

Eliminating Positions:
Syntax and semantics of sentence modification

Published by
LOT
Trans 10
3512 JK Utrecht
the Netherlands

Phone: +31 30 253 6006
fax: +31 30 253 6000
e-mail: lot@let.uu.nl
<http://www.lot.let.uu.nl/>

Cover illustration: Image generated by Owen Ransen's Gliftic graphics software

ISBN 90-76864-34-9
NUR 632

Copyright © 2003 Øystein Nilsen. All rights reserved.

Eliminating Positions:
Syntax and semantics of sentence modification

Posities Verwijderen:
Syntaxis en semantiek van zinsmodificatie
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de Rector Magnificus,
Prof. Dr. W. H. Gispen
Ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen
op vrijdag 24 januari 2003
des middags te 12.45 uur

door

Øystein Nilsen

geboren op 29 maart 1971 te Ringerike

Promotor: Prof. dr. E. J. Reuland

Contents

Acknowledgements	vii
1 Preliminaries	1
1.1 Introduction and overview	1
1.2 Positions	3
1.3 Syntactic vs. ontological hierarchies	7
1.3.1 Cinque's Universal Hierarchy	7
1.3.2 Ernst's Fact-Event object calculus	13
1.4 Antisymmetry and processing	19
1.4.1 Introduction	19
1.4.2 Antisymmetry	19
1.4.3 A processing account	23
1.4.4 GU20 and Penultimate position	26
1.4.5 A&N's arguments for right-adjunction	31
1.5 Summary	35
2 Domains for Adverbs	37
2.1 Introduction	37
2.2 Adverbs and Polarity	39
2.2.1 NPIs and speaker oriented adverbs	39
2.3 Approaches to polarity items	41
2.3.1 Monotonicity and veridicality	41
2.3.2 Why DE?	51
2.3.3 Widen up and strengthen	51
2.4 <i>possibly</i> : Shrink, but don't weaken!	54
2.4.1 Modal bases	54
2.4.2 Entrenched beliefs	56
2.4.3 The status of <i>possibly</i> as a PPI	59

2.5	Prospects and consequences	66
2.5.1	Short Verb Movement	67
2.5.2	Semantic selection	72
2.5.3	Summary	75
3	V2 and Holmberg's Generalization	77
3.1	Introduction	77
3.2	Some data	79
3.2.1	V2-violations with focus particles	79
3.2.2	Pronoun Shift	81
3.3	First approximation	83
3.3.1	Problems	83
3.4	More data: how initial is the initial position?	85
3.5	Second approximation	87
3.5.1	Root clauses	88
3.5.2	More problems	97
3.6	German	99
3.6.1	Hallman's analysis	101
3.6.2	Müller's V2 as <i>v</i> P first	102
3.7	Third approximation: V2 without positions	106
3.7.1	the analysis	108
3.7.2	Σ P fronting	120
3.8	Summary	121
4	Verb movement, Scope and Scrambling	123
4.1	Introduction	123
4.2	A Bobaljik Paradox for SVM	126
4.3	English and French	130
4.4	Italian verb scrambling and VP scrambling	138
4.5	Summary	139
A	Degrees and SOA	141
A.1	The <i>veldig</i> \sim <i>vis</i> Competition	141
A.1.1	A morphological analysis?	142
A.1.2	DegPs and sentence adverbs	144
	Bibliography	147
	Samenvatting in het Nederlands	155

Acknowledgements

Utrecht is a beautiful place to live and a great place to do linguistics. One of the most impressive things about Utrecht is its high density of impressive people. The Utrecht institute of Linguistics OTS is no exception in this respect. Being allowed to do my Ph.D. surrounded by such people has been a privilege.

I would like to express my warm thanks to my supervisor, Eric Reuland, for being a seemingly inexhaustible source of stimulating discussions, support and energy.

Tanya Reinhart is a great inspiration for many people at the institute, including myself. Her courses and her highly constructive comments have been instrumental for my work at several junctures.

My stay in Venice during my Master studies was a turning point in my linguistic training. I am deeply grateful to Guglielmo Cinque for his time and patience in discussions with me about linguistics and adverbial syntax during that stay and on several occasions later.

Several people have been vital for my development as a Ph.D. student by giving courses, having meetings with me or discussing linguistics with me in corridors, pubs, trains, on email, etc. I thank Peter Ackema, Sjeff Barbiers, Raffaella Bernardi, Hagit Borer, Olga Borik, Patrick Brandt, Anne Breitbarth, Balder ten Cate, Lisa Cheng, Crit Cremers, Paul Dekker, Alexis Dimitriadis, Arnold Evers, Christophe Costa Florencio, Anastasia Giannaidou, Taka Hara, Heleen Hoekstra, Anders Holmberg, Janne Bondi Johannessen, Marit Julien, Richard Kayne, Hilda Koopman, Fred Landman, Marika Lekakou, Michael Moortgat, Richard Moot, Iris Mulders, Marie Nilsenova, Rick Nouwen, Tim Stowell, Peter Svenonius, Henriëtte de Swart, Johan Rooryck, Robert van Rooy, Susan Rothstein, Maaïke Schoorlemmer, Kriszta Szendrői, Tarald Taraldsen, Henk Verkuyl, Nadya Vinokurova, Willemijn Vermaat, Yoad Winter, Ton van

der Wouden and Henk Zeevat for contributing to my work in various important ways. I am certain to have left out some people that deserve mention: I apologize for my forgetfulness and thank them, too.

Special thanks are due to Esther Kraak for helping me take care of practical things in the last phase of writing up the dissertation, and to Willemijn Vermaat and Elise de Bree for translating the Dutch summary.

For excellent company, and, in some cases, even home-sharing I thank my friends Patrick, Richard, Raffa, Manuel, Francesco, Frederic, Sue, Sylvain, Kathleen, Reza, Lorna, Karin, Derek, Alireza, Luciano, Federica, Kriszta, Balazs, Rick, Alexis, Amy, Olga and Femke.

Finally, I thank Marika with love for being patient with me, for helping me in all sorts of ways, and for many, many other things.

CHAPTER 1

Preliminaries

1.1 Introduction and overview

This dissertation presents a view of “clause structure” which essentially amounts to denying its existence. It is argued that phenomena like adverb placement, short verb movement and verb second, which form core cases in support of contemporary notions of CP-fields and IP-fields in the clause, are better handled in other ways which do not resort to the notion of arbitrary selectional sequences of (functional) syntactic heads. The alternatives which are pursued here take Chomsky’s notion of a “bare output condition” quite literally (Chomsky, 1995, 1999, 2001). Schematically, the approach can be described as follows: given an expression α with a limited distribution δ , i.e. α can’t occur in a non- δ environment, call these $\bar{\delta}$, one searches for a set of properties of α , δ and $\bar{\delta}$ which, in conjunction with independently motivated assumptions about the conceptual-intentional and sensorimotor interfaces derives α ’s limitation to δ as a theorem. Suppose, for example, that two expressions a and b can only occur in one order, namely ab . Suppose furthermore that the relative ordering of expressions of the relevant kind invariably affects their semantic scope, such that precedence maps directly onto semantic scope. In currently available theories, we essentially have two major ways of dealing with this. On the one hand, we could set up a clausal hierarchy, such as in figure 1.1 and explain the fact that a must precede and outscope b by stating that a must occupy spec-XP, while b must occupy spec-YP (and stating that semantic scope is determined by c-command in the usual way). On the other hand, we could try to identify some semantic

properties of *a* and *b* which would explain why *b* cannot outscope *a*, state that scope = c-command, and thus derive the absence of the order *ba* from that. If tenable, it seems that the latter approach is better, because it involves fewer stipulations than the former. In fact the account in terms assigning positions leaves one with the impression that we are describing facts, not deriving them.

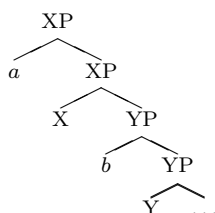


Figure 1.1: Hierarchy for *ab*

The dissertation is organized as follows. The rest of this chapter is devoted to two different topics. First, I will discuss some of the reasons why I think the standard notion of “clausal architecture” must be discarded. Next, I will argue that this does not bear on whether or not one should adopt Kayne’s antisymmetry theory. In fact, I will adopt antisymmetry, because it is the only theory I am aware of which derives the universal absence of certain word orders. I will argue that proposed alternatives, like Ackema and Neeleman (2002), at least as they stand, cannot be taken as a serious competitor. This, in turn, will be important for how I derive the correspondence between certain word orders and certain semantic scopal orders.

Chapter 2 is devoted to arguing that adverb ordering can be given a semantic account. The proposal is essentially that “high” adverbs are positive polarity items, and that this derives the distribution of these adverbs to a surprisingly large extent. Furthermore, adverb orderings which have been reported in the literature to be ungrammatical (including Nilsen (2000)) are often quite acceptable. I believe part of the problem has been that it is quite difficult to conjure up pragmatically felicitous examples containing several adverbs.¹ Hence it might not be surprising that, when one succeeds in doing this, the examples one has created are such that the adverbs are only felicitous in one particular order. In order to sidestep this problem, I have adopted the method of searching for strings of adverbs on the internet. I argue that one can explain why the adverb *possibly* behaves like a positive polarity item by adopting the theory developed for NPIs like *any* by Kadmon and Landman (1993); Krifka (1995); Chierchia (2001).

In Chapter 3, I develop a novel approach to verb second (V2) after arguing that the standard approach in terms of verb movement to C with sub-

¹Readers who are incredulous about this point are invited to come up with pragmatically sensible sentences containing the three adverbs *usually*, *soon* and *already* in all of the six possible orders.

sequent topicalization runs into problems with certain v2-violations involving focus-sensitive operators, like *bare* ‘only’. Then it is argued that the standard approach groups facts together in an incorrect way. For example, it links obligatory subject-verb inversion to verb movement to a high position, i.e. C. However, we will see that there are cases of obligatory inversion where the verb is arguably much lower than this. The approach is developed in three consecutive “approximations”, which can handle the problems discussed above, in addition to leading to a very simple account of Holmberg’s Generalization (Holmberg, 1986, 1999) concerning the interplay between argument shift and verb movement in the Scandinavian languages. The idea is that shifted arguments cannot cross the verb (or other phonetically realized material from the VP) because it is the VP itself (or something bigger) which moves. In other words, it is argued that weak pronoun shift, as well as head movement of the finite verb to C, do not exist. The fact that these elements tend to end up in the “left periphery” is that a larger constituent, dubbed Σ P, must move to the first position. Material occurring lower down in the clause must then have been extracted from Σ P prior to Σ P fronting.

Chapter 4 addresses arguments for positions from verb placement in Romance and English, concluding that, to a large extent, (short) verb movement must be treated as a scope/information structure related phenomenon, to be treated on a par with scrambling. It is suggested that obligatory verb movement around adverbs, as seen in French (Pollock, 1989) does not exist. In other words, when a given verb form *must* precede adverbs, this is to be treated as (remnant) movement of a larger constituent containing the verb to the left periphery, much as fronting of Σ P discussed in chapter 3.

The overarching theme is that phenomena that have been treated in terms of selectional sequences of functional heads need not, and sometimes cannot be analyzed in this way.

1.2 Positions

The notion of a ‘syntactic position’ as used in contemporary generative syntax is a residue of structuralist analysis of clause structure in terms of “fields”, e.g. the analysis of Germanic verb-second languages into a Vorfeld, a Mittelfeld and a Nachfeld. The latter trichotomy can be seen to correspond more or less directly to the more contemporary notions of CP, IP and vP, respectively. Thus, one could say that CP/IP/vP represents an attempt to flesh out the internal structure of the structuralist fields. In the last fifteen years or so, syntactic research has shown, conclusively, I think, that if we want to describe clause structure in these terms, we need to split CPs, IPs and vPs into several distinct functional projections (Pollock, 1989; Rizzi, 1997; Cinque, 1999), generating a plentitude of positions, thus leading again to notions of “CP-fields”, “IP-fields” and “VP-fields”. This line of thought has been pursued intensively over the last few years under the heading of “clausal cartography”. The task, under such a

view, for syntactic theory is to create a detailed “map” of all the positions in the clause, and to assign expressions to these positions. However, one might wonder whether it would not be possible to go further. Given a clausal cartography, one might want to ask why the clause should contain exactly the positions it does, and why they should relate to one another in the way that they do. The present thesis attempts to address the last of these two questions. In fact, the answer I will propose to this question suggests an answer to the other one: The clause does not contain any positions.

In order to see what is meant by this, let us put the notion of a “position” under some scrutiny. In the prevalent theoretical frameworks of the sixties, clausal positions naturally arose as a byproduct of the use of phrase structure (PS) rules. Thus, the PS-rules in (1) give rise to the tree in figure 1.2, with a COMP position, an NP (subject) position, an AUX position, and a VP position.

- (1) a. $\bar{S} \Rightarrow \text{COMP}, S$
 b. $S \Rightarrow \text{NP}, \text{AUX}, \text{VP}$

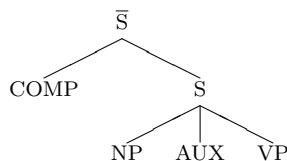


Figure 1.2: Positions generated by (1)

With the advent of \bar{X} -theory, the burden of generating designated positions was shifted to the lexicon. Thus, one made use of subcategorization frames, θ -grids and (s/c-) selectional properties, all of which are qualities of lexical items. According to this view, trees are projected from lexical items in accordance with the ‘Projection Principle’, governing how lexical information is reflected in syntactic structure, and category-neutral PS-rules, like the following, where YP is an “adjunct”, ZP is a “specifier” and UP is a “complement”.

- (2) a. $XP \Rightarrow YP; XP$
 b. $XP \Rightarrow ZP; \bar{X}$
 c. $\bar{X} \Rightarrow X^0; UP$

Now the COMP NP, AUX, etc. positions of figure 1.2 must be projected from lexical items. This lead to the postulation of the functional heads C^0 and I^0 , etc. The idea is that these project positions in accordance with (2) and that C^0 syntactically selects for IP.

\bar{X} -theory is essentially a set of stipulated PS-rules. In the nineties theories were proposed that derive the stipulations from more fundamental properties of syntactic composition (Kayne, 1994; Chomsky, 1994; Brody, 2000), but the core of the approach remains the same, i.e. these approaches all generate positions by means of selectional sequences in interaction with general axioms for

composition.² More often than not, a particular functional head X^0 serves the sole purpose of generating a position by altering the label of the tree, i.e. the label of the complement of the head is different from the label projected from the head. As pointed out by Starke (2001), the empty heads can be omitted if we allow phrases (i.e. “specifiers”) to alter the labels in this fashion themselves. The idea is that since “specifiers” of X invariably share a feature with X , the feature of the specifier could project the label itself. Thus instead of the leftmost tree in figure 1.3, we would have the rightmost one.³



Figure 1.3: \bar{X} -structure and its corresponding Starke structure

In such a system, the positions cannot be ordered by selection in the standard sense. But the selectional properties of an abstract head X^0 is essentially a stipulated property, crucial only inasmuch we use the head to generate a particular position. So Starke suggests that we might as well stipulate the sequence in which labels are allowed to change separately, without any loss of insight or elegance, his “FSEQ”. An FSEQ $\langle f_n, \dots, f_n \rangle$ is essentially an ordering of features. If features are thought of as corresponding to properties (i.e. sets) of expressions, we see that what FSEQ does is essentially to stipulate the distribution of expressions. It would seem overly pessimistic, then, to suppose that FSEQ cannot be derived from more fundamental properties of the classes of expressions that it orders.⁴ The work with this thesis started out as an effort to derive certain parts of FSEQ from independent (semantic) factors. However, after a while it became clear that there are serious problems with the very notion of an FSEQ, problems that seem to indicate that the notion not only is conceptually problematic (i.e. stipulative), but, in fact, empirically wrong.

FSEQ is standardly assumed to be linear. As we shall see, there are distributional phenomena which cannot be accommodated into a linear FSEQ, in fact, not even into a partial order. Hence some other theory of distribution must supplement FSEQ. The alternative to stipulating a linear FSEQ is, of course to derive the distribution of different classes of expressions from independently motivated properties of these expressions.⁵ In other words, we could seek to

²Of course, the question of selectional sequences is orthogonal to the question of how to derive \bar{X} -theory. For example, one can perfectly well adopt Kayne’s antisymmetry framework without assuming selectional sequences to be playing any role.

³See Starke (2001) for technical details of how this is implemented.

⁴Please note that I am not attributing to Starke (or anybody else) the view that FSEQ cannot be derived. I am merely pointing out that, if it were to turn out that must be taken as primitive, this would be a very disappointing state of affairs for linguistic theory. In fact, I take this to be obvious.

⁵Some alternative frameworks generate non-linear “FSEQs”. One example is Type Logical

derive FSEQ from a proper characterization of the features themselves. Then, there would be no need to stipulate an FSEQ or empty heads which are only there to project positions. Given such an approach, there would be no residue of structuralist “positions”, except, perhaps, in a purely descriptive sense.

Apart from being essentially descriptive, positional analyses generate some deep puzzles. As has been noted by Jonathan Bobaljik, the following two observations combine to yield a paradoxical situation if both adverb order and argument order are to be analyzed in terms of FSEQs. i) Scrambling/Argument Shift: In Germanic V2 languages (except Danish⁶), the subject of the clause can occupy any position in a string of adverbs, as long as the relative ordering of the adverbs remains unaltered. The same can be seen to hold of direct and indirect objects. ii) In the same languages, given the string $\langle s, io, do \rangle$, an adverb can occur anywhere in the string, as long as the ordering of the arguments remains unaffected. It should be obvious that no linear FSEQ can accommodate at the same time arguments and adverbs in such a way that both observations hold. Bobaljik, following Áfarli, concludes that adverbs must be Z-axis elements in 3-dimensional graphs. I think one could keep the number of dimensions of syntactic representations low if we abandon the view that we should analyze such distributional phenomena in terms of linear FSEQs in the first place.⁷

Of course eliminating positions in the manner outlined, can only be done by reanalyzing phenomena which have been dealt with in this way. This is obviously more work than one can undertake in one dissertation. The more modest goal of the current dissertation is to propose alternatives to positional analyses in three domains which I take to be crucial cases, namely the analysis of verb second, the analysis of adverb placement and of (short) verb movement.

Grammar (Moortgat, 1996), where the unary connectives give rise to a cube of “positions” which has been put to uses which are similar to the generative notion of FSEQ. See in particular the derivation of polarity phenomena in Bernardi (2002). In this sense, TLG has an “FCUBE” rather than an FSEQ. An important difference is that, while the generative FSEQ is stipulated, the FCUBE of TLG is a theorem of the algebraic laws governing the unary (and binary) grammatical connectives. Given that the base logic of TLG is Curry-Howard isomorphic to the typed lambda calculus, it seems that one could actually have a semantic motivation for which corner of the cube an expression is assigned to. It is possible that such a theory can be developed in ways compatible with the ideas of e.g. adverb ordering to be proposed in the present work. See Nilsen (2001) for some steps in that direction.

⁶In this language, subjects and weak pronouns precede all adverbs in the *mittelfeld*, whereas other kinds of arguments follow all (pre-VP) adverbs. In other words, Danish does not have ‘object shift/scrambling of full DPs’. This is widely and falsely believed to be a property of all Mainland Scandinavian languages. See Nilsen (1997) and chapter 3 for discussion.

⁷The number of dimensions of a graph is obtained by counting the number of primitive relations the graph represents. In the case of syntactic trees, they are 2D inasmuch they encode both precedence and dominance and neither relation is reducible to other theoretical notions. If one is derivative of the other, it is 1D. Adding a third primitive relation, call it “beyond”, would require heavy empirical ammunition. See Nilsen (1997) for discussion of the proposal in Áfarli (1996)

⁹This is the hierarchy he proposes in his (1999) book. Later, he has proposed extensions and refinements of it based on the distributional properties of 'restructuring' verbs and adverbial PPs.

Cinque shows that adverbs obey substantially the same ordering restrictions in languages as diverse as Italian, Serbo-Croatian, Mandarin Chinese, Hebrew and Norwegian (and lots of other languages). He then shows that, in languages which express “adverbial notions” as verbal affixes, the ordering of affixes is entirely consistent with that found for adverbs, given standard assumptions about morphology. Next, he shows that languages where the notions in question manifest themselves as free functional morphemes order these in the same way. All of this goes to show that there is something fundamentally universal about the relative ordering in which functional material occur in the clause. He then shows that, given a rigidly ordered¹⁰ sequence of adverbs, a_1, \dots, a_n , one finds detailed variation within Romance languages with respect to where one can place verbs. In other words, whereas one language L , might allow past participles to occupy any position in the sequence, another language L' would allow anything, except for the very last position. Yet another language L'' might exclude the two last positions in the adverbial sequence, while allowing all other positions and so on. This gives rise to the picture in the following table, where a_i are adverbs and “ \checkmark ,” “ $*$ ” represent possible and impossible positions for the verb, respectively.

L_1	\checkmark	a_1	\checkmark	a_2	\checkmark	a_3	\checkmark
L_2	\checkmark	a_1	\checkmark	a_2	\checkmark	a_3	$*$
L_3	\checkmark	a_1	\checkmark	a_2	$*$	a_3	$*$
L_4	\checkmark	a_1	$*$	a_2	$*$	a_3	$*$

Table 1.1: Verb-adverb distribution

The standard treatment of differential verb placement is head movement to distinct functional heads. Thus, Cinque argues that the pattern in table 1.1 indicates that there must be head positions between each of the adverbs; otherwise there would be no means to describe the *difference* between the languages in question, while capturing the relevant generalizations. In particular, Cinque argues that there is no language where the verb can precede, say, a_1 but none of the other adverbs. Thus, we arrive at a structure like the one in figure 1.4, where L_i now represents the *lowest possible* position for the verb in the corresponding language.

Suppose that we denote the relevant (descriptive) ordering relation of functional material as “ \prec ” where $x \prec y$ iff x can precede y . Cinque’s approach leads one to expect \prec to be a linear ordering, i.e. a transitive ($\forall x, y, z (x \prec y \wedge y \prec z) \rightarrow x \prec z$), asymmetric ($\forall x, y ((x \prec y \wedge y \prec x) \rightarrow x = y)$) and connected ($\forall x, y (x \prec y \vee y \prec x)$) ordering of *classes* of functional material. We need to talk about ordering of classes, rather than individual expressions, because it could happen that two functional expressions represent opposite values for the

¹⁰By “rigidly ordered” I mean that the adverbs in question *can* occur in this order, and, furthermore, that this is the *only* order in which they can cooccur, and that this ordering is constant across languages.

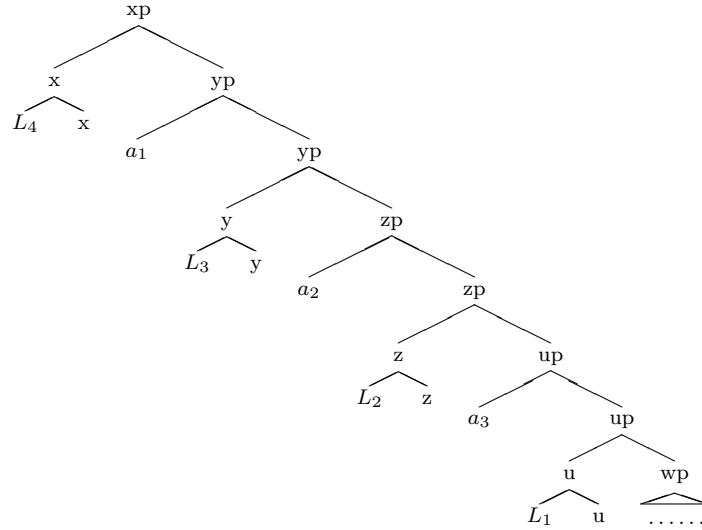


Figure 1.4: syntactic hierarchy for the pattern in table 1.1

same functional notion, e.g. the pair *always* and *never*. These would then be predicted not to be able to cooccur. Linearity comes about because each functional projection is a syntactic complement of another one (except the highest one, of course), and, because of binary branching, each functional head can have at most one complement.

In the face of apparent counterexamples to asymmetry, Cinque would have to argue (and indeed, he does for some cases) that one of the expressions involved can occupy two distinct positions, and therefore must be treated as belonging to two distinct (though possibly semantically related) classes. If two expressions cannot cooccur (counterexample to connectedness), they must either belong to the same class, or one has to show that there are independent, possibly semantic reasons for this. Cinque shows that transitivity holds for some cases. Of course, demonstrating this for every possible combination would be a nearly impossible task. If counterexamples to transitivity exist, they would be particularly hard to handle for the approach under discussion. Such a counterexample would always be of the form that, for some triplet of adverbs a_1, a_2, a_3 , it holds that $(a_1 \prec a_2 \wedge a_2 \prec a_3 \wedge a_1 \not\prec a_3)$. In this case, one would have to develop a complementary theory \mathcal{T} of adverb ordering which would explain why $a_1 \not\prec a_3$. More in particular, no amount of assigning positions to expressions in a linear sequence could accommodate such a triplet. Given that we do not, in general want to explain the same phenomenon twice, this would inevitably lead to a tension: one could be tempted to try to develop \mathcal{T} into a full blown alternative. Depending on its nature, such an alternative could then allow us to return to a syntactically null theory of adverb placement and

the distribution of other functional material. This is, in a nutshell, what the present thesis is all about. The real question here is about the use of syntactic *selection* and *checking*. According to Cinque, adverbs are ordered because there exists a universal inventory of functional heads, and that these are universally ordered by syntactic selection, and because the adverbs enter into checking relations with distinct functional heads. Suppose that \mathcal{T} is a semantic account of adverb ordering. Then one could abandon the selection + checking analysis of adverb ordering, *if* one could also account for patterns like the one in table 1.1, with a possibly different theory of verb placement.

A counterexample to transitivity

The Norwegian triplet of adverbs *muligens* (‘possibly’), *ikke* (‘not’) and *alltid* (‘always’) is not linearly ordered and, hence, cannot be accommodated into a sequence of functional heads. If we consider pairs of adverbs, we find that *muligens* must precede *ikke*; *ikke* must precede *alltid*; but *muligens* does not have to precede *alltid*. Thus, we have a counterexample to transitivity. This behavior is illustrated with the Norwegian examples below.¹¹

- (4) a. Ståle har **muligens ikke** spist hvetekakene sine.
 S has possibly not eaten the-wheaties his
 “Stanley possibly hasn’t eaten his wheaties.”
- b. *Ståle har **ikke muligens** spist hvetekakene sine.
 S has not possibly eaten the-wheaties his
- (5) a. Ståle hadde **ikke alltid** spist hvetekakene sine.
 S had not always eaten the-wheaties his
 “Stanley hadn’t always eaten his wheaties.”
- b. *Ståle hadde **alltid ikke** spist hvetekakene sine.
 S had always not eaten the-wheaties his

Natural examples where *alltid* ‘always’ precedes *muligens* ‘possibly’ are not frequent, but existent. Their low frequency probably has pragmatic reasons: One does not usually talk about the frequency with which something is epistemically possible. The English example (6a) was found on the internet, part of an advertizement for an internet game. It contrasts minimally with (6b).

¹¹In English the sequence *didn’t possibly* may be slightly better than the sharply ungrammatical Norwegian example above, although most English speakers find it degraded. Certainly, English allows examples like (i), with two modals, whereas the corresponding Norwegian example (ii) is still sharply out.

(i) Stanley couldn’t possibly eat his wheaties.

(ii) *Ståle kunne ikke muligens spise hvetekakene. (‘S could not possibly eat the-wheaties’) Similarly, in English the sequence *always not* might be marginally acceptable while the Norwegian counterpart (*alltid ikke*) is sharply ungrammatical.

The Norwegian translation of (6a), i.e. (7a) is also grammatical, and in this case, too, there is a contrast between *alltid* ‘always’ and *aldri* ‘never’ (7b).¹²

- (6) a. This is a fun, free game where you’re **always possibly** a click away from winning \$1000!
- b. ?? This is a fun, free game where you’re **never possibly** further than a click away from winning \$1000!
- (7) a. Dette er et morsomt, gratis spill hvor spillerne **alltid**
this is a fun free game where the-players always
muligens er et klikk fra å vinne \$1000!
possibly are one click from to win \$1000
- b. ?? Dette er et morsomt, gratis spill hvor spillerne **aldri**
this is a fun free game where the-players never
muligens er lenger enn et klikk fra å vinne \$1000!
possibly are further than one click from to win \$1000

This fits our description of counterexamples for transitivity: *Alltid* \prec *muligens* (7a); *muligens* \prec *ikke* (4a), but *alltid* $\not\prec$ *ikke* (5b).

Anticipating somewhat, my explanation why *alltid* ‘always’ can’t outscope the negation is that universal quantifiers never can (Beghelli and Stowell, 1997). Consider for example (8). This example is very odd, unless read with heavy stress on *everybody*, in which case the sentence is an emphatic denial of the sentence *everybody showed up*, i.e. the negation outscopes the universal, not the other way around.

- (8) Everybody didn’t show up.

I will propose that the reason why *muligens* ‘possibly’ can’t occur under negation is that it is a positive polarity item. This proposal is developed in detail in chapter 2.

Bobaljik’s paradox

In a squib in *Glott* Bobaljik (1999) (see also Svenonius (2001)) argues that there are empirical considerations which, when taken together with the idea of a single, linear FSEQ accommodating the distribution of different kinds of expressions, lead to a paradox. His observation is that one finds phenomena where expressions of type *X* are rigidly ordered, and expressions of type *Y* are rigidly ordered, but where any $x \in X$ can precede or follow any $y \in Y$ as long as the ordering requirements internal to *X* and *Y* are both satisfied. It seems to follow that one cannot accommodate both *X*-type expressions and *Y*-type expressions in a unique linear FSEQ. This behavior is illustrated with the

¹²One could argue that one or both of the adverbs are constituent modifiers, rather than sentence modifiers and hence irrelevant for our purposes. This point is taken up in chapter 2 (page 45) for the same examples.

relative ordering of arguments and the relative ordering of adverbs in Norwegian in the examples below.

- (9)
- a. Derfor ga **Jens Kari kyllingen** tydeligvis ikke
therefore gave J K the-chicken evidently not
lenger kald.
any.longer cold
 - b. Derfor ga **Jens Kari** tydeligvis **kyllingen** ikke lenger kald.
 - c. Derfor ga **Jens** tydeligvis **Kari kyllingen** ikke lenger kald.
 - d. Derfor ga **Jens** tydeligvis **Kari** ikke **kyllingen** lenger kald.
 - e. Derfor ga **Jens** tydeligvis **Kari** ikke lenger **kyllingen** kald.
 - f. Derfor ga **Jens** tydeligvis ikke lenger **Kari kyllingen** kald.
 - g. Derfor ga tydeligvis **Jens** ikke lenger **Kari kyllingen** kald.
 - h. Derfor ga tydeligvis ikke **Jens** lenger **Kari kyllingen** kald.
 - i. Derfor ga tydeligvis ikke lenger **Jens Kari kyllingen** kald.
 - j. * Derfor ga **Jens** ikke tydeligvis **Kari** lenger **kyllingen** kald.
 - k. * Derfor ga **Jens** tydeligvis ikke **kyllingen** lenger **Kari** kald.

The problem that patterns like (9) raise is the following. We see that arguments must cooccur in one specific order (modulo topicalization and wh- movement), but that adverbs can intervene between the arguments. Moreover, we see that the adverbs must occur in a specific order, but that arguments can intervene between them. Thus, it seems hopeless to try to account for (9) in terms of specific positions for both adverbs and arguments. In other words, we cannot accommodate patterns like (9) in a linear FSEQ. Bobaljik (1999) shows that similar problems arise with respect to the relative ordering of auxiliaries on the one hand, and the relative ordering of adverbs on the other. I discuss this ‘verb-placement paradox’ in detail in chapter 4. One could rescue the unique FSEQ approach by assuming for (9) that the adverbs occupy unique specifiers, and that the arguments can scramble freely among the adverbs, subject to the requirement that scrambling must preserve the order of the arguments. A similar approach could be proposed for the verb-placement paradox, but that would seem to create problems for Cinque’s argument for functional heads on the basis of differential verb placement in Romance. It is tantamount to giving up the idea of a unique linear FSEQ. Bobaljik (1999) concludes that we should expand the theory to allow for multi-dimensional phrase markers (Áfarli, 1996). I think we can keep the number of dimensions down if we rethink the idea that the distributional phenomena under discussion should be handled in terms of FSEQs. I return to a treatment of (9) in chapter 3, and to the verb placement paradox in chapter 4.

1.3.2 Ernst's Fact-Event object calculus

Ernst (2001) gives a semantic account of adverb ordering by analyzing different classes of adverbs as modifiers of different kinds of semantic objects. By way of example, according to Ernst, an adverb like *completely* modifies EVENTS (henceforth i), whereas e.g. *not* modifies PROPOSITIONS (p), and *paradoxically* modifies FACTS (f). He then imposes the following system of type conversion on his ontological primitives, his FEO-calculus:

$$i \implies p \implies f$$

Ernst does not use type-theoretical notation. He states, for example that an adverb like *completely* is an event modifier, and uses the notation in (10a) for that, indicating that *completely* wants a constituent denoting an EVENT as a syntactic complement, and projects a syntactic node denoting an EVENT. I find this mix of syntactic and semantic notions slightly confusing, so I will assume that what Ernst intends to say could equally well be given the type-theoretical notation in (10b), i.e. *completely* is a function from event descriptions to event descriptions. The types for *not* and *paradoxically* are given in (10c-10d).

- (10)
- a. [EVENT complete [EVENT]]
 - b. completely: $\langle i, i \rangle$
 - c. not: $\langle p, p \rangle$
 - d. paradoxically: $\langle f, f \rangle$

Given the FEO-calculus, this derives the fact that these three adverbs must occur in the following order when they cooccur: *paradoxically* > *not* > *completely*. An adverb of type $\langle x, x \rangle$ can apply to an expression of type y iff $y \implies x$ by the FEO-calculus. Thus the tree in figure 1.5 represents a grammatical structure, because all the type transitions respects the FEO-calculus.

The FEO-calculus looks suspiciously similar to Starke's FSEQ. Given that the latter is just an equivalent reformulation of the selectional sequences that the FEO-calculus is intended to replace, this is a state of affairs that requires some attention. Especially so, since Ernst does assume that there are functional projections like CP, TP and PredP etc. in the clausal structure, and that these are ordered by FSEQ. His idea is that adverbs can attach freely to any functional projection as long as the semantic requirements of the adverbs (FEO-calculus) are respected. Hence, Ernst has it that, while adverbs are ordered by his FEO-calculus, the categories that the adverbs attach to are ordered by an orthogonal relation, FSEQ. Thus, if we suppose that in some language the verb moves to T, as shown in figure (1.6), we expect adverbs to be able to freely precede or follow the verb. A congenial analysis is proposed by Svenonius (2001). Suppose that the language in question allows all of the three adverbs in figure 1.5 to precede or follow the verb. This would then be analyzed as attaching the adverbs to VP or TP in accordance with the calculus. So it is predicted that when the adverbs

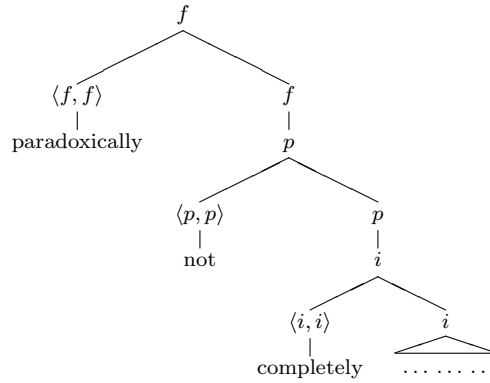


Figure 1.5: Adverbs and FEO-transitions

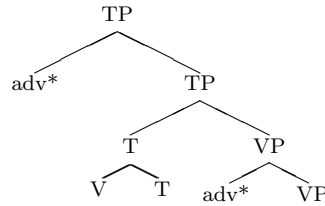


Figure 1.6: verb/adverb placement

cooccur, they must respect the ordering *paradoxically* > *not* > *completely*, but the verb can appear anywhere in the sequence. This allows Ernst and Svenonius to eliminate some functional heads. Interestingly, it also seems to solve the paradoxes discussed by Bobaljik (1999) (see above),¹³ and, as we shall see the machinery is technically capable of handling transitivity failures.

However, there are also some problems. To my mind, the most difficult one is that this setup seems to force us to abandon the view that T has semantic import. For suppose that it does, i.e. that T denotes tense, as standardly assumed. Then T should also have selectional properties, not only with respect to FSEQ, but with respect to the FEO-calculus as well. Given that, in the setup we are considering, *paradoxically* (type $\langle f, f \rangle$) can follow T, and thus must be attached to VP, and that, in that case, VP must denote a fact, it follows that T must be able to apply to a fact. Similarly, given that *completely* (type $\langle i, i \rangle$) can attach above T, it appears to follow that the result of applying T to VP must be a FEO type that *completely* can apply to, i.e. *i*. Hence, T must be of type $\langle f, i \rangle$; it takes a fact and returns an event. But this move destroys the account of adverb ordering, because if T has this type, we could attach *paradoxically* to VP, apply T to the result of that, yielding an event, and then apply *completely*

¹³See Svenonius (2001) for discussion.

to TP. This results in the ungrammatical sequence *completely V paradoxically VP*. In fact, given this type for T, all of the six logically possible orderings of the three adverbs should be fine, as long as one adverb applies below, and one above T. So the FEO-calculus would cease to have an effect. If we constrain the type for T somewhat, we can rule out some adverb orderings, but only at the cost of losing the account of verb placement. In other words, if we do that, we must have more positions for the verb to move to in order to account for the verb placement facts in the hypothetical language we are considering, thus potentially destroying the resolution of Bobaljik’s paradoxes. The only way out seems to be to assume that T mindlessly passes on the FEO-type of its complement. But that creates some problems for the view that T denotes tense, since tense presumably does care about the semantic type of its semantic argument. Hence it is not clear FSEQ and the FEO-calculus can be treated as orthogonal after all. In that case, the account for the different verb positions in terms of figure 1.6 would seem to be much less straightforward than Ernst and Svenonius assume.

Although Ernst’s calculus is linear, it does not rule out non-linearity of adverb ordering. For instance, we have seen that we could assign the type $\langle f, i \rangle$ to some adverbs, and these would then be able to precede or follow any other adverb. Furthermore, since Ernst’s calculus is stipulated, he could in principle impose any other relation on his ontological primitives, and avoid linearity of the calculus in this way. Needless to say, the latter point may also be construed as a weakness of this kind of approach: The FEO-calculus does for Ernst what syntactic selection does for Cinque. Hence Ernst’s approach does not *explain* adverb ordering unless the calculus can be derived from more fundamental properties of the particular ontological primitives he uses.

Let us see how our counterexample to transitivity can be accommodated in an Ernst-type calculus. We refer to the first element in the type $\langle x, y \rangle$ as its *argument* and write $_A(\text{adv})$ for the argument of an adverb. Similarly we write $_R(\text{adv})$ for the second element in the type, i.e. the *result* of the function. As before, we write $\text{Adv}_1 \prec \text{Adv}_2$ for “ Adv_1 can precede Adv_2 ”.¹⁴ Finally, we operate with a calculus $a_1 \leq \dots \leq a_n$ where a_i are types and \leq is a linear ordering of the types. In other words, we abstract away from Ernst’s particular interpretation of his types as “events”, “propositions”, etc. This is in order to be able to focus on the *structural* properties of his calculus. In general, the following biconditional can be seen to hold:

$$\text{Adv}_1 \prec \text{Adv}_2 \Leftrightarrow {}_R(\text{Adv}_2) \leq {}_A(\text{Adv}_1)$$

From this it follows that (contraposition of $[\Leftarrow]$ and $[\Rightarrow]$)

$$\text{Adv}_1 \not\prec \text{Adv}_2 \Leftrightarrow {}_R(\text{Adv}_2) \not\leq {}_A(\text{Adv}_1)$$

¹⁴Strictly speaking, \prec should be read *can apply after*, but since, in the relevant cases, precedence=scope, we ignore this.

and, by linearity of \leq (i.e. connectedness), that

$$\text{Adv}_1 \not\prec \text{Adv}_2 \Leftrightarrow {}_A(\text{Adv}_1) < {}_R(\text{Adv}_2).$$

We can now use the ordering facts above to reason with the types of our triplet of adverbs. The ordering facts are given on the left and the conclusions to be drawn concerning their types are given on the right.

$$\begin{array}{ll} \text{ikke} \prec \text{alltid} & {}_R(\text{alltid}) \leq {}_A(\text{ikke}) \\ \text{alltid} \not\prec \text{ikke} & {}_A(\text{alltid}) < {}_R(\text{ikke}) \\ \text{muligens} \prec \text{ikke} & {}_R(\text{ikke}) \leq {}_A(\text{muligens}) \\ \text{ikke} \not\prec \text{muligens} & {}_A(\text{ikke}) < {}_R(\text{muligens}) \\ \text{muligens} \prec \text{alltid} & {}_R(\text{alltid}) \leq {}_A(\text{muligens}) \\ \text{alltid} \prec \text{muligens} & {}_R(\text{muligens}) \leq {}_A(\text{alltid}) \end{array}$$

By compiling this into one sequence, we get the following:

$${}_R(\text{alltid}) \leq {}_A(\text{ikke}) < {}_R(\text{muligens}) \leq {}_A(\text{alltid}) < {}_R(\text{ikke}) \leq {}_A(\text{muligens})$$

If, in order to minimize the amount of types, we always read \leq as $=$, we need three basic types a, b, c such that $a < b < c$ and the following types for the adverbs which we simply read off the ordering above:

$$\begin{array}{ll} \text{ikke} & \langle a, c \rangle \\ \text{muligens} & \langle c, b \rangle \\ \text{alltid} & \langle b, a \rangle \end{array}$$

As can be seen from our discussion above, it also follows that, if we have just three types this is the *only* type assignment which would work, given a linear type conversion calculus. If we have a category S of type a , we can show that the following facts obtain.

derivable	not derivable
ikke(alltid(S))	*alltid(ikke(S))
muligens(ikke(S))	*ikke(muligens(S))
muligens(alltid(S))	
alltid(muligens(S))	
ikke(alltid(muligens(S)))	

I give the derivation of the three examples (ikke(alltid(S))), (alltid(muligens(S))) and (muligens(ikke(S))) in figure 1.7 below. Premises are written above the lines, and conclusions below. If the conclusion does not follow trivially, the relevant application of the calculus is written to the right of the line. The notation $\alpha : x$ indicates that the expression α is of type x .

The calculus $a < b < c$ is obviously isomorphic to Ernst's FEO calculus, i.e. $a = i$, $b = p$ and $c = f$. Thus, what we have derived is that Ernst could handle our problematic triplet by assigning the type $\langle i, f \rangle$ to *ikke*, $\langle f, p \rangle$ to *muligens* and $\langle p, i \rangle$ to *alltid*. In practice, Ernst usually assigns identity types to

$$\begin{array}{c}
\frac{\frac{\text{alltid} : \langle b, a \rangle \quad \frac{S : a}{S : b} a < b}{(\text{alltid}(S)) : a}}{\text{ikke} : \langle a, c \rangle \quad (\text{ikke}(\text{alltid}(S))) : c} \\
\\
\frac{\frac{\text{alltid} : \langle b, a \rangle \quad \frac{\text{muligens} : \langle c, b \rangle \quad \frac{S : a}{S : c} a < c}{(\text{muligens}(S)) : b}}{(\text{alltid}(\text{muligens}(S))) : a}}{\text{muligens} : \langle c, b \rangle \quad \frac{\text{ikke} : \langle a, c \rangle \quad S : a}{(\text{ikke}(S)) : c}} \\
\\
\frac{\text{muligens} : \langle c, b \rangle \quad (\text{muligens}(\text{ikke}(S))) : b}{}
\end{array}$$

Figure 1.7: Derivations for the non-transitive triplet

adverbs,¹⁵ i.e. they always return the same type as the type of their argument. Nothing prevents him from assigning types like $\langle i, f \rangle$, and given our discussion, it seems that he has to assign just this type to Norwegian *ikke*. This move seems to weaken the plausibility of his account, however. One might find it “plausible” or “intuitive” that e.g. *completely* should be a function from events to events. But how “intuitive” is it that the Norwegian negation *ikke* is a function from events to facts, or that *muligens* ‘possibly’ is a function from facts to propositions rather than from facts to events, for example? Are all possible type-transitions attested?

Another problem with Ernst’s system is that he does not explain what sort of things facts are. At one point he states that they are propositions that the speaker is committed to, and at another, he seems to say that they are factive propositions in the sense of being presupposed. For example, he states that the reason why *paradoxically* cannot occur under negation is that the negation is a propositional modifier, while *paradoxically* takes facts. Then he says that this expresses the intuition that one cannot negate a presupposed proposition, i.e. one cannot presuppose something to be true and and assert its negation at the

¹⁵He argues that evidentials like *obviously* map facts to events. This is to account for examples like (i), where *obviously* occurs under negation (Ernst, 2001, p107).

(i) Sally was not obviously affected by her winning the award.

It seems to me that this example might be an instance of *obviously* used as a manner adverb. Note for example that other cases of evidentials under negation appear to be less good.

(ii) John hasn’t evidently gone home.

(ii) appears to be good with a so-called meta-linguistic or contrastive negation, but these are known to be special in any event (see chapter 2). If evidentials create events from facts, Ernst must make extra assumptions to rule out cases where other event modifiers end up outscoping an epistemic adverb when there is an intervening evidential.

same time. But if we define a presupposed proposition as “a proposition which cannot occur under negation” Ernst’s explanation is circular.¹⁶ If we define it as “a proposition which is required to be part of the common ground”, it seems that *paradoxically* does not trigger this kind of presupposition. It does not appear to give rise to infelicity when the modified proposition is known to be false, for example, which is the hallmark of such presupposition triggers. Thus, (11a) is false, while (11b) is odd, because the definite article presupposes the existence of a president of the UK. Similarly, (11c) is odd because of the strong factivity of *it is sad that*.¹⁷ Is (11d) false or odd? The informants I have consulted find it false, but not odd in the same way as (11b). Interestingly, the adjectival version (11e) *does* seem to give rise to infelicity, so it seems that *paradoxical*, in fact, is a presupposition trigger, more so, apparently, than *paradoxically*. But the adjective occurs happily under negation (11f), so, even if we would grant that the presupposition of the adjective is somehow inherited by the adverb, this would not, in and of itself, explain why the latter cannot occur under negation.

- (11)
- a. There is a current president of the UK playing ping pong in the corridor.
 - b. The current president of the UK is playing ping pong in the corridor.
 - c. It is sad that there is a current president of the UK playing ping pong in the corridor.
 - d. Paradoxically, there is a current president of the UK playing ping pong in the corridor.
 - e. It is paradoxical that there is a current president of the UK playing ping pong in the corridor.
 - f. It is **not** paradoxical that someone is playing ping pong in the corridor.

Thus, although I am very sympathetic to the idea to the idea that the facts discussed by Cinque (1999) should ultimately be derived from more fundamental considerations, it does not seem to me that the primary tool to do so can be enrichment and manipulation of ontological categories. Hence, I try to develop an alternative semantic analysis of adverb ordering in chapter 2, which crucially does not rely on ontology. The analysis pursued there is essentially that several sentence adverbs are (positive) polarity items, and that this suffices to derive surprisingly many adverb ordering effects. This allows us to

¹⁶The standard definition of a presupposition is that it is an implication which is preserved under negation. Although this is clearly compatible with Ernst’s explanation, it does not offer support: the definition does not apply in this case for the very reason that *paradoxically* cannot occur under negation.

¹⁷It is “strongly” factive in the sense that it resists accommodation, thus contrasting with the verb *know* which is “weakly” factive in this sense.

maintain the positive results achieved by Ernst (2000) and Svenonius (2001), while sidestepping the difficulties their approach encounters.

1.4 Antisymmetry and processing

1.4.1 Introduction

From Kayne’s Linear Correspondence Axiom (LCA) (Kayne, 1994), it follows that phrase markers universally conform to specifier-head-complement order, and there can be no “right-adjunction”. The issue remains controversial, in particular with respect to sentence-final adjuncts. In Ackema and Neeleman (2002), they argue that right-adjunction should be maintained as a theoretical option, and that there are alternative ways of deriving the empirical results that antisymmetric analyses can boast of. In particular, Ackema and Neeleman claim that universal word-order asymmetries can be derived as effects of limitations on on-line parsing. I find this proposal a very interesting one that deserves careful consideration. To anticipate, the conclusion will be that, although there may well be other ways of formulating it which would work better, Ackema and Neeleman’s account fails to supply a viable alternative to antisymmetry.

1.4.2 Antisymmetry

The LCA states that for a phrase marker P with the set of terminals T and the set N of nonterminals,

- (12) Linear Correspondence Axiom (LCA)
 $d(A)$ is a linear ordering of T ,
 where

$$A = \{\langle n, m \rangle \in N \times N \mid n \text{ asymmetrically c-commands } m\};$$

for any $n \in N$,

$$d(n) = \{t \in T \mid n \text{ dominates } t\};$$

and for any $n, m \in N$,

$$d\langle n, m \rangle = \{\langle t_1, t_2 \rangle \in T \times T \mid t_1 \in d(n) \text{ and } t_2 \in d(m)\};$$

$$d(A) = \bigcup \{d\langle n, m \rangle \mid \langle n, m \rangle \in A\}.$$

Kayne defines (symmetric) c-command as follows:

- (13) c-command
 n c-commands m iff n, m are categories, n excludes m , and every category that dominates n , dominates m .

A nonterminal n x-projected from terminal t is a *category* iff it reflexively dominates all other x-projections of t , its *segments*. “x-projection” can vary between head-projection and phrasal projection. A node n is dominated by a category c iff all segments of c dominates n . c excludes n iff no segments of c dominates n , and c includes n iff some segment of c dominates n . Kayne shows that a number of stipulative elements of phrase structure theory (\bar{X} -theory) fall out as theorems of this set up, e.g. binary branching, one head per phrasal projection, etc. The notions he uses, i.e. c-command, segments and categories etc., are, of course, independently motivated. However, the system’s reliance on the distinction between terminals and nonterminals has been criticized in Chomsky (1994). Kayne himself has pointed out that the system would be more elegant if it had not relied on this very specific definition of c-command (Kayne, 2000). In Kayne (1994), he also notes that nothing apparently prevents c-command to map onto *postcedence* rather than precedence. His solution to this problem involves some very specific assumptions about time (i.e. that it is linear) which are not uncontroversial in physical and philosophical circles. Thus one might want to look for alternative formulations which rid the system of these features.

I illustrate how the system rules out right adjunction to XP, since this is of crucial importance here. Consider the tree in figure 1.8. The set of pairs $\langle n, m \rangle$

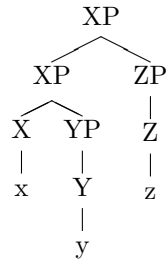


Figure 1.8: Right adjunction

such that n asymmetrically c-commands m in this tree, i.e. A , is given in (14).

$$(14) \quad \{\langle ZP, X \rangle, \langle ZP, YP \rangle, \langle ZP, Y \rangle, \langle X, Y \rangle\}$$

The lower segment of XP does not c-command Z, because it is not a category. YP does c-command X, but not asymmetrically. Finally, ZP asymmetrically c-commands X and YP, because it (vacuously) holds that every category dominating ZP dominates X, YP. Hence, $d(A)$ comes out as (15), which has it that z precedes x, y , rather than follow them.

$$(15) \quad \{\langle z, x \rangle, \langle z, y \rangle, \langle x, y \rangle\}$$

The most compelling evidence for the LCA comes from the fact that it shows promise to derive word-order universals of the following kind: Although there

are many languages that exhibit second position phenomena (e.g. Germanic verb-second, Slavic clitic-second, etc.), there are no known penultimate position phenomena. For instance, there are no languages where the finite verb ends up in the second to last position of the clause. Another asymmetry of this kind is the one noted by Greenberg (1966) (his universal 20), which states that

when any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite.

In other words, there is a (universal) gap in the ordering possibilities: we find the orders (16a)-(16c), but not the one in (16d).

- (16) a. $\langle dem, num, adj, N \rangle$
 b. $\langle N, dem, num, adj \rangle$
 c. $\langle N, adj, num, dem \rangle$
 d. * $\langle adj, num, dem, N \rangle$

The phenomenon at hand appears to be considerably more general than what Greenberg’s original formulation suggests. More specifically, it generalizes to crosslinguistic ordering patterns involving verbal clusters (Koopman and Szabolcsi, 2000; Cinque, 2000b), adverbial PPs (Cinque, 2002), as well as sentence adverbs (Nilsen and Vinokurova, 2000), and verbal affixes (Cinque, 1999). Hence, it seems that one could restate it as a generalization about stacking of modifiers, adequately defined.

Given the restriction that a moved category must c-command its trace, it follows from the absence of right-adjunction (i.e. the LCA) that all movement must be to the right. Hence there could be no derivation of the order $\langle adj, num, dem, N \rangle$, and Greenberg’s U20 would follow as a theorem of the system. Ackema and Neeleman (2002) give one illustration of a system which can derive the GU20, taken from Cinque’s work. We give another illustration of a toy system that derives a generalized U20. Suppose that we distinguish a category of “x-raisers”, (xr) the class of elements that participate in U20-like word-order patterns. When an xr has a complex complement, i.e. a complement which itself has a specifier, it can attract the specifier of its complement, the head of its complement or the entire complement to its own specifier. By iterating attraction of specifier, we end up moving the most deeply embedded specifier to the highest position (climbing). If the complement of an xr is simple, e.g. a simple NP, this NP must raise. This is illustrated in (17).

- (17) $[_{xr1P} \text{ xr1 } [_{NP} N]]$
 ► move NP
 $[_{xr1P} [_{NP} N] \text{ xr1 } t_{NP}]$
 ► merge xr2 and extract spec

$$[_{xr2P} [_{NP} N] \text{ } xr2 [_{xr1P} t_{NP} \text{ } xr1 \text{ } t_{NP}]]$$

► merge $xr3$ and extract spec

$$[_{xr3P} [_{NP} N] \text{ } xr3 [_{xr2P} t_{NP} \text{ } xr2 [_{xr1P} t_{NP} \text{ } xr1 \text{ } t_{NP}]]]$$

Suppose, instead, that we iterate the extract complement option. This leads to “roll-up” structures.

(18) $[_{xr1P} \text{ } xr1 [_{NP} N]]$

► move NP

$$[_{xr1P} [_{NP} N] \text{ } xr1 \text{ } t_{NP}]$$

► merge $xr2$ and extract complement:

$$[_{xr2P} [_{xr1P} [_{NP} N] \text{ } xr1 \text{ } t_{NP}] \text{ } xr2 \text{ } t_{xr1P}]$$

► merge $xr3$ and extract complement

$$[_{xr3P} [_{xr2P} [_{xr1P} [_{NP} N] \text{ } xr1 \text{ } t_{NP}] \text{ } xr2 \text{ } t_{xr1P}] \text{ } xr3 \text{ } t_{xr2P}]$$

Finally, if we consistently extract the head of the complement, we restore the order of merger, except for the two highest elements.

(19) $[_{xr1P} \text{ } xr1 [_{NP} N]]$

► move N

$$[_{xr1P} N \text{ } xr1 [_{NP} t_N]]$$

► merge $xr2$ and extract head

$$[_{xr2P} \text{ } xr1 \text{ } xr2 [_{xr1P} N \text{ } t_{xr1P} [_{NP} t_N]]]$$

► merge $xr3$ and extract head

$$[_{xr3P} \text{ } xr2 \text{ } xr3 [_{xr2P} \text{ } xr1 \text{ } t_{xr2P} [_{xr1P} N \text{ } t_{xr1P} [_{NP} t_N]]]]$$

A system like this one can derive the three observed orders in (16) and many intermediate ones, but not the ungrammatical (16d), that is, if we rename adj , num , dem as $xr1$, $xr2$ and $xr3$, respectively. Assume that it can. Then there must be a derivation conforming to the LCA which yields the order $xr1 \text{ } xr2 \text{ } xr3 \text{ } N$, i.e. where all material is left-adjoined. Since $xr1$ and N are separated, $xr1$ must have moved as a head, i.e. leaving N behind, thus yielding the intermediate structure (20).

(20) $[_{xr3P} \text{ } xr3 [_{xr2P} \text{ } xr1 \text{ } xr2 [_{xr1P} N \text{ } t_{xr1}]]]$

But from in this structure, there is no constituent containing $xr1$, $xr2$, but excluding N , which can be moved to $spec\text{-}xr3$, hence the order $xr1 \text{ } xr2 \text{ } xr3 \text{ } N$ cannot be derived. In other words, the unwanted order could only be derived if there is an element α intervening between $xr3$ and its complement which has the property that it (non-locally) attracts N . Then $xr3$ must (non-locally) attract $xr2P$, as shown in (21).

(21) $[_{\alpha P} \alpha [_{xr2P} \text{ } xr1 \text{ } xr2 [_{xr1P} N \text{ } t_{xr1}]]]$

► move N

$[\alpha_P N \alpha [\text{xr2P } \text{xr1 } \text{xr2 } [\text{xr1P } t_{\text{NP}} t_{\text{xr1}}]]]$
 ► merge xr3 and move xr2P

$[\text{xr3P } [\text{xr2P } \text{xr1 } \text{xr2 } [\text{xr1P } t_{\text{NP}} t_{\text{xr1}}]] \text{xr3 } [\alpha_P N \alpha t_{\text{xr2P}}]]$

In view of ordinary (long distance) displacement facts, we have to acknowledge the existence of non local attractors like α . Furthermore, if α is allowed to be phonetically empty, (21) will give the impression of being a well formed order of the relevant type. If we reserve the term xr for expressions that are not capable of long distance attraction, we have derived that, under the current assumptions, these should conform to the Generalized U20. This system could not derive the U20-behavior of adverbial PPs (Cinque, 2002), given that xrs are treated as heads. I refer the reader to Cinque’s work for a derivation of PPs. The point here is that the LCA gives room for several alternative systems which derives the GU20, i.e. this is a robust feature of LCA-compatible systems.

One point is worth emphasizing. It is often thought that the LCA renders left-branching structures impossible. This, however, is not correct. Left-branching structures can arise when constituents are successively embedded in specifiers, as shown in figure (1.9). This can arise, either by base generation or by (“roll-up”) movement.

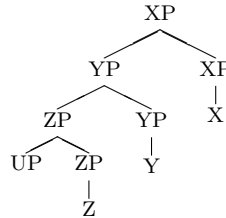


Figure 1.9: Left branching structure

Hence it does not follow from the LCA that if a constituent X follows another constituent Y , X must be “lower in the tree” than Y . What is excluded is right-adjunction and, hence, rightward movement.

1.4.3 A processing account

Ackema and Neeleman (2002) propose that one could exclude head movement to the right (in certain cases) but allow right adjunction, and that this would suffice to derive GU20 and the absence of penultimate position phenomena.¹⁸ They suggest that the absence of rightward head movement could be thought of as a processing phenomenon, because it would force the parser to keep more things in store than leftward movement. Their account rests on the following assumptions about the parser (Ackema and Neeleman, 2002, their (21)):

¹⁸Note that, given the discussion of figure (1.9) in the previous section, this approach may be much closer to the antisymmetric approach than what these authors seem to assume.

- (22)
- a. It scans the input string from the left to the right.
 - b. It constructs a tree, that is, a set of dominance and precedence relations.
 - c. It has no look-ahead.
 - d. It can only postulate a trace after having encountered an antecedent.
 - e. It cannot alter information (dominance and precedence relations) stored in short-term memory for a given parse.

To these assumptions, I would like to add the following one, which they tacitly assume (or something like it).

- (23) Immediate attachment: Incoming material must be integrated immediately into a structure during sentence processing.

Their assumptions (22a)-(22c) could hardly be taken as controversial ones. Unfortunately, this does not hold for their assumptions (22d), (22e) and (23). Let us defer discussion of (22d) for a moment. (22e) seems to say that whenever the parser must reanalyze some already analyzed material there is a processing cost, e.g. a garden path effect. But this is blatantly false, at least if it is to square with assumption (23). Mulders (2002) gives numerous examples where any existing parsing theory would seem to be forced to assume reanalysis, but they do not give rise to detectable garden path effects. Consider the following contrast from Japanese. (24a) gives rise to a garden path, whereas (24b) does not (Mulders, 2002, pp.131-133).

- (24)
- a. *¿ Hurugashi-ga Yumiko-o Ø Ø yobidasita kissaten-ni*
Hurugashi-NOM Yumiko-ACC pro pro summoned tea-room-LOC
nagai koto mata-seta.
long time wait-made
“Hurugashi made Yumiko wait for a long time at the tea room
to which he summoned her.”
 - b. *Yumiko-o Hurugashi-ga Ø Ø yobidasita kissaten-ni*
Yumiko-ACC Hurugashi-NOM pro pro summoned tea-room-LOC
nagai koto mata-seta.
long time wait-made
“Hurugashi made Yumiko wait for a long time at the tea room
to which he summoned her.”

Consider first the parse of (24a). At the stage where the parser has encountered *Hurugashi-ga Yumiko-o yobidasita* (‘Hurugashi yumiko summoned’), it analyzes this as a main clause. But the continuation is not compatible with this. The continuation forces both *Hurugashi* and *Yumiko* to be reanalyzed as arguments of a superordinate clause, and the predicate *yobidasita* to be

reanalyzed as a relative clause on the following locative NP *kissaten-ni* ‘tea-room-LOC’. This is compatible with A&N’s view that “information cannot be altered”, since the reanalysis does, in fact, lead to a garden path effect. However, (24b), which minimally differs from (24a) in that the object *Yumiko* is scrambled around the subject *Hurugashi*, mysteriously does not lead to a garden path effect. But the amount of reanalysis involved is the same. Hence it seems that reanalysis is possible sometimes, after all.

Mulders (2002) gives numerous other cases where reanalysis seems to be costless, not only in Japanese, of course, and she comes up with a more sophisticated restriction on reanalysis (her (T)ROLLC), ultimately relating it to standard syntactic locality constraints on movement.¹⁹ The ease with which (24b) is parsed, indicates that “short term memory” is not as severely limited as Ackema and Neeleman (with many others) suppose. I refer the reader to Mulders’ work to see how contrasts such as that in (24) can be handled. What is crucial here is that Mulder’s approach explicitly rejects assumption (22e).

As for Ackema and Neeleman’s assumption (23), Mulders (2002) has many arguments against that, too. In fact, she devotes an entire chapter (chapter 3) to arguing that all the arguments that have been put forth in the literature in favor of this assumption, and thus against a “ θ -driven” parser (“Maximize satisfaction of the θ -criterion at every stage of the parse” (Pritchett, 1992)), are either inconclusive or wrong. She furthermore shows rather convincingly that a θ -driven parser can handle problems that lead to quite puzzling problems under an “immediate attachment” approach, including the contrast in (24) above. Mulders discusses parsing of Japanese sentences with relative clauses in great detail, showing that whereas “immediate attachment” (of course wrongly) leads us to expect that Japanese speakers should enter into a garden path virtually all the time,²⁰ the θ -driven approach can actually to a large extent correctly predict when Japanese speakers actually do feel a garden path effect, and when they don’t. I refer the reader to Mulders’ work for discussion.

Finally, the assumption (22d), that traces can only be inserted after the antecedent have been encountered, seems to depend in crucial ways on the two assumptions we have just seen to be questionable. In particular, θ -driven parsing algorithms, like the one Mulders is promoting, assume that empty material is freely generated (of course subject to the restrictions of the grammar of the language in question) and do not lead to processing problems. In any case, it is clear that this assumption is rather essential for any algorithm that wants to derive absence of rightward movement from processing difficulties: With rightward movement, the trace precedes its antecedent by definition, and by assumption (22e), it cannot be inserted by the parser before the antecedent is encountered, leading to potential for processing problems because of the assumption that everything must be inserted into a tree at once. On a θ -driven

¹⁹Mulders’ theory is based on work by Pritchett (1992).

²⁰Of course, we do not want to say that Japanese sentence processing follows different principles than, say, English sentence processing.

approach, the crucial assumption is that there is no problem with keeping things in storage: In fact, this is the typical case in strongly “head final” languages like Japanese, where the θ assigner is typically encountered as the very last element.

The point here is that the assumptions that Ackema and Neeleman (2002) make about the parser are not uncontroversial, and, in some cases, likely to be false. Be that as it may, let us look at how their system derives the (G)U20 and the absence of penultimate position phenomena.

1.4.4 GU20 and Penultimate position

In this subsection, I will try to explain how Ackema and Neeleman attempt to derive universal word-order asymmetries from their assumptions about processing. It does not seem to me that their assumptions actually have the results that they claim them to have. In fact, I will argue that they fail to derive both the GU20 and the absence of penultimate position phenomena.

Suppose that the parser has identified an XP. Then, according to their assumptions, it will postulate the existence of some head H that follows XP (it has not yet encountered such a head) and that XP is immediately dominated by some projection H^i of H. They notate this as follows:

$$(25) \quad P(XP, H); ID(H^i, XP); Proj(H, H^i).$$

Here “P” is in short for “precedes”, “ID” for “immediately dominates”, and “Proj” for “is a projection of”. “H” is an “abstract head” which will be given content when an actual head is encountered. If the parser now encounters another phrase YP, we get the following (“D” means “dominates”).

$$(26) \quad \begin{array}{ll} \text{a.} & XP, YP \\ \text{b.} & P(XP, H); ID(H^i, XP); Proj(H, H^i) \\ & P(YP, H); ID(H^j, YP); D(H^i, H^j); Proj(H^j, H) \end{array}$$

So far, our parse corresponds to the tree in figure (1.10). When the parser

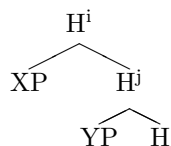


Figure 1.10: Tree for (26)

encounters a head, (ignoring the possibility that it has moved for the moment), we get the following addition to our parse, which, is a monotone increase in the represented information from (26), just as (26) was from (25).

$$(27) \quad \text{a.} \quad XP, YP, V$$

- b. $P(XP, H); ID(H^i, XP); Proj(H, H^i)$
 $P(YP, H); ID(H^j, YP); D(H^i, H^j); Proj(H^j, H)$
 $H=V$

If we now encounter more phrases, these could either be complements of V , or right-adjoined material. Consider the latter option.

- (28) a. XP, YP, V, ZP, WP
 b. $P(XP, H); ID(H^i, XP); Proj(H, H^i)$
 $P(YP, H); ID(H^j, YP); D(H^i, H^j); Proj(H^j, H)$
 $H=V$
 $P(H, ZP); ID(H^k, ZP); Proj(H, H^k)$
 $P(H, WP); ID(H^l, WP); D(H^l, H^k); Proj(H, H^l)$

This is still in accordance with our assumptions, and the parse is unambiguously has it that XP c-commands YP , and that WP c-commands ZP . It does not determine the scope between the preverbal and the postverbal phrases, and Ackema and Neeleman simply assume that this is determined either by the grammar or by the discourse context. Hence we see that the system can handle both left- and rightbranching structures. A&N argue that it cannot handle certain kinds of rightwards movement, however. Consider the tree in figure (1.11) and how it would be parsed incrementally. At the point where the two

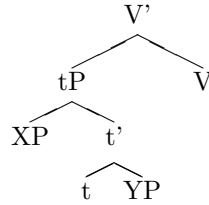


Figure 1.11: Rightward head movement

phrases have been encountered, we would have the following parse:

- (29) a. XP, YP
 b. $P(XP, H); ID(H^i, XP); Proj(H, H^i)$
 $P(YP, H); ID(H^j, YP); D(H^i, H^j); Proj(H^j, H)$

Suppose that the parser, upon encountering the verb tries to insert a trace between the two phrases.²¹ Given that (29) says that H^i dominates XP and H^j , and that H^j immediately dominates YP , the trace must be an H , immediately dominated by H^j . But this contradicts the information in (29) that YP precedes H . Thus insertion of such a trace contradicts their assumption (22e), that the

²¹It cannot hypothesize a trace before, because of assumption (22d).

parser cannot alter information it has already postulated. A&N show that rightward head-movement is, in fact, possible in this approach just in case it does not cross any “dependents” of the head. This is stated in their “Rightward Head Movement Theorem”:

- (30) RHMT (Ackema and Neeleman, 2002, their (38))
 Rightward head movement is possible as long as no dependent of the moving head is crossed.

This formulation arguably requires some attention to the notion “dependent”. In footnote 12, they explain that “dependent” refers to such elements as “specifiers, complements or adjuncts of the heads under discussion.” Thus as long as an element is not thought to be such a dependent, rightward head movement should be able to cross it. They then go through some cases which they admit can be given an antisymmetric account, but can also be analyzed as rightward head movement in accordance with RHMT. This now derives GU20, if it can be shown i) that, e.g. in noun phrases, the noun can only move around dem, num and adj as a head and ii) that these are “dependents” of the noun, i.e. these two auxiliary assumptions would rule out a structure like that in figure (1.12), where N has head-moved to the right. However, A&N do not argue for these two extra assumptions. They explicitly argue that phrasal movement

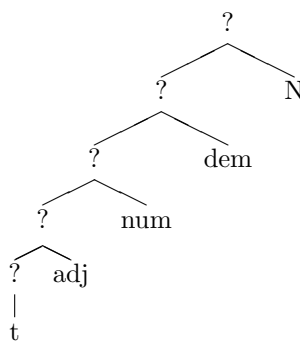


Figure 1.12: Impossible NP

is not subject to the rightward movement restriction, and propose to handle extraposition phenomena in this way, so if N could move as a phrase, the tree in figure (1.12) should be OK, depending on how one analyzes dem, num, adj. The crucial difference between head movement and phrasal movement on their account is that, given “immediate attachment”, the parser is forced to postulate “abstract heads” before the “actual head” is encountered. Postulation of “abstract heads”, in turn, leads to the potential of conflicting information (re-analysis) if a trace is to be inserted inside an already parsed string. Phrases are not postulated before they are encountered, so there’s no potential for conflict.

Thus, it appears that they do derive that head movement, in a technical sense, cannot cross “dependents”, but that phrasal movement can.²² The question is whether, granting, for the sake of the argument, A&N’s assumptions (22-23), this suffices to derive GU20. Apparently, it doesn’t. We will need the two auxiliary assumptions.

To see that this objection should be taken seriously, consider the following phenomenon from Norwegian. Certain “low” adverbs occur in the same order/scope when they follow the VP as when they precede it (Cinque, 1999; Nilsen, 2000). This happens when the last adverb is stressed.²³ Sentential negation (*ikke*) and “high” sentential adverbs do not participate in such orders (31e). If the order/scope of the adverbs is reversed (31f), the result is ungrammatical.

- (31) a. Jens hadde ikke lenger alltid helt **forstått**
 J had not any.longer always completely understood
 problemet,
 the-problem
- b. Jens hadde ikke lenger alltid **forstått** **problemet**
 J had not any.longer always understood the-problem
 HELT.
 completely
- c. Jens hadde ikke lenger **forstått** **problemet** alltid
 J had not any.longer understood the-problem always
 HELT.
 completely
- d. Jens hadde ikke **forstått** **problemet** lenger alltid
 J had not understood the-problem any.longer always
 HELT.
 completely
- e. *Jens hadde **forstått** **problemet** ikke lenger alltid
 J had understood the-problem not any.longer always
 HELT.
 completely
- f. *Jens hadde ikke **forstått** **problemet** helt alltid
 J had not understood the-problem completely always
 LENGER.
 any.longer

It seems hard to avoid the conclusion that, in Norwegian, the VP can move around (left-adjoined) low adverbs. Suppose, for concreteness, that the adverbs

²²It is currently an open question whether head movement in the standard sense is possible in any direction, See a.o. Chomsky (2001); Koopman and Szabolcsi (2000); Müller (2002); Starke (2001); Nilsen (to app.b).

²³If the adverbs are destressed, in a “comma reading”, both orders are possible.

are left-adjoined to vP, and that VP is allowed to move leftward and adjoin to vP. The relevant part of the structure for the examples in (31) would be

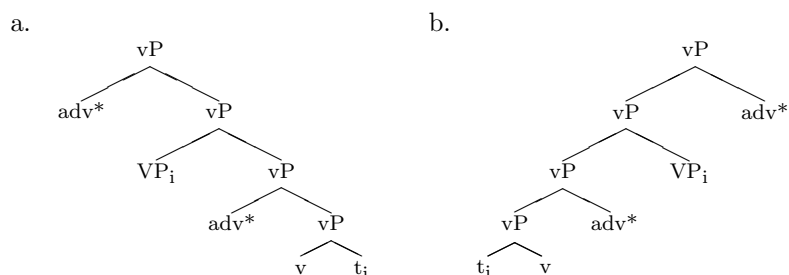


Figure 1.13: VP-scrambling

figure (1.13a). But then what in A&N’s system would prevent a symmetric language, where the adverbs are adjoined to the right, and the VP can move to the right as in (1.13b)? This would be a language where a sequence of adverbs with “inverse” (right to left) scope could precede the VP. Of course, no such language is attested.

In fact there is ample reason to think that the status of the expressions as heads is irrelevant to the proper formulation of GU20. We have just seen that adverb-VP ordering appears to conform to it. Cinque (2002) argues that PP-VP ordering conforms to it, too. Verbal cluster formation also conforms to it, and this has been argued not to involve head-movement, but rather phrasal movement (Koopman and Szabolcsi, 2000).

A&N show that their parser allows string vacuous head movement to the right. In fact, they argue that the finite verb in Japanese moves to C, just as in Dutch, but that in this language, C is final. But then what prevents the existence of a cousin of Japanese where the specifier of C is also final? This would be a verb-second-to-last language. A&N are aware of this problem, of course, and they try to explain this in terms of a Right-Roof effect. The rightward movement of the XP in spec-C would cross XP-barriers, and hence be ruled out by the Right-Roof constraint. They argue that this constraint can be made to follow from an assumption that phrases become “atomic” when the parser has finished parsing them, and one cannot postulate traces inside such atoms. For example, rightward movement of an object to spec-C would necessarily cross VP, so the parser would have to postulate a trace in VP after it has finished analyzing it, i.e. after it has become an atom. It is somewhat strange that the parser, which is otherwise fully aware of the grammar of the relevant language, would “close off” a VP which violates the θ -criterion. Furthermore, it seems rather mysterious that, on A&N’s account, rightwards head movement is *easier* than rightward phrasal movement when it comes to movement out of

IP, but *harder* than phrasal movement when it comes to “extraposition.” Suppose that this can be made to work, however. There could still be a cousin of Japanese where spec-CP is initial when, say, an object is moved there, but final otherwise, e.g. when an adverbial is base-generated there. This would result in a language which is verb-last when something is moved to spec-C, and verb-second-to-last when something is base-generated there. Yet another cousin of Japanese could reserve its final spec-C exclusively for base-generated material, but require it to be filled, sometimes by expletives. This would also be a strict verb-second-to-last language. Hence it is difficult to see that A&N derive the universal absence of penultimate position phenomena. By the LCA it follows, given that the LCA bans rightward movement in principle.

1.4.5 A&N’s arguments for right-adjunction

Ackema and Neeleman point out that there are arguments that circumstantial PPs are right adjoined when they follow the VP. Their arguments are essentially that a cluster of sentence final PPs occur in the mirror ordering of the sentence internal order (Koster, 1974; Barbiers, 1995). This would follow straightforwardly if the PPs can be adjoined to the right or to the left of the VP. The second argument is that the relative scope of the PPs conforms to the right-/left-adjunction structures. However, these facts are also compatible with an antisymmetric approach. What they seem to show is that the structure of a VP with three final PPs, e.g. (32a) must conform to the general layout of (32b) (at some point of the derivation). More in particular, the facts do not pose problems for an antisymmetric account unless it can be demonstrated beyond doubt that the labels rendered here as ‘?’ must be VP. Barbiers (1995) suggests that (32b) is obtained derivationally by successively moving VPs into spec-PP as illustrated in figure (1.14).²⁴

- (32) a. shot him with a gun in the park on Friday
 b. [? [? [? [VP shot him] [PP with a gun]] [PP in the park]] [PP on Friday]]

Part of the evidence Barbiers (1995) has for this kind of derivation comes from the distribution of focus particles like *pas* ‘just’. He shows that, Dutch focus particles must immediately precede the constituent they associate with, with the sole exception of “extraposed” PPs, which cannot be immediately preceded by *pas*. Consider the contrasts in (33):

- (33) a. Jan heeft **pas** in EEN stad gewerkt.
 J has just in one city worked

²⁴Barbiers shows that this is not a case of “sideways movement”, i.e. the moved VPs do end up c-commanding their traces with Kayne’s (or his own) definition of c-command. Kayne’s definition is the following: X c-commands Y iff X, Y are categories, X does not dominate Y, and Y does not dominate X, and every category that dominates X, dominates Y. Note that, with this definition, there would have to be empty heads between each specifier of VP.

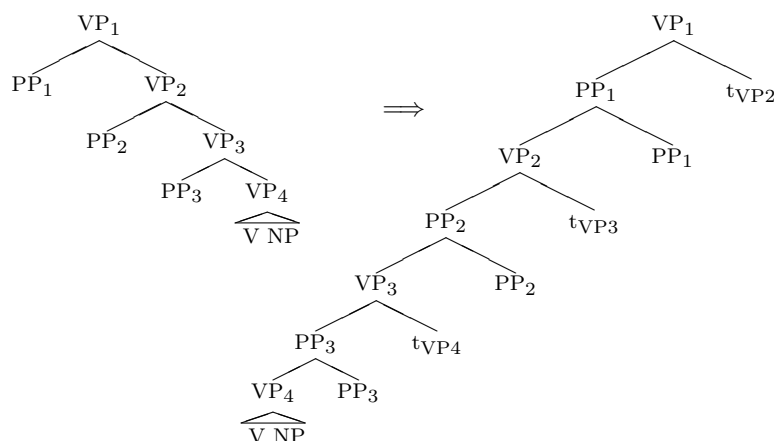


Figure 1.14: VP-intrapolation

- b. * Jan heeft gewerkt **pas** in EEN stad.
J has worked in just one city
- c. Jan heeft **pas** gewerkt in EEN stad.
J has just worked in one city
- d. Jan heeft in EEN stad gewerkt **pas**
J has in one city worked just
- e. **Pas** in EEN stad heeft Jan gewerkt.
just in one city has J worked

A free right/left-adjunction analysis would have to stipulate that (33b) is ungrammatical. Barbiers shows that it follows from his account of PPs and focus particles. Barbiers motivates the movements semantically: The VP moves to spec PP in order to establish appropriate semantic relations with it. I refer the reader to Barbiers' work to see how this is done in detail. In Nilsen (2000), it is argued that the left branching structure can be base-generated directly, still conforming to the LCA, if one treats the PPs as reduced relative clauses on the VP (event variable). According to this view, the structure comes out as something like figure (1.15), without any movements.

The idea according to this view, is that a temporal PP is a reduced relative clause on TP. It projects a TP, and takes a TP in its specifier. Similarly, a locative projects an AspP and takes an AspP in its specifier. The reason for

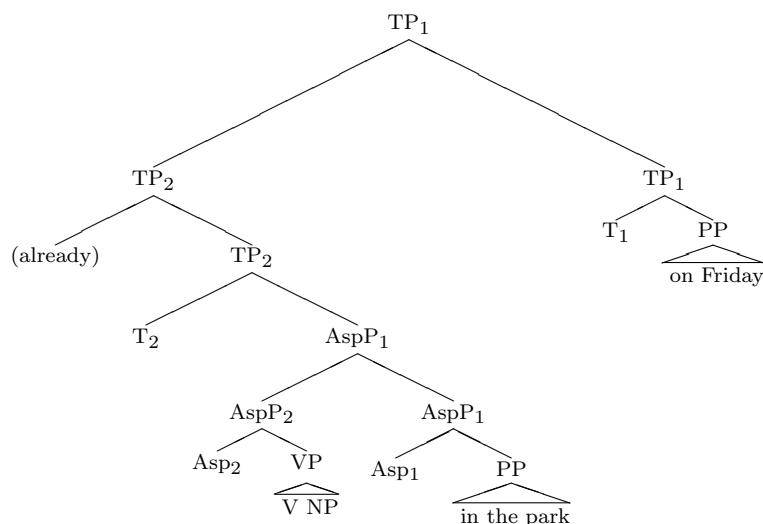


Figure 1.15: Relative clause structure for PPs

differentiating the labels is that temporals and locatives are ordered, i.e. temporals invariably occur higher in the structure than locatives. Some arguments favoring this kind of structure over the right-adjunction structure is that it interacts in interesting ways with the possibility of having sentence final sentential adverbs. Consider the following minimal pairs from Norwegian (capitals indicate prosodic stress):

- (34) a. ... at han har møtt Jens i parken ALLEREDE
 ... that he has met J in the-park already
- b. *... at han har møtt Jens på fredag ALLEREDE
 ... that he has met J on Friday already
- c. ... at han allerede har møtt Jens på fredag
 ... that he already has met J on Friday

The adverb *allerede* ‘already’ can follow the VP if it is modified by a locative PP, but not if it is modified by a temporal one. If we analyze *allerede* as a specifier of TP₂ in figure (1.15), this follows without further stipulation: There is no constituent XP containing the VP and the temporal PP which can shift around the adverb *allerede*. There is, however, a constituent (i.e. AspP₁ which can shift around *allerede*. In Nilsen (2000), several more adverb/PP pairs are shown to (mis-) behave in the same way. If the adverbs and PPs were freely right/left-adjoined to VP, it is hard to imagine a non-stipulative account for contrasts like that in (34).

In order to derive the mirror ordering effects (Koster, 1974) within this framework, one would assume that the predicate of the relative clause can be fronted instead of the relative head. This is in accordance with the derivation of head final relatives in e.g. Japanese proposed by Kayne (1994); Bianchi (1995).

Perhaps the worst problem for the free right/left-adjunction account for circumstantials is what has come to be known as “bracketing paradoxes”: the evidence sometimes favors a left-branching structure, and sometimes a right-branching one, with conflicting results, sometimes for one and the same sentence. These were first discussed by Pesetsky (1995), and recently given an antisymmetric derivation in Cinque (2002). As arguments for a left-branching structure, Cinque mentions i) absence of principle C effects when an argument precedes a coreferential R-expression in an adjunct (35a); ii) constituency diagnostics, e.g. the possibility of stranding the locative in (35b) and the impossibility of fronting the string *John in the park* (35c); iii) the relative scope of VP-final PPs, i.e. VP-final PPs take scope “towards the left”, and not “towards the right” as we would expect on a right-branching structure. For example, in (35d), the *because*-clause takes scope over the entire VP, including the locative PP.

- (35) a. They killed him_i on the very same day John_i was being released from prison.
 b. [Kill John]_i they did [t_i in the park].
 c. * [John in the park]_i they [killed t_i].
 d. John smoked in the car because of the rain.

As arguments favoring a right-branching structure for the PPs, Cinque (2002) mentions i) anaphor binding into an adjunct (36a-36b); ii) variable binding into an adjunct (36c-36d); and iii) negative polarity item licensing into an adjunct (36e-36f).

- (36) a. John spoke to Mary [about [these people]_i] [in [[each others]_i houses]].
 b. * John spoke to Mary [about [each other]_i] [in [[these peoples]_i houses]].
 c. Gideon Kremer performed [in every_i Baltic republic] [on its_i independence day].
 d. * He spent many hours [in its_i memorial] [on every_i independence day].
 e. John spoke to Mary [about **no** linguist] [in **any** conference room].
 f. * John spoke to Mary [about **any** linguist] [in **no** conference room].

This lead e.g. Pesetsky (1995) to suggest that multiple trees can be associated with the same string (simultaneously). Cinque (2002) argues that the evidence is compatible with an antisymmetric account which also incorporates

the word-order facts discussed by Barbiers (1995); Nilsen (2000) in addition to deriving the “koster-effects” (Koster, 1974). This approach takes Pesetsky’s different trees (or something like them) to be different derivational stages of a sequential derivation. All this goes to show that the arguments adduced by Ackema and Neeleman for left-branching structures are not arguments for right-adjunction, and that, once the full complexity of the facts is taken into account, the antisymmetric accounts seem to fare better with the data. The right-adjunction approach would need to be significantly enriched in order to handle the facts, and it is not clear that it would end up being any “simpler” or postulate fewer theoretical entities and operations in need of motivation than the antisymmetric accounts.

A&N argue that the extra heads needed in an antisymmetric account must be motivated independently. This is true as long as one takes them to be functional heads in the standard sense. Aside from answering this challenge directly, by actually trying to independently motivate lots of heads, which is Cinque’s approach, one could think of at least two ways of circumventing this problem: 1) abandoning the extra heads. By minimally adjusting Kayne’s definition of c-command as follows: X c-commands Y iff X, Y are categories and every *segment* dominating X , dominates Y . This leads to an asymmetric system with all the properties of Kayne’s, except that it allows multiple specifiers/adjuncts. 2) One could treat the heads as “dummies” one can insert for free, much as in Larson’s VP-shell analysis. So this is not a deep problem with antisymmetry.

Secondly, they argue that all the extra movements needed in an antisymmetric account must be triggered by (independently motivated) features. But the trigger-based view of movement is controversial, and not the only way of “triggering” movement, and other ways of doing this that have been or will be proposed could be extended to antisymmetric accounts as well. In short, this is not a problem with antisymmetry per se, but rather one with certain aspects of current syntactic theory in general. In fact, I think most syntacticians would agree that there is currently no general account of “triggers” for displacement phenomena. Furthermore, “directionality parameters” and other restrictions on “symmetry” needs motivation as well. What are the general properties (+N, +V, + θ , phonological phrasing etc) according to which directionality is parameterized? Why are there no well documented cases of final specifiers? It does not seem to follow from processing, for example.

1.5 Summary

I have argued that the main challenges for a Cinque-style approach to the distribution of functional material in the clause are transitivity failures and Bobaljik paradoxes. I have argued that an approach along the lines of Ernst (2001); Svenonius (2001) can handle these problems, but only at the cost of abandoning any hope of treating functional heads such as T as semantically contentful expressions. The root of the problem was shown to lie in the use

they make of the two orthogonal orderings of theoretical entities, viz. the FEO-calculus (s-selection) and FSEQ (c-selection). Hence, the problem we are faced with is this: If we want to make use of a single (linear) selectional sequence to account for the distribution of functional material, we face transitivity failures and Bobaljik paradoxes. If we make use of orthogonal sequences, we cannot account for interactions between the kinds of expressions ordered by the orthogonal sequences. I conclude that we should look for alternative accounts of the distribution of functional material in the clause, accounts that do not depend on the notion of a selectional sequence.

We have seen that, although there are problems with antisymmetry, it remains a viable and robust account for word-order asymmetries. At this stage, it seems to me that the account in terms of processing comes with too many problems for it to be taken as a real alternative. As noted in the beginning, there are some problems with the exact formulation of the LCA (complicated definition of c-command, reliance on a linear notion of “time”, etc.), and one would like to look for even more elegant formulations of it. It may well be that such a reformulation will ultimately rely on processing algorithms. So far it seems to me that we have to settle for the original formulation.

CHAPTER 2

Domains for Adverbs

2.1 Introduction

Why are sentential adverbs^{1,2} ordered? This question has become relatively central to linguistic theory since the work of Cinque (1999) and Alexiadou (1997). The present chapter represents an attempt to give it a partial answer. I will mainly concern myself with high (speaker oriented) adverbs, but some discussion of lower (temporal) adverbs and manner adverbs is included.

It is well known that some adverbs are polarity items. For instance, adverbs like *yet* and *any longer* are negative polarity items (NPI). Perhaps less widely acknowledged is the fact that some adverbs are positive polarity items (PPI). In van der Wouden (1997) it is argued that the Dutch adverbs *al* ‘already’ and *niet* ‘not’ are PPIs. In this chapter I will argue that surprisingly many sentential adverbs are PPIs. In particular, I will argue that ‘speaker oriented’ adverbs, like *allegedly*, *fortunately*, *possibly*, *evidently* should be treated as PPIs. I will use this fact to derive the relative ordering in which adverbs can cooccur. The observation is that some adverbs, in addition to being PPIs, themselves induce

¹This chapter will appear in virtually identical form as Nilsen (to app.a).

²Throughout, I use the notions ‘adverb’ and ‘adverbial’ interchangeably and in a purely descriptive sense. Alexiadou (2001) argues that the word class ‘adverb’ is questionable, Dowty (2000) questions the argument–adjunct distinction, and Julien (2000) argues that the notion of grammatical word is irrelevant for grammatical theory. Thus it is not obvious that any class of adverbs or adverbials can be adequately defined and distinguished from e.g. auxiliaries and verbal affixes. On this, see also Nilsen and Vinokurova (2000).

the environments that PPIs are excluded from. This results in a classification of adverbs according to which environments they can/cannot/must appear in, and which expressions can/cannot occur within their scope. It relates adverb ordering to Bellert's observation (Bellert, 1977) that certain adverbs are degraded in e.g. questions, imperatives and antecedents of conditionals.

Classifying adverbs in this fashion already gives us predictions for possible adverb sequences which differ from existing theories (Cinque, 1999; Ernst, 2001). However, I have not answered the initial question, i.e. why sentential adverbs are ordered, unless I can explain why a given adverb should be sensitive to a given property of its environment. In order to approach this question, I will adopt the framework developed in Chierchia (2001) for (negative) polarity items³ and attempt to extend it to positive polarity phenomena. I will focus primarily on the adverb *possibly* here, but I think the approach can be extended to other speaker oriented adverbs as well.

Quantification is quite generally subject to contextual restriction (Westerstahl, 1988). For example, if I say that *Stanley invited everybody*, I do not generally mean that he invited everybody in the whole world. The idea is essentially that negative polarity *any* is lexicalized with a (universally closed) variable ranging over *domain expanding functions*. The application of such domain expansion is governed by a *strengthening condition*: the result of widening the quantificational domain for a quantifier in a proposition must be stronger than (entail) the corresponding proposition without widening.

This derives the major distributional patterns for NPIs in a rather plausible way. I suggest that this approach can be straightforwardly extended to positive polarity items if we assume that PPIs are associated with *domain shrinkage* rather than domain expansion. This approach is then argued to explain the distribution of some (though not all) sentential adverbs.

Before proceeding, I would like to point out one complication, also noted in Cinque (1999, to app.). It is often stated that Cinque's hierarchy (Cinque, 1999) should be given a semantic explanation.⁴ The complication is this: there are adverb–adjective pairs which are apparently synonymous, but, nevertheless, differ in distribution. For instance, the adverb *probably* cannot occur under *never*, whereas the (proposition embedding) adjective *probable* can. Compare (37a) to (37b):

- (37) a. *Stanley never probably ate his wheaties.
 b. It was never probable that Stanley ate his wheaties.

In order to make plausible a *semantic* explanation for the oddness of (37a), it seems one has to be able to establish that *probably* is not synonymous with

³Chierchia bases his theory on that of Kadmon and Landman (1993), and it is closely related to proposals by Krifka (1995) and Lahiri (1997). These proposals are discussed in section 3.2.

⁴Such semantic derivation would not automatically remove the motivation for the hierarchy. Its motivation is that certain *syntactic* phenomena, such as verb placement in Romance, interact with adverb ordering in intriguing ways. I return briefly to this point in section 5.

probable in certain respects, otherwise (37b) should be odd as well. What is more, one has to show that whatever semantic difference there is between the two, that difference must be the culprit of the contrast in (37).⁵ I try to establish such a difference between the pair *possibly*, *possible* below.

The chapter is organized as follows. In section 2, I present the main data, that speaker oriented adverbs are excluded from the same type of environment that license NPIs. In section 3, I present the theories of polarity items that I will employ to derive this behavior and discuss some of their prediction with respect to adverb ordering. Section 4 develops a semantics for the adverb *possibly*, deriving its semantic differences with the adjective *possible* and, at the same time, deriving its distribution. In section 5, the final section, I discuss how the present account bears on current theories of adverb ordering phenomena.

2.2 Adverbs and Polarity

2.2.1 NPIs and speaker oriented adverbs

In Bellert (1977) it was observed that certain sentential adverbs have a narrower distribution than one might expect. For example, she observes that *speaker oriented adverbs* (henceforth SOA), such as evaluatives (*fortunately*), evidentials (*evidently*) and some modals (*possibly*) are degraded in questions (39a). As it turns out, these types of adverbs are also degraded in antecedents of conditionals (39b), imperatives (39c) under negation (39d), under clause-embedding predicates like *hope* (39e), as well as within the scope of monotone decreasing subject quantifiers like *no N* (39f). I use Norwegian examples to demonstrate this. As far as I have been able to determine, English works the same way, i.e. the English translations are also degraded. In (39), ‘ADV’ represents any of the adverbs in (38):⁶

⁵I take this to be a problem for Ernst’s (2002) semantic approach to adverb distribution. This author takes adverb–adjective pairs like the ones under discussion to be, in fact, systematically synonymous.

⁶*Kanskje* ‘maybe’ is better in questions than the other ones. However, when *maybe* occurs in questions, as in (i), the question is not whether or not it is *possible* that S ate his wheaties, rather, it seems to be

equivalent to the same question with *maybe* removed.

(i) Did Stanley (maybe) eat his wheaties (, maybe)?

I have no suggestions as to why this should hold here (but see van Rooy (2002) on polarity items and questions). One can also find occurrences of *probably* in antecedents of conditionals which are not that bad.

(ii) If Le Pen will probably win, Jospin must be disappointed.

I take the slipperiness of some these intuitions to be comparable to that found with relative adverb ordering. Consequently, I will try to stick to phenomena for which the intuitions are sharper.

- (38) **heldigvis, tydeligvis, paradoksalt nok, rlig talt, muligens,**
 fortunately evidently paradoxically honestly possibly
kanskje, sannsynligvis, angivelig, neppe
 maybe probably allegedly hardly

(39) [Norwegian]

- a. Spiste Ståle (*ADV) hvetekakene?
 ate S (*ADV) the-wheaties
 “Did Stanley (*ADV) eat the wheaties?”
- b. Hvis Ståle (*ADV) spiste hvetekakene, ...
 if S (*ADV) ate the-wheaties
- c. (*ADV) Spis (*ADV) hvetekakene!
 (*ADV) eat (*ADV) the-wheaties
- d. Ståle spiste (ADV) ikke (*ADV) hvetekakene.
 S ate (ADV) not (*ADV) the-wheaties
 “Stanley (ADV) didn’t (*ADV) eat the wheaties.”
- e. Jeg håper Ståle (*ADV) spiste hvetekakene.
 I hope S (*ADV) ate the-wheaties
- f. Ingen studenter (*ADV) spiste hvetekakene.
 no students (*ADV) ate the-wheaties

The same adverbs can appear in degree clauses (40a) and under clause embedders like *think* (40b) (compare (39e) with the verb *hope*), as well as, of course, ordinary declaratives (40c).

- (40)
 - a. Ståle var så sulten at han (ADV) spiste hvetekaker.
 S was so hungry that he (ADV) ate wheaties
 - b. Jeg tror Ståle (ADV) spiste hvetekakene.
 I think S (ADV) ate the-wheaties
 - c. Ståle spiste (ADV) hvetekakene.
 S ate (ADV) wheaties

It seems that SOA are excluded from the environments that license negative polarity items (NPI) like English *any*, Greek *tipota* or Dutch *ook maar iets*. Thus, SOA appear to be positive polarity items (PPI). I give Dutch examples in (41) and Greek ones in (42).⁷

(41) [Dutch]

- a. Heeft Jan **ook maar iets** gegeten?
 has J anything eaten

⁷Some Dutch speakers find (5-c) and (5-e) degraded. What is crucial here is that some speakers accept them, so these environments are relevant for NPI licensing.

- b. Als je **ook maar iets** hebt gegeten, ...
if you anything has eaten
- c. Eet **ook maar iets**!
eat anything
“Eat (something)!”
- d. Er heeft geen enkele student **ook maar iets** gegeten.
there has no single student anything eaten
- e. Ik hoop dat Jan **ook maar iets** heeft gegeten.
I hope that J anything has eaten
- f. *Ik denk dat Jan **ook maar iets** heeft gegeten.
I think that J anything has eaten
- (42) [Greek]
- a. Efage o Yiannis **tipota**?
ate J anything
- b. An o Yiannis efage **tipota**, ...
if J ate anything
- c. Fae **tipota**!
eat anything
- d. Elpizo na efage o Yiannis **tipota**.
hope-1sg that ate J anything
- e. *Nomizo pos o Yiannis efage **tipota**.
think-1sg that J ate anything

The emerging generalization is the following: whenever an NPI like *ook maar iets* or *tipota* is licensed, speaker oriented adverbs are degraded. In other words, SOA are PPIs.

2.3 Approaches to polarity items

2.3.1 Monotonicity and veridicality

NPIs like *ook maar iets* and *tipota* are usually classified as ‘weak’ NPIs. These are licensed in all downwards entailing (DE) contexts, as well as certain other contexts, including epistemic *might*. NPIs that are only licensed in a subset of DE environments (viz. antiadditive environments) are called ‘strong’ NPIs. An function f is antiadditive iff

$$f(a \vee b) = f(a) \wedge f(b).$$

It is DE iff it is order reversing, i.e. whenever $a \sqsubseteq b$, it holds that $f(b) \sqsubseteq f(a)$, where “ \sqsubseteq ” denotes semantic strength. It is *Antimorphic* (AMo) iff it is both

AA, and *antimultiplicative*. A function f is antimultiplicative iff

$$f(a \wedge b) = f(a) \vee f(b).$$

Examples of AA operators include *nobody*, *not*, *never*. Classical negation is antimorphic, i.e. it is both AA and antimultiplicative.⁸ DE operators that are not AA include *less than half of the N* and *rarely*.

Giannakidou (1997); Zwarts (1995) argue that there is a well defined class of semantic environments, *non-veridical* (NV), in which weak NPIs are licensed. Informally, a (propositional) operator Op is *veridical* if $\text{Op}(\varphi)$ entails φ for any φ ; *non-veridical* if $\text{Op}(\varphi)$ does not entail φ , and *anti-veridical* if $\text{Op}(\varphi)$ entails $\neg\varphi$. I give a more formal definition below. In general, we have that all antiadditive environments are DE, and that all DE environments are NV, although the converse is not generally true (cf. van der Wouden (1997); Zwarts (1998); Bernardi (2002) for discussion and formal proofs):

$$AMo \subseteq AA \subseteq DE \subseteq NV$$

This is important, because it shows us that, whenever some expression is (anti-)licensed in NV environments, it is (anti-)licensed in DE, AA and AMo environments as well. Thus, an analysis in terms of NV-ness can incorporate results from the analyses in terms of DE-ness.

Formally, veridicality can be defined as follows (Bernardi, 2002):⁹

Definition 2.1 *Let f be a boolean function with a boolean argument,*

1. $f \in D_t^{Dt}$. f is said to be

- (a) veridical iff $\llbracket f(x) \rrbracket = 1 \models \llbracket x \rrbracket = 1$
- (b) non-veridical iff $\llbracket f(x) \rrbracket = 1 \not\models \llbracket x \rrbracket = 1$
- (c) anti-veridical iff $\llbracket f(x) \rrbracket = 1 \models \llbracket x \rrbracket = 0$

2. $f \in D_t^{D\langle a, t \rangle}$ (i.e. f is a generalized quantifier). f is said to be

- (a) veridical iff $\llbracket f(x) \rrbracket = 1 \models \exists y \in D_a(\llbracket x(y) \rrbracket = 1)$
- (b) non-veridical iff $\llbracket f(x) \rrbracket = 1 \not\models \exists y \in D_a(\llbracket x(y) \rrbracket = 1)$
- (c) anti-veridical iff $\llbracket f(x) \rrbracket = 1 \models \nexists y \in D_a(\llbracket x(y) \rrbracket = 1)$

Some examples are in order. *Possibly* is non-veridical because *possibly*(φ) does not entail (“ \models ”) φ . *Obviously* is veridical, because *obviously*(φ) $\models \varphi$. *I hope that* $_$ is non-veridical, because *I hope that* $\varphi \not\models \varphi$. The same appears to hold

⁸Antiadditivity and antimultiplicativity combine to yield the de Morgan Laws for classical negation. See van der Wouden (1997) for discussion of these properties and their relevance for different classes of polarity items.

⁹Bernardi (2002) also generalizes the definition to n -ary boolean functions. I omit this here, as it will not be relevant to our discussion.

for *I think that* φ , but Giannakidou argues that whenever it is true of John that he thinks that φ , then φ must be true in his epistemic model. This does not hold for *hope*. In this sense, then, *think* can be said to be *weakly veridical*. Weak veridicality is thus veridicality relativized to epistemic states.

Giannakidou's motivation for using the notion is, of course, that some NPIs are apparently licensed outside of DE contexts. One example is adversative predicates (43). These are argued by Linebarger (1987) not to be DE.

- (43) a. I'm surprised he has any potatoes.
 b. I'm glad he has any potatoes.

However, Kadmon and Landman (1993) argue rather convincingly that these environments are, in fact, DE, once a certain contextual perspective is fixed and we limit ourselves to cases where their factive presupposition is satisfied. These authors argue along similar lines for antecedents of conditionals, which have also been argued not to be DE (Linebarger (1987)).

Another case for NVness are questions which cannot straightforwardly be said to be DE. In van Rooy (2002) it is shown that questions are, in fact, DE in their subject position under a Groenendijk and Stokhof (1984) style semantics. He gives an analysis of polarity items in questions extending and refining the results of Kadmon and Landman (1993); Krifka (1995), without appealing directly to NVness or DEness, however. It appears that the remaining cases for NVness are existential modals (e.g. *might*, cf. (44a)) imperatives (44b) and generics (44c). Generics was actually given a treatment in Kadmon and Landman (1993), so one might suspect that the other two can be dealt with as well.

- (44) a. John might have bought anything.
 b. Bring anything!
 c. John buys anything he finds.

There are also some problems for the NV-based approach. One obvious one is that languages differ with respect to which NV operators they allow to license weak NPIs. For instance, in Greek, *isos* 'possibly' licenses *tipota*, but the corresponding adverbs do not license Dutch *ook maar iets* or English *any*.

- (45) a. O Yiannis isos efage **tipota**.
 J possibly ate anything
 [Greek]
 b. * John possibly ate anything.
 c. * Perhaps John ate anything.

Similarly, we have seen that the Dutch and Greek verbs for *hope* license *ook maar iets* and *tipota*. According to our informants, English *hope* does not license *any*, as seen in (46).

- (46) * I hope John has any potatoes.

(47) a. ?Less than three students have eaten anything.
b. ?Minder dan drie studenten hebben ook maar iets gegeten.
less than three students have anything eaten

Speaker orientation, NV and DE

(48) SOA are excluded from X environments.
where X ranges over AMo, AA, DE and NV.

(49) [mood_{speech-act} frankly [mood_{evaluative} fortunately [mood_{evidential} allegedly [mod_{epistemic} probably [_{T_{past}} once [_{T_{future}} then [mod_{irrealis} perhaps [mod_{necessity} necessarily [mod_{possibility} possibly [asp_{habitual} usually [asp_{repetitive} again [asp_{freq(I)} often [mod_{vollitional} intentionally [asp_{cclerative(I)} quickly [_{T_{anterior}} already [asp_{terminaitive} no longer [asp_{continuative} still [asp_{perfect(?)} always [asp_{retrospective} just [asp_{proximative} soon [asp_{durative} briefly [asp_{generic/progressive} characteristically(?) [asp_{prospective} almost [asp_{sg.completive(I)} completely [asp_{pl.completive} tutto [voice well [asp_{cclerative(II)} fast/early [asp_{repetitive(II)} again [asp_{freq(II)} often [asp_{sg.completive(II)} completely]]]]]]]]]]]]]]]]]]]]]]

(50) a. allegedly, probably, perhaps, possibly, usually, no longer
b. hardly, never, rarely, not

Let us begin with some relatively clear contrasts. We expect, for example, that there should be a contrast between *often* and *rarely* with respect to their ability to outscope SOA. (51a) is an example found on the internet. It seems to contrast minimally with (51b) which should be a near semantic equivalent.¹⁰ We expect to find a similar contrast between *always* and *never*. (52a), also from the internet, is taken from an advertisement for an internet game. Again there is a fairly sharp contrast with (52b), which is what we expect.¹¹

- (51) a. His retaliations killed or endangered innocents and **often possibly** had little effect in locating terrorists.
 b. ?? His retaliations killed or endangered innocents and **rarely possibly** had an effect in locating terrorists.
- (52) a. This is a fun, free game where you're **always possibly** a click away from winning \$1000!
 b. ?? This is a fun, free game where you're **never possibly** further than a click away from winning \$1000!

In (52a), one could in principle argue either that *possibly* directly modifies the noun phrase *one click away...*, or that *always* directly modifies *possibly* or both. This is rendered implausible by the following facts. The translation of (52a) into Norwegian is equally grammatical (53a), and in (53a), *muligens* 'possibly' is separated from the noun phrase by the copula. Furthermore, the subject of the clause can intervene between the two adverbs (53b). For (53b), in turn, one could object that *alltid* 'always' could directly modify the subject, but this is refuted by the ungrammaticality of (53c), where this supposed constituent is moved to the V2-initial position. One could be tempted to say that the position of the subject in (53b) is due to some PF-reordering mechanism. But this cannot be the case (at least on available ideas of what PF-reordering could do), since the subject takes surface scope with respect to the adverbs: the sentence means that it is always the case that, for at least one player, it is possible that that player is one click away from the prize. Finally, there is a contrast, also in Norwegian between (53a) and (53d), where *aldri* 'never' is substituted for *alltid* 'always', along with some other changes to make the example more pragmatically plausible.

- (53) a. Dette er et morsomt, gratis spill hvor spillerne **alltid**
 this is a fun free game where the-players always
muligens er et klikk fra å vinne \$1000!
 possibly are one click from to win \$1000

¹⁰I did find some examples of the string *rarely possibly* but these seem to involve misspellings. Here is one example:

(i) The receiver requires a line of sight to the satellites that is relatively unobstructed by foliage and buildings, and this is rarely possibly on such a compact campus.

¹¹I have deliberately tried to make the (b)-examples roughly synonymous with the (a) examples. This is to prevent a contrastive interpretation of the (b) examples: Contrastivity is known to suppress (positive) polarity effects (on this, see a.o. Szabolcsi (2002)).

- b. Dette er et morsomt, gratis spill hvor **alltid** [en av
this is a fun free game where always one of
spillerne] **muligens** er et klikk fra å vinne \$1000!
the-players possibly is one click from to win \$1000
- c. * **Alltid en av spillerne** er muligens et klikk fra å
always one of the-players is possibly one click from to
vinne \$1000!
win \$1000
- d. ?? Dette er et morsomt, gratis spill hvor spillerne **aldri**
this is a fun free game where the-players never
muligens er lenger enn et klikk fra å vinne \$1000!
possibly are further than one click from to win \$1000

Another contrast of the same kind holds between the pair *already* and *not yet*¹². Consider (54a) which contrasts with (55a).

(54) a. In Wall Street, Enron was **already allegedly** going bankrupt.

b. In Wall Street, Enron was **allegedly already** going bankrupt.

(55) a. ?? (On Wall Street,) Enron was **not yet allegedly** going bankrupt.

b. On Wall Street, Enron was **allegedly not yet** going bankrupt.

(56a), from the internet, does not contrast sharply with (56b); my informants find the latter example quite acceptable. This requires an explanation.

(56) a. Although Beckham's absence will be felt England are **still probably** the best national team in Europe.

b. (Beckham's absence will be felt and) England are **no longer probably** the best national team in Europe.

Consider what is the semantic contribution of an adverb like *no longer*. In general, *no-longer(p)* is truth-conditionally equivalent to $\neg p$, with the added presupposition that *p* was true at some previous time.¹³ This presupposition makes it virtually impossible to exclude a contrastive reading of the modifiee. Contrastivity, as we have seen, quite generally interferes with polarity effects (Szabolcsi, 2002). The English word *somewhat* is a PPI, so (57a) is odd unless it is read with heavy stress on *doesn't*, thus contrasting it with the corresponding positive assertion. Finally, also *somewhat* can appear under *no longer*, again

¹²See Löbner (1999) for arguments that *not yet* is the negation of *already*, and that *still* is the dual of *already*, whereas *no longer* is the negation of *still*. *Still* and *already* are both PPIs, but they are allowed in a superset of the environments allowing for SOA. I return to this point shortly.

¹³*no longer, still, already, not yet* also come with aspectual requirements pertaining to the modified predicate (Löbner, 1999).

probably due to the contrastivity inherent to the presuppositional content of the adverb. In this way, the fact that (56b) is acceptable is actually expected on the view that *probably* is a PPI. (55a) also improves considerably if interpreted as an emphatic denial of (54a).

- (57) a. ?? Stanley doesn't like it somewhat.
 b. Stanley no longer likes it somewhat.

If examples like (51a, 52a, 54a, 56a) are not very frequent, there are probably pragmatic reasons for this. In any case, we are not trying to account for the statistical distribution of expressions, and I think it is fair to say that these examples are grammatical.

Two SOA in the same sentence rarely give good results. Consider, for instance (58).

- (58) a. ?? Maybe Stanley probably ate his wheaties.
 b. ?? Probably, Stanley possibly ate his wheaties.

We might take this to indicate that these adverbs are excluded from NV environments, rather than merely DE environments. However, I do not think the oddness of these sentences is due to the status of the adverbs as PPIs. Rather, I think it is due to the fact that they are all epistemic adverbs of the same sort, and one cannot epistemically modify the same sentence twice, if the modifiers are of the same epistemic kind. One reason to think that this is true is the following. *Necessarily* is also epistemic, but it certainly isn't a PPI as it occurs felicitously under negation (59a). Nevertheless, it is bad under the scope of *possibly* (59b).

- (59) a. Stanley didn't necessarily eat his wheaties.
 b. ?? Possibly, Stanley necessarily ate his wheaties.

Another reason not to say that SOA are excluded from NV environments is that *allegedly* which is NV, seems to be able to outscope *probably* and *possibly* (60).

- (60) **Allegedly**, Enron was **probably/possibly** going bankrupt.

Thus it does not seem to be the case that *probably/possibly* is excluded from NV environments. Speech-act adverbs like (*briefly*) might seem to be excluded from all NV environments, since these are degraded under NV epistemic adverbs. For example, (61b) does not have a reading according to which "it is possible that I am brief in saying that S ate his wheaties."

- (61) a. **Briefly**, Stanley **possibly** ate the wheaties.
 b. ?? **Possibly**, Stanley **briefly** ate the wheaties.

This might also be treated as a syntactic binding effect. Thus, *briefly* is essentially a manner adverb containing a variable which can be syntactically bound. If it occurs “too high” in the clause for the subject to bind it, and thus to get a “subject” oriented or standard “manner” reading, it becomes speaker oriented by default.¹⁴ This is supported by the fact that even *frankly* becomes subject-oriented in certain cases. (62a) can only mean that Stanley is frank, not the speaker. (62b) is apparently ambiguous between the two readings, while (62c) seems to prefer the speaker oriented reading.

- (62) a. Maybe Stanley frankly doesn’t like fish cakes.
 b. Stanley frankly doesn’t like fish cakes.
 c. Frankly, Stanley doesn’t like fish cakes.

Cinque (to app.) points out that (63a) is degraded, while if one of the adverbs is realized as an adjective (63b), the example becomes good. Cinque does not pose this as a problem for the present account, but, since *surely* is not NV (or DE), it might be thought to be one. However, the two adverbs *surely* and *probably* do not seem to be able to cooccur in any order, as seen by comparing (63a) to (63c), so the relative order of the adverbs cannot be the source of the oddness of these examples.

- (63) a. ?? Stanley **surely probably** ate his wheaties.
 b. It is **surely probable** that Stanley ate his wheaties.
 c. ?? Stanley **probably surely** ate his wheaties.

This behavior might be related to the observation made in Cinque (1999) that cooccurrence of two adverbs ending in *-ly* is quite regularly degraded.¹⁵ I do not have an account for this phenomenon, but I would like to point out that it is equally problematic for all existing accounts of adverb ordering I am aware of. Examples like (63a) seem to improve if the two adverbs are not adjacent. In this case the relative ordering *does* seem to matter, since (64b) is significantly worse than (64a). But in this case, the adjectival version is also degraded, i.e. (64c-64d) are also odd, thus contrasting with (63b).

- (64) a. ? **Surely**, Stanley **probably** ate his wheaties.
 b. ?? **Probably**, Stanley **surely** ate his wheaties.
 c. ?? It is **probably sure** that Stanley ate his wheaties.
 d. ?? I am **probably sure** that Stanley ate his wheaties.
 e. John is **probably sure** that Stanley ate his wheaties.

What seems to go wrong with the examples where *probably* outscopes *surely/sure* is that it makes little or no sense for the speaker to assert that s/he finds it likely

¹⁴For a very similar point of view concerning the manner/subject-oriented distinction, see Ernst (2000, 2001).

¹⁵Similar remarks hold of German adverbs ending in *-weise*, Mainland Scandinavian ones ending in *-vis*, or Italian ones ending in *-mente*.

that s/he is sure that Stanley ate his wheaties. In other words, (64e) is good because the adverb *probably* refers to the speakers epistemic state, whereas *sure* refers to *John's* epistemic state.

In sum, it seems that SOA are excluded from DE environments, but generally allowed in NV environments.

A note on ‘phase quantifiers’ and frequency adverbs

We have noted that *yet* (in its temporal use) and *any longer* are NPIs. This obviously limits their distribution. It has also been argued that *already* and *still* are PPIs. But they do not appear to have the same distribution of adverbs like *possibly*, so I will outline my answer to why this is so here. The answer is that the (positive) phase quantifiers¹⁶ are *not* excluded from DE environments, but merely from antiadditive (AA) environments. The latter, as we have seen, forms a subset of the former. *Few students* and *no students* are DE, but only the latter quantifier is AA. The definition of AA is repeated here.

$$f(a \vee b) = f(a) \wedge f(b).$$

Applying this to the quantifiers in question, we see that (65a) is equivalent to (65b), while (66a) is not equivalent to (66b).¹⁷

- (65) a. No students jumped or danced. \iff
 b. No students jumped and no students danced.

This equivalence goes through, because, if the set of students who danced is empty, and the set of students who jumped is empty, then the set of students who did one or the other must also be empty and vice versa.

- (66) a. Few students jumped or danced \niff
 b. Few students jumped and few students danced.

This equivalence does not go through, because the fact that the set of students who jumped has low cardinality, and the set of students who danced has low cardinality does *not* entail that the union (disjunction) of these two sets has low cardinality. To see this, suppose that “few” is contextually resolved to mean “less than n ”. Let S denote the set of students, D the set of dancers, and J the set of jumpers. For (66b) to entail (66a), we would have to have that

$$|S \cap J| < n \quad \wedge \quad |S \cap D| < n$$

entails that

$$|S \cap (J \cup D)| < n$$

¹⁶This term was introduced (as far as I know) by Löbner.

¹⁷In the a-examples, only the readings where the quantifier outscopes the disjunction are relevant. For some languages this is not possible, because in these languages, the disjunction is itself a PPI (Szabolcsi, 2001). As Szabolcsi points out, this does not appear to hold for English *or*, however.

which it clearly doesn't. Now consider how *already/still* behave with respect to these quantifiers.

- (67) a. Few students are still/already here.
 b. ?? No students are still/already here.
 c. ?? The students aren't still/already here.

The adverbs are apparently degraded under AA (hence under AMo) operators, but good under merely DE ones. This parallels the fact that *still/already* are substantially better within the scope of *rarely* (which is DE) than they are within the scope of *never* (which is AA).

- (68) a. At 9 AM/PM, Stanley is rarely still/already tired.
 b. At 9 AM/PM, Stanley is never still/already tired.

Thus, phase quantifiers like *already/still* appear to be excluded from AA environments. This allows them more freedom than SOA, because, as we have seen $AA \subseteq DE$, and SOA are excluded from all DE environments. In particular, (53-54) demonstrates that *already/still* enjoy considerable freedom.

We have seen that upwards entailing (UE) frequency adverbs, like *always*¹⁸ and *often* have a very free distribution. *Always* is degraded when it outscopes negation. As discussed in Beghelli and Stowell (1997), this is a quite general property of universal quantifiers. Thus, (69a) is an odd sentence unless it is read with heavy stress on *everybody*, in which case the universal takes scope *below* the negation, not the other way around. No such oddness arises with quantifiers like *many people* (69b). An entirely parallel contrast obtains between (69c-69d). (69c) and

- (69) a. ?? Everybody didn't snore.
 b. Many people didn't snore.
 c. ?? John always didn't snore
 d. John often didn't snore.

Apart from this, it seems that frequency adverbs have a very free distribution.

If we have 40 adverbs, there are $40 \times 40 = 1600$ ordered pairs that we would have to consider in order to exhaust their ordering possibilities, that is, if we limit ourselves to pairs of adverbs. If we do not limit ourselves to pairs, but exclude repetitions of the same adverb, the number is $40!$ ($= 1 \times 2 \times \dots \times 40 \approx 10^{48}$), a truly astronomical number. Hence, I am not going to test all possible orderings here. I hope to have made plausible the idea that limitations on adverb distribution can be treated, to a large extent, as a polarity phenomenon.

¹⁸ *Always*, being a universal quantifier, is DE in its restriction and UE in its scope. See Beaver and Clark (2002) for arguments that the restriction of *always* is given by the context, and not by the background, or unfocussed part of the clause. In other words, *always* is not focus-sensitive. Thus a clause (constituent) modified by *always* will be in a UE environment.

2.3.2 Why DE?

Summing up the findings in this section, we can refine our generalization about SOA as in (70). Why should this generalization hold? This question becomes rather acute when we combine it with the observation made in the introduction, that the adjectival counterparts of these adverbs are *not* excluded from DE environments. Compare (71a) to (71b):

- (70) SOA are excluded from DE environments.
- (71) a. *Jospin didn't possibly win.
b. It is not possible that Jospin won.

If we want to explain adverb distribution from the lexical semantics of the different adverbs, we now have to look for some independently motivated semantic difference between *possible* and *possibly* which we can take to be the source of the contrast in (71). We will now see that there is a difference between these two expressions. In the rest of the chapter, I will try to establish that this difference is indeed the culprit. Consider (72).

- (72) a. It's possible that Le Pen will win...
b. # Le Pen will possibly win...
c. # Perhaps Le Pen will win...
d. ... even though he certainly won't.

(72a) followed by (72d) (uttered by the same speaker) appears to make up a consistent statement. (72b)-(72c), on the other hand, cannot consistently be continued with (72d). Restricting our attention to *possible* and *possibly*, we see that there must be a truth-conditional difference between the two. Impressionistically, the difference is this: (72a) simply states that there is some possibility, however remote and implausible, that Le Pen will win. (72b) does not permit such remoteness: It has it that there is a realistic chance for Le Pen to win. In other words, (72b) constitutes a *stronger statement* than (72a).

2.3.3 Widen up and strengthen

According to Kadmon and Landman (1993); Krifka (1995); Lahiri (1997); Chierchia (2001), the ungrammaticality of (73a) is a *pragmasemantic* phenomenon. I follow the implementation in Chierchia (2001) here. The idea is essentially that *any* is synonymous with the indefinite article *a* in the sense that they are both existential quantifiers. Thus the meaning of both (73a,73b) could be represented as (73c), where D subscripted to \exists is a contextual restriction, a quantificational domain (Westerståhl, 1988) and \exists represents an existential generalized quantifier, i.e. $\lambda X \lambda Y. X \cap Y \neq \emptyset$.

- (73) a. * John has any potato.

- b. John has a potato.
- c. $\exists_D(\text{potato})(\lambda x.\text{has}(j, x))$

The difference between the two is argued to be that (73a) involves expansion of D to a different quantificational domain $D' \supseteq D$. In other words, *any* invites us to consider more potatoes; it signals reduced acceptance of exceptions. Thus a more accurate representation of the ungrammatical sentence (73a) would be (74) where g is a function from sets to sets, such that $\forall X(X \subseteq g(X))$.

$$(74) \quad \exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x))$$

So far, this does not explain why (73a) is ungrammatical. The reason for this is that domain expansion is subject to a strengthening condition: The result of domain expansion must entail the same proposition without domain expansion. The fact that there is a potato in some large set such that John has that potato does not entail that he has a potato in a subset of that large set. In order to implement strengthening compositionally, Chierchia defines an operator Op that universally closes the function variable g as follows:

Definition 2.2 *Let Δ be a contextually determined set of domain expansions, let φ be sentential constituent containing a free occurrence of g , and φ' be φ with all free occurrences of g removed. Then*

$$\begin{aligned} \text{Op}(\varphi) &= \forall g \in \Delta[\varphi], \text{ if } \forall g \in \Delta[\varphi] \text{ entails } \varphi', \\ \text{else} &= \text{undefined.} \end{aligned}$$

Op does the following: It takes a formula φ containing a free occurrence of a domain expansion variable g and universally quantifies g just in case the result of this whole operation entails φ without domain expansion. Otherwise it is undefined. Consider the following examples.

$$(75) \quad \text{John doesn't have any potatoes.}$$

Derivation 2.1

$$\neg \exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x))$$

apply **Op** ►

$$\text{Op}(\neg \exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x)))$$

check strengthening ►

$$\begin{aligned} \forall g \in \Delta[\neg \exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x))] &\models \\ \neg \exists_D(\text{potato})(\lambda x.\text{has}(j, x)) & \end{aligned}$$

►

$$\forall g \in \Delta[\neg \exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x))]$$

Derivation 2.2

$$\begin{array}{ll}
\exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x)) & \\
& \text{apply Op} \blacktriangleright \\
\text{Op}(\exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x))) & \\
& \text{check strengthening:} \blacktriangleright \\
\forall g \in \Delta[\exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x))] \not\models & \\
\exists_D(\text{potato})(\lambda x.\text{has}(j, x)) & \\
& \blacktriangleright \\
\text{undefined.} &
\end{array}$$

(76) * John has any potatoes.

In words, the bottom line of Derivation (2.1) says that, for any domain expansion that we are willing to entertain, there is no potato in the expanded domain that John has. This clearly entails that there is no potato in the (unexpanded) quantificational domain that John has, so strengthening is satisfied. In Derivation (2.2), strengthening is not satisfied, hence application of **Op** is undefined. In Derivation (2.2), we actually have that that the “unexpanded” alternative entails the result of application of **Op**, i.e. we have that

$$\exists_D(\text{potato})(\lambda x.\text{has}(j, x)) \models \forall g \in \Delta[\exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x))]$$

This will always happen if domain expansion applies in an upwards entailing (UE) environment. This is because of the fact that if a (small) set A has a non-empty intersection with another set B , then every superset (expansion) A' of A will also have a non-empty intersection with B . Under DE operators, this entailment relation by definition reverses, i.e. the unifying property of DE operators is precisely reversal of entailment relations. Hence, domain expansion is allowed exactly when it occurs in a DE environment. In van Rooy (2002) it is argued that one can derive the stipulative part of this proposal (i.e. that domain expansion requires *strengthening*) from very general and plausible assumptions about the pragmatics of statements, ultimately related to the Gricean notion of *relevance*. The latter notion, he derives from a decision theoretic notion of *utility*. For reasons of space, I cannot go into the details of this here.

This account explains why *any* only occurs in monotone decreasing (entailment reversing) contexts from its lexical semantics. See e.g. Krifka (1995) for a congenial analysis of several other polarity items. In the next section, I will show how the same kind of analysis can be extended to account for the behavior of *maybe* and *possibly*. The idea is essentially that these adverbs are associated with domain *shrinkage*. This explains their contrast with the adjective *possible* noted in the previous section. It also derives their status as weak PPIs.

2.4 *possibly*: Shrink, but don't weaken!

2.4.1 Modal bases

We have seen that there is an intuitive sense in which statements of the form *possibly*(φ) are stronger than statements of the form *it is possible that*(φ). By this we know that, if Γ is a DE operator, $\Gamma(\textit{possibly}(\varphi))$ should be a *weaker* statement than $\Gamma(\textit{it is possible that}(\varphi))$. If the use of *possibly* involves some operation that is subject to a strengthening condition, we can explain the distribution of this adverb in a manner entirely parallel to the explanation of the distribution of *any*. In order to achieve this, we need to look at what *possibly* means.

As a start, I take an information state (belief state) to be a set of possible worlds K , the set of worlds compatible with what we take to be true. If $K = W$, the set of all possible worlds, we are in a state of total ignorance: everything is possible. Upon learning a new proposition p , the new information state K' is given by $K \cap p$. Suppose p is a contradiction. Then $K \cap p = \emptyset$, i.e. the absurd information state. If $|K| = 1$, i.e. it only contains one world, we are in a state of total information. All this is entirely standard. What is the result, given K , of learning that something is possible i.e. *possibly*(φ)? According to a standard view (Groenendijk et al., 1996), the result should be K if $K \cap \varphi \neq \emptyset$, otherwise it should be \emptyset . Groenendijk et al. (1996) themselves point out that this leads to a situation where one can never really learn anything new from possibility statements. Like Groenendijk et al. (1996), I choose to live with this problem for the purposes of this chapter. I add that epistemic possibilities work on *modal bases* Kratzer (1977, 1991). In the default case, I take the modal base to be W . Thus, the meaning of *possible* becomes as in shown in (77):

$$(77) \quad \llbracket \textit{possible} \rrbracket^K = \lambda p[p \cap W \cap K \neq \emptyset]$$

Given that K is always a subset of W , this is equivalent to $\lambda p[p \cap K \neq \emptyset]$. In order to derive the strengthening effect observed with *possibly*, I take this expression to come with domain shrinkage of the modal base. In (78), g is a variable over domain shrinks, i.e. for all X , $g(X) \subseteq X$.

$$(78) \quad \llbracket \textit{possibly} \rrbracket^K = \lambda p \lambda g[p \cap g(W) \cap K \neq \emptyset]$$

Just as with the domain expansion associated with *any*, I assume that our function variable must be universally closed. To this effect, we can use Chierchia's **Op**, but now with respect to domain shrinking functions, and with respect to domains consisting of possible worlds. Thus, *possibly* applied to *Le Pen will win* returns Derivation (2.3). Δ now contains various ways in which we can constrain our modal base in various contexts. For instance, If we are discussing the French election, we can shrink W by intersecting it with plausible assumptions about the behavior of the French electorate. The bottom line of Derivation (2.3) says that the intersection of "win" with such a modal base has a non-empty intersection with K , our information state. Clearly, this can

Derivation 2.3

$$\begin{array}{ll}
\lambda g[\text{win} \cap g(W) \cap K \neq \emptyset] & \\
& \text{apply Op} \blacktriangleright \\
\text{Op}(\lambda g[\text{win} \cap g(W) \cap K \neq \emptyset]) & \\
& \text{check strengthening} \blacktriangleright \\
\forall g \in \Delta[\text{win} \cap g(W) \cap K \neq \emptyset] \models [\text{win} \cap W \cap K \neq \emptyset] & \\
& \blacktriangleright \\
\forall g \in \Delta[\text{win} \cap g(W) \cap K \neq \emptyset] &
\end{array}$$

only happen if “win” has a non-empty intersection with K to begin with, so strengthening is satisfied.

This already derives the fact that *possibly* cannot occur in DE contexts. Before I move on to demonstrate this, however, I need to make sure that this setup actually derives the observed difference between *possible* and *possibly*, i.e. *possible*(win) is consistent with *certainly*(¬win) while *possibly*(win) is not. The standard semantics for *certainly* would be as in (79).

$$(79) \quad \llbracket \text{certainly} \rrbracket^K = \lambda p[K \subseteq p]$$

By this, *it is possible that Le Pen will win, even though he certainly won't* comes out as in (80), which is unfortunately inconsistent.

$$(80) \quad [\text{win} \cap K \neq \emptyset] \wedge [K \subseteq \neg \text{win}]$$

A solution that suggests itself is to make use of modal bases with this adverb as well. Thus, we could try to assume that *certainly* also comes equipped with a W -shrinking variable. There are two options, according to whether the modal base intersects with K or p . I consider them in turn.

$$(81) \quad \llbracket \text{certainly} \rrbracket^K = \lambda p \lambda g[g(W) \cap K \subseteq p]$$

Applying this to ¬win, we get the following.

The problem is that if we have that $A \subseteq B$, it does not follow from $A \subseteq C$ that $B \subseteq C$. In other words, our function variable g is already in a DE environment in Derivation (2.4). In fact, given our meaning for *certainly*, it should behave as a negative polarity item. But this is plainly wrong, given that (82) is quite impeccable.

$$(82) \quad \text{Chirac will certainly win.}$$

The other option would be to intersect our modal base with p rather than with K . This yields (83).

$$(83) \quad \llbracket \text{certainly} \rrbracket^K = \lambda p \lambda g[K \subseteq p \cap g(W)]$$

Derivation 2.4

$$\begin{array}{ll}
\lambda g[g(W) \cap K \subseteq \neg \text{win}] & \\
& \text{apply Op} \blacktriangleright \\
\text{Op}(\lambda g[g(W) \cap K \subseteq \neg \text{win}]) & \\
& \text{check strengthening} \blacktriangleright \\
\forall g \in \Delta[g(W) \cap K \subseteq \neg \text{win}] \not\models [K \subseteq \neg \text{win}] & \\
& \blacktriangleright \\
\text{undefined.} &
\end{array}$$

This would make (84a) a stronger statement than (84b). This does not correspond to our intuitions, I think. If anything, asserting (84b) is a stronger statement than the assertion of (84a). (84a) seems to indicate that there is some (however small) room for doubt.

- (84) a. Chirac has certainly won.
 b. Chirac has won.

Be that as it may, the meaning for *certainly* in (83) does not help us. (85), which is the meaning our example with *it is possible that...* would get now, is still inconsistent.

$$(85) \quad [\text{win} \cap K \neq \emptyset] \wedge \forall g \in \Delta[K \subseteq \neg \text{win} \cap g(W)]$$

2.4.2 Entrenched beliefs

The discussion in the previous subsection seems to indicate that we need something different. I follow van Rooy (2001) in assuming that an information state should be thought of as an ordering relation on the set of propositions $\varphi \subseteq W$. Intuitively, the ordering relation tells us how *plausible* a proposition is, given what we know. To do this, we need a way to compare the plausibility of worlds, given what we know. The definitions below are taken from van Rooy (2001), and ultimately from Harper (1976). They implement what is known as “Harper’s Principle” for determining a similarity relation among worlds:

- (86) Harper’s Principle (HP)
 Only propositions decided by K should count in determining comparative similarity relative to K .

Definition 2.3 Let $S = \wp W$, let $x, y \in W$ and K be a belief state. Then

$$S_y^x K = \{p \in S \mid (K \subseteq p \text{ or } K \subseteq \neg p) \text{ and } ((x \in p \text{ and } y \notin p) \text{ or } (x \notin p \text{ and } y \in p))\}$$

$S_y^x K$ gives us the set of propositions which are decided by K and on whose truth value the two worlds x and y disagree. We can now say that a world x is closer to what we take to be true, or more plausible, than another world y if x disagrees less with what we know than y .

Definition 2.4 Let x, y be worlds. $x \preceq_K y$ iff for all $w \in K$:

$$|S_x^w K| \leq |S_y^w K|$$

The relation \preceq_K gives us a system of spheres of worlds of the kind proposed by Lewis and Stalnaker in the early seventies to account for conditionals and counterfactuals. We now need to know what set of worlds to associate with propositions. This is done by the following *selection* function. It tells us to only consider the most plausible worlds in p .

Definition 2.5 $C_K(p) = \{v \in p \mid \forall u \in p : v \preceq_K u\}$

Given that we can count the number of propositions in $S_y^x K$, we can devise a *quantitative* measure of plausibility.

Definition 2.6 Let $k(w/p)$ be the implausibility of a world w , after updating K with p . Let $k(p/q)$ be the implausibility of a proposition p after updating K with q .

$$\begin{aligned} \text{worlds:} \quad k(w/p) &= |S_u^w C_K(p)|, \text{ for any world in } C_K(p). \\ \text{propositions:} \quad k(q/p) &= \min\{k(w/p) \mid w \in q\} \end{aligned}$$

A proposition p is *accepted* in K iff $k(\neg p/\top) > 0$. p is more implausible than q w.r.t. K if $k(p/\top) > k(q/\top)$. Let $f(p) = k(\neg p/\top)$. Then f measures the level of *plausibility*, or *epistemic entrenchment* of p , given K . We follow van Rooy (2001) in defining an information state (belief state) K as just such an entrenchment relation on a set of possible worlds.

I will now use the entrenchment relation f to give a semantics for epistemically modal statements. In van Rooy (2001), he proposes that what he calls ‘evidential’ attitude reports, like *be certain that*, *be sure that*, *be convinced that* can be treated within the following schema, where f_w^a is the entrenchment function associated with a in w :

$$(87) \quad \llbracket a \text{ that } p \rrbracket^w = 1 \text{ only if } f_w^a(p) = \text{HIGH}$$

“HIGH” is a contextually determined number. I take it to be determined in the following way: One picks some strongly believed proposition p . $\text{HIGH} = f_w^a(p)$ and $\text{LOW} = f_w^a(\neg p)$. Given that $f_w^a(p) = k_w^a(\neg p/\top)$ by definition, this setup ensures that HIGH and LOW are inversely proportional. In order to implement our treatment of positive polarity, I make use of the following epistemic accessibility relation (\mathcal{E}):

$$(88) \quad \mathcal{E}(a, w) = \{w' \in W : k_w^a(w') \leq n\}$$

where n is either HIGH or LOW.

I write $\mathcal{E}^\uparrow(a, w)$ when n is to be understood as HIGH and $\mathcal{E}^\downarrow(a, w)$ when it is to be understood as LOW. Furthermore, I normally skip specification of a and w . For speaker oriented adverbs, a is always the speaker¹⁹ and w is always the world of evaluation. (88) either gives us the set of worlds whose implausibility is at most LOW (i.e. \mathcal{E}^\downarrow) or the set of worlds whose implausibility is at most HIGH (i.e. \mathcal{E}^\uparrow) Our meaning for *certainly* can now be represented as follows:

$$(89) \quad \llbracket \text{certainly} \rrbracket^K = \lambda p[p \cap \mathcal{E}^\downarrow \neq \emptyset]$$

Given that $C_K(p) = \{v \in p \mid \forall u \in p : v \preceq_K u\}$, i.e. the most plausible worlds in p , (89) actually ensures that if *certainly*(p) is true, then $f(p) \geq \text{HIGH}$. This seems to me the right result. It derives the fact that *certainly*(p) leaves some little room for doubt; it says that p is very plausible, given K . The adjective *possible* can now be given the following representation:

$$(90) \quad \llbracket \text{possible} \rrbracket^K = \lambda p[p \cap \mathcal{E}^\uparrow \neq \emptyset]$$

Again, this ensures that the plausibility of p is at least LOW. Our problematic example, *it is possible that Le Pen will win, even though he certainly won't*, now comes out as follows:

$$(91) \quad [\text{win} \cap \mathcal{E}^\uparrow \neq \emptyset] \wedge [\neg \text{win} \cap \mathcal{E}^\downarrow \neq \emptyset]$$

This is consistent, given the way we determine HIGH and LOW. More in particular, (91) is true just in case $f(\text{win}) = \text{LOW}$ and $f(\neg \text{win}) = \text{HIGH}$. As the reader may have anticipated, I will now derive the meaning of *possibly* from that of *possible* by applying domain shrinkage to \mathcal{E}^\uparrow . Such domain shrinkage is still thought to be contextually restricted, but we define a constraint on possible domain shrinks:

Definition 2.7 *Let ∇ be the set of domain shrinks and let n be any natural number, s.t. $n < \text{HIGH}$, then in context c ,*

$$\nabla \subseteq \{g_n \mid g_n(\mathcal{E}^\uparrow) = \{w' \mid k_w^a(w') \leq n\}\}.$$

Applying a domain shrink takes us to a domain of worlds whose plausibilities are strictly greater than LOW. The meaning of *possibly* is the following, where, as before, g has to be universally closed by application of **Op**.

$$(92) \quad \llbracket \text{possibly} \rrbracket^K = \lambda p \lambda g[p \cap g(\mathcal{E}^\uparrow) \neq \emptyset]$$

Op is defined as before, but now operating on $g \in \nabla$. The derivation of (93) is given in Derivation (2.5).

$$(93) \quad \text{Le Pen will possibly win.}$$

¹⁹Except, potentially when they occur embedded under an attitude verb like *believe*. Similarly for w . I ignore this complication here.

Derivation 2.5

$$\begin{array}{lcl}
\lambda g[\text{win} \cap g(\mathcal{E}^\uparrow) \neq \emptyset] & & \\
& \text{apply Op} \blacktriangleright & \\
\text{Op}(\lambda g[\text{win} \cap g(\mathcal{E}^\uparrow) \neq \emptyset]) & & \\
& \text{check strengthening} \blacktriangleright & \\
\forall g \in \nabla[\text{win} \cap g(\mathcal{E}^\uparrow) \neq \emptyset] \models [\text{win} \cap \mathcal{E}^\uparrow \neq \emptyset] & & \\
& \blacktriangleright & \\
\forall g \in \nabla[\text{win} \cap g(\mathcal{E}^\uparrow) \neq \emptyset] & &
\end{array}$$

Strengthening is satisfied. The bottom line of Derivation (2.5) states (indirectly) that the plausibility of “win” must be strictly greater than LOW. This clearly entails that it must be greater than or equal to LOW.

We can now test whether (94a) is inconsistent, which is what we want. (94b) is the meaning assigned to this proposition.

- (94) a. # Le Pen will possibly win, even though he certainly won't.
 b. $\forall g \in \nabla[\text{win} \cap g(\mathcal{E}^\uparrow) \neq \emptyset] \wedge [\neg \text{win} \cap \mathcal{E}^\downarrow \neq \emptyset]$

This is clearly inconsistent. The first conjunct states that the plausibility of “win” is strictly greater than LOW whereas the second conjunct has it that the plausibility of “¬win” is greater than or equal to HIGH. But for any p we have that if $f(p) > \text{LOW}$, then $f(\neg p)$ must be $< \text{HIGH}$. Hence (94b) is inconsistent.

2.4.3 The status of *possibly* as a PPI

I will now show that our independently motivated semantics for *possibly* derives its distribution. We noted above that whenever strengthening is satisfied in some environment, application of a DE operator Γ to that environment reverses its entailment relations, so strengthening is no longer satisfied after application of Γ . I will go through some examples to assure you that this really works. Consider (95), where *possibly* occurs under negation with its derivation in (2.6).

- (95) ?? Stanley didn't possibly eat his wheaties.

Recall that $g(\mathcal{E}^\uparrow) \subseteq \mathcal{E}^\uparrow$. The fact that “eat” has no intersection with the former therefore does not entail that it also has no intersection with the latter. Consider now the alternative Derivation 2.7 for (95), where the negation enters the derivation *after* application of **Op**. Strengthening fails here, too. The fact that not every way of constraining our domain gives us a domain that has an intersection with “eat” does not entail that the unconstrained domain has no intersection with “eat”. The reader might wonder why strengthening applies after the negation has been merged and not before. I clearly need this for the

Derivation 2.6

$\lambda g \neg[\text{eat} \cap g(\mathcal{E}^\dagger) \neq \emptyset]$
apply **Op** ►
 $\mathbf{Op}(\lambda g \neg[\text{eat} \cap g(\mathcal{E}^\dagger) \neq \emptyset])$
check strengthening ►
 $\forall g \in \nabla \neg[\text{eat} \cap g(\mathcal{E}^\dagger) \neq \emptyset] \not\models \neg[\text{eat} \cap \mathcal{E}^\dagger \neq \emptyset]$
►
 undefined.

Derivation 2.7

$\lambda g[\text{eat} \cap g(\mathcal{E}^\dagger) \neq \emptyset]$
apply **Op** ►
 $\mathbf{Op}(\lambda g[\text{eat} \cap g(\mathcal{E}^\dagger) \neq \emptyset])$
MERGE negation ►
 $\neg \mathbf{Op}(\lambda g[\text{eat} \cap g(\mathcal{E}^\dagger) \neq \emptyset])$
check strengthening ►
 $\neg \forall g \in \nabla[\text{eat} \cap g(\mathcal{E}^\dagger) \neq \emptyset] \not\models \neg[\text{eat} \cap \mathcal{E}^\dagger \neq \emptyset]$
►
 undefined.

system to work. I follow Chierchia (2001) in assuming that strengthening is checked at the PHASE level of the syntactic derivation Chomsky (1999, 2001).²⁰

Similar results obtain when *possibly* occurs within the scope of DE subjects like *few students* or DE adverbs like *rarely*. We give the Derivation (2.8) for *few students* here. We are only interested in the wide scope construal for *few students*, i.e. the reading where the subject outscopes *possibly*.

(96) ?? Few students possibly ate their wheaties.

If we take *perhaps* and *maybe* to be synonymous with *possibly*, these results extend to these adverbs as well.

Intervention

In a recent paper, Szabolcsi (2002) discusses the behavior of the English PPI *some*. She analyzes some intriguing properties of this expression there. I will

²⁰We might also say that application of **Op** is restricted to the PHASE level. In that case, Derivation (2.7) would not arise at all.

Derivation 2.8

$$\mathbf{Op}(\lambda g[\text{few}(\text{student})(\lambda x[\text{eat}(x) \cap g(\mathcal{E}^\dagger) \neq \emptyset])])$$

check strengthening ►

$$\forall g \in \nabla[\text{few}(\text{student})(\lambda x[\text{eat}(x) \cap g(\mathcal{E}^\dagger) \neq \emptyset])] \not\models$$

$$[\text{few}(\text{student})(\lambda x[\text{eat}(x) \cap \mathcal{E}^\dagger \neq \emptyset])]$$

►

undefined.

discuss two of them. In Linebarger (1987), she notes that (anti-)licensing of polarity items can be obstructed by certain kinds of interveners. Chierchia (2001) argues that this cannot be reduced to a syntactic Relativized Minimality effect, essentially, because one cannot identify the class of interveners in a natural way. Below is an example with the PPI *somewhat*.²¹

- (97) a. *Stanley didn't like this somewhat.
 b. Stanley didn't always like this somewhat.

When *always* intervenes between the negation and *somewhat*, the sentence becomes good. Recall our internet game examples. We saw that *possibly* cannot follow *never* in such cases. Of course, it also cannot follow sentential negation (98a). The interesting thing is that, also in this case, an intervening *always* has a meliorating effect (98b), i.e. (98b) seems to be much better than (98a).

- (98) a. * This is a fun, free game where you're not possibly further than a click away from winning \$1000!
 b. ...where you're not always possibly a click away from winning \$1000!

In order to see how our setup can derive this, I will give an informal presentation of how Chierchia (2001) derives similar intervention effects for the NPI *any*. First, note that *always* seems to block the licensing of this element (Linebarger, 1987).

- (99) a. Stanley didn't like anything.
 b. ?? Stanley didn't always like anything.

Chierchia notes that the class of interveners seem to share the following characteristic: While they do not (necessarily) introduce a scalar implicature in a UE environment, they *do* introduce one in a DE environment. *Always* participates in a (lexicalized) scale with *sometimes*. Thus, (100a) carries the scalar implicature that Stanley didn't always eat his wheaties. Since *always* is the

²¹This example is good on a so-called 'metalinguistic' or emphatic denial reading. See Szabolcsi (2002) for discussion.

strongest element in the relevant scale, (100b) does not introduce such a scalar implicature. However, under negation, the scale is reversed: (100c) *does* introduce the implicature that Stanley sometimes *did* eat his wheaties. I refer the reader to Krifka (1995); Chierchia (2001) for discussion of an algorithmic way to derive this.

- (100)
- a. Stanley sometimes ate his wheaties.
 ► Stanley didn't always eat his wheaties.
 - b. Stanley always ate his wheaties.
 - c. Stanley didn't always eat his wheaties.
 ► Stanley sometimes ate his wheaties.

Recall that Chierchia assumes that *any* comes with a domain expansion variable g which is universally closed by his **Op**, and that application of **Op** is subject to a strengthening condition. He now argues that if strengthening takes scalar implicatures into account²² we can derive the intervention effects as failure of strengthening. In other words, application of **Op** must yield a proposition which is stronger than the strongest meaning of the proposition without domain expansion. Suppose that we have two sets X, Y of potatoes, such that $X \subseteq Y$. Suppose that John did not always have a potato in Y , though he sometimes did. This does not entail that he did not always (though sometimes did) have a potato in X . Chierchia goes through a number of examples, including ones with numeral interveners (e.g. “*John didn't give two people anything”). For our setup, we would have to show that, while the negation reverses the strengthening relation, as we have shown, interveners would have the effect of blocking such reversal. This is not, in general, possible. In fact, most of the interveners are UE in their relevant argument. Suppose that Γ is DE and that Θ is UE. Then, p is in a DE environment in both of the following examples:

- (101)
- a. $\Gamma(p)$
 - b. $\Gamma(\Theta(p))$

Hence, addition of an implicature triggered by a (UE) intervener, should not have any rescuing effect on strengthening, as long as the implicature is added on both sides. If we only add it on one side, it could still not rescue a failure of strengthening which is what we would need. Suppose that, instead of being subject to strengthening, PPIs are subject to an *anti-weakening* constraint. In other words, application of a domain shrink must never lead to a *weaker* statement than the same statement without shrinkage. In order to implement this compositionally, we could define a universal closure operator Ω as follows:

²²Chierchia argues that there are strong reasons to assume that calculation of scalar implicatures are calculated on subconstituents of the clause, and not on the end result of the derivation as commonly assumed. I refer the reader to Chierchia's work for discussion.

Definition 2.8 Let ∇ be the set of domain shrinks as defined above. φ' is obtained from φ by removing all free occurrences of g and adding scalar implicatures.

$$\begin{aligned} \Omega(\varphi) &= \forall g \in \nabla[\varphi], \text{ if } \varphi' \not\models \forall g \in \nabla[\varphi] \\ \text{else} &= \text{undefined.} \end{aligned}$$

Though this will derive the intervention effects we have seen for PPIs while retaining the result that they cannot be (directly) outscoped by a DE operator, it does introduce an asymmetry between NPIs (subject to strengthening) and PPIs (subject to anti-weakening) which should be further motivated. One piece of support comes from the fact that that PPIs can be outscoped by non-monotone quantifiers like *exactly three* N . These couldn't satisfy strengthening for the simple reason that they are non-monotone. In other words, (102a) has a reading according to which there are exactly three people such that John gave them something. Similarly, (102b) allows the subject to outscope *possibly*, i.e. it has a reading according to which there are exactly three students such that they possibly ate their wheaties. This already shows that there is a problem with assuming strengthening for PPIs. Furthermore, the same quantifier gives rise to an intervention effect with *any* (102c).

- (102) a. John gave exactly three people something.
 b. Exactly three students possibly ate their wheaties.
 c. ?? John didn't give exactly three people anything.

We could not generalize the anti-weakening approach to NPIs, because, then, these should be licensed by non-monotone quantifiers, contrary to fact:

- (103) * Exactly three students ate anything.

In a sense, the asymmetry between strengthening and anti-weakening reflects the fact that, where NPIs are licensed, PPIs are anti-licensed (Giannakidou, 1997). I go through the derivation of our example (98b), but before doing so, I must device a meaning for *always*. I follow, among others, Beaver and Clark (2002) in assuming that *always* is a universal quantifier over events. In other words, it denotes the subset relation over sets of these, and, crucially, the restriction is always contextually given. In our example, the contextual set could be all events of playing the relevant game. Then, (104a) comes out as (104b) and (104c) comes out as (104d).

- (104) a. You're always a click away from winning.
 b. $\{e \mid \text{play}(e)\} \subseteq \{e \mid \text{1click}(e)\}$
 c. You're always possibly a click away from winning.
 d. $\forall g \in \nabla[\{e \mid \text{play}(e)\} \subseteq \{e \mid \text{1click}(e) \cap g(\mathcal{E}^\dagger) \neq \emptyset\}]$

$\{e \mid \text{1click}(e) \cap g(\mathcal{E}^\dagger) \neq \emptyset\}$, where g is an arbitrary domain shrink, is a subset of $\{e \mid \text{1click}(e) \cap \mathcal{E}^\dagger \neq \emptyset\}$. For ease of exposition, I will refer to the former set of events as A , and the latter as B . Furthermore, I will refer to the

restrictor of *always* (in our case $\{e \mid \text{play}(e)\}$), as C . The proposition “You’re not always possibly a click away from winning” thus comes out as $\lambda g[C \not\subseteq A]$ before application of Ω . Adding the scalar implicature to this, we get $\lambda g[C \not\subseteq A \wedge C \cap A \neq \emptyset]$. Applying Ω to this, we need to check that anti-weakening is satisfied. This is done by checking that the following non-entailment holds, recalling that $A \subseteq B$:

$$[C \not\subseteq B \wedge C \cap B \neq \emptyset] \not\models \forall g[C \not\subseteq A \wedge C \cap A \neq \emptyset]$$

Indeed it does. One can also see that it is the implicature introduced by the negated *always* which is responsible for the failure of the entailment. In this sense, then, *always* intervenes precisely because it introduces an implicature which blocks the weakening effect which the negation and other DE operators have on domain shrinkage. If *always* does not intervene, anti-weakening is violated. This is because, as we have seen, the following is true for any proposition p .

$$\forall g[p \cap g(\mathcal{E}^\dagger) \neq \emptyset] \models [p \cap \mathcal{E}^\dagger \neq \emptyset]$$

Since the negation is DE (even AMo), we therefore have the following, which is a violation of anti-weakening.

$$\neg[p \cap \mathcal{E}^\dagger \neq \emptyset] \models \forall g\neg[p \cap g(\mathcal{E}^\dagger) \neq \emptyset]$$

In other words, we have derived that certain interveners can rescue a PPI, just in case they are of the sort that introduce scalar implicatures under DE operators. Without such intervention, the PPI is illicit under DE operators.

Double licensing

Szabolcsi (2002) discusses another intriguing phenomenon about PPIs, namely what she calls ‘double licensing’. She attributes the observation to Jespersen. Consider (105).

- (105) a. *I doubt that Stanley liked it somewhat.
 b. *I think Stanley didn’t like it somewhat.
 c. I doubt that Stanley didn’t like it somewhat.

In this case, the adverbs behave differently from *somewhat*. (106c) does not seem to improve very much, compared to (106a, 106b). In (107), however, there is a detectable improvement.

- (106) a. ??I doubt that Stanley possibly liked it.
 b. ??I think Stanley didn’t possibly like it.
 c. ??I doubt that Stanley didn’t possibly like it.

- (107) I don’t doubt that Stanley possibly liked it.

This bears on the assumption that strengthening/anti-weakening is checked at the PHASE level. For example, for (106a) to come out bad, it is essential that strengthening cannot be checked in the embedded clause. On the other hand, for (108a) to come out good, it seems to be crucial that it *can* be checked in the embedded clause. (108b), in turn seems to require the opposite. It might be that this could be made to follow from properties of the verbs in question. For instance, *think* is known to be a “neg-raising” verb, in the sense that the matrix negation in (108b) seems to be equivalent to negating the embedded clause.

- (108) a. Nobody thinks that Le Pen will possibly win.
 b. ?? John doesn't think that Le Pen will possibly win.

One phenomenon (pointed out to me by Anastasia Giannakidou (p.c.)), which could be taken as a double licensing phenomenon is the following. As we have seen, *possibly* is generally bad under negation. However, (109a,b), involving another modal seem to be much better, in fact, quite perfect. Note, however, that *possibly* seems to be the only adverb for which a second modal has a “double licensing” effect. Thus, there is no meliorating effect of the presence of the modal verb in (109c). Note also that the good examples (109a,b) always get an emphatic/contrastive reading, which, as we know, can rescue a PPI alone. Finally, English is the only language I am aware of that exhibits this phenomenon. This is illustrated by the sharp ungrammaticality of the Norwegian sentence (109d).

- (109) a. Stanley couldn't possibly have eaten his wheaties.
 b. Stanley can't possibly have eaten his wheaties.
 c. *Stanley can't [probably/ maybe/ perhaps/ evidently/ fortunately/ paradoxically/ ...] have eaten his wheaties.
 d. *Ståle [an/ kunne] ikke [muligens/ kanskje] ha spist
 S [can/ could] not [possibly/ maybe] have eaten
 hvetekakene sine.
 the-wheaties his

This sheds doubt on the hypothesis that this phenomenon should be analyzed as double licensing in Szabolcsi's sense. The double licensing phenomena she discusses are crosslinguistically robust, and generalize to more than one item. Furthermore, we have seen that the adverbs are not generally rescued by adding another licenser on top of negation (106c). Thus, I will put the cases in (109a,b) aside for now as a quirky property of one expression in one language. In sum, it seems doubtful that Szabolcsi's double licensing phenomenon applies to adverbial PPIs, although more research is needed to capture the difference between e.g. (106c) and (107).

Summary

In this section, we have seen that the behavior of *possibly* can be derived from its semantic difference with the proposition-embedding adjective *possible*. More in particular, we have seen that *possibly* yields stronger statements than *possible* and that this can be implemented in van Rooy's (2001) framework, by assuming that the adverb comes with a variable over domain shrinking functions, mapping the epistemic accessibility relation \mathcal{E}^\uparrow to one of its subsets. Intuitively, what such domain shrinkage does is increasing the level of plausibility ascribed to the modified proposition. We have furthermore seen that if such domain shrinkage is subject to an anti-weakening constraint, and that such anti-weakening takes scalar implicatures into account, the limited distribution of the adverb follows. Hopefully, the anti-weakening constraint on domain shrinkage can be derived from general pragmatic constraints in the spirit of van Rooy (2002) and the way this author derives the strengthening constraint for NPIs. If so we will have a general explanation of the distribution of SOA and PPIs more generally.

I would like to stress that the machinery employed here for polarity phenomena and for epistemic modality was developed independently, and for other purposes. In other words, the claim is that the distribution of such expressions as *possibly* can be made to follow from already existing theories of modality and polarity.

2.5 Prospects and consequences

I have concentrated on deriving the behavior of *possibly* here. The reason for this is mainly that a similar derivation for all the speaker oriented adverbs we discussed in section 2 would be too ambitious a task for the present thesis. Nevertheless, I think the data discussed in section 2 and 3 allow us to be optimistic about the prospects of explaining the distribution of SOA, phase quantifiers and frequency adverbs in terms of their (independently motivated) scopal requirements. For manner adverbs, and aspectual adverbs like *completely* which generally cannot outscope any other adverbs, such an analysis might seem less adequate. They are not polarity items, for example. However, we noted that some manner adverbs *can* occur high, but then, they receive a subject-oriented or speaker-oriented reading. This can be treated as a scopal effect (Ernst, 2000), if we assume that they come with an implicit variable that must be c-commanded by its binder. For adverbs like *completely*, which do not behave like this, we would need to motivate an analysis according to which such adverbs do *not* come with an implicit variable, or that a subject/speaker-oriented reading would always lead to semantic or pragmatic degradedness. This would explain why they don't receive these readings, but in order to explain why they can't *occur* in high positions, we would need something more. It has been noted quite frequently that these adverbs interact with the argument structure

and temporal constitution of the main VP (Chomsky, 1965; Alexiadou, 1997; Cinque, 1999). Thus, I speculate that, given a proper explanation of that, it might be made to follow that these adverbs cannot apply to a predicate which has already been modified by frequency adverbs, phase quantifiers, or SOA.

There are indications that the reasoning pursued in the present chapter extends to some auxiliaries. Consider, for example the pair *might*, *can*. The former, but not the latter appears to be a PPI. Furthermore, while the former is restricted to an epistemic interpretation, the latter has a wider range of (root) modal uses. Even when *can* is interpreted epistemically it isn't a PPI.²³ This fact seems to show that one can't account for the PPI status of *might* by assuming that there is an epistemic modality head M which c-commands a functional head Neg hosting the negation. Under standard assumptions, epistemic interpretation of *can* should force it to move to M (at LF), if there is such a head, so we would be lead to expect that epistemic *can* should also be a PPI, contrary to fact. Otherwise, one would be forced to say that only the epistemic modals which are, in fact, PPIs occupy M at LF, thus rendering the account vacuous.

- (110) a. The university might be in that direction.
 b. The university might not be in that direction.
 might > not. cf. *mightn't
 c. ?? Nothing interesting might be in that direction.
 d. ? Few interesting sites might be in that direction.
 e. ?? I doubt that the university might be in that direction.
- (111) a. The university can be in that direction.
 b. The university can't be in that direction
 not > can
 c. Nothing interesting can be in that direction.
 d. Few interesting sites can be in that direction.
 e. I doubt that the university can be in that direction.

2.5.1 Short Verb Movement

What are the consequences for a syntactic approach like that of Cinque (1999), if a similar account can be developed for all adverbs? I would like to separate this into two questions. The first one is whether or not adverbs are to be analyzed as specifiers of unique (functional) heads. I do not think the present chapter bears on that question at all. Ultimately, it depends on how one should analyze purely syntactic phenomena like verb placement among adverbs in the

²³Giannakidou (1997) also argues that Greek subjunctive morphology is an NPI. Certain kinds of Germanic (embedded) verb-second clauses appear to be PPIs, i.e. the ("bridge") predicates which allow embedded V2 are never DE.

Romance languages (Pollock, 1989; Cinque, 1999). I return to this question in chapter 4.

The second question is this. Assuming that adverbs are specifiers of unique functional heads, how should we derive the ordering of these heads? The standard answer is that functional heads are ordered by syntactic selection. It seems to me that the present chapter does bear on this question. In particular, if the present account is on the right track, it seems preferable to assume that functional heads are *not* ordered by selection. To make this point clear, I would like to point out some Norwegian facts that plainly cannot be accounted for by assigning positions to adverbs in a linear sequence of functional heads (Nilsen, 2001).²⁴ There are triplets of adverbs that enter into non-linear ordering patterns. We have seen that the Norwegian adverb *muligens* ‘possibly’ has to precede sentential negation. *alltid* ‘always’, on the other hand has to follow it.

- (112) a. Jens hadde ikke alltid pusset tennene sine.
 J had not always brushed the-teeth his
 b. *Jens hadde alltid ikke pusset tennene sine.
 J had always not brushed the-teeth his

The ungrammaticality of (112b) probably relates to the observation (Beghelli and Stowell, 1997) that universals generally don’t like to immediately outscope negation. By linearity, we now expect that *muligens* ‘possibly’ has to precede *alltid* ‘always’. But we have already seen that this is not true. We repeat the relevant example here.

- (113) Dette er et morsomt, gratis spill hvor spillerne alltid
 this is a fun free game where the-players always
 muligens er et klikk fra å vinne \$1000!
 possibly are one click from to win \$1000

Thus the ordering of *muligens*, *ikke*, *alltid* isn’t linear. That is to say, we cannot account for their relative ordering by assigning positions in a linear sequence to them. Now assume that they are to be analyzed as specifiers of functional heads. If these heads are ordered by selection, we would be forced to assume multiple positions for the adverbs. To see this, suppose that we allow Norwegian *alltid* ‘always’ to occupy a “high” position, and let this be responsible for the examples where this adverb precedes *muligens* ‘possibly’. Given that, as we have seen, *muligens* can follow the negation (*ikke*) just in case *alltid* intervenes, we might consider adding an extra, low position for *muligens* as well. This results in the tree in figure 2.1. This accommodates all the grammatical examples, but also many ungrammatical ones: in particular, this tree leads us to expect sentences like (112b) and ones where *muligens* immediately follows the negation to be possible, contrary to fact. Hence, it follows that some extra

²⁴Given binary branching, functional heads ordered by selection necessarily give rise to a linear sequence of specifiers.

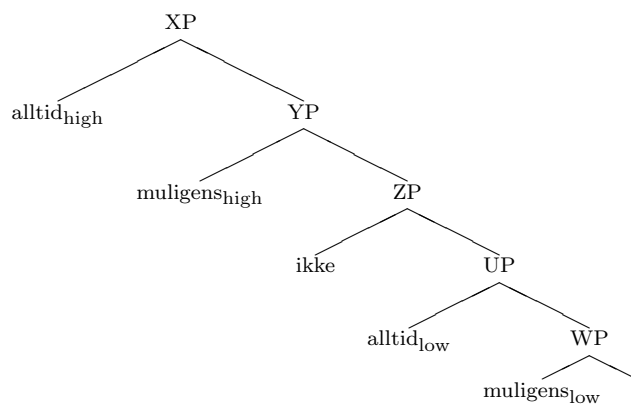


Figure 2.1: transitive hierarchy for non-transitive triplet

theory, such as the one advocated in the present chapter, is needed to rule these out. But the extra theory seems to do quite well on its own, i.e. without extra positions and selectional sequences. In other words, ordering phenomena such as these don't provide arguments for selectional sequences.

An alternative, if one wants to say that adverbs are specifiers of unique functional heads, would be to say that these heads are freely ordered. This would amount to assuming some kind of IP-shell analysis. A particularly intriguing phenomenon to handle for such an approach is the following, discussed by Cinque (2000a, to app.). In the Romance languages, there is considerable variation with respect to which adverbs a given verb form can precede or follow. For instance, Cinque shows that a French past participle can follow manner adverbs like *bien* 'well', whereas Italian past participles can't follow *bene* 'well'.

- (114) a. Il en a bien compris peine la moitié.
 he of-it has well understood hardly the half
- b. Gianni ha (*bene) capito (bene).
 G has (*well) understood (well)

Furthermore, Cinque notes that the following generalization appears to be true: If in a language L, the past participle (or another verb form) can follow (precede) a given adverb, then it can follow all other adverbs that can precede (follow) that adverb. In other words, verb-adverb ordering is transitive.²⁵ If true, I take this to show that whatever is the right account for adverb-adverb ordering, it should somehow carry over to verb-adverb ordering. In other words,

²⁵If our result that adverbs like *still*, *already*, *always*, *often*, etc. can precede "high" adverbs (SOA) carries over to Italian and other Romance varieties, this is problematic for Cinque's transitivity claim, since he shows that the past participle generally cannot precede SOA in Romance, but they generally can precede the Romance translations of *still*, etc.

if we account for adverb distribution in terms of the semantic scopal requirements of the individual adverbs, we should somehow be able to account for “short verb movement” phenomena as scopal in nature as well.

First, if short verb movement is to be treated as a scopal phenomenon, it had better show some scopal effects. In fact, (Cinque, 1999, p49, n8) points out that, whereas (115a) entails that Gianni still has long hair, (115b) is compatible with him now having short hair.²⁶

- (115) a. Gianni ha sempre avuto i capelli lunghi.
 G has always had the long hairs
- b. Gianni ha avuto sempre i capelli lunghi
 G has had always the long hairs

Suppose that past participial morphology is treated as a “modifier” on a par with adverbs like *sempre*, but that it attracts the verb.²⁷ Suppose, furthermore, that it conveys anterior tense of the constituent it modifies.²⁸ Then the position of the verb in (115) would be a function of where the past participial morphology is merged. If it is merged above *sempre*, as would be the case in (115b) anterior tense applies to the constituent [always HAVE long hair], thus allowing for the possibility that the state of affairs denoted by this constituent no longer holds. If it applies below *sempre*, the adverb would result in what has come to be known as the “universal” perfect (Dowty, 1979; Vlach, 1993; Iatridou et al., 2002), which does entail that the state of affairs still holds. In such a theory, one would have to derive the fact that perfect morphology cannot outscope certain adverbs by means of the scopal requirements of the perfect and the adverbs.

In fact, the claim that Romance participles can’t precede “high” adverbs may not even be true. An internet search (google) for exact strings like “dato fortunatamente” (‘given fortunately’), “avuto probabilmente” (‘have-PTC probably’), etc. returned several hundred hits among which were the following sentences.

- (116) a. Due incendi che non hanno **avuto fortunatamente**
 Two fires that not have-3PL had fortunately
 conseguenze rilevanti si sono sviluppati
 consequences relevant SI are developed

²⁶My French informants report that there is no such semantic contrast in this language.

²⁷such attraction may be made to follow on phonological grounds. The affix and the verbal stem are both “phonologically incomplete” in the sense that they do not make up phonological words on their own.

²⁸Iatridou et al. (2002) argue that one should not, strictly speaking, analyze the perfect as “anterior”, i.e. that the Reichenbachian R.E interval is irrelevant for the perfect. See Borik (2002) for a sophisticated implementation of Reichenbachian ideas, which appears to take care of the objections raised in Iatridou et al. (2002) (although Borik herself does not relate her analysis to that of Iatridou et. al.). The precise semantics of the perfect need not concern us in this informal discussion, although ultimately it is, of course, important.

- b. le analisi hanno **dato fortunatamente** esito
 the analyses have-3pl had fortunately output
 negativo
 negative
- c. è stato probabilmente stampato a Roma
 is-3SG been probably printed in Rome

There are two scenarios, depending on which status we assign to examples like (116). I discuss them in turn.

Suppose that these are, in fact, straightforwardly grammatical in Italian. Then past participial movement is allowed around very high adverbs. Then, given that the finite verb can also precede or follow high adverbs, we face the following problem which was pointed out by Bobaljik (1999) for argument ordering versus adverb ordering. Auxiliaries (and arguments) come in a specific order, so, in general, we have that, given a sequence of three auxiliaries aux_1, aux_2, aux_3 and one adverb a of the relevant type, a can occupy any position in the sequence of auxiliaries, as long as the relative ordering of auxiliaries remains constant. Conversely, we have that given a sequence of adverb a_1, a_2, a_3 , and one auxiliary aux , aux can occupy any position in the sequence of adverbs, as long as the relative ordering of the adverbs remains the same. It follows that there can be no single selectional sequence accommodating the relative ordering of both adverbs and auxiliaries.

Suppose, on the other hand, that examples like (116) can be disregarded, for example because the adverb in question is in a so-called comma-intonation. Then, we can maintain Cinque's view that the participle can't move around these adverbs.²⁹ We have seen above that phase quantifiers like *still* can precede "high" adverbs in English. Cinque shows that Italian participles *can* precede such adverbs. If our result that *still* can precede *probably* carries over to Italian *ancora* and *probabilmente*, we have another failure of transitivity: PTC can precede *ancora*, *ancora* can precede *probabilmente*, but (in the current scenario) PTC can't precede *probabilmente*. The following examples were found on the internet, suggesting that our result *does* carry over to Italian.

- (117) a. La risposta è ancora probabilmente no
 the answer is still probably no
- b. Gli americani sono ancora probabilmente in maggioranza,
 the Americans are still probably in majority
 ma non per molto.
 but not for long

A possible defense of functional sequences might be that one can accommodate this by assuming an extra, high position to be available for *ancora*. But adding extra positions for different elements significantly expands the expressive power of the system, and, leads to overgeneration in some cases. Secondly, although

²⁹A relevant question is, of course, what the precise status of the "comma-intonation" is.

it may be made to work for the current problem, we have already seen that it will not do for the transitivity failure we noted for Norwegian.³⁰

A question raised by the suggestion that verb placement can be treated as a scopal phenomenon in the manner outlined here, raises the following question: Why doesn't every language have short verb movement? In fact, I think there is evidence that even languages like Norwegian *do* have short verb movement. Consider (118a) which exhibits crossing scope dependencies between adverbs and verbs. In the sharply ungrammatical (118b) the adverbs immediately precede the constituent they modify.³¹ Crossing scope dependencies are entirely unexpected if absence of SVM is analyzed by letting the verbs remain in-situ.

(118) Norwegian

- a. ...at det ikke lenger alltid helt kunne ha blitt
 ...that it not any.longer always completely could have been
 ordnet.
 fixed
- b. *...at det ikke kunne lenger ha alltid blitt helt
 ...that it not could any.longer have always been completely
 ordnet.
 fixed

Therefore, (118a) may be taken as evidence for a rather elaborate system of short verb movements in Norwegian, inspired by Koopman and Szabolcsi (2000) on verbal complexes in West Germanic and Hungarian.³² In particular, this pattern would arise if scope reflects the order of merger, and each adverb attracts the projection of the closest verb, and each verb attracts the projection of the closest adverb, as illustrated in derivation 2.9.

Such an approach could be extended to phenomena like “affix hopping” in English and other languages. In chapter 4, such a “generalized verb raising” approach to SVM is developed in more detail.

2.5.2 Semantic selection

In Bartsch (1976); Ernst (2001), analyses of adverb distribution have been promoted which share with the present chapter the assumption that it should

³⁰The difference between the two cases is that, in the Norwegian case, we made use of the relation “x *must* precede y”, while the current problem concerns failure of transitivity of the relation “x *can* precede y”. This is why extra positions can rescue the latter, but not the former.

³¹the scope of *alltid* and *lenger* is not easy to determine, so one could also let the former adverb outscope *blitt* ‘been’ or the latter be outscoped by this verb.

³²Bentzen (2002) shows that SVM is visible in some varieties of Norwegian, including my own. For example, examples like (i) are perfectly grammatical.

(i) Jeg har ikke spist **ofte** tran.

I have not eaten often cod.liver.oil

However, in my variety, this is limited to a few adverbs, thus I could not substitute *alltid* ‘always’ for *ofte* ‘often’ in (i).

Derivation 2.9

[completely [fixed]]	
	MOVE VP ►
[fixed [completely]]	
	MERGE <i>been</i> ►
[been [fixed [completely]]]	
	MOVE AdvP ►
[completely [been [fixed]]]	
	MERGE <i>always</i> ►
[always [completely [been [fixed]]]]	
	MOVE VP ►
[[been [fixed]] [always [completely]]]	
	MERGE <i>have</i> ►
[have [[been [fixed]] [always [completely]]]]	
	MOVE AdvP ►
[[always [completely]] [have [been [fixed]]]]	
	MERGE <i>any.longer</i> ►
[any.longer [[always [completely]] [have [been [fixed]]]]]	
	MOVE VP ►
[[have [been [fixed]]] [any.longer [always [completely]]]]	
	MERGE <i>could</i> ►
[could [[have [been [fixed]]] [any.longer [always [completely]]]]]	
	MOVE AdvP ►
[[any.longer [always [completely]]] [could [have [been [fixed]]]]]	
	MERGE <i>not</i> ►
[not [[any.longer [always [completely]]] [could [have [been [fixed]]]]]]	

be given a semantic treatment. The execution of the analyses is quite different, however. While these authors make use of rich semantic ontologies and selectional restrictions, no use has been made here, of notions like “event”, “proposition”, “fact” etc. to account for the distribution of different adverbs. For example, Ernst (2001) assumes that there are different ontological entities

like “events”, “propositions” and “facts” that adverbs can apply to, and that these ontological categories enter into the following system of type-conversion (his “FEO-calculus”)

$$\text{event} \Rightarrow \text{spec-event} \Rightarrow \text{proposition} \Rightarrow \text{fact} \Rightarrow \text{speech act}$$

In this setup, some adverbs are event modifiers (e.g. *completely*), some propositional modifiers (e.g. *not*) etc. The fact that *not* must precede (outscope) *completely* now follows from their (semantic) selectional requirements and the FEO-calculus. The structure of the account is thus remarkably similar to Cinque’s. The FEO-calculus corresponds to Cinque’s hierarchy of functional projections, and the ontological categories to Cinque’s heads. As far as I can see, ontological categories and his FEO-calculus are in no less need of explanation and motivation than Cinque’s sequence of heads. In practice, Ernst always assigns identity types to his adverbs, that is, they always map a constituent of semantic type X to another constituent of type X. If this is taken to be a general restriction on his system, he also predicts adverb ordering to be linear, which we have seen that it is not. If Ernst allows adverbs to have non-identity types, i.e. maps from type X to type Y, he no longer predicts linear ordering of adverbs. However, this move jeopardizes the intuitive “plausibility” his approach might have. For example, the Norwegian non-transitive triplet of adverbs *muligens* (‘possibly’), *ikke* (‘not’), *alltid* (‘always’) can be accommodated in the FEO-calculus if (and only if; see Nilsen (2001)) *muligens* takes a fact and returns a proposition, *ikke* takes an event and returns a fact, and *alltid* takes a proposition and returns an event. See Chapter 1 and Nilsen (2001) for discussion of this point.

I have assumed that adverbs like *always* are temporal and that, say *possibly* applies to propositions, but the fact that the two can occur in either order, with different scopal effects, suggests that this should not be brought to bear on the unavailability of certain adverb orderings. In fact, that would be a surprising claim, which would require an explanation. For example, the fact that *everybody* quantifies over individuals does not prevent it from outscoping (or being outscoped by) modal adverbs or verbs, negation and other non-individual operators. It would be surprising, then, if it were to turn out that the fact that some adverbs operate on times or events should prevent them from behaving similarly. Ernst could maybe maintain his point of view by aluding to the empirical fact that nominal quantifiers tend to give rise to inverse scope effects, whereas adverbs do not seem to do so. However, it is not clear that these phenomena are the same. In other words, it is not clear that ability to generate inverse (non-surface) is what enables nominal quantifiers to relate to their variables at a distance. However, if we assume that VPs denote events, and that some adverbs force conversion to “higher” ontological types in Ernst’s sense, it just seems to be empirically wrong to say that event modifiers must be sisters of event denoting expressions. For example, if *always* is an event modifier, the fact that it can precede and outscope SOA, which we have seen it can, seems to be a counterexample.

2.5.3 Summary

I have argued that SOA should be treated as PPIs, in the sense of being excluded from DE environments and that this suffices to derive their syntactic distribution. Positive phase quantifiers are also PPIs, but only excluded from AA environments. We have seen that the status of the adverb *possibly* as a PPI follows from its lexical semantics if we derive its meaning from the meaning of that of *possible* by means of an operation Ω on its epistemic accessibility relation \mathcal{E} , and making Ω subject to an anti-weakening constraint, sensitive to the scalar implicatures of the modified proposition. We also saw that intervention effects follow in a manner entirely parallel to Chierchia's (2001) derivation of such effects with NPIs. The analysis appears to maintain the advantages that the analysis of adverb ordering proposed by Ernst (2001); Svenonius (2001), while sidestepping the difficulties associated with their ontological approach. In appendix A, some further evidence for the analysis of SOA presented here is adduced from the fact that SOA in many languages are incompatible with degree modification.

CHAPTER 3

V2 and Holmberg's Generalization

3.1 Introduction

The standard¹ (symmetric) analysis of verb second (V2) (Holmberg and Platzack, 1995; den Besten, 1989) is that the finite verb (Vf) head-moves to the highest functional projection of the clause. Then some other constituent, for instance the subject or an adverbial, has to move to the specifier of that same projection. These two movement steps, in addition to a general ban on adjunction to CP, ensure that Vf will always end up in the second position of the clause.

It is the purpose of the present chapter to argue against a head movement analysis of V2. The main argument for an XP-movement analysis will come the fact that certain (apparent) V2-violations in Mainland Scandinavian seem to pose severe problems for a head movement analysis. The problematic data involve focus particles that can intervene between the finite verb and the first constituent. It will be argued that these cannot be treated as a clitic on the verb and that the V2 violations are real. The interaction between V2 violations with focus particles and argument shift of weak pronouns will be used to show that the verb does not move to second position as a head. From this conclusion it follows that weak pronouns, in fact, do not shift. When they appear to have moved, it is a larger constituent containing the VP that has moved.

¹Sections 3.2–3.5 will appear as Nilsen (to app.b). The remaining sections, including the “third approximation” were developed after submission of that paper.

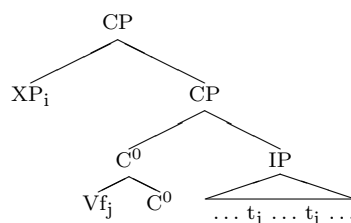


Figure 3.1: standard analysis of V2

Thus, pronoun shift and V2 is treated as surface reflexes of one and the same operation. This gives a very simple explanation for Holmberg's Generalization to the effect that object shift cannot cross phonetically realized material from the VP: it cannot do so because it is the VP itself, or rather an XP containing it, that moves.

The traditional view of V2 is also challenged by the fact that there are topicalization-like processes and wh-movement effects that seem to require a sentence internal landing site. Furthermore, some facts concerning subject-verb inversion are problematic for the standard treatment of V2. Inversion is usually analyzed as V-to-C movement, with the subject in spec-IP or equivalent position. This leads to the expectation that it can only occur when the verb really is in C. However, there are cases in which the verb is arguably much lower than this, and the subject still has to follow it.

The proposed account builds on recent work by Kayne (1998, 1999); Cinque (1999); Rizzi (1997); Koopman and Szabolcsi (2000) and others. In particular, no covert movement is used, all movements are to the left and the analysis relies heavily on the use of 'remnant' movement, i.e. movement of a constituent containing a trace. More specifically, the proposal is that Rizzi's (ibid.) functional projections FocP, TopP and FinP are merged below sentential adverbs. V2 consists in successive raising of TopP around sentential adverbs, carrying the verb-initial FinP along. One of the key features of the analysis is thus that it renders V2 sensitive to the properties of individual classes of adverbs. Finally, it is argued that it may be possible and advantageous to treat V2 phenomena without reference to functional projections, such as FocP, TopP, FinP, etc.

The chapter is organized as follows. In section 2, the basic data concerning the V2 violating focus particles and their interaction with pronoun shift is presented. Section 3 presents a 'first approximation' to an analysis that can handle the facts. Section 4 suggests that the problems encountered with the first analysis is that fronting operations and S-V inversion should be able to access quite low positions. In section 5 the main proposal is developed. Section 6 sums up and concludes.

3.2 Some data

3.2.1 V2-violations with focus particles

As has been discussed by Egerland (1998), there exist certain apparent exceptions to the V2-generalization in Mainland Scandinavian involving so-called 'focus-sensitive adverbs' or 'focus particles' (henceforth fpt). The phenomenon is illustrated in (119) below with data from Norwegian.

- (119) a. Jens bare gikk.
J just left
b. Jens nesten gråt.
J almost cried

It is also possible to have the fpt after Vf, as in (120). Neither of the two orders appear to be marked or degraded in any way.

- (120) a. Jens gikk bare.
J left just
b. Jens gråt nesten.
J cried almost

Other expressions that exhibit the same behavior include *til og med* 'even' (lit. 'to and with'), *minst* 'at least' *utelukkende* 'exclusively' *ikke mer enn såvidt* 'not more than barely', *simpelthen* 'simply'. Thus we are not dealing with a quirk of a couple of words. Below are examples.

- (121) a. Han til og med leste den.
he even read it
b. han utelukkende sover hele dagen
he exclusively sleeps whole the-day
c. Han ikke mer enn såvidt berørte den.
he not more than barely touched it
d. Han simpelthen tok den.
he simply took it

Egerland (ibid.) notes that with *nesten* there is a truth-conditional difference corresponding to its different positions. Consider the following examples.

- (122) a. Jens nesten brølte hurra.
J almost roared hooray
b. Jens brølte nesten hurra.
J roared almost hooray

(122a) can only mean that Jens pronounced the word "hurra" in a manner that almost qualifies as roaring it. Let us refer to this as the 'manner' reading. The most salient reading of (122b) is that he didn't cry "hurra", although he was about to, i.e. a 'modal' reading. It can also get the manner meaning if pronounced with heavy stress on the verb. See Rapp and Von Stechow (1996) for discussion of these and other readings of the German adverb *fast* ('almost'). This pattern can also be taken to indicate that the fpt is not adjoined to C', since the manner reading presumably results from attaching the adverb lower, not higher, than the site responsible for the modal reading.

Egerland (ibid.) analyzes this phenomenon in terms of the Universal Hierarchy of functional projections proposed in Cinque (1999). The cases involving *nesten* and those involving *bare* are given different analyses. Egerland maintains a standard analysis of V2 in terms of head movement to the highest FP in the clause (ForceP in the 'split CP' framework of Rizzi (1997)). He analyses the adverb *nesten* ('almost') as a specifier of a modal projection in the IP-layer. For (122a), he claims that Vf can remain in or below that modal head when *nesten* is in its specifier. The adverb *bare* 'only,' 'just', he treats as a syntactic clitic on Vf. The analysis of *bare* as a clitic is supported by two facts. The first is that the adverb can be phonetically reduced into the monosyllabic *ba* in Swedish (not possible in Norwegian).

- (123) Per *ba'* gick. [Swe]
P just left

The second argument is that, according to Egerland, *bare* cannot appear in front of auxiliaries.

- (124) *Per *bara/ba'* har gått. [Swe]
P just has left

The same applies to Norwegian as long as the example is read with neutral intonation, but if the auxiliary is stressed (125), the result is much better; also if a less semantically impoverished auxiliary is used (126).

- (125) Jens *bare* HAR gått.
J just has left
"It just IS the case that Jens has left."

- (126) Jens *bare* måtte gå.
J just must-past leave
"Jens just HAD TO leave."

There are other problems with assuming that *bare* is a clitic. First, it can be modified. If one does not take *simpelthen* in (127) to directly modify *bare*, the example would present the same kind of problem as (128) below.

- (127) Jens *simpelthen* *bare* gikk.
J simply just left

After Kayne (1975), one of the defining characteristics of syntactic clitics has been taken to be that they cannot be modified. Secondly, other adverbs that cannot plausibly be taken to modify *bare* can also precede Vf when *bare* does; in fact, only when *bare* does.

- (128) a. Jens vanligvis bare svarer ikke.
 J usually just answers not
 b. *Jens vanligvis svarer ikke.
 J usually answers not

This points to the conclusion that *bare* should be treated on a par with *nesten* ‘almost’, so that, when these adverbs are present, Vf can remain in a low position in the IP-field. If the position of *bare* is lower than that of *vanligvis*, we can also explain why the latter adverb actually has to precede Vf when *bare* does. Compare (129) to (128a):

- (129) *Jens bare svarer vanligvis ikke.
 J just answers usually not

Since the negation must follow Vf in the relevant construction, it cannot be regarded as V-in-situ, either. This is illustrated in (130).

- (130) a. Jens bare liker ikke fiskekaker.
 J just likes not fishcakes
 b. *Jens bare ikke liker fiskekaker.
 J just not likes fishcakes

So far, we can conclude that V2-violating *bare* is not a clitic on Vf; that *bare* occupies a position lower than *vanligvis* ‘usually’; and that Vf can remain below that position when *bare* is present, although it cannot remain in situ.

3.2.2 Pronoun Shift

In the next few paragraphs, we will see that there are reasons to think that V2 does not involve head movement of Vf, but rather movement of a phrasal category. A corollary of our observations will be that weak pronouns, in fact, do not shift. In the discussion of weak pronouns, we use the phonetically reduced forms ‘*n*’ ‘he/him/it’ and ‘*a*’ ‘she/her/%it’² since these are unambiguously weak: they must shift to the left. Consider the pattern below.

²Some dialects use ‘*a*’ to replace all feminine nouns, including inanimates. These dialects typically use the “personal” pronouns for inanimates, also in their unreduced forms, like in (i), although, if the pronoun in (i) is prosodically stressed, it must refer to a person Cardinaletti and Starke (1995).

(i) Hu ligger på bordet.

She lies on the-table

‘It [the book] is lying on the table’

Other dialects (including my own) use ‘*n*’ for inanimates, regardless of gender.

- (131) a. Derfor vanligvis bare svarte 'n 'a ikke.
therefore usually just answered he her not
- b. Derfor svarte 'n 'a vanligvis bare ikke.
therefore answered he her usually just not
- c. *Derfor 'n 'a vanligvis bare svarte ikke.
therefore he her usually just answered not
- d. *Derfor svarte vanligvis bare 'n 'a ikke.
therefore answered usually just he her not

In (131) we see that when the subject and the object are both realized as weak pronouns, they must remain immediately right-adjacent to the verb. In (131a), the verb+arguments complex remains below the position of *bare*, whereas in (131b), the entire complex is moved around *bare* and *vanligvis* to the second position. (131c)-(131d) are added to show that the complex cannot be split up. This seems to indicate that it is moving as a constituent. The alternative would be to say that the verb and the arguments move separately, but that the verb somehow blocks further movement of the pronouns (cf. the Shortest Move/Minimal Link Condition Chomsky (1995)). One would need an extra landing site for the pronouns, higher than the negation, but lower than the other adverbs. The pronouns would move as high as they could without crossing the verb in the overt syntax, and then proceed to the higher position(s) covertly. The features attracting the pronouns to the higher positions would have to be 'optionally strong.' It is obviously simpler to say that the Vf and the pronouns are moving as a constituent. That analysis obviates the need for extra landing sites and optionally strong features.

Taking V2 to be derived by XP-movement, we arrive at the following result:

Generalization 1 *Weak pronouns do not shift.*

When a weak pronoun appears to have moved across adverbs, it is something else containing the pronouns that has moved. Consider now the following pattern where the subject and the object appear to be moving as a constituent without Vf.

- (132) a. Derfor svarte Jens 'a vanligvis ikke.
therefore answered J her usually not
- b. Derfor svarte vanligvis ikke Jens 'a.
therefore answered usually not J her
- c. *Derfor svarte Jens vanligvis ikke 'a
therefore answered J usually not her
- d. *Derfor svarte 'a vanligvis ikke Jens.
therefore answered her usually not J

The subject and the weak, pronominal object can follow the adverbs as long as they remain adjacent. In Swedish, (132c),(132d) are also possible. The two

arguments need not remain adjacent in that language. In Danish, only (132a) is grammatical. For Norwegian, then, the subject and the object seem to make up a constituent in these examples.

Assuming that pronoun shift actually does not exist squares well with one of the fundamental properties that pronoun shift appears to have: No effect. It has been noted in the literature that pronoun shift does not create new binding possibilities, it does not license parasitic gaps, it does not block *wh*-movement (relativized minimality), it does not interfere with passivization (relativized minimality); in short, it has no effect at all. This obviously supports generalization 1.³

3.3 First approximation

A simple way to derive these constituents, which will be shown to be inadequate shortly, would be to assume the following. There is an XP dominating VP into which Vf always moves. This XP, in turn, moves to spec-Fin prior to fronting of some constituent to spec-Top. The derivations would go as in derivation 3.1 and 3.2, ignoring the base position of the adverbial *derfor* ('therefore'). Phonetic material is in boldface and cyclicity is ignored in this derivation. This gives us

Derivation 3.1

$$\begin{array}{l}
 [\text{TopP Top } [\text{FinP Fin } [\text{IP not } [\text{XP X } [\text{VP he answered her }]]]]] \\
 \hspace{15em} \text{moves to X } \blacktriangleright \\
 [\text{TopP Top } [\text{FinP Fin } [\text{IP not } [\text{XP answered}_i + \text{X } [\text{VP he } t_i \text{ her }]]]]] \\
 \hspace{10em} \text{XP moves to spec-Fin and some constituent topicalizes } \blacktriangleright \\
 [\text{TopP therefore Top } [\text{FinP } [\text{XP answered}_i + \text{X } [\text{VP he } t_i \text{ her }]]] \text{ Fin } [\text{IP} \\
 \text{not } t_{\text{XP}}]]]
 \end{array}$$

the verb+arguments constituent we demonstrated in (14) above. If the subject bears focus, the remnant VP is extracted into the IP-field prior to movement of XP to spec-Fin.

Thus the constituent made up by the subject and the object in (15) is the remnant VP. If *bare* is merged into the IP, it attracts XP (cf. Kayne 1998). In this case, either XP or IP moves further up to spec-Fin.

3.3.1 Problems

One of the attractive features of an analysis along these lines is that, *prima facie*, it seems to explain Holmberg's Generalization.

³A similar point will be made below for movement of Vf to "second position".

Derivation 3.2

[_{TopP} Top [_{FinP} Fin [_{IP} not [_{XP} answered_i+X [_{VP} Jens t_i her]]]]
 VP scrambles into IP ►
 [_{TopP} Top [_{FinP} Fin [_{IP} not [_{VP} Jens t_i her] [_{XP} answered_i+X t_{VP}]]]
 XP moves to spec-Fin and some constituent topicalizes ►
 [_{TopP} therefore Top [_{FinP} [_{XP} answered_i+X t_{VP}] Fin [_{IP} not [_{VP} Jens
 t_i her] t_{XP}]]

Generalization 2 (Holmberg (1986, 1999) (HG)) *Argument shift cannot cross any phonetically realized material from the VP, i.e. the verb, a verbal particle, a dative preposition, or other arguments of the verb, although it can cross traces of these, as well as sentential adverbs.*

Employing a non-movement analysis of argument shift, HG seems to follow trivially. Weak pronouns cannot cross any phonetically realized material from the VP, simply because it is the VP itself (or something bigger) that moves. Unfortunately, the explanation offered of HG from this account breaks down when one looks at the system more closely. This is because of what would happen in non-V2 contexts. Vf would presumably move to X in these contexts as well, and then nothing prevents the remnant VP to scramble into IP, yielding ungrammatical orders such as the following.

- (133) a. *... at Jens 'a ikke svarer
 ... that J her not answers
 b. *... at Jens ikke fiskekaker liker.
 ... that J not fishcakes likes

In order to prevent such orders, we would have to make extraction of VPs or objects from XP somehow contingent upon subsequent raising of XP to specFin. That is, we would have to reintroduce a notion of HG which is what we set out to derive.

Similarly, one would need to account for the unavailability of subject-verb inversion in non-V2 contexts with weak pronominal subjects. Suppose that the finite subordinator at ‘that’ is generated in Fin and somehow blocks XP-to-spec-Fin movement as well as topicalization. As it stands, our account would lead to the incorrect expectation that the following order should be grammatical with the structure in (134b).

- (134) a. *... at ikke svarte 'n 'a
 ... that not answered he her
 b. [TopP Top [FinP **that** Fin [IP **not** [XP **answered**_i+X [VP **he** t_i
 her]]]]]

Another problem is that it is not clear that the account explains root-V2. The extent to which it succeeds in doing so depends on whether TopP is the only projection dominating FinP. In e.g. Rizzi (1997), which is where the names TopP, FinP originate from, two other FPs are postulated, both of which dominate FinP, i.e. FocP and ForceP. Thus, either we would have to show that these projections do not exist, or that, for independent reasons, they cannot be filled in the relevant cases. These considerations are taken to show that a more radical departure from standard assumptions is needed.

3.4 More data: how initial is the initial position?

Before we turn to our second approximation to the proper analysis of V2, some data will be reviewed that purport to show that the initial position, i.e. spec-CP in traditional analyses, need not be construed as a base-generated initial position. The evidence we will review suggests that operations like wh-movement and topicalization are done in (at least) two separate movement steps, one targeting an IP-internal position, and a second one whose nature we try to elucidate in the remainder of the chapter. Consider the following contrasts:

- (135) a. Al very probably won.
 b. *How probably did Al win?
 c. How probable is it that Al won?
- (136) a. Al quite possibly won.
 b. *How possibly did Al win?
 c. How possible is it that Al won?
- (137) *How [probably/ possibly/ fortunately/ necessarily/ evidently/ maybe/
 frankly/ usually] did Al win?

It seems that wh-movement of higher adverbs (cf. Cinque 1999) is systematically impossible. This contrasts with the behavior of lower adverbs, which do allow this kind of movement:

- (138) How [quickly/ effortlessly/ often/ soon/ frequently] did Al win?

Given that degree modification of the higher adverbs is possible (cf. (135a, 136a)), and that the adjectival counterparts of, e.g. (138) (i.e. "how fortunate/usual/... is it that...") are grammatical, it seems that we are dealing with a syntactic phenomenon, rather than a (purely) semantic one. This suggests that wh-movement is a composite operation, consisting of two parts: One movement step targeting a relatively low, "IP-internal" position, call it P1 and

another step targeting the first position of the clause, (P2). If P1 is generated lower than adverbs like *probably*, the ungrammaticality of examples like (138) would follow from the Extension Condition (Chomsky (1995)). In other words, the only way for them to obey the Extension Condition would be to merge with P1, the *wh*-position directly, and then raise to their base-position.

There are some indications that the same reasoning applies to topicalization. That is, there are topicalization-like processes that seem to target an IP-internal position in embedded clauses only. One case in point is so-called 'stylistic fronting' in Icelandic (cf. a.o. Holmberg (2000)).

Another embedded, "IP-internal" topicalization process is illustrated by the following Norwegian examples. These may seem slightly contrived, but the contrast between (139b) and the other two is quite sharp. In (139a), we see an adverbial modifier (*bak låven*), modifying the most deeply embedded predicate, but appearing displaced from it, in the mittelfeld of the next clause up. In (139b) we see that this kind of displacement is unavailable if the target is the root clause. In this case, the displaced adverbial will appear in the very first position, as illustrated in (139c).

- (139) a. Det jeg sa var at jeg bak låven aldri har
 what I said was that I behind the-barn never have
 skjønt hvorfor han plantet tulipaner.
 understood why he planted tulips
- b. *Jeg har bak låven aldri skjønt hvorfor han
 I have behind the-barn never understood why he
 plantet tulipaner.
 planted tulips
- c. Bak låven har jeg aldri skjønt hvorfor han
 behind the-barn have I never understood why he
 plantet tulipaner.
 planted tulips

It is tempting to say that *bak låven* is, in some sense in the "same position" in (139a) and (139c). Yet another indication that something is wrong with the standard view of the left periphery is what appears to be obligatory subject-verb inversion in a quite low position. Consider the following Norwegian examples where the object has been topicalized around a V2-violating fpt.

- (140) a. Meg vanligvis bare svarte ikke Jens.
 me usually just answered not J
- b. Meg vanligvis bare svarte Jens ikke.
 me usually just answered J not
- c. *Meg vanligvis bare Jens svarte ikke.
 me usually just J answered not

- d. *Meg vanligvis Jens bare svarte ikke.
 me usually J just answered not
- e. *Meg Jens vanligvis bare svarte ikke.
 me J usually just answered not

Given the discussion in section 2, it appears that Vf cannot be in C. One could try to say that, in (140a) the subject is inside VP. This would not work for (140b), since here, the subject precedes the negation as well. Thus, the standard 'V-to-C' analysis of subject-verb inversion cannot handle this phenomenon. The same point could be made with 'distributive' conjunctions like (141) (cf. Zamparelli 2000).

- (141) a. Meg vanligvis både slo de og sparket.
 me usually both beat they and kicked
 'They usually both beat and kicked me.'
- b. *Meg vanligvis både de slo og sparket.
 me usually both they beat and kicked
- c. *Meg vanligvis de både slo og sparket.
 me usually they both beat and kicked
- d. *Meg de vanligvis både slo og sparket.
 me they usually both beat and kicked

For reasons of space, we will not enter into a discussion of distributive conjunction here. We refer the reader to Zamparelli's work for an analysis that is congenial to the analysis to be presented here.

3.5 Second approximation

Taking our conclusions so far quite literally, we arrive at the following picture of 'clausal architecture':

- (142) [advP adv* [WP [FocP [TopP [FinP VP]]]]]]

Here, the heads Foc, Top, Fin are the ones argued for by Rizzi (1997). W is the head introduced by Kayne (1998) to deal with scope phenomena involving the fpt only. We do not invoke Rizzi's (ibid.) head 'Force', mainly because it is not necessary for our purposes. Adverbs attract TopP to their specifiers. We assume that, in case nothing is focus, or the entire sentence is, TopP is attracted to spec-Foc, and subsequently to spec-W. Thus in these cases, the sequence W, Foc will simply be omitted. Let us see how the system works by going through some derivations.

3.5.1 Root clauses

We begin by deriving simple root clauses like the following.

- (143) Jens svarte meg vanligvis.
J answered me usually

Derivation 3.3

[_{VP} **John answered me**]
MERGE Fin and MOVE V ►
[_{FinP} **answered** [_{VP} **John me**]]
MERGE Top and MOVE *John* ►
[_{TopP} **John** [_{FinP} **answered** [_{VP} **me**]]]
MERGE *usually* and MOVE TopP ►
[_{AdvP} [_{TopP} **John** [_{FinP} **answered** [_{VP} **me**]]] **usually**]

In this derivation, the entire sequence John answered me climbs around the adverb. The only movement step which is always necessary is movement of Vf to Fin. The object could also have moved to spec-Top. In that case, the subject would either remain in-situ, or move to spec-Foc.

Consider now a sentence with an indefinite object noun phrase. In these, the indefinite will move to spec-Foc, prior to movement of TopP to spec-W.

Derivation 3.4

[_{VP} **John read a book**]
MERGE Fin and MOVE V ►
[_{FinP} **read** [_{VP} **John a book**]]
MERGE Top and MOVE *John* ►
[_{TopP} **John** [_{FinP} **read** [_{VP} **a book**]]]
MERGE Foc and MOVE *a book* ►
[_{FocP} **a book** [_{TopP} **John** [_{FinP} **read** [_{VP}]]]]
MERGE W and MOVE TopP ►
[_{WP} [_{TopP} **John** [_{FinP} **read**]]] [_{FocP} **a book**]]
MERGE *usually* and MOVE TopP ►
[_{AdvP} [_{TopP} **John** [_{FinP} **read**]]] **usually** [_{WP} [_{FocP} **a book**]]]

- It is easy to see that addition of more adverbs would only lead to iteration of this last step, so V2 is derived for these cases. If the subject bears focus, we now expect the ungrammatical (145a), as shown in derivation 3.5.

- ### Derivation 3.5

[AdvP [TopP **therefore** [FinP **recognized** [VP **me**]]] **usually** [WP [FocP **a student**]]

In Danish subjects must precede all adverbs. Objects follow them unless they are weak pronouns, in which case they must also precede them. I would like to suggest that in this language, spec-Foc has been 'grammaticalized' as a case position for objects. Weak pronouns, being marked for case, do not have to move there. This means that there are but two positions the subject can choose between: it can move to spec-Top, or it can remain in-situ. In the first case, it will end up in the V2-initial position, as what happens to *therefore* in

Derivation 3.6

$[_{VP} \text{ a student recognized me}]$
 MERGE Fin and MOVE V ►
 $[_{FinP} \text{ recognized } [_{VP} \text{ a student me}]]$
 MERGE Top and *therefore* ►
 $[_{TopP} \text{ therefore } [_{FinP} \text{ recognized } [_{VP} \text{ a student me}]]]$
 MERGE Foc and MOVE VP ►
 $[_{FocP} [_{VP} \text{ a student me}] [_{TopP} \text{ therefore } [_{FinP} \text{ recognized}]]]$
 MERGE W and MOVE TopP ►
 $[_{WP} [_{TopP} \text{ therefore } [_{FinP} \text{ recognized}]] [_{FocP} [_{VP} \text{ a student me}]]]$
 MERGE *usually* and MOVE TopP ►
 $[_{AdvP} [_{TopP} \text{ therefore } [_{FinP} \text{ recognized}]] \text{ usually } [_{WP} [_{FocP} [_{VP} \text{ a student me}]]]]]$

derivation 3.6. In the latter, it will remain immediately right-adjacent to the verb, much as the weak object pronoun in derivation 3.3. Thus we have three levels of freedom with regard to subjects and spec-Foc. In Swedish, the subject can move there alone; in Norwegian, it must pied-pipe the VP, and in Danish, it cannot move there at all.

bare

Let us now turn to the problematic cases with V2-violating focus particles. In general, we expect these to appear in the Adv* area of (142). I suggest that they differ from other adverbs in that, instead of attracting TopP, they attract their focus associate, and come with a W-projection into which *bare* itself and TopP moves, i.e. we essentially adopt the treatment in Kayne (1998). Suppose that we add *bare* to the end result of derivation 3.6. Then, if *en student* is the associate, we get derivation 3.7, corresponding to the grammatical sentence in (146).

- (146) Derfor gjenkjente *bare* EN student meg.
 therefore recognized just one student me

Adding *usually*, we get the continuation in derivation 3.8, corresponding, again, to the grammatical sentence (147).

- (147) Derfor gjenkjente vanligvis *bare* EN student meg.
 therefore recognized usually only one student me

Derivation 3.7

$[_{VP} \text{ one student recognized me}]$
 MERGE *Fin* and MOVE *V* ►
 $[_{FinP} \text{ recognized } [_{VP} \text{ one student me}]]$
 MERGE *Top* and *therefore* ►
 $[_{TopP} \text{ therefore } [_{FinP} \text{ recognized } [_{VP} \text{ one student me}]]]$
 MERGE *Foc* and MOVE *VP* ►
 $[_{FocP} [_{VP} \text{ one student me}] [_{TopP} \text{ therefore } [_{FinP} \text{ recognized}]]]$
 MERGE *W* and MOVE *TopP* ►
 $[_{WP} [_{TopP} \text{ therefore } [_{FinP} \text{ recognized}]] [_{FocP} [_{VP} \text{ one student me}]]]$
 MERGE *just* and MOVE *FocP* ►
 $[_{bareP} [_{FocP} [_{VP} \text{ one student me}]] \text{ just } [_{WP} [_{TopP} \text{ therefore } [_{FinP} \text{ recognized}]]]]]$
 MERGE *W* and MOVE *just* and *TopP* ►
 $[_{WP} [_{TopP} \text{ therefore } [_{FinP} \text{ recognized}]] \text{ just } [_{bareP} [_{FocP} [_{VP} \text{ one student me}]]] [_{WP}]]]]$

Derivation 3.8

MERGE *usually* and MOVE *TopP* ►
 $[_{AdvP} [_{TopP} \text{ therefore } [_{FinP} \text{ recognized}]] \text{ usually } [_{WP} \text{ only } [_{bareP} [_{FocP} [_{VP} \text{ one student me}]]]]]]]$

Consider now derivation 3.9 of (128a) above, repeated here, in which the verb is the associate. In this case, *FinP* moves to spec-bare.

- (148) Jens vanligvis bare svarer ikke.
 J usually just answers not

The essential difference between these examples and 'ordinary' V2 sentences is thus that the verb has been pulled out of *TopP*. Therefore, when *TopP* moves around higher adverbs, the verb is left behind. Of course, the last step of derivation 3.9) can be skipped, yielding the equally grammatical sentence *Jens bare svarer ikke* ('J just answers not').

Since weak pronouns do not move, they will be left inside the *VP*. This is what causes them to stay adjacent to the verb. The partial derivation 3.10 for (149) (=131a), illustrates this.

Derivation 3.9

$[_{VP} \text{ John answers}]$
 MERGE Fin and MOVE V ►
 $[_{FinP} \text{ answers } [_{VP} \text{ John}]]$
 MERGE Top and MOVE *John* ►
 $[_{TopP} \text{ John } [_{FinP} \text{ answers } [_{VP}]]]$
 MERGE *not* and MOVE TopP ►
 $[_{AdvP} [_{TopP} \text{ John } [_{FinP} \text{ answers}]] \text{ not }]$
 MERGE just and MOVE FinP ►
 $[_{bareP} [_{FinP} \text{ answers}] \text{ just } [_{AdvP} [_{TopP} \text{ John}] \text{ not }]]$
 MERGE W and MOVE just and TopP ►
 $[_{WP} [_{TopP} \text{ John}] \text{ just } [_{bareP} [_{FinP} \text{ answers}] \text{ [AdvP not }]]]]$
 MERGE usually and MOVE TopP ►
 $[_{AdvP} [_{TopP} \text{ John}] \text{ usually } [_{WP} \text{ just } [_{bareP} [_{FinP} \text{ answers}] \text{ [AdvP not }]]]]]]$

Derivation 3.10

$[_{AdvP} [_{TopP} \text{ therefore } [_{FinP} \text{ answered } [_{VP} \text{ he her}]]] \text{ not }]$
 MERGE *just* and MOVE FinP ►
 $[_{bareP} [_{FinP} \text{ answered } [_{VP} \text{ he her}]] \text{ just } [_{AdvP} [_{TopP} \text{ therefore}] \text{ not }]]$
 MERGE W and MOVE just and TopP ►
 $[_{WP} [_{TopP} \text{ therefore}] \text{ just } [_{bareP} [\text{FinP answered } [_{VP}_{he} \text{ her}]] \text{ [AdvP not }]]]]$
 MERGE usually and MOVE TopP ►
 $[_{AdvP} [_{TopP} \text{ therefore}] \text{ usually } [_{WP} \text{ just } [_{bareP} [_{FinP} \text{ answered } [_{VP} \text{ he her}]] \text{ [AdvP not }]]]]]]$

- (149) Derfor vanligvis bare svarte 'n 'a ikke.
 therefore usually just answered he her not

Nothing special needs to be said about the low subject-verb inversions noted in section 4. There are two cases to consider. If the subject is a weak pronoun, it remains in the VP, and MOVES with the verb, as in the previous derivation. If it is a full noun phrase, as in (150) (=140a) it can move to spec-Foc, giving rise to the derivation 3.11.

- (150) Meg vanligvis bare svarte ikke Jens.
 me usually just answered not J

Derivation 3.11

[_{VP} John answered me]
 MERGE Fin and MOVE V ►
 [_{FinP} answered [_{VP} John me]]
 MERGE Top and MOVE *me* ►
 [_{TopP} me [_{FinP} answered [_{VP} John]]]
 MERGE Foc and MOVE VP ►
 [_{FocP} [_{VP} John] [_{TopP} me [_{FinP} answered]]]
 MERGE W and MOVE TopP ►
 [_{WP} [_{TopP} me [_{FinP} answered]] [_{FocP} John]]
 MERGE *not* and MOVE TopP ►
 [_{AdvP} [_{TopP} me [_{FinP} answered]] not [_{WP} [_{FocP} John]]]
 MERGE *just* and MOVE FinP ►
 [_{bareP} [_{FinP} answered] just [_{AdvP} [_{TopP} me] not [_{WP} [_{FocP} John]]]
 MERGE W and MOVE *just* and TopP ►
 [_{WP} [_{TopP} me] just [_{bareP} [_{FinP} answered] [_{AdvP} not [_{WP} [_{FocP} John]]]]]
 MERGE *usually* and MOVE TopP ►
 [_{AdvP} me usually [_{WP} just [_{bareP} [_{FinP} answered] [_{AdvP} not [_{WP} [_{FocP} John]]]]]]]

As the reader may have noticed, nothing has been said, so far, about examples like (131b), repeated in (151), where Vf does appear in the second position even though there is an fpt modifying it.

- (151) Derfor svarte 'n 'a vanligvis bare ikke.
 therefore answered he her usually just not

In order to handle these cases, we propose that FinP, when it moves to *spec-bare*, can pied-pipe TopP. This has the effect of making *bare* behave as other adverbs with respect to V2. This is illustrated in the partial derivation 3.12.

Auxiliaries and the derivation of HG

There is a large literature on the syntactic treatment of auxiliaries (cf. Cinque 2000, Koopman and Szabolcsi 2000, Julien 2000 for recent discussion). One

Derivation 3.12

$[_{\text{AdvP}} [_{\text{TopP}} \text{therefore} [_{\text{FinP}} \text{answered} [_{\text{VP}} \text{he her}]]] \text{not}]$
 MERGE *just* and MOVE TopP ►
 $[_{\text{bareP}} [_{\text{TopP}} \text{therefore} [_{\text{FinP}} \text{answered} [_{\text{VP}} \text{he her}]]] \text{just} [_{\text{AdvP}} \text{not}]]$
 MERGE W and MOVE *just* and TopP ►
 $[_{\text{WP}} [_{\text{TopP}} \text{therefore} [_{\text{FinP}} \text{answered} [_{\text{VP}} \text{he her}]]] \text{just} [_{\text{bareP}} [_{\text{AdvP}} \text{not}]]]$
 MERGE *usually* and MOVE TopP ►
 $[_{\text{AdvP}} [_{\text{TopP}} \text{therefore} [_{\text{FinP}} \text{answered} [_{\text{VP}} \text{he her}]]] \text{usually} [_{\text{WP}} \text{just} [_{\text{bareP}} [_{\text{AdvP}} \text{not}]]]]$

controversy is whether or not to treat e.g. participial constructions as biclausal (Kayne 1993). We remain agnostic about this question here, simply analyzing auxiliaries by stacking them onto the VP. We do treat them as 'raising verbs' in the sense that they attract the subject from the inner VP. This is in order to exclude sentences like (49) where the subject follows an auxiliary.

- (152) *Derfor kan ikke ha 'n sett Jens.
 therefore can not have he seen J

We assume that participial VPs move to spec-Foc unless the participial verb is a topic. Apart from this, auxiliaries do not pose any special problems for our account. We need to discuss them, however, in order to show that the account really derives HG which we repeat here for convenience.

- (153) Holmberg's Generalization (HG)
 Argument shift cannot cross any phonetically realized material from the VP, i.e. the verb, a verbal particle, a dative preposition, or other arguments of the verb, although it can cross traces of these, as well as sentential adverbs.

We have said that weak pronouns never move. What is crucial is that they never shift. Otherwise, they are free to move, as it were. In particular, they can move to spec-Top, which ultimately places them in the initial position of the clause. They cannot move to spec-Foc for the simple reason that they are weak, and, in Danish, because they bear morphological case. Consider now the contrast between (154) and (155), a typical example of HG.

- (154) a. Jeg så 'n ikke.
 I saw him not
 b. *I så ikke 'n.
 I saw not him

- (155) a. *Jeg har 'n ikke sett.
 I have him not seen
 b. Jeg har ikke sett 'n.
 I have not seen him

We have already seen how it comes about that the weak pronoun must precede the negation in (154a). We need to demonstrate that it cannot do so if there is an auxiliary. The first steps in the derivation of (155b) are as follows.

Derivation 3.13

[_{VP} I have [_{PtcP} seen him]]

MERGE Fin and MOVE V ►

[_{FinP} have [_{VP} I [_{PtcP} seen him]]]

At this point, one of two things can happen: either the subject moves to spec-Top, or the object does. As we shall see below, the participial verb can also do this. If the subject moves, we get the following continuation, deriving (155b).

Derivation 3.14

MERGE Top and MOVE *John*

[_{TopP} I [_{FinP} have [_{VP} [_{PtcP} seen him]]]]

MERGE Foc and MOVE PtcP ►

[_{FocP} [_{PtcP} seen him] [_{TopP} I [_{FinP} have [_{VP}]]]]

MERGE W and MOVE TopP ►

[_{WP} [_{TopP} I [_{FinP} have]] [_{FocP} [_{PtcP} seen him]]]

MERGE *not* and MOVE TopP ►

[_{AdvP} [_{TopP} I [_{FinP} have]] not [_{WP} [_{FocP} [_{PtcP} seen him]]]

In the alternative case, the object moves to spec-Top, yielding the following continuation of derivation 3.13, corresponding to the grammatical sentence in (156).

- (156) Han har jeg ikke sett.
 him have I not seen

If the participial verb is a topic, it moves to spec-Top on its own. This implies that heads can move to specifier positions, contra standard assumptions. It is possible that this problem would disappear if we adopt a biclausal structure for participial constructions, (Kayne 1993). Another alternative would be to

Derivation 3.15

MERGE Top and MOVE *him*

[_{TopP} **him** [_{FinP} **have** [_{VP} **I** [_{PtcP} **seen**]]]]

MERGE Foc and MOVE PtcP ►

[_{FocP} [_{PtcP} **seen**] [_{TopP} **him** [_{FinP} **have** [_{VP} **I**]]]]

MERGE W and MOVE TopP ►

[_{WP} [_{TopP} **him** [_{FinP} **have** [_{VP} **I**]]] [_{FocP} [_{PtcP} **seen**]]]

MERGE *not* and MOVE TopP ►

[_{AdvP} [_{TopP} **him** [_{FinP} **have** [_{VP} **I**]]] **not** [_{WP} [_{FocP} [_{PtcP} **seen**]]]]

say that the participle adjoins to Top, and that this serves the same purpose as moving it to spec-Top. We choose to live with this problem for the purposes of this section. Consider the contrast below.

- (157) a. Sett har jeg 'n ikke.
 seen have I him not
 b. *Sett har jeg ikke 'n.
 seen have I not him

When the participle is fronted, the pronoun must precede the negation. This fact is problematic for the traditional account which treats weak pronoun distribution as pronoun shift and V2 as V-to-C with subsequent topicalization. In order to rule out (155a), such an account must block pronoun shift over the verb, but allow it over its trace. In order to account for the grammaticality of (157a), the same account must allow movement over the verb. In fact, in view of (157b) such movement must be forced just in case the verb subsequently topicalizes or undergoes V-to-C movement.

Suppose that the proper formulation of HG is in entirely phonological terms. The fact that pronoun shift is allowed around traces might follow rather elegantly from such a formulation, since traces do not have phonological content. But the fact that they can cross adverbs seems rather mysterious; the more so because they must cross adverbs if they can. It seems unlikely that there should be any phonological property distinguishing adverbs from all other material. To be on the safe side, we will show that there isn't. The Norwegian word *fortsatt* ambiguously represents the adverb 'still' and the participial verb 'continued'. In the adverb interpretation it forces pronoun shift and in the participial interpretation it blocks it.

- (158) a. Han har fortsatt det.
 he has continued it

- b. Han har det fortsatt.
he has it still

Rejecting a phonological formulation, there are two scenarios to consider according to whether or not bare heads are allowed to topicalize. If they are, one would need to assume that there is some domain extension mechanism (cf. Chomsky's 1995 notion of 'equidistance') which would be triggered by such topicalization. If head movement to spec-CP is not endorsed, one would need to assume that pronoun shift cannot cross the participle for syntactic reasons, but that this can be violated if the (remnant) VP subsequently topicalizes. On our account, nothing special needs to be said. We assume that heads can move to spec-Top with the caveat noted above. The derivation of (157a) runs as follows.

Derivation 3.16

[_{FinP} have [_{VP} I [_{PtcP} seen him]]]
MERGE Top and MOVE Ptc ►
[_{TopP} seen [_{FinP} have [_{VP} I [_{PtcP} him]]]]
MERGE *not* and MOVE TopP ►
[_{AdvP} [_{TopP} seen [_{FinP} have [_{VP} I [_{PtcP} him]]]] not]

This concludes our discussion of Mainland Scandinavian root clauses.

3.5.2 More problems

One obvious problem with the second approximation is that it leads us to expect sentences like (159) to be underivable, contrary to fact.

- (159) Sannsynligvis har Jens gått hjem.
probably has J gone home

The adverb *sannsynligvis* 'probably', being generated in the "adv*" area of (142), would have to be lowered into spec-TopP in order to end up in the first position. Assuming that lowering is impossible for principled reasons, this must be wrong.

A second problem, pointed out to me by Richard Kayne (p.c.) is that the account of topicalized bare participles in terms of head-movement to spec-TopP fails to generalize to cases like the following, where the fronted material is syntactically complex, but nevertheless stranding the pronoun.

- (160) a. [lagt på bordet] har jeg dem ikke.
put on the-table have I them not

- b. Jeg har (*dem) ikke (*dem) lagt *(dem) på bordet
 I have (*them) not (*them) put *(them) on the-table
 (*dem).
 (*them)

Given that the subject-initial counterpart of (160a) (i.e. (160b)) has the pronoun sandwiched between the participle and the PP, it seems that we have to assume that the weak pronoun is moved out of the VP prior to VP-topicalization. Thus, (160) suggests a derivation like derivation 3.17, which contradicts our assumption that weak pronouns don't shift.

Derivation 3.17

[not [put them on the table]]

MOVE *them* ►

[them_i [not [put t_i on the table]]]

MERGE *I* and *have* ►

[I have [them [not [put t_i on the table]]]]

MOVE *have* and VP ►

[[put t_i on the table]_j have_k [I t_k [them_i [not t_j]]]]

However, there are reasons to think that the position of the pronoun in (160) is due to an operation different from “ordinary” pronoun shift (henceforth PS; I will refer to the pronoun shift of examples like (160) as “stranded PS”, in short SPS). It is well-known that PS does not license parasitic gaps, thus (161a) is ungrammatical.

- (161) a. *Jeg kysset henne aldri uten å danse med *pg* først.
 I kissed her never without to danse with *pg* first
 b. Jeg kysset henne aldri uten å danse med **henne** først.
 I kissed her never without to danse with her first.

At first sight, SPS behaves the same, i.e. (162) is also ungrammatical.

- (162) *Kysset har jeg henne aldri uten å danse med *pg* først.
 kissed have I her never without to danse with *pg* first

However, when the adverbial PP is moved along with the topicalized participle, the parasitic gap becomes possible, if fact, even *obligatory*: (163b) is sharply ungrammatical.

- (163) a. [kysset uten å danse med *pg* først] har jeg henne aldri.
 [kissed without to danse with *pg* first] have I her never

- b. * [Kysset uten å danse med **henne** først] har jeg henne
 [kissed without to danse with her first] have I her
 aldri.
 never

For the standard theory of PS, this would have to be taken to indicate that the dependency between the pronoun and the VP-internal trace in SPS (163a) must be of a different kind than the dependency between the pronoun and the trace in “ordinary” cases of PS, such as (161b). SPS exhibits \bar{A} -properties, while PS doesn’t. In our setup, the difference would lie in the presence (163a) versus absence (161b) of a dependency. That is, I am suggesting that there is movement in the former case but not in the latter.⁴

A third problem is that we have stipulated that e.g. the subject must pied-pipe the VP (or vP) when it moves to spec-TopP. This is in order to rule out examples where the object is carried along with Vf to the initial position, crossing the subject, i.e. orders like (145), grammatical in Swedish. Suppose, putting aside Swedish for the moment, that only VPs can move to spec-FocP.⁵ Then these stipulations would follow from that. There is evidence that it is at least possible to move VPs to spec-Foc, with subsequent raising of the remnant EVP, thus yielding “VP-extraposition” structures.

- (164) a. ...at han hver dag møtte en ny pike
 ...that he every day met a new girl

In Nilsen (2000), these are interpreted as fronting of the adverbial *hver dag* to a high, left peripheral position. The grammaticality of (165), where a sentence adverb intervenes between the matrix subject and *hver dag* might be taken to call that analysis into question, however. Instead, I will interpret it as leftward movement of the VP [*met a new girl*] prior to leftward movement of the remnant EVP [every day t_{VP}], as illustrated in the derivation 3.18 below.

- (165) ...at han tydenigvis hver dag møtte en ny pike
 ...that he evidently every day met a new girl

3.6 German

German does not have V2-violations with *fpt* (166a). Thus, it might appear that our argument in section 2, that Mainland Scandinavian V2 is not derived by head movement, does not carry over to German. Of course, we would not want to say that German V2 is derived in a fundamentally different way

⁴This does not explain the properties of sentences like (163). In particular, it does not explain the obligatoriness of the parasitic gap. Although I find this question intriguing, I will leave it aside for now. What is important here is just that SPS behaves differently from PS in important respects, so we should not expect both to be reducible to the same operation.

⁵In case one needs further projections over VP, one could generalize this statement to “extended V-projections”.

Derivation 3.18

[_{EVP} [_{VP} met a new girl] every day]

MERGE Foc^0 and MOVE VP ►

[_{FocP} [_{VP} met a new girl] Foc^0 [_{EVP} _{t_{VP}} every day]]

MERGE W^0 and MOVE EVP ►

[_{WP} [_{EVP} _{t_{VP}} every day] W^0 [_{FocP} [_{VP} met a new girl] Foc^0 _{t_{EVP}}]]

from Mainland Scandinavian V2. After all, this is one property that the two languages obviously have in common. We would like to know, however, what underlies the contrast between the German sentence (166a) and the Norwegian one (167), and how the analysis of V2 defended here would fare with German facts.

- (166) a. *Gerhard nur weiß es nicht.
 G only knows it not
- b. Gerhard weiß es nur nicht.
 G knows it only not.
 “Gerhard just doesn’t know it.”
- (167) Jens bare vet det ikke.
 J just knows it not

However, as discussed in Meinunger (to app.); Fanselow (2002),⁶ other expressions, such as the operator *mehr als* ‘more than’ must precede the finite verb (when this is the modifiee). Consider the examples in (168).

- (168) a. daß Hans seinen Profit letztes Jahr mehr als verdreifachte
 that H his profit last year more than tripled
- b. *Hans verdreifachte seinen Profit letztes Jahr mehr als.
 H tripled his profit last year more than
- c. %Hans mehr als verdreifachte seinen Profit letztes Jahr.
 H more than tripled his profit last year.
- d. %Seinen Profit mehr als verdreifachte Hans letztes Jahr.
 his profit more than tripled H last year.

⁶Meinunger (to app., 2001) gives examples like (168c) as ungrammatical. Fanselow (2002) reports that in a survey, 6 out of 20 speakers actually accept them. For Dutch, I found that 4 out of 8 speakers accept them. The differences were very sharp: the ones who accept them, find them perfectly grammatical, while those who don’t find them sharply ungrammatical.

- (169)
- a. %Hij meer dan verdubbelde zijn scoretotaal
 hemore than dubbled his score-total
 vorig jaar.
 - b. *Hij verdubbelde zijn scoretotaal meer dan.
 he dubbled his score-total more than
 - c. %De winst meer dan verdubbelde.
 The gain more than dubbled
 - d. *De winst verdubbelde meer dan.
 The gain dubbled more than

In this case, the v2-violation is actually obligatory for the people who accept it. The speakers who don't accept them have to use a periphrastic construction to express the same (or a similar) meaning. I take this to show that some varieties of Dutch and German do allow for the relevant pattern. More research is required to determine the precise generalizations concerning the limitations on this phenomenon.

3.6.1 Hallman's analysis

In a recent paper, Peter Hallman (2001) proposes to tie together the verb-finality of German embedded clauses and the root-embedded asymmetry in the same language (i.e. root clauses are V2, embedded clauses are typically verb-final) in the following way. He follows, more or less the standard approach to root V2, i.e. the verb moves to some high F^0 , T^0 in his case, and then some XP moves to spec- F^0 .⁷

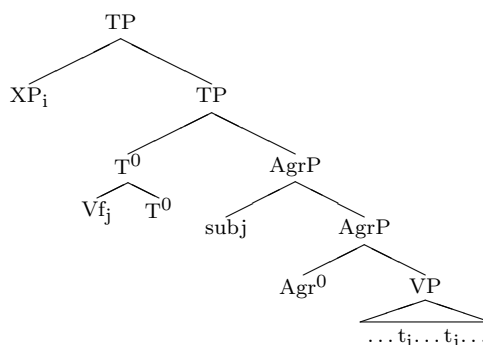


Figure 3.2: Hallman's root V2

Hallman's innovation is to treat embedded (V-final) clauses as, in a sense, V2 as well. His derivation of a V-final structure is given in Figure 3.3, where

⁷Hallman has AgrP, hosting the subject of the clause below TP.

Vf moves to T^0 as in Figure 3.2, and then the entire AgrP is moved to spec-T. In this sense, Vf is actually in second position, even in V-final structures.

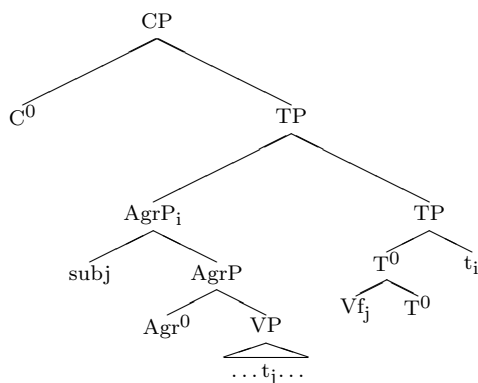


Figure 3.3: Hallman's V-final

3.6.2 Müller's V2 as *v*P first

Müller (2002) proposes an analysis of German V2 which, as we shall see, shares crucial assumptions with the analysis developed in section 3.5. However, the execution is quite different, and some of the facts discussed in the present work, notably HG, does not fall under Müller's analysis. I will use his analysis as a spring board to extend my analysis to German.

Müller casts the analysis in terms of PHASE theory (Chomsky, 1999, 2001). He adopts the following definitions to that end.⁸

- (170) Strict Cycle condition SCC
Within the current XP α , a syntactic operation may not target a position that is included within another XP β that is dominated by α .
- (171) Phase Impenetrability Condition (PIC)
Material that is dominated by a phase XP is not accessible to operations at ZP (the next phase) unless it is part of the edge domain of X.
- (172) Edge Domain
A category is in the edge domain of a head X if it is at an edge of the minimal residue of X

⁸Phases are utilized in derivations of island constraints. See Starke (2001) for a recent, comprehensive account of island constraints without resorting to phases.

- (173) a. Minimal Residue
The minimal residue of X includes X, and the head minimally c-commanded by X (the head residue, HR), and the specifiers of X (the spec-residue, SR).
- b. Edge
A category is at an edge of the minimal residue of X iff it is the highest phonologically overt item in HR or SR.

Müller then assumes that German clause structure conforms to (174);⁹ he treats “scrambling” as movement of some XP to an outer specifier of *v*; he follows Diesing (1992) and others in assuming optional subject raising to spec-T; he assumes that weak pronouns are obligatorily fronted within TP; that there is *wh*-movement to the specifier of “filled C”; and, finally, that main verbs always remain in situ, since head movement is ruled out in principle.

- (174) [CP C [TP T [_vP NP [_vP [VP ... V] *v*]]]]

V2 results from attraction of *v* by C. Since head movement is not possible, *v* must move as a phrase, thus (potentially) pied piping other material. Müller proposes the following constraint on *v*P movement to C:

- (175) Edge Domain Pied Piping Constraint (EPC)
A moved *v*P contains only the edge domain of its head.

This forces massive evacuation of *v*P prior to movement of *v*P to CP. He notes that such evacuation cannot be triggered by features, and therefore suggests that Last Resort (LR)¹⁰ should be weakened to a ‘soft’ (i.e. violable) constraint. Thus, his EPC must be ranked higher than Last Resort, implying an Optimality Theoretic (OT) evaluation procedure.¹¹ Given all this, Müller’s derivation for (176) is given in derivation 3.19.

- (176) Die Maria hat den Fritz geküßt
The-NOM M has the-ACC F kissed

Müller assumes that adverbs can be merged as specifiers of *v*P. If an adverb is merged before the subject, the subject will be in the edge domain of *v*, so the EPC forces adverb to move out before *v*P fronting takes place. This would lead to a subject-initial sentence again. If the adverb is higher than the subject, the latter is forced out of *v*P, again because of the EPC, leading to an adverb-initial sentence. One must then assume that *v*P evacuation is order preserving: If not, one would be able to derive sentences where the VP precedes the subject.

⁹Note that this structure seems to imply that Müller has head directionality.

¹⁰LR states that every movement must result in checking of a feature.

¹¹He also suggests that one could reformulate LR to the effect that movement must either be driven by feature checking or by the need to satisfy a constraint like EPC.

Derivation 3.19

$[_{vP} [\text{Die Maria}] [_{vP} [_{VP} [\text{den Fritz}] \text{geküßt}] \text{hat}]]$
 MERGE T and MOVE VP ►
 $[_{TP} [_{VP} [\text{Den Fritz}] \text{geküßt}]_i [_{TP} \text{T} [_{vP} [\text{die Maria}] [_{vP} t_i \text{hat}]]]]$
 MERGE C and MOVE vP ►
 $[_{CP} [_{vP} [\text{die Maria}] [_{vP} t_i \text{hat}]]_j C_{[+vP]} [_{TP} [_{VP} [\text{den Fritz}] \text{geküßt}]_i [_{TP} \text{T} t_j]]]$

Müller suggests that this is a general property of movement operations that are not feature driven.¹² I give his derivation for (177) in derivation 3.20.

- (177) Gestern hat die Maria den Fritz geküßt.
 Yesterday has the-NOM M the-ACC F kissed

Derivation 3.20

$[_{vP} \text{gestern} [_{vP} [\text{die Maria}] [_{vP} [_{VP} [\text{den Fritz}] \text{geküßt}] \text{hat}]]]$
 MERGE T and MOVE VP and *die Maria* ►
 $[_{TP} [\text{die Maria}]_i [_{VP} [\text{den Fritz}] \text{geküßt}]_j \text{T} [_{vP} \text{gestern} [_{vP} t_i [_{vP} t_j \text{hat}]]]]]$
 MERGE C and MOVE vP ►
 $[_{CP} [_{vP} \text{gestern} [_{vP} t_i [_{vP} t_j \text{hat}]]]_k C [_{TP} [\text{die Maria}]_i [_{VP} [\text{den Fritz}] \text{geküßt}]_j \text{T} t_k]]$

Object-initial sentences are analyzed by scrambling the object to spec-*vP* after merging the subject. He analyzes such scrambling as an instance of feature-driven movement, and he denotes the feature (or bundle of features) responsible for scrambling as $[\Sigma]$. Thus, (178) is derived in derivation 3.21.

- (178) Den Fritz hat die Maria geküßt.
 the-ACC F has the-NOM M kissed

This kind of derivation generalizes to other material which can be moved to spec-*v*. in other words, VP-topicalizations and long-distance topicalizations will be derived in the same way, by first moving the constituent in question to the highest specifier of *v*, then evacuating *vP* in order to satisfy the EPC and finally moving *vP* to spec-CP.

¹²One would still like to know why. Note also that, given that the subject and the VP move separately, this would have to be stated as a constraint, not on a single application of MOVE, but on a set of applications, all of which are not triggered by a feature. The fact that the two movements must cooperate in this way seems to me to suggest that we are rather dealing with one movement. I return to this point below.

Derivation 3.21

$$\begin{array}{l}
[v_P [die\ Maria] [v_P [VP [den\ Fritz]_{[\Sigma]} geküßt] hat_{[+\Sigma]}]] \\
\text{MOVE } den\ Fritz \blacktriangleright \\
[v_P [den\ Fritz]_{\Sigma} [v_P [die\ Maria] [v_P [VP t_{\Sigma} geküßt] hat_{[+\Sigma]}]]] \\
\text{MERGE } T, \text{ and MOVE } VP \text{ and } die\ Maria \blacktriangleright \\
[TP [die\ Maria]_i [VP t_{\Sigma} geküßt]_j T [v_P [den\ Fritz]_{[\Sigma]} [v_P t_i [v_P t_j hat_{[+\Sigma]}]]]] \\
\text{MERGE } C \text{ and MOVE } vP \blacktriangleright \\
[CP [v_P [den\ Fritz]_{[\Sigma]} [v_P t_i [v_P t_j hat_{[+\Sigma]}]]]_k C [TP [die\ Maria]_i [VP t_{\Sigma} geküßt]_j T t_k]]
\end{array}$$

Müller’s account is similar to the one developed in section 3.5 (Nilsen, to app.b) in the sense that, whereas Müller has *vP* first, I have *TopP* first. In fact, movement to spec *TopP* and movement to an outer specifier of *vP* (triggered by Σ) are also similar, modulo labelling and Müller’s EPC-driven *vP*-evacuation is analogous to our movement to spec-Foc. But there are also differences. An obvious one is that it is unclear how, on Müller’s account, we could account for Holmberg’s Generalization (HG) in the way that we did in section 3.5. Since the pronouns follow the finite verb, they would have to reside in *VP* on Müller’s account. But *VP* could only participate in Müllerian *vP*-fronting in the relevant cases at the cost of incurring an EPC-violation. Perhaps one could treat the EPC as a soft constraint as well, which is what Müller suggests for Last Resort. But assuming an OT evaluation mechanism on top of the sort of syntax that we are considering seems to me slightly off the parsimonious track. Another problem, pointed out by Müller, is that the analysis relates topicalization to scrambling. In other words, in this setup, an object noun phrase can only end up preceding the finite verb by scrambling past the subject to the left edge of *vP*. This creates problems for all the other Germanic V2 languages which do not allow such scrambling in the first place, but do allow objects to occupy the first position.¹³ Müller’s account also has some advantages over mine. For example, it does not have a problem with generating V2-initial adverbs, which, as we saw, pose problems for my account. Secondly, and related to the previous point, Müller’s account correctly leads on to expect there to be an asymmetry between subjects and adverbs in first position on the one hand, and objects and other arguments on the other. The latter kind of expression can only occupy the first position in certain discourse functions (see below), whereas subjects and adverbs do not require special discourse function to occupy the first position.

¹³Dutch, Norwegian, Swedish and Icelandic allow “argument-shift” of weak pronouns and full noun phrases, as long as the relative ordering of the arguments is unaffected. Danish only allows shifting of weak pronouns.

3.7 Third approximation: V2 without positions

In order to incorporate the advantages of Müller's account, I will assume that there is a head Σ which can merge above or below the subject. If it is merged below the subject, as in the leftmost tree in figure (3.4), the subject will count as a "specifier" of Σ . If it is merged above, any adverb merged to the result of that will count as the "specifier" of Σ . This is illustrated in the rightmost tree in figure (3.4). I will furthermore assume that Σ has an EPP-feature: it must have a phonetically visible specifier.

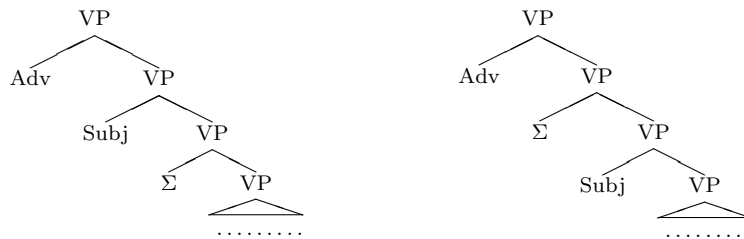


Figure 3.4: Unmarked Σ

Finally, Σ attracts the verb. If it has a marked feature, it attracts some phrase functioning as a contrastive topic. Before we go on, we must address how the verb moves to Σ .

Head Movement: an aside

Müller's analysis is motivated by the desire to rule out head movement in principle. If this is to be possible, phenomena like V2 must be analyzed without recourse to HM. Needless to say, I fully agree with Müller that V2 can and should be so analyzed. However, I would like to point out some aspects of Müller's approach that I take to be problematic. First, the machinery that he employs to rid the system of HM is rather elaborate (Phases, massive remnant movement, optimality theory, etc.). Thus it is not clear that eliminating HM in this way represents a simplification of the system. This becomes even less clear if we look at how Müller derives the unavailability of HM. Suppose that we treat the unavailability of HM as an empirical generalization in need of explanation, and then look at how it is explained in the setup under consideration.

Generalization 3 (Müller (2002)) *Heads stay.*

Müller's explanation is that adjunction of a head to another one would violate the Extension Condition Chomsky (1995), stated roughly in (179)

- (179) Extension Condition
 Merger extends the tree at the root.

This would rule out head movement only if it follows from independent considerations that moved heads cannot be merged to the root. In other words, the tree in figure 3.5 must be ruled out if α is a moved head (not if it isn't moved, of course, because then heads could never be merged to anything). We now

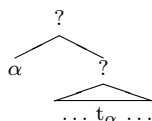


Figure 3.5: Illicit merger

need to explain why it matters that α is *moved* and why it matters that it is a *head*. To this end, we adopt some contextual definition of the notion “head”, crucially involving that all heads project. In other words any syntactic object that does not project further is a maximal projection, hence not a head. We then assume some version of the Chain Uniformity Condition, such as (180):

- (180) Chain Uniformity Condition (CUH)
 Chains are uniform with respect to the feature $[\pm\text{max}]$ (and, perhaps other features)

If the label of the root node is not allowed to project from the moved head, this setup rules out the configuration in question, because the trace of α is a head (does project further), while α is a maximal projection (does not project). Hence the configuration violates the CUH. As noted in Fanselow (2002), the CUH would be satisfied if α projects, since, in that case, both the trace and its antecedent are heads. In order to rule out this option, Müller could make reference to his Unambiguous Domination constraint (UD):

- (181) Unambiguous Domination (UD)
 An α -trace cannot be α -dominated.

Müller motivates this constraint by independent considerations.¹⁴ This rules out the version of figure 3.5 where the root label is projected from α , because the root node dominates the trace of α .

In sum, HM is ruled out, because the Extension Condition rules out adjunction of a head to a head; the CUH rules out attachment of a non-projecting, moved head to the root node, given the $[\pm\text{max}]$ distinction; and the UD rules out attachment of a projecting, moved head to the root node.

It seems to me that this is just a roundabout way of stating generalization 3. It seems to say that heads can't move because they have a feature $[-\text{MAX}]$ that prevents them from moving. Furthermore, this feature doesn't seem to do anything else than just that.

¹⁴It is needed to explain e.g. why remnant VPs can be topicalized but not scrambled in German.

Suppose that we don't allow ourselves all this machinery. In other words, we do away with the distinction between heads and phrases. Then the CUH becomes vacuous, so it can be dismissed, too. Thus, we are left with the extension condition. Then we would be closer to the system discussed in Koenenman (2000); Fanselow (2002). However, whereas, e.g. Fanselow (2002) uses the CUH to allow head movement *only* if the head projects, we do not have this option, since we do not have the CUH (or any distinction to state it on). Therefore, we would be led to allow heads to move to specifiers.

Holmberg (2000) argues on the basis of empirical facts that head movement to specifiers should be allowed, just in case it is movement of the phonological features of the head alone. In other words, if it is driven by purely phonological considerations. Bobaljik and Brown (1997) point out that, even with the elaborate machinery employed in the scenario above to rule out head movement, one could think of ways to allow it. Their idea is that, in a theory of MOVE as "copy+merge", one could merge two heads before they are merged to the main tree. In other words, if the derivation has reached a stage such as (182a), where x is a head, and the next thing to be merged y is also a head, one can copy x and merge it to y , yielding the intermediate stage (182b) with two distinct syntactic objects. Then these are merged in (182c).

- (182) a. $[xp \dots x \dots]$
 b. $[y \ x \ y] \ [xp \dots x \dots]$
 c. $[yp [y \ x \ y] [xp \dots x \dots]]$

One might object that this would allow "sideways" movement, i.e. movement to a non-c-commanding position, but Bobaljik and Brown (1997) argue that there are obvious ways around this.

I will not dwell on the issue here. I will assume that movement of V to Σ is head-movement in the traditional sense that the verb adjoins to Σ . I take it to be driven by the phonetic emptiness of Σ .

3.7.1 the analysis

In a sense, Müller's massive vP evacuation is forced by the fact that he assumes that v and the left-edge XP may be separated by other material. This leads to the situation that lower specifiers of v and VP do not make up a constituent. Hence, when they must leave vP , they must do so separately. But the mere fact that they must land in the same order as they started may lead one to suspect that they move as one constituent, rather than massive parallel movement. The elegance of that would be that it would derive the order preservation property, rather than stipulating it as an extra constraint on the output. Suppose, therefore, that our head Σ corresponds to the standard v . As before, it attracts the highest verb. It must have a specifier, i.e. it has an "EPP" feature. If it also has a marked Σ feature, it furthermore requires its specifier to be a (contrastive) topic. Suppose, furthermore that, Σ merges above or below the subject, and

that an adverb merged immediately after Σ will satisfy its EPP. This would be the same for all the V2 languages. They differ with respect to what happens next.

Main clauses

For Main clauses, I assume very much the same analysis as before. Any VP-node which contains focused material must move out of the scope of Σ , to the left of its specifier. Next, Σ pied pipes it's "specifier" to the left of that again. For ease of reference, I will refer to the smallest node containing both Σ and its specifier (and potentially more material following Σ) as ΣP . Thus, we end up with the following kind of structure. Note now that whatever material

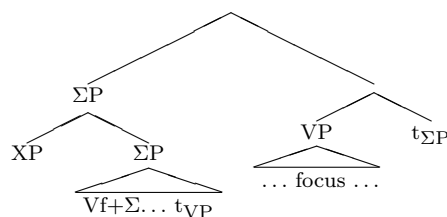


Figure 3.6: V2 structure

follows Σ inside ΣP will end up preceding the extracted VP, as it did before movement. Furthermore, no reordering of material will take place within VP. Hence, Müller's order preservation constraint follows.

Before I demonstrate this, some words on what can occupy spec- Σ when this head has an unmarked feature are in order. Subjects definitely can, and some though not all adverbial type elements. Thus, the examples in (183) do not require any marked intonation pattern, while those in (184) do.

- (183)
- a. Han så ikke Jens.
he saw not J
 - b. Kanskje han ikke så Jens.
maybe he not saw J
 - c. Derfor så 'n ikke Jens.
Therefore saw he not J
- (184)
- a. Alltid så han Jens.
always saw he J
 - b. Jens så han ikke.
J saw he not
 - c. Så Jens gjorde han ikke.
saw J did he not

- d. muligens så han Jens.
 possibly saw he J

(183a) is the clearest case of an unmarked XP. In (183b) we see the well-known fact that the adverb *kanskje* ‘maybe’ can serve as both the first constituent and the finite verb simultaneously, as it were (Platzack, 1986).¹⁵ I take this to be a special case, attributable to the fact that the adverb is “verbal” and “phrasal” or “adverbial” at the same time, so it can check both the EPP feature and the V feature of Σ . Finally, discourse relation markers, like *derfor* ‘therefore’, *likevel* ‘nevertheless’ etc. can (but need not) occupy the first position. These may clearly relate to a marked, rather than an unmarked Σ , if marked Σ is to be related to topichood. I will leave these aside. Hence it really seems that the subject is the core unmarked XP. Everything else requires a marked intonation pattern. Consider now potentially problematic cases where an unstressed object pronoun occupies XP. In particular, consider (185b) and (185c) as replies to the question in (185a).

- (185) a. Så du Jens?
 Saw you J
 b. Nei, han så jeg ikke.
 no him saw I not
 c. Nei, jeg så 'n ikke.
 no I saw him not

(185b), construed as an answer to (185a), suggests that you did see somebody else, although you didn't see *Jens*. (185c) does not give rise to such a suggestion. In other words, even the unstressed object pronouns seem to receive a contrastive interpretation when they occur in first position. Thus, I take this to suggest that the only way for a non-subject (discourse markers aside) to occupy spec- σ is by attraction of a marked value. Now, note also that the discourse value of the first (marked) constituent is not that of “new information focus”. This can be seen by the fact that an object in first position is very clumsy as an answer to an object wh-question. Suppose that, upon meeting John outside the cinema, I ask him (186a). (186b) does not seem to be an adequate answer to the question, or at the very least it requires that we have been talking about the movie “Mulholland Drive” before. In other words, it must be in our common ground. (186c) would be the most straightforward answer.

- (186) a. Hvilken film har du sett?
 which movie have you seen?
 b. # Mulholland Drive har jeg sett.
 M D have I seen

¹⁵In Swedish, the corresponding adverb works slightly differently: It behaves as a finite verb, i.e. may (but need not) occupy the second position in traditional terms. I refer the reader to Platzack's work for discussion.

- c. Jeg har sett Mulholland Drive.
 I have seen M D

It seems to me that non-canonical topicalization often (if not always) leads to switching of topic to another already accessible one. Thus, if the discourse has been about topics x, y, z , and z is the current one, topicalization of a constituent denoting y makes y the current topic again. Non-canonical topicalization of an x which *is* the current topic, like in (185b) seems to suggest that x is no longer the topic. In other words, marked Σ denotes “switch topic”.

The analysis for simple main clauses generalizes straightforwardly to German and Dutch. I return to the slightly more complicated question of periphrastic constructions and verb finality shortly.

Order preservation and scrambling

Let us now see how our setup derives order preservation. I give the derivation for (187) in derivation 3.22. The crucial fact is that the two objects must occur in the same order as they were merged.

- (187) Jeg ga alltid Jens en kylling.
 I gave always J a chicken

Derivation 3.22

[I gave Jens a chicken]
 MERGE Σ and MOVE *gave* and *I* ►
 [_i [_j gave_j + Σ [_i _j Jens a chicken]]
 MOVE VP ►
 [[_i _j Jens a chicken]_k [_i [_j gave_j + Σ _k]]]
 MERGE *always* and move Σ P ►
 [[_i [_j gave_j + Σ _k]_l [_i [_j always [_i _j Jens a chicken]_k _l]]]

Suppose that we move the two objects separately instead of as a constituent. Clearly, this could reverse their order, if we can move the indirect object first, and then move the direct object across it. Of course, this is perfectly possible in languages like German which has argument–argument scrambling. However, Mainland Scandinavian, Icelandic and Dutch do not allow argument reordering. I suggest that the difference lies in whether the language in question allows the arguments themselves, or only VPs to “scramble” out of Σ P for focus reasons. If only VPs are allowed, the Extension Condition will actually rule out reordering. Consider the following partial derivation 3.23 for (188).¹⁶ I assume a VP-shell

¹⁶As argued extensively in Nilsen (1997). Norwegian and Swedish do allow for “object shift” of full DPs, indirect, as well as direct objects, even simultaneously. However, the

(188) Jeg ga Jens alltid en kylling.
I gave J always a chicken

$$\begin{array}{l}
\text{[I gave[Jens [a chicken]]]} \\
\text{MERGE } \Sigma \text{ and MOVE } gave \text{ and } I \blacktriangleright \\
\text{[I}_i \text{ [gave}_j \text{+}\Sigma \text{ [t}_i \text{ t}_j \text{ [Jens [a chicken]]]]]} \\
\text{MOVE } a \text{ chicken } \blacktriangleright \\
\text{[[a chicken]}_k \text{ [I}_i \text{ [gave}_j \text{+}\Sigma \text{ [t}_i \text{ t}_j \text{ [Jens t}_k \text{]]]]]} \\
\text{MERGE } always \text{ and MOVE } Jens \blacktriangleright \\
\text{[[Jens t}_k \text{]}_l \text{ [always [[a chicken]}_k \text{ [I}_i \text{ [gave}_j \text{+}\Sigma \text{ [t}_i \text{ t}_j \text{ t}_l \text{]]]]]} \\
\text{MOVE } \Sigma P \blacktriangleright \\
\text{[[I}_i \text{ [gave}_j \text{+}\Sigma \text{ t}_l \text{]]}_m \text{ [[Jens t}_k \text{]}_l \text{ [always [[a chicken]}_k \text{ t}_m \text{]]]}
\end{array}$$

Derivation 3.24

$$\begin{array}{l}
[[a \text{ chicken}]_k \text{ [I}_i \text{ [gave}_j + \Sigma \text{ [t}_i \text{ t}_j \text{ [Jens t}_k \text{]]]]}] \\
\text{MERGE } \textit{always} \blacktriangleright \\
[\text{always } [[a \text{ chicken}]_k \text{ [I}_i \text{ [gave}_j + \Sigma \text{ [t}_i \text{ t}_j \text{ [Jens t}_k \text{]]]]}]] \\
\text{MOVE } \Sigma P \blacktriangleright \\
[[\text{I}_i \text{ [gave}_j + \Sigma \text{ [t}_i \text{ t}_j \text{ [Jens t}_k \text{]]}]]_1 \text{ [always } [[a \text{ chicken}]_k \text{ t}_i \text{]]}]
\end{array}$$

analysis presented there does not derive the difference between pronoun shift and object shift of full DPs.

hence more orderings ensue. Whether DP scrambling, rather than merely VP scrambling is allowed, could be related to the presence of morphological case, but addressing this question is beyond the scope of the present work. It also seems to me that the present analysis would extend to other “second position” phenomena, either by analyzing the second position elements by attraction to Σ , or by treating them on a par with weak pronouns, i.e. elements that cannot trigger extraction from Σ P on their own, and hence, will tend to be tagged along when Σ P moves to the first position. This explains Norwegian patterns like the following, discussed in chapter 1 as “Bobaljik’s Paradox” (Bobaljik, 1999; Nilsen, 1997), where three arguments appear to be scrambling on their own, among several adverbs, where the ordering of the adverbs *and* the ordering of the arguments must remain unaltered.

- (189) a. Derfor ga **Jens Kari** **kyllingen** tydeligvis ikke
therefore gave J K the-chicken evidently not
lenger kald.
any.longer cold
- b. Derfor ga **Jens Kari** tydeligvis **kyllingen** ikke lenger kald.
- c. Derfor ga **Jens** tydeligvis **Kari kyllingen** ikke lenger kald.
- d. Derfor ga **Jens** tydeligvis **Kari** ikke **kyllingen** lenger kald.
- e. Derfor ga **Jens** tydeligvis **Kari** ikke lenger **kyllingen** kald.
- f. Derfor ga **Jens** tydeligvis ikke lenger **Kari kyllingen** kald.
- g. Derfor ga tydeligvis **Jens** ikke lenger **Kari kyllingen** kald.
- h. Derfor ga tydeligvis ikke **Jens** lenger **Kari kyllingen** kald.
- i. Derfor ga tydeligvis ikke lenger **Jens Kari kyllingen** kald.
- j. * Derfor ga **Jens** ikke tydeligvis **Kari** lenger **kyllingen** kald.
- k. * Derfor ga **Jens** tydeligvis ikke **kyllingen** lenger **Kari** kald.

Such patterns are obviously problematic if one wants to assume fixed landing sites for scrambled arguments, and at the same time assume fixed positions for the adverbs. In the present account, then, we don’t have to assume fixed positions for either. The ordering of the adverbs follows from the scope requirements of the different adverbs,¹⁷ while the ordering of the arguments follows from the order of merger in the VP (linking) in addition to the restriction of scrambling to VP nodes.¹⁸

Again, Dutch behaves essentially the same. For German, which allows argument reordering, we would simply loosen up the requirement that only VP

¹⁷according to the analysis developed in chapter 2, *lenger* ‘no.longer’ must follow the negation because it is a negative polarity item, while *tydeligvis* ‘evidently’ must precede it because it is a positive polarity item.

¹⁸It should be pointed out here that my solution to Bobaljik’s paradox is independent of the question whether adverbs occur in fixed positions. Thus, one could claim that the adverbs occur in fixed positions and that remnant VP-nodes can scramble in the fashion outlined among the FPs hosting the adverbs.

nodes can scramble. As we noted above, if arguments can scramble on their own, they may not end up in the same order as they started. I give a derivation for the German sentence below, where the direct object has scrambled around an adverb and the subject.

- (190) Gestern küßte den Fritz wahrscheinlich die Maria.
 Yesterday kissed the-ACC F probably the-NOM M

Derivation 3.25

$$\begin{aligned}
 &[\Sigma_P \text{ gestern } [\text{küßte}_i + \Sigma \text{ } [[\text{die Maria}] \text{ } t_i [\text{den Fritz}]]]] \\
 &\hspace{20em} \text{MOVE } \textit{die Maria} \blacktriangleright \\
 &[[[\text{die Maria}]_j [\Sigma_P \text{ gestern } [\text{Küßte}_i + \Sigma \text{ } [t_j \text{ } t_i [\text{den Fritz}]]]]]] \\
 &\hspace{20em} \text{MERGE } \textit{warscheinlich} \text{ and MOVE } \textit{den Fritz} \blacktriangleright \\
 &[[[\text{den Fritz}]_k [\text{warscheinlich } [[[\text{die Maria}]_j [\Sigma_P \text{ gestern } [\text{küßte}_i + \Sigma \text{ } [t_j \text{ } t_i \text{ } t_k]]]]]]]] \\
 &\hspace{20em} \text{MOVE } \Sigma_P \blacktriangleright \\
 &[[[\Sigma_P \text{ gestern } [\text{küßte}_i + \Sigma \text{ } [t_j \text{ } t_i \text{ } t_k]]]_l [[[\text{den Fritz}]_k [\text{warscheinlich } [[[\text{die Maria}]_j \\
 &\text{ } t_l]]]]]]
 \end{aligned}$$

It can be seen that it is the possibility of extracting the argument [die Maria] without pied piping its dominating VP node that results in the potential for reordering. A remaining problem is Danish, where all subjects and weak pronouns must precede all adverbs, and all other arguments line up following the adverbs. In other words, Danish does not have argument shift or scrambling. This would follow if Danish must always move the VP sister of the subject out of Σ_P immediately after XP movement to spec- Σ . I leave open the interesting question why this should be so.

Periphrastic constructions In chapter 2, (page 72) I noted that sequences of adverbs and auxiliaries in Norwegian enter into crossing scope dependencies. Hence, I concluded that there must be a rather elaborate set of movements in order to generate this. The sentence we considered was the following, where the four adverbs and the four auxiliaries are linearly separated, but semantically interspersed, as it were. In other words, the linear order in (191b) reflects the semantic scope of the adverbs and auxiliaries in (191a). However, while (191a) is perfectly grammatical, (191b) is sharply ungrammatical.

- (191) Norwegian
 a. ...at det ikke lenger alltid helt kunne ha blitt
 ...that it not any.longer always completely could have been
 ordnet.
 fixed

- b. *...at det ikke kunne lenger ha alltid blitt helt
 ...that it not could any.longer have always been completely
 ordnet.
 fixed

The same observation can be made on the basis of Dutch data. Thus, (192) is a similar example in this language.

- (192) ...dat het niet meer helemaal kon worden gemaakt
 ...that it not any.longer completely could be fixed

In chapter 2, I suggested that this can be derived by letting adverbs attract projections of verbs, and verbs attract projections of adverbs. I repeat the derivation I gave there for the Norwegian sentence.

What could be driving all these movements? Of course, we could stipulate that all these expressions are lexicalized with uninterpretable V-features and Adv-features, but that does not lead to any further understanding of why the movements should apply. Another possibility is that they are driven by an adjacency requirement on the auxiliaries. This also seems rather unsatisfactory; we should probably rather derive the adjacency requirement from something more fundamental. The auxiliaries enter into selectional relations with each other, and it might be that the reason for all the movements is that each auxiliary must be adjacent to the auxiliary (verb) it selects. For example, auxiliary *ha* ‘have’ selects for a participial complement, while *kunne* ‘could’ selects for an infinitival complement. If an intervening adverb would block the selection relation, this could suffice to drive the movements. This would leave some room for variation with respect to how exactly a language chooses to satisfy the requirement. Thus, for lack of any better account, I assume that this is how the movements are triggered. Note that, if the adverb projections had not moved in derivation 3.26, the verbs would not end up being adjacent.

Suppose that, in a configuration like figure 3.7, the four nodes xp_1 , xp_2 , yp , x are ‘equidistant’ to α . If we raise yp to spec- α , and subsequent attractors will iterate this option, we will derive climbing of yp to the first position. If we, instead, move the entire complement of α , xp_1 , and later iterate this option, we derive roll-up structures. Both options are needed to derive the ordering patterns of verbal clusters. If we extract the head of xp , and later on iterate this, we end up with the order of merger. This is illustrated in derivation 3.27. This derives the English pattern, where adverbs and verbs are interspersed. The verbs are adjacent at the point of merger. The fourth option is to move xp_2 . This option we have already illustrated in derivation 3.26 for Norwegian facts. Given what we have seen so far, these derivational options seem to be parameters that are fixed for entire categories.¹⁹

¹⁹Some questions arise with respect to interaction of two kinds of x-raising at the same time. For example, one could investigate whether some otherwise possible orderings of Dutch verbal clusters are ruled out in the presence of a preceding crossing-scope adverb cluster.

Derivation 3.26

[completely [fixed]]	
	MOVE VP ►
[fixed [completely]]	
	MERGE <i>been</i> ►
[been [fixed [completely]]]	
	MOVE AdvP ►
[completely [been [fixed]]]	
	MERGE <i>always</i> ►
[always [completely [been [fixed]]]]	
	MOVE VP ►
[[been [fixed]] [always [completely]]]	
	MERGE <i>have</i> ►
[have [[been [fixed]] [always [completely]]]]	
	MOVE AdvP ►
[[always [completely]] [have [been [fixed]]]]	
	MERGE <i>any.longer</i> ►
[any.longer [[always [completely]] [have [been [fixed]]]]]	
	MOVE VP ►
[[have [been [fixed]]] [any.longer [always [completely]]]]	
	MERGE <i>could</i> ►
[could [[have [been [fixed]]] [any.longer [always [completely]]]]]	
	MOVE AdvP ►
[[any.longer [always [completely]]] [could [have [been [fixed]]]]]	
	MERGE <i>not</i> ►
[not [[any.longer [always [completely]]] [could [have [been [fixed]]]]]]	

An approach along these lines could be extended to roll-up structures like the ones found in German. I will not do that here, but see Koopman and Szabolcsi (2000) for an analysis of verbal clusters along similar lines for several languages.

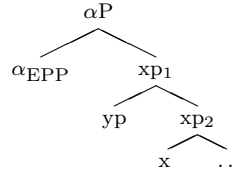


Figure 3.7: x-raising configuration

Derivation 3.27

[_{vp2} completely [_{vp1} fixed]]

MOVE *fixed* ►

[_{vp2} fixed_i [_{vp2} completely [_{vp1} t_i]]]

MERGE *been* and MOVE *completely* ►

[_{vp3} completely_j [_{vp3} been [_{vp2} fixed_i [_{vp2} t_j [_{vp1} t_i]]]]]

MERGE *always* and MOVE *been* ►

[_{vp4} been_k [_{vp4} always [_{vp3} completely_j [_{vp3} t_k [_{vp2} fixed_i [_{vp2} t_j [_{vp1} t_i]]]]]]]]

MERGE *have* and MOVE *always* ►

[_{vp5} always_l [_{vp5} have [_{vp4} been_k [_{vp4} t_l [_{vp3} completely_j [_{vp3} t_k [_{vp2} fixed_i [_{vp2} t_j [_{vp1} t_i]]]]]]]]]]

MERGE *any.longer* and MOVE *have* ►

[_{vp6} have_m [_{vp6} any.longer [_{vp5} always_l [_{vp5} t_m [_{vp4} been_k [_{vp4} t_l [_{vp3} completely_j [_{vp3} t_k [_{vp2} fixed_i [_{vp2} t_j [_{vp1} t_i]]]]]]]]]]]]

MERGE *could* and MOVE *any.longer* ►

[_{vp7} any.longer_n [_{vp7} could [_{vp6} have_m [_{vp6} t_n [_{vp5} always_l [_{vp5} t_m [_{vp4} been_k [_{vp4} t_l [_{vp3} completely_j [_{vp3} t_k [_{vp2} fixed_i [_{vp2} t_j [_{vp1} t_i]]]]]]]]]]]]]]

MERGE *not* and MOVE *could* ►

[_{vp8} could_o [_{vp8} not [_{vp7} any.longer_n [_{vp7} t_o [_{vp6} have_m [_{vp6} t_n [_{vp5} always_l [_{vp5} t_m [_{vp4} been_k [_{vp4} t_l [_{vp3} completely_j [_{vp3} t_k [_{vp2} fixed_i [_{vp2} t_j [_{vp1} t_i]]]]]]]]]]]]]]]]

Dutch and Mainland Scandinavian would follow derivation 3.26. In Dutch, it seems that non-verbal material from the most deeply embedded VP ends up in the adverb cluster, rather than in the verb cluster. Hence, we get “climbing” of verbal particles etc. leading to the characteristic Dutch pattern in (193) (Koopman and Szabolcsi, 2000).

- (193) ...dat Jan Marie op zal willen bellen
 ...that J M up shall want call
 "... that Jan will want to call Mary up"

Once we have formed the adverb cluster and the verb cluster, and merged the finite verb, we merge Σ and execute the analysis as before.

Embedded clauses

Finite embedded clauses are normally not V2. In particular, only so-called "bridge" predicates allow for embedded V2. I return to embedded V2 shortly. Embedded (non-V2) clauses must have a different Σ than root clauses. I assume that the complementizer is such a Σ . It has no EPP-feature, so no topicalization is possible. Furthermore, it does not attract the verb. Other than that, everything goes as before. However, since the verb is not a weak element like the pronouns, it will always trigger VP extraction from ΣP . I give the derivation for (194) in derivation 3.28.

- (194) ... at Jens ofte spiser tran
 ... that J often eats cod.liver.oil

Derivation 3.28

[that [VP₁ Jens [VP₂ eats cod liver oil]]]
 MERGE *often* and MOVE VP₂ ►
 [[VP₂ eats cod liver oil]_i [often [that [VP₁ Jens t_i]]]]
 MOVE ΣP ►
 [[that [Jens t_i]]_j [often [[eats cod liver oil]_i t_j]]]

The question now arises whether the direct object could have extracted alone, so that the verb could become separated from it by the adverb, much as in the cases discussed for arguments in the previous subsection. The short answer is that it can. Bentzen (2002) notes that sentences like the following are, in fact, grammatical in Norwegian.

- (195) ...at han spiser ofte tran.
 ...that he eats often cod.liver.oil

However, this option is quite restricted. The adverb *ofte* is one of the few which can occupy this position, and substituting e.g. *alltid* 'always' or other adverbs for *ofte* leads to degradedness. It remains to be understood what governs the availability of orders like (195).

For OV languages I will assume the following, inspired by the analysis proposed by Hallman (2001). The verb moves to Σ in these cases as well, but there is no EPP-feature, hence no topicalization. In this case, what is

attracted to C is the VP dominating the complementizer. Thus, essentially, the verb will be left behind in the final position, as wanted. Before attraction to C, scrambling works as before, i.e. with VPs in Dutch, hence leading to order preservation, and with the arguments themselves in German, leading to potential reordering of the arguments. I give the derivation for the Dutch embedded sentence (196) in derivation 3.29.

- (196) ...*dat* Jan Marie kuste
 ...*that* J M kissed

Derivation 3.29

$[\Sigma \text{ that } [\text{VP Jan kissed Mary}]]$

MOVE V ►

$[\text{kissed}_i [\Sigma \text{that } [\text{VP Jan } t_i \text{ Mary}]]]$

MERGE C and MOVE Σ ►

$[\text{CP } [\Sigma \text{ that } [\text{VP Jan } t_i \text{ Mary}]]_j [\text{kissed}_i t_j]]$

The bracket to the immediate left of *kissed* in derivation 3.29 is not labelled. Given that the verb is attracted there by Σ , it should be a Σ P. Hence, one might wonder why we do not get Vf first, and complementizer second. We need the Σ to pied-pipe its specifier just in case it is not realized by *dat*, and similarly for German. This problem is similar to the question raised by the traditional analysis of v2, namely why CPs do not allow multiple adjunction (in V2 languages). I speculate that the answer should ultimately be given in phonological terms. In other words, *dat* wants to be leftmost in an intonation phrase as do topic switchers. Such prosody-semantics/pragmatics correspondences have been explored for focus and stress by Reinhart (1995); Szendrői (2001). Working out such an account in detail for topic switch and intonation phrasing is beyond the scope for the present dissertation.

Note that we can use the same account for pronoun-shift and scrambling in Dutch as the one we explored for Norwegian above. In particular, weak pronouns are expected to immediately follow the complementizer, unless they have been extracted along with other material from Σ P prior to Σ P fronting. In either case they will end up preceding the verb. Our account also explains why one part of HG holds for Dutch, namely that objects cannot scramble around subjects. This follows if Dutch, like Norwegian, does not allow scrambling of arguments on their own, but only of VP nodes. Finally, we have an explanation why the other part of HG does *not* appear to hold for Dutch: Scrambling is not contingent on verb movement in this language. This now follows from the fact that the verb (in embedded clauses) is attracted and then stranded by Σ_{dat} . In Norwegian, Σ_{at} does not attract the verb, so, since only VP-nodes can move,

objects must end up following the verb.²⁰

3.7.2 Σ P fronting

Why does Σ P have to move to the beginning of the clause? Are there any languages where it does not? Suppose that there are. Such a language would have sentences with the word-order of penultimate steps of the derivations above for Norwegian. In other words, it would have a designated XP position left adjacent to the finite verb, it would generally move arguments to the left of this XP position, and, finally destressed material, like weak pronouns would occur to the right of the verb. If wh-movement is thought of as movement to spec- Σ , this language would also have a wh-phrases to the immediate left of the finite verb, though clause internally. In fact, Jayaseelan (2001) shows that the Dravidian language Malayalam has precisely such properties and proposes an analysis which is congenial to the one I am proposing in several respects. In other words, Malayalam is generally OV, and wh-phrases must immediately precede Vf.²¹ Compare (197) to (197b) (Jayaseelan, 2001, p40).

- (197) a. ninn-e aare aTiccu?
you-ACC who beat-PAST
b. iwiTe aare uNTe?
here who is
c. awan ewiTe pooyi?
he where went
d. nii aa pustakam aar-kke kiDuttu?
you that book who-DAT gave
- (198) a. *aare ninn-e aTiccu?
who you-ACC beat-PAST
b. *aare iwiTe aare uNTe?
who here is
c. *ewiTe awan pooyi?
where he went
d. *nii aar-kke aa pustakam kiDuttu?
you who-DAT that book gave

More or less focused material occurs to the left of the verb, as seen in (199), and, finally, destressed material can occur to the right of Vf (200b). The latter option is unavailable or indefinite noun phrases (200c).

²⁰An alternative formulation would be to say that also the Norwegian complementizer attracts the verb, but that the verb pied-pipes the VP in this language. The choice between the two alternatives hinges on the proper explanation of examples like (195).

²¹Jayaseelan (2001) points out that Vf left-adjacent wh-phrases of this sort are also found in Hungarian, Basque, and several African languages, like Aghem, Chadic and Kirundi. For this, I refer the reader to Jayaseelan's work and references cited there.

- (199) ñaan innale Mary-k'k'ε oru kattε ayaccu
 I yesterday Mary-DAT a letter sent
- (200) a. aarum kaND-illa, aana-ye
 nobody saw-NEG elephant
- b. aarε ayaccu, ninn-e?
 who sent you-ACC
- c. ?* ñaan awan-εayaccu, oru kattε
 I he-DAT sent a letter

Jayaseelan (2001) analyzes the postverbal destressed material as occupying a high ToP, with IP moving around it, but does not give arguments for this. Malayalam is a pro-drop language, so one cannot test for weak pronouns. The postverbal topic position could in principle also be analyzed in the same way as I have analyzed Norwegian weak pronouns. Finally, the XP position is reserved for focused material in Malayalam. It is typically occupied by an indefinite noun phrase, and if there are both definites and indefinites in the clause, the indefinites must occupy XP. I take it that this gives an important clue to why ΣP does not move to the left in Malayalam: Germanic Σ is associated with topichood or switch topic, hence it must move out of the “focus area of the clause.” If we assume with Cinque (1993) that the main stress of the clause falls on the most deeply embedded constituent on the recursive side of the tree, we can make sense of this. I follow Reinhart (1995) in assuming that the focus of a sentence must contain the main stress of the clause. This explains why focused material must extract from ΣP prior to ΣP fronting. The reason why weak pronouns are allowed to stay in ΣP must then be related to the fact that weak pronouns are not good switch topics. Weak pronouns function like discourse variables; they can never cause topic switch, and, as is well known, can never be focused or contrasted (Kayne, 1975). Other material does not have this deficiency. they can be contrasted and focused, and they can also cause topic switch. The fact that all potential topic switchers must evacuate ΣP suggests the following generalization:

Generalization 4 (Unambiguous Topic Switch) *ΣP may contain at most one potential topic switcher.*

Whether this is the right formulation remains to be seen. If it is, it would have to be derivable from more fundamental considerations about topics and topic switch. I will leave that for future investigation.

3.8 Summary

In this chapter, I have argued that there are strong reasons to reject the standard analysis of V2 in terms of head movement to C with subsequent topicalization. I argued that Mainland Scandinavian V2-violations with focus particles

, when seen in conjunction with the behavior of weak pronouns, in particular Holmberg's Generalization, strongly suggest that the verb ends up in the left periphery of the clause by means of an XP-movement, rather than a head movement operation.

On the proposed account, the part of HG concerning weak pronouns is handled by assuming that weak pronouns do not move on their own. They do not trigger VP-movement out of ΣP , hence they must stay inside the fronted ΣP when they can, i.e. when nothing else triggers movement out of ΣP of a VP-node containing them. HG concerning argument shift of full noun phrases follows if verbs that do not move to Σ trigger extraction of their dominating VP-node out of ΣP prior to ΣP -fronting. Such extraction will necessarily take arguments following the verb along. Hence, full DP can "scramble" among sentential adverbs, just when they have not been taken along in ΣP -extraction triggered by a verb. When arguments *can* scramble, they must always end up in the same order as before scrambling. This follows from the assumption that scrambling of arguments in Scandinavian and Dutch is really movement of VP-nodes. Hence, because of the extension condition, one cannot reverse the relative ordering of scrambled arguments in these languages. The possibility in German of such order reversal is due to the fact that arguments in this language can scramble on their own, i.e. DPs rather than VPs scramble in German.

I have argued that no reference needs to be made to specific Topic Phrases and Focus Phrases, as was done in section 3.5. Instead a "dynamic" interpretation of notions like topic and focus, along with default stress assignment rules can drive the movement operations required to derive the observed word-order patterns. This suggests that Last Resort should not be stated in terms of feature checking, which only indirectly affects the interfaces. Rather, it should be stated as interface requirements directly. This has the potential of solving (or at least reducing) a frequently noted problem with 'remnant movement' analyses, namely that it is hard to see how all the movements could be triggered. If the current proposal is on the right track, this suggests that one should not look for formal (uninterpretable) features that trigger each movement operation. Rather, it seems that several operations can be triggered in order to satisfy one interface requirement, such as generalization 4.

CHAPTER 4

Verb movement, Scope and Scrambling

4.1 Introduction

In this chapter, I will argue that the phenomena discussed in the literature under the heading of ‘(Short) Verb Movement’ (SVM) must be seen, in part as a scopal phenomenon, and in part as a phenomenon similar to scrambling as seen with arguments in the Germanic languages. In this way, we can, in principle, extend the approach we pursued in the previous chapters for adverb placement to verb placement. The leading idea is that verbal morphology attaches freely to VPs much in the same way as adverbs and auxiliaries. The relative order with which such elements are stacked onto the VP is then governed by the scopal requirements of the individual expressions. Just what the scopal requirements of a given verbal form would be is often a difficult question, often compounded by the fact that the surface position of an element may be the result of remnant movements of the kind we demonstrated for Norwegian auxiliary sequences which entered into crossing scope interactions with adverbs. Because of this, the analysis presented in this chapter at times leaves much to be said. The idea is not as much to present a fleshed out analysis of SVM, as to point to a novel interpretation of it, and discuss some of the questions that would have to be solved under this view. The idea is illustrated by the tree in figure 4.1. Affixes are different from adverbs in that they are phonologically incomplete, i.e. they are affixes; similarly for bare verb stems. Therefore,

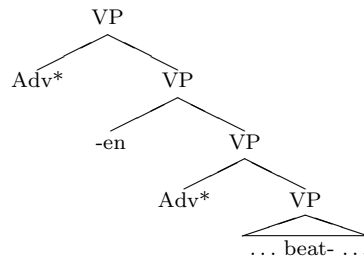


Figure 4.1: SVM without positions

affixes will attract a verb stem for purely phonological reasons (cf. the “stray affix filter”). In this way, Chomsky’s (1995) suggestion that head movement is a PF-phenomenon is partly correct, albeit not in precisely the way Chomsky had in mind. The approach also leads to an understanding of why it should be that SVM is optional within a range of adverbs: Suppose that some range of adverbs a_1, \dots, a_n allows a verb form $v + \text{affix}$ to occupy any position in the sequence. In our approach this means that the affix can be merged above or below each of the adverbs, and then attracts the verb stem for phonological reasons. In other words, we do not have to postulate “optionally strong” features or other mechanisms to accommodate the apparent optionality of SVM, the point being that the movement operation is obligatory, but the attractor can be merged in several “positions”.

The approach also has advantages over the approach to SVM proposed by Svenonius (2001); Ernst (2001), according to which the verb always occupies the functional head T (or some other head) and the adverbs can be merged to spec-TP or spec-VP, subject to s-selectional requirements of the adverbs. This approach also derives the apparent optionality of SVM, but it leads to

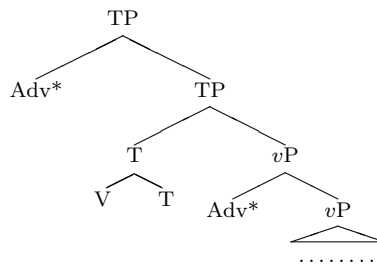


Figure 4.2: Svenonius/Ernst-type SVM

other problems. As discussed in chapter 1, it seems to force us to view T as a semantically vacuous head. Consider how adverb ordering is dealt with in this setup: Some adverbs (e.g. *completely*) semantically select for constituents denoting EVENTS, and the resulting structure [completely XP] also denotes an

event. Other adverbs select for PROPOSITIONS and return propositions (e.g. *not*), while yet others select for FACTS and return facts (e.g. *paradoxically*), etc. The crucial idea is that *v*P_s and TP_s can denote any of these ontological entities, and that events can be turned into propositions, which, in turn can be turned into facts, etc. but not vice versa. Hence, if *paradoxically* is merged to *v*P, this constituent must first be lifted so as to denote a fact, then the adverb can apply to this constituent, resulting in a *v*P denoting a fact. It follows that *completely* or *not* cannot apply after this, because they select for events and propositions, respectively and facts cannot be turned into either of these. Now consider what happens when T is merged to our fact-denoting *v*P. Suppose that T is not semantically vacuous, which seems natural, i.e. it conveys tense. For the Svenonius/Ernst setup to work, T must be remarkably open minded with respect to the semantic denotation of its complement *v*P. Furthermore, whatever *v*P denotes, the result of applying T to this must denote the same. Thus, if *v*P denotes a fact, [T *v*P] must also denote a fact, for, if it were allowed to denote a proposition or an event, then *not* and *completely* would wrongly be predicted to be able to precede and outscope *paradoxically*. Finally, if *not* and *completely* are ever to be allowed to merge to TP, this constituent must be allowed to denote propositions or events in some cases, and a fortiori, *v*P must also denote propositions or events in these cases. Hence, it seems to follow that T cannot have a denotation.

Again, in our system, where semantic selection plays no role, the problem doesn't arise. T is a modifier just like the adverbs, and they can be merged in any order, as long as scopal requirements are satisfied, so T is allowed to have semantic content, as wanted. In other words, we sidestep Svenonius/Ernst's problem by denying the existence of TP as syntactically distinct from *v*P or VP.

Suppose that, instead of their s-selection approach with type-conversion systems, Svenonius/Ernst would adopt something along the lines of the analysis developed in chapter 2 to handle adverb ordering. Then, they could maintain the label TP as distinct from *v*P etc. and assume that T has semantic content. If certain adverbs must precede T when it is specified for, say, past, this can now be treated as a scopal phenomenon, much as in the present chapter. However, in such a theory, the question arises what is the motivation, or even what *could* be the motivation for changing the label from *v*P to TP, except for purely theory internal considerations. In other words, what could be the motivation for claiming that T (or any other expression) occurs in a fixed position when everything else is floating around it: there's nothing to be fixed with respect to.¹ One could try to say that tense is special, because every sentence has tense. But this is certainly not a universal property of natural language, as

¹Incidentally, abandoning TP does not create problems for implementing the insight that nominative case and tense are intimately connected. The connection has never been explained anyway, so we could stipulate the EPP with respect to a "floating" T head just as much as with a fixed one.

(201) a. A quadratic equation always has more than one solution.
 b. Time is branching.
 c. Giraffes have long necks.

In Chapter one, we discussed the ordering paradox pointed out by Bobaljik (1999); Svenonius (2001) for adverb ordering versus argument ordering in languages like Dutch and Norwegian. The problem is can be stated as follows. Given a sequence of arguments A_1, A_2, A_3 , and an adverb a , the adverb can occupy any position in the sequence of arguments as long as the relative ordering of the arguments remains unaltered. Thus, we have the following, where “ \surd ” indicates possible adverb positions among the arguments A_i .

Conversely, given a sequence of adverbs a_1, a_2, a_3 and one argument A , the argument can occupy any position among the adverbs, as long as the ordering of the adverbs remains the unaltered. Thus, we have the following, where “ \sqrt{A} ” represents possible argument positions among the adverbs.

It follows that there cannot be a single linear sequence of functional heads (FSEQ) which accommodates the relative ordering of both kinds of elements. Bobaljik also argues that such a paradox can be found in the case of verb placement. I return to Bobaljik's argument after presenting Cinque's evidence for verb movement.

(202) [mood_{speech-act} frankly [mood_{evaluative} fortunately [mood_{evidential} allegedly
[mod_{epistemic} probably [T_{past} once [T_{future} then [mod_{irrealis}
perhaps [mod_{necessity} necessarily [mod_{possibility} possibly [asp_{habitual} usually
[asp_{repetitive} again [asp_{freq(I)} often [mod_{volutional} intentionally [asp_{celerative(I)}
quickly [T_{anterior} already [asp_{terminative} no longer [asp_{continuative} still
[asp_{perfect(?)} always [asp_{retrospective} just [asp_{proximative} soon [asp_{durative}
briefly [asp_{generic/progressive} characteristically(?) [asp_{prospective} almost
[asp_{sg.completive(I)} completely [asp_{pl.completive} tutto [voice well [asp_{celerative(II)}
fast/early [asp_{repetitive(II)} again [asp_{freq(II)} often [asp_{sg.completive(II)} com-
pletely]]]]]]]]]]]]]]]]]]]]

Cinque argues that an active past participle optionally precedes or follows all the adverbs above *tutto* and below *possibly* in the hierarchy. He illustrates this with the following examples.

- (203)
- a. Da allora, non hanno **rimesso** di solito mica più
since then not the-have put usually non any.longer
sempre completamente tutto bene in ordine.
always completel everything well in order
 - b. Da allora, non hanno di solito **rimesso** mica più sempre completamente tutto bene in ordine.
 - c. Da allora, non hanno di solito mica **rimesso** più sempre completamente tutto bene in ordine.
 - d. Da allora, non hanno di solito mica più **rimesso** sempre completamente tutto bene in ordine.
 - e. Da allora, non hanno di solito mica più sempre **rimesso** completamente tutto bene in ordine.
 - f. Da allora, non hanno di solito mica più sempre completamente **rimesso** tutto bene in ordine.
 - g. *Da allora, non hanno di solito mica più sempre completamente tutto **rimesso** bene in ordine.
 - h. *Da allora, non hanno di solito mica più sempre completamente tutto bene **rimesso** in ordine.

Thus, he shows that the PTC cannot follow the elements, *tutto* ‘everything’ and *bene* ‘well’. He shows that in Logurdesse Sardinian, PTC, can follow *tottu* ‘everything’, but not *bene* ‘well’, and that in French, PTC can follow both *tout* and *bien*. Hence, he argues that the difference between these languages can be captured by stating that PTC must move to *AspPlCompletive* in Standard Italian, whereas in Logurdesse Sardinian, it only has to move to the *Voice* head, and in French, it may stay even lower. Movement to the higher heads is then optional. Cinque also shows that finite verbs can precede or follow high adverbs. His examples are given below.

- (204)
- a. **Mi ero** francamente purtroppo evidentemente formato
me was frankly unfortunately evidently formed
una pessima opinione di voi.
a bad opinion of you
 - b. Francamente **mi ero** purtroppo evidentemente formato una pessima opinione di voi.
 - c. Francamente purtroppo **mi ero** evidentemente formato una pessima opinione di voi.
 - d. Francamente purtroppo evidentemente **mi ero** formato una pessima opinione di voi.

- (205) a. Evidentemente **mi ero** probabilmente allora formato una
 evidently me was probably then formed a
 pessima opinione di voi.
 bad opinion of you
- b. Evidentemente probabilmente **mi ero** allora formato una pes-
 sima opinione di voi.
- c. Evidentemente probabilmente allora **mi ero** formato una pes-
 sima opinione di voi.
- (206) a. Allora **aveva** forse saggiamente deciso di non presentarsi
 Then he-had maybe wisely decided to not go
- b. Allora forse **aveva** saggiamente deciso di non presentarsi
- c. Allora forse saggiamente **aveva** deciso di non presentarsi

The question now arises whether there is an overlap in the range of finite verb movement and participle movement. If there is, we have a problem. Bobaljik (1999) argues that Cinque, in fact, has data which demonstrates such an overlap. In particular (203a-203b) show that PTC can precede both *mica* ‘not_{emphatic}’ and *di solito* ‘usually’. Cinque also shows (pp. 50-51) that Vf can follow *mica*. He argues that this is not due to a higher position of the adverb, but, rather movement of Vf around *mica*.²

- (207) Gianni (*non) **mica** (*non) gli **telefonerà**.
 G (*not) not (*not) to-him will-tepephone

Below is an example I found on the internet, where the Vf follows *di solito* ‘usually’ as well.

- (208) Aumenti di dosatura **di solito sono** molto rapidamente effettivi,
 increases in dosage usually are very quickly effective
 spesso in un giorno.
 often in a day

Thus, we have that *di solito* must precede *mica*. Vf must precede PTC. Otherwise, the two pairs can be combined in all and only the orderings in (209)

- (209) a. Vf PTC di solito mica
 b. Vf di solito PTC mica
 c. Vf di solito mica PTC
 d. di solito Vf PTC mica
 e. di solito Vf mica PTC

²*Mica* is an emphatic negation which, when it follows Vf, must cooccur with the sentential negation *non*. As (207) shows, it cannot cooccur with sentential negation when it precedes Vf. Hence it seems that negative concord, perhaps more generally, is restricted to “negative” sentences where Vf would not be c-commanded by a negative element in the absence of *non*.

f. di solito mica Vf PTC

Hence, there is no linear sequence of positions that can accommodate all the orderings in (209) without also allowing ungrammatical orders, i.e. orderings where Vf follows PTC, or *mica* precedes *di solito*. As briefly discussed by Bobaljik (1999), Cinque could solve the problem by stipulating that multiple verb movements must preserve the ordering of the verbs. This is similar to the order preservation principle adopted by Müller (2002) for the massive evacuation of *vP* he assumes (see chapter 3.6.2 for discussion). But such order preservation would be a curious stipulation in need of explanation. In other words, we would need an explanation why Vf needs to precede PTC anyway. One might suspect, then, that this explanation, in conjunction with an approach to adverb ordering along the lines of chapter 2 would suffice to explain the limitation to the orderings in (209).

Just how high can the PTC raise? Cinque states that PTC do not raise across “higher” adverbs, but he does not give examples of ungrammatical sentences. I searched the internet (google) for exact strings like “*avuto fortunatamente*” ‘have-PTC fortunately’, *stato probabilmente* ‘been probably’. Below are some of the examples I found.

- (210)
- a. Due incendi che non hanno **avuto fortunatamente**
Two fires that not have-3PL had fortunately
conseguenze rilevanti si sono sviluppati
consequences relevant SI are developed
 - b. le analisi hanno **dato fortunatamente** esito
the analyses have-3PL had fortunately output
negativo
negative
 - c. è stato probabilmente stampato a Roma
is-3SG been probably printed in Rome

The native speakers I have consulted report that these sentences are grammatical, and do not require a ‘comma intonation’ on the adverbs.³ It seems, then that the Italian PTC can precede very “high” adverbs. It could be that our examples in (210) are derived by moving Vf and PTC around the adverb as a remnant XP. In other words, it could turn out that the correct constituency for these examples is the one given in figure 4.3. Such an analysis is actually supported by the fact that the relevant examples are reported to be ungrammatical if the two verbs are separated by a “high” adverb. Thus, (211) is reported to be bad.

- (211) È purtroppo stampato probabilmente a Roma.
is-3SG unfortunately printed probably in Rome

³If they did, the examples could be disregarded. Comma intonation is known to license all sorts of adverb orders (Jackendoff, 1972).

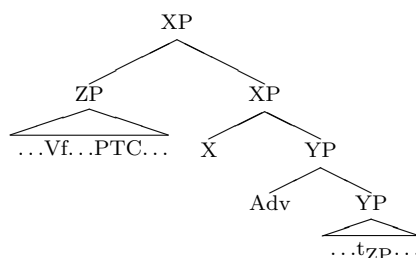


Figure 4.3: Remnant movement analysis for (210)

Given this, I tentatively conclude that movement of the participle around high adverbs is indeed impossible. I return below to a suggestion as to why this should be so.

4.3 English and French

Pollock (1989) notes that English differs from French in disallowing verb movement of finite lexical verbs around adverbs, while allowing it for *be* (and for some varieties) *have*. He gives examples like the following (his (17-19)). (212c) is possible in British English, but generally unavailable in American English.⁴ In French, such verb movement is even obligatory (213).

- (212)
- a. John is not happy.
 - b. *John seems not happy.
 - c. John hasn't a car.
 - d. *John owns not a car.
 - e. John has often kissed Mary.
 - f. *John kisses often Mary.
 - g. John often kisses Mary.
- (213)
- a. Jean (n') aime pas Marie.
J (NE) loves not M
 - b. Il est rarement satisfait.
He is rarely satisfied
 - c. *Jean ne pas aime Maire.
J NE not loves M
 - d. *Jean souvent embrasse Marie.
J often kisses M

⁴Note, incidentally, that the ungrammaticality of examples like (212b) shows that the negation cannot modify the adjective *happy* directly; similarly for *often* and *sarcastic* in (214b).

- e. Jean embrasse souvent Marie.
 J kisses often M

He also notes that English exhibits a similar restriction on SVM of infinitives as illustrated in (214) (Pollock, 1989, p382), while French again allows this across the board.

- (214) a. (?) I believe John to be often sarcastic
 b. *To look often sad during ones honeymoon is rare.

Pollock points out that one obvious difference between lexical verbs and auxiliary have/be is that, whereas the former assign θ -roles the latter do not. Hence, he suggests that verb movement, in English, though not in French, has the property of blocking θ -assignment. He implements this idea by assuming that his functional head Agr is opaque for theta-assignment in English (though not in French). Thus, if the verb moves to Agr, as shown in figure 4.4, whether or not it moves further to T, it cannot assign a theta role to its arguments inside the VP. While I agree with Pollock that the restriction to these verbs should

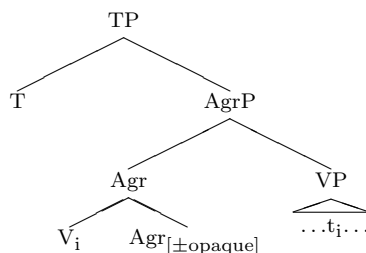


Figure 4.4: Pollock's explanation for the restriction to *have/be*

be stated in terms of the relation between a θ -assigner and its arguments, I obviously cannot adopt his execution of the idea. An alternative to his opaque Agr analysis would be to blame the intervening adverb for the blocking effect. In other words, rather than stipulating that English Agr is opaque, we could stipulate that English theta assigners must surface adjacent to (the trace of) their arguments (Stowell, 1981).⁵ As Pollock points out, it is crucial that this adjacency requirement be stated on *all* arguments of the verb, so as to block verb movement ([V adv] orders) for all verbs, including unergatives. Just as with Pollock's opacity of certain kinds of Agr, we would hope that this adjacency requirement can ultimately be reduced to more primitive notions. I will treat adjacency as a requirement on PF. I will treat English sentences like

⁵A problem for this formulation, pointed out to me by Richard Kayne (p.c.) is that English verbs *can* be separated from their arguments by verbal particles, as in *let out the dog*. The particle does seem to project its own phonological word. One way out would be to adopt a small-clause analysis for particle constructions and claim that the theta relation holds between the verb and the [Prt DP] constituent.

(212f) in the following way. We base generate the sequence [often kiss- Mary] in the standard way. Now, we merge the affix *-s* which is phonologically complete, i.e. it needs to be adjacent to a host (the verb) which does not end in a word boundary. Hence, we could generate the configuration in figure 4.5 if it had not been for the fact that the verb must be adjacent to the object. Thus, this is what happens in French. In English, we must find a way to satisfy both the stray affix filter and the adjacency requirement. It seems that the way to do this is to split up the lowest VP constituent by *Mary* prior to movement of the entire complement of the affix around it. The derivation is given in derivation 4.1 What is the sense in which the verb and *Mary* are adjacent in the bottom

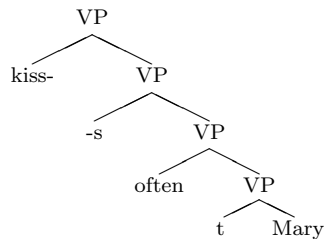


Figure 4.5: verb/affix/adverb placement

Derivation 4.1

[often kiss- Mary]

MOVE *Mary* ►

[Mary_i [often kiss- t_i]]

MERGE *-s* ►

[-s [Mary_i [often kiss- t_i]]]

MOVE *often kiss* ►

[[often kiss t_i]_j [-s [Mary_i t_j]]]

line of derivation 4.1? In other words, how does this count as “adjacent”, while the configuration in figure 4.5 does not? The question can be rephrased as follows: why is V adjacent to O in the following configuration just in case X is an affix?

(215) V X O

It seems obvious that this calls for a phonological explanation. An affix by definition phonologically incorporates into the preceding stem, resulting in a single phonological word. In other words, adjacency should be stated as follows:

Generalization 5 (Adjacency) *No phonological word may intervene between X and Y.*

If the adverb in derivation 4.1 could be merged above tense, we would have the simpler derivation in derivation 4.2. I assume that this is done for other adverbs, but, as we shall see below, there are reasons to assume that frequency adverbs like *often* cannot be merged above tense. Later on, I will try to derive this restriction from the semantics of the adverbs.

Derivation 4.2

[-s [kiss- Mary]]

MOVE *kiss* and MERGE *often* ►

[often [kiss_i [-s [t_i Mary]]]]

I treat auxiliaries like *have* as modifiers of the VP, just like the adverbs and the affixes. Movement of such auxiliaries is not subject to adjacency, because the auxiliaries do not assign θ -roles (Pollock, 1989). Thus, the derivation of (216) is given in derivation 4.3.

(216) John has often kissed Mary.

Derivation 4.3

[kiss Mary]

MERGE *-ed* and *move kiss* ►

[kiss_i [-ed [t_i Mary]]]

MERGE *have* ►

[have [kiss_i [-ed [t_i Mary]]]]

merge often ►

[often [have [kiss_i [-ed [t_i Mary]]]]]

MERGE T and MOVE *have* ►

[have_j [-s [t_j [often [kiss_i [-ed [t_i Mary]]]]]]]

Orders where the auxiliary follows the frequency adverb is also possible, especially if the auxiliary is stressed, leading to a “verum focus” interpretation (217).

(217) John often HAS kissed Mary.

I take the stress requirement to indicate that this is a marked option. Hence, I take them to be derived by extracting the lowest VP [has kissed Mary] in the

fashion of derivation 4.1. This is illustrated in derivation 4.4. Crucially, tense is still merged after the frequency adverb.

Derivation 4.4

$$\begin{array}{l}
 [\text{VP}_2 \text{ often } [\text{have } [\text{VP}_1 \text{ kiss}_i \text{ [-ed } [t_i \text{ Mary}]]]]] \\
 \text{MOVE VP}_1 \blacktriangleright \\
 [[\text{VP}_1 \text{ kiss}_i \text{ [-ed } [t_i \text{ Mary}]]]_j [\text{VP}_2 \text{ often } [\text{have } t_j]]] \\
 \text{MERGE T and MOVE VP}_2 \blacktriangleright \\
 [[\text{VP}_2 \text{ often } [\text{have } t_j]]_k \text{ [-s } [[\text{VP}_1 \text{ kiss}_i \text{ [-ed } [t_i \text{ Mary}]]]_j [t_k]]]]
 \end{array}$$

The range of possible generation sites for finite T is limited. This is illustrated by contrasts like the following, where, as before I take the “low” position of the main verb in (218) (cf. Jackendoff (1972); Pollock (1989)) to reflect remnant movement of [completely V t] around the affix.

- (218)
- a. John completely lost his mind
 - b. *John completely will lose his mind
 - c. John will completely lose his mind.
 - d. *John completely is losing his mind.
 - e. John is completely losing his mind.
 - f. *John completely has lost his mind.
 - g. John has completely lost his mind.

Similarly, while speakers allow both (219a,219b), they find (219b) marked, (219a) being the neutral option.

- (219)
- a. John has always done his homework.
 - b. ? John always has done his homework.

We thus have to account for the complete ungrammaticality of adverbs like *completely* in the pre-aux position contra the relative markedness of frequency adverbs in this position. Moreover, we must account for the markedness of *always aux_f* contra the unmarkedness of *aux_f always*.

The impossibility of (218f) would follow if not only T, but also aux must merge outside of the adverb. This could be made to follow from the fact that *completely* is an aspectual adverb, applying to ‘inner’ (or predicate) aspect in the sense of Verkuyl (1993); Borik (2002). *Completely* can only apply to gradable predicates for which there is a well-defined maximal degree (Doetjes et al., 1998). Thus, (220a) is odd, because there is no well defined maximal degree of snoring. There are degrees of it, though, as seen with *John snored too much/very much/more than Bill*. There is, however, a maximal degree of *recovering*; hence (220b) is good.

- (220) a. John (*completely) snored (*completely).
 b. John (completely) recovered (completely).

I assume that such “maximal” degrees can be implemented in terms of a Verkuyl (1993) style path structure. In other words, *completely* can apply to a predicate only if it denotes a path with a termination point, beyond which there can be no more change of the relevant kind.⁶ Predicates denoting terminated path structures are “telic” in the sense of being non-homogeneous for the reason that no proper subpart of a terminated path is itself a terminated path (see Verkuyl (1993); Borik (2002)). It now suffices to note that [aux VP] is always homogeneous, regardless of the aspectual properties of the VP. In other words, if the predicate [have recovered] holds of John at interval *i*, then, presumably it holds of John at every subinterval *i'* of *i*. Hence this larger constituent is not telic, i.e. does not denote a terminated path, and *completely* cannot apply to it. In other words, the only possible order of merger would be [have[completely VP]], and this, in turn, requires the modified VP to be telic. Along with the adjacency requirement on lexical verbs, this derives the fact that lexical verbs can follow *completely*, while auxiliaries can't. I show the derivation of (220b) in derivation 4.5.

Derivation 4.5

[recover]

MERGE *completely* ►

[completely [recover]]

MERGE T and MOVE recover ►

[recover_i [-ed [completely t_i]]]

For sentences like (218g), the correct derivation depends on whether *completely* can apply outside of participial morphology or not. For concreteness, I assume that it can't, although I do not know of any evidence from word order in either direction. Given that, and the adjacency requirement, we get the derivation in derivation 4.6 for (218f).⁷ Of course, if *completely* can merge outside of PTC, we get a simpler derivation, where the main verb moves to the

⁶This only partially characterizes the behavior of *completely*: in English, (i) is not a good sentence, even though [go to the store] clearly does denote a terminating path in the relevant sense. Interestingly, the Norwegian translation of *completely*, i.e. *helt* can modify the corresponding predicate, as seen in (ii).

(i) *John went completely to the store.

(ii) Jens gikk helt til butikken (Norwegian, same)

Similarly, *completely* (and Norwegian *helt* seems to require the presence of a verbal particle in some cases, e.g.

(iii) John has eaten it ?(up) completely.

⁷The fact that the progressive in (218e) seems to obligatorily outscope the adverb, suggests, that, at least in this example, the adverb must attach before the affix.

left of PTC without prior extraction of the object. In either case, the auxiliary will necessarily end up preceding the adverb.

Derivation 4.6

[loose his mind]
MERGE *completely* and MOVE *his mind* ►
[[his mind]_i [completely [loose t_i]]]
MERGE PTC and *move completely loose* ►
[[completely loose t_i]_j [-ed [[his mind]_i t_j]]]
MERGE *have* ►
[have [[completely loose t_i]_j [-ed [[his mind]_i t_j]]]]
MERGE PRES and MOVE *have* ►
[have_k [-s [t_k [[completely loose t_i]_j [-ed [[his mind]_i t_j]]]]]]

Why is movement of Vf around frequency adverbs obligatory in French, and not just marked? It is standardly assumed that Vf is in the highest position within “IP” in French. Hence it seems that French is, in some sense V2. This is further supported by the “stylistic” inversion cases discussed recently by Kayne and Pollock (1999), whereby one gets obligatory subject–verb inversion triggered by fronting of a non-subject.

- (221) A qui a téléphoné ton ami?
to whom has telephoned your friend

Kayne and Pollock (1999) argue that this does not involve a low position of the subject, but rather, extraction of the subject out of IP, with subsequent fronting of IP around it. Let us assume that French follows the derivational pattern in derivation 3.27 in the previous chapter. In other words, it allows adverbs to occur interspersed between the auxiliaries. Suppose, furthermore that French also has a Σ head, which normally attracts the subject to its specifier. Then, it could be that stylistic inversion involves movement of some non-canonical XP to spec- Σ , and that this triggers extraction of the subject from Σ P and subsequent raising of Σ P around the subject, much as what happens in the Germanic V2 languages. In other words, I adopt the analysis of Kayne and Pollock (1999), just stated in the terminology of my chapter 3. But now we can understand why the finite verb must precede adverbs in French. It is because Σ P attracts it, and that, as we have argued, Σ P must precede adverbs. Thus the difference between French and English lies in the attraction of Vf to Σ in French, versus absence of such attraction in English. Whether extraction of non-clitic material from Σ is triggered in French in the same way as Germanic now depends on whether Σ is categorized with the same features as Germanic Σ . In other words,

extraction from ΣP was argued to be triggered by the generalization that ΣP could only contain one potential topic switch. But if French Σ does not encode topic switch, no extraction is predicted. These considerations suggest that verb movement in French is really ΣP fronting. Hence, (213e) is derived as in derivation (4.7). Now, the obligatoriness of $Vf > adv$ orders is expected in

Derivation 4.7

[Σ [Jean embrasse Marie]]	MOVE <i>Vf</i> and <i>Jean</i> ►
[Jean _i [embrasse _j + Σ [t _i t _j Marie]]]	
	MOVE <i>Marie</i> ►
[Marie _k [Jean _i [embrasse _j + Σ [t _i t _j t _k]]]]	
	MERGE <i>souvent</i> and MOVE ΣP ►
[[Jean _i [embrasse _j + Σ [t _i t _j t _k]]] _I [souvent [Marie _k]]]	

French. It does not indicate the presence of a high functional head attracting the verb. In fact, it is not the verb itself that moves around the adverb, but rather a bigger constituent containing it.

For English, it would seem that auxiliary movement around adverbs is similar to scrambling of arguments in other languages. This is actually supported by the fact that it has information structural effects. Consider the pattern in (222).

- (222)
- a. John's always done his homework.
 - b. * John always's done his homework.
 - c. John has always done his homework.
 - d. ? John always has done his homework.
 - e. John always HAS done his homework.

A clitic auxiliary is completely ungrammatical following an adverb. With non-clitic, unstressed auxiliaries, the $Adv > Aux$ order is possible, but marked, whereas with a stressed auxiliary, such orders are perfect. I take this to reflect the fact that the auxiliary generally avoids stress, perhaps because it is semantically quite bleached. The point here is that this is highly reminiscent of the patterns holding of scrambling phenomena. A second observation supporting the scrambling interpretation of English auxiliary movement is that it may have scopal effects. Thus, (223a) was found on the internet. The adverb seems to outscope the epistemic auxiliary, and similarly for (223b), also found on the internet.

- (223)
- a. It is safe to conclude that mariners' thorough knowledge of the river **always must** have been essential for a safe passage.

- b. Damages caused by these items can sometimes be extensive and costly to repair. It **often might** result in the total refinishing of the top surface.

This suggests the following generalization 6 and, perhaps, more surprisingly, generalization 7.

Generalization 6 (Optional Verb Movement) *Optional verb movement around adverbs is a scrambling phenomenon.*

Generalization 7 (Obligatory Verb Movement) *Obligatory verb movement around adverbs does not exist.*

In cases where the verb seems to move obligatorily around adverbs, such as in Germanic and French, it is something else, containing the verb which has moved. In other words, obligatory verb movement and obligatory pronoun shift can both be reduced to fronting of Σ .

4.4 Italian verb scrambling and VP scrambling

(Cinque, 1999, p31) notes that the Italian Vf cannot precede high adverbs unless some other material follows them, or they occur in a so-called comma-reading. This is illustrated with the contrast between (224a) and (224bc) (Cinque, op. cit.).

- (224) a. * Gianni lo merita [francamente/ fortunatamente/
G it deserves [frankly/ fortunately/
evidentemente/ probabilmente/ forse].
evidently/ probably/ perhaps]
- b. Gianni lo merita, [francamente/ fortunatamente/
G it deserves [frankly/ fortunately/
evidentemente/ probabilmente/ forse].
evidently/ probably/ perhaps]
- c. Gianni lo merita [francamente/ fortunatamente/
G it deserves [frankly/ fortunately/
evidentemente/ probabilmente/ forse] per più di una
evidently/ probably/ perhaps] for more than one
ragione.
reason

I take this to reflect the following situation. Finite tense can be merged above or below high adverbs. If it is merged above, the Vf will occur to the left of the adverbs, because it is attracted by the affix. But since the high adverbs are stress avoiders, this leads to the situation where they must be read with a comma-intonation in case there is nothing further down in the clause which

can receive the default stress (cf. Cinque (1993)). If there is lower material, nothing special needs to be done with the adverb.

“Lower” adverbs are not stress avoiders, and Cinque shows that these actually do participate in stress-driven movements (pp. 13-16). For example, he notes that *ancora* and other low adverbs can follow the entire VP, sometimes requiring a slight pause. This is illustrated with *di già* ‘already’ in (225) (Cinque, 1999, p14).

- (225) Gianni ha ricevuto la notizia DI GIÀ.
 G has received the news already

This is interpreted as movement of the VP (or a somewhat larger constituent) around the adverb, rendering the adverb the lowest element on the recursive side of the tree, thus the recipient of prosodic stress. It is striking that the domain for such VP scrambling out of the focus domain has exactly the same range as movement of the participial verb. In other words, it seems that the Italian and French PTC can scramble around all and only adverbs which are not stress avoiders.

4.5 Summary

This concludes my tentative discussion of (short) verb movement phenomena. I have suggested that apparent cases of obligatory verb movement should be treated as pied piping by a larger constituent, and hence, that it does not exist. Optional verb movement as observed with English auxiliaries, and participles in French and Italian, as well as ordinary finite verbs in Italian is treated, in part, as base generation of the verbal morphology above or below the adverbs in question, and in part as a scrambling (i.e. stress/focus driven) operation. For languages where all verb forms follow all adverbs, as is apparently the case in Spanish, but also Germanic, modulo V2, an analysis in the sort proposed in the previous chapter (derivation 3.26 is assumed. Hence such languages are predicted to exhibit crossing scope dependencies between adverbs and verbs.

The present chapter leaves the following hard questions to be solved. Why must finite tense apparently be merged above certain adverbs? Why can’t participial morphology be merged above stress avoiding adverbs? I do not have an answer to these questions at present, but I think one should try to address them before rejecting an account along the present lines.

If correct, the analysis presented here suggests that very few distributional phenomena, if any at all, should be handled in terms of selectional sequences of functional heads. This does not mean that there cannot be any “functional” heads. What I have argued against in this thesis is to order them by (syntactic) selection.

APPENDIX A

Degrees and SOA

In this appendix, I will argue that there is independent evidence for the analysis of speaker-oriented adverbs (SOA) proposed in Chapter 2.

A.1 The *veldig* ~ *vis* Competition

Unlike English, Norwegian does not allow degree modified sentence adverbs, like (226).

- (226) a. *Jens har veldig sannsynligvis gått hjem.
 J has very probably gone home.
 b. *Jens har ganske muligens gått hjem.
 J has quite possibly gone home
 c. *Jens er helt tydeligvis ikke forbryteren.
 J is completely evidently not the-perpetrator

In some cases, it is possible to circumvent this problem by omitting the derivational suffix *-vis*.

- (227) a. Jens har veldig sannsynlig gått hjem.
 J has very probable gone home
 b. ?Jens har veldig mulig gått hjem.
 J has very possible gone home
 c. Jens er helt tydelig ikke forbryteren.
 J is completely evident not the-perpetrator

This gives the impression that “bare” adjective phrases can be used as sentence adverbials in Norwegian. However, if the adjective is not degree modified, the result is ungrammatical, i.e. in this case, *-vis* is obligatory.¹

- (228) a. Jens har sannsynlig*(-vis) gått hjem
 J has probable gone home
 b. Jens har mulig*(-ens) gått hjem.
 J has possible gone home
 c. Jens er tydelig*(-vis) ikke forbryteren.
 J is evident not the-perpetrator

In other words, *-vis* and degree modification are in complementary distribution.

A.1.1 A morphological analysis?

One way one may want to think about this is to say that the derivational suffix *-vis* turns the adjectives into adverbs presyntactically, and that the degree modifiers *veldig* ‘very’, *ganske* ‘pretty’ *helt* ‘completely’ syntactically select for adjectives.² I will refer to this as the “morphological analysis” (MA). It is compatible with the fact that *adj+vis* cannot occur in other positions where adjectives typically occur, e.g. predicative position, and adnominal position, i.e. they are not adjectives.

- (229) a. *Det er tydeligvis at Jens er forbryteren.
 it is evidently that J is the-perpetrator
 b. *Den mest sannsynligvise løsningen er at Jens er
 the most probably-DEF the-solution is that J is
 forbryteren.
 the-perpetrator

However, it leaves open some other questions, like the following ones: 1) why can “bare” APs occur in sentence-adverbial position just in case they are degree modified? 2) Why are English sentence adverbs ending in *-ly* compatible with degree modification (230), while Norwegian *-vis* adverbs are not? The fact that English does allow for degree modified *-ly* adverbs actually suggests that the phenomenon cannot be reduced to the status of a modifier as an adjective/adverb. English *-ly* adverbs are also ungrammatical in predicative and adnominal position so they are not adjectives.

- (230) a. John has very probably gone home.
 b. John has quite possibly gone home.
 c. John has very evidently gone home.

¹the ending *-ens* on *mulig* ‘possible’ is irregular: I take it to be an allomorph of *-vis*.

²On the selectional properties of *very* etc., see Corver (1997); Doetjes et al. (1998); Kayne (2002).

- (231) a. *It is evidently that John has gone home.
 b. *A probably solution is that John is the perpetrator.

The worst problem for MA is that some of the degree modifiers incompatible with *-vis* have a relatively free distribution, syntactically. Corver (1997); Doetjes et al. (1998) argue that one should distinguish two classes of degree modifiers, namely those that select for adjectives, and those that do not. English *very* is in the first class (class 1), whereas *more* is in class 2. Thus, *more*, but not *very* can modify nouns (232a), verb phrases (232b)-(232c) and prepositional phrases (232d). Hence, it seems that *more* is capable of modifying any syntactic category.

- (232) a. Stanley has [more/*very] money (than Bill).
 b. Stanley [[likes her] [more/*very]]
 c. *Stanley very [likes her].
 d. Stanley is [more/*very] into syntax.

The examples with *very* would be rescued by insertion of the dummy *much* (Corver, 1997), i.e. *very much*. In (233), we see that the distribution of *veldig* ‘very’ is something in between English *very* and *more*, i.e. *veldig* can to some extent modify verb phrases and PPs, but not nouns. Norwegian *mer* ‘more’ behaves like its English cognate.

- (233) a. Ståle har [mer/ *veldig] penger (enn Willy).
 S has [more/ *very] money (than W)
 b. Ståle liker henne [mer/ veldig ?(mye)]
 S likes her [more/ very ?(much)]
 c. *... at Ståle veldig liker henne
 ... that S very likes her
 d. Ståle er [mer/ veldig (mye)] mot hvalfangst.
 S is [more/ very (much)] against whale-hunting

If MA is to have any plausibility, there should be some clear cases of class 2 degree modifiers that can modify adverbs ending on *-vis*, since these are indifferent to the morpho-syntactic class of the category they modify. There are none.³

- (234) a. *Ståle har [veldig/ altfor/ litt/ mer/ nok/ tilstrekkelig/
 S has [very/ all-too/ a.little/ more/ enough/ sufficiently/
 utrolig/ kjempe-/ ...] sannsynligvis gått hjem.
 incredibly/ giant-/ ...] probably gone home

³In colloquial Norwegian, the noun *kjempe* ‘giant’ productively acts as a prefixal degree modifier to adjectives, as in *kjempe-stor* ‘giant-big’ (‘very big’), *kjempe-smart* ‘giant-smart’ (‘very smart’), etc.

- b. Ståle har [veldig/ altfor/ ?litt/ mer/ ?nok/
 S has [very/ all-too/ ?a.little/ more/ enough/
 tilstrekkelig/ utrolig/ ?kjempe-/ ...] **sannsynlig** gått
 sufficiently/ incredibly/ ?giant-/ ...] probable gone
 hjem.
 home

Even *mer* ‘more’, which is a clear class 2 degree adverb, and certainly can modify the adjective *sannsynlig* ‘probable’ (i.e. there shouldn’t be any semantic incompatibility), is strongly incompatible with *-vis*. In short, it does not seem to be the case that the complementarity of *veldig* and *-vis* is reducible to the s-selectional properties of the former and the morphosyntactic category of the latter.

A.1.2 DegPs and sentence adverbs

Having dismissed MA, let us have a second look at the Norwegian facts. The contrast between (227) and (228) indicates that degree modification (of *sannsynlig* and *tydelig*; we return to *mulig*) suffices to turn these adjectives into sentence adverbs. The fact that the suffix *-vis* is capable of turning an adjective into an adverb, i.e. without extra degree modification, in addition to the complementary distribution between *-vis* and other degree modifiers, suggests that *-vis* itself is a degree modifier.⁴ In this way, the facts in (226)-(228) support the following generalization:

Generalization 8 (Norwegian) *Sentence adverbs are degree modified adjectives.*

Let us now turn to some facts that support this generalization. A very productive way of turning adjective into “evaluative” adverbs, is to add the post-adjectival degree modifier *nok* ‘enough’.⁵ Below is a non-exhaustive list of some adverbs formed in this way found in a search in the Oslo Corpus of Tagged Norwegian Text (OCTN).

⁴ftn4w-*vis* appears on some adverbs, like *delvis* ‘partly’ which are not sentence adverbs on their most common use. It seems that when *-vis* attaches to an adjective ending on *-(l)ig*, the result is invariably a sentence modifier.

⁵see Barbiers (2001) on this phenomenon in Dutch with *genoeg* ‘enough’.

(235)	beklagelig nok	regettable enough	‘regrettably’
	besnærende nok	captivating enough	‘captivatingly’
	ironisk nok	ironical enough	‘ironically’
	merkelig nok	strange enough	‘strangely’
	mirakuløst nok	miraculous enough	‘miraculously’
	naturlig nok	natural enough	‘naturally’
	overraskende nok	surprising enough	‘surprisingly’
	paradoksalt nok	paradoxical enough	‘paradoxically’
	pussig nok	funny enough	‘funnily enough’
	rettferdig nok	fair enough	‘fairly enough’
	rett nok	right enough	‘true enough’
	rimelig nok	reasonable enough	‘reasonably’
	sant nok	true enough	‘true enough’
	typisk nok	typical enough	‘typically enough’
	utrolig nok	unbelievable enough	‘unbelievably’

In fact, it seems that this is *the* productive way of making evaluatives. Note for example that most of the examples in (236) are incompatible with this ending *-vis*, even the ones that end in *-lig*, i.e. **merkeligvis* ‘strangely’, **rimeligvis* ‘reasonably’, **utroligvis* ‘unbelievably’, etc. On the other hand, evaluatives which have forms ending on *-vis*, like *heldigvis* ‘fortunately’, can occur with *nok*, as in *heldig nok* ‘fortunately’ (lit. ‘fortunate enough’).

I illustrate the use of these “adverbs” in (236). Note that they are always “evaluative”, i.e. they impose some evaluation of the asserted proposition on the part of the speaker.⁶

- (236) a. på skolen gikk det naturlig nok til helvete.
on school went it natural enough to hell
- b. Utrolig nok føk pucken inn bak Steve Allman i
incredible enough darted the-puck in behind S A in
gjestenes mål
the-guests’ goal
- c. Ironisk nok har hun det best når mamma og pappa
ironic enough has she it best when mum and dad
forsvinner til Syden på ferie.
vanish to the-south on vacation

I take this to support the analysis of SOA proposed in chapter 2. There I claimed that adverbs like *possibly* differ from the corresponding adjective *possible* in that the former give rise to stronger statements than the latter. This, in turn, was implemented by applying a domain-shrinking function to the epistemic accessibility relation inherent to the meaning of the adjective. From the perspective of the present appendix, it seems that the domain shrink

⁶The examples are also from the OCTN.

is, in fact, the ending *-vis*, and that domain shrinkage should be related to degree modification, perhaps more generally.

Bibliography

- Ackema, Peter, and Ad Neeleman. 2002. Effects of short-term storage in processing rightward movement. In *Storage and computation in the language faculty*, ed. S. Nootboom, F. Weerman, and F. Wijnen, 219–256. Dordrecht: Kluwer. 2, 19, 21, 23, 26, 28
- Åfarli, Tor A. 1996. Dimensions of phrase structure: the representation of sentence adverbials. Ms. University of Trondheim. 6, 12
- Alexiadou, Artemis. 1997. *Adverb placement a case study in antisymmetric syntax*. Amsterdam: John Benjamins. 37, 67
- Alexiadou, Artemis. 2001. On the status of adverb in a grammar without a lexicon. Ms. University of Stuttgart. 37
- Barbiers, Sjef. 1995. The syntax of interpretation. Doctoral Dissertation, Leiden University. 31, 35
- Barbiers, Sjef. 2001. Is vreemd genoeg genoeg? In *Kerven in een rots*, ed. B. Dongelmans et al., 15–28. Stichting Neerlandistiek Leiden Reeks 7. 144
- Bartsch, Renate. 1976. *The grammar of adverbials*. North-Holland: Elsevier. 72
- Beaver, David, and Brady Clark. 2002. Always and only – why not all focus sensitive operators are alike. Ms. Stanford. 50, 63
- Beghelli, Filippo, and Timothy Stowell. 1997. Distributivity and negation. In *Ways of scope taking*, ed. Anna Szabolcsi, 71–109. Dordrecht: Kluwer. 11, 50, 68

- Bellert, Irena. 1977. On semantic and distributional properties of sentential adverbs. *Linguistic Inquiry* 8:337–51. 38, 39
- Bentzen, Kristine. 2002. Independent V-to-I movement without morphological clues. Handout from paper presented at Grammatik i Fokus, Lund, 2002., Feb 2002. 72, 118
- Bernardi, Raffaella. 2002. Reasoning with polarity in categorial type logic. Doctoral Dissertation, Utrecht University. 6, 42
- den Besten, Hans. 1989. *Studies in Western Germanic syntax*. Amsterdam: Rodopi. 77
- Bianchi, Valentina. 1995. Consequences of antisymmetry for the syntax of headed relative clauses. Doctoral Dissertation, Scuola Normale Superiore, Pisa. 34
- Bobaljik, Jonathan. 1999. Adverbs: the hierarchy paradox. *Glott International* 4:27–28. 11, 12, 14, 71, 113, 126, 128, 129
- Bobaljik, Jonathan, and Samuel Brown. 1997. Interarboreal operations: Head movement and the extension requirement. *Linguistic Inquiry* 28:345–356. 108
- Borik, Olga. 2002. Aspect and reference time. Doctoral Dissertation, Utrecht institute of Linguistics OTS. 70, 134, 135
- Brody, Michael. 2000. Mirror theory. *Linguistic Inquiry* 31:29–56. 4
- Cardinaletti, Anna, and Michal Starke. 1995. The typology of structural deficiency. on the three grammatical classes. *ZAS Papers in Linguistics* 1:1–55. 81
- Chierchia, Genarro. 2001. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. Ms. University of Milan – Bicocca. 2, 38, 51, 60, 61, 62, 157
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Massachusetts: MIT Press. 67
- Chomsky, Noam. 1994. Bare phrase structure. Ms. MIT. 4, 20
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Massachusetts: MIT Press. 1, 82, 106
- Chomsky, Noam. 1999. Derivation by phase. Ms. MIT. 1, 60, 102
- Chomsky, Noam. 2001. Beyond explanatory adequacy. Ms. MIT. 1, 29, 60, 102

- Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24:239–297. 121, 139
- Cinque, Guglielmo. 1999. *Adverbs and functional heads – a crosslinguistic perspective*. Oxford: Oxford University Press. 3, 7, 18, 21, 29, 37, 38, 44, 48, 67, 68, 70, 78, 80, 138, 139, 155
- Cinque, Guglielmo. 2000a. On Greenberg’s U20 and the Semitic DP. Ms. University of Venice. 69
- Cinque, Guglielmo. 2000b. ”Restructuring” and functional structure. Ms. University of Venice. 21
- Cinque, Guglielmo. 2002. Complement and adverbial PPs: Implications for clause structure. Paper presented at GLOW 2002, Amsterdam/Utrecht. 21, 23, 30, 34
- Cinque, Guglielmo. to app. Issues in adverbial syntax. *Lingua* Special edition on adverbs, ed. by Artemis Alexiadou. 7, 38, 48, 69
- Corver, Norbert. 1997. The internal syntax of the dutch extended adjectival projection. *Natural Language and Linguistic Theory* 15:289–368. 142, 143
- Diesing, Molly. 1992. *Indefinites*. Cambridge, Massachusetts: MIT Press. 103
- Doetjes, Jenny, Ad Neeleman, and Hans van de Koot. 1998. Degree expressions. *UCL Working Papers in Linguistics* 10:323–367. 134, 142, 143
- Dowty, David. 1979. *Word meaning and montague grammar : the semantics of verbs and times in generative semantics*. Dordrecht: Reidel. 70
- Dowty, David. 2000. The dual analysis of adjuncts/complements in categorial grammar. *ZAS Papers in Linguistics* 17. 37
- Egerland, Verner. 1998. On verb-second violations in Swedish and the hierarchical ordering of adverbs. *Working Papers in Scandinavian Syntax* 61. 79
- Emonds, Joseph. 1976. *A transformational approach to English syntax: Root, structure-preserving and local transformations*. New York: Academic Press. 155
- Ernst, Thomas. 2000. Manners and events. In *Events as grammatical objects*, ed. Carol Tenny and James Pustejovsky, 335–358. Stanford: CSLI. 19, 48, 66
- Ernst, Thomas. 2001. *The syntax of adjuncts*. Cambridge: Cambridge University Press. 7, 17, 35, 38, 48, 72, 73, 75, 124

- Fanselow, Gisbert. 2002. Münchhausen-style head movement and the analysis of verb second. Ms. University of Potsdam. 100, 107, 108
- Giannakidou, Anastasia. 1997. The landscape of polarity items. Doctoral Dissertation, University of Groningen. 42, 63, 67
- Greenberg, J. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of language*, ed. J. Greenberg, 73–113. Cambridge, Massachusetts: MIT Press. 21
- Groenendijk, Jeroen, and Martin Stokhof. 1984. Studies on the semantics of questions and the pragmatics of answers. Doctoral Dissertation, University of Amsterdam. 43
- Groenendijk, Jeroen, Martin Stokhof, and Frank Veltman. 1996. Coreference and modality. In *Handbook of contemporary semantic theory*, ed. S. Lappin, 179–216. Oxford: Blackwell. 54
- Hallman, Peter. 2001. On the derivation of verb-final and its relation to verb-second. Ms. U. of Michigan. 101, 118
- Harper, W. L. 1976. Ramsey test conditionals and iterated belief change (a response to Stalnaker). In *Foundations of probability theory, statistical inference and statistical theories of science*, ed. W. L. Harper and C. A. Hooker. Dordrecht: Reidel. 56
- Holmberg, Anders. 1986. Word order and syntactic features in the Scandinavian languages and English. Doctoral Dissertation, University of Stockholm. 3, 84
- Holmberg, Anders. 1999. Remarks on Holmberg's generalization. *Studia Linguistica* 53:1–39. 3, 84
- Holmberg, Anders. 2000. Scandinavian stylistic fronting: How any category can become an expletive. *Linguistic Inquiry* 31:445–483. 86, 108
- Holmberg, Anders, and Christer Platzack. 1995. *The role of inflection in Scandinavian syntax*. Oxford: Oxford University Press. 77
- Iatridou, Sabine, Elena Anagnostopoulou, and Roumyana Izvorski. 2002. Observations about the form and meaning of the perfect. In *Ken Hale. a life in language*, ed. Michael Kenstowicz, 189–238. Cambridge, Massachusetts: MIT Press. 70
- Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Massachusetts: MIT Press. 129, 134
- Jayaseelan, K. A. 2001. IP-internal topic and focus phrases. *Studia Linguistica* 55:39–75. 120, 121

- Julien, Marit. 2000. Syntactic heads and word formation : a study of verbal inflection. Doctoral Dissertation, University of Tromsø. 37
- Kadmon, Nirit, and Fred Landman. 1993. Any. *Linguistics and Philosophy* 16:353–422. 2, 38, 43, 51
- Kayne, Richard S. 1975. *French syntax: The transformational cycle*. Cambridge, Massachusetts: MIT Press. 81, 121
- Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, Massachusetts: MIT Press. 4, 19, 20, 34
- Kayne, Richard S. 1998. Overt vs. covert movement. *Syntax* 1:128–191. 78, 90
- Kayne, Richard S. 1999. Prepositional complementizers as attractors. *Probus* 11:39–73. 78
- Kayne, Richard S. 2000. Recent thoughts on antisymmetry. Talk presented at the conference on antisymmetry, Cortona, Italy. 20
- Kayne, Richard S. 2002. On the syntax of quantity in english. Ms. NYU. 142
- Kayne, Richard S., and Jean-Yves Pollock. 1999. New thoughts on stylistic inversion. Ms. NYU and CNRS-Lyon. 136
- Koenenman, Olaf. 2000. The flexible nature of verb movement. Doctoral Dissertation, Utrecht University. 108
- Koopman, Hilda, and Anna Szabolcsi. 2000. *Verbal complexes*. Cambridge, Massachusetts: MIT Press. 21, 29, 30, 72, 78, 116, 117
- Koster, Jan. 1974. Het werkwoord als spiegelcentrum. *Spektator* 3:603–618. 31, 34, 35
- Kratzer, Angelika. 1977. What ‘must’ and ‘can’ must and can mean. *Linguistics and Philosophy* 1:337–355. 54
- Kratzer, Angelika. 1991. Modality. In *Semantics: An international handbook of contemporary research*, ed. A. von Stechow and D. Wunderlich, 639–650. Berlin: de Gruyter. 54
- Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25:209–257. 2, 38, 43, 51, 53, 62
- Lahiri, Utpal. 1997. Focus and negative polarity in Hindi. *Natural Language Semantics* 6:57–123. 38, 51
- Linebarger, Marcia-C. 1987. Negative polarity and grammatical representation. *Linguistics and Philosophy* 10:325–387. 43, 61

- Löbner, Sebastian. 1999. Why German *schon* and *noch* are still duals: A reply to van der Auwera. *Linguistics and Philosophy* 22:45–107. 46
- Meinunger, André. 2001. Adjacency requirement blocks verb raising. Paper presented at GLOW 2001, Portugal. 100
- Meinunger, André. to app. Restrictions on verb raising. *Linguistic Inquiry* . 100
- Moortgat, Michael. 1996. Categorical type logics. In *Handbook of logic and language*, ed. J. van Benthem and A. ter Meulen. Cambridge, Massachusetts: MIT Press. 6
- Mulders, Iris. 2002. Transparent parsing – head-driven processing of verb-final structures. Doctoral Dissertation, Utrecht University. 24, 25
- Müller, Gereon. 2002. Verb-second as vP-first. Ms. IDS Mannheim. 29, 102, 106, 129
- Nilsen, Øystein. 1997. Adverbs and A-shift. *Working Papers in Scandinavian Syntax* 59:1–32. 6, 111, 113
- Nilsen, Øystein. 2000. *The syntax of circumstantial adverbials*. Oslo: Novus Press. 2, 29, 32, 33, 35, 99
- Nilsen, Øystein. 2001. Adverb order in type logical grammar. In *Proceedings of the Amsterdam Colloquium 2001*, ed. R. van Rooy and M. Stokhof, 156–161. Amsterdam. 6, 68, 74
- Nilsen, Øystein. to app.a. Domains for adverbs. *Lingua* Special edition on adverbs, ed. by Artemis Alexiadou. 37
- Nilsen, Øystein. to app.b. Verb second and Holmberg’s generalization. In J.W. Zwart and W. Abraham (eds.) *Studies in Comparative Germanic Syntax*. 29, 77, 105
- Nilsen, Øystein, and Nadezhda Vinokurova. 2000. Generalized verb raisers. In *Proceedings of the 2000 International Workshop on Generative Grammar*, ed. Young Jun Jang and Jeong-Seok Kim, 167–176. Hansung university, Seoul. 21, 37
- Pesetsky, David. 1995. *Zero syntax: Experiencers and cascades*. Cambridge, Massachusetts: MIT Press. 34
- Platzack, Christer. 1986. The position of the finite verb in Swedish. In *Verb second phenomena in germanic languages*, ed. Hubert Haider and Martin Prinzhorn, 27–47. Dordrecht: Foris. 110
- Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20:365–424. 3, 68, 130, 131, 133, 134, 155

- Pritchett, B. 1992. *Grammatical competence and parsing performance*. Chicago: University of Chicago Press. 25
- Reinhart, Tanya. 1995. Interface Strategies. OTS working Papers. 119, 121
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. L. Haegeman, 281–337. Dordrecht: Kluwer. 3, 78, 80
- van Rooy, Robert. 2001. Attitudes and context change. Ms. of book ILLC, Amsterdam. 56, 57
- van Rooy, Robert. 2002. Negative polarity items in questions. Ms. ILLC, Amsterdam. 39, 43, 53, 66
- Starke, Michal. 2001. Move dissolves into merge: a theory of locality. NYU. 5, 29, 102
- Stowell, Tim. 1981. Origins of phrase structure. Doctoral Dissertation, MIT. 131
- Svenonius, Peter. 2001. Subject positions and the placement of adverbials. In *Subjects, expletives, and the EPP*, ed. Peter Svenonius, 199–240. Oxford/New York: Oxford University Press. 7, 11, 13, 14, 19, 35, 75, 124, 126
- Szabolcsi, Anna. 2001. Hungarian disjunctions and positive polarity. Ms. NYU. 49
- Szabolcsi, Anna. 2002. Positive polarity–negative polarity. Ms. NYU. 45, 46, 60, 61, 64
- Szendrői, Kriszta. 2001. Focus and the syntax-phonology interface. Doctoral Dissertation, University College London. 119
- Verkuyl, Henk. 1993. *A theory of aspectuality: The interaction between temporal and atemporal structure*. Cambridge: Cambridge University Press. 134, 135
- Vlach, Frank. 1993. Temporal adverbials, tenses and the perfect. *Linguistics and Philosophy* 16:231–283. 70
- Westerståhl, Dag. 1988. Quantifiers in formal and natural languages. In *Handbook of philosophical logic*, ed. D. Gabbay and F. Günthner, 1–131. Dordrecht: Kluwer. 38, 51
- van der Wouden, Anton. 1997. *Negative contexts : collocation, polarity and multiple negation*. London: Routledge. 37, 42
- Zwarts, Frans. 1995. Nonveridical contexts. *Linguistic Analysis* 25:286–312. 42
- Zwarts, Frans. 1998. Three types of polarity. In *Plurality and quantification*, ed. F. Hamm and E. Hinrichs, 177–238. Dordrecht: Kluwer. 42

Samenvatting in het Nederlands

De notie van zinsarchitectuur Wat is de betekenis van de bewering dat een zinsstructuur bestaat uit een sequentie FSEQ van $n \geq 1$ functionele hoofden F^0 ? Op welke gronden zou deze bewering afgewezen worden? Eén van de de hoofdstellingen van deze dissertatie is dat deze bewering in feite fout is. Er wordt beargumenteerd dat de essentiële empirische motivatie voor FSEQ, namelijk bijwoordverplaatsing, werkwoordverplaatsing en de tweede positie van het werkwoord (verb-second, V2) beter verklaard kunnen worden in alternatieve benaderingen die geen gebruik maken van FSEQ.

De basisargumentatie voor het bestaan van functionele hoofden, F^0 , bestaat uit variatie in woordvolgorde tussen talen waar voor sommige uitdrukkingen een syntactisch hoofd en ander materiaal, veelal bijwoordelijke vervoegingen, voorkomen. Kortom, Emonds (1976); Pollock (1989) maken gebruik van het feit dat in het Frans finiete werkwoorden (Vf) vooraf moeten gaan aan bijwoorden zoals *souvent* ‘vaak’ (237), terwijl in het Engels de finiete werkwoorden juist na het correponderende bijwoord van ‘vaak’, *often*, komen.

- (237) a. Jean embrasse souvent Marie. [Fr]
 Jan kust vaak Marie.
 b. * Jean souvent embrasse Marie.
 Jan vaak kust Marie.
- (238) a. John often kisses Mary.
 b. * John kisses often Mary.

Met een vergelijkbare argumentatie, komt Cinque (1999) tot de conclusie dat er verscheidene hoofdposities P_i in een zinsdeel moeten bestaan waar finiete werkwoorden en andere werkwoordsvormen naar verplaatst kunnen worden.

Allereerst merkt hij op dat alleen bijwoorden in bepaalde volgordes kunnen voorkomen en dat de restricties op deze bijwoordvolgorde universeel zijn. Vervolgens merkt hij op dat er een meer gedetailleerde variatie is tussen (Romaanse) talen waarin het mogelijk is om werkwoorden te plaatsen tussen een sequentie van bijwoorden. Uiteindelijk beargumeert hij dat als in een taal L , een werkwoordsvorm V vooraf kan gaan aan een willekeurig bijwoord a , dan moet V ook voorafgaan aan alle bijwoorden a_i die in een sequentie na het bijwoord a komen. In andere woorden, de volgorde van werkwoorden en bijwoorden volgt de mathematische notie van *transitiviteit*.

Dit patroon zou gevolgd worden mits de verschillende bijwoorden specificateurs (=‘specifiers’) zijn van de specifieke functionele hoofden F_j en de werkwoorden aan de verschillende hoofden F_j gekoppeld kunnen worden. Dit wordt geïllustreerd in figuur A.1.

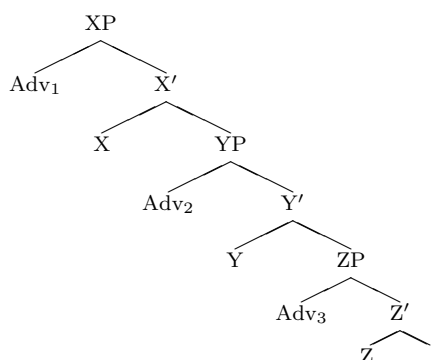


Figure A.1: Architectuur voor zinsdelen met een transitief volgorde patroon

Gegeven een dergelijke hiërarchische architectuur, zou iedere taal werkwoordsverplaatsing naar een andere hoofdpositie vereisen of toestaan. Echter, als aangenomen wordt dat de positie van het werkwoord constant blijft, en dat bijwoorden boven of onder deze positie vastgehecht kunnen worden, dan zou *transitiviteit* niet opgaan. In dit geval kunnen de sequentiele relaties tussen de bijwoorden en werkwoorden niet onafhankelijk van elkaar gegeven worden.

Dit illustreert een cruciaal aspect van de betreffende bewering: crosslinguïstische variatie van de hier beschouwde distributiefenomenen, kunnen worden ondergebracht in een *lineaire* sequentie. Als we empirisch bewijs kunnen vinden voor niet-lineaire patronen van woordvolgordes, hebben we een tegenvoorbeeld gevonden. In dit onderzoek wordt aangetoond dat het voorkomen van de Noorse bijwoorden: *muligens* ‘mogelijk(erwijs)’, *ikke* ‘niet’ and *alltid* ‘altijd’ niet transitief geordend is. Dit laat zien dat een relatieve ordening zoals in figuur 1 niet mogelijk is.

Een gerelateerd probleem is dat twee verschillende soorten uitdrukkingen zich soms houden aan orthogonale eisen voor de woordvolgorde. Dit gaat op

voor de onderlinge volgorde van argumenten en bijwoorden in het Scandinavisch en het Nederlands. Het komt ook voor met betrekking tot de onderlinge volgorde van de werkwoorden in het Italiaans, vergeleken met de volgorde van de bijwoorden. Aangezien de volgorde patronen orthogonaal zijn, kunnen ze niet worden ondergebracht in één enkele FSEQ.

Het afschaffen van FSEQ Het is aangetoond dat de onderlinge volgorde van bijwoorden in zinsdelen voor een groot gedeelte volgen uit het feit dat heel veel bijwoorden (positieve) ‘polarity items’ (PPI) zijn. Bovendien zorgen verschillende bijwoorden ervoor dat er een omgeving ontstaat waar andere bijwoorden gevoelig voor zijn. Het is bijvoorbeeld aangetoond dat spreker-geörienteerde bijwoorden, zoals *evidently* (=duidelijk genoeg), *paradoxically* (= paradoxaal genoeg), *fortunately* (= gelukkig(-erwijs)), *possibly* (=mogelijk(-erwijs)), etc. positieve polarity items zijn, in de zin dat deze woorden uitgesloten zijn in neerwaards implicerende omgevingen. Zodoende, terwijl zin (239a) (Internet-bron) als grammaticaal wordt bevonden, wordt zin (239b) ongrammaticaal beoordeeld door autochtone speakers van het Engels.

- (239) a. His retaliations killed or endangered innocents
 Zijn wraaknemingen doodde of bedreigde onschuldigen
 and **often possibly** had little effect in
 en vaak mogelijk(-erwijs) hadden weinig effect op
 locating terrorists.
 het localiseren van terroristen
 “Zijn wraakneming had tot gevolg dat onschuldigen gedood en bedreigd werden en hadden mogelijkerwijs vaak weinig effect op het localiseren van de terroristen.”
- b. ?? His retaliations killed or endangered innocents
 Zijn wraaknemingen doodde of bedreigde onschuldigen
 and **rarely possibly** had an effect in
 en vaak mogelijk(-erwijs) hadden een effect op
 locating terrorists.
 het localiseren van terroristen

Het PPI gedrag van bijwoorden zoals *possibly* wordt afgeleid door de analyse van negatieve polarity items, zoals gegeven door Chierchia (2001), aan te passen aan positieve polarity items. *Possibly* wordt afgeleid van *possible* door toepassing van een functie waar de modale basis die met het bijvoegelijke naamwoord wordt geassocieerd te verkleinen. Een dergelijke verkleining wordt geregeerd door een pragmatische anti-verzwakkingsrestrictie. De output van de domeinverkleining is niet logisch afleidbaar uit de input. Er wordt beargumenteerd dat dit voldoende is om de meeste van de geobserveerde volgorde patronen met bijwoordelijke zinsdelen af te leiden.

Een vergelijkbare analyse wordt voorgesteld voor de volgorde van zowel verbale morfologie als hulpwerkwoorden. Bovendien volgt hieruit dat zulke sequentiële fenomenen geen ondersteuning bieden voor een FSEQ.

Er wordt beargumenteerd dat de standaardanalyse van werkwoordplaatsing op de tweede positie ('verb-second', V2) in termen van hoofdverplaatsing naar C tal van problemen oplevert met betrekking tot verb-second schendingen door focus-gevoelige partikels in het Scandinavisch. De voorgestelde analyse beschouwt verb-second als de uitkomst van de verplaatsing van één XP naar de eerste positie, waarbij de XP zowel de constituent die voorafgaat aan het werkwoord als het werkwoord zelf, als ook de onbeklemtoonde voornaamwoorden bevat. Deze analyse lost de problemen omtrent de focus partikels op en verklaart Holmberg's generalisatie met betrekking tot de wisselwerking tussen argumentsverschuiving en werkwoordsverplaatsing in het Scandinavisch. Argumentsverschuiving kan geen werkwoord, of enig ander fonetisch waarneembaar materiaal dat onderdeel is van de VP, kruisen. In sommige gevallen kunnen en moeten argumenten echter bijwoorden kruisen. Deze generalisatie volgt als het argument en het werkwoord niet apart verplaatsen, maar als een XP, die al deze elementen bevat, verplaatst. In andere woorden, argumentsverschuiving kan niets anders kruisen dan bijwoorden, omdat er geen argumentsverschuiving bestaat: er bestaat slechts verplaatsing van VP's over bijwoorden.

Er wordt beargumenteerd dat fenomenen die korte werkwoordsverplaatsing beschrijven, zoals in de Romaanse talen, in weze een 'scrambling' fenomeen is. Dit wordt ondersteund door het feit dat zulke verplaatsingen semantische ('scopal') effecten hebben. Deze verplaatsingen zijn gevoelig voor informatie-structurele eigenschappen van de betrokken expressies. Bovendien zijn ze optioneel voor een beperkte reeks bijwoorden. Vandaar dat korte werkwoordsverplaatsing geen ondersteuning biedt voor het bestaan van verschillende functionele hoofden, zoals algemeen wordt aangenomen.

Door dit alles samen te nemen, kan vastgesteld worden dat de distributie van werkwoorden en bijwoorden in zinsdelen geanalyseerd kan worden zonder gebruik te maken van 'posities' in een sequentie van functionele hoofden. Door dat werkwoorden en bijwoorden alleen onderhevig zijn aan 'scope' en andere interface vereisten, maar verder onbeperkt samengevoegd kunnen worden, wordt het probleem van transitiviteit en orthogonale sequentiële patronen opgelost.

Curriculum vitae

Øystein Nilsen was born in Norway on March 29, 1971. He studied linguistics, philosophy and psychology at the University of Tromsø between 1992 and 1996, and completed his M.Phil. in linguistics at the same university in 1998. After working as a lecturer at the Tromsø Institute of Linguistics, he enrolled as a Ph.D. student at the Utrecht institute of Linguistics OTS in August 1999. The present dissertation is the result of work he carried out there.