

Studies on Basic Word Order, Word Order Variation and Word Order Change in Germanic

Roland Hinterhölzl

University of Venice

1. Introduction

With these studies, I present 14 papers dedicated to exploring different issues concerning basic word order and synchronic and diachronic variation in word order, written in the years 2000-2007.¹ After my dissertation on restructuring infinitives in West Germanic, I have started to examine the nature of variation in the head-complement parameter focussing on (synchronic) comparative investigations on word order in German and English. On the basis of the wide-held opinion (cf. Kemenade 1987) that English started out as an OV-language, I have extended these investigations with diachronic studies, since the development of English promised to yield interesting insights into the nature of this parameter. As it turns out, both English and German started out with mixed OV/VO orders (cf. Pintzuk 1999 for English and Hinterhölzl 2010 for German), raising new and interesting questions about the nature of the variation and about the factors responsible for the development of word order in the two languages.

In accounting for word order variation within one language, it has proven very fruitful to consider the role of information structure. Since information structural categories in German and English typically are correlated with word order alternation (syntactic displacement) and alternation of the prosody, the study of the complex

¹ Obviously, these works appear in the bibliography not in their original versions but in their final published form.

University of Venice

Working Papers in Linguistics

Vol. 22, 2012

interaction between syntax, phonology and information structure represents a key issue in accounting for the pertinent word order changes in Germanic.

From these informal considerations three issues emerge as main objectives of these studies according to which the papers presented here can be grouped as given below. In the following sections, I will discuss these issues in more detail and summarize what the individual papers presented contribute to them².

1. Which factors determine the basic word order in a language and what is the nature of variation in the head complement parameter?

- 01 Hinterhölzl, R. (2000): Licensing Movement and Stranding in the West Germanic OV-languages.
- 02 Hinterhölzl, R. (2002a): Remnant Movement and Partial Deletion.
- 03 Hinterhölzl, R. (2001): Event-related Adjuncts and the OV/VO Distinction.

2. The role of Information structure in accounting for word order variation within one grammar

- 04 Hinterhölzl, R. (2002b): Parametric Variation and Scrambling in English.
- 05 Hinterhölzl, R. (2004a): Scrambling, Optionality and Non-lexical Triggers.
- 06 Hinterhölzl, R. and D. Pili (2004): Argument shift across languages types.
- 07 Frascarelli, M. and R. Hinterhölzl (2007): Types of Topics in German and Italian.
- 08 Hinterhölzl, R. and S. Petrova (2005): Rhetorical Relations and Verb placement in Old High German.
- 09 Hinterhölzl, R. and S. Petrova (2010): From V1 to V2 in older Germanic.

² **A note on the joint publications:** There are four joint publications in the file of papers that I submit here. The papers Hinterhölzl & Pili (2004) and Frascarelli & Hinterhölzl (2007) each contain an Italian and a German part. The German parts were written by Roland Hinterhölzl and the Italian parts were written by Pili and Frascarelli, respectively.

Introduction and conclusions in these papers were joint work. Hinterhölzl & Pili (2004) in particular represents an application of the scrambling account of Hinterhölzl (2002b) on Italian data.

Hinterhölzl & Petrova (2005) and Hinterhölzl & Petrova (2010) result from cooperation of the two authors in the SFB-project B4. The empirical research was done by Svetlana Petrova, while the interpretation and implementation of these data is due to Roland Hinterhölzl, except for the idea that the use of V1/V2 clauses in OHG signals two types of discourse relations which is due to Svetlana Petrova.

- 10 Hinterhölzl, R. (2010): Information Structure and Unmarked Word order in (older) Germanic

3. **The complex interaction between syntax, prosody and information structure that can be exploited in models and explanations of word order change**

- 11 Hinterhölzl, R. (2004b): Language Change versus Grammar Change: What diachronic data reveal about the distinction between core grammar and periphery.
- 12 Hinterhölzl, R. (2006b): The Phase Condition and Cyclic Spell out: Evidence from VP-topicalization.
- 13 Hinterhölzl, R. (2009a): The Role of Information Structure in Word order Variation and Word order Change.
- 14 Hinterhölzl, R. (2009b): A Phase-based Comparative Approach to Modification and word order in Germanic.

Part 1: The Nature of the Variation in the Head-Complement Parameter

2. The Factors Determining Basic Word Order

The main question pertaining to differences in basic word order is whether the distinction between OV- and VO-languages is a basic or derived property in the grammar. The standard assumption in the GB-framework (cf. Haegeman 1991), expanded in the principles and parameter approach (cf. Chomsky and Lasnik 1993) is that basic word order is fixed (arbitrarily) by the head complement parameter. Kayne's (1994) proposal within the antisymmetry approach that languages exhibit the universal base order Specifier – Head – Complement (UBH) triggered anew substantial work on word order phenomena, in general, and on the issue of what the criteria are according to which a particular word order is called basic or derived, in particular. Within the Minimalist Program (MP) with its objective to reduce the inventory of grammatical objects to a minimum, the question – independently of the assumption of whether languages exhibit only asymmetric structures – arises of how to account for differences in basic word order in the absence of simple parameters, the desideratum being that differences between languages should reduce to differences in the interfaces.

2.1. Absence of direct Evidence for basic OV-order in West Germanic

German and the continental West Germanic languages in general have been analyzed as OV-languages since in the unmarked case the direct object precedes the verb that selects it as a complement. However, a closer inspection reveals that if a head is modified, complement and head are not adjacent in the complement-head order, as is the case in German APs and VPs, while they are typically adjacent in the head-complement order, as is the case in German PPs and English APs and VPs, if we abstract away from cases of heavy NP-shift in English.

- (1) a. * sehr über den Hans verärgert (AP)
 very about John angry
 b. über den Hans sehr verärgert
 about John very angry
 ‘very angry about John’

- c. ?? genau einen Brief lesen (VP)
succinctly a letter read
 - d. einen Brief genau lesen
a letter succinctly read
'succinctly read a letter'
 - e. hoch über dem Feld (PP)
high above the field
 - f. * über hoch dem Feld
above high the field
'high above the field'
- (2)
- a. very angry about John
 - b. *angry about John very
 - c. *read carefully a letter
 - d. carefully read a letter
 - e. read a letter carefully

The data in (1) and (2) provide evidence for the underlying structures in (3a-b), indicating that complements leave the containing APs and VPs in German. This is argued in more detail in Hinterhölzl (2000), where it is shown with the help of manner adverbs and the infinitival marker that all VP-internal material except for the verb is evacuated from the VP, barring any direct evidence of basic word order in the German VP.

- (3)
- a. [Modifier [Head Complement]]
 - b. [[Head Complement] Modifier]

For instance, it is argued in Hinterhölzl (2000) that a true manner reading of the adjunct is only obtained in the order complement adjunct head (cf. (1c-d) above), indicating that also indefinite NPs, contrary to Diesing's (1992) influential mapping hypothesis, move out of the VP obligatorily.

Since event-related adjuncts, that is Time, Place and Manner adverbs can also appear postverbally in VO-languages (cf. (2e) above) and since it has been argued by Larson (1988), Stroik (1990) and Pesetsky (1995) that these orders cannot be accounted for by the basic structure in (3b), but call for the presence of VP-shells in which these adjuncts are c-commanded by VP-internal arguments, the test with manner adverbs is

strengthened with the analysis of the infinitival marker in West Germanic in Hinterhölzl (2000).

The position of the infinitival marker in clauses like (4a) poses a problem for the traditional SOV approach to the syntax of West Germanic, since the sequence *zu lesen* (to read) looks very much like a head-initial right-branching structure in which the infinitival marker heading a functional position takes the infinitival VP as its complement to the right.

- (4) a. ohne gestern das Buch zu lesen
without yesterday the book to read-INF
- b. [CP ohne [IP PRO gestern [VP das Buch [V zu-lesen]]]]
- c. [CP ohne [IP PRO gestern [VP das Buch t_i] [I zu+lesen_i]]]

To accommodate this order with the particular assumptions about the word order in the IP within the standard approach, where VP and IP are analysed as head-final projections, it is either assumed that the infinitival marker in the West Germanic languages is not an independent head but rather a verbal affix (Haider 1993), as is illustrated in (4b), or that the infinitival verb undergoes rightward head-movement to (right-) adjoin to the head-final infinitival marker in the IP domain, as is illustrated in (4c).

The crucial empirical evidence that disqualifies these assumptions comes from the position of the infinitival marker in non-finite IPP (infinitivus pro participio) constructions in West Flemish (5a) and Afrikaans (5b).

- (5) a. mee Valere te [willen [dienen boek kuopen]] een
with Valere to want that book buy have
'with Valere having wanted to buy that book'
- b. Die banke moes oop gewees het, om dit gister te [kan betaal] het .
the bank should open been have it yesterday to can buy have
'the bank should have been open to have been able to buy it yesterday.'

Since material, given in square brackets in (5), can intervene between the infinitival marker and the corresponding infinitival verb, it follows that the infinitival marker in the West Germanic SOV-languages cannot be analyzed as a verbal affix, but, like in English, has to be analyzed as occupying a functional position within the IP-domain. That this functional position is not head-final and that the sequence *te+V* cannot be

accounted for by assuming rightward head-movement of the verb within the traditional approach, also follows directly from (5a). In (5a), the constituent that has been moved to the right of the infinitival marker cannot possibly be a head since it contains the DP *that book*.

Hinterhölzl (2000) thus argues that the infinitival marker occupies a functional head to the left of VP (that is later argued to be the head of an Aspect phrase within the VP-phase) and then goes on to show with the help of the infinitival marker that all VP-internal constituents, including APs, verb-particles and PPs move out of the VP in the West Germanic OV-languages.

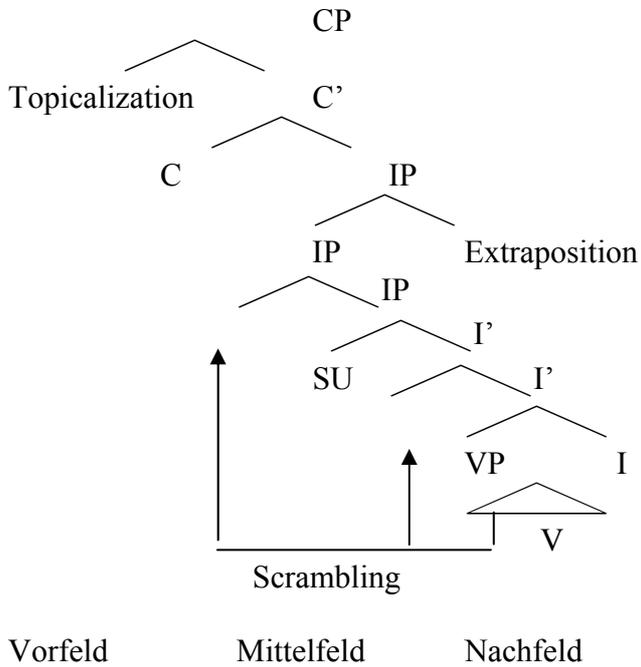
Given that there is direct empirical evidence that in West Germanic OV-orders are derived, while VO-orders exemplify traits of base properties, it seems cogent to adopt the UBH and explore whether its implications trigger coherent results in a wider domain within the syntax of the Germanic languages. This will constitute our approach throughout these studies.

Since movement out of the VP has to be analysed as an obligatory operation that derives the unmarked word order in the West Germanic languages, while scrambling, which is traditionally held to evacuate arguments from the VP, is arguably an operation that derives (more) marked word orders, several questions arise at this point. The first question is: What is the rationale behind these licensing movements out of the VP? Second, is there any independent evidence for the distinction between licensing movement and scrambling, necessary within the UBH-approach? Third, if scrambling is not an operation that moves strong NPs out of the VP for interpretational reasons, what is the nature of scrambling and what are its triggers? The first two questions will be addressed in the remainder of Part I, while the third question is deferred till Section 3 in Part II, where word order variation in the middle field is investigated in more detail.

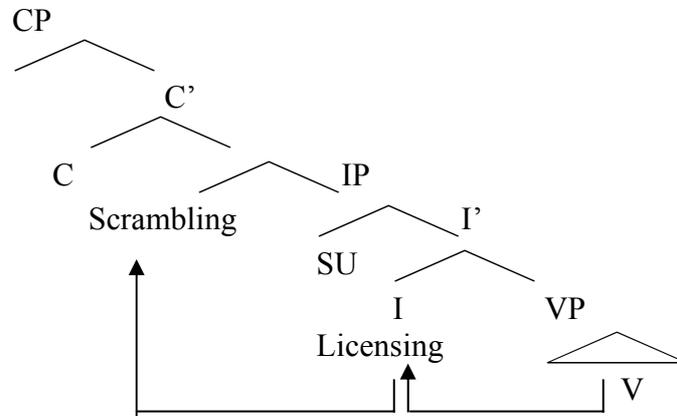
2.2. Motivating licensing movement

The differences between the standard approach to the syntax of the Germanic OV-languages and the UBH-based approach is illustrated in (6). While movement of argumental DPs out of VP is in accord with early minimalist assumptions about Case-licensing (cf. Chomsky and Lasnik 1993, Zwart 1993), the movements of predicates, verb particles and PPs (both argumental and predicatives ones) out of the VP raises questions about their motivation that are addressed in Hinterhölzl (2006a, Chapter 4.2).

(6) a. Standard Approach



b. UBH-Approach



At a closer inspection it turns out, however, that problems arise already with the (Case-) licensing of DP arguments within simple minimalist accounts. First, the derivation of the unmarked order subject > object requires that the arguments of the verb move out of the VP in a parallel fashion. However, there is no principle within MP that would require parallel movement (in standard treatments within MP, the object is licensed in a

secondary Specifier of v^3). Second, since German does not have a unique unmarked word order but displays a number of verb class-dependent unmarked word orders (cf. Haider 1993), individual Cases cannot simply be identified with specific functional positions, as is evidenced by the fact that Nominative can appear before or after an internal argument in the unmarked word orders in (7).

- (7) a. weil der Hans dem Freund half
 since the Hans the friend helped
 ‘since Hans helped the friend’
 b. weil dem Hans die Schallplatte gefiel
 since to-Hans the record pleased
 ‘since Hans liked the record’
 c. weil [_{Agr1P} [der Hans]₁ [_{Agr2P} [dem Freund]₂ [_{vP} t₁ half t₂]]]]
 d. weil [_{Agr2P} [dem Hans]₁ [_{Agr3P} [die Schallplatte]₂ [_{vP} t₁ gefiel t₂]]]]

In Hinterhölzl (2006a), I outline a flexible system of Case-licensing, given in (8) below, that allows to derive different unmarked word orders from thematically determined underlying orders in the vP via parallel movement, as is illustrated in (7c-d). In (7d), since the psych-predicate *gefallen* (please) does not project a (causative) vP, Nominative can be assigned as default Case in Agr3P.

- (8) a. (1) Agr1 is only projected if a (causative) v is present
 (2) Agr2 is only projected if the verb assigns lexical Case
 b. (1) Agr3 checks Accusative if (causative) v is present
 (2) Agr2 checks lexical Case (usually Dative)
 (3) Nominative is a default Case

Next, the issue of parallel movement is addressed. The key economy principle in MP guiding movement processes is *Attract Closest*, which enforces that the subject in (9) is

³ In Chomsky (1995ab), adopting a proposal by Hale & Keyser (1993), transitive VPs are analysed as vPs, in which an abstract v that thematically licenses the subject argument in its Specifier takes the verb and its object argument (DO), the traditional VP in approaches where the subject is taken to be base-generated in [Spec, IP], as its complement. Since the distinction between vP and VP is mostly irrelevant for our purposes here, I will use the term VP when referring to the verb phrase in work written before 1995 and the term vP when referring to minimalist accounts and whenever the distinction is relevant.

moved to the lower Agreement phrase (AgrOP) rather than to the higher one (AgrSP), as is required.

To account for parallel movement - which is not only necessitated by licensing movement out of the vP, but also evident in many other domains within the grammar (witness the presence of minimality effects in scrambling operations or in cases of object shift) - and to preserve the effects of Attract Closest where appropriate, an extension of this economy principle is proposed that gives preference to an element in the smallest cyclic domain as long as the selection of an element in a larger domain does not yield a more economic derivation, as given in (9).

- (9) a) Cyclic Attraction (Hinterhölzl 2006a):
 a probe selects a suitable target in the smallest cyclic domain unless an alternative selection (of a target in a bigger domain) is more economic
- b) Economy:
 the selection of an alternative target in a larger domain is more economic if this selection leads to shorter movements in the overall derivation within the phase containing the probe

What remains to be addressed is the rationale behind licensing movement. There are two notions of licensing that both seem to be relevant here: formal licensing and functional licensing. Chomsky (1995) assumes that only thematic properties are satisfied (through merge) in the vP. This assumption provides a rationale for formal licensing. It implies that all other properties of the verb, its subcategorization of a specific lexical property, traditionally called c-selection, is checked in the extended projections of the verb, that is, in Spec-head relations within the I-domain. In this account licensing movement of PPs becomes on a par with Case checking: They have to move out of the vP to satisfy the selection of the verb for a specific preposition in the Spec-head relation with a functional head that represents this selectional property of the verb. This principle is called c-selection via upwards projection in Hinterhölzl (2006a) and involves the complete syntactification of the idea behind the projection principle: a verb that c-selects a certain morphological property on its arguments projects a functional head with the relevant feature specification, explaining why DPs and PPs, arguments and predicates alike, must move out of the vP.

A second motivation for licensing movement, equally important as formal licensing, is constituted by what may be called functional licensing. As is argued in Hinterhölzl (2006a) functional licensing involves the decision mediated by the

functional structure of the middle field of whether a formal complement of the verb is licensed as an argument or as (co-) predicate of the verb and if licensed as an argument whether it will be realized as a structural, oblique or implicit argument of the verb. This is a generalization of the observation that active and passive verbs, for instance, do not differ in their argument structure, but solely in the way they realize their arguments. For more discussion of this issue see Hinterhölzl (2006a).

Envisaging licensing movement out of the vP, the following generalizations concerning the structure of the lower middle field in German emerge. Nominal arguments (including prepositional arguments) are Case-licensed in Agr-heads, as illustrated in (7) and (8), above the position of manner adverbs. The negative marker dominates manner adverbs. Small clauses, idiomatic expressions and directional PPs are licensed in PredP directly below the position of VP-adverbs. This Predicate Phrase dominates AspP which hosts the infinitival marker and in the Specifier of which verb particles are licensed. AspP itself immediately dominates vP (10). (10) does not display the complete structure of the lower middle field. It just indicates a few landmark positions that I have chosen as criterial for differentiating licensing movement out of the vP from additional movement that scrambles arguments higher up in the structure. I assume that the licensing positions of arguments and their related scrambling positions are interspersed with functional projections licensing adverbs (cf. Frey and Pittner (1998) for some discussion).

(10) [**Scr-ed DPs** [Neg [**DPs** Agr [Manner adverbs [**Predicates** Pred⁰ [**Particles** Asp [vP]]]

2.3. Motivating Licensing Movement and Scrambling

Hinterhölzl (2002a) shows that the distinction between licensing movement and scrambling can solve two long standing problems with the standard account of VP-topicalization in terms of remnant movement and scrambling that goes back to the famous account of Den Besten and Webelhuth (1987). Provided that only XPs can move into XP-positions, it follows that what has been moved into [Spec,CP] in (11) is not simply a verb, but must be minimally a full VP. Thus, they propose to analyze (11a) parallel to cases of VP-preposing (cf. (11b)) in which the direct object has been scrambled out prior to VP-to-CP movement. The moved VP is called a remnant category since it contains, at least, the trace of the direct object as is indicated in (11c).

- (11) a. gelesen hat Hans das Buch
 read-PART has Hans the book
 ‘Hans has read the book’
 b. [_{VP} das Buch gelesen] hat Hans t_{VP}
 c. [_{VP} t_{SCR} gelesen] hat Hans [das Buch]_{SCR}

It is well-known that scrambling typically leads to a freezing effect. A freezing effect occurs if extraction of constituent *a* out of a constituent *b* takes place in a derived position of *b*. This is illustrated in (12a-b). In (12a), extraction of the wh-PP arguably takes place from the base position of the direct object. In (12b), the direct object has been scrambled to a higher position from which extraction is excluded. As is shown in (12c), remnant movement typically leads to an anti-freezing effect, since the direct object in (12c) occurs in a derived position and contains a trace created by extraction. Hinterhölzl (2002, 2006c) shows that the difference between (12b) and (12c) can be derived from the Principle of Attract Closest.

- (12) a. Worüber_i hat keiner [ein Buch t_i] gelesen
 where-about has nobody a book read
 ‘about what has nobody read a book’
 b. *Worüber_i hat [ein Buch t_i] keiner gelesen
 where-about has a book nobody read
 ‘Which topic has nobody read a book about’
 c. [Ein Buch t_i] hat darüber_i keiner gelesen
 a book has there-about nobody read
 ‘Nobody has read a book about that’

However, there is one class of cases of remnant movement that are illicit even if they obey the principle Attract Closest. They involve the topicalization of remnant categories created by wh-movement, as is illustrated in (13). (13a) is a case of topicalization of a clause across a wh-island which leads to a mild, subjacency-like violation. (13b), where the wh-word is extracted from the embedded clause (creating a remnant category), however, is ungrammatical. The same contrast can be observed in English, as is shown in (14) taken from Pesetsky (2000). The relative grammaticality of (13c) is interesting in this respect. If the trace of the wh-word were contained in the fronted VP, as is standardly assumed, then (13c) should be on a par with (13b) and (14b). However, it is just as good as (13a). I will come back to this difference and argue that this contrast can

be resolved with the assumption of licensing movement out of the VP below.

- (13) a. ??[dass Fritz Peter liebt] weiss ich nicht wer gesagt hat
that Fritz loves Peter know I not who has said
'I don't know who has said that Fritz loves Peter'
- b. *[dass Fritz t liebt] weiss ich nicht wen er gesagt hat
[that Fritz loves t] know I not who he has said
'I don't know who he said that Fritz loves'
- c. ??[geküßt] weiss ich nicht wen sie hat
[kissed] know I not whom she has
'I don't know whom she has kissed'
- (14) a. [give a book to John] I can guess who will
b.* [give a book to t] I can guess who Mary will
(I can guess who Mary will give a book to)

Now we can address the two problems connected with the account of verb preposing by Den Besten & Webelhuth (1987): the extraction paradox (A) and the evacuation paradox (B).

A) The assumption that there is licensing movement of VP-internal material to specific positions in the middle field is corroborated by the existence of so-called string vacuous scrambling as is illustrated in (15). In (15), the underlined phrases seemingly occur in their canonical position (in (15b) the direct object follows the subject) but have to be assumed to have been scrambled out of VP in order to be exempted from VP-topicalization.

- (15) a. [_{VP} t gerechnet] hat wie immer keiner damit
counted has as always noone there-with
'as usual nobody has reckoned with that'
- b. [_{VP} t gelesen] hat gestern der Fritz ein Buch darüber
read-PART has yesterday the Fritz a book there-about
'Fritz has read a book about this yesterday'

As (16) shows, these presumed scrambling positions do not behave like regular scrambling positions, since they do not exhibit any freezing effect (cf. (12) above) in that they allow for further extraction which scrambles part of the phrase to a higher

position. Thus, the data in (16) pose the following problem: If scrambling viewed as an operation that moves VP-internal material into the middle field were a unitary operation, then why is that scrambled phrases allow for further extraction in certain positions but not in others?

- (16) a. gerechnet hat da wie immer keiner mit
 counted has there as always noone with
 ‘as usual nobody has reckoned with that’
- b. gelesen hat darüber gestern der Fritz ein Buch
 read has there-on yesterday the Fritz a book
 ‘Fritz has read a book about this yesterday’

B) Another problem for the standard approach is the fact that elements that resist scrambling can be left behind by VP-topicalization. This is illustrated for small clause predicates and indefinite W-words in (17a) and (17b), respectively. I will show in Section 3 that scrambling (proper) may not affect (small clause) predicates. Furthermore, I will demonstrate that indefinite w-words may only scramble to extend their scope domain. This is not the case in (17b). If scrambling were the only operation that can evacuate material from the VP, then it is hard to understand how these elements can be left behind by VP-topicalization.

- (17) a. gegessen hat der Karl das Fleisch roh
 eaten has the Karl the meat raw
 ‘Karl has eaten the meat raw’
- b. gelesen hat die Maria erst gestern was
 read-PART has the Maria only yesterday something
 ‘Its only yesterday that Mary read something’
- c. ??[t_i geküßt]_j weiß ich nicht [wen_i sie hat t_j]
 kissed I don’t know whom she has
 ‘I don’t know whom she has kissed’

A possible solution to these problems is the assumption of licensing movement (prior to scrambling) that moves VP-internal material, irrespective of its quantificational or referential potential out of the VP. This assumption immediately solves the evacuation problem--elements that do not scramble nevertheless leave the VP to be licensed in specific positions in the lower middle-field--but also opens up the way towards a solution to the

extraction problem. Given the distinction between licensing movement and scrambling, we may assume that licensing movement out of the VP does not give rise to a freezing effect and that this freezing effect is connected with the *Specificity effect* of the semantically motivated scrambling operation. We know independently that extraction out of specific DPs is illicit. Since DPs that have scrambled across sentential adverbs or negation (crucially not ones that have scrambled across VP-adverbs) exhibit a specificity effect, the freezing effect associated with these scrambling operations follows without further ado.

Finally note that the trace within the VP in (13c) above, repeated here as (17c), does not behave like a Case-marked wh-trace, otherwise we would expect (13c) to be on a par with (13b) and (14b). However, (17c) gives rise only to a mild, subjacency-like violation and is on a par with (13a), as we have seen above. This fact too follows straightforwardly, if we assume that there is licensing movement of DPs (and of VP-internal predicates) out of the VP that leaves A-movement traces in the VP such that the wh-word in the embedded clause can bind a wh-trace outside of the fronted VP within its clause, as is indicated in (18a). The analysis of (17c) in (18a) makes this case of remnant movement parallel to cases of movement of remnant categories that are created by movement of the subject to [Spec,TP] or another licensing position for subject, as is illustrated in (18b). (18a) thus only differs from (18b) in that it involves extraction of the topicalized phrase out of a weak wh-island which accounts for its slightly marked grammatical status.

- (18) a. ??[t_i geküßt]_j weiß ich nicht [wen_i sie hat t_i [VP t_j]]
 kissed I don't know whom she has
 'I don't know whom she has kissed'
 b. [fired t by the company] John indeed was

To summarize, I have argued that the assumption of movement of DPs and VP-internal predicates to specific licensing positions in the middle-field, as illustrated in (10) above, can provide a solution to the extraction problem as well as the evacuation problem and gives us an explanation for the otherwise rather mysterious contrast between (13b) and (13c).

3. Base Order and the Syntax of Event-related Adverbs

A property that correlates with the position of the object with respect to the verb and that has received little attention in this connection so far, is the position of so-called event related adverbs, that is, Time, Place and Manner adverbs. These adverbs occur preverbally

in the order T>P>M in OV-languages but postverbally in the exact mirror image in VO-languages, as is illustrated in (19).⁴

- (19) a. C T P M- V OV-languages
 b. C V- M P T VO-languages

The properties of event-related adjuncts raise various interesting questions. First, their distribution within OV- and VO-languages raises the question of what makes exactly these adjunct types special such that their positioning, but not the positioning of other adverbs, seems to be correlated with the positioning of arguments. Since these adjuncts might provide us with an insight into the relevant factors that determine basic word order, the investigation of the syntax of event-related adjunct features prominently in these studies on basic word order and word order variation in search of an alternative to a simple parametric approach, as we will show in this section and in section 6.

Secondly, given Cinque's seminal work on adverbs, these adverbials are peculiar in several respects (cf. Cinque 1999). A) They appear to differ from adverb phrases (AdvPs) proper in not being rigidly ordered. B) In contrast to AdvPs proper, they can be interchangeably in the scope of each other, as is illustrated in (20). C) They differ from AdvPs proper in being typically realized in form of PPs or bare NPs. And D) Scope may go from right to left (cf. (21a)), but binding only from left to right, as is illustrated by the contrast in (21bc).

- (20) a. They met students every day of the week in a different university
 b. They met students in each university on a different day

- (21) a. John met Mary in a (different) park every Sunday
 b.* Sue met Mary in his house on everybody's birthday
 c. Sue met Mary on everybody's birthday in his house

⁴ We have seen in (3) above that manner adverbs can occur both preverbally and postverbally in English. This holds for all event-related adjuncts. Preverbal occurrences, though, are only licit if the adjunct is light – in a sense to be made precise. We will come back to this generalization in the following sections.

3.1. Against a Larsonian Approach

The standard approach to the syntax of adverbs in English assumes that adverbs are right-adjoined to VP or IP, as is indicated in (22). Right-adjunction structures, either base-generated or derived by movement, are incompatible with the universal base hypothesis. Independently of the universal base hypothesis, Larson (1988) and Pesetsky (1995) have argued that the standard approach to the syntax of adverbs is mistaken, since it fails to account for basic c-command relations between them and the complements of the verb. Typical c-command diagnostics, as NPI-licensing (23) and quantifier-bound pronouns (24), indicate that postverbal adjuncts are in the c-command domain of postverbal complements.

(22) [_{IP} SU [_{VP} [_{VP} V DO] Adjunct]]

(23) John saw no student in any classroom

(24) John visited everyone on his birthday

Since in the representation in (22) the direct object fails to c-command the postverbal adjunct, Larson (1988) proposed that event related adverb(ial)s are part of a (multi-)layered VP-shell in which these elements are deeper embedded than the complements of the verb as is indicated in (25).

(25) [_{VP} Subject V [_{VP} Object t_V Adjunct]]]

In the Larsonian approach, event related adverbs are analyzed as a sort of (optional) complements in the VP. While the Larsonian approach neatly accounts for c-command effects illustrated in (23) and (24), it fails to account for standard constituency tests such as VP-fronting. The latter process indicates, contrary to the state of affairs in (25), that the verb and the direct object form a constituent which excludes adjuncts (26c). On the other hand, constituents of Larsonian shells motivated by binding do not permit fronting (26d). These observations led Pesetsky (1995) into proposing an additional - layered – structure.

(26) John promised that he would visit them in Vienna on Friday, and...

- a. visit them in Vienna on Friday, he did
- b. visit them in Vienna, he did on Friday
- c. visit them, he did in Vienna on Friday
- d.* them in Vienna on Friday, he visited

Hinterhölzl (2001, 2003) argues that the Larsonian approach to the syntax of event-related adverbs is untenable. Here I will briefly summarize the two decisive arguments. First, the Larsonian approach raises questions about the proper interpretation of these elements. Note that in a Larsonian shell, temporal adverbs are deeper embedded than manner adverbs, as is shown in (27).

- (27) a. John wrote the letter carefully today
 b. [_{VP} John wrote [_{VP} the letter t_V [_{VP} carefully t_V today]]]

Hinterhölzl (2001, 2003) assumes, following Ernst (1998) and Haider (1998), that the attachment of adverbs is determined by their scopal properties. The scopal requirements of an adverb include the selection for a clausal argument of a particular type. Ernst (1998) specifies a schema of abstract clausal entities relevant for the interpretation of adverb(ial)s, given in (28).

- (28) Speech Act > Fact > Proposition > Event > Specified Event

From (28) it follows, for instance, that evaluative adverbs like *unfortunately* selecting for a fact cannot attach lower to the clausal skeleton than modal adverbs like *probably* selecting for a proposition, though they can otherwise occupy various positions in the clause, as is illustrated in (29).

- (29) a. (Unfortunately) Eddie (unfortunately) has (?unfortunately) left
 b. *Probably Eddie unfortunately has left

From a semantic point of view, manner adverbs specify an aspect of only part of the event, namely the process component of the event, while temporal adverbs situate the entire event with respect to the speaking time. Thus, standard assumptions about the interaction of syntactic structure and semantic interpretation predict that temporal adverbs should attach to the clause higher, not lower than manner adverbs, as in the Larsonian approach.

Secondly, here is what is called the comparative argument in Hinterhölzl (2001, 2003). If the English order is basic, then it is not clear how the German order is to be derived. A roll-up structure that moves a constituent containing the temporal adverb in front of the manner adverb and subsequently moves that larger constituent in front of the final position of the verb fails to account for the scopal properties of these adverbs in the middle field. In

the German middle field an adverb always scopes over the adverb to its right. A derivation in terms of individual movement of the adverbs by themselves raises several questions. First, the question arises, why the hierarchy of adverbs in German is different from the hierarchy of adverbs in English, as is illustrated in (30).

(30) [TP \uparrow [VP SU [V [DO V1 [Manner [V2 Temp]]]]]]

The second question that arises is what the motivation of adverb movement in German is. That is, even if we follow Cinque (1999) and assume that the base structure of these adverbs in English is as given in (31), such that the hierarchy of adverbs in German and English is identical, the question arises why these adverbs are licensed in situ in English, but have to undergo licensing movement in German. To the extent that we cannot find a satisfactory answer to this question, the Larsonian approach is rendered unattractive.

(31) [[[VP [e Manner]] e Loc] e Temp]

Given the above considerations, in particular, taking serious the semantic argument, it seems that the order of event related adverbs observed in German, namely T>P>M, is closer to the base than the English order. Thus, Hinterhölzl (2001, 2003) proposes that the English order is derived from the German order via successive cyclic intraposition of verbal projections. In accordance with the observation on the order between infinitival marker and manner adverbs in Section 2.2 (cf. (10) above), we can assume that Time, Place and Manner adverbs are base-generated above VP, as is indicated in (32).

(32) [Temp ... [Loc ... [Manner [V DO]]]]

- (33) a. John visited them in Vienna on Friday
 b. [IP John [on Friday [in Vienna [VP visited them]]]]
 c. [IP John [[[visited them] in Vienna] on Friday]]

Under these assumptions, the English sentence in (33a) is derived from the base structure in (33b) via successive intraposition as indicated in (33c). The representation in (33c) accounts for the VP-fronting facts illustrated in (26), but fails to account for the binding facts. To explain the data in (23) and (24), Hinterhölzl (2001, 2003) proposes that English has *silent* scrambling. Silent scrambling means that there is an overt operation that moves the direct object of the verb into the middle field but spells out the lower copy in the vP and will be treated in detail in sections 3 and 6.

3.2. Base Order and Stylistic Rules

In this section I will address the question of how obligatory VP-intrapolation came about in the history of English. Hinterhölzl (2001) presents a study of word order variation in Old English (OE) and Middle English (ME).

The predominant word order in Old English (OE) is verb-final in subordinate clauses and verb-second in main clauses. Thus OE has been analyzed as an OV-language akin to modern German or modern Dutch. In fact, OE does display a couple of other features typical for OV-languages, according to typological criteria (cf. Hawkins (1983)). Next to the finite verb appearing in final position in embedded clauses, it is a typical feature of OV-languages that verb-particles precede the verb and non-finite verbs precede the auxiliary, as is illustrated in (34).

- (34) a. *Þæt he his stefne up ahof*
 that he his voice up raised
 ‘that he raised his voice’
- b. *forþon of Breotone nædran on scippe lædde wæron*
 because from Britain adders on ships brought were
 ‘because snakes were brought on ships from Britain’

However, OE also displays a number of VO-features. First, we find a considerable number of VO-orders in cases of Verb Raising and Verb Projection Raising. Secondly, we find a considerable number of extraposed PPs, CPs and DPs. While extraposed CPs and PPs are unproblematic, extraposed DPs do not occur in Modern Dutch and Modern German, though they did occur in older varieties of German. Pintzuk & Kroch (1989) show on the basis of a metrical analysis of *Beowulf* that these DPs receive stress. A case in question is given in (35a).

However, there is further evidence that points against a pure OV-character of OE. A) Verb-particles, though they cannot occur after a non-finite verb, can be moved along with the verb in V2-contexts, be stranded somewhere in the middle field and as such be followed by a DP (cf. (35e) below). B) small clauses can follow the finite verb in embedded clauses and can also follow the particle in V2 clauses, as is shown in (35b) and (35c), respectively. C) OE allows postverbal adverbs in embedded clauses, as is illustrated in (35d).

- (35) a. Ðæt ænig mon atellan mææge [ealne one demm]
 that any man relate can all the misery
 ‘that anyone can tell all this misery’
- b. forðam ðe he licettað hie unscyldige
 because that they pretended themselves innocent
 ‘because they pretended to be innocent’
- c. he ahof Ðæt cild up geedcucod and ansund
 he raised the child up quickened and healthy
 ‘he raised the child up strengthened and healthy’
- d. Ðæt hie Ðæt unaliefede dod aliefedlice
 that he the unlawful did lawfully
 ‘that he committed a crime following the laws’
- e. Ða ahof Paulus up his heafod (van Kemenade 1987:33)
 then raised Paulus up his head
 ‘then Paulus raised his head’

The conclusion to be drawn from these observations is that the finite verb moves leftward even in embedded clauses (cf. Pintzuk 1996). Note, however, that this rule will not account for all cases of VO-order. The corpus of W. Koopman (1994) contains, next to 340 OV sentences, all sorts of VO-structures involving non-finite verbs, 94 in totum.

In order to account for these and similar facts, Kroch & Taylor (1997) assume, what I call, a doubly double base. They assume that not only the VP can be head-initial or head-final, but also that the IP can be head-initial or head-final in OE. Not only is this proposal cumbersome, it also allows for a grammar that overgenerates. We do not find sentences of the form V XP Aux, which would result from combining a head-initial VP with a head-final IP.

Alternatively, Hinterhölzl (2001) assumes a VO-based grammar plus licensing movement of arguments and VP-internal predicates to designated positions in the middle field (cf. Zwart (1993). This will derive all the OV-properties of OE. To account for the apparent VO-properties of OE, Hinterhölzl (2001) proposes optional XP-movement of the verb to a medial position (light predicate raising). Assuming this rule has the following advantages: a) it can apply to finite as well as to non-finite verbs and b) it can pied-pipe verb-particles.

Finally, Hinterhölzl (2001) presents the results of a corpus analysis arguing that the loss of OV orders is due, not to the loss of scrambling, as is traditionally assumed (cf. Roberts 1997, Fuß 2002 among many others), but due to a rising amount of light

predicate raising (VP-intrapolation) that as a stylistic rule places heavy arguments and adjuncts into the postverbal field. VP-intrapolation is relevant for fixing basic word order, since this process arguably masks movement of the direct object out of the VP and therefore blurs any evidence (that, for instance, is available for the Dutch child in the absence of Case morphology) of whether the object is to be spelled out VP-externally or VP-internally, as is illustrated in (36).

- (36) a. [IP Adv [VP V O]] base order
 b. [IP O Adv [VP V]] licensing movement
 c. [IP [VP V] [O Adv t_{VP}]] VP-intrapolation
 d. [IP [VP V O] Adv t_{VP}] “reanalysis”

The corpus study investigates the distribution of DPs, PPs and event-related adverbs with respect to the non-finite verbs and yields two important results. A) There is a clear difference in postverbal occurrences according to syntactic category which can be explained most coherently as a length effect. B) While the distribution of PP arguments remains more or less constant across the period investigated, more and more adverbs and DPs appear in postverbal position. The important fact is that the number of postverbal adverbs, which start from a lower level due to their lightness, rises much quicker than the number of postverbal DPs. In other words postverbal adverbs *pull* postverbal DP arguments rendering support to the assumption that the loss of OV-orders is connected with VP-intrapolation.

The corpus study also explains why it is exactly event-related adjuncts and not higher adverbs that appear postverbally in VO-languages like English. Since these adjuncts are typically realized as rather heavy NPs and PPs, they were prime candidates to be affected by the stylistic rule of VP-intrapolation.

The paper closes with a question that highlights the potential role of prosody and information structure (IS) within an approach that tries to explain changes in basic word order as a result of the complex interaction between syntax, prosody and IS which will be the subject of Sections 5 and 6. If it is true, as is claimed in Hinterhölzl (2001), that the postverbal occurrence of event-related adverbs is a characteristic property of VO-languages, how likely is it then that this property came about due to the application of a stylistic rule in the history of all VO-languages? Not very likely, one is inclined to say.

There is another property distinguishing OV- and VO-languages that I have not addressed yet. VO-languages, as a rule, place the main stress at the right edge of an intonational domain, in particular, at the right boundary of the intonational phrase of the

clause, while in OV-languages main stress can be quite removed from the right edge of the intonational phrase of the clause. It is not immediately clear whether light predicate raising is to be considered the consequence or the trigger of this intonational requirement in VO-languages. In the remainder of these studies, we will investigate in more detail the relationship between the two that promises to yield interesting insights into the factors determining unmarked word order.

3.3. Base Order and Unmarked Word Order: the Role of Prosodic Information

Before we can address different issues concerning word order variation in Sections 3 and 4, that is, before we discuss the consequences and triggers of derived or marked word orders, we have to identify the factor(s) that determine(s) the basic or unmarked word order in a language in the absence of a head complement parameter. In Hinterhölzl (2004b), it is proposed that the unmarked word order is determined by prosodic conditions, on the basis of the following argumentation.

We have come to identify as marked or unmarked word order in a language what goes against or conforms to the particular setting of the head complement parameter in this language. When we dispense with the head complement parameter and adopt the UBH (Kayne 1994), then what counts as an unmarked order in a language is not a basic property anymore, but has to be derived from other properties in the language.

We may assume that the unmarked word order is defined by the amount of obligatory movement of the different constituents in a language. For instance, one can say that German has unmarked OV-order because objects obligatorily move out of the vP and the verb stays low, presumably in the vP (except in cases of V2). This is a coherent notion which involves dissociating the unmarked order and the base order.

Dissociating base order and unmarked order may be problematic when it comes to the acquisition of the unmarked order in OV-languages. German children virtually make no word order mistakes and use the correct word orders from very early on. If the child needs to learn all or the majority of movement operations in its language before it can produce unmarked word order patterns, one should expect that in early utterances the German child produces VO-structures, which conform to the presumed base order. But this simply does not happen.

Note, however, that the standard account to the acquisition of the head complement parameter faces a similar problem as well. The common assumption is that in order for the child to set this parameter, it needs to be able to segment the speech

stream into words. Thus, the head complement parameter cannot be set before the child has attained a substantial part of the lexicon.

Alternatively, Nespor, Guasti and Christophe (1996) propose that the head-complement parameter is set on the basis of prosodic information (the rhythmic activation principle). The advantage of this proposal is that the unmarked word order can be determined very early indeed, that is, before any lexical knowledge is attained. More specifically, they propose that the decisive information for the child is the placement of main prominence within the phonological phrase, a constituent of the prosodic hierarchy.

I will adopt this account and propose that the unmarked word order in the phrases of a language is determined, independently of obligatory movement operations, by the predominant, that is, unmarked prosodic patterns in a language.⁵ If a language displays only one prosodic pattern, this pattern will define the unmarked word order. If a language displays several prosodic patterns, the predominant pattern or the predominant patterns will define the unmarked word order(s),⁶ which have to be derived in the core grammar by obligatory movements, while less frequent patterns will be relegated as marked word orders to the periphery, to be derived by stylistic rules.

As we will see in Section 5, the frequency component within the definition of a marked prosodic pattern is important for explaining changes in unmarked word order, with the basic word order always remaining Spec-head-complement.

Having determined that the unmarked word order in a phrase is fixed by the unmarked prosodic pattern in that phrase, the question arises of which factors influence the establishment of unmarked prosodic patterns. As far as the clausal domain is concerned, IS plays a major role in determining what counts as the unmarked prosodic pattern. Therefore, we will investigate the role that IS plays in explaining word order variation in Part II and in influencing the mapping between syntax and prosody in Part III of these studies.

⁵ To get prosodic bootstrapping off the ground, we have to assume that the child has access to the following universal principle that if a head and a non-head are combined in a single phrase it is the head (or the functional element) that is weak. So if the child encounters the pattern (weak strong), it knows that the head precedes the complement or that a functional element precedes a lexical category.

⁶ Note that a language may have several unmarked prosodic patterns. In German, the unmarked pattern with Nouns and Prepositions is weak-strong (w s), but with verbs it is (s w).

Part II: Word order Variation and the Role of Information Structure

This part contains two sections. Section 4 presents comparative studies of scrambling in German, Dutch, English and Italian. We will see that a decisive factor of explaining this type of word order variation is discourse-givenness of a syntactic constituent. Section 5 is dedicated to word order variation in the left and right peripheries in the clause and investigates the role of information structure in fixing word orders in the C-domain and in the lower middle field in the history of German and English.

4. Information Structure and the Nature of Scrambling

In section 1, I have introduced a distinction between licensing movement and scrambling. Licensing movement is an obligatory operation that moves vP-internal constituents into positions in the lower middle field (cf. (10) above) and checks and identifies their lexical and grammatical properties. Scrambling is an optional operation that moves constituents into the higher middle field (above negation and sentence adverbs) according to their semantic and discourse functional properties, as I will argue below.

4.1 A Typology of Scrambling Operations and Evidence for Silent Scrambling

In section 2, I have introduced the notion of silent scrambling to account for the (unexpected) c-command effect between postverbal argument and postverbal adjuncts in English, illustrated again for convenience in (37) below. In this subsection, we will take a closer look at the nature of silent scrambling in the wider perspective of different scrambling operations in Germanic and provide some independent evidence for the existence of silent scrambling in English.

Assuming that event-related adjuncts are base-generated preverbally in the German order and that the English order is derived via successive cyclic intraposition of V-projections (simply called VP-intraposition), raises the question of how to account for the c-command effects in (37).

- (37) a. John saw no student in any classroom
 b. Mary visited everyone on his birthday

Hinterhölzl (2002b) proposes that the data in (37) can be accounted for by an operation of silent scrambling, which like scrambling in German moves the direct object to a c-commanding position above the adjunct in the middle field, but spells out the lower copy in the vP, as is illustrated in (38a) by underlining the relevant copy. VP-intrapolation will then establish the correct word order V-DO-Adjunct, as is illustrated in (38b). In this scenario, bound readings involve the interpretation of the *silent* copy at LF.

- (38) a. DO [Adv [V DO] silent scrambling
 b. [V DO] DO Adv [t_{VP}] VP-intrapolation

First note that scrambling, being an operation of A-movement, creates new binding positions to license bound pronouns and NPIs. As is shown in (39ab) scrambling of the direct object across the temporal adverb can license the bound pronoun contained in the adverb. (40) shows that scrambling can license an NPI.

- (39) a.* weil die Maria an seinem_i Geburtstag jeden_i Freund besuchte
 since the Maria on his birthday every friend visited
 ‘since Mary invited every friend on his birthday’
 b. weil die Maria jeden_i Freund an seinem_i Geburtstag besuchte
 since the Maria every friend on his birthday visited
 ‘since Mary invited every friend on his birthday’
- (40) a.* weil Hans jemals keinen Studenten traf
 since Hans ever no student met
 ‘since Hans never met any student’
 b. weil Hans keinen Studenten jemals traf
 since Hans no student ever met
 ‘since Hans never met any student’

Also note that if we dispense with the Larsonian approach to the syntax of event-related adverbs, then scrambling is the only option to derive the binding facts in (37) : the scrambling analysis is superior to an analysis in terms of a covert operation at LF, since quantifier raising would lead to a WCO-effect in the case of (37b).

4.1.1 Parametric variation and scrambling in English

The assumption that English has scrambling is not an innocent one at all. The traditional explanation of the pertinent word order change is that English lost OV orders due to the loss of its Case morphology (incidentally these two developments temporally overlap, cf. Fuß 2002, Roberts 1997).

As I have outlined above, in the Principles and Parameters approach, word order differences between languages have been explained by assuming parametric variation in the head-complement order. In a framework subscribing to the UBH (Kayne 1994, Chomsky 1995), word order differences have to be related to other properties. In the early versions of the Minimalist Program (Chomsky 1993, 1995), the desideratum is established that parametric differences between languages be reducible to morphological differences of the lexical items of these languages. Along these lines, Vikner (1995) tried to relate the extent of overt V-movement in languages to the *strength* of the morphological paradigm of verbal endings in those languages.

Zwart (1993) proposed that OV-orders are the result of overt movement of the object into AgrOP, while VO-orders result from covert movement of the object to a licensing agreement position. In a similar vein, the difference in the position of the direct object vis-a-vis the verb between English and German can then be related to the strength of the case morphology of the nouns in these languages. Along these lines, Roberts (1997) has accounted for the pertinent word order change in English by arguing that English lost scrambling due to its loss of strong Case morphology.

However, if one looks at the situation in Dutch – OV-order, but no Case-morphology – one realizes that the strength of Case-features in Dutch comes out as a rather arbitrary property. Moreover, if we look at the different positions that verb-particles, small clause predicates and argument-PPs occupy in, say, English and German, then it becomes doubtful whether any morphological property of these elements can be found that could be held responsible for these differences. Thus, I have argued in Section 1 that licensing movement out of vP serves to check different lexical requirements of the selecting verb and can therefore be taken not to be directly driven by exclusively morphological properties of these elements (though, Case morphology may be part of the c-selection in certain languages).

For sure, head final structures can be taken to be triggered by morphology in languages with agglutinative morphology, but it is not evident at all that morphology is so tightly connected with the head complement parameter also in inflectional languages like German, Icelandic or English.

Dispensing with the head complement parameter, Hinterhölzl (2002b) proposes that languages are essentially alike differing only in how they satisfy universal constraints. In other words, the syntax allows a limited amount of options, called macro-parameters, that are fixed by interface requirements which include morphological, prosodic and semantic conditions. These options are specified in (41).

- (41) A) Each feature may be checked via XP- or X^0 -movement
 B) In checking a feature either the higher or the lower copy may be spelled out

With respect to option A, Hinterhölzl (2002b) assumes that Attract Closest implies that XP-movement is the unmarked option in feature checking. X^0 -movement is dependent on special requirements of the attracting head, for instance, on the attracting head being an inflectional affix (a requirement of the interface to word-syntax).

Then it is argued that option A is relevant for the positioning of event-related adverbs in German and English. In Section 2 above, I have proposed that VP-intrapolation applies in English to account for the placement of event-related adjuncts. In Hinterhölzl (2002b), it is proposed that VP-intrapolation applies in German and English alike and is due to licensing movement of the verb: When the English verb undergoes licensing movement, it does so by XP-movement (the unmarked option) which will result in inverted orders. When the German verb undergoes licensing movement, it does so by X^0 -movement, followed by remnant XP-preposing (cf. Hinterhölzl 1998, Haegeman 2001), which leaves preexisting orders unchanged. This proposal unifies the syntax of event-related adjuncts in the two languages as much as possible, but remains unsatisfactory as long as no good motivation is given for why the verb should need to undergo licensing movement. What triggers licensing movement of the verb is discussed in detail in Section 6.

With respect to option B, Hinterhölzl (2002b) proposes that in checking an uninterpretable feature, the higher copy must be spelled out (the lower copy with the unchecked uninterpretable feature being deleted on account of convergence (cf. Nunez 1995)). In checking an interpretable feature either copy may be spelled out, unless the attracting head has a positional feature. A positional feature requires that the attractee is spelled out in the checking domain of the attractor (this is different from the notion of strong features). Typical examples of positional features are [wh], [neg] and [focus].

In retrospect, a positional feature in Hinterhölzl (2002b) is a cover term for different interface conditions, here mostly pertaining to LF, which require that the scope of a

particular element or its IS-status is made visible positionally. A case in question concerns the distribution of definite NPs in German and is discussed below.

Definite NPs always have to move across negation, as is shown in (42ab). We may assume that this movement occurs to check the Specificity feature of the NP. However, if the definite NP is focussed, checking of the Specificity feature may seemingly be dispensed with (42c). This is a problem, since a feature that needs to be checked, should be checked alike in all contexts. A neat analysis of the facts in (42) is obtained, if we assume that the NP has in fact moved into the domain where specifics are licensed (that is above sentence adverbs including negation) but due to the positional feature of focus, is spelled out in [Spec,FocP] taken to occur below negation. Which head in a given language has a positional feature is subject to crosslinguistic variation.

- (42) a. weil der Hans das Buch nicht gelesen hat
 since the Hans the book not read-Part has
 ‘since Hans has not read the book’
- b. *weil der Hans nicht das Buch gelesen hat
 since the Hans not the book read-Part has
 ‘since Hans has not read the book’
- c. weil der Hans nicht das BUCH gelesen hat
 since the Hans not the book-Focus read-Part has
 ‘since Hans did not read the BOOK’

Generalizing the case in (42), Hinterhölzl (2002b) proposes that option B is relevant for the instantiation of scrambling. In this way, silent scrambling in English is not an idiosyncratic property of this language anymore, but instantiates a universal option according to which the spell out of copies is determined by (language specific) interface requirements. In the following subsection, I will provide a typology of scrambling operations in West Germanic and show how scrambling in English fits into the picture.

3.1.2 Silent scrambling: a typology of scrambling operations

A careful investigation of scrambling operations in German and Dutch reveals that there are (at least) three types of scrambling operations whose properties are summarized in the table given below.

<i>Name</i>	<i>Motivation</i>	<i>Permutation of arguments</i>	<i>Recon- struction</i>	<i>moved item bears stress</i>	<i>creates new binding position</i>	<i>Depends on Case</i>
Permutation	Familiarity/ Scope	Yes	No	no	yes	yes
Extension	Familiarity/ Scope	No	No	no	yes	no
F-scrambling	Contrast	Yes	Yes	yes	?	no

First, there is the type of scrambling known from German which allows movement of the object across the subject, as is illustrated in (43ab). It is this operation that can be held responsible for voiding WCO-effects and *Superiority*-effects between subjects and objects in German, as is illustrated in (43cd).

- (43) a. weil den Hans die Maria liebt
since the Hans-ACC the Maria-NOM loves
'since Mary loves Hans'
- b. weil jeden_i seine_i Mutter liebt
since everyone-ACC his mother-NOM loves
'since his mother loves everyone'
- c. Wen_i liebt seine_i Mutter nicht?
Whom loves his mother not
'Whom does his mother not love?'
- d. Wen hat wer geliebt?
Whom has who loved
'Who has loved whom?'
- e. weil jeden eine Frau liebt
since everyone-ACC a woman-NOM loves
'since a woman loves everyone'

I call this operation permutation since it allows for the arguments of the verb to be generated in any order. It is an operation of A-movement since it creates a new binding position (cf. (43b) that cannot be reconstructed as can be seen from the fact that sentences like (43e) are unambiguous. Furthermore, permutation is dependent on the presence of morphological Case.

Permutation is unavailable in Dutch, as is illustrated in (44ab). But Dutch allows for scrambling of the arguments of the verb across adverbs that occur in the middle field as long as scrambling preserves the order of the arguments. I call this operation extension. Extension is a type of A-movement, as we have seen in (39) above, which like permutation does not allow for reconstruction, since the resultant structures are unambiguous.

- (44) a. dat Jan de boeken niet koopt
 that Jan the books not buys
 ‘that John does not buy the books’
 b. *dat de boeken Jan niet koopt
 that the books Jan not buys
 ‘that John does not buy the books’

With permutation and extension, the moved item is unstressed. That there is a third type of scrambling is most clearly shown by Dutch. As was shown in (44), the direct object can generally not move across the subject. However, there is a type of scrambling which Neeleman (1994) called focus-scrambling, but really involves movement of a contrastive topic, which exactly does that, as is shown in (45a). This operation comes with a special intonational contour (the so-called hat-contour, with a rise on the moved item and a fall on the focussed element), obligatorily requires a (contrastively) focussed element in the remainder of the clause and is to be conceived as an A'-movement operation, since it may not only affect arguments, as permutation and extension exclusively do, but also predicates, as is illustrated in (45b).

- (45) a. dat zulke boeken zelfs JAN niet koopt
 that such books even Jan-focus not buys
 ‘that not even John buys such books’
 b. dat rood zelfs JAN de hekje niet verfd
 that red even Jan-focus the fence not paints
 ‘that not even John paints the fence red’

One may wonder why two types of A-scrambling have to be distinguished and what their respective motivations are. As has been noted before (cf. Diesing (1992), scrambling often comes with a specificity effect (in the table above I use the term *familiarity* (cf. Corver & Delfitto 1997) to encompass DPs that are familiar from the linguistic as well as the non-linguistic context).

That A-scrambling cannot be reduced to the sole trigger of specificity/familiarity is shown in (46). In (46), the direct object has scrambled across the indirect object but remains in the scope of the subject. If A-scrambling necessarily came with a specificity effect, then the direct object should only have widest scope in (46). Since an intermediate scope-reading is possible A-scrambling must be able to have an additional trigger which Hinterhölzl (2002b) identifies with scope-taking.

- (46) weil jeder zwei Bücher einer Frau schenkte (intermediate scope possible)
 since everyone-NOM two books-ACC a woman-DAT gave
 ‘since everyone gave two books to a woman’

Hinterhölzl (2002b) assumes that permutation and extension can target the same positions and proposes that these positions are to be identified with the Specifiers of functional heads hosting clitics. This proposal is developed in more detail in Hinterhölzl (2004) and will be discussed in detail in section 3.2 below.

Older stages of Dutch and English had scrambling of the German type. Dutch lost permutation when it lost its Case-distinctions in the nominal system (which are partially preserved in the pronominal system). Given these observations, it seems reasonable to assume that English preserved scrambling of the Dutch type, that is, extension and F-scrambling.

In the following subsection, I present some independent evidence for the conjecture that English has scrambling, as well as some observations that speak for the silent character of scrambling in English.

4.1.3 Silent scrambling and the role of QR

In this subsection, I present an argument developed in Hinterhölzl (2002b) that supports the conjecture that English has scrambling, namely antecedent-contained deletion. More specifically, I will argue that an account of ACD-resolution in terms of scrambling is superior to an account in terms of QR.

The assumption of (silent) scrambling has the advantage that it can account for all cases of antecedent-contained deletion (ACD). There are two types of explanations that posit covert movement for ACD-resolution: QR and covert Case-driven movement to AgrO.

Within these approaches it is agreed that movement of the direct object out of the VP is necessary to create the proper antecedent, the V-trace-configuration, that can be copied into the ellipsis site, as is illustrated in (47).

- (47) a. Mary invited everyone that I did
 b. Mary [everyone that I \hat{I}] invited $t \hat{I}$ = [invited t]

May (1985) has argued that the phenomenon of ACD provides good evidence for the existence of QR, since only quantified expressions but not referential expressions can resolve cases of ACD, as the contrast between (47a) and (48a) shows.

- (48) a. *Mary invited Peter who I did
 b. Mary invited Peter, who Susan also did

However, Cases of ACD involving names are not impossible in general. (48a) improves considerably if the relative clause expresses some contrast to the previous clause, that is, if it contains some focused element, as is illustrated in (48b). Since names are not the type of elements that can reasonably be assumed to be in need of QR for the sake of their interpretation, QR-based approaches have problems with ACD involving referential expressions like names. This observation led Hornstein (1994) into proposing that cases of ACD are resolved by Case-driven movement into AgrO in English, the argument being that both referential and quantificational DPs are in need of Case. Note, however, that Case movement-based explanations have problems with NP-contained ACD, as has been pointed out by Kennedy (1997) and Pesetsky (2000). The interpretation of (49a), given in (49b), involves movement of the PP out of the containing NP *a report*, which cannot be motivated with the need for checking Case.

- (49) a. Mary read a report on every murder the police did
 b. every murder Mary read a report on, the police read a report on
 c. weil Maria über jeden Mordfall einen Bericht las
 since Maria about each murder a report read
 ‘since Maria read a report about each murder’

Note, however, that scrambling can be argued to be able to resolve all the cases of ACD illustrated in (47-49). Names and QPs scramble alike in German, though due to different motivations as we have seen in the previous section. And scrambling may also move a PP out of an NP, as is illustrated in (49c), as long as the NP is indefinite, which is exactly the restriction that applies to cases of NP-contained ACD.

The assumption of silent scrambling may even provide us with an explanation for the restriction on ACD involving names that is illustrated in (48ab) above. Let us make the reasonable assumption that each movement operation must be either PF- or LF- interpretable. Silent scrambling never has a PF-effect and will thus be only licit if it has an LF-effect. However, scrambling of names is not interpretable in terms of scope, since a name, independently of its position, is always interpreted in the same way, namely, with widest scope. It thus follows that silent scrambling of names cannot be extension-scrambling but must be F-scrambling which, as we showed in the previous section, necessarily involves a focused constituent which sets up the required contrast.

Since I proposed that cases of ACD are better not treated as being resolved by the operation of QR, the question arises whether QR plays any role in English and in language in general. In the spirit of Kayne (1998), it would be advantageous to have a grammar without postcyclic rules, that is to say, a grammar in which all operations occur overtly in a cyclic fashion. Two phenomena immediately come to one's mind which were traditionally taken as evidence for the existence of QR: cases of inverse scope, as illustrated in (50), and cases of inverse linking, as illustrated in (51).

(50) Someone loves everyone

(51) Some inhabitant of every city hates its traffic

In the remainder of the paper, Hinterhölzl (2002b) shows that both cases can be better explained in terms of scrambling rather than in terms of QR, rendering further support for the existence of silent scrambling in English.

4.2. What triggers scrambling in German: a Minimalist Account

In Section 1, we have introduced a distinction between licensing movement and scrambling, providing a rationale for licensing movement in terms of checking of c-selectional features and proposing that scrambling is driven by semantic and discourse related properties. In Hinterhölzl (2002b), as discussed in the previous section, the

relevant triggers, namely specificity and scope, were already identified. Hinterhölzl (2004a) formally implements these ideas in an LCA-based framework which dispenses with adjunction and assumes that a) all movement operations involve feature checking and b) are not optional.

The first issue pertains to the long standing debate about the A/A'-nature of scrambling in German. In the previous section we have distinguished three types of scrambling operations: f-scrambling, extension and permutation. We have argued that extension and permutation are triggered by the same features. Moreover, both operations share the property that the moved item bears no special stress or is deaccented and can be considered as A-movement operations, while the operation of scrambling in which the moved item is stressed shows clear properties of A'-movement. It can affect arguments as well as predicates and is not clause-bound. This operation moves contrastive topics and comes with a special intonation, the so-called hat contour comprising a fall-rise tone in the moved topic and a fall tone on the (contrastive) focus-element in the remainder of the clause, as we have seen in the previous section. This operation has been called focus scrambling by Neeleman (1994), I-topicalization by Jacobs (1997) and T-scrambling by Haider and Rosengren (1998). Because the moved item bears stress, I will use the term S-scrambling. In this section, I will put S-scrambling on the side and concentrate on the operations extension and permutation, simply called scrambling (proper) here.

4.2.1 Properties of Scrambling in German

In this subsection, I will argue that extension and permutation are the same type of operation that targets the same kind of positions in German and Dutch and embodies a case of A-movement. That scrambling should be treated as A-movement is not a novel idea at all and has been entertained by many researchers starting with Van den Wyngaerd (1989). However, these authors considered scrambling to constitute (overt) movement of an argument out of the vP into a Case-licensing position in the middle field. Since I distinguish between licensing movement and scrambling, we have to envisage the existence of A-movement beyond movement into Case positions. But before we can address this issue, we must firmly establish that scrambling (proper) indeed constitutes a case of A-movement.

This involves providing arguments for A-movement and at the same time discarding the arguments that have been brought forward in the literature in favor of an account in terms of A'-movement.

Arguments for A-movement

Hinterhölzl (2004a) shows that scrambling proper is restricted to arguments (52) and strictly clause-bound (53). In (52), the infinitival predicate has been scrambled across negation leading to ungrammaticality. In (53), long distance scrambling of the embedded object results in ungrammaticality.

- (52) a. weil jeder oft gewinnen möchte
 since everyone often win wants
 ‘since everyone wants to win often’
 b. *weil jeder gewinnen oft möchte
 since everyone win often wants
 ‘since everyone wants to win often’
- (53) a. Gestern hat niemand geglaubt, dass er die Maria einladen wird
 yesterday has nobody believed that he the Maria invite will
 ‘yesterday nobody believed that he will invite Maria’
 b. *Gestern hat niemand die Maria geglaubt, dass er einladen wird
 yesterday has nobody the Maria believed that he invite will
 ‘yesterday nobody believed that he will invite Maria’

Scrambling can create new binding possibilities. In (54a), the quantifier cannot A-bind the pronoun simply for lack of c-command. However, if the direct object quantifier is scrambled across the subject containing the pronoun, the latter can be bound with no WCO-effect being noticeable.

- (54) a. *weil seine_i Mutter jeden_i liebt
 since his mother-NOM everyone-AKK loves
 ‘since his mother loves everyone’
 b. weil jeden_i seine_i Mutter liebt
 since everyone-AKK his mother-NOM loves
 ‘since his mother loves everyone’

Scrambling CAN lead to scope ambiguities. Stress in the former statement is on the modal *can*, since there seem to be two groups of speakers. For speakers of the first group, scope is solely a matter of surface relations. For these speakers, including myself, the scrambled structure in (55b) is as unambiguous as the base structure in

(55a), though the scope relations have been inverted by scrambling. For speakers of the second group, the base structure is unambiguous as well, but the scrambled structure is ambiguous. In (55b), these speakers not only get the reading that results from surface scope (as speakers of the first group do), but also get the reading that results from reconstructing the scrambled object into its base position.

- (55) a. weil [mindestens eine Frau] [fast jeden Mann] liebt
 since at least one woman-NOM almost every man-ACC loves
 ‘since at least one woman loves almost every man’
- b. weil [fast jeden Mann] [mindestens eine Frau] t liebt
 since almost every man-ACC at least one woman-NOM loves
 ‘since at least one woman loves almost every man’

It is important to note that also for the second group binding relations are strictly read off from surface relations. In German, scrambling may not only create new binding possibilities, it may also destroy binding possibilities, as the example adopted from Haider & Rosengren (1998) shows. Even for speakers of the second group, the unbound anaphor in (56b) cannot be saved by reconstructing it into its base-position.

- (56) a. dass der Mann die Bilder einander anglich
 that he the pictures-ACC each other-DAT made-alike
 ‘that the man made the pictures similar to each other’
- b. *dass der Mann einander die Bilder t anglich
 that he each other-DAT the pictures-ACC made-alike
 ‘that the man made the pictures similar to each other’
- c. einander hat der Mann die Bilder angeglichen
 each other-DAT has he the pictures-ACC made-alike
 ‘since the man made the picture similar to each other’

To sum up, the fact that scrambling is a) restricted to arguments, b) clause-bound and c) can create new binding possibilities, speaks in favor of an analysis in terms of A-movement. Moreover, the fact that scrambling cannot be reconstructed for reasons of binding speaks against an analysis in terms of A'-movement, since as is shown in (56c), an anaphor that has been A'-moved into [Spec,CP] can be bound via reconstruction. Thus, there is little reason to assume that scrambling in German should not be A-movement.

Discarding Arguments for A'-movement

One argument that played an important role in the debate on the nature of scrambling was the observation that scrambling can license parasitic gaps. If parasitic gaps can only be licensed by A'-movement, as seems to be the case in English, then scrambling ought to be A'-movement, so the argument went.

Neeleman (1994) provides an excellent discussion of this argument and convincingly argues that this fact has been overrated. It is sufficient here to present his strongest argument, namely the observation that so-called parasitic gaps can also be licensed by undisputed cases of A-movement namely by passivization. This holds for Dutch as well as for German.

In an influential paper, Grewendorf and Sabel (1999) investigate scrambling in German and Japanese and argue that clause-internal scrambling in German, contrary to clause internal scrambling in Japanese, must be A'-movement. Their claim is decisively based on two arguments/observations. First, scrambling of a potential binder in German, contrary to scrambling in Japanese, cannot license an anaphor contained in a DP. Secondly, they argue that the lack of WCO-effects with scrambling in German should not be taken as evidence for an A-movement analysis, since clear instances of A'-movement in German can be found which do not give rise to a WCO-effect either. I will take up the latter issue in the following subsection.

To illustrate their first point, let us look at the data in (57). In (57a), the anaphor contained in the direct object is licensed by the c-commanding subject. In (57b), the anaphor contained in the subject cannot be licensed by the direct object, a potential antecedent, for lack of c-command. If scrambling were an instance of A-movement in German, so they argue, then the direct object in (57c) should be able to license the anaphor, contrary to fact. Since in the parallel case in Japanese, where scrambling is to be treated as A-movement, the anaphor is licensed, clause internal scrambling in German must be an instance of A'-scrambling.

- (57) a. weil Peter_i einen Freund von sich_i eingeladen hat
 since Peter-NOM a friend of himself-ACC invited has
 'since Peter invited a friend of his'
- b. *weil ein Freund von sich den Peter eingeladen hat
 since a friend of himself the Peter-ACC invited has
 'since a friend of himself has invited Peter'
- c. ??weil den Peter ein Freund von sich eingeladen hat
 'since a friend of himself has invited Peter'

Note, however, that the postnominal anaphor in German behaves like a long distance anaphor in being strongly subject-oriented, as is shown in (58). In (58), only the subject qualifies as an antecedent for the anaphor within the complex noun phrase. Thus, I would like to contend that (57c) is ungrammatical not because scrambling is not an instance of A-movement in German but because an object does not qualify as an antecedent for the subject-oriented anaphor in German.

- (58) weil der Peter_i den Hans_j zu einem Freund von sich_{i/??j} geschickt hat
 since the Peter-Nom den Hans-ACC to a friend of himself sent has
 ‘since Peter sent Hans to a friend of himself’

Since the binding of anaphors is the only empirical argument they advance in favour of an A'-movement analysis and since clause-internal scrambling in German in their account would stand out compared to clause internal scrambling in Hindi (which has been shown by Mahajan (1994) to be a clear instance of A-movement) and Japanese (which they themselves argue to be an instance of A-movement as well), Hinterhölzl (2004a) concludes that the observation in (58) weakens their argument to a degree that it seems ill-advised to base such a strong claim on the sole data of (57).

Vikner's Argument

Vikner (1994) puts forward an important comparative argument in favor of an A'-movement analysis of scrambling in German. Vikner points out that scrambling in Dutch and object shift in Scandinavian, contrary to scrambling in German, may not permute arguments. For instance, in Icelandic, a definite direct object cannot move across the indirect object, as is shown in (59).

- (59) a. Petur keypti bokina ekki
 Peter bought book-the not
 ‘Peter did not buy the book’
 b. *Petur syndi bokina oft Mariu
 Peter showed book-the often Mary-DAT
 ‘Peter often showed the book to Mary’
 c. Peter zeigte das Buch oft der Maria
 Peter showed the book often the Maria-DAT
 ‘Peter often showed the book to Mary’

If object-shift is analyzed as A-movement into AgrO, then the contrast in (59) can be explained as a violation of (relatized) minimality (cf. Rizzi 1990): the intervening indirect argument in an A-position would block object-shift of the direct object in (59b). If scrambling in German is A-movement as well, so Vikner argues, then it is surprising that no minimality effect can be observed in this language (59c).

Vikner concludes that free permutation of arguments in German calls for an account in terms of A'-movement: not only would this account explain the lack of minimality effects, adjunction would also be the most elegant way to derive the great number of possible word-orders in German. (60) shows for the case of three arguments that all possible permutations, namely six, are grammatical. The permutations of the base order in (60a) are most easily derivable by (multiple) adjunction to VP and IP.

- (60) a. weil Peter der Maria das Buch zeigte
 since Peter the Maria-DAT the book-AKK showed
 b. weil Peter das Buch der Maria zeigte
 c. weil das Buch Peter der Maria zeigte
 d. weil der Maria Peter das Buch zeigte
 e. weil der Maria das Buch Peter zeigte
 f. weil das Buch der Maria Peter zeigte
 ‘since Peter showed the book to Mary’

Vikner then goes on to challenge the arguments that favor an analysis of scrambling in terms of A-movement. He notes that almost all the arguments are based on the lack of WCO-effects in German scrambling. Given that wh-movement does trigger WCO-violations (61a), scrambling and wh-movement cannot be of the same kind.

However, so Vikner argues, and this brings us back to the second argument of Grewendorf and Sabel (1999), it is not possible to have this kind of WCO-violation at all in German and concludes the crucial difference would not seem to be between scrambling and wh-movement but between German and English. As is shown in (61b) not even (local) wh-movement does trigger a WCO-violation in German.

- (61) a.* Who_i does his_i mother love?
 b. Wen_i liebt seine_i Mutter nicht?
 Who-AKK loves his mother-NOM not
 ‘who is such that his own mother does not love him’
 c. Wen_i liebt [t_i [seine Mutter nicht t_i]]

The observation is correct, but the above conclusion is invalid. It is true that no WCO-effects are observable in a single clause in German. However, when it comes to long distance movement, we can detect an interesting contrast (cf. Frey 1993). Long distance wh-movement of the embedded object does give rise to a WCO-effect, if the pronoun is contained in the matrix subject, but not, if the pronoun is contained in the embedded subject, as is illustrated in (62). If A'-movement in German, as Vikner argues, were exceptionally not subject to WCO, then the contrast in (62) remains unexplainable.

- (62) a. Wen_i glaubt Peter dass seine_i Mutter t nicht liebt
 Who-AKK believes Peter that his mother not loves
 ‘who does Peter believe is such that his own mother does not love him’
- b. *Wen_i glaubt seine_i Mutter daß der Peter t nicht liebt
 Who-AKK believes his mother that the Peter not loves
 ‘who does his own mother believe is such that Peter does not love him’

However, if we assume that the WCO-effect can be circumvented by prior scrambling of the wh-word (cf. 61c), then the contrast in (62) follows simply from the fact that scrambling is clause-bound. Given this assumption, scrambling can provide an A-binder for the pronoun in the embedded subject but not for the pronoun in the matrix subject. Consequently, wh-movement in (62b) only leads to an A'-bound pronoun in the matrix clause and therefore does give rise to a WCO-violation in as much as wh-movement does in English.

On the other hand, the (well-motivated) assumption that German scrambling is A-movement gives us for free the property that scrambling, contrary to S-scrambling, is clause-bound. Were scrambling indeed an operation of A'-movement, then it would be hard to explain why long distance scrambling, for instance in (53b) above, cannot make use of [Spec,CP] of the embedded clause as a licit escape hatch. If scrambling is treated as A-movement, then movement via the embedded [Spec,CP] falls prey to the uniformity condition on chains, an option that we may assume is open for the A'-movement operation of S-scrambling.

4.2.2 A-movement beyond movement into Case positions

Though Vikner's refutation of the A-movement approach to German scrambling fails, the issues that he raises are important. Its comparative aspects will be addressed in the

following section. Here we will show how the data in (60) can be accounted for in the absence of (multiple) adjunction.

In order to dissociate A-movement from Case-licensing movement, Hinterhölzl (2004a) proposes to make the nature of movement dependent on the type of feature that is checked in the target head. If the feature checked is an operator feature, [wh], [foc] or [neg] to name a few, then the movement operation will have the properties of A'-movement. If a non-operator feature is checked, for instance, [specificity] or [topicality] of an argument, then the movement operation will have properties of A-movement.

Now we can address the issue of what the landing positions and the triggers of this type of A-movement are. In section 3.1, we have identified two triggers for A-scrambling in German, namely specificity in the sense of Enc (1991) and scope. Specificity in the sense of Enc (1991) is now often called discourse-givenness or simply givenness. In the following, I will use these terms interchangeably. In Hinterhölzl (2004a), it is proposed that scrambling of specific (given) DPs is movement into the Specifier-position of heads licensing weak pronouns. Since in German there are three functional positions relevant for licensing weak pronouns: one for subject clitics and two for object clitics; one below the subject and one above the subject, as is illustrated in (63), it can be shown that all the orders in (60) can be derived without adjunction by assuming that scrambled DPs move into the Specifiers of clitic-licensing heads in (63).

(63) [C [Cl-S [Cl-O1[Su [Cl-O2 [...]]]]]]

One argument that could be launched against this account of scrambling is the observation that weak pronouns are subject to conditions not observed by scrambled DPs (cf. (64)).

- (64) a. weak pronouns appear in the order NOM>ACC>DAT
 b. weak subject pronouns precede scrambled DPs

Weak pronouns reach their licensing positions by either XP-movement only, or by XP-movement and subsequent head movement. In the latter case, they form a cluster and I would like to propose that condition (64a) pertains to clitic clusters, which are licensed in the highest licensing position for weak pronouns. Since scrambled DPs do not form clusters, they are not subject to condition (64a). Furthermore, if weak pronouns scramble by XP-movement only, they can appear in various orders and different positions, very much like scrambled full DPs. This latter option is available in West Flemish (cf. Haegeman (1994)) and various German dialects (cf. Gärtner and Steinbach (2000)).

To account for condition (64b), I will assume that weak subject pronouns always have to move to the highest licensing position for weak pronouns and that these licensing heads cannot simultaneously license an element in the head position and another element in their Specifier.

The proposal that one type of scrambling is movement into the Specifiers of heads licensing weak pronouns is supported by cross-linguistic considerations. Alexiadou and Anagnostopoulou (1997) argue that there are some striking resemblances between scrambling in Germanic and clitic doubling constructions in Greek and Romance (cf. also Cecchetto (2000)) that pertain to the obviation of WCO-effect with clitic doubled wh-phrases.

4.2.3. The Trigger Problem

Haider and Rosengren (1998) (henceforth H&R) argue that scrambling may not be considered as being triggered by a feature that needs to be checked in a designated position. They point out that trigger accounts are often circular in the sense that they postulate features which apparently are only introduced into the theory to just trigger scrambling. Moreover, they argue that accounts which introduce substantive, independently motivated features prove to be either too weak, too strong or both.

A trigger account is too weak if the trigger just involves DP-type features, like Case or a strong [D]-feature, since it does not cover scrambling of PPs and CPs in German. What is needed are substantive features that are independent of the syntactic category of an argument. A number of features that relate to the semantic or discourse properties of an argument have been proposed in the literature. Such accounts, so H&R argue, are too strong by necessity since they entail that if the respective feature is present scrambling is obligatory, and then they go on to show that scrambling is optional.

First, they argue that the semantic and/or pragmatic effects induced by scrambling cannot be taken to be the triggering factor of scrambling, since the interpretational effect that is induced by scrambling is found in unscrambled structures as well. Scrambling seems to reduce, but not to replace, the interpretation potential of a phrase.

As evidence they provide examples in which a generic interpretation (65a), a specific interpretation of an indefinite (65b) and a specific definite interpretation (65c) are applicable to DPs in situ. For instance, in (65a), the definite DP object follows an indefinite pronoun subject. So the generically interpretable DP is likely to be in situ (the assumption being that nonspecific indefinites resist scrambling). Analogous considerations apply to (65b-c). In (65b), the indefinite is interpreted as specific, given that its reference is picked

up by *Maria*, though it seems to occur in its base-position following an indefinite nonspecific subject. The same holds for the definite DP *her dress* in (65c), which--though occurring in its base position--can have a specific interpretation referring to *Maria's dress*. I share their judgements and agree that the data in (65) are problematic for a simple minded trigger account, but I do not agree with their conclusions, as we will see below.

- (65) a. dass wer die Pockenviren ausrotten sollte
that someone the pockvirus exterminate should
'that someone should exterminate the pockvirus'
- b. wenn wer eine rothaarige Frau sucht dann ist das Maria
if someone a red-haired woman seeks then is it Maria
'if someone is looking for a red-haired woman then it is Maria'
- c. dass er wem ihr Kleid gezeigt hat, hat Maria nicht gefallen
that he someone her dress shown has, has Maria not pleased
'that he has shown her dress to someone did not please Maria'

Other triggers that have been proposed involve semantic-driven movement, that is, movement of bare plurals into strong DP positions as in De Hoop (1992) or *familiarity* as in the account of Corver and Delfitto (1997). But again, so H&R argue, these accounts prove to be too strong. If indeed scrambling was triggered by a specific interface feature, (66a-b) should differ to the extent defined by the absence or presence of that feature. However, (66a) can have the same interpretation as (66b), namely that *Max in general admires primaballerinas*. Therefore, so H&R argue, the claim that a strong *generic* feature triggers scrambling cannot be correct.

- (66) a. dass ja Max Primaballerinas bewundert
that PRT Max primaballerinas admires
that Max admires primaballerinas, as everyone knows'
- b. dass ja Primaballerinas Max bewundert
that PRT primaballerinas Max admires
'that Max admires primaballerinas, as everyone knows'

In this case, I find the evidence less convincing. In my judgment, I get an existential reading of the bare plural if the direct object in (66a) receives nuclear stress and a generic reading if the verb is assigned nuclear stress. This latter fact would indicate that the direct object in this reading of (66a) does not occupy its base-position. In this case, it also seems

plausible that the definite subject DP *Max* has moved to a high position above the generically interpreted object. This analysis of (66a) presupposes that the modal particle *ja*, which is generally assumed to mark the VP-boundary (cf. Diesing (1992), Kratzer (1995)) can occupy a position high up in the clausal domain. Such an analysis, however, is not available to H&R since they exclude string vacuous scrambling quite generally.

In a similar vein, if scrambling were triggered by a discourse-type feature, like [+familiar] as in Corver and Delfitto's account, one would expect that the reading associated with scrambled constituents is absent in the base order. (67) provides a context in which the existence of a unique white ball is given or familiar. Nevertheless, scrambling is not obligatory. In fact, the unscrambled order is the preferred one.

- (67) In diesem Sack sind zwei rote und eine weisse Kugel
 in this bag are two red and one white ball
- a. Wenn jemand die weisse Kugel zieht, hat er gewonnen
 if someone the white ball draws, has he won
 'if someone draws the white wall, he has won'
- b. ?Wenn die weisse Kugel jemand zieht, hat er gewonnen
 if the white ball someone draws, has he won
 'if someone drawas the white ball, he has won'

Hinterhölzl (2004a) discusses in detail the arguments brought forward by H&R against a checking account and discards them one by one. Here, I will only point out the two major flaws in their argumentation and sketch the solution to these problems argued for in Hinterhölzl (2004a). The first wrong assumption is that indefinite w-words do not scramble. However, (68) shows that w-indefinites scramble for reasons of scope taking. In (68a), the w-indefinite is in its expected position following the indefinite (negative) subject. (68a) is unambiguous, meaning 'that nobody met anyone does not surprise me.' In (68b), the w-indefinite has scrambled across the subject yielding the interpretation 'that there is someone that nobody met does surprise me.'

- (68) a. dass keiner wen getroffen hat, überrascht mich nicht
 that noone-NOM someone-ACC met has, surprises me not
 'that noone met anyone does not surprise me'
- b. dass wen keiner getroffen hat, überrascht mich
 that someone-ACC noone-NOM met has, surprises me
 'that there is someone that noone met surprises me'

To utter (68b), the speaker does not need to have a specific individual in mind, it suffices that he has some evidence that there is a person with the relevant property. Thus, the *w*-indefinite is only specific in the sense that it is the DP with widest scope, but it is not specific in the sense that it is presupposed, known to the speaker or in any other way anaphorically anchored in the context. This indicates that scope is a trigger for scrambling. Second, note that *w*-indefinites may not occur in the domain of negation in German (69a-b).⁷ The test with negation shows that *w*-indefinites in contradistinction to other indefinite and definite DPs may not even stay in the scope of negation if they are focussed (cf. (72b) and (73b) below). As is shown in (69c), the *w*-pronoun must be reinforced by an additional element suggesting that *w*-indefinites as weak elements can only occur in scrambled positions and thus rendering the data in (65) less forceful.

- (69) a. *dass jemand nicht wen jagte
 that someone not somebody chased
 ‘that someone did not chase anyone’
- b. dass jemand wen nicht jagte
 that someone somebody not chased
 ‘that someone did not chase anyone’
- c. dass Peter nicht *(irgend-)WEN jagte, sondern Fantomas
 that Peter did not chase just anyone but Fantomas
 ‘that Peter did not chase anyone, but Fantomas’

The second flaw in their argumentation is that they assume too broad a trigger. Definite DPs cannot be equated with discourse given elements. For instance, the definite DP *ihr Kleid* in (65b) refers to a specific individual that is nonetheless discourse-new. Though anchored to the discourse (which is responsible for licensing the definite determiner), the

⁷ H&R argue that this case of alleged obligatory scrambling rests on a controversial premise, namely that negation universally c-commands the whole vP. They argue that the relevant condition for German and Dutch is that negation only needs to c-command the finite verb in its base position. This condition is fulfilled if negation is adjoined to *v* below the base position of arguments, rendering scrambling of indefinite *w*-pronouns superfluous in order to derive the grammatical order in (67b). However, Hinterhölzl (2004a) shows that negation in German, as has been argued for many other languages, occurs in a high position in the middle field (to be precise, is positioned in the T-domain).

definite DP introduces a new discourse referent and receives stress. Since the DP is not specific in the relevant sense, it will not undergo scrambling in my account that is based on discourse-givenness.

The most relevant counterexamples to my account to scrambling is constituted by cases involving apparent non-scrambling of discourse given DPs, as in (67) above. Though I think that the context in (67) is not well chosen, we will discuss it since it indicates the role of focus in the spell-out of scrambled elements.

First, note that the indefinite in (67) is interpreted as universally quantified (for every *x*, if *x* draws the white ball, *x* wins), so that it is far from clear that the indefinite is in its base-position. Secondly, the context itself may not be appropriate to trigger scrambling of a familiar discourse referent, as the parallel example (70) with a pronoun shows. In (70), even the pronoun is awkward in a scrambled position, possibly because the discourse referents have been introduced as a group (or sum) and the choice of one member over the other seems to require some kind of focal stress (selecting but not contrastive focus).

(70) In diesem Zimmer befinden sich ein Mann und eine Frau:

in this room there-are a man and a woman

Wenn jemand ihn sprechen will, wählt er die 9

if someone him speak wants, dials he the 9

‘if someone wants to talk to HIM, he has to dial number 9’

The observation in (67) and (70) accords well with the distribution of discourse given, focused and non-focused DPs with respect to negation in German. To explain the contrastive behavior of the discourse-given DPs in (71a-b), it is proposed in Hinterhölzl (2004a) that all discourse-given elements scramble, but are spelled out according to their interface properties.

(71) Hat Hans die Maria getroffen?

Has Hans the Maria met?

‘Did Hans meet Maria?’

a. Hans hat die Maria nicht getroffen`

Hans has the Maria not met

‘Hans has not met Maria’

b. Hans hat nicht die Maria getroffen, sondern die Sabine

Hans has not the Maria met, but the Sabine

‘Hans has not met Maria, but he has met Sabine’

The same holds for quantified DPs. They appear in front of negation when they have wide scope with respect to it, but appear after the negation if they are non-specific and have narrow scope with respect to it, as is illustrated in (72). Also quantifiers with a specific interpretation regularly precede negation but appear after negation if they have narrow scope with respect to it, as is illustrated in (73).⁸

- (72) a. weil der Hans viele Bücher nicht gelesen hat (only specific interpretation)
 since the Hans many books not read has
 ‘since for many of the books it holds that Hans did not read them’
- b. der HANS hat **nicht** viele Bücher gelesen, der PETER hat viele gelesen
 the Hans has not many books read, the Peter has many read
 ‘it was not Hans but Peter who read many books’
- (73) a. weil der Hans viele der Bücher nicht gelesen hat
 since the Hans many of the book not read has
 ‘since Hans did not read many of these books’
- b. Der HANS hat nicht viele der Bücher gelesen, der PETER hat viele davon
 gelesen
 the Hans has not many of the books read, the Peter has many thereof read
 ‘it was not Hans that read many of these books, but Peter’

H&R have interpreted (73) and other similar data as showing that specifics do not need to undergo scrambling to receive a specific interpretation, the interpretation being available in the base position as well. I do not disagree with this observation, but (73) must be interpreted in the light of the data in (71-72), which show that specifics do undergo scrambling but that this process is overlaid by an interface requirement that operates on the spell out of the resultant copies of scrambling. Hinterhölzl (2004a) concludes that all specific DPs (DPs that are given in the discourse) move to check this feature, but like discourse-anaphoric pronouns that are focused can or must be spelled out in the lower

⁸ To show that scrambling is not optional, but adheres to specific requirements that are correlated with interpretation, Hinterhölzl (2004a) investigates the distribution of indefinite and definite, focused and non-focused, specific and quantified DPs with respect to sentence negation. In order to rule out any confusion with constituent negation, another constituent but the one following the negative marker is focussed in these examples, as indicated by capitals.

position. In this way, scrambling can be treated as an obligatory operation that checks a discourse feature (that is interpreted at the interface as *do not introduce a new discourse referent*) in a position above the negative marker, while the spell out of this operation is determined by the conditions summarized in (74).

- (74) a. contrastive DPs are spelled-out in the focus position
 b. quantified DPs are spelled-out in their scope position
 c. de-accented DPs are spelled-out before accented DPs

The positions relevant for scrambling and its spell out are indicated in (75) (recall that according to (63) above scrambled (specific) DPs can be licensed below or above the subject). The notion scope position that is used in (74) will be defined in the following subsection.

- (75) [_{CP} [_{IP} Specifics [**Subject** Tense [Specifics [**Negation** [Focus AgreePs [V]]]]]]]

While the conditions (74a-b) are conditions relevant at the LF-interface (if we assume that focused elements are interpreted in the nuclear scope of the sentence at LF, then they uniformly concern the scope of an element in the semantic representation), the condition in (74c) is related to the PF-interface. While the LF-related conditions are inviolable - we may assume that the heads licensing contrastive focus und scope have a phonological EPP-feature - the PF-related condition is soft. This is illustrated in (76). For the working of condition (74c), I assume that backgrounded and discourse-anaphoric DPs are de-accented. The question in (76) can either be answered with (76a) or with (76b). While (76a) is completely unmarked and the preferred option, (76b) is slightly marked but completely grammatical.

- (76) Q: Wem hat Otto das Buch gegeben?
 Who has Otto the book given?
 ‘Whom did Otto give the book to?’
- A: a. Otto hat das Buch dem PETER gegeben
 b. Otto hat dem PETER das Buch gegeben
 Otto has (the book) to Peter (the book) given
 ‘Otto has given the book to Peter’

The marked/unmarked status of the examples of (76) can be ascribed to the workings of an interface condition that determines the mapping between information-structure and prosodic structure. I assume that in both answers the direct object has scrambled (across the indirect object) with the difference following from spelling out either the higher or the lower copy.

(77) Interface Condition:

The phonological phrase containing the focus (main accent) must be rightmost within its intonational phrase (cf. Chierchia 1986, Hayes & Lahiri 1991, Frascarelli 2000)

As is illustrated in (78), the different status of the answers follows from the prosodic condition in (77). (78) shows the prosodic structure of both answers, where round brackets indicate phonological phrases and *iP* indicates an intonational phrase. We see that (78a) optimally fulfills the prosodic condition in (77), while (78b) violates this prosodic condition. I propose that this is the reason why (78a) is preferred over (78b). (78b) is repaired by being assigned a stronger pitch accent, while in (78a) the assignment of normal sentence accent suffices to mark the focused constituent. Thus (78b) is prosodically more marked than (78a), but speakers are free to use the more marked forms for their communicative purposes, whatever they are.

- (78) a. [_{iP}(Otto hat) (das Buch) (dem PEter gegeben)]
 b. [_{iP} (Otto hat) (dem PEter) (das Buch gegeben)]

To summarize, we may assume that (74c) is a statistical consequence of the workings of the Interface Condition in (77). Note that we managed to restrict optionality to the workings of condition (77) only. By this manoeuvre, optionality is confined to a PF-interface condition that specifies prosodic requirements on the linearization of phonological material. The syntactic computation, however, including the branch leading to the LF-interface is deterministic throughout.

4.2.4 Feature checking and scope

The other trigger for scrambling as we have seen also in the previous subsection is scope. Scrambling for scope taking is uncontroversially non-optional (cf. (55) above). Nevertheless, it poses a problem for an account that relies on checking privative

features of syntactic elements. The difficulty arises since it is quite inappropriate to assign the respective scopal feature to any specific functional head in the clause. Scopal features – if we introduce the features [w], [n], [i] (for wide, narrow and intermediate scope) for the sake of concreteness – by their very nature are not absolute properties. Scopal properties are relative properties: a DP has wide scope only in relation to another DP. Also since it was argued that scrambling proper is an A-movement operation, a flexible mechanism is needed that replaces adjunction.

In order to achieve this, Hinterhölzl (2004a) proposes an extension of the minimalist framework which allows for the assignment of **non-lexical** features to any head in the course of the derivation according to the mechanism given in (79).

- (79) a) assign the feature to an existing structure (the head at the root) in the course of the derivation
 b) assign the feature to (a copy of) a bare functional head and merge the head with the existing structure

The idea is that arguments and other scope bearing elements enter the derivation equipped with interpretable scope features (w, n, i) according to the communicative intentions of the speaker. These features must be checked with their formal equivalents in a Spec-head relation. The assignment of scopal features to functional heads has to meet the interface condition in (80) which enforces that scopal elements appear in the relevant order.

(80) Scopal Filter

- 1) a head assigned the feature [w] must c-command a head assigned the feature [n]
- 2) a head assigned the feature [i] must c-command a head assigned the feature [n] and be c-commanded by a head assigned the feature [w]

The assignment of scopal features according to (79) defines the scope positions relevant for spell out. Note that *economy of derivation* guarantees that operation b) which is more complex than operation a) will only apply in case the derivation using only a) does not converge. Furthermore, Hinterhölzl (2004a) shows that the standard economy conditions constrain the mechanism in (79) sufficiently to render it a viable alternative to an optional operation of adjunction. *Fewest steps* insures that these features are not assigned repeatedly to a structure and *shortest step* requires that only the smallest extension that guarantees a convergent derivation is taken.

For the sake illustration of this mechanism, let us finally have a look at the sample derivations of two cases: one in which the scopal properties accord with the thematic hierarchy (no scrambling) and one in which the scopal properties of arguments differ from their thematic ordering and scrambling obtains, as is illustrated in (81).

- (81) a. weil jeder mindestens eine Frau liebt
 since everyone-NOM at least one woman-ACC loves
 ‘since everyone loves at least one woman’
- b. weil mindestens eine Frau jeder liebt
 since at least one woman-ACC everyone-NOM loves
 ‘since every one loves at least one woman’

First, the arguments are merged in their thematic position in the VP. Then the Case-checking heads are merged and the arguments move into their Case-licensing positions in a parallel fashion. In (81a), where the subject is to be read with wide scope with respect to the object, the scopal features can be directly assigned to and checked in the Case-positions, with Agr3 being assigned and checking [n] and Agr1 being assigned and checking [w]. In (81b), where the object is to be read with wide scope with respect to the subject, only one scopal feature can be assigned and checked in the Case position, since the assignment of [w] to Agr3 and [n] to Agr1 would violate the scopal filter in (80). Thus, the assignment of [n] to Agr1, which according to (79) is more economic than not using any Case-position for the checking of scopal features, forces the assignment of the remaining scopal feature [w] to a higher head. Since the direct object in (81) is non-specific – if it were specific the scopal features could be assigned to the head licensing weak pronouns (cf. Section 3.2.2) – and since the object does not have any other features to check in the I-Domain, the scopal feature is assigned to a bare functional head which is merged with the existing structure. Then, the direct object moves across the subject to check its scopal feature and is spelled out in its scope position according to (74b). Finally, the complementizer is merged to complete the derivation of the clause.

4.3. A Case of A-scrambling in Italian

Sofar we have discussed cases of A-scrambling in German, Dutch and in the Scandinavian languages (called object shift). Furthermore, I have proposed that A-

scrambling also exists in English with the difference that it spells out the lower copy in the vP. Vikner (1994) has pointed to an important difference between scrambling in German on the one hand, and scrambling and object shift in Dutch and Scandinavian on the other hand. The latter languages observe a minimality effect which is not observable in German; a difference that is still up to be accounted for.

In this context, one may wonder whether A-scrambling can only be found in Germanic languages or whether similar cases of discourse-related movement phenomena can also be found in other languages families. Hinterhölzl and Pili (2004) investigate non-clitic resumed dislocation (NRD) in Italian and compare it with scrambling operations in West Germanic. They show that NRD constitutes a case of A-movement that is triggered by the discourse property [givenness] (called [+ familiar] in Hinterhölzl and Pili 2004) and provide a unified analysis in terms of movement of discourse-given elements into the Specifiers of functional heads licensing clitics. A typical case of NRD is given in (82).

- (82) C: Mario ha appena incontrato Marta. Anche se la conosce poco, credo che la chiamerà presto.
 [Mario has just met Marta. Though he still doesn't know her very well, I think he'll call her soon]
 No, io penso invece che PAOLO a **Marta** telefonerà presto (non Mario)
 [No, instead, I think that PAOLO to Marta will phone soon (not Mario)]
 'No I think that PAULO will soon call Marta, not Mario'

An idiosyncratic property of NRD, not shared by scrambling operations in Germanic, is that it requires the presence of a left peripheral focused element, *Paolo* in (82). Otherwise it would be more natural to realize familiar elements in the right periphery, in other words, to have them right-dislocated in Italian (personal communication Mara Frascarelli).

Another important property of NRD, that is illustrated in (83), is that NRD of direct objects is blocked by preverbal subjects (by subjects that are not dislocated into a focus or topic position), while no such blocking effect is observed with indirect objects, illustrated in (84ab). The properties of NRD are summarized in (85).

- (83) a. *Io credo che DOMANI Mario Marta rivedrebbe volentieri
 [I think that tomorrow Mario_{subj} Marta_{obj} would see again with pleasure]
 b. *Io credo che DOMANI Marta Mario rivedrebbe volentieri

- (84) a. Mi chiedo se IERI Mario a Maria avesse telefonato
[I wondered if YESTERDAY Mario to Mary had phoned]
b. Mi chiedo se IERI a Maria Mario avesse telefonato
- (85) a. NRD of the direct object cannot cross the subject (this is parallel to scrambling in Dutch)
b. NRD cannot be reconstructed for reasons of binding (this is parallel to scrambling in German and Dutch)
c. NRD cannot be reconstructed for reasons of scope (this is parallel to scrambling in German with speakers of Dialect 1)
d. NRD of the indirect object can cross the subject (this is parallel to scrambling in German)
e. NRD is clause-bound

The properties (a-d) in (85) and the differences between scrambling in German and scrambling and object shift in Dutch and Scandinavian can be captured by the definition of relativized minimality in (86).

- (86) a) A-movement of an argument across an overt categorically non-distinct argument is blocked
b) Case (that distinctively marks grammatical functions) can render DP-arguments categorically distinct (cf. Hinterhölzl 2000)

The role of Case in (86) will explain the difference in scrambling between German and Dutch. DPs in Dutch have lost their Case-distinction. Thus a scrambled object in Dutch may never cross another argument, including the subject. Along the same lines, NRD of the direct object targeting a position above the subject will only be licit if the subject is removed, for instance, by moving into a higher Focus-position.

Incidentally, there is independent evidence that Case is relevant for obviating minimality violations. A) Crucially scrambling in German is blocked as well in cases in which two arguments receive the same Case, as is the case in ECM-constructions. This is illustrated in (87). While the embedded subject can scramble across the Case-distinct matrix subject (87b), the embedded object cannot scramble across the non-distinct embedded subject. This is shown by the fact that (87a) cannot be interpreted as meaning that ‘Hans saw Maria kiss Otto’.

- (87) a. weil Hans den Otto die Maria küssen sah
 since Hans-Nom the Otto-ACC the Maria-ACC kiss saw
 ‘since Hans saw Otto kiss Maria’
- b. weil den Otto Hans die Maria küssen sah
 since the Otto-ACC Hans-NOM the Maria-ACC kiss saw
 ‘since Hans saw Otto kiss Maria’

B) In Westflemish and Dutch, Case-marked pronouns can scramble across full DP-arguments. Furthermore, scrambled pronouns can appear in various orders in Westflemish, pretty much like scrambled DPs in German.

C) Object shift in Scandinavian is correctly predicted to be subject to minimality effect, as has been pointed out by Vikner, since Danish, Norwegian and Swedish, like Dutch, have lost their Case distinction. The only exception to this generalisation is Icelandic, which has preserved all its Case distinctions but nevertheless observes minimality effects in object shift. Here, the extra condition that only Case that distinctively marks grammatical function can obviate minimality effects becomes relevant. Note that Case distinctions in Icelandic though robust and numerous do not have this property. Icelandic does have Dative subjects as well as Nominative objects. In German, a DP assigned Nominative, even if it is a deep object, always represents the subject of the clause and the subject of the clause can be no other phrase (except for clauses) than the DP marked Nominative.

D) In Italian, indirect objects do not give rise to minimality effects with respect to the subject, as is illustrated again in (88). (88a) is ungrammatical because the direct object has scrambled across the non-distinct subject. The parallel case in (88b) involving an indirect object is fine, however. We can account for the grammaticality of (88b) by assuming that the preposition ‘a’ which introduces indirect objects in Italian has the same function as Case has in German. It makes the indirect object distinct from both the subject and the direct object, as is indicated in (88c).

- (88) a. *Penso che Mario un libro vorrebbe comprare
 I think that Mario a book would like to buy
- b. Penso che Mario a Marta vorrebbe telefonare
 I think that Mario to Marta would like to phone
- c. Mi chiedevo quando MARIO a Maria quella lettera avrebbe spedito

Summing up, we have shown that scrambling in German and Dutch and object shift in Scandinavian as well as NRD in Italian can be accounted for as instances of A-movement triggered by a discourse-related feature and we argued that the different restrictions on this operation in Italian, Dutch, German and Scandinavian can be subsumed under a uniform minimality condition as given in (86) above.

5. Word order Variation in the left and right peripheries of the clause

So far we have focused on synchronic issues of word order variation and presented comparative studies of word order variation driven by IS-categories in various languages in Section 4. In this section, I will focus on diachronic issues pertaining to word order variation and address the grammatical factors and historical processes that led to the development of rather strict word order regularities in the left and right periphery in Germanic. Before we can address the issues concerning the synchronic and diachronic variation in word order in the left periphery in Germanic, I would like to sketch how the grammatical representation of the left periphery of the clause in German looks like in minimalist treatments that encode information structural categories in the syntax (cf. Rizzi 1997). This is the objective of the following section.

5.1. The Fine Structure of the Left Periphery: Topics in German and Italian

Having seen in the previous section that one type of topic, namely discourse given elements, undergo the same kind of displacement operation in German and Italian, the question arises whether also other kinds of topics are realized in a similar (or quite different) way in the two languages. The corpus study of Frascarelli and Hinterhölzl (2007) investigates the syntactic and phonological realization of topics in natural spoken discourse and reveals that there is a systematic correlation between discourse properties and grammatical properties of topics. The basis of this study is the insight that topics have different functions in the discourse. So far at least three types of topics have been identified in the literature, namely:

- a) ABOUTNESS TOPIC: “what the sentence is about” (Reinhart 1981, Lambrecht 1994), “what is a matter of standing and current interest or concern” (Strawson 1964);

- b) CONTRASTIVE TOPIC: an element that induces alternatives which have no impact on the focus value and creates oppositional pairs with respect to other topics (Kuno 1976, Büring 1999);
- c) FAMILIAR TOPIC: a given, D-linked constituent, which is typically destressed and realised in a pronominal form (Pesetsky 1987), generally used for “topic continuity” (Givón 1983).

Frascarelli and Hinterhölzl (2007) show that that the two languages distinguish the same three types of topics in their syntax and phonology: shifting topics (new aboutness topics), contrastive topics and familiar topics (including continuing or non-new aboutness topics). However, since the grammars of the two languages are organized differently, they also differ considerably in the syntactic realization of these topics. While non-clitic resumed topics in the left and right periphery of Italian exhibit the same properties as scrambling in German (confirming the result of the previous section), clitic resumed topics in Italian display different properties. It is argued that German uses Case to identify the grammatical function of topics and can thus exclusively resort to scrambling. Italian, where (morphological) Case is not available, resorts predominantly to clitics in left-hand and right-hand topic constructions to signal the syntactic function of the dislocated element. Since clitics are related with *pros* in an A-position and pronouns cannot be locally A-bound, Frascarelli and Hinterhölzl (2007) argue that Italian clitic-resumed topics are merged in an A'-position.

One important result of this study, which is also relevant for our analysis of the C-domain in Germanic, is that topics in the C-domain are not fully recursive, as originally proposed by Rizzi (1997) and as is illustrated in (89), but appear in a strict order in the C-domain, as is given in (90), with familiar topics corresponding to discourse-given arguments in the previous section.⁹

(89) [TopP Topic* [FocP [TopP Topic* [IP ...[VP]]]]]

(90) Topic hierarchy:

Shifting topic [+ aboutness] > Contrastive topic > Focus > Familiar topics

⁹ Since the distinction between new (shifting) and non-new (continuing) aboutness topics is irrelevant for the purposes of this paper, we will simply use the term aboutness topic for both of them in the following sections of this paper.

One important question addressed in the paper is how Rizzi's split CP-hypothesis (cf. (89)), as refined in (90), can be reconciled with the V2-rule in modern German. In principle, there are two types of solutions for this problem. A) One could assume that the verb always lexicalizes the highest functional head in the C-domain that is activated. This would mean that – depending on their information structural roles – clauses come in different sizes. In this account, a clause containing a SHIFTING Topic would be larger than a clause with a FAMILIAR Topic, with the verb moving into different functional projections in each case. The other possibility is to assume that B) the verb always moves to the same functional head that has a phonological EPP-feature. This head occurs even higher than the functional projections responsible for licensing Topics in (90) and can be identified with the head expressing clausal Force. One can then assume that it is this position in which a Complementizer is merged and the finite verb is moved to in V2-clauses, as is illustrated in (91). The highest Specifier in the clause will then count as A'-position, since the corresponding head hosts an Operator feature (clausal Force) (cf. the discussion of defining A/A'-positions in Section 3.2 above). In the analysis in (91), topics, including pronouns, that are licensed between sentence adverbs and the verb in ForceP, already belong to the C-domain.

(91) $V_{\text{fin}}\text{-Force}^0$ SHIFTING Topics CONTRAST Topics FAMILIAR Topics **S-Adverbs**
Subject

The above question is in fact an old research issue that readdresses the question of whether subject initial clauses should be analysed as IPs or CPs. Travis (1987) and Zwart (1993) proposed that subject initial clauses are IPs on account of the fact that object initial clauses that must be analysed as CPs are subject to restrictions absent in subject initial clauses, as is illustrated in (92). (92) is problematic for an analysis that assumes that V2-clauses are uniformly represented by CPs.

- (92) Q: What happened?
 A: Der Hans hat gestern die Maria getroffen
 The Hans has yesterday the Maria met
 'Hans has met Maria yesterday'
 B: ?? Die Maria hat gestern der Hans getroffen
 The Maria has yesterday Hans met
 'Hans has met Maria yesterday'

How can we distinguish between these two approaches? The first approach can be discarded immediately since it makes wrong predictions about word order. It predicts that in a sentence with a SHIFTING Topic, no other material can precede this topic. It is very easy to find or construct examples that prove the contrary. (93) is a constructed and completely natural example in which a frame adverbial precedes a constituent that is marked with the typical intonational contour of a (shifting) ABOUTNESS Topic, the L+H* tone. (94a) and (94b) are taken from the corpus. (94a) shows that a focused element marked with the H* tone may precede an ABOUTNESS Topic, while (94b) indicates that a time adverbial *da* may precede a FAMILIAR Topic that continues the topic established in the matrix clause.

(93) Gestern hat der Hans die Maria getroffen.

yesterday have.3SG the Hans the Maria met

L+H*

Yesterday Hans met Maria

(94) a. wir wissen im Westen hier immer alles besser.

we know.1PL in the West here always everything better

H*

L+H*

We here in the West always know everything better.

b. (Putin hat gemerkt) da hab ich mich falsch verhalten.

Putin have.3SG realized there have.1SG I me wrongly behaved

FLAT FAMILIAR Topic

Putin realized in this situation I behaved wrongly.

Assuming that approach B is the correct one, how can we account in this approach for the data in (92)? It is important to note that the V2-constraint can also be fulfilled by formal movement (cf. Fanselow 1993). Formal movement is an operation that allows for the satisfaction of an EPP-feature without being triggered by a specific syntactic or interpretational feature. Formal movement, however, may only attract the closest element within a certain domain, explaining why the subject in (92), as the highest element in clause can satisfy the V2-constraint (that boils down to the requirement of a head that its Specifier is filled with phonological material) without any specific discourse feature, while the object in order to take precedence over the subject needs a specific discourse license, in the case of (92), it has to be given information which in the above context is not indicated.

Now that we have developed an analysis of the C-domain of modern German and given an account of the V2-constraint in a split-CP system (cf. (91)), we are in a position to address issues of word order variation in the C-domain, including the important question of how the V2-constraint came about in the history of Germanic in the following section.

5.2. Verb Placement in the Left Periphery and the (Non-)Development of V2

This section is dedicated to an investigation of the development of the generalized V2-rule in German, in contrast to the maintenance of residual V2 (in questions and negative declaratives) in English.

5.2.1 Variation between V1 and V2 declaratives in OHG

While the V2 pattern was already generalized in the oldest OHG texts in interrogative clauses, we find systematic variation between V1 and V2 orders in declarative clauses in OHG, as is illustrated in (95-98). Taking into account the discourse properties of these sentences, we can distinguish four different patterns relevant for the distribution of V1 and V2 clauses. First, while V1-clauses serve to introduce a new discourse referent and therefore are typically used in presentational sentences in the beginning of texts or episodes, V2-clauses serve to provide additional information about an already established discourse referent that is interpreted as aboutness topic (in the sense of Reinhart 1981) of the sentence, as is illustrated in (95) and schematized in (96).

- (95) a. Et pastores **erant** In regione eadem
uuarun thô hirta In thero lantskeffi (ahd T 35, 29)
 were there shepherds in that area
 ‘there were shepherds in that place’
- b. [ego sum pastor bonus.] bonus pastor / animam suam **dat** pro ouibus suis
 [ih bin guot hirti.] guot hirti / **tuot** sina sela furi sinu scaph. (ahd. T 225, 16f.)
 I am good shepherd. good shepherd gives his soul for his sheep
 ‘I am a good shepherd. The good shepherd does everything for his flock’

- (96) a. FOC[Vfin ... DR_{new} ...] V1
 b. TOP[DR_{giv}] FOC[Vfin ...] V2 *elaboration*

Second, we also find V1-clauses with given discourse referents that introduce a new situation into the discourse, as is illustrated in (97a), and we find an alternative pattern in the same contexts that regularly license occurrences of V1-clauses that involves a V2-clause with the sentence initial adverb *thô* ‘then’ (*thô*+V2), as is illustrated in (97b). The latter two patterns can be schematized as given in (98) below.

- (97) a. **uuard** *thô* giheilit ther kneht in thero ziti (T 84, 7)
 became then healed this boy at this moment
 ‘Then the boy was healed at this very moment’
 & sanatus est puer in illa hora
- b. & ecce homo erat In hierusalem. (T 37, 23)
 senonu tho uuas man In hierusalem.
 and look there was a man in Jerusalem

The demonstrative pronoun in the DP *ther kneht* in (97a) clearly signals that the discourse referent has been introduced in the previous context (note that a definite determiner system has not been grammaticalized yet in this language state).

- (98) a. FOC[Vfin ... DR_{giv}] V1 *narration*
 b. *thô/thar* FOC[Vfin ... DR_{giv/new}] V2

Hinterhölzl and Petrova (2005, 2007) account for the variation between V1-clauses and topic-V2-clauses by showing that their occurrence in the text systematically correlates with the formal expression of two main types of discourse relations in the theory of Asher and Lascarides (2003). Asher and Lascarides (2003) distinguish between coordinating and subordinating discourse relations that play a major role in discourse organization. In Hinterhölzl and Petrova (2005, 2007) it is shown that V1-clauses realize coordinating discourse relations which serve to push the main story line forward on the same discourse level and express a relation of temporal sequence (*narration*), while V2-clauses realize subordination relations which serve to create a hierarchical structure in the discourse. A typical example of subordination is the discourse relation of *elaboration* in which a unit *a* provides further information on a unit *b* situated on higher level of discourse hierarchy.

This implies that verb placement in the left periphery was originally a means of discourse organization, where subordinating discourse relations signalled the establishment of a topic-comment structure in which the V2-clause serves as comment on an already established discourse referent (that is referred to by the clause initial aboutness topic) on a lower discourse level, while V1-clauses signal that a new discourse situation is introduced on the same level of discourse organization.

If we assume that discourse situations are indexed for variables like speaker, speaking time, location and participants, then we predict the occurrence of V1-clauses with the following classes of verbs: *verba dicendi* and verbs of movement, that is, verbs that signal a change of speaker or location. This is the case in OHG, but is also found in the oldest stages of the other Germanic languages (cf. Hinterhölzl, Solf & Petrova 2005 for details). V1-clauses occur with the same discourse functions in Old English (OE), Old Saxon (OS) and Old Norse (ON). This is illustrated in (99) below (for V1 in ON see Donhauser et al. 2006). (99a) is a V1-clause that appears text initially in OE, (99b) displays a V1-clause that signals a change of speaker in OE, while (99c) shows a typical occurrence of a V1-clause in a presentational sentence in OS.

- (99) a. **Nolde** eorla hleo [...] / cwealm-cuman cwicne forlætan (Beo 791f.) OE
 not wanted noblemen_{Gen} protector murderous visitor_{Akk} alive let go
 ‘The protector of the warriors did not wish to let the murderer go alive’
- b. **Spræc**/ ða ides scyldinga (Beo 1168) OE
 spoke then [the] queen [of the] Danes_{GenPl}
 ‘Then the queen of the Danes spoke’
- c. **Lag** thar ên felis bioþan (Hel 4075) OS
 lay there a stone upon
 ‘A stone lay there upon [the entry of the tomb]’

5.2.2 The development of V2 in German

If the distribution of V1 and V2 was ruled by discourse-organizational principles and each of these patterns was associated with one particular, well-defined functional field in the system of early German, then the question arises of how and why this functional opposition was lost in the course of language development and V2 became generalized in main clauses.

I propose that the reason for this development is already present in the system of OHG. Note that V2 has already been generalized in *wh*-interrogatives at the stage of development represented in the Tatian text (cf. Petrova and Solf to appear b). Moreover,

as we have seen above, there is variation in one functional domain of the opposition described for V1 and V2, namely in the domain of the coordinative type of discourse relations. As illustrated in (96a) and in (98b), V1-clauses and V2-clauses with a sentence-initial adverbial, mainly *thô* ‘then’ as a connective marking the coordinative relation to the previous event, co-occur, signalling that there is competition between these two patterns in OHG.

In Hinterhölzl and Petrova (2010) the development of a generalized V2 rule in German in contradistinction to the maintenance of residual of V2 in English is investigated in the light of Rizzi’s (1997) split CP-hypothesis. A possible analysis of this variation in OHG is that the V1-clauses (present with the same discourse functions in all older stages of Germanic) represent the older pattern and that the pattern of subordinating V2-clauses is used as the basis of analogy in a process in which the reference to a previously introduced individual in topic comment structures is extended to reference to the previous discourse situation by the adverbial *thô* in coordinating discourse relations, as is illustrated in (100).

(100) coordination in discourse:

- | | | | |
|----|--|-------|-------------------|
| a. | [V _{fin} ...DR _{new/giv} ...] | FOCUS | (V1) |
| b. | <i>thô</i> [V _{fin} ...DR _{new/giv} ...] | FOCUS | (<i>thô</i> +V2) |

If the use of the pattern (100b) in coordinating discourse relations stabilizes and reaches a certain threshold, it will have two major effects: a) it will blur the formal distinction in the expression of coordinating and subordinating discourse relations and b) it will lead to the information-structural neutralization of the initial position, for the following reasons.

In interrogatives the initial position is occupied by a *wh*-element that is IS-categorized as [+focus]. In subordinating declaratives the initial position is occupied by an element that is IS-categorized as [+topic] and in coordinating declaratives the initial position is occupied by an adverbial that can neither be classified as focal or topical element, that is, is categorized as [-focus, -topic].

This scenario can be summarized in the following way: the V2-pattern was brought about by aboutness topics but the extension of the V2-pattern from subordinating discourse relations to coordinating discourse relations led to the information-structural neutralization of the sentence-initial position, paving the way to the modern German V2-construction: as we have argued in the previous section, the information-structurally defined positions occur below the position of the finite verb in

the Force head with [SpecForceP] being filled with any category capable of satisfying its phonological EPP-feature (see (91) above).

Since the IS-neutralization of the sentence initial position did not lead to the loss of a specialized position for aboutness topics (cf. Frey 2000, Hinterhölzl & Frascarelli 2007) in German, the simplest analysis of verb placement in coordinating and subordinating clauses in OHG in a framework that adopts Rizzi's 1997 split CP-hypothesis is as given in (101) below.

- (101) a. [Force V_{fin} [familiar topics [IP new discourse referents [VP t]]]]
 b. [ForceP Aboutness_i V_{fin} [t_i familiar topics [IP new discourse referents [VP t]]]]

In a V1-clause (101a), the finite verb moves into the highest head in the C-domain, that is, Force⁰ and precedes both given (familiar topics) and new discourse referents. The sentence lacks an aboutness topic and serves to express a new discourse situation on the same level of discourse organization. In a V2-clause, the aboutness topic is promoted into a sentence initial position followed by the finite verb that separates it from the rest of the clause.

Possibly this movement analysis of the aboutness topic in (101b) is the result of the reanalysis, triggered by the neutralization of the clause initial position, of an older more basic structure with a base-generated topic followed by a V1-clause, but about such a stage we lack any empirical evidence in Germanic.

5.2.3 The development of verb placement in English and Low German

While coordinating discourse relations in OE and in OS are expressed by V1-clauses and alternatively by thô+V2-clauses as in OHG, there is a categorial difference between the former two languages and OHG in the expression of subordinating discourse relations, as is illustrated in (102) below. While the finite verb in OHG serves to separate a special kind of topic, namely the aboutness topic, from the rest of the clause that functions as comment on the topic, the finite verb in OE and OS serves to separate all background elements from the focus domain of the clause.

(102) Patterns of Verb placement in older Germanic

discourse relation	syntactic realisation		
	(OHG)	(OS)	(OE)
coordination	i) V1 ii) Partikel+V2 (Part) [_{FOC} Vfin...]	i) V1 ii) Partikel+V2 (Part) [_{FOC} Vfin...]	i) V1 ii) Partikel+V2 (Part) [_{FOC} Vfin...]
subordination	i) Topik+V2 [_{TOP} XP] [_{FOC} Vfin...]	i) Topik+V2 ii) Hintergrund Vspät Fokus [_{BGR} XP ₁ ...XP _n] [_{FOC} Vfin...]	i) Topik+V2 ii) Hintergrund Vspät Fokus [_{BGR} XP ₁ ...XP _n] [_{FOC} Vfin...]

This implies that V2-clauses in OHG on the one hand and V2-clauses in OE and OS on the other hand had a different status in the grammar. V2-clauses in the latter two languages in subordinating discourse relations come about accidentally just in case the sentence happens to contain only one background element. While very little is known about the syntax of OS, this characterisation of V2-clauses fits very well with observations that have been made about the status of V2 in OE by many researchers. Already Kemenade (1987) - although she claims that OE is a V2-language like modern German or modern Dutch - observes that pronouns regularly give rise to V3-orders. Haeberli (2002) observes that V3-orders are not restricted to the occurrence of preverbal pronouns, as is illustrated in (103) and shows that one finds systematic variation between V2- and V3-orders with full DPs.

Bech (2001) and Westergaard (2005) then resolve this puzzle in noting that if the subject is focused or new information, V2 (XVS)-order arises, while when the subject is given V3-orders are found, independently of whether the subject is a full DP or a pronoun. This squares well with Kemenade and Los's (2006) observation that there are two subject positions in embedded clauses in OE, with the higher one (above *thal/thonne*) preferred by informationally given elements, as is illustrated in (104).

(103) *Ðysne yrming æfter his forðsiðe wurðodon þa hæðenan*
 this poor-wretch after his decease worshiped the heathens
eac for healicne god
 also instead-of high God (Wulfstan, 223.58)

‘After his decease, the heathens also worshiped this poor wretch instead of God’

- (104) a. He ne mihte swatheah aefre libban, theah de hi hine **tha** ut alysd
 He not could nevertheless ever live, though that they him then released
 He could not have lived for ever, even though they released him then’
- b. Gif him **thonne** God ryhtlice & straellice deman wile
 if him then God justly and strictly judge will
 ‘if God will judge him justly and strictly then’

These observations suggest that the finite verb in OE (and OS) moves into the C-domain only in clauses expressing coordinating discourse relations, while the verb stays in a deeper position (within the IP-domain) in clauses expressing subordinating discourse relations, as is illustrated in (105).

- (105) a. [_{ForceP} thô V [familiar topics [_{TP} t DR-new]]] coordination
 b. [_{ForceP} [Aboutness [familiar topics [_{TP} V DR-new]]] subordination

The case of OS is very interesting since it patterns with OE in the formal expression of subordinating relations, rather than with OHG. In the course of its history Low German has developed a generalized V2 rule. This development may have been affected by influence from High German. This constitutes a hypothesis which is all the more likely since Low German can not be considered the direct successor of OS, but results from a mix of Saxonian and Low Franconian dialects.

For sure, the development of English has taken a different path that can be characterized as an impoverishment of the C-domain, as proposed in Hinterhölzl and Petrova (2010). What we already see in OE/ME is that the particle/AdvP +V2-construction that replaces V1-clauses also gives rise to V3-clauses. Bech (2001) reports that V3-clauses in OE/EME come in two varieties: a) the pattern XP_{old} XP_{old} V which conforms to (102) and b) the pattern AdvP XP_{old} V in which the adverb may provide new information and which does not conform to any of the patterns in (102).

Here is a scenario of how the second pattern may have come about. Assume that a speaker wants to use a more specific expression instead of *then* to signal a certain discourse relation, say, the adverbial phrase *after the battle*. Furthermore, assume that both *thô* and the aboutness topic in (24) are uniformly interpreted as discourse linkers, then a our speaker that wants to utter a clause with an initial AdvP has two patterns according to which he can construct his sentence: (105a), giving rise to a V2-clause and (105b), where the adverbial phrase occupies the special topic position giving rise to a V3-clause. Preference for the pattern in (105b) would indicate that most speakers

interpret an adverbial phrase like *after the battle* as instantiating a subordinating discourse relation.

In this scenario, we can envisage the following development: Westergard (2005) argues in favor of a development in which English lost the lower subject position. In the present account, we can assume that [Spec,FamP] - after the loss of Case and the failure of object pronouns to appear in this position – was reanalysed as the default subject position (= Spec,TP). The modern English pattern in (106) then follows if topicalized constituents and sentence-initial adverbials are analysed as occupying [Spec, ForceP] on the basis that both of them are interpreted as discourse linkers in non-negative declaratives in English.

(106) [_{ForceP} AdvP [_{TP} Su V...]]

While the above observations about the development of V2 in English are still a mere hypothesis that needs to be confirmed or disproven by empirical investigations, I hope to have shown in the last two sections how the consideration of IS can open alternative research paths to the development of V2 in German and English that offer new interesting insights into processes that are little understood to the present day.

5.3. Word order Variation and IS in the Right Periphery of German

We have seen in section 2 above that the older Germanic languages displayed mixed OV/VO word orders. Furthermore, we have seen in that section that heavy and focussed elements appear postverbally in OE and proposed that it is this factor – a type of stylistic rule – that is responsible for the postverbal placement of event-related adjuncts in modern English, since these adjuncts are typically realized as rather heavy NPs and PPs.

These facts are well-known in the traditional literature about word order in older Germanic. A much quoted point of view is constituted by Behaghel's (1932) observation that pronouns and unmodified nouns tend to precede the verb, while modified nouns, PPs and other heavy material tend to follow the verb that gave rise to the generalization in (107).

(107) Light elements precede heavy elements in OE, OI and OHG.
(Behaghel 1932: *Das Gesetz der wachsenden Glieder*)

At this point the question arises which principle this tendency derives from. This will be one of the objectives in Part III of these studies. In this section, we will have a closer look at word order variation in OHG and show that the generalization in (107) in so far as it is correct, derives from a deeper generalization that involves the interaction between IS and word order.

Hinterhölzl (2010) provides a detailed study of deviations in word order from the Latin origin of the OHG Tatian translation and argues that these deviations can be explained in terms of IS-requirements that point towards an original OHG-system that differs considerably from the NHG-system. In this subsection, I will focus on word order regularities in the right periphery, that is, on the placement of arguments and adjuncts with respect to the verb in embedded clauses.

There is another generalization that emerges from an IS-analysis of the Tatian translation (to be revised below) that derives (107) as a mere corollary. Given that discourse-given elements are typically realized as light elements while focused constituents may count as prosodically heavy elements since they receive stress, (108) derives the tendency expressed in (107).

(108) C background V focus

The generalisation in (108) has the nice property that it allows us to abstain from positing the existence of more than one grammar. With it we can account for word order variation within one grammar by taking into account the information-structural contribution of a constituent in the discourse: for instance, a direct object will precede the verb (in embedded clauses) if it is discourse-given but will follow it, if it is discourse-new.

According to the generalisation in (107) a direct object will be placed preverbally, if it is realized as a pronoun or single noun, but postverbally, if it is made heavy by modification. As I have indicated above, the two conditions are not independent of each other, but it would be interesting to see which one of them is more basic (see also Hroarsdottir (2004) for similar observations in Old Icelandic).

In the following, I will analyse the information structural role of constituents that have been either preposed (into the middle field) or postposed into the postverbal domain in embedded clauses. Needless to say that preposed heavy constituents and non-preposed light constituents are of particular interest here. At this point we must make precise what is meant with a prosodically light or heavy element. For reasons that will

be made explicit in Section 6.2 we assume here that a non-branching constituent (a pronoun, a single noun) counts as prosodically light while a branching constituent counts as prosodically heavy. The following small scale empirical investigation (in Hinterhölzl to appear b) is based on Dittmer and Dittmers (1998) study of word order regularities in the Tatian translation. Their study comes handy for our purposes, since they have categorized and listed all cases of deviations from the original Latin text.

According to Dittmer and Dittmer (1998), there are 142 cases in which a constituent is moved from the postverbal field into the preverbal field in an embedded clause. The majority of these cases involves a pronominal subject or object that can be equally well analysed as prosodically light and as discourse anaphoric. Among the 30 cases with a nominal subject, there are 23 cases that involve a branching constituent that must be analysed as heavy (the rest are biblical names like *David*, *Herodes*, *Johannes*, etc.). These heavy subjects, however, are to be analysed as belonging to the background and involve combinations of a determining pronoun and a common noun like *ther heiland*, *min fater* and so on. Two cases are illustrated in (109). In both cases, *thas magatin* and *thiu fluot* have been introduced in the previous discourse.

- (109) a. ubi erat **puella** iacens
 thar **thas magatin** lag (T 96, 25)
 where that girl was lying
- b. donec venit **diluuium**
 unc **thiu fluot** quam (T 257, 7)
 until that flood came

It is interesting to note that the number of preposed objects is relatively small compared to the number of preposed subjects. Dittmer & Dittmer (1998) only list 8 such cases. Two cases involve the demonstrative pronoun *then*, translating the Latin personal pronoun *eum* ('him'), which is necessarily discourse anaphoric. Among the remaining 6 cases, 3 involve the preposing of a (light) noun and 3 involve the preposing of a (heavy) noun phrase. The latter, however, are mentioned in the previous discourse.

If, in turn, we investigate the deviations in which a constituent is postposed from the middle field, then it must be noted that among the few 10 cases to be found (in very few cases a middle field can be identified in Latin), there are 7 cases that involve a non-branching constituent which cannot possibly be analysed as prosodically heavy. However, if we consider the information-structural role of these constituents, we note that they must be analysed as being focused, as is illustrated in (110).

- (110) Haec locutus sum vobis thisu sprahih iu
 ut in me **pacem** habeatis thaz in mir habet **sibba** (T 290, 10)
 in mundum **presuram** habebitis in therru weralti habet ir **thrucnessi**
 ‘this I tell you that in me you-have peace, in the world you-have unrest’

In (110), the non-branching direct object *sibba* is postposed, while the branching PP remains in the middle field. The information structural analysis, however, reveals that this configuration is to be expected, since we are dealing with contrastive statements in which the PPs represent contrastive topics and the direct objects provide the corresponding focused constituents.

The example in (111) shows that word order in the Tatian translation cannot be explained with the simple rule that pronominals are preposed while DPs according to the Latin origin remain in their base position. In (111), a weak pronoun has been placed in the postverbal domain against the Latin origin. However, the context makes clear that the postposed pronoun is focused. The text passage requires that the pronoun be stressed. While modern German would use a focus particle (to strengthen the reflexive pronoun that cannot be stressed) – *wer sich selbst erniedrigt, der wird erhöht werden* – in OHG this focus is expressed syntactically by moving the pronoun into a right-peripheral focus position.

- (111) Quia omnis bidiu uuanta iogiuuelih (T 195, 16)
 qui se axaltat thiedar sih arheit
 humiliabitur uuiridit giotmotigot
 qui **se** humiliat inti therdar giotmotigot **sih**
 exaltabitur wirdit arhaban
 ‘therefore everyone that exalts himself will be humiliated but the one that humiliates himself will be uplifted’

In conclusion, we have seen that Behaghel’s law is fulfilled if the syntax including information structure does not impose more specific restrictions on its own. In this case, light elements are moved into the preverbal domain, while heavy elements move to or remain in the postverbal area.

This conclusion is further strengthened by the following observation. A careful analysis of the data in the Tatian translation also reveals that the generalisation in (108) needs to be refined in as much as contrastive foci are concerned. While new information

focus is typically realized in postverbal position, as we have seen above, contrastive focus is realized in preverbal position, as is illustrated in (112) and (113).

- (112) tu autem cum ieiunas/ unge caput tuum/ & faciem tuam laua/ ne uideatis
hominibus/ ieiunans. Sed patri tuo
thane thu fastes/ salbo thin houbit/ Inti thin annuzi thuah/ zithiu thaz thu
mannon nisís gisehan/ fastenti. úzouh **thimeno fater**
[when you fast, do not be like the hypocrites ...]when you fast, anoint your head
and wash your face so that you do not appear to men to be fasting but to your
Father

Note that in (112) the contrastive element is placed preverbally against the order in the Latin origin signifying that we are dealing with an independent requirement of OHG. In this context it is interesting to note that PPs that are placed predominantly postverbally in accordance with Behaghel's law (due to their heaviness) appear preverbally when contrastively focussed, as is illustrated in (113). The PP in (113) comprising three words is rather heavy but nevertheless appears preverbally due to its interpretation as a constituent bound by a focus operator like 'only'.

- (113) orantes autem. nolite multum loqui/ sicut &hnici.'/ putant enim quia **in multiloquio**/ exaudiantur.
b&onte nicur& filu sprehan/ sósó thie heidanon mán/ sie uuanen thaz sie **in iro filusprahhi** / sín gihórte T(67, 23-26)
'And when you pray, do not use vain repetitions as the heathens do. For they think that they will be heard (only) for their many words'

Therefore, the generalization in (108) that explains (in tandem with Behaghel's law) the variation in word order found in the Tatian translation needs to be slightly revised, as given in (114).

- (114) C background contrastive focus V information focus

In this respect, OHG seems to pattern with Yiddish, which is a West Germanic language that has preserved mixed word orders. According to the description of Diesing (1997), contrastive foci pattern with background elements in occupying a preverbal position.

To summarize, we have seen that mixed word orders can be accounted for easily within one grammar. We have shown that word order in the right periphery of OHG is predominantly determined by IS-requirements. Several questions, though, remain open at this point. First, the question arises of where Behaghel's law - to extent that it is relevant for guiding word order variation - can be placed in a minimalist grammar. As we will see in Section 6.2, it corresponds to a mapping condition between syntactic structure and prosodic structure. Second we would like to know from which principles of grammar the generalisation in (114) can be derived from. These questions will be addressed in Part III, where we will also develop a model of language change based on the complex interaction between syntax, prosody and information structure within which the development of strict word order regularities in Germanic can be described and accounted for.

Part III The complex interaction between syntax, prosody and IS

In this final part, I will investigate different issues pertaining to the interaction between syntax, prosody and information structure and develop a model of language change within which the pertinent word order changes can be described. In this way, the theoretical question of how these modules interact with each other - that is important by itself - is supplemented by the empirical demand that the resulting model must be such that it can help us understand the grammatical developments in question.

6. Word Order Change and the Distinction between Core Grammar and Periphery

In this section, I will develop a model of language change that is apt to account for word order change, including changes in basic word order, in the absence of a head-complement parameter. More specifically, I will investigate the following two questions in the light of the change from mixed OV/VO-order to pure VO-order in the history of English and before the background of the development of the German sentence bracket: A) How can we explain that language change often proceeds in a slow, gradual fashion, extending over long periods of time in which the old and the new pattern coexist side by side? This fact conflicts with the idea that grammar change via parameter resettings should give rise to abrupt changes.

B) What is the nature of word order change? In a framework subscribing to the UBH, changes in word order cannot be relegated to changes in the head complement parameter. These two questions, which are discussed in detail in Hinterhölzl (2004b), will be dealt with in the following two subsections.

6.1. Grammar change versus language change

Diachronic processes are often characterized by a slow gradual change that eventually is followed by an abrupt change that marks the new parameter settings. Typically, before the innovative pattern replaces the old pattern, there can be a relatively long period of time in which both patterns are used in the language.

Examples of diachronic processes that proceed in this way actually abound. Take, for instance, the slow rise of VO-orders in the history of Icelandic. As Hroarsdottir (1998) has shown, OV and VO orders coexisted in Icelandic for many centuries. A

similar case is the slow decline of V-XP orders in embedded clauses in the history of German. This process starts in the 9th century (cf. Näf 1979 who diagnoses a major decline of postverbal accusative objects in the writings of Notker), but the declining pattern is still available in the present day language in restricted contexts. Another example that does not come from the field of word order is the slow rise of expletive subjects in the history of Swedish, as it has been described by Falk (1993).

Under the standard assumption that reanalysis and parameter resettings give rise to abrupt change (cf. Lightfoot 1991, 1999), the period of gradual change must be viewed as a period of language change without a change in grammar which sets the stage for the abrupt change instantiating the grammar change proper.

This scenario, plausible from an empirical point of view, raises the question of how language change is possible without a change in grammar. One approach that pays attention to the gradualness of change and the parallel availability of two (or more) alternative forms is the concept of grammar competition (cf. Kroch 1989, Pintzuk 1996). This approach was proposed to account for the development of English from an OV- to a VO-language, but clearly has applications beyond this particular historical process and beyond word order changes in general.

An alternative would be to assume change in language use within one grammar. In order for adult speakers to be innovative in their language use within the limits of the grammar they acquired, we need to assume that a grammar defines a certain space of limited options that can be exploited by innovative speakers to achieve certain communicative effects in specific discourse situations. Hinterhölzl (2004b) proposes that the area of limited optionality pertains to the field of stylistic form, more particularly to what has been called information packaging (cf. Vallduvi 1992), that is, to the arrangement of information according to features like new/old or prominent/non-prominent and the like. Stylistic or information-structural rules (in short IS-rules) involve mainly alternation of word order and alternatives in prosodic phrasing, as we will see below.

To delimit the space of restricted options in grammar, Hinterhölzl (2004b) reinvoles the old distinction between the core grammar and the periphery and the notion of markedness as the distinguishing criterion (cf. Chomsky 1965). What is derived in the core grammar is unmarked and stylistically or information-structurally marked forms are derived by peripheral /stylistic rules.

Having introduced the distinction between core rules and peripheral/stylistic rules in an intuitive way, the question arises whether the distinction can be made in more general terms. The standard notion of a stylistic rule so far has been that they apply

postcyclically, after the core derivation (hence also the name of a peripheral rule), that is, at PF. As PF-rules, stylistic rules may change the word order and alter the prosodic form of the syntactic output. Thus, it seems that relegating these rules to the PF-component derives the relevant properties that we would like to ascribe to stylistic rules (cf. Zubizarreta 1997, Szendrői 2001).

However, a closer inspection reveals that this conception of “peripheral” rules is untenable. The case of stylistic fronting in Icelandic reveals that this rule cannot apply at PF, since it is subject to a syntactic restriction, namely that the subject position be empty (cf. Maling & Zaenen 1990). Since syntactic information cannot be taken to be available at the PF-interface, the effect can only be obtained if the rule applies in narrow syntax. This implies that so-called peripheral rules, a misnomer as it turns out, apply in the core derivation as well and that the distinction between core rules and stylistic rules should possibly be made on grounds of the type of features that they check respectively – morphosyntactic features on the one hand and information-structural features on the other hand. Hinterhölzl (2004b) explores the question of whether the distinction between core and periphery can be made in terms of prosodic information and argues on the basis of historical data that a peripheral/stylistic rule can be defined as a rule that leads to a marked prosodic output.

6.2. What is the Nature of Word Order Change?

Hinterhölzl (2004b) addresses the change of word order that occurred in the transition between the Old English period (OE) and the Middle English period (ME). The standard answer to the question of what the pertinent change consists in is that there has been a change in the head complement parameter. This analysis is based on two assumptions: A) OE was an OV-language akin to modern Dutch or German and B) the base order changed from OV to VO. The first assumption we have already shown to be wrong: In section 2.4, we have shown that OE was a language with mixed OV/VO orders. In the traditional literature, these state of affairs is accounted for by generalizations according to which the distribution of arguments and adjuncts is determined crucially by stylistic factors in the older stages of Germanic, such as Behaghel’s law, repeated for convenience in (115).

(115) Light elements precede heavy elements in OE, OI and OHG.

(Behaghel 1932: *Das Gesetz der wachsenden Glieder*)

Within the generative tradition, word order variation in OE and EME is recently modelled in terms of competing parametric choices. One such approach is the so-called Double Base Hypothesis (Pintzuk 1999) according to which word order variation follows from the co-existence of competing grammars that differ with respect to the head parameter of VP and IP. One of the central assumptions of the double base hypothesis which is based on the idea of grammar competition is that VO-orders are an EME-innovation that was brought about by language contact between Anglo-Saxons and the Scandinavian settlers in the 10th century. But we have seen before that the grammar of OE already contained a number of VO-features. Second the assumption that the grammar of the Scandinavian settlers was (already) VO is equally unwarranted: There is simply no evidence that the grammar of the Scandinavian settlers in England was predominantly VO. The oldest written Old Nordic documents are Icelandic texts from the 12th century. These documents show a preponderance of OV-orders¹⁰. A recent corpus studies of OI confirms the observations that have been made by traditional grammarians (cf. Hroarsdottir 2002). A) There is variation between OV and VO structures and B) the distribution of syntactic elements is determined by stylistic/information-structural factors. Hroarsdottir (2002) shows that the preverbal or postverbal occurrence of DP-arguments is determined by the interaction of two factors: the heaviness of the argument and its discourse status (whether it represents old or new information).

As we have seen in the previous section, a similar picture arises for OHG. Also OHG displays mixed OV/VO word orders with the stronger factor, however, being information structural requirements that warrant a strengthening of Behaghel's observation in terms of information structural categories, as repeated in (116).

(116) C background V focus

¹⁰ One may argue that the presence of OV-orders as such (in OI) does not necessarily constitute a problem for the language contact scenario of OV-to-VO in English, the rationale being that these orders are derivable from a VO base by leftward movement operations. If the learner would be able to detect this VO base, this might still influence his grammar (that develops in a bilingual environment), so the argument goes. But this begs the question. In this line of reasoning, also the grammar of OE would have contained evidence for a VO base. Without any further argument as to why the contact situation in general would favour VO-orders (rather than OV-orders) or as to which feature of OI would favour a VO analysis, the contact scenario seems mute to me.

(116) expresses the hypothesis that in OHG, which did not have a determiner system or was about to grammaticalize one, the discourse status of an argument was signalled by its positioning with respect to the verb. Hroarsdottir (2002) argues that similar restrictions obtain in OI, while word order in OE seems to be guided predominantly by prosodic restrictions pertaining to the weight of a constituent. In the following sections, we will investigate what the nature of these restrictions is and how they interact in order to derive the observed mixed word order patterns in OHG, OE and OI.

Summing up, it is unwarranted to assume that VO-orders came about in the history of English by language contact with Scandinavians, since we find both OV- and VO-orders in the older stages of German and Icelandic as well, suggesting that the variation in argument positioning in OE, OHG and OI is simply part of the common Germanic heritage.

The second assumption can be shown to be insufficient in so far as there are differences between German and English that cannot possibly be subsumed under the head complement parameter and that point towards a more global distinction that involves the syntax-prosody interface, as we will see in Section 6.

There are two other properties that distinguish German and English and that seem to be correlated with the head complement parameter. One, that we have already discussed in Section 2.4 above, involves the placement of event-related adjuncts in the two languages, as is illustrated again in (117). These adverbs occur preverbally in the order T>P>M in OV-languages but postverbally in the exact mirror image in VO-languages (cf. Haider 2000, Hinterhölzl 2002).

- (117) a. C T P M-V OV-languages
 b. C V- M P T VO-languages

The second property also concerns adverbs in VO languages. A property that is in need of an explanation is the fact that adjuncts that can occur between the subject and the VP in VO-languages are subject to restrictions absent in OV-languages, as is illustrated in (118) and (119).

- (118) a. John (more) often (* than Peter) read the book
 b. Hans hat öfter (als der Peter) das Buch gelesen
 Hans has more-often (than Peter) the book read
 ‘Hans read the book more often than Peter’

Descriptively speaking, the head of the adjunct must not have material to its right. This is only possible if the adjunct appears in sentence final position. An option, on the other hand, that is not available in OV-languages, as the contrast illustrated in (2) shows.

- (119) a. John read the book more often than Peter
 b. *Hans hat das Buch gelesen öfter (als Peter)
 Hans has the book read more-often than Peter
 ‘Hans has read the book more often than Peter’

From (118) and (119), it can be concluded that only light adverbs (of a restricted class that have been called adverbs of indefinite time by Ellegard 1953) can appear in the middle field in English. This property can of course not be derived by the head complement parameter.

This distinction between English and German may seem to be just one of many stylistic differences between the two languages, but I will argue that it is part of a larger development that also affected the order of object and verb and can be given a uniform explanation. The restriction observable in (118-119) is reminiscent of Behaghel’s law in (115): the English middle field does not tolerate heavy material. This condition can now be made a bit more precise: the English middle field does not tolerate phrases that branch to the right, while the German middle field tolerates heavy, right-branching constituents. This property, I will consider the main difference between English and German that accounts for the different distribution of both arguments and adjuncts in the two languages. Therefore the main objective of a diachronic account must involve investigating how these grammatical characteristics of German and English came about. This question is addressed in the following three subsections.

6.3. Peripheral Rules and Word Order Change in English

To account for mixed word orders in OE, Hinterhölzl (2004b) proposes a VO-based grammar plus licensing movement of arguments and VP-internal predicates to designated positions in the middle field as outlined in Section 1 above. Licensing movement will derive all the OV-properties of OE, as is illustrated in (120). (120a) is derived from the VO-base structure (120b) in a cyclic fashion by XP-movement of the particle into its licensing position in [Spec, AspP] (120c) and by movement of the arguments into Case-licensing positions (120d).

- (120) a. Ðæt he his stefne up ahof
 that he his voice up raised
 ‘that he spoke up’
- b. [_{CP} Ðæt [_{VP} he ahof his stefne up]]
- c. [_{CP} Ðæt [up [_{VP} he ahof his stefne]]]
- d. [_{CP} Ðæt [he [his stefne [up [_{VP} ahof]]]]]

To account for the apparent VO-properties of OE, Hinterhölzl (2004b) assumes optional movement of an XP containing the verb to a medial position (light predicate raising). Thus, an OE sentence like (121a) is derived within a grammar that yields basic OV-structures (121b) by successive application of light predicate raising (LPR) and V2 (verb movement to the Force head in the C-domain), as illustrated in (121c).

- (121) a. Ða ahof Paulus up his heafod (van Kemenade 1987:33)
 then raised Paulus up his head
 ‘then Paulus lifted his head (up)’
- b. [_{CP} Ða [_{IP} Paulus his heafod up ahof]]
- c. [_{CP} Ða ahof_i [_{IP} Paulus [_{VP} up t_i] his heafod]]

If OE had a rule of LPR that is triggered by the requirement that the middle field may not contain right-branching material, then it is to be expected that event-related adverbs, which are primarily realized as right-branching NPs and PPs appear preferably postverbally in OE. In section 2 above, we have shown that the number of postverbal event-related adjuncts rises in EME and in turn triggered the rise of the postverbal occurrence of arguments. What is still missing is a general model of word order change that accommodates these developments.

In section 2.3, I have argued that base order and unmarked word order are to be distinguished and that the unmarked word order in the phrases of a language is determined by the predominant, that is, unmarked prosodic pattern in those phrases, defined by the headedness of phonological phrases. If a language displays only one prosodic pattern, this pattern will define the unmarked word order. If a language displays several prosodic patterns, the predominant pattern or the predominant patterns will define the unmarked word order(s), which have to be derived in the core grammar by obligatory movements, while less frequent patterns will be relegated as marked word orders to the periphery, to be derived by stylistic rules.

Old English must be analysed as exhibiting various marked and unmarked word orders. OV-orders would give rise to the prosodic pattern (s w), with the direct object forming the prosodic head of the phonological phrase comprising the verb, and VO-orders giving rise to the prosodic pattern (w s).

The frequency component within the definition of a marked prosodic pattern is important for explaining language change. If a peripheral rule like LPR is extended and more and more widely used (when it becomes popular), it will be less and less marked and may, when crossing a certain (statistical) threshold give rise to a new or an additional unmarked word order. When this happens a peripheral rule will be reanalysed as a rule of the core grammar. In other words, LPR, that is, VP-intrapolation will become obligatory triggering the need for a new stylistic (peripheral) rule that optionally moves light (non-branching) adjuncts into the middle field in Modern English (cf. (118) above).

This model of language change in terms of the distinction between core grammar and periphery can account for the gradualness of word order change and can explain a change in unmarked word order without invoking word order parameters. The model relies on optional rules that are triggered by prosodic requirements (no heavy elements may occur in the middle field). In the following two sections, we will investigate whether this model can also explain the pertinent word order change in the history of German.

6.4. Modelling Word Order Variation in OHG: The Grammar of Focus

We have seen in section 4 that word order in OHG is predominantly determined by information-structural requirements and obeys the generalization in (116) above. Furthermore, we have seen that word order in OHG seems to be obey similar constraints that have been reported to hold in Yiddish. Diesing (1997) claims that definite and specific indefinite NPs in Yiddish precede the verb while non-specific definites follow it. In so far, as definites and specific indefinites can be taken to be background (given) elements and indefinites to be focused, this confirms our hypothesis. However, there is a third group of elements which disturb this picture. Diesing claims that arguments that are contrastively focused also precede the verb. As we have seen in section 4, the same is true in OHG, yielding the generalisation in (123). Three notions of focus are relevant for the generalization in (123): broad and narrow presentational focus and contrastive focus. These are illustrated in (122). In (122), brackets mark the focussed constituents

and capital letters mark a high pitch accent, which is typical for contrastively focussed elements.

- (122) a. What did John do? (broad presentational focus)
 John [gave a book to Mary]
 b. What did John give to Mary? (narrow presentational focus)
 John gave [a book] to Mary
 c. John gave Mary [a BOOK], not a pen (contrastive focus)

(123) C background contrastive focus V presentational focus

Hinterhölzl (2004b) proposes that the distribution in (123) can be implemented into the following phrase structure representation. According to the UBH, arguments move out of VP to be licensed in (Case-) Agreement positions. A structural focus position is located above these licensing positions. The word order facts in (123) then follow from the following assumptions about the syntax of focus: The verb moves into the head position of a Focus phrase (FocP). Contrastively focussed constituents move into [Spec,FocP]. Constituents that are part of the presentational focus remain in the scope of the head of FocP. Background elements move out of the scope of Foc into licensing positions in the C-domain. This is illustrated in (124).

(124) [C background [FocP ContrastF V [Agp PresentationF [VP]]]]

It is interesting to see that the grammar of Yiddish and OHG implements a syntactic distinction between contrastive and non-contrastive, that is, presentational focus. Note in this respect that prosody makes a distinction between narrow and broad focus. Constituents that are broadly focused do not restructure with the verb. They are mapped into prosodic constituents according to the syntactic structure alone, while narrowly focussed constituents integrate into the phonological phrase of the verb (cf. Nespor & Vogel 1986, Frascarelli 2000).

Given the generalisation in (124), the development of German must be described as one in which the postverbal focus position is lost. In the given model, we can envisage the following development. If speakers of late OHG, for whatever stylistic purposes, moved more and more narrowly focussed constituents into [Spec,FocP], then there will be a greater number of phonological phrases that, due to focus restructuring, instantiate the pattern (*s w*), with *s* being the focused element and *w* being the verb.

Since according to the rhythmic activation principle, unmarked order is determined on the basis of the metric structure of phonological phrases (cf. Nespor, Guasti & Christophe 1996), such a stylistic rule, if it catches on and reaches a certain threshold, will have a profound effect on what counts as unmarked or marked word order in the language.

More specifically, we may assume that if the application of the stylistic rule crosses a certain threshold, it will cause the IS-prosodic parameter in (125) to be set to the value (F>V), that is, Focus restructures to the right. If this happens, postverbal focussed constituents will be literally exbraciated prosodically: They are excluded from the intonational phrase containing the verb and will thus be highly marked and be useable only for certain stylistic effects.

(125) IS-requirements on the prosodic mapping of syntactic structures:

A focussed constituent introduces an IP-boundary on one side and restructures with the adjacent verb on its other side (generalized from Frascarelli 2000).

Harris & Campbell (1995) in their comparative survey of types of word order changes in the languages of the world, point out that it frequently happens that an originally pragmatically defined position provides the basis for the fixing of word order through reanalysis (p. 235). For example, Comrie (1988) argues that an earlier stage of Armenian placed focussed constituents in preverbal position and that in Modern East Armenian this word order has been fixed as a base order.

Thus, there is crosslinguistic evidence for the process of the neutralization of a focus position that is envisaged in Hinterhölzl (2004b). In other words we have developed a general model for word order change without a change in parameter and outlined the grammatical mechanism for such a development: namely prosodic bootstrapping in the acquisition of unmarked word order and focus restructuring.

To summarize, I have argued that the head-complement parameter is dispensable. I have shown that German and English underwent changes that cannot be subsumed under the head-complement parameter. Instead, I have sketched an account in terms of the distinction between core grammar and peripheral rules embodying a dynamic concept of (prosodic) markedness and outlined the conditions and factors that led German and English to develop from a common source into opposite directions.

Along these lines, German can be claimed to be a VO-language, whose unmarked word order happens to be OV, because the IS-prosodic parameter was fixed to F>V. Finally, I have argued that in the acquisition process, prosody helps the child to decide

which structures belong to the core grammar and which ones are peripheral. This entails that the child acquires the prosodic patterns of its language first and ranks them in their markedness according to their frequency.

Though this model is promising and the above account is theoretically sound, it is important to note that from an empirical point of view our account of the development of the German sentence bracket remains unsatisfactory as long as we do not identify a plausible factor that led to the preference of preposing narrowly focused constituents.

6.5. The Development of the German Sentence Bracket: a Case of Stylistic Change?

As already mentioned in section 5.1, postverbal occurrences of arguments and adjuncts were slowly diminishing in the history of German. The process has started in the middle of the 9th century, where scholars have diagnosed a sharp decline of postverbal accusative objects in the later texts of Notker (cf. Bolli 1975, Näf 1979 und Bortler 1982) and has not been fully completed till today. To illustrate the development, let us look at the statistics of Kavanagh (1979), cited by Lenerz (1984), who studied the development of the verb final position, traditionally called the sentence bracket, in German. His counts show that the verb final pattern has risen from 50% occurrence in OHG, via an average number of 80% in the period of Middle High German (MHG) to more than 90% in Modern German.

If this statistics is correct, it shows two things. First, as we have argued that OE cannot be considered a pure OV-language, 50% verb final patterns indicate that German in its oldest accessible stage cannot be considered an OV-language either. Secondly, the change was rather slow and gradual and we may assume that starting from the MHG period, V XP-orders represented marked options in the grammar.

However, the evaluation of word order regularities in OHG is a quite intricate and difficult task since the majority of texts are translations from Latin. Thus many researchers consider only those examples which deviate from the Latin word order as revealing real properties of the grammar of OHG. Based on this rationale, it was argued that already the earliest OHG records display much stronger OV properties than comparable OE texts (cf. Robinson 1997, Dittmer & Dittmer 1998, Fuss 2002).

It is interesting to note but compatible with the above rationale that while the change from OE to Modern English word order has been described as a grammatical

change (a change in the head complement parameter), the inverse change in German has been characterized as mere stylistic in the German traditional literature.

Also Lenerz (1984) states that with respect to the development of verb positioning from OHG to ModG, no syntactic change in the strict sense has taken place: What has changed is the functional relation of certain verb orders to specific pragmatic types. He claims that the changes that you find should be treated as the variation of parameters which are not part of the core grammar (p.126).

He backs up his position with the claim that in general the relative proportions of extrapositionable constituents have remained constant over the whole period. Citing Kavanagh (1979), he states that a) adjuncts are most frequently postponed and arguments least frequently and b) longer and heavier constituents are more easily postponable than short and light ones. Furthermore, he argues that the change cannot be a categorical one, since even arguments can still be postponed in Modern German, as is illustrated in (126).

- (126) Auf Gleis 5 fährt ein | der IR nach Straubing
at platform 5 comes in the Interregio to Straubing

The question only is whether this (traditional) picture of the development of German is correct. I think it is at least incomplete. Note also that the study of Robinson (1997) reveals that certain embedded clauses display only 50% verb final order: “in consecutive and purposive sentences it (non-final word order) occurs 12 times out of 24”(p.83). If OHG was a simple OV-language, this result is rather surprising. Secondly, Robinson (1997) also points out that “predicative elements (predicate adjectives, nouns, prepositional phrases, etc.) regularly follow any non-finite element at the end of the clause.” Postverbal predicates in Modern German are completely ungrammatical in any context and style. Thus, if the development of the sentence bracket in German, was simply a matter of stylistic change, this fact cannot be explained.

Our findings about word order in the OHG Tatian text (cf. Section 4) confirm this judgment: OHG cannot be treated as a simple OV-language. Let us now return to Lenerz’s argument that extrapositioned arguments in Modern German are indicative of a mere change in style rather than of a change in grammar proper. The sentence in (126), which involves a postponed subject, is okay, but it is very rare and highly marked in modern German, as the corresponding prosodic structure in (127) indicates. The sentence with a postponed subject cannot be phrased with just one intonational phrase. Hence it is marked since DPs normally do not form intonational phrases on their own.

As a result, the focussed subject may not be contained in the intonational phrase that contains the verb and thus violates the prosodic condition in (128).

(127) [_{IP} (Auf Gleis 5) (fährt ein)] [_{IP} (der Interregio) (nach Straubing)]

(128) Focus constituents are mapped into the intonational phrase which contains the verb. (Nespor & Vogel 1986)

Based on the above discussion, I conclude, contrary to Lenerz and the tradition, that there was indeed a change in grammar, that involves the prosodic mapping of focussed constituents. In modern German, only a preverbal focused constituent can integrate into the intonational phrase containing the verb, while postverbal focus was possible in OHG. The question that arises is which factor led to this loss of postverbal foci and is addressed in Hinterhölzl (2009a).

Given the discussion of the previous section, the development of German is to be described as one in which certain factors led to the non-application of the prosodic constraint in the German middle field allowing for the spell-out of branching constituents in the preverbal domain, while the development of English is to be described as one in which certain (other) factors not only led to the retention of the prosodic constraint in the middle field but also caused light (non-branching) constituents like pronouns to be spelled out in the postverbal domain.

I will concentrate here on discussing the potential factors that can help us understand the development of German. As far as the development of English is concerned, I will only discuss it in so far as to make plausible why the pertinent factors working in German did not have the same effect in the history of English.

Given the presence of both a preverbal and a postverbal focus position, we can assume that OHG had two unmarked word order patterns: due to focus restructuring, preverbal focus will give rise to prosodic phrases of the type (s w) and postverbal focus will give rise to prosodic phrases of the type (w s). Hence, if we can identify a factor that led to the occurrence of more and more focussed (stressed) preverbal constituents, then we can envisage a slow development in which certain word order patterns become more and more marginalized to a point where postverbal stressed constituents are prosodically highly marked such that they can only be used for specific communicative purposes.

In the following, I would like to argue that this factor was the grammaticalization of the definite determiner. Following this line, we have to consider a development from

a system in which noun phrase interpretation was signalled by syntactic position, as in the OHG Tatian text (cf. (116) above), to a determiner system which signals noun phrase interpretation with morphological means.

First, note, that the morphological system that became grammaticalized is not congruent with the distinction between given and new discourse referents (that was embodied in the syntactic system in (116)), since the definite determiner came to signal that the discourse referent - whether given in the discourse or discourse-new - is uniquely identifiable in the context. Thus, the development of this morphological determiner system has blurred the original information-structural distinction of (nominal) arguments.

The original distinction between definites, however, is still visible in the scrambling rule of modern German. In section 2, I discussed Haider & Rosengren's (1998) take on scrambling in which they note that neither definiteness nor specificity can be taken to be triggers of scrambling, citing examples like (129), where a specific indefinite NP in (129a) and a definite NP in (129b) seemingly occur in their base positions.

- (129) a. wenn wer eine rothaarige Frau sucht dann ist das Maria
 if someone a red-haired woman seeks then is it Maria
 'if some is looking for a woman with red hair, then it must be Maria'
- b. dass er wem ihr Kleid gezeigt hat, hat Maria nicht gefallen
 that he someone her dress shown has, has Maria not pleased
 'that he has shown her dress to someone did not please Maria'

Note, however, that their examples only involve NPs that must be characterized as discourse-new. This shows again that definiteness is not a trigger for scrambling. But, as we have argued in Section 3, specificity can be taken as the relevant trigger for scrambling, if it is understood as membership in a set that is given in the discourse (cf. Enc 1991). In conclusion, only discourse-given DPs may scramble; DPs that are discourse-new, even when they are uniquely identifiable (by binding or bridging, for instance) stay in the focus-domain and receive stress (cf. Hinterhölzl 2004b)

But why should the grammaticalization of a definite determiner have an influence on the unmarked word order? There are three observations that fit nicely into an explanatory pattern of this kind. First, recent studies of the grammaticalization of the definite determiner (cf. Oubouzar 1992, Demske 2001) agree that the definite determiner - derived from the demonstrative pronoun - first appears in contexts with

pragmatic definite interpretations. They refer to Löbner (1985) who distinguishes between pragmatic and semantic definites. While semantic definites are uniquely identifiable on the basis of their lexical meaning (functional concepts), pragmatic definites only become uniquely identifiable in the contexts they are given in.

The grammaticalization of the definite determiner starts in the earliest OHG-texts and is concluded in Notker (early 11th century), in which text all semantic groups of nouns (including abstract nouns and uniquely referring expressions) appear with the definite determiner. The slow stepwise process of grammaticalization can be exemplified by the occurrences of the definite determiner in Otfrid. In this text that is placed between the Tatian translation and the writings of Notker, the definite determiner regularly appears with pragmatic definites, that is, in discourse-anaphoric uses, as is illustrated in (130), while with semantic definites the determiner is still missing, as is illustrated in (131). Examples are taken from Demske (2001).

- (130) a. ein burg ist thar in lante... (O.I 11.23)
 a town is there in (the) country
 ‘there is a town in the country’
 zi theru steti fuart er thia druhtines muater (O.I.11.26)
 to this city leads he the Lord’s mother
 ‘he lead the Lord’s mother to this city’
- (131) a. tho ward himil offan (O.I.25.15)
 there was (the) sky open
 ‘the sky was open then’
- b. inti iz hera in worolt sante (O.I.13.5)
 and it here into (the) world sent
 ‘and he sent it into the world here’
- c. in ira barm si sazta [barno bezista] (O.I.13.10)
 in her lap she placed child most-beloved
 ‘she placed her beloved child on her lap’

The second observation concerns another statement by Behaghel (1932), given in (132), who notes that noun phrases with a determiner are introduced preverbally.

(132) Behaghel (p. 79): "Substantiva mit Pronomen stehen auf der Seite der einfachen Wörter; zum Teil mag das daher rühren, dass ihnen der Artikel früher fehlte."

'Nouns with determiner are to be found on the side of single nouns; the reason may be that initially they lacked a determiner.'

Behaghel's observation is in line with the stepwise grammaticalization from pragmatic definites to semantic definites, since discourse anaphoric NPs are typically realized preverbally in OHG (see (116) above). We now can envisage a development in two stages, as sketched in (133).

- (133) stage 1: the determiner is introduced for discourse-given referents preverbally
 stage 2: the use of the determiner is extended to all uniquely identifiable NPs, which are placed according to the pattern in phase 1 preverbally

There are two consequences of interest in such a scenario. A) More DPs appear preverbally, which due to focus restructuring strengthen the prosodic pattern (s w) and B) preverbal DPs have a profound effect on the prosodic make-up of the language, since they introduce right-branching constituents on a left branch.

We can assume that when the determiner is first introduced in discourse-anaphoric contexts, it will carry stress (with the discourse anaphoric noun being destressed), such that the initial introduction of determiners in the preverbal domain does not lead to heavy right-branching constituents in the middle field (in section 6.1 I will introduce a constraint from which the relevant distinction between heavy left- and right-branching constituents follows). However, as the determiner is grammaticalized it becomes deaccented and stress will be placed on the noun per default (on the basis that lexical heads are stronger than functional heads) giving rise to more and more right-branching constituents in the preverbal domain.

The scenario sketched in (133) makes the following prediction: there should be a stage where an increased number of focused DPs appear preverbally but focused predicates (predicative adjective, nouns and participles) still appear predominantly in postverbal position. This prediction is borne out since the verbal cluster including predicative elements predominantly remained right-branching till the Early New High German period (cf. Ebert 1986).

The third observation that supports this scenario concerns the time frame of these two developments. Bolli (1975) and Bortler (1982), investigating the development of the German sentence bracket note that in the late OHG Notker text a larger number of

Accusative objects start to be placed in preverbal position. This is exactly the same period in which the definite determiner has become fully grammaticalized.

If this scenario for German is correct, then the question arises why the grammaticalization of the definite determiner did not have the same effect in the history of English and in the development of the Scandinavian languages.

First note that the Scandinavian languages have developed a clitic determiner. Thus, we can assume that at the stage of grammaticalization in which the determiner loses its accent (and gives rise to right-headed phonological phrases in German), it cliticizes onto the noun voiding any violation of the stylistic condition in (115) above.

The case of the development in English is more difficult to deal with. Philippi (1997) notes that the literature on OE suggests that the language did not have a definite or indefinite article, making use of demonstratives and numerals instead. Also van Kemenade (p.c.) reports that the definite determiner was grammaticalized in English at least 100 years later than in German. So, while the grammaticalization of the definite determiner in German was completed at the end of the OHG-period, the grammaticalization of the determiner in English must have started in the beginning of the ME-period.

OE had two demonstrative pronouns *se* ('that') and *þes* ('this'). Hroarsdottir (2006) reports that the demonstrative *se* which had different Case forms in OE was split into two invariant forms, *the* (article) and *that* (demonstrative) at the turn of the OE to the EME-period. Thus, it is possible that the English definite determiner took a different path of grammaticalization than the German one. We can assume that *the* was unaccented and therefore placed with its containing noun phrase postverbally in accordance with the condition in (115). More research on the development of the definite determiner in English is necessary to evaluate the validity of this scenario, but I think it has enough initial plausibility to provide an interesting alternative to the improbable standard scenario that tries to relate word order change to the loss of Case morphology in the history of English.

Above, I have provided a plausible scenario for why the English middle field, contrary to the German middle field, does not tolerate right-branching constituents, but what is still missing at this point is an explanation for why English has also started to postpone light elements, like pronouns.

At this point, I can only speculate about the origin of this development. But it stands to reason that language contact was at issue in this case, since at the end of the OE-period the Anglo-Saxon pronoun system was partially replaced with Scandinavian pronouns. Weak Scandinavian pronouns (nowadays), as is evidenced by the

phenomenon of object shift, are enclitic elements that are spelled out in the smallest domain that contains a suitable host, with a suitable host being the head of the phase that the pronoun belongs to, that is, the verb in the vP and the preposition in the PP (cf. Hinterhölzl 2009b).

This explanation is supported by the fact that V2-second (before its loss) is generalized in the northern contact area in the EME-period in that the pattern *Topic pronoun finite verb*, in which the Old Saxon pronoun can be analyzed as proclitic element on the verb, is replaced with the pattern *Topic finite verb pronoun*, in which the EME-pronoun can be analysed as enclitic element on the verb (see Kroch & Taylor 1999 for more details on this development).

To summarize my model of word order change, I have made the following proposals: I have argued that unmarked word order is defined by prosodic properties. Since information-structural constraints play a crucial role in the mapping between syntactic structure and prosodic structure, IS can help us to gain new insights into processes of word order change.

For instance, we have seen that what counts as unmarked word order in a language may be influenced by focus articulation, since the order focus > verb can strengthen the prosodic pattern (s w) in the middle field.

Furthermore, we have seen that the expression of IS-categories may lead to marked and unmarked syntactic structures. Since speakers, on the one hand, generally prefer unmarked structures and, on the other hand, tend to use marked structures to achieve particular communicative goals in specific communicative situations, we derive that in most stages there will be some sort of competition of marked and unmarked forms within one grammar.

While the strive towards prosodic unmarkedness might be the source for word order tendencies, the strive towards prosodic markedness might be one explanation for why we find variation in the domain of syntax, particularly in terms of word order and prosodic phrasing.

In general, a marked form *a* will be weeded out by an unmarked form *b* if it does not give rise to a different interpretation from *b*. It follows that a marked **syntactic** form will only be grammaticalized if it can be identified with a distinct **pragmatic** interpretation. Possibly, this is the reason for why information structural categories are imported into the syntax in grammaticalization processes. In conclusion, the above observations on word order change indicate that the minimalist conception of the grammar as a computational system that only operates on formal features like Case and phi-features is misguided and needs to be extended with information-structural features

like givenness, aboutness, contrast focus and the like. In short the historical developments that we investigated cannot be accounted for and understood in any deep sense without considering the role of information structure in grammar.

What is still missing is an answer to the following question: What is the place in grammar of the stylistic condition in (115) that we argued played a major role in the development of English into a pure VO-language? This will be one of the issues in Section 7.

7. Phase Theory and the Interaction between Syntax and Prosody

Although the model laid out in the previous section has many advantages over traditional models in terms of parameter change or grammar competition, it leaves open a number of questions. First, note that the model relies on a distinction between core grammar and periphery that is non-standard. For instance, to account for the VO-properties of OE, I have proposed a stylistic rule of VP-intrapolation that applies in the core derivation (see the derivation in (121) above, where VP-intrapolation feeds V2). This raises questions about their exact nature and their status in the grammar. Chomsky (2001) introduces the concept of a phase that allows for interaction between syntax and prosody in the course of the derivation. As we will see, this is exactly the type of interaction that is needed to model our stylistic rules. Therefore, I will adopt a phase-based model of the syntax-prosody interface that enables stylistic preferences to have an impact on the core derivation. This implies that stylistic preferences like the condition in (115) above are better modelled in terms of prosodic conditions that act as output filters on the syntactic derivation in a cyclic phase-based fashion. This will allow us to dispense with peripheral rules of the kind assumed in the previous section.

As an example for such an approach, we can take the account of word order variation in the middle field developed in Section 3. There, I have argued that variation of word order should be treated in terms of prosodic constraints that determine the spell-out of copies. The relevant case is illustrated again in (134).

(134) Q: Wem hat Otto das Buch gegeben?
 Who has Otto the book given?
 ‘to whom has Otto given the book’

A: a. Otto hat das Buch dem PETER gegeben

- b. Otto hat dem PETER das Buch gegeben
 Otto has (the book) to Peter (the book) given
 ‘Otto has given the book to Peter’

- (135) a. [_{IP} (φ Otto hat) (φ das Buch) (φ dem PEter gegeben)]
 b. [_{IP} (φ Otto hat) (φ dem PEter) (φ das Buch gegeben)]

(136) Interface Condition:

The phonological phrase containing the focus (main accent) must be rightmost within its intonational phrase. (cf. Chierchia 1986, Hayes & Lahiri 1991)

We have argued that both options in (134) involve scrambling of the discourse-given DP *das Buch* and that the preference for the pattern (134a) derives from a prosodic condition that crucially involves information structural notions (cf. (135) and (136)). The idea behind this account of (134) is that the grammar allows for a limited amount of options (in this case, the spell out of the copy in the higher or the lower position) that is fixed by interface conditions, like that in (136). This is the approach that I will develop in more detail in this section. I will investigate which interface constraints are relevant for the pertinent word order variation and describe how their application is linked to phases in the derivation.

In the following section, I will investigate whether and how Behaghel’s law can be modelled as a prosodic constraint in a minimalist grammar.

7.1 Prosodic Constraints and Phases

In section 5, we have seen that Behaghel’s law can be overwritten by syntactic requirements in OHG. For instance, we have noted that heavy PPs, which are regularly placed postverbally according to their weight, have to appear preverbally if contrastively focused. These syntactic conditions ultimately derive from information-structural conditions that are encoded in syntactic structure in terms of scope conditions, as outlined in (124) above.

In this section, I will argue that Behaghel’s law can be derived from a (violable) interface condition that applies in the mapping between syntactic structure and prosodic structure. The first question that arises with Behaghel’s law is the issue of when a constituent counts as heavy. So far, I have proposed that heaviness should be identified

with a branching prosodic constituent and we have come across two other phenomena where we argued that heaviness plays a role and which are relevant for the definition of the prosodic condition to be developed here. A) Adjuncts that appear between the subject and VP in English must not be right-branching, as is illustrated again in (137) and (138) below. B) Event-related adverbs that appear preverbally in German, are placed postverbally in English in the exact mirror order (cf. (139) below).

- (137) a. John (more) often (* than Peter) read the book
 b. Hans hat öfter (als der Peter) das Buch gelesen
 Hans has more-often than the Peter the book read
 ‘Hans has more often read the book than Peter’

- (138) a. John read the book more often than Peter
 b. *Hans hat das Buch gelesen öfter (als Peter)
 Hans has the book read more-often (than Peter)
 ‘Hans has read the book more often than Peter’

- (139) a. C T P M-V OV-languages
 b. C V- M P T VO-languages

In Section 3, I have argued that the order found in German is basic and that the English order is to be derived from the German word order in terms of successive cyclic VP-intrapolation that pied-pipes the adjunct at each step¹¹. Furthermore, I have argued that VP-intrapolation came about due to a stylistic rule of light predicate raising that was operative in OE and affected typically event-related adjuncts since they were primarily realized as rather heavy NPs and PPs.

In this respect, it is interesting to note that a similar restriction also applies in German verb clusters, as is shown in Hinterhölzl (2006b). German verb clusters are predominantly left-branching, but right-branching verb clusters are possible as long as

¹¹ In this approach, a modern English sentence like (ia) is derived as is illustrated in (ib-e).

- (i) a. John visited them in Vienna on Friday
 b. [...[on Friday [in Vienna [John visited them]]]]
 c. [...[on Friday [[John visited them] in Vienna t_{VP}]]]
 d. [...[[John visited them] in Vienna] on Friday]
 e. [_{IP} John_i [[[_{VP} t_i visited them]_k in Vienna t_k]_j on Friday t_j]]

the most deeply embedded cluster is left-branching (cf. Hinterhölzl 2006a). A case in question is given in (140a). However, once a right-branching verb cluster is introduced, the verb cluster must be also right-branching at the next level up, as is illustrated by the contrast in (140b) and (140c).

- (140) a. weil er den Text muß lesen können
 since he the text must read can
 ‘since he must be able to read the text’
 b. ??weil er den Text [[müssen [lesen können]] wird]
 since he the text must read can will
 ‘since he will have to be able to read the text’
 c. weil er den Text [wird [müssen [lesen können]]]

The formation of verb clusters is motivated by the following two licensing requirements (cf. Hinterhölzl (2006a): dependent verbs move into dedicated positions in the V-domain to be temporally linked and to check the subcategorisation of the selecting verb. F2⁰ is responsible for temporal linking and Asp⁰, the highest head in the V-domain, is responsible for checking the subcategorisation of the matrix verb.

In this approach, left-branching verb clusters are derived if the dependent verbs are spelled out in the highest specifier, as is illustrated in (141a), while right-branching verb clusters are derived if the dependent verbs are spelled out in the lower specifier (141b). The only exception to this rule build verb clusters comprising an IPP-infinitive (an infinitive that replaces a part participle). These verb clusters are obligatorily right-branching, since in this case there is a null morpheme that moves to the highest head in the V-domain, stranding the dependent verbs in the lower specifier, as is illustrated in (141c).

- (141) a. [AspP [**lesen können**] muss [F2P [*lesen können*] [VP]]]
 b. [AspP [*lesen können*] muss [F2P [**lesen können**] [VP]]]
 c. [AspP 0-hat [F2P [*lesen können*] [VP]]]

The interesting question now is what the generalisation illustrated in (140bc) results from. In any event, the generalisation cannot be derived from a hard syntactic condition in West Germanic, since we can find many instances that violate it, as is the case with IPP-infinitives in West Flemish. (142) illustrates a right-branching verb cluster headed by *willen* which itself sits on a left-branch with respect to the selecting auxiliary *een*.

- (142) dan ze toch kosten [[willen [dienen boek kuopen]] een] (West Flemish)
 they they but could want-IPP that book buy have
 ‘if they only could have wanted to buy that book’

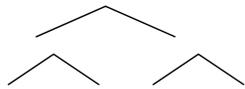
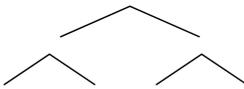
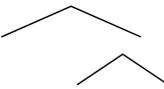
A possible solution is to relegate the contrast in (140bc) to a violable interface condition that determines the best match between a given syntactic structure and a prosodic output structure. Given that left- and right-branching verb clusters are mapped onto left- and right-headed phonological phrases, a possible candidate for such an interface condition is (143).

- (143) Mapping Condition between syntactic structure and prosodic structure:

A right-headed phonological phrase (in a verb cluster) must sit on a right branch with respect to the syntactic head that is to become its prosodic sister

The mapping condition in (143) can probably be formulated in a more elegant way. But it is meant to account for the patterns in (144). The patterns in (144) indicate the deeper reason that probably lies behind the condition in (143): once fixed, stress tends to stay in a peripheral position within a certain domain (this will be reworked in the final section).

(144b) is okay, since stress remains left-peripheral and (144c) is okay, since stress remains right-peripheral, while stress in (144a) is neither left- nor right-peripheral unless H itself is stressed. (140b) violates the condition in (143). The violation can be circumvented if the (already) right-branching verb cluster is spelled out in the lower specifier, as is the case in (140c).

- (144) a.* 
 ((A B) H)
- b. ok 
 ((A B) H)
- c. ok 
 (H (A B))

Note, however, that condition (143) is a condition that can be taken to apply in the V-domain and in the I-domain (excluding the subject) in VO-languages. It can explain why VP-intrapolation should pied-pipe event-related adjuncts: pied-piping ensures that these adjuncts end up to the right of the verb, yielding phonological structures like (144c). The condition in (144) can also be used to rule out (137a), since *more often than Peter* is a right-headed prosodic constituent that sits on left branch with respect to the

verb phrase that is to become its prosodic sister. Since German tolerates heavy adjuncts and arguments in the middle field, this condition cannot be taken to apply in the German middle field, but - as we have seen above- is still active in the German V-domain. Why should that be so?

In this respect it is interesting to note that prosodic constraints can apply phasewise and can thus hold in one phase while not being operative in another phase. The crucial evidence comes from restrictions on VP-topicalization in German, as is argued in Hinterhölzl (2006b). As is illustrated in (145), topicalized right-branching verb clusters exhibit an interesting contrast. In general, the topicalization of a right-branching verb cluster leads to ungrammaticality, unless it comprises an IPP-infinitive. In this case the topicalization is rather marked but grammatical (145b).

- (145) a. ?*[müssen [lesen können]] wird er den Text (ok lesen können müssen wird)
 must read can will he the text
- b. ? [haben [lesen wollen]] wird er den Text (* lesen wollen haben wird)
 have read want-IPP will he the text
 he will have wanted to read the text

Hinterhölzl (2006b) shows that the contrast in (145) can be explained in a phase-based account in which the application of the Phase Impenetrability Condition (cf. Chomsky 2001) leads to a violation of the prosodic condition in (143) in the left edge of the V-domain. The violation of (143) leads to ungrammaticality if there is no alternative spell out option, as is the case in (145a), but only to a marked prosodic structure if there is no alternative spell out option, as is case with IPP-infinitives (145b).

This solution, however, is only feasible in a framework in which spell out applies cyclically, including the cyclic, that is, phasewise application of conditions that map syntactic structure onto prosodic structure. More specifically, we have seen above that interface conditions can apply phasewise in the sense that they may hold in one phase (the V-domain in German) without holding in another phase (the I-domain, or middle field, in German).

In conclusion, word order variation can be accounted for by the differential expression of information structural categories within one grammar and word order preferences (Behaghel's law) are due to violable interface conditions that define the ideal mapping between syntactic structure and prosodic structure in the course of the derivation.

In the following section, I will show that it is the (non-)application of the condition in (143) that explains the major differences between German and English.

7.2. A Unified Approach to the Syntax of OV- and VO-languages

The main objective of the following two sections is to develop a model of the complex interaction between syntax, prosody and information structure that can explain the pertinent change in English word order without the help of stylistic rules. In other words, an account is needed that can explain the co-occurrence of OV- and VO-orders without optional VP-intrapolation. The solution that I am going to propose is rather simple. VO-orders, according to the UBH, are base generated orders. OV-orders, in principle, result from licensing movement (and scrambling). Pure OV-orders, pure VO-orders and mixed OV/VO-orders on surface are derived from spell out options that are fixed by prosodic conditions like (143) above. This takes care of the default placement of arguments and (vP-internal) predicates. The approach that we are going to develop, however, must also be able to account for the default placement of event-related adjuncts.

Remember that we described the development of English to a pure VO-language as a process in which an originally peripheral rule of VP-intrapolation (also called LPR) became reanalysed as an obligatory rule of the core grammar after it has crossed a certain threshold in language usage. This account raises the question of what it means that a rule suddenly becomes obligatory. If VP-intrapolation is obligatory in modern English, and thus must be taken to serve some licensing purpose, some version of V-movement must have also been obligatory in the older stage.

A possible solution was already pointed out in Section 3.1 (cf. Hinterhölzl 2003). There two macro-parameters were discussed: a) the spell-out of copies and b) XP-movement versus head movement in feature checking operations. There, it was also proposed that the differential placement of event-related adjuncts follows from the macro-parameter b) and the assumption that the licensing of event-related adjuncts requires V-movement. In German, it was argued, the verb undergoes head-movement through the row of adjuncts that leaves them in the original order, while in English the verb undergoes XP-movement which pied-pipes the adjuncts at each step and leads to inverted orders of these adjuncts in the postverbal domain.

In this scenario we could assume that adjunct licensing in OE originally involved head movement of the verb, which - due to an increased use of LPR - was reanalysed as XP-movement of the verb, that is, as obligatory VP-intrapolation. This scenario is plausible, but raises a number of questions, addressed in Hinterhölzl (2009b), that are

the main subject in the following subsections: What is the rationale of adjunct licensing and what is the reason behind TP-movement that is necessary to account for the preverbal occurrence of event-related adjuncts in German?

7.2.1 The licensing of event-related adjuncts

Hinterhölzl (2009b) proposes that event-related adjuncts introduce separate phases into the clause and are interpreted as predicates on events that require the movement of a verbal projection into their licensing domain that serves as derived subject.

In the standard account to modification, it is assumed that adjuncts are adjoined to the maximal category of the head they modify. Thus, the attachment site of the adjunct is determined by its interpretation (it minimally has to attach to the constituent it modifies). Secondly, the syntactic operation of adjunction is interpreted as identification of the individual variables introduced by adjunct and modified head, as is illustrated in (146).

- (146) a. meet in the park
 b. $[_{VP} [_{VP} V(e_1)] PP (e_2)]^{12}$
 c. identification: $e_1 = e_2 \rightarrow \exists e$ such that meet (e) & in the park (e)

In Cinque's (1999) proposal adjuncts are introduced as specifiers of functional heads that are ordered according to a universal hierarchy in the extended projections of the modified head. In this approach, the order of event-related adjuncts (T-P-M) can be viewed as representing a small section of the universal hierarchy of modifiers pertaining to the verb (see Schweikert (2005) for a more complete picture of this hierarchy). Cinque's proposal can thus be taken to provide an alternative account to the question of how adjuncts are to be attached to the head they modify, but fails to address the question of how the individual variables of adjunct and modified head are identified.

In other words, if we want to dispense with adjunction altogether, we must address the question of how an adjunct, being base generated as the specifier of a functional head is interpreted and how, for instance, the event variable of the verb is identified with the individual variable of an adjunct that sits higher up in the tree in the

¹² Davidson (1966) argued that events should be treated as referential entities in the logical description of sentences (cf. also Higginbotham 1985, Parsons 1990 and Kratzer 1995 for applications of this proposal).

specifier of a functional head, possibly separated from the verb by various heads dedicated to the licensing of the arguments of the verb.

Hinterhölzl (2009b) makes the following proposal. The adjunct introduced as the specifier of a functional head is interpreted as a predicate on the category it is taken to modify in the standard theory. Assuming that every predicate provides a licensing domain for its arguments, introducing an adjunct in the functional skeleton of the CP will always involve two functional heads: the one that introduces the adjunct as an additional predicate (called F1 in (147)) and the other that licenses the argument of this predicate (called F2 in (147)). In the course of the derivation the vP moves into [Spec F2] and the two individual variables are identified via predication, with the vP acting as subject of the predication, as is illustrated in (147). In this approach, the semantic interpretation of event related adjuncts as predicates on the verb phrase is already represented in the syntax as a derived subject-predicate relation between vP and adjunct.

(147) $[[_{vP} V (e_1)] F2 [PP (e_2) F1 \dots [t_{vP}]]]$

In Hinterhölzl (2009b), it is argued that F1 and F2 in (147) constitute projections of a separate phase and are not considered as being part of the extended projection of the verb. F1 introduces an additional predicate in the clause that has its own licensing domain, namely F2. In other words, [Spec,F2] can be compared with [Spec, IP] in the clause. vP-intrapolation, therefore, has to be considered as a case of A-movement that serves to license the adjunct as a (secondary) predicate. Some evidence for the assumption that vP-intrapolation is a case of A-movement is given in Hinterhölzl (2009b).

That the projections F1 and F2 and their respective specifiers constitute separate phases follows from the following typology of phases. I propose that the main phases (the CP in the clausal domain and the DP in the nominal domain) comprise the following sub-phases: a predicate domain (roughly the vP in the clause) that introduces a predicate and its arguments, the I-domain, in which the (properties of the) arguments of the predicate are licensed and a C-domain (or completing domain) that embeds the predicate in another clause or in the relevant context. I will call these sub-phases *homorganic*, since they are projected by the same phase predicate.

According to this typology, adjuncts comprise a predicate domain and an I-domain but lack a completing domain, which bars them from being embedded like complements. Instead of being embedded they are superimposed in the I-domain of another predicate. To be licensed event-related adjuncts must enter into a predication

relation with the vP in the clause. Furthermore, I propose that Tense is a separate predicate that projects its own sub-phases (which are non-homorganic with respect to the phases projected by the verb) in the I-domain of the verb.

To summarize, vP-intrapolation is triggered by the licensing requirement of adjuncts. The intraposed vP acts as subject of predication and vP-movement is thus to be considered as obligatory A-movement. Furthermore, I have proposed that the I-domain in the clause is interspersed with non-homorganic sub-phases that are projected by adjuncts and Tense.

7.3.2 TP-movement and vP-movement into the C-domain

If the account of adjunct licensing sketched above is correct, then vP-movement in the clause has to be taken to occur in OV- and VO-languages alike and word order differences in OV- and VO-languages should follow from an elementary choice connected with vP-intrapolation. In the following, I will argue that this basic choice involves vP-extraction or pied-piping induced by vP-movement (and cannot involve the choice between XP and head movement, as envisaged above).

In the present account, adjuncts are introduced preverbally in OV- and VO-languages alike and their preverbal or postverbal placement should follow from vP-extraction or vP-pied piping at each step, respectively. The default English order is derived, if the vP pied-pipes the relevant PPs at each step. Furthermore, vP-extraction in this approach must be taken to obtain as a rule in OV-languages, where vP-movement leaves the original order of adjuncts intact.

This account raises two questions. First, there is the issue of which factors decide which option is taken. Second, if the vP extracts at each step in the process of licensing adjuncts in the middle field in OV-languages, then there must be an additional step that moves the entire middle field in front of the vP again before the end of the derivation. The latter will be the subject in this subsection, the former issue will be taken up in following section.

In Hinterhölzl (2006a), it is proposed on the basis of restructuring infinitives that the extended vP (AspP)¹³ and the TP undergo licensing movement into the C-domain in

¹³ In Hinterhölzl (2006a) it is argued on the basis of VP-topicalisation data that AspP constitutes the edge of the V-domain. As will become evident below, I propose that there is an AspP in the V-domain that defines different event-types (cf. Vendler 1967) and there is an AspP in the T-domain (so-called viewpoint aspect) that together with an abstract Tense predicate defines different Tenses (cf. Smith 1991,

German, as is illustrated in (148). These movements are argued to follow from a general theory of sentential complementation, in which the complementizer acts as a placeholder for the selectional requirements of the matrix verb. In particular, it is argued that movement of the AspP into FinP (cf. Rizzi 1997) (called Status phrase in Hinterhölzl 2006a) serves to check the morphological subcategorisation of the matrix verb and that movement of the TP into MoodP serves to temporally link the embedded event to the matrix event time. In (148), ForceP encodes clausal force and represents the highest head in the C-domain.

- (148) a. [CP=FP Force [MP Mood [FinP Fin [TP T [AspP V]]]]] base structure
checking finiteness ->
- b. [CP=FP Force [MP Mood [FinP [AspP V] Fin [TP T]]]]
temporal anchoring ->
- c. [CP=FP Force [MP [TP T] Mood [FinP [AspP V] Fin]]]

In Hinterhölzl (2009b), it is proposed that this account be extended to non-restructuring contexts. In particular, it is proposed that the dependency relations between C⁰ and T⁰ (cf. Chomsky 2005) and between Fin⁰ and v⁰ (cf. Rizzi 1997) are embodied via XP-movement of TP and AspP into the C-domain in English and German. The rationale behind these movements is that different speech act (forces) are connected with different verbal moods that determine the situational and temporal anchoring of the event in TP and different verbal moods select different finite and non-finite verbal forms that are expressed in the V-domain.

On its way to the C-domain the extended vP moves into the specifier of (viewpoint) Aspect in the T-domain, as is illustrated in (149). I will argue in the following section that interface conditions determine that the extended vP pied-pipes the containing Aspect phrase in English, while in German the extended vP extracts from the Aspect phrase when moving on into the C-domain.

- (149) [I-domain [T-domain (Spec AgrS) [Spec PRES/PAST [vP Asp]]] [V-domain t_{vP}]]¹⁴

Kratzer 1998). For example, the English simple past tense (as in *Peter ran*) expresses past tense **and** perfective viewpoint.

¹⁴ In (149), the specifier of the tense predicate contains a referential temporal argument (cf. Stowell 1996) with respect to which the event denoted by vP is situated.

Given this scenario, we can assume that on its way to the T-domain, the extended vP moves through all the *predication* positions introduced by modifying adjuncts in the middle field. The modifying adjuncts will remain in the original order in preverbal position, if the vP is subextracted at each step, since TP-movement will then move the entire middle field anew in front of the extended vP in the C-domain. On the other hand, the adjuncts will appear in the mirror order, that is typical of VO-languages, if the extended vP at each step on its way up to the C-domain pied-pipes the respective functional projections containing the adjuncts. In this case, as is typical for VO-languages, the entire middle field will follow the verb in the C-domain, with only the subject and possibly some higher adverbs being moved via TP-movement to MoodP in front of the verb (phrase) again.

Note, however, that we must assume that pied-piping in a VO-language like English is the preferred but not the only option. To derive the correct word order in (150), in which case the original hierarchical relationship between temporal adverbial and local adverbial is preserved, we have to assume that vP-extraction may also take place in the derivation of (150a).

- (150) a. John met Mary on everybody's birthday in his house
 b. They met students every day of the week in a different university

In the derivation of (150a), the vP extracts after having moved into the licensing position of the lower locative adverbial and will induce pied-piping only after having moved into the licensing position of the higher temporal adverb. The rationale could be that extraction in this case is allowed in order to preserve the binding relation between the temporal adverbial and the locative adverbial. This implies that pied-piping is the default option in a VO-language. However, this default must be taken to be overruled by interface requirements, like the availability of certain binding relations. Another interface requirement that can be taken to enforce extraction instead of the default operation of pied-piping is focus articulation. Since VO-languages like English demand that focused constituents occupy the right edge within the intonational phrase, focussing of the lower adjunct is expected to induce extraction in order to remain in clause final position, as is illustrated in (150b). In the following section, I argue that the default option of pied-piping is due to a prosodic requirement that only allows light material in the middle field of English, as given in (151).

(151) Prosodic Mapping Condition (final version):

A right-headed prosodic constituent (w s) may not sit on a left branch with respect to the head of the predicate phase

7.3. Interface Conditions and the OV/VO-Parameter

Before we come back to the issue of how the distribution of event-related adjuncts in German and English is derived exactly, we will have to address the issue of the differential placement of arguments and predicates in the two languages. In section 1, I have argued at length that XV-orders are derived from base-generated VX-orders by different types of licensing movements. These obligatory licensing movements must also obtain in VO-languages. In the present framework we have two options to account for VO-orders on the surface: a) spell out in the lower base position or b) spell out in the licensing position with vP-movement around it.

The second option can be excluded, since there is ample empirical evidence that VO-orders in English cannot be derived by object movement that spells out the higher copy plus vP-movement around it. First note that the vP cannot be topicalized excluding the direct object, as is illustrated in (152). Furthermore, note that the object cannot be separated from the verb and appear in its *scope position* between adverbs, as is illustrated by the contrast between German and English in (153a-b). The intended reading of (153a) is possible in the order given in (153c), where the direct object arguably occupies a vP-internal position.

(152) a. John wanted to buy something yesterday

... * and buy John did a book

b. Hans wollte gestern etwas kaufen

... und gekauft hat er heute ein Buch

Hans wanted to buy something yesterday

and bought has he today a book

‘Hans wanted to buy something yesterday and today he bought a book’

(153) a. * John met every day two girls in their classrooms (Temp > DO > Loc)

b. Hans traf jeden Tag zwei Mädchen in ihren Klassenzimmern

Hans met every day two girls in their class rooms

- c. John met two girls in their classrooms every day

The only option that remains is to assume that DPs that undergo Case-licensing movement are spelled out in the vP in English. This option can be derived from the condition in (151), if we assume that (151) requires the spell out of arguments in the lower position. Furthermore, I have argued in Section 3 that arguments that undergo scrambling for reasons of scope taking spell out the lower copy as well. This type of movement was called silent scrambling in Section 3.¹⁵ Now, we have to assume that also licensing movement of argumental DPs and PPs is silent in English. On the other hand, vP-external adjuncts must be taken to undergo scrambling (for reasons of scope taking) that spells out the higher copy, since, as is illustrated in (154), a locative PP can take scope over a temporal PP that is base-generated in a higher position.

(154) They met students in each university on a different day

The challenge for the present account is to provide a principled explanation of when an A-movement chain spells out the lower or the higher copy in German and English. The generalization that emerges is the following: A-movement of arguments spells out the lower copy in VO-structures (and the higher copy in OV-structures) while A-movement of adjuncts is not affected by the OV/VO parameter and always spells out the higher copy. A principled solution to this problem is given in Hinterhölzl (2009b). Here I will only outline the basic tenets of the account.

The basic idea that is developed there is that this generalization follows from two basic modes of prosodic composition of arguments, adjuncts and heads. Let us start with the more complex case that English represents. Above I have argued that A-movement of arguments in English spells out the lower copy since the English middle field does not tolerate heavy constituents and formulated the prosodic mapping condition in (151). Furthermore, I argued that this prosodic condition motivates also pied-piping of vP-external material, that is, of PP adjuncts.

¹⁵ Silent scrambling was argued to be necessary to account for the fact that postverbal complement demonstrate evidence of c-command relations into postverbal adjuncts. In order to account also for cases of c-command exerted by argument and adjunct PPs, Hinterhölzl (2009b) proposes that there is also silent scrambling within PPs. But in this paper I will abstract away from this fact since it is not directly relevant for our purposes here.

I propose that a constituent is spelled out at the point at which all its features have been checked, guided by the prosodic constraints that apply in the phase that contains it. The crucial distinction between argument and adjunct scrambling follows from the point in time at which the condition in (151) applies to right-headed constituents on a left branch in the I-domain. I would like to propose that this mapping condition applies at the time at which an argument or an adjunct is joined with the head of the predicate domain into a single prosodic constituent.

Many researchers noted that there is an asymmetry in prosodic domain formation between arguments and adjuncts (cf Gussenhoven 1984, Krifka 1984, Jacobs 1992 among others) and pointed out that adjunct and verb form separate prosodic constituents. In the present account this follows since they belong to non-homorganic phases as stated above. No problem arises between an argument and the verb: a verb may restructure with an argument to form a single prosodic constituent since they belong to homorganic phases. In analogy to recent work by Wagner (2005), I propose that there are two modes of prosodic composition which are phase-based: subordination and coordination.¹⁶ Subordination pertains to arguments and their selecting head, that is, material in homorganic phases: when an argument and a head are combined the result is a single prosodic constituent (of a certain type), as is illustrated in (155a). Coordination pertains to the composition of an adjunct and a head, that is, material in non-homorganic phases: the combination of an adjunct and a head yields two prosodic constituents of the same type (that can be combined at the next cyclic level), as is illustrated in (155b).

(155) Modes of prosodic composition

- a. subordination: (DP) + V → ((DP) V)
- b. coordination: (PP) & V → (PP) (V)

When a right-headed prosodic constituent that is an argument is licensed in the I-domain, the mapping condition applies immediately since according to (155a) a single prosodic constituent with the verb is computed at this point. When a right-headed prosodic constituent that is an adjunct is licensed in the I-domain, the mapping condition does not apply immediately, since according to (155b) no single prosodic constituent is computed at this point in the derivation. Thus, an adjunct may also be

¹⁶ Wagner (2005) calls them subordination and sister-matching and invokes a directionality parameter. Here I assume that no directionality parameter is necessary to account for the data.

spelled out in the higher position. As we will see below, spell out in the higher position will follow from the interface constraint in (156). Only when the adjunct is joined into a prosodic unit containing the verb at the sentence level (that is, at the level of an intonational phrase), the prosodic condition in (151) applies and guarantees that the English middle field will only contain light, that is, non-right-branching adjuncts.

In German, the prosodic condition (151) does not apply in the I-domain of the verb such that arguments that move into the middle can be spelled out in the higher position and adjuncts -independently of their weight- can be stranded by vP-movement in the clause. This raises the question of which factor is responsible for the spell out of the higher copy in German and in all those cases in English in which the mapping condition (151) does not apply at the relevant point in the derivation

Note that while we can say that VO-structures (with right-headed arguments and adjuncts) guarantee a monotonous, more perspicuous mapping of word order onto prosodic structure on the PF-side, OV-languages are more transparent with respect to the mapping of word order onto scope relations on the LF-side of the computation. Arguments and adjuncts in German appear in their scope positions allowing scope to be read off directly from their respective surface positions. Therefore we can envisage, parallel to the interface condition on the mapping to PF in (151), an interface condition on the mapping to LF, as given in (156).

(156) Scope Transparency:

If *a* scopes over *b*, the spell out copy of *a* must c-command the spell out copy of *b*

Let us now address the important question of how these two interface conditions interact with each other. The data that we have discussed so far indicate that we do not need to resort to a mechanism like competition or constraint ranking in OT. Their application can be described with an elsewhere-condition, where Scope Transparency presents the general condition which can be overruled by more specific prosodic requirements, as is stated in (157).

(157) The spell out of an A-chain in a given phase is determined by prosodic constraints that hold in that phase, if applicable; if not, spell out is fixed according to Scope Transparency

In this context, it is interesting to note that the Germanic languages started out from a similar basis and developed into different directions. In Old High German, Old Norse and Old English light elements tended to precede the thematic verb while heavy (branching constituents) followed the verb (cf. Behaghel 1932). While the placement of arguments and adjuncts in modern German is thoroughly scope transparent (also PP-stranding by vP-movement serves to preserve pre-existing c-command relations among adjuncts) and their placement in modern English is predominantly determined by prosodic constraints, Icelandic has opted for a split solution: referential DPs are spelled out in a low position obeying the prosodic condition, while quantificational DPs (including negative quantifiers) are spelled out in their scope positions in the middle field.

The diachronic development of the Germanic languages suggests that these interface constraints also serve as restrictions in the acquisition process that guide the child in fixing spell out positions of universally given checking/movement operations resulting in a uniform analysis across DP-types, unless there is strong evidence to the contrary (as we must assume was the case in Icelandic).

To sum up this section, I am proposing that the properties of OV- and VO-languages are not determined by directionality parameters but follow from different choices in obeying two types of transparency constraints at the interfaces. In this way, we have eliminated the head complement parameter. We have shown that OV- and VO-languages differ in a number of respects that can be better explained as the result of the interaction of two transparency constraints at the interfaces.

8. Summary

I have argued that the base order and the unmarked word order in a language are to be distinguished. While the base order is determined universally as the sequence specifier-head-complement according to the UBH, the unmarked word order in a language is given by the unmarked prosodic pattern in this language. I have developed a model of word order change, in which stylistic preferences over time may change the marked and unmarked prosodic patterns in a language. I have argued that such a model is superior to a model employing a change in the head-complement parameter, since it can account better for the global differentiation that German and English have undergone in their history.

Furthermore, I have argued and shown with historical data that the unmarked prosodic patterns in a language are crucially determined by the expression of

information structural categories. I have identified as one key process focus restructuring that influences the formation of phonological phrases. In addition, I have argued that information structural requirements determine word order not only in that they trigger displacement, but also in that they directly influence the spell-out of movement copies. Therefore, I have laid out a phase-based model of the syntax-prosody interface in which stylistic and information-structural requirements are satisfied in a cyclic fashion in the course of the derivation.

Finally, I have investigated in detail the differences in word order between German and English. I have argued that these differences can be reduced to the workings of two interface conditions, showing that while word order is determined by scope transparency in German, it is fixed predominantly by requirements that ensure PF-transparency in English and I have outlined the historic conditions and processes that resulted in this state of affairs.

The empirical studies of OHG and OE/ME revealed that although the effects of PF-transparency are also observable in OHG, word order was predominantly determined by information structural requirements that influence the syntax in terms of scope conditions in this stage of the language, while the development of English reveals that word order in the earliest accessible stage of this language was predominantly determined by conditions guaranteeing PF-transparency.

References

- Alexiadou, A. and E. Anagnostopoulou. 1997. "Towards a uniform account of scrambling and clitic doubling". In W. Abraham and E. van Gelderen (eds.) *Syntactic Problems - Problematic Syntax*. 142 – 161. Tübingen, Max Niemeyer Verlag.
- Behaghel, O. 1932. *Deutsche Syntax*. Band 4. Heidelberg, Carl Winters Universitätsbuchhandlung.
- Belletti, A. 2001. "Aspects of the low IP area". In L. Rizzi (ed) *The Structure of CP and IP. The Cartography of Syntactic Structures*, Vol.2, 16-51. Oxford: Oxford University Press.
- Benincà, P. 2001. "The Position of Topic and Focus in the Left Periphery". In G. Cinque and G. Salvi (eds:) *Current Studies in Italian Syntax. Essays Offered to Lorenzo Renzi*. 39-64. Amsterdam and London: Elsevier.

- Benincà, P. and C. Poletto. 2002. in L. Rizzi (ed.) *The Structure of IP and CP, The Cartography of syntactic structures*, Vol.3. Oxford, OUP.
- Bolli, E. 1975. *Die verbale Klammer bei Notker. Untersuchungen zur Wortstellung in der Boethius-Übersetzung*. Berlin/New York, Mouton de Gruyter.
- Borner, A. 1982. *Syntaktische Klammerbildung in Notkers Psalter*. Berlin/New York, Mouton de Gruyter.
- Bowers, J. 1994. The Syntax of Predication. *Linguistic Inquiry* 24: 591-655.
- Büring, D. 1999. "Topic". In P. Bosch and R.v.der Sand (eds.) *Focus – Linguistic Cognitive and Computational Perspectives*, 142-165. Cambridge: Cambridge University Press.
- Cardinaletti, A. 2002. "Towards a cartography of subject positions". In L. Rizzi (ed.) *The Structure of IP and CP, The Cartography of syntactic structures* vol.3. Oxford, OUP.
- Cecchetto, C. 2000 "Doubling structures and reconstruction." *Probus* 12,1, 1-34.
- Cecchetto, C. 1999. "A Comparative Analysis of Left and Right Dislocation in Romance". *Studia Linguistica* 53: 40-67.
- Chafe, W. 1987. "Cognitive constraints on information flow". In R. Tomlin (ed) *Coherence and grounding in discourse*. 21-51. Amsterdam, Benjamins.
- Chierchia, G. 1986. "Length, Syllabification and the Phonological Cycle in Italian". *Journal of Italian Linguistics* 8, 5-34.
- Cinque, G. 1990. *Types of A'-dependencies*. Cambridge, Mass.: MIT Press.
- Cinque, G. 1999. *Adverbs and functional heads. A Crosslinguistic Perspective*. New York: Oxford University Press.
- Cinque, G. 2004. "Complement and adverbial PPs: Implications for Clause Structure". Ms. Written version of the GLOW Presentation in Amsterdam (April 9-11, 2002).
- Cinque, G. 2005a. "The prenominal origin of relative clauses". Seminar presented at *Linguaggio a Roma Tre*, Università degli Studi Roma Tre (January 2005).
- Cinque, G. 2005b. "Deriving Greenberg's Universal 20 and its Exceptions". *Linguistic Inquiry* 36, 3: 315-332.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: MIT Press.
- Chomsky, N. and H.Lasnik. 1993. "The theory of principles and parameters". In J. Jacobs, A. von Stechow, W. Sternefeld and T. Vennemann (eds.) *Syntax: an international handbook of contemporary research*. 506-569. Berlin, Walter de Gruyter.
- Chomsky, N. 1995a. *The Minimalist Program*. Cambridge, Mass., MIT Press.

- Chomsky, N. 1995b. "Bare Phrase Structure". In G. Webelhuth (ed.) *Government and Binding Theory and the Minimalist Program*. Cambridge, MA, Blackwell.
- Chomsky, N. 1998. "Minimalist Explorations". Ms. Cambridge, MA, MIT.
- Chomsky, N. 2001a. "Derivation by phase". In M. Kenstowicz (ed.) *Ken Hale: a life in language*. 1-52. Cambridge, MA, MIT Press.
- Chomsky, N. 2001b. "Beyond Explanatory Adequacy". In A. Belletti (ed.) *Structures and Beyond. The Cartography of Syntactic Structures, volume 3*. 104-131. Oxford, Oxford University Press.
- Chomsky, N. 2005. "On phases". Manuscript. Cambridge, MA, MIT.
- Comrie, B. 1988. "Topics, grammaticalized topics and subjects". In: S. Axmaker, A. Jaisser and H. Singmaster (eds.) *General Session and Parasession on Grammaticalization*. 265-79. Berkeley Linguistic Society.
- Corver, N. and H. van Riemsdijk. 1994. *Studies on Scrambling*. Berlin/New York, Mouton de Gruyter.
- Corver, N. and D. Delfitto. 1997. "Feature Primitives and the Syntax of Specificity." *GLOW Newsletter* 34.
- Delbrück, B. 1911. *Zur Stellung des Verbuns*. (Germanische Syntax 2.) Leipzig, Teubner.
- Den Besten, H. and G. Webelhuth. 1987. "Remnant Topicalization and the Constituent Structure of the VP in the Germanic SOV Languages" (abstract). *GLOW Newsletter* 18, 15-16.
- Diesing, M. 1992. *Indefinites*. Linguistic Inquiry Monographs. Cambridge, Massachusetts, MIT Press.
- Diesing, M. 1997. "Yiddish VP order and the Typology of Object Movement in Germanic". *Natural Language and Linguistic Theory* 15, 369-427.
- Dittmer, A. and E. Dittmer. 1998. *Studien zur Wortstellung-Satzgliedstellung in der althochdeutschen Tatianübersetzung*. Göttingen, Vandenhoeck & Ruprecht.
- Dobrovie-Sorin, C. 1990. "Clitic Doubling, Wh-movement and Quantification in Romanian". *Linguistic Inquiry* 21:351-397
- Ellegard, A. 1953. *The Auxilliary DO: the Establishment and Regulation of its Use in English*. Stockholm, Almqvist & Wiksell.
- Enc, M. 1991. "The Semantics of Specificity". *Linguistic Inquiry* 22, 1-25.
- Ernst, T. 1998. "The scopal basis of adverb licensing". *NELS* 28:127-142.
- Falk, C. 1993. "Non Referential Subjects in the History of Swedish". Doctoral dissertation, University of Lund.

- Frascarelli, M. 2000. *The Syntax-Phonology Interface in Focus and Topic Constructions in Italian*. Dordrecht, Kluwer Academic Publishers.
- Frascarelli, M. and R. Hinterhölzl. 2007. "Types of Topics in German and Italian". In S. Winkler and K. Schwabe (eds.) *On Information Structure, Meaning and Form*. 87-116. Amsterdam & Philadelphia, John Benjamins.
- Frey, W. 2000. "Über die syntaktische Position der Satztopiks im Deutschen". *ZAS Papers in Linguistics* 20, 137-172.
- Frey, W. 1993. *Syntaktische Bedingungen für die Interpretation*. Berlin, Studia Grammatika 35.
- Frey, W. and K. Pittner. 1998. "Zur Positionierung der Adjunkte im deutschen Mittelfeld". *Linguistische Berichte* 176, 489-534.
- Fuss, E. 2002. "Wortstellungsvariation in den älteren germanischen Sprachen". Ms., University of Frankfurt.
- Fuss, E. and C. Trips. 2002. "Variation and Change in Old and Middle English. On the validity of the Double Base Hypothesis". *Journal of Comparative Germanic Linguistics* 4, 1771-224.
- Gärtner, H. M. and M. Steinbach. 2000. "What do reduced pronominals reveal about the Syntax of Dutch and German". *LIP* 9:7-62.
- Givón, T. 1983. "Topic Continuity in Discourse: An Introduction". In T. Givón (ed.) *Topic Continuity in Discourse: A Quantitative Crosslanguage Study*. 5-41. Amsterdam and Philadelphia, John Benjamins.
- Greenberg, J. 1963. "Some universals of grammar with particular reference to the order of meaningful elements". In J. Greenberg (ed.) *Universals in Language*. 33-59. Cambridge, Massachusetts, MIT Press.
- Grewendorf, G. 2002. "Left Dislocation as Movement". *Georgetown University Working Papers in Theoretical Linguistics* 2: 31-81.
- Grewendorf, G. and J. Sabel. 1999. "Scrambling in German and Japanese". *NLLT* 17:1-65. Dordrecht, Academic Publishers.
- Grewendorf, G. and W. Sternefeld. 1990. *Scrambling and Barriers*. Amsterdam, John Benjamins.
- Haegeman, L. 2001. "Remnant Movement and OV order". In P. Svenonius (ed.) *The Derivation of OV and VO*. 69-96. Amsterdam, John Benjamins.
- Haegeman, L. 1991. *Introduction to Government and Binding Theory*. Oxford and Cambridge, Blackwell.

- Haegeman, L. 1994. "The Typology of Syntactic Positions:L-Relatedness and the A/A'-Distinction". *Groninger Arbeiten zur Germanistischen Linguistik* 37, 115-157.
- Haegeman, L. and R. Zanuttini 1994. "Negative Heads and the Neg Criterion". *The Linguistic Review* 8, 233-252.
- Hale, K. and J. Keyser. 1993. "On Argument Structure and the Lexical Expression of Syntactic Relations". In K. Hale and J. Keyser (eds.) *The View from Building 20*. 53-110. Cambridge, Mass., MIT Press.
- Haider, H. 2000. "Adverb Placement-Convergence of Structure and Licensing". *Theoretical Linguistics* 26: 95-134.
- Haider, H. 1993. *Deutsche Syntax, generativ: Vorstudien zur Theorie einer projektiven Grammatik*. Tübingen, Narr.
- Haider, H. and I. Rosengren. 1998. "Scrambling". *Sprache and Pragmatik* 49. Germanistisches Institut der Universität Lund, 1-104.
- Harris, A. and L. Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge, CUP.
- Hayes, B. and A. Lahiri. 1991. "Bengali Intonational Phonology". *NLLT* 9, 47-96.
- Hawkins, J. A. 1983. *Word Order Universals*. New York, Academic Press.
- Hinterhölzl, R. and S. Petrova (2012) "From V1 to V2 in older Germanic". *Lingua* 120, 315 - 328.
- Hinterhölzl, R. and S. Petrova (2011). "Rhetorical relations and verb placement in Old High German". In C. Chiriarcos, B. Claus and M. Grabski. (eds.) *Salience: Multidisciplinary Perspectives on its Function in Discourse*. 173-201. Berlin, Mouton de Gruyter.
- Hinterhölzl, R. 2010. "Information Structure and Unmarked Word order in (older) Germanic". In C. Féry and M. Zimmermann (eds.) *Information Structure from different Perspectives*. 282-304. Oxford, OUP.
- Hinterhölzl, R. 2009a. "The Role of Information Structure in Word Order Variation and Word order Change". In R. Hinterhölzl and S. Petrova (eds.) *New Approaches to Word Order Variation and Change in Germanic*. 45-67. Berlin, Mouton de Gruyter.
- Hinterhölzl, R. 2009b. "A Phase-based Approach to Modification and Word Order in Germanic". *Syntax* 12:3, 242-284.
- Hinterhölzl, R. 2006a. *Scrambling, Remnant Movement and Restructuring in West Germanic*. Oxford, OUP.
- Hinterhölzl, R. 2006b. "The Phase Condition and cyclic Spell out: Evidence from VP-

- topicalization”. In M. Frascarelli (ed.) *Phases of Interpretation*. 237-259. Berlin, Mouton de Gruyter.
- Hinterhölzl, Roland and S. Petrova (2005). Rhetorical Relations and Verb Placement in Early Germanic Languages: Evidence from the OHG Tatian Translation. In M. Stede (ed.) *Saliency in Discourse: Multidisciplinary Approaches to Discourse*. Münster, Stichtus/Nodus. 71-79.
- Hinterhölzl, R. 2004a. “Scrambling, Optionality and Non-lexical Triggers”. In A. Breitbarth and H. van Riemsdijk (eds.) *Triggers*. Studies in Generative Grammar 75. 173-203. Berlin, Mouton de Gruyter..
- Hinterhölzl, R. 2004b. “Language Change versus Grammar Change: What diachronic data reveal about the distinction between core grammar and periphery”. In E. Fuss and C. Trips (eds.) *Diachronic Clues to Synchronic Grammar*. 131-160. Linguistik Aktuell 72. Amsterdam, John Benjamins.
- Hinterhölzl, R. and D. Pili. 2004. “Argument shift across language types”. In P. Bottari (ed.) *Proceedings of XXVIII IGG*, Lecce, Congede Editore.
- Hinterhölzl, R. 2002a. “Remnant Movement and Partial Deletion”. In A. Alexiadou et al. (eds.) *Dimensions of Movement: From Features to Remnants*. 127-150. Amsterdam, John Benjamins.
- Hinterhölzl, R. 2002b. “Parametric Variation and Scrambling in English”. In W. Abraham and J.W. Zwart (eds.) *Studies in Comparative Germanic Syntax*. 131-150. Amsterdam, John Benjamins.
- Hinterhölzl, R. 2001. “Event-related Adjuncts and the OV/VO Distinction”. In K. Megerdooimian and L. Bar-el. (eds) *WCCFL 20 Proceedings*. 276-289. Somerville, Cascadilla Press.
- Hinterhölzl, R. 2000. “Licensing Movement and Stranding in the West Germanic OV-languages”. In P. Svenonius (ed.) *The Derivation of VO and OV*. 293-326. *Linguistics Today*, 31. Amsterdam, John Benjamins.
- Hoop, H. de. 1992. “Case Configuration and Noun Phrase Interpretation”. Phd. Dissertation, University of Groningen.
- Hornstein, N. 1994. “An argument for minimalism: the case of antecedent-contained deletion”. *Linguistic Inquiry* 25, 455-480.
- Hroarsdottir, T. 1998. *Verb Phrase Syntax in the History of Icelandic*. Doctoral dissertation: University of Tromsøe.
- Hroarsdottir, T. 2002. “Explaining language change: A three step process”. *LIP* 19, 103-141. Papers from the workshop “Language change from a Generative Perspective”. Thessaloniki, February 2002.

- Jacobs, J. 1997. „I-Topikalisierung“. *Linguistische Berichte* 168, 91-133.
- Kavanagh, J. 1979. “The Verb Final Rule in German: A Diachronic Analysis of Surface Structure Constraints”. Doctoral dissertation: University of Michigan.
- Kayne, R. (1994). *The Antisymmetry of Syntax*. Linguistic Inquiry Monograph 25. Cambridge, Massachusetts, MIT Press.
- Kayne, R. 1998. “Overt versus covert movement”. *Syntax* 1:128-191.
- Kemenade, A. van. 1987. *Syntactic Case and Morphological Case in the History of English*. Dordrecht, Foris.
- Kemenade, A. van and N. Vincent. 1997. *Parameters of Morphosyntactic Change*. Cambridge, CUP.
- Kennedy, C. 1997. “Antecedent-contained deletion and the syntax of quantification”. *Linguistic Inquiry* 28, 662-688.
- Koopman, W. 1994. “The Order of Dative and Accusative Objects in Old English”. Ms., University of Amsterdam.
- Koster, J. 1995. “Predicate Incorporation and Word Order in Dutch”. In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi and R. Zanuttini (eds.) *Paths towards Universal Grammar*. 255-277. Washington, D.C., Georgetown University Press.
- Krifka, M. 1984. “Fokus, Topik, syntaktische Struktur und semantische Interpretation”. Universität Tübingen.
- Kroch, A. 1989. “Reflexes of Grammar in Patterns of Language Change”. *Language Variation and Change* 1:199-244.
- Kroch, A. and A. Taylor 1997. “Verb Movement in Old and Middle English: Dialect Variation and Language Contact”. In A. van Kemenade and N. Vincent (eds.) 1997. 297-325.
- Kuno, S. 1976. “Subject, theme, and the speaker’s empathy – A reexamination of relativization phenomena”. In C. Li (ed.) *Subject and Topic*. 417-444. New York: Academic Press.
- Lambrecht, K. 1994. *Information Structure and Sentence Form: Topic, Focus and the mental representations of discourse referents*. Cambridge, CUP.
- Larson, R. 1988. “On the Double Object Construction”. *Linguistic Inquiry* 19: 335-391.
- Lenerz, J. 1984. *Syntaktischer Wandel und Grammatiktheorie*. Tübingen, Niemeyer.
- Lightfoot, D. 1991. *How to Set Parameters*. Cambridge, Massachusetts, MIT Press.
- Lightfoot, D. 1999. *The Development of Language. Acquisition, Change and Evolution*. Oxford, Blackwell.

- Mahajan, A. 1994. "Toward a Unified Theory of Scrambling". In N. Corver and H. v. Riemsdijk (eds.) *Studies on Scrambling*. 301-330. Berlin, Mouton de Gruyter
- Maling, J. and A. Zaenen. 1990. *Modern Icelandic Syntax*. Syntax and Semantics 24. San Diego/New York, Academic Press.
- May, R. 1985. *Logical Form: its structure and derivation*. Cambridge, Massachusetts, MIT Press.
- Molnar, V. 2002. "Contrast – from a contrastive perspective". In H. Hallelgard, S. Johansson, B. Behrens and C. Fabricius-Hansen (eds.) *Information Structure in a cross-linguistic perspective*. 147-162. Amsterdam and New York: Rodopi.
- Näf, A. 1979. *Die Wortstellung in Notkers Consolation*. Berlin/New York: Mouton de Gruyter.
- Neeleman, A. 1994. "Complex Predicates". PhD. Dissertation, Utrecht University.
- Nespor, M. and I. Vogel. 1986. *Prosodic Phonology*. Dordrecht, Foris Publications.
- Nespor, M., M. T. Guasti and A. Christophe. 1996. "Selecting Word Order: The Rhythmic Activation Principle". In: U. Kleinhenz (ed.) *Interfaces in Phonology*. 1-26. Berlin, Akademie Verlag.
- Nuñez J. 1995. "The Copy Theory of Movement and Linearization of Chains in the Minimalist Program". Ph.D. Dissertation, University of Maryland.
- Parsons, T. 1990. *Events in the semantics of English: A study in subatomic semantics*. Cambridge, MA, MIT Press.
- Pesetsky, D. 2000. *Phrasal Movement and its kin*. Cambridge, Massachusetts, MIT Press.
- Pesetsky, D. 1995. *Zero Syntax*. Cambridge, Massachusetts, MIT Press.
- Pesetsky, D. 1987. "Wh-in-situ: Movement and unselective binding". In E. Reuland and A. ter Meulen (eds.) *The representation of (In)Definiteness*. 98-129. Cambridge, Massachusetts, MIT Press.
- Pierrehumbert, J. 1980. "The Phonology and Phonetics of English Intonation". Ph.D. Dissertation, MIT.
- Pierrehumbert, J. and J. Hirschberg. 1990. "The Meaning of Intonational Contours in the Interpretation of Discourse". In P. R. Cohen, R. Philip, J. Morgan and M. E. Pollack (eds.) *Intentions in Communication*. 271-311. Cambridge, Massachusetts, MIT Press.
- Pintzuk, S. and A. Kroch. 1989. "The Rightward Movement of Complements and Adjuncts in the Old English of *Beowulf*". *Language Variation and Change* 1:115-43.

- Pintzuk, S. 1991. "Phrase Structures in Competition: Variation and Change in Old English Word Order". Doctoral dissertation, University of Pennsylvania.
- Pintzuk, S. 1996. "Old English Verb-Complement Word Order and the Change from OV to VO". *York Papers in Linguistics* 17:241-264.
- Pintzuk, S. 1999. *Phrase Structures in Competition: variation and change in Old English word order*. New York, Garland.
- Prince, E. 1981. Towards a Taxonomy of Given-New Information. In P. Cole (ed.). *Radical Pragmatics*. 223-255. New York, Academic Press.
- Reinhart, T. 1981. "Pragmatics and Linguistics: An Analysis of Sentence Topics". *Philosophica* 27:53-94.
- Rizzi, L. 1997. "The Fine Structure of the Left Periphery". In L. Haegeman (ed.) *Elements of Grammar. Handbook in Generative Syntax*. 281-337. Dordrecht, Kluwer Academic Publishers.
- Rizzi, L. 1990. *Relativized Minimality*. Cambridge, Massachusetts, MIT Press.
- Roberts, I. 1997. "Directionality and word order change in the history of English". In A. van Kemenade and N. Vincent (eds.) *Parameters of morphosyntactic change*. 397-426. Cambridge, CUP.
- Robinson, O. W. 1997. *Clause Subordination and Verb Placement in the Old High German Isidor Translation*. Heidelberg, C. Winter.
- Selkirk, E. 1984. *Phonology and syntax: The Relation between Sound and Structure*. Cambridge, Massachusetts, MIT Press.
- Selkirk, E. 1995. "The prosodic structure of function words". *University of Massachusetts occasional papers* 18: *Papers in Optimality Theory*, 439-469. GLSA, Amherst.
- Schweikert, W. 2005. *The order of prepositional phrases in the order of the clause*. Amsterdam, John Benjamins.
- Smith, C. 1991. *The Parameter of Aspect*. Dordrecht, Kluwer.
- Stockwell, R. 1977. "Motivations for exbraciation in Old English". In C. Li (ed.) *Mechanisms of Syntactic Change*. 291-314. Austin, University of Texas Press.
- Stowell, T. 1996. "The Phrase Structure of Tense". In J. Rooryck and L. Zaring (eds.) *Phrase structure and the lexicon*. 271-286. Dordrecht, Kluwer.
- Stroik, T. 1990. "Adverbs as V-sisters". *Linguistic Inquiry* 21:654-661.
- Strawson, P. 1964. "Identifying reference and truth values". *Theoria* 30: 96-118.
- Suner, M. 1988. "The role of AGR(eement) in Clitic Doubled Constructions". *NLLT* 6: 391-434.

- Truckenbrodt, H. 1999. "On the Relation between Syntactic Phrases and Phonological Phrases". *Linguistic Inquiry* 30: 219-255.
- Vallduvi, E. 1992. *The informational component*. New York, Garland.
- Van den Wyngaerd, G. 1989. "Object Shift as an A-movement Rule". *MIT Working Papers in Linguistics* 11, 256-271.
- Vendler, Z. 1967. "Verbs and Times". In Z. Vendler (ed.) *Linguistics in Philosophy*. 97-121. Ithaca, Cornell University Press.
- Vikner, S. 1995. "V to I-movement and inflection for person in all tenses". *Working Papers in Scandinavian Syntax* 55, 1-27.
- Vikner, S. 1994. "Scandinavian Object Shift and West Germanic Scrambling". In N. Corver and H.v. Riemsdijk (eds.) *Studies on Scrambling*. Berlin, Mouton de Gruyter
- Wagner, M. 2005. "Asymmetries in Prosodic Domain Formation". *MIT Working Papers in Linguistics* 49, 329-367. Cambridge, Massachusetts.
- Wackernagel, J. 1892. "Über ein Gesetz der indogermanischen Wortstellung". *IF* 1: 333-435.
- Zwart, J. W. 1993. "Dutch syntax: a minimalist approach". Doctoral dissertation, University of Groningen.