Alternative paths toward null arguments

In the theoretical literature on Creole languages, it is occasionally claimed that Creoles (similar to Pidgins) generally lack null pronouns (cf. e.g. Muysken 1981, Roberts 1999, among others). However, there is actually quite some work on Creoles that directly contradicts this claim. To mention only a few, Kouwenberg (1990) and Kouwenberg and Muysken (1995: 215f.) show that Papiamento exhibits at least non-referential empty pronouns (in impersonal constructions and with weather verbs), DeGraff (1993) argues for the existence of pro-drop in Haitian Creole, and Veenstra (1994) claims that Saramaccan has developed agreement-related referential pro-drop due to the reanalysis of pronominal subjects, which is reminiscent of the changes discussed in section 4 above. The following discussion draws heavily on Lipski (2001), who provides a detailed overview of the evolution of null arguments in Romance-based Creoles. Taking a brief look at the development of discourse-oriented pro-drop in Mauritian Creole and Philippine Creole Spanish, I argue that in these Creoles, null arguments evolved on the model of substrate influence from (and/or intense contact with) Austronesian languages.

Mauritian Creole

Mauritian Creole (MC) is a French-based Creole that developed after slaves from different parts of Africa and Madagascar were brought to Mauritius roughly between 1715 and 1810 (when the slave trade was abolished). Present-day MC exhibits a variety of pro-drop phenomena (Syea 1993, Adone 1994a, 1994b). First, it exhibits null subjects in impersonal constructions, where the missing argument corresponds to an expletive (or quasi-argument), or an indefinite (generic) pronoun:

(52) Ti fer fre yer.
 TNS make cold yesterday
 '[It] was cold yesterday.'
 (Adone 1994a: 114)

(53) Lôtâ, ti degrad karo kan ar pios.
 long ago TNS cleared canefields with a pickaxe.
 'Long ago, [people] cleared cane fields with a pickaxe.'
 (Baker and Corne 1982: 89f.)

In addition, MC allows referential null subjects under certain conditions. Apparently, a referential (or, definite) interpretation of the null element is only possible if the identity of the missing argument can be readily recovered from the immediate discourse context, typically in answers to questions, as in (54):⁴⁷

(54) Question: ki Pyer pe fer?
 what Peter ASP do
 'What is Peter doing?'
 Answer: pe petir labutik.
 ASP paint shop
 '(He) is painting the shop.'
 (Syea 1993: 93)

However, according to Adone (1994a, 1994b), null subjects may also occur in contexts other than answers, as long as the missing argument can be identified with a prominent discourse topic, mostly the speaker (see also Syea 1993: 93). But, as shown by (57), 3rd person subjects may also be left out. That is, there is no principled contrast between different persons, in contrast to what has been observed for Bavarian and Non-Standard French above.⁴⁸

(55) Pu return dañ peis bieñto.
 MOD return in country soon
 '[I] will go back to the country soon.'
 (Adone 1994b: 33)

⁴⁷ Adone (1994a) notes that MC also exhibits referential null objects, albeit to a lesser extent.

⁴⁸ Note that null subjects may only refer to humans in MC (Syea 1993: 93).

(56) Pu repar sa sime la dimeñ.
 MOD repair DET road DET tomorrow
 '[We] will repair this road tomorrow.'
 (Adone 1994a: 114)

(57) Ti boykot en paket kreol dañ travay.
 TNS boycott QUA many Creole in work
 '[He] boycotted many Creoles in his work.'
 (Adone 1994b: 33)

If there is no appropriate antecedent available in the discourse context, the missing argument is interpreted as a specific indefinite (such as English *someone*), cf. Syea (1993: 93):

(58) fin koke Pyer so loto.
 ASP steal Peter his car
 '(Someone) stole Peter's car.'
 (Syea 1993: 92)

There is general agreement that the licensing and interpretation of null subjects is dependent on the presence of preverbal Tense/Mood/Aspect (TMA) particles, which presumably realize an inflectional head (*pu*, *fin*, and *ti* in the above examples).⁴⁹ In the absence of an overt TMA particle, a missing subject can only be interpreted as a generic indefinite pronoun (such as English *one* or the generic use of *people*, see also (53) above), cf. Syea (1993: 94):

(59) van puasõ dã bazar.
 sell fish in market
 'Fish is sold in the market.'
 (Syea 1993: 92)

⁴⁹ Mauritian Creole has an elaborate system of TMA markers, which is made up of six basic markers (*ti* [+anterior/past], *pe* [progressive], *pu* [definite future], *ava* [indefinite future], *fin* [completive], and *fek* [immediate completive]) that can be used to express at least twelve fine-grained temporal and aspectual differences, see Adone (1994a: ch. 6) for details.

If a generic interpretation is not possible, an overt pronoun must be used in the absence of a TMA marker (Syea 1993: 94f.):

(60) a. *van puasō.
sell fish
b. li van puasō.
he sell fish
'He sells fish.'

Accordingly, we can conclude that in MC, the licensing of referential null subjects is connected to the presence of an overt TMA marker (i.e., an overt realization of Infl).⁵⁰ The missing argument is identified in relation to a prominent discourse topic. In more formal terms, this can be analyzed in terms of a coindexation relation between the null element and a discourse topic, presumably mediated by an abstract operator that occupies a left-peripheral A'-position (cf. Adone 1994a). Furthermore, we may ask whether the empty category can be subsumed under the analysis proposed above, that is, whether it can be analyzed as a null realization of a regular weak/clitic pronoun. Unfortunately, it is not clear to me whether MC exhibits overt clitic pronouns (which may block a null spell-out) or not. However, there is another piece of evidence that suggests that we do not deal with null pronouns here. In embedded clauses, we can observe a curious restriction on the interpretation of null arguments. As shown in (61), an embedded null subject cannot be coreferential with the subject of the matrix clause:

(61) *Zañ_i dir [Ø_i fin al lakaz].
John say ASP go home
'John_i says (he_i) has gone home.'
(Adone 1994a: 114)

⁵⁰ The hypothesis that referential null subjects are licensed by the TMA markers is supported by facts from language acquisition. Adone (1994a) identifies three stages in the acquisition of null subjects in Mauritian Creole. At the first stage, children use a lot of empty subjects (>60%), often in contexts where they are not allowed in the target grammar. The second stage shows a sharp decline in the frequency of null subjects (between 10% and 30%). This change is accompanied by the rise of various TMA markers. The third stage is characterized by a slight increase of null subjects and a more systematic use of TMA markers. Interestingly, from this stage on, null subjects and TMA markers systematically cooccur, that is, the children have correctly acquired the licensing conditions on null subjects of the target grammar.

If the missing argument were a null spell-out of a regular pronoun, we would expect that it can be coreferential with the matrix subject (in fact, that is the preferred interpretation of the English translation of (61)). We can therefore conclude that the empty category must be another kind of element. Adone (1994a) argues that it is a variable bound by an abstract operator that has moved into the left periphery of the embedded clause. As a result, the ungrammaticality of (61) can be attributed to a violation of Principle C of the Binding theory. However, as pointed out by Lipski (2001), it is perhaps more adequate to analyze the null argument as a null constant (*nc*, Rizzi 1994). According to Rizzi (1994), *nc* is an empty category with the properties [-pronominal, -anaphoric, -variable]. It differs from a *wh*-trace (i.e., a variable) in that it does not range over a set of values; rather, its interpretation is fixed to an antecedent given in the immediate discourse context (presumably mediated via an abstract operator). Hence, it is also a referential expression and may not be A-bound, ruling out its use in embedded contexts such as (61).

Turning now to the historical development of null arguments in MC, we can observe that early stages of MC⁵¹ exhibited empty expletives, but lacked the kind of referential null subjects found in the present-day language (cf. Adone 1994b). This suggests that the rise of pro-drop is a rather recent development. In other words, it appears that the pro-drop properties in question did not develop during the original genesis of MC, but rather are the result of a later change. Pro-drop in MC cannot be attributed to its lexifier language (17th and 18th century French), cf. Adone (1994b). Furthermore, Lipski (2001) claims that it cannot be the result of substrate influence, since the relevant languages (several Bantu languages and Malagasy) do not exhibit null subjects. However, it what follows, I am going to argue that there are some facts which suggest that the presence of null subjects in MC can be attributed to (substrate) influence from Malagasy, contra Lipski (2001). To substantiate this claim, let's first review some basic properties of Malagasy.

Malagasy is an Austronesian language with basic VOS word order. It is characterized by the voice system typical of many Austronesian languages (cf. e.g. Keenan 1976 on Malagasy, Schachter 1976, 1990, Kroeger 1993 on Tagalog): distinctive verbal morphology triggers the promotion of one of the verb's arguments to clause-final position. The relevant affixes on the verb indicate the thematic role of the promoted argument. The promoted argument is usually interpreted as a

⁵¹ The creolization of MC took place roughly between 1730 and 1770, cf. Baker and Corne (1986).

familiarity topic (in this way, the voice system serves to implement topic continuity in a discourse, cf. e.g. Hopper 1979, Cooreman, Fox and Givón 1988). In the following examples, the promoted argument and the relevant parts of voice morphology are marked by underlining (AT=actor topic; TT=theme topic; CT=circumstantial topic):

(62) a. Man-asa ny lamba amin' ny savony ny reny.
 AT-wash the clothes with the soap the mother
 b. Sasa-n' ny reny amin' ny savony ny lamba.
 wash-TT the mother with the soap the clothes
 c. An-asa-n' ny reny ny lamba ny savony.
 CT-wash-CT the mother the clothes the soap
 'The mother washes the clothes with the soap.'
 (Sabel 2003: 229f.)

Importantly, the special voice system seems to make available a particular type of discourse-oriented pro-drop. As shown in (63), the promoted argument (but no other argument) can be left out in Malagasy (Pearson 2005 and Hyams et al. 2006: 21):

(63) a. Mamangy an'i Tenda (izy).
 AT.visit OBJ-DET Tenda he
 '(He) is visiting Tenda.'
 b. Mamangy *(azy) i Naivo.
 AT.visit him DET Naivo
 'Naivo is visiting (him).'
 c. Vangian' i Naivo (izy).
 TT.visit DET Naivo he
 '(Him), Naivo is visiting.'
 d. Vangian- *(-ny) i Tenda.
 TT.visit he DET Tenda
 'Tenda, (he) is visiting.'

In other words, it appears that the null subjects of Malagasy (i) are licensed by a special morphology (the voice morphology on the verb which indicates the thematic role of the missing argument) and (ii) are identified in relation to an element which figures prominently in the discourse. Note that this is reminiscent of the conditions on referential pro-drop in MC, where definite null subjects (i) are licensed by a

special inflectional morphology (TMA markers) and (ii) are identified in relation to a prominent discourse topic (e.g., subject of a previous clause, speaker etc.).

Bearing these similarities in mind, one might entertain the idea that the null arguments in MC evolved on the model of the particular type of discourse-oriented pro-drop illustrated in (63) (it is a well-known fact that Creoles often exhibit strong structural similarities with their substrate languages, rather than with their lexifier languages, cf. e.g. Crowley 1992: 268). The pro-drop properties found in MC can then possibly be attributed to substrate influence from Malagasy in the following way.⁵² When learners of MC continued to be confronted with Malagasy (or, rather, Malagasy-influenced) input data that exhibited null arguments, they adapted the licensing mechanism (via distinctive verbal morphology that indicates the thematic role of the missing argument) to the impoverished inflectional system of a Creole language. In the absence of an elaborate voice system, the TMA markers became associated with the formal licensing of pro-drop. In a similar way, the mechanisms of identifying the relevant null element (presumably a null constant) carried over from Malagasy to MC, with the missing argument being interpreted as coreferent with the most prominent discourse topic. In Malagasy, this process is facilitated by structural means (by promoting the discourse topic to clause-final position, together with the distinctive voice morphology), while MC has to resort to conditions that limit the search space to the immediate discourse context (the speaker, or the subject of the previous clause, most often a question). The next section shows that a related development can be observed in Chabacano, a Spanish-based Creole spoken in the Philippines.

Chabacano

‘Chabacano’ is a cover term for a number of different Spanish-based Creoles spoken in the Philippines. The most well-known variety of Chabacano is Zamboangueño, the local vernacular of Zamboanga City in southwestern Mindanao. Other areas where

⁵² See Lipski (2001) for an alternative explanation based on the assumption that null subjects initially developed in embedded contexts via the reanalysis of a variable bound by a left-dislocated element (*e* in (i)):

(i) [sa madam la]_i mo rapel *e_i* ti vini.
 this lady DET I remember TNS come
 ‘This lady, I remember she came.’
 (Adone 1994a: 115)

Chabacano is (or was) wide-spread include the Manila Bay, in particular Cavite and Ternate. The following description of Chabacano is based on Lipski (2001) (see also Steinkrüger 2004, 2006).

As many other Creoles, Chabacano lacks verbal inflection apart from a small class of prefixal TMA markers.⁵³ However, it exhibits two remarkable traits which are quite rare among Creoles, namely basic VSO order and (referential) null subjects, as illustrated by the following examples:

(64) *Null expletive and indefinite subjects*

- a. Ya tiene gente na mundo.
TNS be people in world
'(There) were already people in the world.'
(Lipski 2001: 2)
- b. Ta siña kanila "English".
TNS/ASP teach them English
'(One) teaches them English.'
(Lipski 2001: 6)
- c. Ya tira konele.
TNS shoot him
'He was shot.' (lit., '(One) shot him.')(Lipski 2001: 6)

(65) *Null referential subjects*

- a. Ya man-engkwentro konele na tyangge.
TNS meet her in market
'(I) met her in the market.'
- b. Ya abla kon el muher...
TNS say to the woman
'(He) said to the woman...'

⁵³ The set of TMA markers consists of *ta* (imperfective), *ya* (perfective), *ay/di* (irrealis), and *kabá* (completive). See Steinkrüger (2006) for details.

c. Despues ay anda na eskwela.
 then MOD go to school
 'Then (we) would go to school.'

d. Tiene mas di nobenta años, pero fuerte pa.
 be more than ninety years but strong still
 '(They) are more than ninety years old, but (they) are still strong.'

(Lipski 2001: 4f.)

Similar to MC, null subjects are available for all persons and numbers.⁵⁴ Another important parallel consists in the fact that the missing argument must be identified in relation to an element in the immediate discourse context, compare the following quote taken from Lipski (2001: 3):

"In each case, the referent of the null subject is recoverable from the preceding context, usually being the same as the last-occurring overt pronoun. The usage of null subjects is most common in response to a question, with appropriate shift of pronominal reference."

However, in contrast to MC, it seems that the TMA-markers are not instrumental in the licensing of referential null subjects, as indicated by (65d), where the missing argument can only be interpreted as referring to a certain group of people (despite the lack of a preverbal TMA marker).

According to Lipski (2001), Chabacano exhibits a restriction on the interpretation of embedded null subjects that resembles the relevant constraint in MC: An embedded null subject may not be coreferential with the (overt) subject of the matrix clause when the latter occurs in immediate postverbal position (i.e., the canonical subject position). Again, this can be taken to indicate that the empty category cannot be a null pronoun, but must rather be analyzed as a null constant bound by an abstract (discourse) operator occupying a left-peripheral A'-position.⁵⁵

⁵⁴ Lipski (2001) shows that Chabacano exhibits null direct objects as well.

⁵⁵ Lipski (2001) claims that coreference of an embedded null subject and a matrix subject becomes possible if the latter is fronted to a preverbal (clause-initial) position (similar to Tagalog, fronting is used to focus or (re-) introduce a discourse referent in Chabacano). He then proposes an analysis according to which an embedded null subject (which is not treated as an *nc*) must be c-commanded by the matrix subject, which by assumption is only possible if the latter is fronted to preverbal (clause-initial) position. However, it is not clear to me why the relevant licensing condition cannot

It is generally assumed that a number of (morpho-) syntactic properties of Chabacano (such as basic VSO order, aspects of the inventory of pronouns and the system of marking grammatical functions, cf. Steinkrüger 2006 and Barrios 2006) can be traced back to substrate influence of and, more recently, language contact with the neighboring Austronesian languages, in particular Tagalog and Cebuano. Moreover, Lipski (2001) suggests that the kind of discourse-oriented pro-drop exhibited by Chabacano is also due to influence from Tagalog and Cebuano. Both these languages exhibit the typical Austronesian voice system (cf. Schachter 1976, 1990, Kroeger 1993), that is, the promoted argument's thematic role is indicated by voice morphology on the verb. In contrast to Malagasy, however, the promoted argument does not occupy a designated position, but is marked by (case) particles (*ang* for common nouns and *si* for personal names), compare the following examples from Tagalog:⁵⁶

also be fulfilled by matrix subjects in postverbal position (from which they should also be able to command the lower null subject). Furthermore, it is doubtful whether the example cited by Lipski in favor of this claim (p. 4f., his example (3ah)) actually shows what it is supposed to show. More precisely, (i) seems to be rather a paratactic structure without real embedding. Therefore, the missing arguments are actually not embedded subjects, and the possibility of coreference with 'those kids' is compatible with an analysis of the empty categories in terms of null constants (which must be A-free):

(i) Aquel *mga bata* *sabe man-comprehend, entendé* *kosa ki ta* *lé,*
 those kids know understand understand what they TNS/ASP read
y *sabe* *eskribi.*
 and know write

'Those kids know how to understand, (they) understand what they read, and (they) know how to write.'

⁵⁶ For expository reasons I labeled the relevant case particles SUBJ=subject, OBJ=object, and OBL=oblique. Note that this is slightly misleading, since the *ang*-marked NP arguably does not represent the grammatical subject of the clause, but rather is to be identified with the discourse topic (cf. e.g. Schachter 1990).

(66) a. B-um-ili ang lalake ng isda sa tindahan
 buy.AT SUBJ man OBJ fish OBL shop
 'The man bought fish in a/the shop.'

b. B-in-ili ng lalake ang isda sa tindahan
 buy.TT OBJ man SUBJ fish OBL shop
 'A/the man bought the fish in a/the shop.'

c. B-in-ili-an ng lalake ng isda ang tindahan
 buy.LOCT OBJ man OBJ fish SUBJ shop
 'A/the man bought fish in the shop.'

The *ang/si* marked phrase is normally interpreted as definite and familiar (and as the, continuing topic of the discourse, cf. e.g. Hopper 1979, McGinn 1988, and Cooreman, Fox and Givón 1988).⁵⁷ As in Malagasy, the promoted argument (marked by *ang/si*) can be left out in Tagalog, giving rise to a similar type of discourse-oriented pro-drop (McGinn 1988: 278), where the null realization of a given argument seems to be licensed by morphological means that serve to identify the thematic role of the missing element:

(67) B-um-ili (siya) ng isda sa tindahan
 buy.AT (SUBJ-he) OBJ fish OBL shop
 '(He) bought fish in a/the shop.'

Given that Chabacano 'borrowed' quite a number of grammatical traits from its Austronesian neighbors, it is quite possible that the null arguments found in Chabacano also developed on the model of the kind of discourse-oriented pro-drop that we can observe in Tagalog (and Cebuano), where the licensing (and identification) of the argument gap is linked to the voice morphology indicating the argument's thematic role. This conjecture is further supported by the observation that Chabacano and MC exhibit similar restrictions on the identification of null pronouns via the immediate discourse context (subject of the previous clause, speaker etc.). As noted above, this is possibly related to the absence of structural means to mark the discourse topic (i.e., the typical Austronesian voice system) in the Creole language.

⁵⁷ Non-topic themes are interpreted as indefinite, while other non-topic arguments may be interpreted as definite or indefinite, cf. Schachter (1990: 940f.).

In contrast to MC, however, it is apparently not possible to link the licensing of referential null subjects to the presence of TMA markers in Chabacano (cf. (65d) above). This raises the question of whether there is an alternative structural means in Chabacano that can take up the role of the Austronesian voice morphology in the licensing/identification of null arguments. A possible candidate that comes to mind is the set of prenominal markers (or, case particles) that are used to identify the grammatical function of the verb's arguments in Chabacano (*si* for agents/subjects that are personal names, *kon* for direct objects, *para di* for datives, *na* for locations; furthermore note that there are different series of pronouns for subjects/agents and objects, the latter carrying the marker *kon*, e.g. *konele* 'him/her'). Due to the fact that Chabacano has no grammatical function changing devices such as passive, these markers do not only indicate the grammatical function of the element they modify, but also (at least roughly) its thematic role. In this way, they serve a function which is quite similar to the combined effects of voice morphology and case particles in languages like Tagalog (i.e., indicating the thematic role of a given argument). Let's suppose that this suffices to license a null realization of arguments (as a null constant) in Chabacano, which mimics the relevant licensing conditions that hold in Tagalog/Cebuano, albeit with the impoverished inflectional means of a Creole language (see also Lipski 2001).

The historical developments in Chabacano and Mauritian Creole contrast with the changes affecting Bavarian and Non-Standard French in at least two ways: first, the rise of null arguments does not involve the grammaticalization of agreement markers. Second, the relevant type of discourse-oriented pro-drop develops in an across-the-board manner for all persons and numbers at once (including objects), with a preference for arguments that can readily be recovered from the immediate discourse context (again giving rise to a special role of 1st person pronouns), in contrast to null subjects which arise due to the grammaticalization of agreement morphology. The observed parallels between MC and Chabacano suggest that the kind of pro-drop characteristic of Malagasy and Tagalog, where the topic/null argument is marked by structural means (via verbal voice morphology), represents a very salient feature which is possibly more easily adopted under language contact/substrate influence than other forms of pro-drop.

Chapter 5: Concluding summary

This work has investigated aspects of the relationship between language change and language acquisition, focusing on phenomena at the interface between syntax and the PF-branch of grammar, such as word order change, the emergence of inflectional paradigms, and the rise and loss of null subjects. I have argued that the formal study of language change should be based on a special notion of change, namely *grammar change*, that is, clearly identifiable, discrete differences between the target grammar and the grammar acquired by the learner. In particular, I have tried to show that only under this idealized notion of the proper object of scientific investigation we can develop a restrictive linguistic theory of change, which may eventually lead to a deeper understanding of the ways in which the set of possible changes is restricted and shaped by (i) universal properties of grammar and (ii) the workings of the language acquisition device. In the course of this study, I have sought to identify a couple of factors involving both (i) and (ii) that guided the historical development of the empirical phenomena under investigation.

On the one hand, it is quite obvious that grammar change is constrained by 'hard-wired' properties of grammar that delimit the set of parametric options that must be taken into consideration when the learner tries to reconstruct the target grammar underlying the utterances he/she is exposed to. However, it is also clear that the range of possible changes that we predict is intimately linked to properties of our theory of grammar. In other words, a restrictive theory of the ways in which a certain linguistic property may change over time must be based on a restrictive synchronic theory of that property. Focusing on the phenomenon of word order change, I have argued that the set of possible grammar changes affecting the base order of the verb and its arguments is constrained by universal properties of the

mapping from syntax to PF. More precisely, I have proposed that word order is created by a set of post-syntactic operations that map the cyclic output of the syntactic computation to linear orderings of phonological exponents. This approach makes crucial use of a mechanism called *Edge Replacement* that establishes a linear ordering between neighboring Spell-Out domains/phonological domains via substituting the right edge of a higher domain with the string of exponents created by previous applications of Vocabulary Insertion. Importantly, *Edge Replacement* is subject to conditions such as the *No-Tampering Condition on Edge Replacement* that not only define the set of surface strings that may be created from a given syntactic structure, but also impose restrictions on possible pathways of change via delimiting the range of parametric choices in which the grammar eventually acquired by the learner may possibly differ from the target grammar. In particular, the relevant restrictions predict that the often observed change from a strict SOV grammar to a SVO grammar proceeds in a ‘top-down’ fashion, in the sense that a change in the setting of the (phonological) Head Parameter must first affect exponents of higher functional heads before it can affect exponents of lower functional heads.

On the other hand, universal properties of grammar may also act as causal factors in processes of grammar change. For example, we have seen that the workings of *Edge Replacement* may impede the acquisition of a OV setting for v in grammars that have developed systematic movement of v to head-initial T (see chapter 2 on the loss of OV orders in the ME period). Another example of change driven by universal properties of grammar has been discussed in chapter 3, where I have argued that in early OHG, conditions on the locality of syntactic operations (e.g., strict cyclicity) first promoted the development of an EPP-feature in C (when clauses with initial *thō* ‘then’ were reanalyzed as involving a spec-head relation in the C-domain) and then led to the loss of multiple specifiers in the C-system after *thō* had been reanalyzed as an expletive element.

In addition, I have suggested that another potential source of grammar change are acquisition strategies that the learner applies to the input data if the evidence provided by the PLD does not suffice to trigger a certain parametric choice. Due to the fact that the relation between linear orderings and hierarchical structures is non-unique (i.e., a given string of words may be compatible with different underlying structures), the learner must often take into account additional information to decide whether a given order is the result of displacement or reflects a base generated structure. One such type of information involves the surface interpretation of the utterance. The child has to decide whether the pattern in question is systematically

linked to certain surface-related meaning properties (scope, information structure etc.), which are typically implemented by displacement/internal Merge. Grammar change may result if the learner cannot detect the original semantic/pragmatic trigger of a movement operation (due to ‘noise in the channel’, overuse etc., see chapter 1). More precisely, I have argued that the task of the learner is eased by an acquisition strategy that enables the child to accommodate dislocation phenomena for which no substantial trigger can be detected in PLD. Thus, learners prefer to blindly replicate patterns produced by the target grammar, instead of discarding the relevant word orders encountered in the input, in particular if the latter would give rise to obvious deviations from the set of sentences generated by the target grammar. By assumption, the grammar makes available semantically vacuous EPP-features as a means to imitate the relevant displacement phenomena, giving rise to ‘fossilized’ syntactic patterns that fail to be associated with a particular surface meaning. In chapter 3, I have discussed a set of relevant examples, among them the rise of an EPP-feature in T in the history of English (mimicking movement of anaphoric expressions originally triggered by a [*D*] feature) and the development of an EPP-feature in C in the history of German, which was triggered (amongst other things) by the need to cope with XP-fronting originally triggered by information-structural factors.

Another set of acquisition strategies have been discussed in chapter 4 where I have examined how learners identify the inventory, shape and specification of inflectional markers on the basis of the linguistic input they receive. In particular, I have argued that the cyclic nature of morphological change can be attributed to (i) blocking effects that favor the acquisition of more specified lexical items over less specified lexical items (the *Blocking Principle*, Fuß 2005) and (ii) a preference for minimizing the number of features (or, lexical items) mentioned in the lexicon (*Minimize Feature Content*). Moreover, I have claimed that grammar change may also be determined by blocking in terms of a universal principle grammar that guides the workings of Vocabulary Insertion (i.e., the *Elsewhere Condition*, or Halle’s 1997 *Subset Principle*). In particular, we have seen that a null-spell out of weak pronominal forms may become available as a result of *deblocking* (due to changes that lead to gaps in the paradigm of overt weak pronouns), while the grammaticalization of overt pronouns may block a former null realization. In what follows, I briefly summarize the content of the individual chapters.

Chapter 1 has introduced the notion of *grammar change* as the proper object of a theoretically informed study of language change (i.e., discrete differences between

the grammar acquired by a learner and the grammar(s) that generated the input on the basis of which the learner constructed his/her grammar). I have argued that only under the restricted understanding of language change as grammar change we can hope to discover systematic restrictions on possible changes imposed by properties of UG and the workings of language acquisition. In addition, I have addressed the *logical problem of language change*, showing that upon closer inspection, change is not a rare and paradoxical phenomenon, but rather a logical necessity if we assume that language acquisition is a highly deterministic process.

Chapter 2 has examined the interface of syntax and the PF-branch of grammar, focusing on the question of how the cyclic output of the syntactic computation is mapped to linear orderings of phonological exponents. I have proposed that the linearization process is part of the operation of Vocabulary Insertion, which supplies syntactic terminal nodes with phonological material and thereby incrementally builds a linear string of phonological exponents. The decision whether to add a phonological exponent to the left or to the right of the existing string of elements is determined by a phonological Head Parameter which is taken to ignore a subset of the symmetric c-command relations established in the syntax. Assuming a model of cyclic Spell-Out, I have argued that the phonological component recombines the cyclic output of the syntax into larger and partially overlapping phonological domains. The overlap between neighboring phonological domains is exploited to establish a linear ordering between the chunks of structure transferred to the phonological component. The central proposal I have put forward is that the linearization of separate phonological domains involves a process called *Edge Replacement* that substitutes the right edge of a phonological domain with the string of exponents created so far. This process is subject to the *No-Tampering Condition on Edge Replacement*, which requires that the substitution operation preserve ordering relations that have been established for elements at the overlap. I have shown that an approach in terms of *Edge Replacement* is superior to recent LCA-based analyses (Biberauer et al. 2007, 2008) when it comes to deriving a set of cross-linguistic word order generalizations (e.g., the cross-linguistic absence of VO-Aux, or the correlation between complementizer position and the position of complement clauses). Based on the assumption that there is a basic parametric difference between OV and VO grammars (*The Root Raising Parameter*), I have then developed a typology of possible and impossible grammars, arguing that this model of linearization not only imposes a number of restrictions on possible combinations of parametric choices, but also

makes a number of predictions on possible pathways of grammar change (as demonstrated by a discussion of the OV-VO change in the history of English).

Chapter 3 has examined a particular historical phenomenon in some more detail, focusing on the nature of the V2 constraint in early Germanic. It has become clear that V2 in early Germanic is not a unitary phenomenon. Rather, we can identify a number of different structural configurations that may give rise to surface V2 orders. First, we have seen that contexts such as questions, imperatives, and neg-fronting trigger regular V-to-C movement as early as Gothic, which suggests that these contexts presumably constitute the historical core of the V2 phenomenon in Germanic (see also Eyþórsson 1995, 1996; Kiparsky 1995). Another context that triggers systematic verb fronting in all early Germanic languages are clauses introduced by certain temporal adverbs such as Gothic *þanuh*, OE *þa/bonne* and OHG *thô*, all roughly meaning 'then'. It appears that these elements share a common discourse function across early Germanic, in that they typically introduce new actions/events along the main time line of the story, often in connection with the introduction of new discourse topics. I have presented evidence suggesting that the anaphoric character of 'then' was linked to the specifier of TP, which by assumption served to express the discourse-related property of anaphoricity in the early Germanic languages. In OE, we can observe an additional source of surface V2 orders (so-called 'pseudo V2'), which involves a configuration in which the fronted XP (which occupies SpecCP) and the finite verb (in T) are merely linearly adjacent. Due to the fact that T did not carry an EPP-feature in OE, pseudo-V2 could also give rise to inversion with full DP subjects, which stay behind in their vP-internal theta-position. In contrast, inherently anaphoric elements such as pronouns and *þa/bonne* are always located in SpecTP, giving rise to V3 orders. Finally, there is the parametric option of 'generalized V2', in which the finite verb occupies C in all main clauses, accompanied by EPP-triggered XP-fronting in declaratives. This option is found already in early OHG, gaining a wider distribution in the course of the OHG period. I have argued that across Germanic, the historical development of the V2 property was shaped by an acquisition strategy in which learners may freely posit EPP-features to imitate dislocation phenomena for which they cannot identify a substantial (morphological, or semantic/pragmatic) trigger. Under this scenario, the loss of surface 'pseudo V2' patterns in English can be attributed to the development of a generalized EPP-feature in T that served to mimic patterns that had formerly been derived by merging anaphoric elements in SpecTP. The same acquisition strategy gave rise to generalized V2 in OHG when the learner could no longer

detect the original information-structural triggers of XP-fronting (after the earlier development of obligatory verb fronting in main clauses). I have argued that the rise of a strict V2 grammar was in addition promoted by a set of reanalyses that affected the syntax of clause-initial *thô* 'then'. First, after the development of generalized V-to-C movement, main clauses introduced by *thô* could only be parsed as involving fronting of *thô* to SpecCP. This induced an EPP-feature in C, since *thô*-fronting could not be attributed to strong topic or focus features. Second, when its original discourse function (which was originally linked to SpecTP) became unclear, *thô* was reanalyzed as a semantically light expletive-like element directly merged in SpecCP. I have suggested that the latter change was decisive for the development of generalized V2 in German since the presence of expletives signaled to the learner that a functional head may project only a single specifier, which led to the loss of V3 orders.

In chapter 4, I have widened the scope of this investigation by taking into account changes affecting the phonological realization of inflectional categories. Again, we have seen that the specific course taken by grammar change can be attributed to the workings of acquisition strategies that enable the learner to determine structural properties of the grammar in case the input is ambiguous or defective. In particular, I have argued that the cyclic nature of morphological change can be attributed to the workings of two apparently conflicting acquisition strategies that help the learner to identify phonological exponents of abstract inflectional categories and their feature specifications on the basis of the linguistic input he/she receives. Focusing on the development of verbal agreement marking in Bavarian and Alemannic, I have shown that the course of grammaticalization processes giving rise to new and more distinctive inflectional formatives is shaped by a learning strategy (the *Blocking Principle*, Fuß 2005) that selects the most specified variant in case the input contains more than a single potential realization of a given inflectional category. It has become clear that the workings of blocking-induced change are often countered by a strategy that aims at minimizing the number of elements/features stored in the lexicon (*Minimize Feature Content*, Halle 1997). The tendency to posit an economical system of featural distinctions may lead learners to innovate 'less costly' variants that are not part of the input. I have argued that *Minimize Feature Content* offers a new and more explanatory account of changes traditionally subsumed under the label of analogical change. In particular, we have seen that over time, repeated applications of *Minimize Feature Content* may lead to paradigms in which each phonological exponent is uniquely paired with a single function/meaning (which

has inspired numerous functionalist explanations of analogical change). I have then argued that that it is possible to reconcile the apparently conflicting strategies as separate devices employed by the learner to select an optimal inventory of Vocabulary items based on the evidence available to him/her. In addition, I have taken a brief look at the workings of blocking effects in another major cyclic change concerning the availability of null subjects. Focusing on the transition from pronominal clitics to agreement markers in Bavarian and Non-Standard French, I have suggested that referential null subjects may emerge as a side-effect of the loss of overt realizations of a weak pronominal D-head. The rise of (partial) pro-drop can then be analyzed as an instance of deblocking, where a (universally available) null realization becomes available due the loss of a more specific spell-out (i.e., the former clitic forms). This approach implies that the loss of null subjects may be linked to the development of new overt weak pronominal forms. A relevant case in point seems to be Colloquial Finnish, where the loss of partial pro-drop is accompanied by the development of a new series of weak subject pronouns, which block the less distinctive null realization of a weak pronominal D-head.

* * *

It is a truism that the data set available to the historical linguist is very small if we compare it with the empirical sources available to linguists working on living languages. There are only a restricted number of historical records, and we do not have access to speaker judgments, or any kind of negative evidence. Thus, no matter how carefully we make use of the evidence available to us, we still have to face the fact that there are major gaps and discontinuities in the historical records. Given the state of the empirical evidence, a key question of historical linguistics is how to bridge the gaps in our knowledge of the past. In this work, I have explored how formal approaches to language and language change can help us to (partially) fill in the gaps left by the historical evidence via formulating precise analyses and, possibly, restrictive constraints on what a possible change in a human language can be. Ideally, this approach will not only sharpen our understanding of the past; in addition, the study of language change can inform us about properties of grammar/UG as well, via making available information that cannot be gathered by purely synchronic investigations (Lightfoot 1979, 1991). However, it is important to keep in mind that the historical linguist should not strive to explain too much (cf. e.g. Lightfoot 1999). Many facets of language change are chaotic and unpredictable in nature. Therefore, we should be cautious when it comes to the formulation of

universal pathways of change driven by general considerations of harmony, efficiency or economy (both in functionalist terms or general tendencies shaped by inborn properties of UG). Change is not a necessity, as becomes clear from the observation that certain properties of grammar or even whole grammars can be fairly stable over large stretches of time (cf. e.g. Nichols 2003). Furthermore, the actual outcome of change is shaped by extra-linguistic factors. That is, while it is certainly true that some properties of grammar are more prone to change than others, the question of which innovations prevail via being adopted by the speaker community is largely governed by accidental sociolinguistic factors such as prestige, power etc. Still, that does not mean that anything goes. Rather, crucial aspects of change can be shown to be governed by factors that can be modeled by formal linguistic theory. Accordingly, it has been one goal of this work to explore aspects of grammar and its acquisition that delimit the set of possible changes and therefore provide us with a first approximation of a theory of language change. So while we may not be able to predict the ways in which a given language may change in the future, we may still be able to complete a less ambitious task, namely predicting the ways in which it may *not* change at a given point in time.

References

Abney, Steven. 1987. *The English Noun Phrase in Its Sentential Aspect*. Doctoral dissertation, MIT.

Abraham, Werner. 1993. Null subjects in the history of German: From IP to CP. *Lingua* 89, 117-142.

Abraham, Werner. 1995. *Deutsche Syntax im Sprachenvergleich*. Tübingen: Gunter Narr.

Abraham, Werner. 2003. Autonomous and non-autonomous components of 'grammaticalization': Economy criteria in the emergence of German negation. *Sprachtypologie und Universalienforschung* 56/4, 325-365.

Abraham, Werner. 2004. The grammaticalization of the infinitival preposition – toward a theory of 'grammaticalizing reanalysis'. *Journal of Comparative Germanic Syntax* 7, 111-170.

Ackema, Peter and Ad Neeleman. 2003. Context-sensitive Spell-out. *Natural Language and Linguistic Theory* 21, 681-735.

Ackema, Peter and Ad Neeleman. 2004. *Beyond Morphology. Interface Conditions on Word Formation*. Oxford: Oxford University Press.

Adams, Marianne. 1987. From Old French to the theory of pro-drop. *Natural Language and Linguistic Theory* 5.1, 1-32.

Adger, David. 2003. *Core Syntax. A Minimalist Approach*. Oxford: Oxford University Press.

Adone, Dany. 1994a. *The Acquisition of Mauritian Creole*. Amsterdam: John Benjamins.

Adone, Dany. 1994b. Creolization and language change in Mauritian Creole. In: Dany Adone and Ingo Plag (eds.), *Creolization and Language Change*, 23-43. Tübingen: Niemeyer.

Albright, Adam C. 2002. The Identification of Bases in Morphological Paradigms. Doctoral dissertation, UCLA.

Alexiadou, Artemis. 1997. *Adverb Placement: A Case Study in Antisymmetric Syntax*. Amsterdam: John Benjamins.

Alexiadou, Artemis. 2000. On the syntax of temporal adverbs and the nature of Spec,TP. *Rivista di Linguistica* 12: 53-73, special issue on Adverbs edited by Norbert Corver and Denis Delfitto.

Alexiadou, Artemis and Elena Anagnostopoulou. 1998. Parametrizing AGR: word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16: 491-539.

Allen, Cynthia. 1995. *Case marking and reanalysis: grammatical relations from Old to Early Modern English*. Clarendon, Oxford.

Altmann, Hans. 1984. Das System der enklitischen Personalpronomina in einer mittelbairischen Mundart. *Zeitschrift für Dialektologie und Linguistik* 51.2, 191-211.

Anderson, Stephen R. 1977. On mechanisms by which languages become ergative. In *Mechanisms of Syntactic Change*, C. N. Li (ed.), 317-363. Austin: University of Texas Press.

Anderson, Stephen R. 1980. On the development of morphology from syntax. In *Historical Morphology*, J. Fisiak (ed.), 51-69. Berlin: Mouton de Gruyter.

Anderson, Stephen R. 1986. Disjunctive ordering in inflectional morphology. *Natural Language and Linguistic Theory* 4, 1-31.

Anderson, Stephen R. 1992. *A-morphous Morphology*. Cambridge: Cambridge University Press.

Anderson, Stephen R. 1993. Wackernagel's revenge: Clitics, morphology, and the syntax of second position. *Language: Journal of the Linguistic Society of America* 69.1, 68-98.

Ariel, Mira. 2000. The development of person agreement markers: from pronoun to higher accessibility markers. In *Usage-based Models of Language*, S. Kemmer and M. Barlow (eds.), 197-261. Stanford: CSLI.

Árnason, Kristján. 1985. Icelandic word stress and metrical phonology. *Studia Linguistica* 2/85, 93-129.

Árnason, Kristján. 2002 Kuhn's Laws in Old Icelandic Prose and Poetry. *Journal of Germanic Linguistics* 14.3, 201-241.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, Mass.: The MIT Press.

Arregi, Karlos. 1999. Person and number inflection in Basque. In *MIT Working Papers in Linguistics* 33: *Papers on Morphology and Syntax, Cycle One*, K. Arregi, B. Bruening, C. Krause, and V. Lin (eds.), 229-264. Cambridge, Mass.: Department of Linguistics and Philosophy, MIT.

Ashby, William. 1977. *Clitic Inflection in French. An Historical Perspective*. Amsterdam: Rodopi.

Auger, Julie. 1993. More evidence for verbal agreement marking in Colloquial French. In *Linguistic Perspectives on the Romance Languages*, W. J. Ashby, M. Mithun, G. Perissinotto, and E. Raposo (eds.), 177-98. Amsterdam: John Benjamins.

Auger, Julie. 1994a. *Pronominal clitics in Quebec Colloquial French: A morphological analysis*. PhD dissertation, University of Pennsylvania (IRCS report 94-29).

Auger, Julie. 1994b. On the nature of subject clitics in Picard. In M. L. Mazzola (ed.), *Issues and Theory in Romance Linguistics*, 159-179. Washington, DC: Georgetown University Press.

Auger, Julie 2003. Pronominal clitics in Picard revisited. In R. Núñez-Cedeño, L. López, and R. Cameron (eds.), *A Romance Perspective on Language Knowledge and Use*, 3-20. Amsterdam: Benjamins.

Axel, Katrin. 2001. Althochdeutsche Modalverben als Anhebungsverben. In *Modalität und Modalverben im Deutschen*, R. Müller and M. Reis (eds.), 37-60. *Linguistische Berichte* Sonderheft 9. Hamburg: Buske.

Axel, Katrin. 2007. *Studies on Old High German Syntax: Left Sentence Periphery, Verb Placement and Verb-Second*. Amsterdam: John Benjamins

Axel, Katrin and Helmut Weiß. 2007. Pro-drop in the history of German. From Old High German to the modern dialects. Ms., Universities of Saarbrücken and Frankfurt.

Bacquet, Paul. 1962. *La structure de la phrase verbale a l'époque alfrédienne*. Paris: Belles Lettres.

Baker, Mark C. 1985. The Mirror Principle and morphosyntactic explanation. *SvenoLinguistic Inquiry* 16.3, 373-415.

Baker, Mark C. 1988. *Incorporation. A Theory of Grammatical Function Changing*. Chicago: Chicago University Press.

Baker, Philip and Chris Corne. 1986. Universals, substrata and the Indian Ocean Creoles. In P. Muysken and N. Smith (eds.), *Substrata versus Universals in Creole Genesis*, 163-183. Amsterdam: Benjamins.

Bally, Charles. 1965. *Linguistique générale et linguistique française* (4th edition). Bern.

Barrios, Aireen. 2006. Austronesian elements in Philippine Creole Spanish. Paper presented at Tenth International Conference on Austronesian Linguistics. 17-20 January 2006. Puerto Princesa City, Palawan, Philippines. (<http://www.sil.org/asia/philippines/ical/papers.html>)

Bauer, Gero. 1970. *Studien zum System und Gebrauch der 'Tempora' in der Sprache Chaucers and Gowers*. (Wiener Beiträge zur englischen Philologie, 73). Vienna: Wilhelm Brahmüller.

Bauer, Walter. 1979. *A Greek-English Lexicon of the New Testament and Other Early Christian Literature*. 2nd ed. Chicago: University of Chicago Press.

Bayer, Josef. 1984. COMP in Bavarian syntax. *The Linguistic Review* 3, 209-274.

Bayer, Josef. 1996. *Directionality and Logical Form: On the Scope of Focusing Particles and Wh-in-situ*. Dordrecht: Kluwer.

Bayer, Josef. 1999. Final complementizers in hybrid languages. *Journal of Linguistics* 35.2, 233-271.

Beard, Robert 1988. The separation of derivation and affixation: Toward a Lexeme-Morpheme Base Morphology. *Quarderni di semantica* 5.277-287.

Beard, Robert 1995. *Lexeme-Morpheme Base Morphology; a General Theory of Inflection and Word Formation*. Albany, NY: SUNY Press.

Behaghel, Otto. 1928, *Deutsche Syntax. Eine geschichtliche Darstellung*. Vol. 3: *Die Satzgebilde*. Heidelberg: C. Winter.

Behaghel, Otto. 1932, *Deutsche Syntax. Eine geschichtliche Darstellung*. Vol. 4: *Wortstellung*. Heidelberg: C. Winter.

Belletti, Adriana. 1990. *Generalized Verb Movement: Aspects of Verb Syntax*. Turin: Rosenberg and Sellier.

Belletti, Adriana. 2004. Aspects of the low IP area. In *The Structure of CP and IP*, Luigi Rizzi (ed.), 16-51. Oxford: Oxford University Press.

Bennis, Hans and Liliane Haegeman. 1984. On the status of agreement and relative clauses in West Flemish. In: Wim de Geest and Yvan Putseys (eds.), *Sentential complementation*, 33-55. Dordrecht: Foris.

Benveniste, Emile. 1950. La phrase nominale. *Bulletin de la Société Linguistique de Paris*, 46: 19-36.

Benveniste, Emile. 1966. *Problèmes de linguistique générale*. Paris: Editions Gallimard.

Benveniste, Emile. 1968. Mutations of Linguistic Categories. In *Directions for Historical Linguistics: A Symposium*, W. Lehmann and Y. Malkiel (eds.), 83-94. Austin: University of Texas Press.

Berwick, Robert C. 1985. *The Acquisition of Syntactic Knowledge*. Cambridge, Mass.: MIT Press.

Besch, Werner. 1967. *Sprachlandschaften und Sprachausgleich im 15. Jahrhundert*. München: Francke Verlag.

Besten, Hans den. 1986 Decidability in the syntax of verbs of (not necessarily) West-Germanic languages. *Groninger Arbeiten zur Germanistischen Linguistik* 28, 232-256.

Betten, Anne. 1987. Zur Satzverknüpfung im althochdeutschen Tatian. Textsyntaktische Betrachtungen zum Konnektor *thô* und seinen lateinischen Entsprechungen. In: Rolf Bergmann, Heinrich Tiefenbach, Lothar Voetz (eds., in association with Herbert Kolb, Klaus Matzel and Karl Stackmann), *Althochdeutsch. Band I: Grammatik. Glossen und Texte*, 395-407. Heidelberg: Carl Winter.

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.

Biberauer, Theresa and Ian Roberts. 2006. The Loss of residual 'head-final' orders and remnant fronting in Late Middle English: Causes and consequences. *Comparative Studies in Germanic Syntax: From Afrikaans to Zurich German*. J. M. Hartmann and L. Molnarfi (eds.), 263-297. Amsterdam: Benjamins.

Biberauer, Theresa and Ian Roberts. 2008. Subjects, tense and verb movement in Germanic and Romance. *Cambridge Occasional Papers in Linguistics* 3: 24-43.

Biberauer, Theresa, A. Holmberg and I. Roberts. 2007. Disharmonic word-order systems and the Final-over-Final-Constraint (FOFC). To appear in *Proceedings of the Incontro di grammatica generativa* 33.

Biberauer, Theresa, A. Holmberg and I. Roberts. 2008. Structure and Linearization in Disharmonic Word Orders. In *Proceedings of the 26th West Coast Conference on Formal Linguistics*, Ch. B. Chang and H. J. Haynie (ed.), 96-104. Somerville: Cascadilla Proceedings Project.

Bierwisch, Manfred. 1967. Syntactic features in morphology: general problems of so-called pronominal inflection in German. In: *To Honor Roman Jakobson: Essays on the Occasion of His Seventieth Birthday*, 239-270. The Hague: Mouton.

Blood, Cindy. 1992. Subject-verb agreement in Kisar. *NUSA. Linguistics Studies of Indonesia and Other Languages in Indonesia*, vol. 34, 1-21. Jakarta: Badan Penyelenggara NUSA.

Bobaljik, Jonathan. 1994. What does adjacency do? In Heidi Harley and Colin Phillips (eds.), *The Morphology-Syntax Connection*. MITWPL 22, 1-32. Cambridge, Mass.: Department of Linguistics and Philosophy, MIT.

Bobaljik, Jonathan. 1995. *Morphosyntax: the syntax of verbal inflection*. Doctoral dissertation, MIT.

Bobaljik, Jonathan. 2002. A-chains at the PF interface. *Natural Language and Linguistic Theory* 20, 197-267.

Bobaljik, Jonathan. 2003. Realizing Germanic inflection: why morphology does not drive syntax. *Journal of Comparative Germanic Linguistics* 6.2: 129-167.

Boeckx, Cedric, and Kleanthes Grohmann. 2007. Putting phases into perspective. *Syntax* 10.2, 204-222.

Bohnacker, Ute. 2006. When Swedes begin to learn German: From V2 to V2. *Second Language Research* 22.4, 443-486.

Bohnacker, Ute and Christina Rosén. 2007. How to start a declarative V2 clause: Transfer of syntax or information structure in L2 German. In press in *Nordlyd* 34.3: *Papers from the*

Language Acquisition Workshop, 22nd SCL 2006, M. Anderssen and M. Westergaard (eds.). Tromsø: CASTL.

Bokamba, Eyamba G. 1976. *Question Formation in Some Bantu Languages*. Doctoral dissertation, Indiana University.

Bonet, Eulalia. 1991. *Morphology After Syntax. Pronominal Clitics in Romance*. Doctoral dissertation, MIT.

Bopp, Franz. 1816. *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprache*. Frankfurt am Main: Andreäische Buchhandlung.

Borer, Hagit. 1984. *Parametric Syntax*. Dordrecht: Foris.

Borer, Hagit. 1998. Morphology and Syntax. In *The Handbook of Morphology*, A. Spencer and A. Zwicky (eds.), 151-189. Oxford and Malden, MA: Blackwell.

Bouchard, Denis. 2002. *Adjectives, Number, and Interfaces*. Amsterdam: Elsevier/North Holland.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24.4, 591-656.

Bowers, John. 2002. Transitivity. *Linguistic Inquiry* 33.2, 183-224.

Brandi, Luciana and Patrizia Cordin. 1989. Two Italian dialects and the null subject parameter. In *The Null Subject Parameter*, O. Jaeggli and K. Safir (eds.), 111-142. Dordrecht: Kluwer.

Branigan, Phil. 2005. The phase-theoretic basis for subject-aux inversion. Ms., Memorial University.

Braune, Wilhelm and Ingo Reiffenstein. ¹⁵2004. *Althochdeutsche Grammatik I*. Tübingen: Niemeyer.

Breivik, Leiv. 1989. On the causes of syntactic change in English. In Leiv Breivik and Ernst Håkon Jahr (eds.), *Language change: contributions to the studies of its causes*. Berlin: Mouton de Gruyter, 29-70.

Breivik, Leiv. 1990. *Existential there: a synchronic and diachronic study*. Oslo: Novus Press.

Bresnan, Joan. 1973. The syntax of the comparative clause construction in English. *Linguistic Inquiry* 4, 275-343.

Brinkmann, Hennig. 1931. *Sprachwandel und Sprachbewegungen in althochdeutscher Zeit*. Jena: Frommann.

Brodführer, Eduard. 1906. Beiträge zur Syntax Willirams. University of Halle-Wittenberg.

Brugmann, Karl. 1916. *Vergleichende Laut-, Stammbildungs- und Flexionslehre der indogermanischen Sprachen* (2nd edition). Vol. 2.3: *Lehre von den Wortformen und ihrem Gebrauch: Verbum finitum*. Strassburg: Trübner.

Brugmann, Karl. 1917. *Der Ursprung des Scheinsubjekts 'es' in den germanischen und den romanischen Sprachen*. Leipzig: Trübner.

Brugmann, Karl and Berthold Delbrück. 1900. *Grundriss der vergleichenden Grammatik der Indogermanischen Sprachen* (Vol. 5). Strassburg: Trübner.

Büring, Daniel. 1995. On the base position of embedded clauses in German. *Linguistische Berichte* 159, 370-380.

Büring, Daniel and Katharina Hartmann. 1997. Doing the right thing. *The Linguistic Review* 14, 1-42.

Butler, Milton Chadwick. 1980. *Grammatically motivated subjects in Early English*. Doctoral dissertation, University of Texas, Austin.

Bybee, Joan L. 1985. *Morphology*. Amsterdam: John Benjamins.

Bybee, Joan L., William Pagliuca and Revere Perkins. 1990. On the asymmetries in the affixation of grammatical material. In *Studies in Typology and Diachrony*, W. Croft, K. Denning and S. Kemmer (eds.), 1-42. Amsterdam: John Benjamins.

Campbell, Alistair. 1970. Verse influences in Old English prose. In *Philological Essays: Studies in Old and Middle English Language and Literature in Honor of H.D. Merritt*, J. Rosier (ed.), 98-108. The Hague: Mouton.

Campbell, Lyle. 1991. Some grammaticalization changes in Estonian. In *Approaches to Grammaticalization*, Vol. I: Theoretical and methodological issues, E. C. Traugott and B. Heine (eds.), 285-99. Amsterdam: John Benjamins.

Cardinaletti, Anna and Ian Roberts. 1991. Clause structure and X-second. Ms., Università di Venezia and Université de Genève. Published in Guglielmo Cinque (ed.). 2002. *Functional Structure in DP and IP: The Cartography of Syntactic Structures, Volume 1*. Oxford: Oxford University Press, 123-166.

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*, 145-232. Berlin: de Gruyter.

Carstens, Vicki. 2003. Rethinking complementizer agreement: Agree with a Case-checked goal. *Linguistic Inquiry* 34.3, 393-412.

Ćavar, Damir and Chris Wilder. 1999. 'Clitic third' in Croation. In *Clitics in the Languages of Europe*, H. van Riemsdijk (ed.), 429-467. Berlin: Mouton de Gruyter.

Chafe, Wallace L. 1977. The evolution of third person verb agreement in the Iroquoian Languages. In *Mechanisms of Syntactic Change*, C. N. Li (ed.), 493-524. Austin: University of Texas Press.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, Mass.: The MIT Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1986a. *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.

Chomsky, Noam. 1986b. *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1991. Some notes on economy of derivations and representations. In *Principles and Parameters in Comparative Grammar*, R. Freidin (ed), 417-454. Cambridge, Mass.: The MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: The MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In Roger Martin, David Michaels, and Juan Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, 89-155. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale. A Life in Language*, M. Kenstowicz (ed.), 1-52. Cambridge Mass.: MIT Press.

Chomsky, Noam. 2002. An interview on minimalism. In *Noam Chomsky: On Nature and Language*, A. Belletti and L. Rizzi (eds.), 92-161. Cambridge: Cambridge University Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and Beyond. The Cartography of Syntactic Structures, Vol. 3*, 104-131. Oxford: Oxford University Press.

Chomsky, Noam. 2005. On phases. Ms., MIT.

Chomsky, Noam and Howard Lasnik. 1995. The theory of Principles and Parameters. In *The Minimalist Program*, N. Chomsky, 13-127. Cambridge, Mass.: The MIT Press.

Chung, Sandra, William Ladusaw, and James McCloskey. 1995. Sluicing and logical form. *Natural Language Semantics* 3, 239-282.

Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24.2, 239-297.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. Oxford: Oxford University Press.

Clark, Eve V. 1998. The acquisition of morphology. In *Handbook of morphology*, Andrew J. Spencer and Arnold M. Zwicky (eds.), 374-389. Oxford: Blackwell.

Clark, Robert and Ian Roberts. 1993. A computational model of language learnability and language change. *Linguistic Inquiry* 24: 299-345.

Collins, Chris. 1997. *Local Economy*. Cambridge, Mass.: The MIT Press.

Cooreman, Ann, Barbara Fox, and Talmy Givón. 1988. The discourse definition of ergativity: A study in Chamorro and Tagalog texts. In R. McGinn (ed.), *Studies in Austronesian Linguistics*, 387-425. Athens, Ohio: Center for Southeast Asia Studies.

Corbett, Greville. 1991. *Gender*. Cambridge: Cambridge University Press.

Corbett, Greville. 1995. Agreement. In *Syntax: An International Handbook of Contemporary Research* Vol. 2, J. Jacobs, A. von Stechow, W. Sternefeld, and T. Vennemann (eds.), 1235-1244. Berlin: Walter de Gruyter.

Corbett, Greville. 2000. *Number*. Cambridge: Cambridge University Press.

Counts, David R. 1969. *A Grammar of Kialai-Kove*. *Oceanic Linguistics Special Publication No. 6*. Honolulu: University of Hawaii Press.

Craenenbroeck, Jeroen van. 2004. *Ellipsis in Dutch dialects*. LOT Dissertation Series 96. Utrecht: LOT.

Craenenbroeck, Jeroen van and Marjo van Koppen. 2002. The locality of agreement and the CP-domain. Paper presented at GLOW 2002, Amsterdam.

Crowley, Terry. 1992. *An Introduction to Historical Linguistics*. Oxford: Oxford University Press.

Curnow, Timothy Jowan. 2001. What language features can be borrowed? In *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, A. Aikhenvald and R. M. W. Dixon (eds.), 412-436. Oxford: Oxford University Press.

Cysouw, Michael. 2003. *The Paradigmatic Structure of Person Marking*. Oxford: Oxford University Press.

Dal Negro, Silvia. 2004. *The Decay of a Language. The Case of a German Dialect in the Italian Alps*. Frankfurt: Peter Lang.

Dasgupta, Probal. 2007. The ubiquitous complementizer. In *Linguistic Theory and South Asian Languages. Essays in Honor of K.A. Jayaseelan*, J. Bayer, T. Bhattacharya, and M.T. Hany Babu (eds.), 163-173. Amsterdam: John Benjamins.

Davis, Henry. 1998. Person splits, φ -features and temporal architecture. *GLOW Newsletter* 40, 83-84.

Davison, Alice. 2007. Word orders, parameters, and the extended COMP projection. In *Linguistic Theory and South Asian Languages. Essays in Honor of K.A. Jayaseelan*, J. Bayer, T. Bhattacharya, and M.T. Hany Babu (eds.), 175-198. Amsterdam: John Benjamins.

De Cat, Cécile. 2005. French subject clitics are not agreement markers. *Lingua* 115, 1195-1219.

Déchaine, Rose-Marie and Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33, 409-442.

DeGraff, Michel. 1993. Is Haitian Creole a pro-drop language?. In F. Byrne and J. Holm (eds.), *Atlantic Meets Pacific*, 71-90. Amsterdam: Benjamins.

de Haan, Germen and Fred Weerman. 1986. Finiteness and verb fronting in Frisian. In Hubert Haider and Martin Prinzhorn (eds.), *Verb Second Phenomena in Germanic Languages*, 77-110. Dordrecht: Foris.

Denison, David. 1993. *English Historical Syntax*. London: Longman.

Dewey, Tonya. 2007. *The Origins and Development of Germanic V2*. Doctoral dissertation, University of California, Berkeley.

Dikken, Marcel den. 1995. *Particles: On the Syntax of Verb-Particle, Triadic, and Causative Constructions*. Oxford University Press, Oxford.

Dimmendaal, Gerrit Jan. 1983. *The Turkana Language*. Dordrecht: Foris.

Dittmer, Arne and Ernst Dittmer. 1998. *Studien zur Wortstellung – Satzgliedstellung in der althochdeutschen Tatianübersetzung*. Göttingen: Vandenhoeck & Ruprecht.

Dobashi, Yoshihito. 2003. *Phonological Phrasing and Syntactic Derivation*. Doctoral dissertation, Cornell University.

Doherty, Cathal. 2000. Residual verb second in Early Irish: On the nature of Bergin's Construction. *Diachronica* 17, 5-38.

Donohue, Mark. 2002. A most agreeable language. Ms., University of Sydney.

Dresher, Elan. 1999. Charting the learning path: cues to parameter setting. *Linguistic Inquiry* 30.1: 27-67.

Dresher, Elan. and Jonathan Kaye. 1990. A computational learning model for metrical phonology. *Cognition* 34: 137-195.

Dryer, Matthew. 1980. The positional tendencies of sentential noun phrases in Universal Grammar. *The Canadian Journal of Linguistics* 25.2, 123-195.

Dryer, Matthew. 1992. The Greenbergian word order correlations. *Language* 68.1, 81-138.

Dryer, Matthew 2007 (to appear). Verb-Object-Negative Order in Central Africa. In *Negation Patterns in West Africa*, N. Cyffer, E. Ebermann, and G. Ziegelmeyer. Amsterdam: John Benjamins.

Dürscheid, Christa. 1989. *Zur Vorfeldbesetzung in deutschen Verbzweit-Strukturen*. Trier: Wissenschaftlicher Verlag.

Eggers, Hans. 1964. *Der althochdeutsche Isidor*. Tübingen: Niemeyer.

Embick, David. 1997. *Voice and the Interfaces of Syntax*. Doctoral dissertation, University of Pennsylvania.

Embick, David. 2000. Features, syntax, and categories in the Latin perfect. *Linguistic Inquiry* 31.2: 185-230.

Embick, David and Alec Marantz. 2006. Architecture and blocking. Ms., University of Pennsylvania and MIT.

Embick, David and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32, 555-595.

Embick, David and Rolf Noyer. 2007. Distributed Morphology and the syntax/morphology interface. In *The Oxford Handbook of Linguistic Interfaces*, G. Ramchand and C. Reiss (eds.), 289-324. Oxford: Oxford University Press.

Enkvist, Nils Erik and Brita Wårwik. 1987. Old English *þa*, temporal chains, and narrative structure. In: Anna Giacalone Ramat, Onofrio Carruba and Giuliano Bernini (eds.), *Papers from the 7th International Conference on Historical Linguistics*, 221-237. Amsterdam: Benjamins.

Epstein, Samuel D. 2007. On I(nternalist)-functional explanation in minimalism. *Linguistic Analysis* 33(1-2): 20-53.

Epstein, Samuel D., Erich M. Groat, Ruriko Kawashima, and Hisatsugu Kitahara. 1998. *A Derivational Approach to Syntactic Relations*. Oxford: Oxford University Press.

Erdmann, Oskar. 1874. *Untersuchungen über die Syntax Otfrieds. Erster Teil: Die Formationen des Verbums in einfachen und zusammengesetzten Sätzen*. Halle/Saale: Verlag der Buchhandlung des Waisenhauses.

Eyþórsson, Thórhallur. 1995. *Verbal syntax in the early Germanic languages*. Doctoral dissertation, Cornell University.

Eyþórsson, Thórhallur. 1996. Functional categories, cliticization, and verb movement in the Early Germanic languages. Höskuldur Thráinsson, Samuel. D. Epstein and Steve Peter (eds.), *Studies in comparative Germanic syntax II*. Dordrecht: Kluwer, 109-139.

Faarlund, Jan Terje. 2004. *The Syntax of Old Norse*. Oxford: Oxford University Press.

Falk, Cecilia. 1993. Non-referential subjects and agreement in the history of Swedish. *Lingua* 89, 143-180.

Fanselow, Gisbert. 2002. Quirky 'subjects' and other specifiers. In *More than Words*, B. Stiebels and I. Kaufmann (eds.), 227-250. Berlin: Akademie Verlag.

Fanselow, Gisbert. 2004. Cyclic phonology-syntax-interaction: Movement to first position in German. *Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure (ISIS) 1*. S. Ishihara, M. Schmitz and A. Schwarz (eds.), 1-42. Potsdam: Universitätsverlag Potsdam.

Fanselow, Gisbert. 2006. On pure syntax (uncontaminated by information structure). In *Form, Structure, and Grammar. A Festschrift presented to Günther Grewendorf on Occasion of his 60th Birthday*, P. Brandt and E. Fuß (eds.), 137-157. Berlin: Akademie Verlag.

Fanselow, Gisbert and Damir Ćavar. 2002. Distributed deletion. In *Theoretical Approaches to Universals*. Alexiadou, A. (ed). 65-107. Amsterdam: John Benjamins.

Ferraresi, Gisella. 1997. *Word order and phrase structure in Gothic. A comparative study*. Doctoral dissertation, Universität Stuttgart.

Fintel, Kai von. 1995. The formal semantics of grammaticalization. *Proceedings of NELS 25: Workshop on Language Change*, 175-189, Amherst, Mass.: GLSA.

Fischer, Olga, Ans van Kemenade, Willem Koopman, Wim van der Wurff. 2000. *The Syntax of Early English*. Cambridge: Cambridge University Press.

Fischer, Silke. 2004. *Towards an Optimal Theory of Reflexivization*. Doctoral dissertation, University of Tübingen.

Fodor, Janet. 1998. Unambiguous triggers. *Linguistic Inquiry* 29.1: 1-36.

Fonseca-Greber, Bonnie. 2000. *The Change from Pronoun to Clitic to Prefix and the Rise of Null Subjects in Spoken Swiss French*. Doctoral dissertation, University of Arizona, Tucson.

Fonseca-Greber, Bonnie and Linda R. Waugh. 2003. The subject clitics of Conversational European French: Morphologization, grammatical change, semantic change, and change in progress. In R. Núñez-Cedeño, L. López, and R. Cameron (eds.), *A Romance Perspective on Language Knowledge and Use*, 99-117. Amsterdam: Benjamins.

Foster, Robert. 1975. The use of *þa* in Old and Middle English narratives. *Neuphilologische Mitteilungen* 77, 404-414.

Fourquet, Jean. 1938. *L'Ordre des éléments de la phrase germanique ancien. Etudes de syntaxe de position*. Paris: Les Belles Lettres.

Fox, Danny and David Pesetsky. 2003. Cyclic linearization and the typology of movement. Paper presented at GLOW XXVI, Lund University.

Fox, Danny and David Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31, 1-46.

Frey, Werner. 2004. A medial topic position for German. *Linguistische Berichte* 198, 153-190.

Frey, Werner. 2006. How to get an object-es into the German prefield. In *Form, Structure, and Grammar. A Festschrift presented to Günther Grewendorf on Occasion of his 60th Birthday*, P. Brandt and E. Fuß (eds.), 159-185. Berlin: Akademie Verlag.

Friedemann, Marc-Ariel. 1997. *Sujets Syntaxiques: Positions, Inversions et pro*. Frankfurt: Peter Lang.

Fukui, Naoki. 1986. *A Theory of Category Projection and Its Application*. Doctoral dissertation, MIT.

Fukui, Naoki and Yuji Takano. 1998. Symmetry in Syntax: Merge and Demerge. *Journal of East Asian Linguistics* 7.1, 27-86.

Fuß, Eric. 2003. On the historical core of V2 in Germanic. *Nordic Journal of Linguistics* 26.2, 195-231.

Fuß, Eric. 2004. Diachronic clues to pro-drop and complementizer agreement in Bavarian. In *Diachronic Clues to Synchronic Grammar*, Eric Fuß and Carola Trips (eds.), 59-100. Amsterdam: John Benjamins.

Fuß, Eric. 2005. *The Rise of Agreement. A Formal Approach to the Syntax and Grammaticalization of Verbal Inflection*. Amsterdam: John Benjamins.

Fuß, Eric. 2007. Cyclic spell-out and the nature of post-syntactic operations. *Linguistic Analysis* 33.3-4, 267-302.

Fuß, Eric. 2008. Multiple agreement and the representation of inflection in the C-domain. *Linguistische Berichte* 213, 78-108.

Fuß, Eric and Carola Trips. 2002. Variation and change in Old and Middle English. On the validity of the Double Base Hypothesis. *Journal of Comparative Germanic Linguistics* 4, 171-224.

Fuß, Eric and Carola Trips. 2003. *þa, þonne*, and V2 in Old and Middle English. Paper presented at the annual LAGB meeting, Oxford University.

Gabelentz, Georg von der. 1891. *Die Sprachwissenschaft. Ihre Aufgaben, Methoden, und bisherigen Ergebnisse*. Leipzig: Weigel.

Gelder, Elly van. 2004. *Grammaticalization as Economy*. Amsterdam: John Benjamins.

Gerritsen, Marinus. 1984. Divergent word order developments in the Germanic languages: a description and a tentative explanation. In *Historical Syntax*, Jacek Fisiak (ed.), 107-137. Berlin: Mouton.

Gerlach, Birgit. 2001. Optimality aspects of clitic doubling in Romance. In *Clithques et cliticisation*, P. de Carvalho, L. Labrune, F. Lambert, CL. Muller and K. Ploog (eds.), 343-363. Paris: Champion.

Gerlach, Birgit. 2002. *Clitics Between Syntax and Lexicon*. Amsterdam: John Benjamins.

Giegerich, Heinz. 2001. Synonymy blocking and the elsewhere condition: Lexical morphology and the speaker. *Transactions of the Philological Society*, 99(1), 65-98.

G. M. Gilligan. 1987. *A Cross-Linguistic Approach to the Pro-Drop-Parameter*. Doctoral dissertation. University of Southern California.

Givón, Talmy. 1976. Topic, Pronoun, and Grammatical Agreement. In *Subject and Topic*, C.N. Li (ed.), 149-188. New York: Academic Press.

Glasbey, Sheila. 1993. Distinguishing between events and times: Some evidence from the semantics of *then*. *Natural Language Semantics*. 1: 285-312.

Green, Georgia. 1980. Some wherewords of English inversions. *Language* 56, 582-601.

Greenberg, Joseph. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In: J. Greenberg (ed.), *Universals of Grammar*, 73-113. Cambridge, Mass.: The MIT Press.

Grewendorf, Günther. 1988. *Aspekte der deutschen Syntax*. Tübingen: Narr.

Grewendorf, Günther. 1989. *Ergativity in German*. Dordrecht: Foris.

Grewendorf, Günther. 1995. Präsens und Perfekt im Deutschen. *Zeitschrift für Sprachwissenschaft*, 14.1, 72-90.

Grewendorf, Günther. 2001. Multiple wh-fronting. *Linguistic Inquiry* 32.1, 87-122.

Grewendorf, Günther. 2002. *Minimalistische Syntax*. Tübingen: A. Francke.

Grewendorf, Günther. 2005. The discourse configurationality of scrambling. In *The Free Word Order Phenomenon: Its Syntactic Sources and Diversity*, J. Sabel and M. Saito (eds.), 75-135. Berlin: Mouton de Gruyter.

Grewendorf, Günther and Joachim Sabel. 1999. Scrambling in German and Japanese: Adjunction versus multiple specifiers. *Natural Language and Linguistic Theory* 17, 1-65.

Grimm, Jacob. 1890. *Deutsche Grammatik* Vol. 3. Gütersloh.

Grohmann, Kleanthes. 2003. *Prolific Domains: On the Anti-Locality of Movement Dependencies*. Amsterdam: John Benjamins.

Grohmann, Kleanthes and Michael T. Putnam. 2007. Dynamic stress assignment. *Linguistic Analysis* 33.3-4, 337-374.

Guiraud, Pierre. 1968. *L'ancien français* (3rd edition). Paris.

Haeberli, Eric. 1999. *Features, categories and the syntax of A-positions. Synchronic and diachronic variation in the Germanic languages*. Doctoral dissertation, University of Geneva.

Haeberli, Eric. 2000. Observations on the loss of verb second in the history of English. Ms., University of Reading.

Haeberli, Eric. 2002a. *Features, Categories and the Syntax of A-Positions. Cross-Linguistic Variation in the Germanic Languages*. Dordrecht: Kluwer.

Haeberli, Eric. 2002b. Inflectional morphology and the loss of V2 in English. In David Lightfoot (ed.), *Syntactic effects of morphological change*. Oxford: Oxford University Press, 88-106.

Haeberli, Eric. 2004. Syntactic effects of inflectional morphology and competing grammars. In *Diachronic Clues to Synchronic Grammar*, E. Fuss and C. Trips (eds.), 101-130. Amsterdam: John Benjamins.

Haegeman, Liliane. 1990. Subject pronouns and subject clitics in West Flemish. *The Linguistic Review* 7, 333-363.

Haegeman, Liliane. 1992. *Theory and Description in Generative Syntax: A Case Study in West-Flemish*. Cambridge: Cambridge University Press.

Haegeman, Liliane and Henk van Riemsdijk. 1986. Verb Projection Raising, scope and the typology of rules affecting verbs. *Linguistic Inquiry* 17, 417-466.

Haider, Hubert. 1982. *Dependenzen und Konfigurationen: Zur deutschen V-Projektion*. *Groniger Arbeiten zur Germanistischen Linguistik* 21, 1-60.

Haider, Hubert. 1993. *Deutsche Syntax – generativ*. Tübingen: Gunter Narr.

Haider, Hubert. 1994. (Un-)heimliche Subjekte – Anmerkungen zur Pro-drop Causa, im Anschluß an die Lektüre von Osvaldo Jaeggli and Kenneth J. Safir (eds.), *The Null Subject Parameter*. *Linguistische Berichte* 153, 372-385.

Haider, Hubert. 1995. Downright down to the right. In *On Extraction and Extraposition in German*, U. Lutz and J. Pafel (eds.), 245-271. Amsterdam: John Benjamins.

Haider, Hubert. 1997. Extraposition. In *Rightward Movement*, D. Beerman, D. LeBlanc, and H. van Riemsdijk (eds.), 115-151. Amsterdam: John Benjamins.

Haider, Hubert. 2000a. OV is more basic than VO. In *The derivation of VO and OV*, P. Svenonius (ed.), 45-67. Amsterdam: John Benjamins.

Haider, Hubert. 2000b. Adverb placement – convergence of structure and licensing. In *Adverbs and Adjunction*, A. Alexiadou and P. Svenonius (eds.), 50-77. Potsdam: Universitätsverlag Potsdam.

Haider, Hubert. 2000c. Scrambling – what's the state of the art? In *The Acquisition of Scrambling and Cliticization* S.M. Powers and C. Harman (eds.), 19-40. Dordrecht: Kluwer.

Haiman, John and Paola Benincà. 1992. *The Rhaeto-Romance Languages*. London and New York: Routledge.

Hale, Mark. 1996. Theory and method in historical linguistics. Ms., Concordia University, Montreal.

Hale, Mark. 1998. Diachronic syntax. *Syntax* 1.1: 1-18.

Hale, Mark. 2003. Neogrammarian sound change. In *The Handbook of Historical Linguistics*, B. Joseph and R. D. Janda (eds.), 343-368. Oxford: Blackwell.

Hale, Mark. 2007. *Historical Linguistics. Theory and Method*. Oxford & Malden: Blackwell.

Halle, Morris. 1990. An approach to morphology. *Proceedings of the North East Linguistics Society* 20, Vol. 1, 150-184. Amherst: GLSA, University of Amherst, Massachusetts.

Halle, Morris. 1997. Distributed Morphology: Impoverishment and Fission. In *MIT Working Papers in Linguistics* 30: *PF: Papers At the Interface*, B. Bruening, Y. Kang, and M. McGinnis (eds.), 425-450. Cambridge, Mass.: Department of Linguistics and Philosophy, MIT.

Halle, Morris and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In *The View from Building* 20, S.J. Keyser and K. Hale (eds.), 111-176. Cambridge: The MIT Press.

Halle, Morris and Alec Marantz. 1994. Some key features of Distributed Morphology. In *MIT Working Papers in Linguistics* 21: *Papers on Phonology and Morphology*, A. Carnie, H. Harley and T. Bures (eds.), 275-288. Cambridge, Mass.: Department of Linguistics and Philosophy, MIT.

Halpern, Aaron. 1992. *Topics in the placement and morphology of clitics*. Doctoral dissertation, Stanford University.

Harbour, Daniel. 2003. The Kiowa Case for Feature Insertion. *Natural Language and Linguistic Theory* 21: 543-578.

Harley, Heidi. 1994. Hug a tree: Deriving the morphosyntactic feature hierarchy. In *MIT Working Papers in Linguistics* 21: *Papers on phonology and morphology*, A. Carnie and H. Harley (eds.), 275-288. Cambridge, Mass: Department of Linguistics and Philosophy, MIT.

Harley, Heidi. 2007. *One-replacement, unaccusativity, actegorial roots, and Bare Phrase Structure*. Ms., Harvard University/University of Arizona.

Harley, Heidi and Rolf Noyer. 1999. State-of-the-article: Distributed Morphology. *Glot International* 4.4, 3-9.

Harley, Heidi and Elizabeth Ritter. 2002. Structuring the bundle: A universal morphosyntactic feature geometry. In *Pronouns – Grammar and Representation*, H. J. Simon and H. Wiese (eds.), 23-39. Amsterdam: John Benjamins.

Harnisch, Rüdiger. 1989. Die sogenannte 'sogenannte Flexion der Konjunktionen'. Ein Paradigma aus der Bavaria thuringica. In *Bayerisch-österreichische Dialektforschung*, E. Koller, W. Wegstein, and N. R. Wolf (eds.), 283-290. Würzburg.

Harris, Alice and Lyle Campbell. 1995. *Historical Syntax in Cross-linguistic Perspective*. Cambridge: Cambridge University Press.

Harris, Martin. 1978. *The Evolution of French Syntax. A Comparative Approach*. London: Longman.

Hawkins, John. 1990. A parsing theory of word order universals. *Linguistic Inquiry* 21, 223-261.

Heine, Bernd. 2008. Are there any generalizations on grammatical change in language contact? Paper presented at the annual meeting of the DGfS, University of Bamberg.

Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd and Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.

Heine, Bernd, Ulrike Claudi and Friederike Hünnemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.

Hoekstra, Eric and Caroline Smits. 1999. Everything you always wanted to know about complementizer agreement. In E. van Gelderen and V. Samiian (eds.), *Proceedings of WECOL 1998*. Fresno, CA: California State University Press.

Hoekstra, Jarich and Laszlo Marácz. 1989. On the position of inflection in West Germanic. *Working Papers in Scandinavian Syntax* 44, 75-88.

Holmberg, Anders. 1986 *Word Order and Syntactic Features in the Scandinavian Languages and English*. Doctoral dissertation. University of Stockholm.

Holmberg, Anders. 1999. Remarks on Holmberg's Generalization. *Studia Linguistica* 53: 1-39.

Holmberg, Anders. 2000. Deriving OV Order in Finnish. In *The Derivation of VO and OV*, P. Svenonius (ed.), 123-152. Amsterdam: Benjamins.

Holmberg, Anders. 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36.4, 533-564.

Holmberg, Anders and Urpo Nikanne. 2006. Subject doubling in Finnish: The role of deficient pronouns. http://www.meertens.knaw.nl/projecten/edisyn/Online_proceedings/Paper_Holmberg-Nikanne.pdf

Holmberg, Anders and Christer Platzack. 1995. *The Role of Inflection in Scandinavian Syntax*. Oxford: Oxford University Press.

Hopper, Paul. 1975. *The Syntax of the Simple Sentence In Proto-Germanic*. The Hague: Mouton.

Hopper, Paul. 1979. Aspect and foregrounding in discourse. In Talmy Givón (ed.), *Syntax and Semantics 12: Discourse and Syntax*, 213-241. New York: Academic Press.

Hopper, Paul and Elizabeth Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

Hornstein, Norbert. 1990. *As Time Goes by*. Cambridge, Mass.: MIT Press.

Hróarsdóttir, Thorbjörg. 2000. *Word Order Change in Icelandic. From OV to VO*. Amsterdam: John Benjamins.

Hinterhölzl, Roland. 2001. Event-related adjuncts and the OV/VO Distinction. *WCCFL 20 Proceedings*, 276-289. Somerville: Cascadilla Press.

Hinterhölzl, Roland. 2002. Parametric variation and scrambling in English. In *Studies in Comparative Germanic Syntax*, W. Abraham and J.W. Zwart (eds.), 131-150. Amsterdam and Philadelphia: John Benjamins.

Hinterhölzl, Roland. 2004. Language change versus grammar change: What diachronic data reveal about the distinction between core grammar and periphery. In *Diachronic Clues to Synchronic Grammar*, E. Fuß and C. Trips (eds.), 131-160. Amsterdam: John Benjamins.

Hinterhölzl, Roland, Svetlana Petrova, and Michael Solf. 2005. Diskurspragmatische Faktoren für Topikalität und Verbstellung in der althochdeutschen Tatianübersetzung (9.Jh.). *Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure (ISIS)* 3. S. Ishihara, M. Schmitz and A. Schwarz (eds.), 143-182. Potsdam: Universitätsverlag Potsdam.

Hirt, Hermann. 1929. *Indogermanische Grammatik. Teil V: Der Akzent*. Heidelberg: Carl Winter.

Hirt, Hermann. 1934. *Handbuch des Urgermanischen. Teil 3: Abriss der Syntax*. Heidelberg: Winter.

Hulk, Aafke and Ans van Kemenade. 1995. Verb second, pro-drop, functional projections and language change. In Adrian Battye and Ian Roberts (eds.), *Clause structure and language change*. Oxford: Oxford University Press, 227-256.

Hyams, Nina. 1986. *Language Acquisition and the Theory of Parameters*. Dordrecht: Reidel.

Hyams, Nina, Dimitris Ntelitheos, and Cecile Manorohanta. 2006. Acquisition of the Malagasy Voicing System: Implications for the adult grammar. Ms., UCLA.

Inaba, Jiro. 2007. *Die Syntax der Satzkomplementierung. Zur Struktur des Nachfeldes im Deutschen*. Berlin: Akademie Verlag.

Ingham, Richard. Negated subject and object positions in 15th century non-literary English. *Language Variation and Change* 14, 291-322.

Jackendoff, Ray and Peter Culicover. 2005. *Simpler Syntax*. Oxford: Oxford University Press.

Jaeggli, O. and K. Safir. 1989. The null subject parameter and parametric theory. In Jaeggli, O. and K. Safir (eds.), *The Null Subject Parameter*. Dordrecht: Kluwer, 1-44.

Jakobson, Roman 1936 [1971]. Beitrag zur allgemeinen Kasuslehre: Gesamtbedeutung der russischen Kasus. Reprinted in: *Selected Writings* (vol. II: *Word and Language*). The Hague: Mouton (originally in *Travaux du Cercle Linguistique de Prague* VI, 240-299).

Janda, Richard D. and Brian D. Joseph. 2003. Introduction. In *The Handbook of Historical Linguistics*, B. D. Joseph and R. D. Janda (eds.), 1-180. Oxford & Malden: Blackwell.

Jayaseelan, K.A. 2001. IP-internal topic and focus positions. *Studia Linguistica* 55.1, 39-75.

Julien, Marit 2002. *Syntactic Heads and Word Formation*. Oxford: Oxford University Press.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.

Kayne, Richard. 1998. Overt vs. covert movement. *Syntax* 1, 128-191.

Keenan, Edward. 1976. Remarkable subjects in Malagasy. In: Charles N. Li (ed.), *Subject and Topic*, 249-301. New York: Academic Press.

Keenan, Edward. 2002. Explaining the creation of reflexive pronouns in English. In *Studies in the History of English: A Millennial Perspective*, D. Minkova and R. Stockwell (eds.), 325-355. Berlin: Mouton de Gruyter.

Kemenade, Ans van. 1987. *Syntactic case and morphological case in the history of English*. Dordrecht: Foris.

Kemenade, Ans van. 1997. V2 and embedded topicalization in Old and Middle English. In Ans van Kemenade, and Nigel Vincent (eds.), *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 326-351.

Kemenade, Ans van. 1999. Sentential negation and clause structure in Old English. In Ingrid Tieken-Boon van Ostade, Gunnel Tottie, and Wim van der Wurff (eds.), *Negation in the History of English*. Berlin and New York: Mouton de Gruyter, 147-165.

Kemenade, Ans van. 2002. Word order in Old English prose and poetry: the position of finite verbs and adverbs. In *Studies in the History of the English Language: A Millennial Perspective*, D. Minkova and R. Stockwell (eds.), 355-373. Berlin: Mouton de Gruyter.

Kemenade, Ans van and Bettelou Los. 2006. Discourse adverbs and clausal syntax in Old and Middle English. In *The Handbook of the History of English*, A. van Kemenade and B. Los (eds.), 224-248. Oxford: Blackwell.

Kemenade, Ans van and Tania Milicev. 2005. Syntax and discourse in Old and Middle English word order. To appear in *Articles from the 8th Diachronic Generative Syntax Conference*, S. Anderson and D. Jonas (eds.), Oxford University Press.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Oxford: Blackwell.

Kiparsky, Paul. 1973. 'Elsewhere' in phonology. In *A Festschrift for Morris Halle*, S. Anderson and P. Kiparsky (eds.), 93-106. New York: Holt, Rinehart and Winston.

Kiparsky, Paul. 1982. Word-formation and the lexicon. In *Proceedings of the 1982 Mid-America Linguistics Conference*, F. Ingemann (ed.), 2-29. University of Kansas.

Kiparsky, Paul. 1995. Indo-European origins of Germanic syntax. In *Clause Structure and Language Change*, A. Battye and I. Roberts (eds.), 140-169. Oxford: Oxford University Press.

Kiparsky, Paul. 1996. The shift to head-initial VP in Germanic. In *Studies in Comparative Germanic syntax II*, H. Thráinsson, S.D. Epstein and S. Peter (eds.), 140-179. Dordrecht: Kluwer.

Kiparsky, Paul. 1997. The rise of positional licensing. In *Parameters of Morphosyntactic Change*, A. van Kemenade and N. Vincent (eds.), 460-493. Cambridge: Cambridge University Press.

Kiparsky, Paul. 2003. The phonological basis of sound change. In: Joseph, Brian D. and Richard D. Janda (eds.), *The Handbook of Historical Linguistics*, 313-342. Oxford: Blackwell.

Kitahara, Hisatsugu. 1997. *Elementary Operations and Optimal Derivations*. Cambridge, Mass.: MIT Press.

Klein, Jared. 1994. Gothic *þaruh*, *þanuh* and -(u) *þan*. *Indogermanische Forschungen* 99, 253-276.

Klein, Jared and Nancy Condon. 1993. Gothic -(u)h: a synchronic and comparative study. *Transactions of the Philological Society* 91, 1-62.

Koeneman, Olaf. 2007. Deriving the difference between full and partial pro-drop. In: Ackema, Peter, Patrick Brandt, Maaike Schoorlemmer and Fred Weerman (eds.), *Arguments and Agreement*, 76-100. Oxford: Oxford University Press.

Koizumi, Masatoshi. 1995. *Phrase Structure in Minimalist Syntax*, Doctoral dissertation, MIT.

Kollmer, Michael. 1987. *Die schöne Waldersprach*. Bd. I-III. Prackenbach: Kollmer (Eigenverlag).

Kolmer, Agnes. 2005. Subjektklitika als Kongruenzmarkierer: Ein Vergleich zwischen bairischen und alemannischen Sprachinseldialekten in Norditalien (Zimbrisch und Walserdeutsch). In: Bidese, Ermenegildo/Dow, James R./Stolz, Thomas (eds.), *Das Zimbrische zwischen Germanisch und Romanisch*. (Diversitas Linguarum 9), 164-189. Bochum: Brockmeyer.

Koopman, Hilda. 1984. *The Syntax of Verbs: from Verb Movement rules in the Kru Languages to Universal Grammar*. Dordrecht: Foris.

Koopman, Willem. 1991. The order of dative and accusative objects in Old English and scrambling. *Studia Anglica Posnaniensia* 25-27, 109-121.

Koopman, Willem. 1997. Another look at clitics in Old English. *Transactions of the Philological Society* 95.1, 73-93.

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

Koppen, Marjo van. 2005. *One Probe – Two Goals: Aspects of Agreement in Dutch Dialects*. Utrecht: LOT.

Kornfilt, Jaklin. 1990. Turkish and the Turkic languages. In *The Major Languages of Eastern Europe*, B. Comrie (ed.), 227-252. London: Routledge.

Kouwenberg, Silvia. 1990. Complementizer PA, the finiteness of its complements and some remarks on empty categories in Papiamentu. *Journal of Pidgin and Creole Languages* 5.1, 39-51.

Kouwenberg, Silvia and Pieter Muysken. 1995. Papiamento. In J. Arends, P. Muysken, and N. Smith (eds.), *Pidgeons and Creoles. An Introduction*, 205-218. Amsterdam: Benjamins.

Kratzer, Angelika 1996. Severing the external argument from its verb. In *Phrase Structure and the Lexicon*, J. Rooryck and L. Zaring (eds.), 109-137. Dordrecht: Kluwer.

Kremers, Joost. 2003. *The Noun Phrase in Arabic: A Minimalist Approach*, PhD thesis, University of Nijmegen. LOT Dissertation Series 79.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Journal of Language Variation and Change* 1.3, 199-244.

Kroch, Anthony. 1994. Morphosyntactic variation. In *Proceedings of the Thirthieth Annual Meeting of the Chicago Linguistics Society*. Vol. II, K. Beals, J. Denton, B. Knippen, L. Melnar, H. Suzuki, and E. Zeinfeld (eds.), 180-201. Chicago: Chicago Linguistics Society.

Kroch, Anthony. 2001. Syntactic change. In *The Handbook of Contemporary Syntactic Theory*, M. Baltin and C. Collins (eds.), 699-729. Oxford: Blackwell.

Kroch, Anthony and Ann Taylor. 1997. Verb movement in Old and Middle English: Dialect variation and language contact. In *Parameters of Morphosyntactic Change*, A. van Kemenade and N. Vincent (eds.), 297-325. Cambridge: Cambridge University Press.

Kroch, Anthony and Ann Taylor. 2000. Verb-complement order in Middle English. In: S. Pintzuk, G. Tsoulas, and A. Warner (eds.), *Diachronic Syntax: Models and Mechanisms*, 132-163. Oxford: Oxford University Press.

Kroeger, Paul. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. Stanford: CSLI Publications.

Kuen, Heinrich. 1957. Die Gewohnheit der mehrfachen Bezeichnung des Subjekts in der Romania und die Gründe ihres Aufkommens. In *Syntactica und Stilistica. Festschrift für Ernst Gamillscheg zum 70. Geburtstag*, G. Reichenkron (ed.), 293-326. Tübingen: Niemeyer.

Kuhn, Hans. 1933. Zur Wortstellung und -betonung im Altgermanischen. Reprinted 1969, *Kleine Schriften*, D. Hoffmann (ed.), Vol. 1, 18-103. Berlin: de Gruyter.

Kural, Murat. 1997. Postverbal constituents in Turkish and the Linear Correspondence Axiom. *Linguistic Inquiry* 28.3, 498-519.

Kuroda, S. Y. 1988. Whether We Agree or Not: A Comparative Syntax of English and Japanese. In *Papers from the Second International Workshop on Japanese Syntax*. W. J. Poser (ed.), 103-143. Stanford: Center for Study of Language and Information.

Labov, William. 1978. On the use of the present to explain the past. In *Readings in Historical Phonology*, P. Baldi and R. Werth (eds.), 275-312. University Park: Pennsylvania State University Press.

Labov, William. 1994. *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell.

Lahne, Antje. 2007. A multiple specifier approach to left peripheral architecture. Ms., University of Leipzig. To appear in *Linguistic Analysis*.

Lambrecht, Knud. 1981. *Topic, antitopic and verb-agreement in non-standard French*. Pragmatics and beyond vol. II: 8. Amsterdam: John Benjamins.

LaPolla, Randy. 1992. On the dating and nature of verb agreement in Tibeto Burman. *Bulletin of the School of Oriental and African Studies* 55.2, 298-315.

Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19, 335-391.

Lasnik, Howard. 1999. *Minimalist Analysis*. Malden: Blackwell.

Lasnik, Howard. 2000. *Syntactic Structures Revisited. Contemporary Lectures on Classic Transformational Theory*. Cambridge, Mass.: MIT Press.

Lasnik, Howard. 2001. When can you save a structure by destroying it? In Minjoo Kim and Uri Strauss (eds.), *Proceedings of the North East Linguistic Society* 31, 301-320. Amherst, Mass.: Graduate Linguistics Students Association.

Lawson, Richard H. 1980. Paratactic 'thô' in Old High German 'Tatian'. *Neuphilologische Mitteilungen* 81: 99-104.

Lechner, Winfried. 2001. Reduced and phrasal comparatives. *Natural Language and Linguistic Theory* 19, 683-735.

Legate, Julie. 1999. The morphosyntax of Irish agreement. In *MIT Working Papers in Linguistics* 33, K. Arregi, B. Bruening, C. Krause and V. Lin (eds.), 219-240. MIT: Department of Linguistics and Philosophy.

Lehmann, Christian. 1984. *Der Relativsatz. Typologie seiner Strukturen - Theorie seiner Funktionen - Kompendium seiner Grammatik*. Tübingen: Narr

Lehmann, Christian. 1988. On the function of agreement. In *Agreement in Natural Language*, M. Barlow and Charles A. Ferguson (eds.), 55-65. Stanford: CSLI.

Lehmann, Christian. 2002. *Thoughts on Grammaticalization*. Second edition. Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt.

Lenerz, Jürgen. 1984. *Syntaktischer Wandel und Grammatiktheorie. Eine Untersuchung an Beispielen aus der Sprachgeschichte des Deutschen*. Tübingen: Niemeyer.

Lenerz, Jürgen. 1985a. Diachronic Syntax: Verb Position and COMP in German. In *Studies in German Grammar*, J. Toman (ed.), 103-132. Dordrecht: Foris.

Lenerz, Jürgen. 1985b. Zur Theorie syntaktischen Wandels: Das expletive *es* in der Geschichte des Deutschen. In *Erklärende Syntax des Deutschen*, W. Abraham (ed.), 99-136. Tübingen: Narr.

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

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

Lightfoot, David. 1997. Catastrophic change and learning theory. *Lingua* 100, 171-192.

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

Lightfoot, David. 2003. Grammatical approaches to syntactic change. In *The Handbook of Historical Linguistics*, B. Joseph and R. Janda (eds.), 495-508. Oxford & Malden, Mass.: Blackwell.

Linder, Karl Peter. 1987. *Grammatische Untersuchungen zur Charakteristik des Rätoromanischen in Graubünden*. Tübingen: Gunter Narr.

Lippert, Jörg. 1974. *Beiträge zur Technik und Syntax althochdeutscher Übersetzungen*. München: Wilhelm Fink.

Lipski, John. 2001. Null subjects in (Romance-derived) creoles: Routes of evolution. Ms., University of New Mexico.

Lobeck, Anne. 1995. *Ellipsis: Functional Heads, Licensing, and Identification*. Oxford: Oxford University Press.

Lockwood, W.B. 1968. *Historical German Syntax*. Oxford: Oxford University Press.

Longobardi, Giuseppe. 1994. La posizione del verbo gotico. *Miscellanea di studi linguistici in onore di Walter Belardi*. Rome.

Longobardi, Giuseppe. 2001. Formal syntax, diachronic minimalism, and etymology: The history of French *chez*. *Linguistic Inquiry* 32.2: 275-302.

Lüdtke, Helmut. 1980. Auf dem Wege zu einer Theorie des Sprachwandels. In *Kommunikationstheoretische Grundlagen des Sprachwandels*, H. Lüdtke (ed.), 182-252. Berlin: de Gruyter.

Lühr, Rosemarie. 1984. Reste der athematischen Konjugation in den germanischen Sprachen. In *Das Germanische und die Rekonstruktion der Indogermanischen Grundsprache*, J. Untermann and B. Brogyanyi (eds.), 25-90. Amsterdam: Benjamins.

MacKenzie, D. N. 1990. Pashto. In *The Major Languages of Southeast Asia, the Middle East and Africa*, B. Comrie (ed.), 132-150. London: Routledge.

Marantz, Alec. 1992. Case and licensing. In: *Proceedings of ESCOL 1991*, 234-253.

Marantz, Alec. 1984. *On the Nature of Grammatical Relations*. Cambridge, Mass.: MIT Press.

Marantz, Alec. 1988. Clitics, morphological merger, and the mapping to phonological structure. In *Theoretical Morphology: Approaches in Modern Linguistics*, M. Hammond and M. Noonan (eds.), 253-270. San Diego: Academic Press.

Marantz, Alec. 1995. 'Cat' as a phrasal idiom: Consequences of late insertion in Distributed Morphology. Ms., MIT.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own Lexicon. In *Proceedings of the 21st Annual Penn Linguistics Colloquium. Penn Working Papers in Linguistics 4.2*, A. Dimitriadis, L. Siegel, C. Surek-Clark and A. Williams. (eds.), 201-225. University of Pennsylvania, Philadelphia.

Masser, Achim. 1994. *Die lateinisch-deutsche Tatianbilingue des Cod. Sang. 56*. Göttingen: Vandenhoeck and Ruprecht.

Masser, Achim. 1997. Syntaxprobleme im althochdeutschen Tatian. In Yvon Desportes (ed.), *Semantik der syntaktischen Beziehungen. Akten des Pariser Kolloquiums zur Erforschung des Althochdeutschen*, 123-140. Heidelberg: C. Winter.

Matthews, Peter H. 1991. *Morphology* (2nd edition). Cambridge: Cambridge University Press.

Maurer, Friedrich. 1924. Zur Anfangsstellung des Verbs im Deutschen. In *Beiträge zur germanischen Sprachwissenschaft. Festschrift für O. Behaghel*, W. Horn (ed.), 141-184. Heidelberg: Winter.

Mayerthaler, Willi. 1980. Aspekte der Analogietheorie. In *Kommunikationstheoretische Grundlagen des Sprachwandels*, H. Lüdtke (ed.), 80-130. Berlin: de Gruyter.

McCawley, James. 1998. *The Syntactic Phenomena of English*. 2nd edition. Chicago: Chicago University Press.

McGinn, Richard. 1985. A principle of text coherence in Indonesian languages. *The Journal of Asian Studies* 44.4, 743-753.

McGinn, Richard. 1988. Government and case in Tagalog. In R. McGinn (ed.), *Studies in Austronesian Linguistics*, 275-293. Athens, Ohio: Center for Southeast Asia Studies.

McKnight, George H. 1897. The primitive Teutonic order of words. *Journal of English and Germanic Philology* 1, 136-219.

Meillet, Antoine. 1912. L'évolution des formes grammaticales. *Scientia (Rivista di Scienza)* 12, No. 26.6. Reprinted in A. Meillet 1958, *Linguistique historique et linguistique générale*, 130-148. Paris: Champion.

Merchant, Jason. 2001. *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford: Oxford University Press.

Mitchell, Bruce. 1985. *Old English Syntax*. Oxford: Clarendon.

Mitchell, Bruce and Fred C. Robinson. 1988). *A Guide to Old English*. Oxford: Blackwell.

Mithun, Marianne. 1991. The development of bound pronominal paradigms. In *Language Typology 1988: Typological Models in Reconstruction*, W. P. Lehmann and H.-J. Jakusz Hewitt (eds.), 85-104. Amsterdam: John Benjamins.

Moulton, William. 1944. Scribe γ of the Old High German Tatian translation. *PMLA* 59:2, 307-334.

Müller, Friedrich Max. 1875. *Vorlesungen über die Wissenschaft der Sprache*, Bd. 1. Leipzig: Julius Klinkhardt.

Müller, Gereon. 1995. On extraposition and successive cyclicity. In *On Extraction and Extrapolation in German*, U. Lutz and J. Pafel (eds.), 213-243. Amsterdam: John Benjamins.

Müller, Gereon. 1997. Extrapolation as remnant movement. In *Rightward Movement*, D. Beerman, D. LeBlanc, and H. van Riemsdijk (eds.), 215-246. Amsterdam: John Benjamins.

Müller, Gereon. 2004. Verb-Second as vP-First. *Journal of Comparative Germanic Linguistics* 7.3, 179-234.

Müller, Gereon 2006a. Subanalyse verbaler Flexionsmarker. In *Grammatische Untersuchungen, Analysen und Reflexionen. Festschrift für Gisela Zifonun*, E. Breindl, L. Gunkel, and B. Strecker (eds.), 183-204. Tübingen: Narr.

Müller, Gereon. 2006b. Pro-drop and Impoverishment. In: Patrick Brandt and Eric Fuß (eds.), *Form, Structure, and Grammar. A Festschrift Presented to Günther Grewendorf on Occasion of His 60th Birthday*, 93-115. Berlin: Akademie Verlag.

Müller, Gereon. 2007. Toward a relativized concept of cyclic linearization. Ms., University of Leipzig.

Müller, Stefan. 2005. Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte* 203: 297-330.

Muysken, Pieter. 1981. Creole tense/mood/aspect systems: The unmarked case. In P. Muysken (ed.), *Generative Studies on Creole Languages*, 181-199. Dordrecht: Foris.

Näf, Anton. 1979. *Die Wortstellung in Notkers Consolatio. Untersuchungen zur Syntax und Übersetzungstechnik*. Berlin: de Gruyter

Neeleman, Ad and Kriszta Szendrői. 2007. Radical pro-drop and the morphology of pronouns. *Linguistic Inquiry* 38.4, 671-714.

Neidle, C., D. MacLaughlin, R.G. Lee, B. Bahan, and J. Kegl. 1997. Rightward *wh*-movement in American Sign Language. In *Rightward Movement*, D. Beerman, D. LeBlanc, and H. van Riemsdijk (eds.), 247-278. Amsterdam: John Benjamins.

Nespor, Marina and Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.

Nespor, Marina, Maria Teresa Guasti, and Anne Christophe. 1996. Selecting word order: the Rhythmic Activation Principle. In Ursula Kleinhenz (ed.), *Interfaces in Phonology*, 1-26. Berlin: Akademie Verlag.

Newmeyer, Frederick. 1998. *Language Form and Language Function*. Cambridge, Mass.: MIT Press.

Nissenbaum, Jonathan W. 2000. *Investigations of Covert Phrase Movement*. Doctoral dissertation, MIT.

Niyogi, Partha and Robert C. Berwick. 1998. The logical problem of language change: A case study of European Portuguese. *Syntax* 1.2, 192-205.

Nordlinger, Rachel. 1998. *A Grammar of Wambaya, Northern Territory (Australia)*. (Pacific Linguistics C-140). Canberra: ANU.

Noyer, Rolf. 1997. *Features, Positions, and Affixes in Autonomous Morphological Structure*. New York: Garland.

Nübling, Damaris. 1992. *Klitika im Deutschen. Schriftsprache, Umgangssprache, alemannische Dialekte*. Tübingen: Gunter Narr.

Nunes, Jairo. 1999. Linearization of chains and phonetic realization of chain links. In *Working Minimalism*, S.D. Epstein and N. Hornstein (eds.), 217-249. Cambridge, Mass.: The MIT Press.

Nunes, Jairo. 2002. VO or OV? That's the underlying question: commentary on Pintzuk. In *Syntactic Effects of Morphological Change*, D. Lightfoot (ed.), 300-306. Oxford: Oxford University Press.

O'Grady, William. 1997. *Syntactic Development*. Chicago: Chicago University Press.

Ohala, Diane. 1996. Cluster reduction and constraints in acquisition. Doctoral dissertation, University of Arizona.

Ohala, Diane. 1999. The influence of sonority on children's cluster reductions. *Journal of Communication Disorders* 32, 397-422.

Ohala, John J. 2003. Phonetics and historical phonology. In *The Handbook of Historical Linguistics*, B. Joseph and R. D. Janda (eds.), 669-686. Oxford: Blackwell.

Önnerfors, Olaf. 1997. *Verb-erst-Deklarativsätze. Grammatik und Pragmatik*. Stockholm: Almqvist & Wiskell.

Ortiz de Urbina, Jon. 1989. *Parameters in the grammar of Basque : a GB approach to Basque syntax*. Dordrecht: Foris.

Ouhalla, Jamal. 1991. *Functional Projections and Parametric Variation*. London: Routledge.

Partee, Barbara. 1973. Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy* 70.18, 601-609.

Pater, Joe and Jessica Barlow. 2003. Constraint conflict in cluster reduction. *Journal of Child Language* 30, 487-526.

Paul, Hermann. 1879. Beitraege zur Geschichte der Lautentwicklung und Formenassociation. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 6, 538-560.

Paul, Hermann. 1880. *Prinzipien der Sprachgeschichte*. (5th edition, published 1920, *Konzepte der Sprach- und Literaturwissenschaft*, 6). Tübingen: Niemeyer.

Paul, Hermann. 1952. *Deutsche Grammatik II.3: Flexionslehre*. Halle: Max Niemeyer.

Paul, Hermann. 1998. *Mittelhochdeutsche Grammatik*. Tübingen: Niemeyer.

Partee, Barbara, Alice ter Meulen, and Robert Wall. 1993. *Mathematical Methods in Linguistics*. Dordrecht: Kluwer.

Pearson, Matthew. 2005. The Malagasy subject/topic as an A'-element. *Natural Language and Linguistic Theory* 23, 381-457.

Pesetsky, David. 1995. *Zero Syntax. Experiencers and Cascades*. Cambridge, Mass.: The MIT Press.

Pesetsky, David and Esther Torrego. 2001. T-to-C movement: causes and consequences. In: Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, 355-426. Cambridge, Mass.: MIT Press.

Petrova, Svetlana. 2006. A discourse-based approach to verb placement in Early West germanic. *Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure (ISIS)* 5. S. Ishihara, M. Schmitz and A. Schwarz (eds.), 1-32. Potsdam: Universitätsverlag Potsdam.

Petrova, Svetlana and Michael Solf 2007. Syntaktischer Wandel und Satzmodus. Beobachtungen zur Wortstellung in direkten Fragesätzen des Althochdeutschen. To appear in *Beiträge zur Geschichte der deutschen Sprache und Literatur*.

Petronio, K. and D. Lillo-Martin. 1997. *Wh-movement and the position of SpecCP: Evidence from American Sign Language*. *Language* 73: 18-57.

Pfalz, Anton. 1918. Suffigierung der Personalpronomina im Donaubairischen. Republished 1983 in *Die Wiener dialektologische Schule. Grundsätzliche Studien aus 70 Jahren Forschung*. (Wiener Arbeiten zur germanischen Altertumskunde und Philologie 23), P. Wiesinger (ed.), 217-235. Wien: Peter Lang.

Phillips, Colin. 1996. *Order and Structure*. Doctoral dissertation, MIT.

Phillips, Colin. 2003. Linear Order and Constituency. *Linguistic Inquiry* 34.1, 37-90.

Pintzuk, Susan. 1991. *Phrase Structures in Competition: Variation and Change in Old English Word Order*. Ph.D. dissertation, University of Pennsylvania.

Pintzuk, Susan. 1993. Verb-seconding in Old English: Movement to Infl. *The Linguistic Review* 10, 5-35.

Pintzuk, Susan. 1999. *Phrase Structures in Competition: Variation and Change in Old English Word Order*. New York: Garland.

Pintzuk, Susan. 2002. Verb-object order in Old English: Variation as grammatical competition. In: D. Lightfoot (ed.), *Syntactic Effects of Morphological Change*, 276-299. Oxford: Oxford University Press.

Pintzuk, Susan. 2005. Arguments against a Universal Base: Evidence from Old English. *English Language and Linguistics* 9.1, 115-138.

Pintzuk, Susan and Ann Taylor. 2006. The loss of OV order in the history of English. In *Handbook on the History of English*, Kemenade, A. van and B. Los (eds.), 249-278. Oxford: Blackwell.

Poletto, Cecilia. 1995. The diachronic development of subject clitics in North-Eastern Italian dialects. In *Clause Structure and Language Change*, A. Battye and I. Roberts (eds.), 295-324. Oxford: Oxford University Press.

Poletto, Cecilia. 1997. Pronominal syntax. In *The Dialects of Italy*, M. Maiden and M. Parry (eds.), 137-145. Cambridge: Cambridge University Press.

Pollock, Jean-Yves. 1989. Verb movement, UG and the structure of IP. *Linguistic Inquiry* 20.3: 365-424.

Poppe, Nicholas. 1960. *Buriat Grammar*. Bloomington: Indiana University Publications.

Postal, Paul. 1969. On so-called 'pronouns' in English. In *Modern Studies in English*, D. Reibel and S. Schane (eds.), 201-224. Englewood Cliffs: Prentice-Hall.

Prasada, Sandeep and Steven Pinker. 1993. Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes* 8, 1-56.

Radford, Andrew. 2004. *Minimalist Syntax. Exploring the Structure of English*. Cambridge: Cambridge University Press.

Ramat, Anna Giacalone and Paul Hopper (eds.). 1998. *The Limits of Grammaticalization*. Amsterdam: John Benjamins.

Rauch, Irmengard. 1992. *Old Saxon Language: Grammar, Epic Narrative, Linguistic Interference*. Frankfurt: Peter Lang.

Reinhart, Tanya and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24, 657-720.

Reis, Hans. 1901. Über althochdeutsche Wortfolge. *Zeitschrift für deutsche Philologie* 33, 212, 238, 330-349.

Renzi, Lorenzo. 1992. I pronomi soggetto in due varietà substandard: fiorentino e français avancé. *Zeitschrift für Romanische Philologie* 108, 72-98.

Renzi, Lorenzo and Laura Vanelli. 1983. I pronomi soggetto in alcune varietà romanze (Subject pronouns in some Romance varieties). In *Scritti linguistici in onore die Giovan Battista Pellegrini (Linguistic Papers in Honor of G. B. Pellegrini)*, P. Benincà et al. (eds.), 121-145. Pisa: Pancini.

Reuland, Eric. 2001. Primitives of Binding. *Linguistic Inquiry* 32.3, 439-492.

Richards, Marc. 2004. *Object Shift and Scrambling in North and West Germanic: A Case Study in Symmetrical Syntax*. PhD thesis, University of Cambridge.

Richards, Marc. 2007. Dynamic linearization and the shape of phases. *Linguistic Analysis* 33.3-4, 209-237.

Richards, Norvin. 2001. *Movement in Language: Interactions and Architectures*. Oxford, New York: Oxford University Press.

Rizzi, Luigi. 1982. Issues in Italian Syntax. Dordrecht: Foris.

Rizzi, Luigi. 1986a. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17, 501-557.

Rizzi, Luigi. 1986b. On the status of subject clitics in Romance. In *Studies in Romance Linguistics*, O. Jaeggli and C. Silva-Corvalan (eds.), 391-419. Dordrecht: Foris.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, MA: The MIT Press.

Rizzi, Luigi. 1994. Early null subjects and root null subjects. In: Teun Hoekstra and Bonnie Schwartz (eds.), *Language Acquisition Studies in Generative Grammar*, 151-176. Amsterdam: John Benjamins.

Rizzi, Luigi. 1996. Residual Verb Second and the Wh-criterion. In Adriana Belletti and Luigi Rizzi (eds.), *Parameters and functional heads*. Oxford: Oxford University Press, 63-90.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar: Handbook in Generative Syntax*, 281-337. Dordrecht: Kluwer.

Rizzi, Luigi. 2001. On the position Int(errogative) in the left periphery of the clause. In G. Cinque and G. Salvi (eds.), *Current Studies in Italian Syntax: Essays Offered to Lorenzo Renzi*, 287-296. Amsterdam: North Holland.

Roberge, Yves. 1990. *The Syntactic Recoverability of Null Arguments*. Kingston and Montreal: McGill-Queen's University Press.

Roberts, Ian. 1993a. *Verbs and Diachronic Syntax. A Comparative History of English and French*. Dordrecht: Kluwer.

Roberts, Ian. 1993b. A formal account of grammaticalization in the history of Romance futures. *Folia Linguistica Historica* 13: 219-258.

Roberts, Ian. 1994. Second position effects and agreement in Comp. Ms., University of Wales, Bangor.

Roberts, Ian. 1996. Remarks on the Old English C-system and the diachrony of V2. In Ellen Brandner and Gisella Ferraresi (eds.), *Language change and Generative Grammar*. Opladen: Westdeutscher Verlag, 154-167.

Roberts, Ian. 1997. Directionality and word order change in the history of English. In *Parameters of Morphosyntactic Change*, A. van Kemenade and N. Vincent (eds.), 397-427. Cambridge: Cambridge University Press.

Roberts, Ian. 1998. *Have/be raising, Move F, and Procrastinate*. *Linguistic Inquiry* 29.1, 113-125.

Roberts, Ian. 1999. Verb movement and markedness. In M. DeGraff (ed.), *Language Creation and Language Change: Creolization, Diachrony and Development*, 287-329. Cambridge, Mass.: MIT Press.

Roberts, Ian. 2007a. *Diachronic Syntax*. Oxford: Oxford University Press.

Roberts, Ian. 2007b. A deletion analysis of null subjects: French as a case study. Ms., University of Cambridge.

Roberts, Ian and Anna Roussou. 1999. A formal approach to 'grammaticalization'. *Linguistics* 37: 1011-1041.

Roberts, Ian and Anna Roussou. 2001. The EPP as a condition on the Tense dependency. In *Subjects, Expletives, and the EPP*, Peter Svenonius (ed.), 125-155. Oxford: Oxford University Press.

Roberts, Ian and Anna Roussou. 2003. *Syntactic Change. A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge University Press.

Robinson, Orrin W. 1994. Verb-first position in the Old High German 'Isidor' translation. *Journal of English and Germanic Philology* 93, 356-373.

Robinson, Orrin W. 1997. *Clause Subordination and Verb Placement in the Old High German Isidor Translation*. Heidelberg: C. Winter.

Rögnvaldsson, Eiríkur. 1996. Word order variation in the VP in Old Icelandic. *Working Papers in Scandinavian Syntax* 58, 55-86

Rohrbacher, Bernhard. 1999. *Morphology-driven Syntax: A theory of V-to-I Raising and Pro-drop*. Amsterdam: Benjamins.

Rood, David S. 1973. Aspects of Subordination in Lakhota and Wichita. In *You Take the High Node and I'll Take the Low Node. Papers from the Comparative Syntax Festival*. C. Corum, T. C. Smith-Stark and A. Weiser (eds.), 71-88. Chicago: Chicago Linguistic Society.

Roßdeutscher, Antje. 2005a. 'On-line'-inferences in the semantics of *dann* and *then*. In *Proceedings of SuB9*, E. Maier, C. Bary and J. Huitink (eds.), 325–339. Nijmegen Center for Semantics. [Http://www.ru.nl/ncs/sub9](http://www.ru.nl/ncs/sub9).

Roßdeutscher, Antje. 2005b. 'On-line'-interpretation and 'cause' in the semantics of temporal *dann/then* and German *da*. Submitted to *Journal of Semantics*.

Roßdeutscher, Antje and Christiane von Stutterheim. 2006. Semantische und pragmatische Prinzipien der Positionierung von *dann*. *Linguistische Berichte* 205, 29–60.

Ross, John R. 1969. Guess who? In R. Binnick, A. Davison, G. Green, and J. Morgan (eds.), *Papers from the 5th regional meeting of the Chicago Linguistic Society*, 252-286. Chicago Linguistic Society: Chicago, IL. [R]

Sabel, Joachim. 1996. *Restrukturierung und Lokalität*. Berlin: Akademie Verlag.

Sabel, Joachim. 1998. *Principles and Parameters of Wh-movement*. Habilitation thesis, University of Frankfurt.

Sabel, Joachim. 2001. Deriving multiple head and phrasal movement: The Cluster Hypothesis. *Linguistic Inquiry* 32.3, 532-547.

Sabel, Joachim. 2003. Malagasy as an optional multiple wh-fronting language. In: Cedric Boeckx and Kleanthes K. Grohmann (eds.), *Multiple Wh-Fronting*, 229-254. Amsterdam: John Benjamins.

Saito, Mamoru. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1.1, 69-118.

Saito, Mamoru and Naoki Fukui. 1998. Order in phrase structure and movement. *Linguistic Inquiry* 29.3, 439-474.

Sapir, Edward. 1921. *Language*. New York: Harcourt, Brace and World. Smith, Carlotta. 1981. Semantic and syntactic constraints on temporal interpretation. In Philip. Tedeschi and Annie Zaenen (eds.), *Syntax and Semantics 14: Tense and Aspect*. New York: Academic Press, 213-237.

Schachter, Paul. 1976. The Subject in the Philippine Languages: Topic, Actor, Actor-Topic or None of the Above. In Charles N. Li (ed.), *Subject and Topic*, Academic Press, New York, 491-518.

Schachter, Paul. 1990. Tagalog. In B. Comrie (ed.), *The World's Major Languages*. Oxford: OUP, 936-958.

Schiffrin, Deborah. 1992. Anaphoric *then*: Aspectual, textual, and epistemic meaning. *Linguistics* 20, 753-792.

Schirmunski, V.M. 1962. *Deutsche Mundartkunde. Vergleichende Laut- und Formenlehre der deutschen Mundarten*. Berlin: Akademie Verlag.

Schlachter, Eva. 2004. Satzstruktur im Althochdeutschen. Eine Skizze zur Position des Verbs im Isidor-Traktat des 8. Jahrhunderts. In *Beiträge zu Sprache und Sprachen 4. Vorträge der Bochumer Linguistik-Tage*, : K. Pittner et al. (eds.), 179-188. München: Lincom.

Schlegel, August Wilhelm von. 1818. *Observations sur la langue et la littérature provençales*. Paris: Librairie grecque-latine-allemande.

Schrodt, Richard. 2004. *Althochdeutsche Grammatik II*. Tübingen: Niemeyer.

Schütze, Carson. 1994. Serbo-Croatian second position clitic placement and the phonology-syntax interface. In: Andrew Carnie, Heidi Harley and Tony Bures (eds.), *Papers on phonology and morphology: MITWPL 21*, 373-473. Cambridge, Mass.: Department of Linguistics and Philosophy, MIT.

Seidl, A. 2001. *Minimal Indirect Reference: A Theory of the Syntax-Phonology Interface*. London: Routledge.

Selkirk, Elisabeth. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, Mass.: MIT Press.

Selkirk, Elisabeth and Angelika Kratzer. 2005. Focuses, phases and phrase stress. Paper presented at the Mediterranean Syntax Meeting in Rhodes.

Shlonsky, Ur. 1994. Agreement in Comp. *The Linguistic Review* 11, 351-375.

Sievers, Eduard. 1893. *Altgermanische Metrik*. Halle: Niemeyer.

Sievers, Eduard (ed.). 1961. *Tatian*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Siewierska, Anna. 1999. From anaphoric pronoun to grammatical marker: why objects don't make it. *Folia Linguistica* 33.2, 225-251.

Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.

Simpson, Andrew. 2004. The EPP, fossilized movement and reanalysis. In *Diachronic Clues to Synchronic Grammar*, E. Fuß and C. Trips (eds.), 161-189. Amsterdam: John Benjamins.

Simpson, Andrew and Zoe Wu. 2002. Agreement, shells and focus. *Language* 78.2: 287-313.

Smith, Carlotta. 1981. Semantic and syntactic constraints on temporal interpretation. In Philip. Tedeschi and Annie Zaenen (eds.), *Syntax and Semantics 14: Tense and Aspect*. New York: Academic Press, 213-237.

Sommer, Thomas. 1994. *Flexionsmorphologie des Verbs im althochdeutschen Tatian*. München: Tuduv.

Spiess, Federico. 1956. *Die Verwendung des Subjektpersonalpronomens in den lombardischen Mundarten. (Romanica Helvetica 59)*. Bern: A. Francke.

Sprouse, Rex A. and Barbara S. Vance. 1999. An Explanation for the Decline of Null Pronouns in Certain Germanic and Romance Languages. In *Language Creation and Language Change: Creolization, Diachrony, and Development*. M. DeGraff. (ed.), 257-284. Cambridge, MA, MIT Press.

Stechow, Arnim von. 2005. LF in einem Phasenmodell: Bemerkungen anhand von Fischers Bindungstheorie. Paper presented at GGS 2005, University of Tübingen.

Steele, Susan. 1975. On some factors that affect and effect word order. In *Word Order and Word Order Change*, C. N. Li (ed.), 197-268. University of Texas Press, Austin.

Steinkrüger, Patrick. 2004. Philippine Spanish Creoles. In: Philipp Strazny (ed.), *Encyclopedia of Linguistics*. New York: Routledge.

Steinkrüger, Patrick. 2006. The puzzling case of Chabacano: Creolization, substrate, mixing and secondary contact. Paper presented at Tenth International Conference on Austronesian Linguistics. 17-20 January 2006. Puerto Princessa City, Palawan, Philippines. (<http://www.sil.org/asia/philippines/ical/papers.html>)

Sternfeld, Wolfgang. 2007. *Syntax. Eine morphologisch motivierte generative Beschreibung des Deutschen*. Tübingen: Stauffenberg.

Stockwell, Robert. 1977. Motivations for exbraciation in Old English. In *Mechanisms of Syntactic change*, Charles N. Li (ed.), 291-314. Austin: University of Texas Press.

Stockwell, Robert. 1984. On the history of the verb-second rule in English. In *Historical Syntax*, J. Fisiak (ed.), 575–592. Berlin: Mouton Publishers.

Stockwell, Robert and Donka Minkova. 1994. Kuhn's laws and the rise of verb-second syntax. In *Language Change and Language Structure: Older Germanic Languages in a Comparative Perspective*, T. Swan, E. Mørck, and O. Jansen Westvik (eds.) , 213-233. Berlin: Mouton de Gruyter.

Stowell, Tim. 1995. The phrase structure of tense. In: Johan Rooryck and Laurie Zaring (eds.), *Phrase Structure and the Lexicon*, 277-291. Dordrecht: Kluwer.

Stroik, Thomas. 1990. Adverbs as V-Sisters. *Linguistic Inquiry* 21.4, 654-661.

Stump, Gregory T. 1998. Inflection. In *The Handbook of Morphology*, A. Spencer and A. Zwicky (eds.), 13-43. Oxford and Malden, MA: Blackwell.

Svenonius, Peter. 1994. *Dependent Nexus: Subordinate Predication Structures in English and the Scandinavian Languages*. PhD dissertation, UCSC, Santa Cruz, Ca.

Svenonius, Peter. 1996. The optionality of particle shift. *Working Papers in Scandinavian Syntax* 57, 47-75.

Swan, Toril. 1994. Old English and Old Norse initial adverbials and word order. In Toril Swan, Endre Mørck and Olaf Janse Westvik (eds.), *Language change and language structure: Older Germanic languages in comparative perspective*. Berlin: Mouton de Gruyter, 233-270.

Syea, Anand. 1993. Null subject in Mauritian Creole and the pro-drop parameter. In: Francis Byrne and John Holm (eds.), *Atlantic Meets Pacific. A Global View of Pidginization and Creolization*, 91-102. Amsterdam: John Benjamins.

Taylor, Ann. 1994. The change from SOV to SVO in Ancient Greek. *Language Variation and Change* 6: 1-37.

Thim-Mabrey, Christiane. 1988. Satzadverbialia und andere Ausdrücke im Vorvorfeld. *Deutsche Sprache* 16: 52-67.

Thomason, Sarah Grey and Terrence Kaufmann. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley, Los Angeles, London: University of California Press.

Thompson, Ellen. 1999. The temporal structure of discourse: the syntax and semantics of temporal *then*. *Natural Language and Linguistic Theory* 17: 123-160.

Thompson, William G. 1974. An historical perspective in the Gospel of Matthew. *Journal of Biblical Literature* 93.2, 243-262.

Thráinsson, Höskuldur. 1996. On the (non-)universality of functional categories. In *Minimal Ideas. Syntactic Studies in the Minimalist Framework*, W. Abraham, S.D. Epstein, H. Thráinsson and J.-W. Zwart (eds.), 253-281. Amsterdam: John Benjamins.

Thráinsson, Höskuldur. 1997. The chapters by Kiparsky, Roberts and Weerman: an epilogue. In Ans van Kemenade and Nigel Vincent (eds.), *Parameters of morphosyntactic change*. Cambridge: Cambridge University Press, 495-508.

Tomaselli, Alessandra. 1995. Cases of verb third in Old High German. In *Clause Structure and Language Change*, A. Battye and I. Roberts (eds.), 345-369. Oxford: Oxford University Press.

Traugott, Elizabeth and Bernd Heine (eds.). 1991. *Approaches to Grammaticalization*. Amsterdam: Benjamins.

Traugott, Elizabeth and Bernd Heine (eds.). 1991. *Approaches to Grammaticalization*. Amsterdam: Benjamins.

Travis, Lisa deMena. 1984. *Parameters and Effects of Word Order Variation*. Doctoral dissertation, MIT.

Travis, Lisa deMena 1985. The role of INFL in word order change. In *Papers from the 4th International Conference on English Historical Linguistics*, R. Eaton et al. (eds.), 331-341. Amsterdam: John Benjamins.

Trips, Carola. 2002. *From OV to VO in Early Middle English*. Amsterdam: John Benjamins.

Trips, Carola and Eric Fuß. 2008. The syntax and semantics of the temporal anaphor 'then' in Old and Middle English. To appear in *Proceedings of CGSW 21 & 22*, T. McFadden and F. Schäfer (eds.). Amsterdam: John Benjamins.

Trommer, Jochen. 2001. *Distributed Optimality*. Doctoral dissertation, University of Potsdam.

Trommer, Jochen. 2003. Feature (non-)insertion in a minimalist approach to spellout. In *Proceedings of CLS 39*, 469-480. University of Chicago: Chicago Linguistic Society.

Trommer, Jochen. 2006. Third-person marking in Menominee. Ms., University of Leipzig.

Truckenbrodt, Hubert. 1995. *Phonological phrases: their relation to syntax, focus, and prominence*. Doctoral dissertation, MIT.

Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30:2, 219-255.

Uriagereka, Juan. 1995. Aspects of the syntax of clitic placement in western Romance. *Linguistic Inquiry* 26, 79-123.

Uriagereka, Juan. 1998. *Rhyme and Reason*. Cambridge, Mass.: The MIT Press.

Uriagereka, Juan. 1999. Multiple spell-out. in Samuel D. Epstein and Norbert Hornstein (eds.), *Working Minimalism*, 251-282. Cambridge, Mass.: MIT Press.

Vainikka, Anne and Yonata Levy. 1999. Empty subjects in Finnish and Hebrew. *Natural Language and Linguistic Theory* 17, 613-671.

Vanelli, Laura. 1987. I pronomi soggetto nei dialetti italiani settentrionali dal Medio Evo a oggi. *Medioevo Romanzo XII*.

Veenstra, Tonjes. 1994. The acquisition of functional categories: The creole way. In Dany Adone and Ingo Plag (eds.), *Creolization and Language Change*, 99-117. Tübingen: Niemeyer.

Vennemann, Theo. 1974. Topics, subjects, and word order: From SXV to SVX via TVX. In: J. Anderson and C. Jones (eds.), *Historical Linguistics: Proceedings of the First International Congress of Historical Linguistics*, 339-376. Amsterdam: North Holland.

Vikner, Sten. 1997. V-to-I movement and inflection for person in all tenses. In *The New Comparative Syntax*, L. Haegeman (ed.), 189-213. London: Longman.

Vikner, Sten. 2001. *Verb Movement Variation in Germanic and Optimality Theory*. Habilitationsschrift, University of Tübingen.

Vogel, Irene. 1999. Subminimal constituents in prosodic phonology. In *Issues in Phonological Structure*, S.J. Hannahs and M. Davenport (eds.), 251-267. Amsterdam: John Benjamins.

Vogelaer, Gunther de, Annemie Neuckermans and Guido vanden Wyngaerd. 2002. Complementizer agreement in the Flemish dialects. In S. Barbiers, L. Cornips and S. van der Kleij (eds.): *Meertens Institute Electronic Publications in Linguistics (MIEPiL) II: Syntactic Microvariation*. Amsterdam, 2002.

Wackernagel, Jakob. 1892. Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1, 333-436.

Warner, Anthony. 1997. The structure of parametric change, and V-movement in the history of English. In *Parameters of Morphosyntactic Change*, Ans van Kemenade and Nigel Vincent (eds.), 380-393. Cambridge: Cambridge University Press.

Wartburg, Walther von. 1970. *Einführung in die Problematik und Methodik der Sprachwissenschaft* (3rd edition). Tübingen: Niemeyer.

Wårwik, Brita. 1995. *þa* and *bonne* in Middle English. In Andreas Jucker (ed.), *Historical pragmatics*. Amsterdam: Benjamins, 345-357.

Watanabe, Akira. 2000. Feature copying and binding: Evidence from complementizer agreement and switch reference. *Syntax* 3.3, 159-181.

Webelhuth, Gerd 1992. *Principles and Parameters of Syntactic Saturation*. Oxford: Oxford University Press.

Webber, Bonnie, Matthew Stone, Aravind Joshi, and Alistair Knott. 2003. Anaphora and discourse structure. *Computational Linguistics* 29.4, 545-587.

Weber, Albert. 1987. *Zürichdeutsche Grammatik. Ein Wegweiser zur guten Mundart*. Zürich: Hans Rohr.

Weerman, Fred. 1989. *The V2 Conspiracy*. Dordrecht: Foris.

Weerman, Fred and Jaqueline Evers-Vermeul. 2002. Pronouns and case. *Lingua* 112, 301-338.

Wexler, Kenneth and Peter Culicover. 1980. *Formal Principles of Language Acquisition*. Cambridge, Mass.: The MIT Press.

Wexler, Kenneth and Rita M. Manzini. 1987. Parameters, Binding theory, and learnability. *Linguistic Inquiry* 18: 413-444.

Weinhold, Karl. 1863. *Alemannische Grammatik (Grammatik der deutschen Mundarten. Erster Theil: Das Alemannische Gebiet)*. Berlin: Ferdinand Dümmler's Verlagsbuchhandlung. Reprint 1967: Amsterdam: Rodopi.

Weiß, Helmut. 1998. *Die Syntax des Bairischen. Studien zur Grammatik einer natürlichen Sprache*. Tübingen: Niemeyer.

Weiß, Helmut. 2002. Agr-in Comp and partial pro-drop. Ms., Universität Regensburg.

Weiß, Helmut. 2005. Inflected complementizers in Continental West Germanic Dialects. *Zeitschrift für Dialektologie und Linguistik* 72, 148-166.

Weiß, Helmut 2006. Pro-drop im Deutschen. Vom AHD zu den rezenten Dialekten. Paper presented at the Kolloquium Nullpronomen (12.09.2006), University of Jena.

Wenzlaff, Michaela and Harald Clahsen. 2004. Tense and agreement in German Agrammatism. *Brain and Language* 89, 57-68.

Wiese, Bernd. 1994. Die Personal- und Numerusendungen der deutschen Verbformen. In *Funktionale Untersuchungen zur deutschen Nominal- und Verbalmorphologie*, K.-M. Köpcke (ed.), 161-191. Tübingen: Niemeyer.

Wiesinger, Peter. 1989. *Die Flexionsmorphologie des Verbums im Bairischen*. Wien, Verlag der Österreichischen Akademie der Wissenschaften.

Wiklund, Anna-Lena. 2005. *The Syntax of Tenselessness: On Copying Constructions in Swedish*. Ph.D. dissertation, University of Umeå.

Wischer, Ilse and Gabriele Diewald (eds.). 2002. *New Reflections on Grammaticalization*. Amsterdam: John Benjamins.

Wratil, Melani. 2008. Uncovered *pro*. On the development and identification of null subjects. Ms., University of Düsseldorf.

Wright, Joseph. 1924. *Grammar of the Gothic language and the Gospel of St. Mark, selections from the other Gospels and the second epistle to Timothy*. Oxford: The Clarendon Press.

Wu, Zoe. 2004. Restructuring and the development of functional categories. In *Diachronic Clues to Synchronic Grammar*, E. Fuß and C. Trips (eds.), 191-217. Amsterdam: John Benjamins.

Wunderlich, Dieter. 1984. Zur Syntax der Präpositionalphrase im Deutschen. *Zeitschrift für Sprachwissenschaft* 3, 65-99.

Wurff, Wim van der. 1990. *Diffusion and Reanalysis*. Doctoral dissertation, University of Amsterdam.

Wurff, Wim van der. 1997. Deriving object-verb order in Late Middle English. *Journal of Linguistics* 33, 485-509.

Wurff, Wim van der. 1999. Objects and verbs in Modern Icelandic and fifteenth-century English: A word order parallel and its causes. *Lingua* 109, 237-265.

Wurff, Wim van der and T. Foster. 1997. Object-verb word order in 16th century English: A study of its frequency and status. In: R. Hickey and S. Puppel (eds.), *Language History and Linguistic Modelling: A Festschrift for Jacek Fisiak on his 60th Birthday*, 439-453. Berlin: Mouton de Gruyter.

Wurmbrand, Susi. 2001. *Infinitives: Restructuring and Clause Structure*. Berlin: Mouton de Gruyter.

Zec, Draga and Sharon Inkelas. 1990. Prosodically constrained syntax. In *The Phonology-Syntax Connection*, Sh. Inkelas and D. Zec (eds.), 365-378. Chicago: The University of Chicago Press.

Zeller, Jochen. 1994. *Die Syntax des Tempus*. Opladen: Westdeutscher Verlag.

Zribi-Hertz, Anne. 1994. La syntaxe des clitiques nominatifs en français standard et en français avancé. In *Travaux de linguistique et de philology*. G. Kleiber and G. Roques (eds.), 131-147. Strasbourg-Nancy: Klincksieck.

Zwart, Jan-Wouter. 1993a. *Dutch Syntax: A Minimalist Approach*. Doctoral dissertation, University of Groningen.

Zwart, Jan-Wouter. 1993b. Clues from dialect syntax: Complementizer agreement. In: Werner Abraham and Josef Bayer (eds.), *Dialektsyntax. Linguistische Berichte Sonderheft* 5, 246-270. Opladen: Westdeutscher Verlag.

Zwart, Jan-Wouter. 1997. *Morphosyntax of Verb Movement. A Minimalist Approach to the Syntax of Dutch*. Dordrecht: Kluwer.

Electronic corpora

TITUS (*Thesaurus Indogermanischer Text- und Sprachmaterialien*). University of Frankfurt. <http://titus.uni-frankfurt.de/>

Kroch, Anthony and Ann Taylor. 2000. *The Penn-Helsinki Parsed Corpus of Middle English, Second Edition (PPCME2)*. Philadelphia: Department of Linguistics, University of Pennsylvania.

Taylor, Anne, Anthony Warner, Susan Pintzuk, and Frank Beths. 2002. *The York-Toronto-Helsinki Parsed Corpus of Old English Prose (York Corpus)*.