Münchhausen-style head movement and the analysis of verb second*

Gisbert Fanselow
University of Potsdam

1 Introduction

Head movement has gained a bad reputation. It is accused of being incompatible with fundamental laws of movement theory. The minimum penalty is banishment to phonology (Chomsky 1999), but more radical prosecutors (Mahajan 2001) have pleaded for capital punishment. The head movement constructions of previous models are analyzed as involving remnant movement (see Koopman & Szabolcsi 2000, Mahajan 2001, Müller 2003).

The present paper subscribes to such a reductionist view as well, but it argues that the substitution type of head movement exemplified, e.g., by verb second movement cannot be replaced by remnant movement. For these constructions, we develop a restrictive concept of head movement that arises from a slight extension of assumptions made in Chomsky (1995). Our approach differs from others in confining head movement to true substitutions within the limits of extended projections in the sense of Grimshaw (1991).

The paper is organized in two parts. The first part is dedicated to a theoretical analysis of the *pro*s and *con*s of head movement. A slight relaxation of the conditions of feature checking opens up a tiny and highly specialized niche for head movement. This version of

*Some of the ideas in this paper have been presented at the University of California at Los Angeles, the Aristotle University at Thessaloniki, and the University of Wuppertal. I am grateful to the audiences for criticism and helpful comments. Thanks also go to Artemis Alexiadou, Sjeff Barbiers, Hans Broekhuis, Joanna Blaszczak, Eva Engels, Caroline Féry, Susann Fischer, Werner Frey, Jane Grimshaw, Liliane Haegeman, Andreas Haida, Gunnar Hrafnbjargarsson, Hilda Koopman, Anoop Mahajan, Gereon Müller, Vieri Samek-Lodovici, Matthias Schlesewsky, Halldór Sigurðsson, Arthur Stepanov, and Ralf Vogel. The research reported here was partially supported by a DFG grant to the Research Group "Conflicting Rules" at the University of Potsdam, and a Transcoop grant from the Alexander-von-Humboldt Foundation.

Linguistics in Potsdam 22 (2004): 9-49
Ralf Vogel (ed.):
Three papers on German verb movement

head movement is immune to the general criticism alluded to above, and some core properties of constructions such as verb second or V-to-Infl movement are immediate consequences of our approach. In the second part of the paper, we corroborate the resulting model by an in-depth analysis of verb second constructions in a variety of languages.

2 A restrictive theory of head movement

2.1 Preliminary Remarks

Recent typologies of head movement such as Roberts (1994) and Riemsdijk (1998) suggest that three different types of constructions can be distinguished in which an element with the phonetic properties of a word is displaced syntactically. In a pre-theoretic sense, the verb moves *into* an independently existing position in the case of, say, German verb second (V2) constructions, as illustrated in (1) [=substitution]. Verbs (and other heads) can also be *adjoined* to other verbs, as shown in (2) [=adjunction]. This distinction between substitution and adjunction is independent of the issue of the existence of so-called long head movement, as illustrated in (3) for Croatian, which differs from (1) and (2) in that the Head Movement Constraint[1] of Travis (1984) is, apparently, violated.

Substitution: V-to-C movement, V-to-I movement

(1) er hat$_i$ ihn gesehen t_i (German)
he has him seen

Head Adjunction: "Restructuring" in V-V-contexts

(2) dass er [sie t_i [$_V$ [$_V$ zu küssen]$_i$ wagt]] (German)
that he her to kiss dared
"that he dared to kiss her"

[1] According to the Head Movement Constraint, head movement can only target the next head position up in the structure.

Long Head Movement

(3) dao$_i$ mu ga je Ivan t_i (Croatian)
given him it is Ivan t
"Ivan has given it to him"

In a restrictive model of movement such as Chomsky (1993, 1995), a substitution operation cannot exist. Categories that undergo movement are either adjoined to the root node of the phrase marker currently under construction, or to the head projecting that root node. In any event, a head position H would have to be radically empty if a category C is to move *into* H. Consequently, H could not possess a feature triggering the attraction of C. In a restrictive model of grammar, in which movement is a last resort operation serving the need of feature checking, nothing can move to a radically empty head position. Substitution operations of earlier models thus have to be reanalysed as adjunctions to phonetically empty heads, as illustrated in (4) for V-to-C movement.

(4) $[_{CP}\ \alpha\ [_C[_T[_V\ \text{VERB}]\ \varnothing]\ \varnothing]\ [_{TP}\ \beta\ t_T\ [_{VP}\ t_V\ \gamma\ ]]]$

According to Riemsdijk (1998), an attracting head can be specified phonetically only if it is strictly adjacent to the attracted head before movement[2]. The fact that the attractor must be empty, otherwise, is a key generalization to be captured in a movement theory.

"Long head movement" as in (3) maps words into a position that is otherwise occupied by maximal projections The position preceding the finite verb can be occupied by focused objects (5a) and subjects (5b) in Breton, but in pragmatically unmarked clauses, it is occupied by the non-finite verb (5c). Similarly, Icelandic Stylistic Fronting as in (6) can place a non-finite verb into [Spec,IP] (according to

[2] Given that the two heads are adjacent before movement, any phonetic or morphological effect of movement can be taken care of in the morphological component alone. To the extent that movement of the head H is, thus, primarily motivated by the absence of an island status of the XP projected from H (in the spirit of Baker 1988), the development of an alternative theory of barrierhood might in fact eliminate the motivation for movement. I will not pursue this issue here.

Holmberg, 2000) when [Spec,IP] is empty, as in an impersonal passive construction, or when the subject has moved to the left or the right.

(5) a. E bark en deus aret Yann (Breton)
his field PRT have-3m ploughed Yann
b. Yann en deus aret e bark
c. *Aret* en deus Yann e bark
"Yann has ploughed his field"

(6) ég helt að *kysst* hafðu hana margir stúdentar (Icelandic)
I believed that kissed have her many students
"I believe that many students have kissed her"

The idea thus suggests itself that "long head movement" belongs to the paradigm exemplified in (7). Full verb phrases may be moved to [Spec,CP] as in (7a), but scrambling can remove one or more phrases from that verb phrase before it goes to [Spec,CP]. This leads to structures such as (7b-e), as Thiersch (1985) and den Besten & Webelhuth (1987, 1990) argue[3]. (7e) is particularly interesting: in phonetic terms, what occupies [Spec,CP] is a single word, but syntactically, the position is filled by a maximal verbal projection that is full of traces. See Müller (1998) for an elaborate theory of remnant movement.

(7) a. [gestern hier dem Kind den Stern gezeigt] hatte sie (German)
yesterday here the child the star shown had she
b. hier dem Kind den Stern gezeigt hatte sie gestern
c. dem Kind den Stern gezeigt hatte sie gestern hier
d. den Stern gezeigt hatte sie gestern hier dem Kind
e. gezeigt hatte sie gestern hier dem Kind den Stern
„she had shown the star to the child here yesterday"

(5c) and (6) differ from (7e) in the pragmatic conditions, and in terms of the obligatoriness of extracting all elements but the verb from the verb phrase, but structurally, they are similar. Thus, "long head movement" at least reduces to remnant phrasal movement. Furthermore, Mahajan (2001) shows that a simplification of the syntax of OV languages is possible when one assumes remnant movement, because, e.g., rightward scrambling can be dispensed with. For Hindi (8), it seems

[3] But see Fanselow (in press, a) for critical remarks.

more reasonable that the *leftward* extraction of the object *saare phal* out of VP/IP is followed by a further *leftward* movement of VP/IP.

(8) Raam-ne [$_{IP}$[$_{VP}$ t khaaye] the t_{VP}] saare phal t_{IP} (Hindi)
Raam-ERG eat.PERF.MASC.PL be.MASC.PL.PST all fruits.MSC
"Raam had eaten all the fruits"

As Mahajan points out, *all* apparent instances of head movement might in principle be reanalysed as remnant phrasal movement. This is mandatory if head movement is untenable from a theoretical point of view.

2.2 Theoretical Problems of Head Movement

The first charge against head movement is based on the structure given in (9), with X having moved from the head position in XP to Y, involving head adjunction, the minimalist way of spelling out head movement.

(9) [$_{YP}$[$_{Y}$ Y] [$_{XP}$ …. X ….]] ==>
[$_{YP}$[$_{Y}$ X Y] [$_{XP}$ …. ~~X~~ ….]]

The movement in (9) fails to meet the extension requirement of Chomsky (1995). In principle, movement should be an operation that picks an element α in Σ, and adjoins it to Σ, such that [αΣ] arises. Moved material must be merged at the root. In (9), this condition is not fulfilled: X is adjoined to a daughter of the root, not the root itself. Head movement is, therefore, counter-cyclic as well, because it affects two positions internal to a structure that has already been built. The head moved in (9) fails to c-command its trace under a strict definition of the term: α c-commands β if the first node above α also dominates β. After movement, the first node above X in (9) is Y, and Y does not dominate the trace of X.

This summary of three of the four arguments[4] Mahajan (2001) brings forward against head movement shows that the charge is based on serious offences, and it

[4] The fourth argument is that head movement appears to be semantically vacuous. To the extent that the claim is true at all (see Engels, in prep., for counterexamples) it is not really related to the issue under consideration: the problem does not disappear when head movement is replaced by phrasal movement.

seals the fate of the *standard practise* of carrying out head movement (adjunction to another head) if an alternative is at hand. Note the arguments rule out adjunction to a non-root positions in general. They are valid quite independently of whether this unacceptable operation adjoins a head to a further head, or a phrase, to a specifier (as has been suggested for multiple wh-movement such as (10) in Bulgarian or Romanian by Rudin 1988 and Grewendorf 2001).

(10) koj kogo mislis (Bulgarian)
who what bought
„who bought what“

Pointing out that there are other culprits does not eliminate the guilt. A solution of the problems identified by Mahajan needs to avoid adjunction to a non-root position. It need need not avoid head movement, though.

A second set of problems arises in the context of identifying the “traffic rules” for head movement. Suppose that Tense has a strong V-feature (triggering V-to-I movement) and a strong D-feature (triggering movement to the subject position), as may be true in French, but see below. The question is why such requirements are always met by moving DP to [Spec,T] and V to Tense (11b), and by not by moving D to T and VP to [Spec,TP] (11c).

(11) a. Tense {D, V} [$_{VP}$ DP_1 [$_{V'}$ V DP_2]]

b. [$_{TP}$ DP_1 [$_T$ V] [$_{VP}$ ~~DP_1~~-[$_{V'}$ ~~V~~ DP_2]]]

c. [$_{TP}$ [$_{VP}$ [$_{DP1}$ ~~D~~ NP] [$_{V'}$ V DP_2]] [$_T$ D] [~~VP~~]]

This difficulty is unavoidable in *any* system in which a head can possess two attracting features, independent of whether these lead to the creation of multiple specifiers, or one specifier and one head. One might add some traffic rules, as encoded by, say, accessibility in the sense of Zwart (1993): feature f can be checked only if feature f’ has previously been erased. Pesetsky & Torrego (2000) offer a more principled solution:

(12) HEAD MOVEMENT GENERALIZATION

Suppose a head H attracts a feature of XP as part of a movement operation.

(i) If XP is the complement of H, copy the head of XP into the local domain of H.

(ii) Otherwise, copy XP into the local domain of H.

(12) implies a very strict version of the Head Movement Constraint of Travis (1984): head movement can only target the closest head. This follows from (12) in an obvious way: if the attracting head is higher, it could not trigger the movement of a *head*. (12) is attractive, but one would like to be able to derive it from some general property in the theory of movement.

Koopman (1994) proposes a version of Kayne's (1994) Linear Correspondence Axiom LCA that has far-reaching consequences. Like Chomsky (1995) she restricts the effects of the LCA to overt material. Making the assumption that intermediate projections count when c-command relations are computed, it follows that α and X cannot be linearized in (13). X' asymmetrically c-commands α, so that all material dominated by X' –in particular, X itself- should precede α, given the LCA. Furthermore, YP asymmetrically c-commands X. Therefore, all material dominated by YP –in particular, α– should precede X. Thus, we have derived a contradiction which is resolvable only if either the head or the specifier of a projection is phonetically empty.

(13) $[_{XP}[_{YP}\ \alpha\ ]\ [_{X'}\ X\ BP]]$

As a consequence, one can assume that each head can have at most one attracting feature[5]. This eliminates the traffic rule problem for head movement, or, rather, translates it into a problem of the sequencing of functional heads. Unlike (12), it does not eliminate the need of deciding which features trigger head movement, and which

[5] French seems to be a counterexample if the subject moves to [Spec,TP] and V moves to T. However, as Koopman (1996) points out, negation and clitics may intervene between the subject and the verb in French (*Jean le voit* John him sees), and to the extent that clitics land in a projection of their own, such data show that the subject moves to a position in a higher projection than the one hosting the finite verb.

lead to phrasal movement. Taking these two points together, it is not entirely clear that real progress has been made,

The constellation created in (9) also violates the Chain Uniformity Condition of Chomsky (1995). In a minimalist grammar, projection levels cannot be primitive entities (they violate the inclusion requirement because they are not specified in the lexicon), rather, they are relational concepts that can be read off structural representations. Following Speas (1990) and Chomsky (1995), a configurational definition of projection levels amounts up to the following: Σ is a maximal projection unless its mother is a projection of Σ. Σ is a head if Σ does not dominate further (non-terminal) material. If a head H adjoins to another category α, its mother fails to be a projection of H in the resulting structure $[_{\alpha}$ H α]. Therefore, in $[_{YP} [_{Y}$ X Y] $[_{XP} \ldots$ ~~X~~ ….]], the trace of X is not maximal, while the moved head acquires that status in its landing site. This violates the Chain Uniformity Condition that requires that the phrasality status of a category must not change after movement, that is, the members of a chain agree in terms of maximality. Chomsky (1995) circumvents the problem resulting for head movement by assuming that elements adjoined to a head are not subject to the syntactic mechanisms that determine phrasal level status.

Finally, we need to explain why the attracting head is always empty in head movement constellations (at least in the contexts identified by Riemsdijk 1998), if that property does not characterize attracting heads in *all* movement constellations (as Koopman 1996 suggests, see above). One might be able to derive this property from a Chomskyan interpretation of Kayne's (1994) Linear Correspondence Axiom. Chomsky (1995) proposes that the LCA affects overt categories only (because it holds at PF). In the constellation $[_{\alpha}$ H α] arising from head movement, H and α c-command each other symmetrically, so that the LCA does not imply anything concerning their serialization. The LCA only requires that an element α *asymmetrically* c-commanding β precedes β. If elements can *only* be serialized by

the LCA, the structure [$_{\alpha}$ H α] cannot surface (because the linear position of H is not determinable) unless α is phonetically empty. Chomsky (1995) stipulates that elements dominated by a word-level category are serialized by principles different from the LCA. If this stipulation is abandoned, the phonetic properties of the attractor in head movement contexts are derived.

2.3 Remnant Phrasal Movement

Remnant phrasal movement of XP can create constellations in which the head X is the only overt category that undergoes movement. This has been noted when the concept "remnant movement" was introduced. That remnant movement might replace head movement in general is a recent suggestion, see Koopman & Szabolcsi (2000), Koopman (2001), Mahajan (2001), among others, and Fanselow & Ćavar (2001) for a different execution of the same idea. In a straightforward way, the replacement of head movement by phrasal movement solves some of the difficulties discussed in the preceding section. It does not solve other problems, and creates fresh ones. Therefore, we will develop a new model for head movement below.

Remnant phrasal movement of YP maps a phrase to an (inner) specifier of an XP. This movement can be compatible with the extension requirement, so that the problems that arise when an element is not adjoined to the root are avoided. YP is maximal both in its pre-movement position and in its landing site, so that the Chain Uniformity Condition is respected as well.

(14) [$_{XP}$ [$_{YP}$ ... Y …] X [$_{ZP}$ …. t_{YP} ….]

The traffic rule problem seems non-existent, too (since one does not have to determine which instances of attraction imply *head* movement), but it reappears in a –perhaps- sharper form: now, there must be a component of grammar that decides under which conditions the moved phrase must not contain more phonetic material than a head.

Such a component might consist of complexity filters in the sense of Koopman & Szabolcsi (2000) that restrict the phrasal makeup of elements appearing in certain specifier positions, or we might state the constraints in phonological terms, as suggested by Fanselow & Ćavar (2001). Such approaches embody the claim that (a) complexity restrictions may lead to phonetic constellations different from those arising by head movement, and that (b) the complexity restrictions are uncorrelated with the "traditional phrase structural" position of the material in question. Let us begin with (a).

Fanselow & Ćavar (2001) argue that the German paradigm (15) shows the need for complexity restrictions independent of head status. Verbs pied-pipe their unstressed particles when they undergo V2 movement (15a,c), while stressed particles are stranded (15b,d). (15) indeed establishes the need for a morpho-phonological complexity filter for the second position. (15) also shows that *lexical entries* can be split up in a V2 construction. (15) does not show that elements other an X° category can occupy the second position, however. The paradigm in (15) constitutes no reason for abandoning the idea that X° elements only undergo V2 movement in German.

(15)	a.	dass	er	den	Brief	beginnt	(German)
	b.	dass	er	den	Brief	an.fängt	
		that	he	the	letter	begins	
	c.	er	beginnt	den	Brief	Ø	
	d.	er	fängt	den	Brief	an	
	e.	"(that) he begins with the letter"					

The (non-)existence of constructions in which *more* material than a single lexical item appears in a slot reserved for X° in head movement accounts allows to draw stronger conclusions. A brief consideration of the empirical evidence suggests that there is no compelling evidence for giving up the generalization that it is exactly X° elements which are displaced in head movement constellations. Thus, the remnant movement theory faces a serious overgeneration problem.

Confining our attention to the substitutional type of operation, clitics could be pied-piped in head movement constellations. To the extent that clitics form an

incorporated part of an X°-category, however, their pied-piping does not tell us much about the upper limits of what can appear in a position targeted by head movement.

There are only very few examples in which material that appears to have been pied-piped in head movement does not allow an incorporation analysis. According to Tuller (1992), the focus position of Kanakuru is postverbal, and since V moves to Infl in focus constructions, the focus position immediately follows Infl. When the subject is in focus, and the verb is transitive, the object is placed between the lexical verb and the focus (16 = (5a) of Tuller 1992). Tuller (1992) argues that the structure involves V-to-Infl movement as well, but the object has been incorporated into V before movement to Infl.

(16) are lowoi **jewoi** la lusha (Kanakuru)
bury boy.def slave.DEF in bush
"it was the slave who buried the boy in the bush"

As Tuller (1992:320) notes, one also finds examples such as (17) in which the object is more complex, but still precedes the subject in subject focus constructions. Standard insights on incorporation make it unlikely that a sequence of a noun, a relative marker, and an adjective could incorporate into V. If the postverbal position of a focal subject is, in fact, a consequence of a movement to Infl, (17) would instantiate a construction in which more material than X° shows up in a head position – an analysis considered in work in progress of Vieri Samek-Lodovici and myself. It is not entirely clear, however, whether the verbal projection is really displaced to Infl in examples such as (16) and (17). Tuller offers no independent evidence for the claim that the fronting of verbal material in focus constructions must go to a head position. In fact, (16) and (17) may be used as an argument for a movement of VP to [Spec,IP] or an adjunct position of IP.

(17) nai gwa m ?wali nani (Kanakuru)
drank water RM cold.DEF I
"it is me who drank cold water"

In German, verbs cannot move out of the syntactic scope of certain operators such as *mehr als* "more than", see, e.g., Meinunger (2001), as the contrast between (18a) and

(18b) shows. For most speakers, this constraint implies that (18a) has no matrix counterpart, but others find (18c-d) only mildly ungrammatical.[6] (18c-d) might be analyzed as involving the pied-piping of V' or a larger verbal projection to Comp - but this analysis is far from being the only one available. Given that *mehr als* freely combines with all kinds of categories[7], it might be attached to C' in (18c-d), squeezing itself between Comp and [Spec,CP].

(18) a. dass Hans seinen Profit letztes Jahr mehr als *verdreifachte* (German)
that Hans his profit last year more that tripled
b. *Hans *verdreifachte* seinen Profit letztes Jahr mehr als t
c. ?Hans mehr als *verdreifachte* seinen Profit letztes Jahr
d. ?Seinen Profit mehr als *verdreifachte* Hans letztes Jahr
e. "Hans more than tripled his profit last year"

It seems fair to conclude, then, that there are no strong reasons for giving up the generalization that only X°-elements may be displaced overtly in the core cases of "substitutional" head movement. The remnant movement theory has no answer to the question of why this generalization holds, if movement always involves the displacement of a phrasal category.

There are two aspects of this point which render it a strong objection against remnant movement theories. First, whenever a category C is moved to a domain in sentence structure which would be a head position under standard assumptions concerning phrase structure (viz., between [Spec,CP] and IP, or between [Spec,IP] and the verbal projections), it cannot consist of more than an X° overtly. A head movement theory has a straightforward answer to the question as to why this should be so (because C moves to a head position, after all, which cannot host more than an X°). In a remnant theory, it is a mystery why the linear slots that can be filled by X° elements only on phrase structural grounds coincide with those that satisfy this requirement as a consequence of additional restrictions imposed on specifiers filled

[6] Six out of 20 native speakers of German accepted (18c-d) in an informal survey.

[7] Compare (i) and (ii), where *mehr als* "more than", takes scope over the verb in (ii), and over the whole (VP) in (ii).
(i) er hat seineKinder mehr als geschlagen (German) - he has his children more than beaten
(ii) er hat mehr als seine Kinder geschlagen

by movement. Similarly, the Pesetsky-Torrego-generalization (12) is a mystery under a remnant movement theory: why should the linear distance between the attractor and the attractee be correlated with complexity restrictions on the specifier position of the attractor?

Approaches replacing head movement by remnant phrasal movement have to deal with yet a further difficulty. In minimalist syntax, it is not sufficient to formulate complexity restrictions for positions P targeted by phrasal movement. The derivational steps that take enough material out of XP before it moves to P (so that the complexity restrictions can be satisfied) must be licensed themselves. This condition is not always fulfilled.

For example, recall that stressed verbal particles must be stranded in German (and Dutch) V2 constructions. If V2 movement is remnant movement of VP or IP, the particle *an* has to be moved out of VP in (19b), and there seems little motivation for this operation besides the need to create a remnant VP category that contains the head *kommt* only.

(19) a. dass der Zug pünktlich ankommt (German)
that the train punctally at.comes
b. der Zug kommt pünktlich an
the train comes punctually at
c. *dass der Zug an pünktlich kommt
"(that) the train arrives in time"
d. angekommen ist der Zug pünktlich
at.come is the train punctually
e. *gekommen ist der Zug pünktlich an
"the train has arrived in time"

The particle extraction preceding remnant VP fronting would have to be one of those operations that never change linear order: the verbal particle must not precede any other constituent in the clause but the verb (19c). Likewise, the particle cannot be stranded in clear cases of remnant VP-movement such as the fronting of VP to [Spec,CP] in (19d,e). It is unclear why the particle should be strandable in doubtful cases of remnant movement, but never in undisputed ones. By allowing movement to be already licensed by the need to satisfy constraints restricting the complexity of

certain phrases (see Müller 2003), the technical problems are solved, but it is hard to see what kind of evidence could then ever refute such a version of a remnant movement reinterpretation of head movement.

This type of problem also arises in an approach such as the one proposed in Fanselow & ?avar (2001), which exploits the descriptive potential of the copy-and-deletion theory of movement. In their model, the formation of a full copy as a first step in movement may either be followed by a complete deletion of the lower copy (overt movement) as in (20a), the deletion of the upper copy (covert movement) as in (20b), or partial deletion affecting both copies, as in (20c). The impression of "head movement" arises when everything but the head is deleted in the upper copy of a phrase. This model is in need of being complemented by strong principles restricting partial deletion.

(20) a. αβγ [... ~~αβγ~~ …]
b. ~~αβγ~~ [... αβγ …]
c. ~~α~~βγ [... α~~βγ~~ …]

Summing up, we have observed that the Chain Uniformity difficulty and the problems resulting from the fact that the head does not adjoin to the root in standard accounts of head movement are circumvented in a remnant movement reinterpretation. However, this model fails to offer an explanation for a number of generalizations of head movement, viz., those that characterize the conditions under which a moved phrase must not contain more visible material than a head. In a convincing account of head movement, these generalization should not just be stipulated. Furthermore, it is unclear whether the set of movement operations licensed independently is sufficiently powerful to be able to extract the necessary amount of material for creating a remnant XP in which the head is the only overt category.

2.4 Münchhausen Style Head Movement

Given the results of the preceding section, we propose to *modify and restrict* the standard account of head movement rather than replacing it by remnant phrasal movement. Recall that many problems arising with head movement result from the fact that heads are not adjoined to the root in the standard way of carrying out head movement. Let us therefore assume that heads adjoin to the root as well – just as phrases do.[8] This yields a structure such as (21). Head X is adjoined to a non-minimal projection of some head Y attracting it, rather than to this head itself.

(21) $[_{YP}[X\,[Y\,]\,[_{XP}\ldots.\,t_X\,\ldots.]]]$

There is, thus, a way of carrying out head movement that is innocuous in terms of extension and cyclicity. It is reminiscent of the original substitution idea, but differs from it in that X does not move to a position that was occupied by something else before movement. Of course, (21) is not yet the structure we are looking for, in spite of the fact that X c-commands is trace and that adjunction to the root involves a cyclic operation only. If nothing is changed, (21) violates the Chain Uniformity Condition: since its mother is not projected from X, X is a maximal projection in (21), while its trace is not.

This difficulty disappears if we make the theory of movement more minimalist, in the sense of reducing the number of assumptions made concerning movement and checking. In particular, let us change the theory of movement as indicated in (22):

(22) After the attraction of α to the root of Σ, either α or Σ may project.

If the target of movement projects as in (23a), α is a maximal projection, because its mother node is projected from Σ. Given the Chain Uniformity Condition, α must be a maximal projection in the root position, too. Thus, we are confronted with phrasal movement, that is, (23a) represents the standard case. If the moved category α projects after movement as in (23b), it cannot be maximal in either its root or its

[8] A similar assumption is made in Koeneman (2000), who traces back the idea to Ackema et al (1993). The idea seems to have first been formulated by Anders Holmberg, in his 1991 GLOW talk *Head Scrambling*, as was pointed out to me by Gereon Müller.

target position. We have thus identified a head movement constellation[9], in which none of the problems discussed by Mahajan (2001) arises. Let us therefore assume that movement is strictly cyclic, and governed by Chain Uniformity. (23) represents the only two constellations that can arise.

(23) a. $[_{\Sigma}\ \alpha\ \Sigma]$

b. $[_{\alpha}\ \alpha\ \Sigma]$

In the constellation (23b) created by head movement, the category Σ that α has been merged with inevitably becomes the complement of α. If this is interpreted in a proper theory of complementation, the strict locality of head movement can be derived, which constitutes a major argument in favor of the approach proposed here. To see why, suppose that (24) holds

(24) a. α can merge with head H as a specifier or complement only if α checks a feature of H

b. If a strong [- interpretable] feature f is checked in Hα or αH^{1}, it is a feature of H.

Both assumptions are common in current versions of minimalism. Notice that the slight deviation (24a) constitutes from the system of Chomsky (1995) licenses the head movement constellation (23b). As Chomsky (1995:256-260) points out, the moved category cannot project if feature checking is confined to specifier-head-relations. This is so because Σ is a complement in (23b) if α projects. Therefore, it could not function as a feature checker, that is, movement of α could not take place at all. In the more general approach (24a), (23a) and (23b) are licensed - but nothing else.

We now have to figure out which constellations lead to the creation of (23b) rather than (23a). It will turn out that head movement can arise under extremely restricted circumstances only. Given (24b), heads only possess the strong uninterpretable features triggering syntactic processes such as movement. Suppose that H has a strong uninterpretable feature f, and suppose that α itself possesses the matching feature f^{+}. Then the most economical way of checking f arises by just *merging* α with

[9] We continue to assume that intermediate projections cannot be addressed at all by grammatical processes, so that it is only heads and phrases that can move under the new perspective.

H as a complement, or a specifier (if the complement position is already filled). In this case, either [Hα] or [αH[1]] are generated, and no movement (in particular, no head movement) is licensed.

Suppose, then, that the feature f* that matches a requirement of H is embedded in α, appearing there on some β headed by a k. If locality requirements allow it, f may still be checked by merging α, but if α has already been merged with H on independent grounds, checking must proceed by movement. The constellation βH[1] (=phrasal movement) may arise in this context, but apparently not kHP (=head movement). It is easy to see why. First, given that the strong [-interpretable] feature is, *ex hypothesi*, a feature of H, the specifier-head constellation βH[1] is in line with (24b). If f* on β headed by K is not strong, the head movement constellation kHP violates (24b): a strong uninterpretable feature of the complement, and not of the head k, is checked in this configuration. So suppose that f* of β headed by K is strong (too). Recall that the overall structure we are considering is a constellation H[$_{\alpha}$ … [$_{\beta}$ …k …] …]. Could k move in this structure? The answer is negative. If α is a projection of β, then it is a projection of k. Consequently, the feature f* would appear on α, too, and f* would have already checked α and H merged. So α cannot be a projection of k.

This, however, is irreconcilable with the assumption that k possesses a strong feature. The cyclicity of movement is guaranteed by the requirement that at least strong uninterpretable features cannot be tolerated for long in a derivation. They must be checked as early as possible. Assume that the proper way of spelling this out is to say that all strong features of a head X must be eliminated before XP is merged with a further category that projects (=Chomsky 1995), that is, all strong features of X must be checked within the maximal projection of X. Therefore, (23b) cannot arise because the strong feature f* of k failed to have been checked when β was embedded in α not projected from k.

What we have just derived appears to be a counterproductive result! The additional freedom created by (24a) cannot be made use of because (24b) excludes it. In order to create a constellation like (23b) by movement, H would have to possess a strong feature checked by α, but either that feature is checked automatically by merger, or it has to appear too deep in the structure for being tolerable.

But notice that we have so far overlooked exactly one possible constellation. When the strong feature f is embedded too deeply in the structure, the situation cannot be remedied, but there is a constellation in which a strong feature of k or H could **not** be checked by merger already in H [$_{kP}$ … k …] – this is impossible when k and H are *identical*. Therefore, (25) is the only constellation in which head movement is licensed by (24): the head in question possess the checking feature and the feature to be checked at the same time.

(25) [$_{XP}$ … X …] →
X [$_{XP}$ … t_X …]

On obvious grounds, (25) does not violate the requirement that strong features of X must be checked before the projection of X is embedded in a projection of a different element. (25) also satisfies the strict cycle condition and the Chain Uniformity Condition. When X undergoes head movement, there is no attractor different from X present in the structure. This is equivalent in its net effect to the generalization that the attracting category must always be invisible in head movement. Finally, in a constellation leading to the head movement of X, the relevant feature must not be too deeply embedded. (25) implies that the head cannot move too far. In fact, it can only place itself immediately above (one of) its own projection(s). In its net effect, this is equivalent to the Head Movement Generalization uncovered by Pesetsky & Torrego (2000). Our restricted theory of head movement thus meets *all* requirements a model of head movement must fulfil. It is quite exceptional in this respect.

There are various types of heads for which (25) might arise, that is, for which one may assume the simultaneous presence of selecting and selected features. The most

restricted domain of verb movement applies within the so-called Larsonian shell (Larson 1988). In the context of somewhat different assumptions, Anders Holmberg (in his 1992 GLOW talk) proposed an analysis of movement within VP that is similar to the one defended here, as was pointed out by Müller (p.c.). "V-to-v movement" takes place when a lexical entry is categorized as a v and a V, with v possessing a strong V feature, see below.

To sum up, one can be quite content with (25) resulting from (24): the objections raised against head movement in section 2.2. do not hold for (25), and (25) avoids the difficulties identified for remnant phrasal movement. UG opens only a limited niche for head movement, and (25) seems a good characterization of this niche. The feature structure necessary for (25) is unobjectionable under closer inspection. (25) arises when a head X possesses a strong feature f *and* the matching feature f^+ at the same time. The potentially offending feature of X is eliminated by X itself. Feature checking is thus always Münchhausen[10]-style in head movement. There is *nothing* in the theory of features that excludes that situation on principled grounds. Strong features triggering movement are abstract entities, uncorrelated with any "objective" morphology, at least as far as we know (see Alexiadou & Fanselow, in press, for this point).

3 Verb Second Movement

3.1 Introductory Remarks

In this section, we apply the model developed above to one particular instance of head movement, viz. V2 constructions in German and other languages.

Let us begin by asking what would be an example of the feature structure leading to (25). A lexical element such as French *aime* "loves" combines feature of both a verb and an Infl. Recent approaches to morphology do not assume that this entity has been composed in the syntax. Therefore, it seems safe to assume that this element

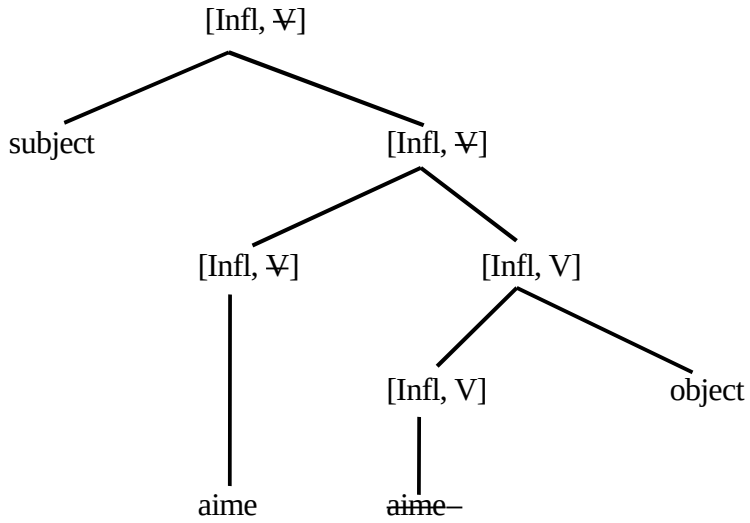
[10] Recall that according to popular wisdom, the legendary count of Münchhausen managed to pull himself out a swamp by pulling his own hair.

comes directly from the lexicon, and that it is categorized as an element that is an Infl (like English *will*) and a verb (like English *love*). If there is no other verb present in the clause-relevant numeration, as is necessary on independent grounds, the V feature of the Infl aspect of *aime* can be checked only by moving *aime* itself.

(26) <*aime*, {Infl (_V), V (_D)}, ? ♫☼>

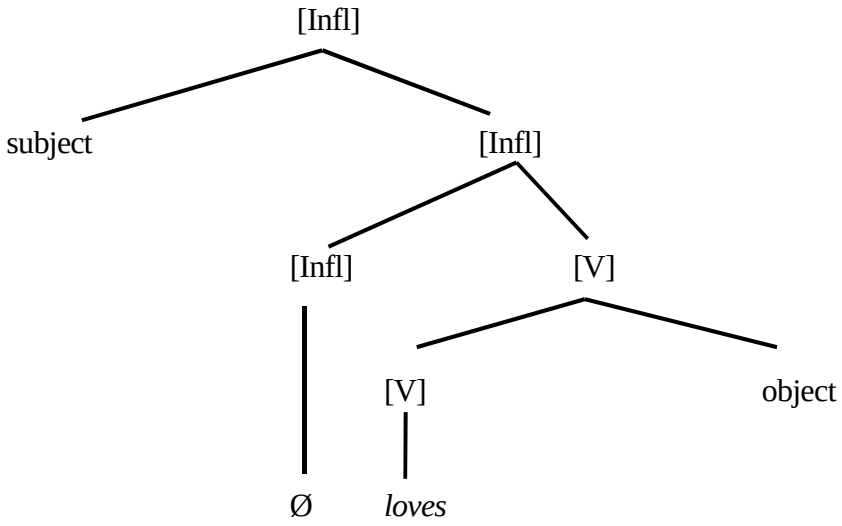
In languages that are traditionally analysed as allowing V-to-Infl movement, structures like (27) arise by Münchhausen-movement of the verb analysed as a V-Infl complex. In this structure, two feature complexes co-project (at least up to a certain level in structure). Our proposal is thus much reminiscent of the matching projection idea introduced into generative discussion by Haider (1987).

(27)



In a language like English, *love* is entered as a verb (and not as a V-Infl) into the syntactic representation. Therefore, an empty Infl element must be selected in the numeration if an IP is to be generated. The resulting structure (28) is quite classical.

(28)



If nothing else is said (but see the final subsection), the V-Infl entities we assume for languages with verb movement just need to move *at some point* in order to get rid of the strong V feature they possess, but they may do so *at any point* in the derivation. What we expect to find in the empirical data is thus possibly exemplified by (29) – (32) taken from Cinque (1999): the verb undergoes movement, but it may place itself between any two specifiers/ adjuncts related to its projection.

(29)	da allora,	non	hanno	*rimesso*	di solito	mica	più	sempre
	since then	not	have-3pl	put	usually	not	any longer	always
	completamente		tutto	bene	in ordine			
	completely		all	well	in order			

(30) da allora, non hanno di solito *rimesso* mica più sempre completamente tutto bene in ordine
da allora, non hanno di solito mica *rimesso* più sempre completamente tutto bene in ordine
da allora, non hanno di solito mica più *rimesso* sempre completamente tutto bene in ordine
da allora, non hanno di solito mica più sempre *rimesso* completamente tutto bene in ordine
da allora, non hanno di solito mica più sempre completamente *rimesso* tutto bene in ordine

(31)	mi	*ero*	francamente	purtroppo	evidentemente	formato	una	pessima
	me	is	frankly	unfortunately	obviously	formed	a	very bad
	opinione	di	voi					
	opinion	of	you					

(32) francamente *mi ero* purtroppo evidentemente formato una pessima opinione di voi
francamente purtroppo *mi ero* evidentemente formato una pessima opinione di voi
francamente purtroppo evidentemente *mi ero* formato una pessima opinione di voi

This analysis is tenable if the adverbs in (29) – (32) are adjuncts or secondary specifiers of a V-Infl-projection, as is implied by the work of Ernst (2001). Alexiadou (1997) and Cinque (1999) argue that adverbs are specifiers of their own projections. Their analysis is (primarily) based on the insight that adverbs need to appear in a specific order, and that there is a landing site for heads like verbs between the adverbs. Ernst and Engels (in prep.) show that adverb order can be states in terms of semantic selection. The present proposal implies that verbs create their own landing site when they move anywhere in their own projection. (29) – (32) are compatible with this view.

3.2 Problems of V/2 constructions: Does V really move to Comp?

V2 clauses as exemplified by German (33) support our head movement model in quite a number of respects, but they also point to some shortcomings. We will focus our discussion on German data that directly bear on where V moves to in this subsection. Section 3.3. is concerned with what one can learn from the nature of the element in preverbal position about the nature of V2. Broadening the perspective to other languages in section 3.4. will suggest some amendments.

In our approach, V2 movement is triggered by the simultaneous presence of a strong feature to be checked (say, a feature checking finiteness) and the matching feature (fin) on the finite verb. This constellation leads to a convergent derivation only if the finite verb moves within its own projection, to check the feature.

According to the standard analysis proposed by den Besten (1989) the finite verb moves to Comp in German (and Dutch) in sentences like (33). This is incompatible with the present analysis, because heads cannot move to pre-defined positions. Rather, they are displaced within their own projection, creating the landing site in the attraction process themselves.

(33) a. dass der Mann den Wagen sah (German)
that the man the car saw
b der Mann sah den Wagen
c. "(that) the man saw the car"

A standard argument[11] for the claim that V moves to Comp lies in the complementarity of overt complementizers and verb movement. In German (and Dutch), V2 movement takes place in clauses without an overt complementizer only, as (34) illustrates. This argument is invalid in all grammatical approaches that do not assume substitution operations in the strict sense, that is, in *all* current models. In approaches that adjoin heads to others heads, or work with remnant movement, additional mechanisms (discussed above) need to be invoked in order to to guarantee the complementary distribution of overt complementizers and V2.

(34) a. ich denke er hat sie eingeladen (German)
I think he has her invited
b. ich denke dass er sie eingeladen hat
I think that he her invited has
c. "I think (that) he has invited her"

Furthermore, the assumption that V moves to Comp and that movement is blocked when Comp is filled does not suffice to explain the full array of facts. V does not move in indirect questions and relative clauses, as shown by (35), although the Comp position is empty. This cannot be derived from a doubly-filled-Comp-filter constraint that rules out that Comp and [Spec,CP] be filled at the same time. Many German dialects allow the optional insertion of complementizers in (35), but none of them tolerates V2 movement in these contexts.

(35) a. ich weiss wen sie eingeladen hat (German)
I know who she invited has
"I know who she has invited"
b. ein Mann den sie eingeladen hat
a man who she invited has
"a man who she has invited"

Quite in general, (34) might turn out to be one of the most misguiding patterns in the recent history of syntax. V2 movement in embedded clauses is a process frequently

[11] The other argument in den Besten (1989) involves the placement of clitic pronouns, that follow the verb in main clauses, but are placed immediately after the complementizer in embedded clauses. The assumption that clitics are placed after the uppermost head in a clause suffices to explain the data. No reference to a particular head is necessary. There are also differences in the agreement form of the verb in Dutch that depend on whether the finite verb precedes or follows the subject. The explanation of this fact need not involve a distinction of Infl vs. Comp as the landing site of the verb, see, e.g., the analysis proposed in Ackema & Neeleman (2001).

attested in the world's languages (see Yiddish (36), but also Mainland Scandinavian, Icelandic, Kashmiri, and, if you wish to analyze them in this way, Hebrew, Spanish, Hungarian), and German is *exceptional* in needing to eliminate the complementizer in embedded V2 clauses. For theories that assume that the verb moves to Comp, the (abundant) existence of structures in which V2 movement and overt complementizers co-occur constitutes a major problem, as the discussion in Vikner (1995) shows. (36) and related structures constitute strong evidence against the idea that V2 movement targets Comp.

(36) Jonas bedoyert az dos bukh hob ikh geleyent (Yiddish)
Jonas regrets that this book have I read
"Jonas regrets that I have read this book"

Our approach is not so much influenced by parochial properties of German. V2 movement is triggered when the verb-Infl complex possesses a strong feature (say, fin) triggering movement and the matching feature at the same time. Then fin is checked by moving the finite verb within its own projection. There is no principled reason why the presence and nature of Comp should be relevant for this.

On the other hand, our approach does not *rule it out* that Comp (irrespective of its phonetic specification, however) may exert an effect on the applicability of head movement. Suppose that German and Dutch complementizers check a fin-feature of their complement. They are thus able to check this fin feature on the verb-Infl-complement when IP merges with Comp. Suppose that features that stand in a checking relation are, practically, identical. Then, the two occurrences of fin on verb-Infl can *both* be checked by Comp when IP merges with Comp[12]. Checking the fin-feature on verb-Infl by merging IP with Comp is, however, less costly than Münchhausen-style V2 movement, because one application of movement less is required. Complementizers therefore can, but need not, block the application of V2 movement.

[12] This presupposes a distinction between checking and erasure of the feature that seems standard.

3.3 The preverbal position

The assumption that the finite verb goes to Comp creates a further descriptive problem. If the verb is in Comp, the element preceding it is [Spec,CP]. In German (37) and in Kashmiri (38) (see Bhatt, 1999), the element preceding the finite verb can be a wh- or a focus phrase. Similar facts appear to characterize Hungarian (39) and Breton (40). If the position preceding the verb is [Spec,CP], this situation is expected, because [Spec,CP] is an operator position.

(37) a. den Fritz hat sie eingeladen (German)
the.ACC Fritz has she invited
"it is Fritz who she has invited"
b. wen hat sie eingeladen
who has she invited

(38) raath khyav tem batI (Kashmiri)
yesterday ate he food
"it was yesterday that he ate food"

(39) Kevés filmet néztem meg (Hungarian)
few film.ACC saw-I prt
"I saw few films"

(40) E bark en deus aret Yann (Breton)
his field PRT have-3m plowed Yann
"it is his field which Yann has ploughed"

However, the preverbal position can be filled by others elements, too, a fact that is hard to reconcile with the idea that the preverbal position is (always) [Spec,CP]. Thus, in German, topical elements may appear in preverbal position. We return to this observation below. In addition, German (and Kashmiri, see Bhatt 1999) tolerate the subject in preverbal position in out of the blue utterances, that is, even when the subject bears no pragmatic force, when it is neither a topic nor a focus. In fact, the clause-initial position of subjects is mandatory in unmarked sentences projected from transitive predicates. But since [Spec,CP] is an operator position, it can be reached by operator movement only. If the subject has no operator features, it cannot undergo operator movement. Subject-initial clauses thus at least suggest that main clauses are not *always* CPs, in line with a proposal made first by Travis (1984).

(41) [$_{IP}$ [Der Hans] [[$_{I}$ hat$_{k}$] [$_{VP}$ den Peter eingeladen t_{k}]]] (German)
the.NOM Hans has the.ACC Peter invited
"Hans has invited Peter"

(41), or, more precisely, the idea that V2 movement goes to Comp in some cases, but to Infl in others, raises two types of problems – one related to the analysis of verb movement, the other linked to the interpretation of [Spec,IP] and [Spec,CP]. Let us first discuss the verb movement issue, and turn then to the specifier position.

The verb movement problem is easy to explain, but difficult to resolve. In a minimalist system, the idea that V raises to Comp presupposes that Comp has a strong feature attracting the finite verb. In the light of (41), we would also need to assume that Infl possesses such a strong feature. Then, the question arises why the strong feature of Infl cannot attract the verb to the position following the subject in embedded clauses as well (see (34b) and (35)).

The discussion in Zwart (1993)[13] shows that a rather complex set of additional assumptions concerning feature checking is necessary if one wants to account for the difference in the attraction behavior of Infl between root and embedded clauses in terms of feature strength. Zwart (2001) follows a different route. He assumes that *abstract chains* linking V, Infl, and Comp are always formed in overt syntax. The uppermost element of such a chain must have a phonetic realization, but it does not matter whether that phonetic realization is, say, a complementizer, or a verb. The displacement of *phonetic features* is a last resort operation that applies only when there is no other way by which the uppermost position of the chain receives a phonetic matrix. Thus, an overt complementizer prevents the phonetic matrix of the finite verb from being displaced to Comp, and to any other position between V and Comp. There is no displacement of phonetic features but the one needed for lexicalizing the uppermost position in a chain. Thus, Infl receives a phonetic matrix by movement only if is the uppermost element of a chain.

[13] Zwart (1993) formulates the problem in a different way, however.

One difficulty of this model is that it fails to capture the data in which neither Comp nor Infl are phonetically filled (35). Furthermore, it is hard to imagine that Zwart's approach can be generalized. Thus, in order to explain why V2 movement is possible in the presence of a complementizer in, say, Yiddish, Zwart assumes that no chain is formed between Comp and Infl in such structures. Then, Infl is the highest element in the chain, and must be spelt out, which is done by moving the verb there. However, there are more examples of verb movement where the phonetic matrix of the verb shows up in an *intermediate* position. In Polish, V moves to Asp, but not higher. V moves out of VP since it precedes the clitic *go*, located in AgroP (42), but V cannot precede adverbs of the type associated with Aspect and higher ones. We find a pattern similar to English (43). It is hardly likely that such data show that no chain between the verb and AGR-S or Tense is created in Polish.

(42) ty widziałes go w parku (Polish)
you saw- him in the park

(43) a. *wy skończyliście prawie swoją pracę (Polish)
you finished almost your work

b. Jan by prawie skończyl swoją pracę
Jan would almost finish his work

The idea that V sometimes goes to Comp, and sometimes to Infl, thus creates descriptive difficulties. Let us turn to the second problem: the idea that non-operators are moved to a *subject* position is not convincing. This point was made by Bhatt (1999), and by Fanselow (in press, b), among others.

For German, it has been observed by Lenerz (1977) that unaccusative and psychological predicates, and passives of ditransitive verbs, have a normal word order pattern of their own: the dative precedes the nominative in pragmatically unmarked clauses. For verbal projections, this is easy to explain: in the cases under consideration, the nominative noun phrase is an underlying direct object. If noun phrases need not be raised in order to receive nominative Case in German, and if normal order in VP reflects thematic hierarchies, the dative before nominative order

causes no surprise. However, Hubert Haider was the first to point out[14] that the parallel order facts in German *main* clauses create a descriptive problem:

(44) Einem Kind wurde das Fahrrad gestohlen (German)
a.DAT child was the.NOM bike stolen
"a bike was stolen from a child"

(45) Einem Schauspieler ist der Text entfallen (German)
a.DAT actor is the.NOM text forgotten
"an actor forgot the text"

In an unmarked context, (44) – (45) are perfect. In fact, the nominative noun phrase could not be fronted in such a context. The fact that dative noun phrases appear in preverbal position is surprising, however, because they have no operator features in (44)-(45) (they need not bear focus or topic features for being well-formed), so that they cannot have moved in front of the verb by operator movement. Likewise, it seems to be standard wisdom that they cannot go to [Spec,IP], because German has no quirky subjects (see Fanselow, in press b, for a discussion). Even if we were willing to accept that datives *can* be subjects in German, this would not help us to explain why temporal and sentence level adverbs can also appear in clause-initial position in German, without bearing any specific pragmatic force. The idea is unattractive that a sentential adverb can be a "subject" in German in any interesting interpretation of the term.

(46) Am Sonntag hat ein Eisbär einen Mann gefressen (German)
on Sunday has a polar bear a man eaten
"On Sunday, a polar bear ate a man"

(47) Vielleicht hat der Schauspieler seinen Text vergessen (German)
perhaps has the actor his text forgotten
"Perhaps, the actor has forgotten his text"

The special behavior of sentence level adverbs had already been noted by Koster (1978) for Dutch. He also observes that sentence level adverbs of a complement clause cannot be placed into the matrix clause (48b). Since Dutch operator movement is of the long distance type, the ungrammaticality of (48b) suggests that sentence level adverbs cannot undergo operator movement. Therefore, *waarschijnlijk* and

[14] In a talk at the 1998 GGS meeting in Passau, Germany.

vielleicht have not reached the preverbal position by operator movement. But at the same time, they are not subjects.

(48) a. Waarschijnlijk is hij ziek (Dutch)
probably is he sick
b. *Waarschijnlijk zegt Jan dat hij ziek is
probably says Jan that he sick is

Bhatt (1999) notes that temporal adverbs resemble subjects in Kashmiri as well, in not having to be in focus when they occupy the preverbal position:

(49) rameshas cha azkal shiila khosh karaan (Kashmiri)
Ramesh is these days Sheila happy do
azkal cha rameshas shiila khosh karaan
"Ramesh likes Sheila these days"

What do these examples have in common? Bhatt (1999) observes that the element preceding the verb in V2 clauses would also appear in first position in clauses without verb movement. Subjects of transitive predicates and the dative arguments of unaccusative and psychological predicates are the highest arguments in the verbal projection. Thus, in case the prefield is not filled by a focal or [+wh]-element, the uppermost argument in (50a) that is present in the clause will move to the preverbal position. Frey (2001) shows that temporal adverbs may precede subjects in base order, that is, the order arising through merger, and the same holds for sentence level adverbs (50b). The observations concerning (46) – (49) suggest, then, the generalization in (51), which was first proposed by Bhatt for Kashmiri, but which seems to hold for German as well: it is always the element that would be the uppermost category in a "normal" clause that moves to preverbal position in a verb second clause – unless a focal or wh-element needs to go to the preverbal position.

(50) a. $[_{vP}\ \alpha\ v\ [_{VP}\ \beta\ [V\ \gamma]]]$

b. (sentence level adverb) (temporal adverb)$[_{vP}\ \alpha\ v\ [_{VP}\ \beta\ [V\ \gamma]]]$

(51) In $[\alpha\ V.FIN\ [_{\Sigma}\ \ldots]]$, α is the uppermost element of Σ, or bears a [+wh] or [+foc] feature.

Before we turn to additional data, the optimal way of capturing (51) should be identified.

Bhatt proposes to derive (51) as a Minimal Link effect. Suppose the verb is placed into some head position F. If F imposes no further requirements on its specifier, the Minimal Link Condition (53) implies that α of (52) can only be targeted by the highest element of XP. In languages like Kashmiri and German, this element may be the highest argument, or a high adverb. In languages like Icelandic or Breton, the highest element that moves may be a non-finite verb as well (because of the different position the verb occupies in vP), leading to Stylistic Fronting (see Holmberg 2000 for an MLC-based account) or to default verb-initiality, as in Breton (54).

(52) α F [$_{XP}$ ...]

(53) Minimal Link Condition: α cannot move to Σ if there is a β that could also move to Σ, such that β c-commands α

(54)

aret	en	deus	Yann	e	bark	(Breton)
ploughed	PRT	have-3m	Yann	his	field	

"Yann has ploughed his field"

How are V2 clauses with a focus or wh-element in preverbal position accounted for? Still following Bhatt (1999), we may assume that F may *optionally* carry a focus-feature. If so, α in (52) must be able to check that feature. Consequently, (53) requires in such a constellation that *the closest focus element* moves to preverbal position. It can skip any phrase that has no focus feature. Focus movement to clause-initial position is thus accounted for, and so is wh-movement, if we assume that it is a subcase of focus fronting.[15]

This analysis translates easily into the model developed here. Instead of assuming that a particular head F (representing finiteness, as in Fanselow (in press, b) or Mood, as in Bhatt 1999) attracts the finite verb, we postulate that the inflected verb comes from the lexicon as a complex category, bearing the features of V, Infl, and, say, M (ood)[16]. As an M, it possesses a strong feature checking Infl. Since the checking feature and the feature to be checked reside on the same head, Münchhausen-style

[15] Or, if we assume that F may carry an optional [+wh]-feature as well.

[16] Since the strong Infl feature is checked by a Comp when the latter is present, we can assume that all finite verbs bear this feature in German.

head movement is the only way of getting rid of the strong Infl feature. Therefore, V2 movement is triggered in the relevant structures.

The checking of the Infl-feature of the finite verb is, in principle, independent of the rest of the checking process. Thus, the derivation we propose is slightly different from the one in the references just mentioned, and more in line with Wunderlich (in press). In Bhatt (1999) and Fanselow (in press, b), the final steps in the derivation of a simple (non-operator) V2 clause are: (1) α is merged with some K and Fin merges with αK, (2) V is moved to Fin(Mood), and (3) α is attracted to [Spec,Fin]. What we propose here is one step shorter: when K has been formed, the Infl feature of the finite verb is checked by moving the verb, creating the constellation [verbK]. In the second step, the feature residing on the verb-infl-mood-complex related to α is checked, by either merging α with [verbK], or by moving α from K to the preverbal position. Thus, the preverbal α *moves* to its position in a verb-second clause only if it would do so in a non-verb-second context, too. In a sense, then, the verb literally "squeezes itself in" between α and K in V2 contexts.

Three further aspects need to be discussed before the explanation of verb second order may be considered complete. First, we mentioned it above that the preverbal phrase may be a focus- or wh-operator in German, the subject, or any other element that may be merged in the highest position of a clause - but a topical element is also licensed, as in (55a). This additional option is, in fact, an expected one, given what we have said so far:

In German, topical material may be placed into clause initial position by scrambling (55b), see Fanselow (2001), Grewendorf & Sabel (1994), Haider & Rosengren (1998), Müller & Sternefeld (1993), among others. Whatever is responsible for (55b) implies that (55a) is grammatical, too – both in the approach pursued here, and the more "traditional" one of Bhatt (1999) and Fanselow (in press, b), because these models (and only such types of models) imply the generalization (51).

(55) a. den Fritz mag niemand (German)
the.ACC Fritz likes nobody
b. dass den Fritz niemand mag
that the Fritz nobody likes
"(that) nobody likes Fritz"

From a semantic point of view, topichood does not correspond to an operator feature under a strict interpretation of that term. Thus, (55a) is not easily captured in V2 theories in which the initial element either must be the subject, or a category moved to an operator position. That the present model faces no difficulty with (55a) is an argument in its favor.

Holmberg (2000) mentions that non-subjects are focal in Icelandic when they show up in preverbal position. This is implied by our proposal, because Icelandic has no scrambling operation. Thus, an object cannot get in front of the subject on the basis of an A-movement -like processes such as the one exemplified in German (55b).

Dutch, however, seems to contradict the expectations derivable here. It has limited options for scrambling only– focused material may be placed into a preverbal position as in (56) under very restricted conditions (see, e.g., Neeleman 1994). However, any constituent (except unstressed pronouns and perhaps negation) can precede the verb in V2 position in Dutch, while no constituent can occur between the complementizer and a definite subject, see (57).

(56) dat ZO'n boek zelfs JAN niet zou lezen (Dutch)
that such a book even John not would read
"that even John would not read such a book"

(57) a. Het boek heeft Jan niet gelezen (Dutch)
the book has Jan not read
b. *dat het boek Jan niet heeft gelezen
that the book John not has read
"(that) Jan has not read the book"
c. Het meisje hebben we het boek gegeven
the girl have we the book given
d. *dat het meisje we het boek hebben gegeven
"(that) we have given the book to the girl"

Since (57b,d) are not well formed, we seem to have no source from which to generate (57a,c), because the preverbal elements are not focal. Instead of assuming a topic-operator feature (in spite of its semantic implausibility), we may, however, analyze

(57a,c) in terms of a left-dislocated phrase (as in (58)). The left dislocated phrase could be base-generated in a topic position, and be linked to the rest of the clause by an invisible operator (see, e.g., Zwart 1993).

(58) dat boekje dat leg ik even neer (Dutch)
that book that pit I adv down
"I will just put down that book"

This analysis eliminates the descriptive problem posed by Dutch, but it raises the issue of why Icelandic topics cannot be placed in preverbal position by the same route. We will leave this issue open here.

Müller (2003) argues for (51) from a different perspective. His observation concerning the distribution of clausal complements in clause initial position (which is also independent of V2) supports our analysis.

The second array of data that we need to discuss concerns the fact that (51) cannot be strengthened into a bi-conditional. It is *not* the case that whatever can appear in the first position following a complementizer in an embedded clause may also appear in the initial position of a V2 clause. The first set of structures does not pose a serious problem for our analysis First, unstressed elements may be clause-initial in CPs with a complementizer, but they cannot occupy the preverbal position in a verb second clause:

(59) a. dass sich jeder irren kann (German)
that refl everybody err can
"that everyone can be wrong"
a'. *sich kann jeder irren
b. es weint
it weeps
"(s)he weeps"
c. dass wer gekommen ist
that indef come is
"that someone came"
c'. *wer ist gekommen
c". wer aus Hamburg ist nicht gekommen
indef from Hamburg is not come
c". "someone from Hamburg has not come"
d. dass ja niemand damit rechnen konnte
that ptc. nobody there-with reckon could
"that nobody could reckon with that"

d`. *ja konnte niemand damit rechnen

(59a) illustrates that unstressed pronouns cannot be placed into preverbal position, unless they are subjects (59b). This observation figured prominently in Travis (1984) when she tried to establish the CP-IP distinction for German main clauses. Data such as (60) suggest, however, that the ban against weak object pronouns in first position is not an absolute one (see also Gärtner & Steinbach 2001).

(60) Ihr Geld ist ja nicht weg. Es haben jetzt nur andere
your money is yes not gone it have now only others
"Your money isn't really gone. It is only others that have it now"

(59a,b) and (60) can be captured along the following lines: placing weak pronouns into the so-called Wackernagel position (following the first head) is not obligatory. Pronouns may be merged in their argument positions, and remain there. In contrast to accusative pronouns, subject pronouns can be the first element in a clause by virtue of being merged there. Thus, because of (51), they can also be placed into preverbal position.

Object pronouns cannot be merged as arguments in clause-initial positions. When they are weak, they cannot be focal, so a focus feature cannot transport them into clause-initial position either. To a limited degree, they may undergo scrambling, which may be responsible for (60). Normally, however, they are preposed by a movement that places them into the "Wackernagel"-position. If this movement yields a well-formed result only when the clitic ends up after the uppermost head of a clause, (59a') simply cannot arise.

There appears to be an additional (weaker?) ban against stressless elements in preverbal position that affects (non-pronominal) subjects as well, as (59c') shows. Elements such as indefinite *wer* and particles like *ja* cannot appear in preverbal position, but whenever stress may go to a different entity (as in (59c'')), the sentences become fine. Thus, when intonation is taken into account, (59) can be explained (see Müller 2003, for a different analysis).

Consider, finally, the data in (61) and (62). In an impersonal passive (61a) or with sone nominal predicates (62a), a clause may consist of a finite verb and a non-finite predicate only. If (51) were a biconditional, one would expect that an unmarked main clauses might look like (61b) and (62b). This expectation is not fulfilled. The participle must be interpreted as focused in (61b), while (62b) is hardly acceptable at all. In a pragmatically unmarked clause, an expletive needs to be inserted into the preverbal position.

(61) a. dass getanzt wird (German)
that danced was
"that one danced“
b. getanzt wird
c. es wird getanzt
there is danced

(62) a. dass Krieg ist
that war is
"that there is war"
b. ?Krieg ist
c. es ist Krieg
there is war

Icelandic Stylistic Fronting and Breton show that there is no universal ban against a non-finite predicate showing up in preverbal position in an unmarked clause. The ban against (62b) is thus a mystery in any approach in which the highest element of Σ is moved to [Spec,FP] after the finite verb was moved to F.

3.4 The Second Position

The property of V2 constructions that still calls for an explanation is the verb *second* property itself. While our model predicts that the verb must squeeze itself in between two positions of a clause, it does not predict that it must go behind exactly the first constituent. The Münchhausen-feature of the verb must be checked, but it can be so at any time in the derivation. Its checking is, in principle, independent of any other processes creating specifiers by merging or moving categories.

Note that this is far from being a problem that is confined to our theory. Fir example. approaches that assume that there is more than one head position above the

verb phrase (see, e.g., Rizzi 1997) run into a similar difficulty. The preverbal element of a V2 clause may have quite different semantic/pragmatic functions (see above). In a model following Rizzi, it will appear in different specifier positions α, ..,γ in (63). One then needs to make sure that the verb moves to exactly the head position corresponding to the highest specifier filled. Otherwise, something different from a V2 structure would be generated.

(63) [α A [β B ... [γ [C vP]]]]

Likewise, remnant movement theories must make sure that a single phrase must become "light" in exactly the right type of position.

A (partial) answer to the question as to why the finite verb moves to second (and not to third or first) position can be found by a reconsideration of the behavior of strong uninterpretable features. From Chomsky (1995:234), we have taken over the assumption that strong features of H cannot be tolerated in a projection not headed by H itself. When a complex structure such as (64) is created by Münchhausen-style movement of a head K out of Σ, both α and β are projections of K, so that it might seem irrelevant how many other strong features are still present on K when K leaves Σ. This property is fatal if one wants to explain that K must be the *second* element in the final structure.

(64) $[_{\alpha}$ K $[_{\beta}$ Σ]]

For obvious reasons, the assumption that K must not possess any strong features but the one licensing its own movement at the point when (64) is formed would not help either, because this would result in a *head-first* structure. Rather, a closer look at the featural composition of finite verbs seems to be called for.

In a V2 construction, the finite verb enters the derivation as a(n) M(ood)-element that has the categorical specification Infl as well. M has a strong uninterpretable Infl-feature that must be checked by raising the finite verb itself. It is natural to assume that this set of features is a structured object, such that the feature structure of the finite verb is an array of features linked to M (among them a strong Infl-checking

feature), of features linked to Infl (among which there may be various other features which need to be checked), and perhaps of feature bundles linked to verbal properties. The finite verb enters the derivation with a feature structure [[[[Σ] f] g] h], and we have a chance of coming close to deriving the V2 effect if the feature checking process in the syntax respects this feature structure, such that features related to "lower" bundles of features are checked before higher feature bundles are addressed[17]. If correct, this means that the position of the finite verb is determined by the features related to M only, as required. If the highest feature complex of a finite verb with Münchhausen-movement properties selects one specifier, the verb places itself into either second or first position, depending on the order in which its EPP-feature and its Münchhausen-feature are checked.

This results falls short of exactly deriving second position placement (because it does not exclude clause-initiality), but perhaps, it is not an incorrect one. It may well be that the choice among the two options comes from a further source. Second position effects can be observed in a variety of languages. There are verb-second phenomena that cannot be accounted for in terms of movement to Comp in a direct sense. Rather, it seems that the verb moves to Infl only – yet, a category different from the subject precedes the verb. Baylin (to appear) argues that the second position effect we observe in Russian for non-focal (non-wh-) element in clause initial position involves the placement of the verb in Infl:

(65) gazety darjat profesora studentam (Russian)
newspapers gave professors students
"the professors gave the newspapers to students"

(66) étu knigu Ivan c`itaet c'asto
this book Ivan reads often
"It is this book that Ivan reads often"

In Hebrew (see Borer 1995) there is an optional verb second effect in main and embedded clauses involving topicalization. Borer argues that the subject is not in [Spec,IP] in verb second clauses.

[17] In effect, this means that the c-command relations among the features of a head must correspond to those we find realized in the syntactic tree.

(67) (Ran amar se-) (Hebrew)
Ran said that
a 'et ha-gvina ha-xatul *xisel* lifney Se-higanu
the cheese the cat finished before we-arrived
b. 'et ha-gvina *xisel* ha-xatul lifney Se-higanu

Spanish data are most revealing. Spanish shows a verb-second effect in clauses involving focus movement and wh-movement, in the sense that the subject does not intervene between the verb and the operator (as it should, given that Spanish is an SVO language).

(68) a. diez cafecitos toma Drea todos las mañanas (Spanish)
10 coffes takes D all the mornings
b. *diez cafecitos Drea toma todos los mañanas
"Drea takes 10 coffees every morning
c. Briana preguntó (que) qué habiá comprado Mara ayer
Briana asked that what has bought Mara yesterday
"Briana asked what Mara bought yesterday"

Note that Spanish appears to lack at least "long" V to Infl movement, because adverbs such as *apenas* precede the verb, as illustrated by (69). The relative order of verb and adverb is, however, preserved in a verb-second structure! Thus, (70) must not involve *any* verb movement at all, although it illustrates *a second position* effect (if we disregard adverbs).

(69) La viejita apenas puede leer los periódicos (Spanish)
the old rarely can read the newspapers
(70) a quién casi nunca le manda regalos la abuela
to whom almost never cl sends presents the grandma
"to whom does Grandma almost never send presents?"

How can we explain these data? One may follow standard wisdom and assume that topical elements may occupy [Spec,IP] in Russian (see Baylin, to appear) or Hebrew (see Borer, 1995), and that wh- and focus elements may go to the very same position (Suñer 1994). Technically, there is not much one can object to this, but such approaches imply that [Spec,IP] is a notion quite devoid of content. One also would have to explain why, e.g., wh-phrases go to [Spec,IP] in certain languages, but to a different specifier position in others.

The obvious alternative is to assume that there may be *surface constraints* on the complexity (and nature) of the second element in a clause. Given the particle stranding facts of German discussed above, these seem unavoidable. If the second element must be a finite category, the constraint may require that the subject does not move to [Spec,IP] when the verb would end up in third position otherwise in languages that lack verb movement to Infl (Spanish, and, perhaps, Russian), or that the EPP-feature is checked last in language that have verb movement (German), creating a specifier-first constellation.

4 References

Ackema Peter, Ad Neeleman & Fred Weerman. 1993. Deriving Functional Projections. Paper, presented at NELS 23.

Ackema, Peter & Ad Neeleman 2001. Context-sensitive Spellout and Adjacency. Paper, presented at the 5th Workshop on OT Syntax, Potsdam.

Alexiadou, Artemis. 1997. *Adverb Placement.* Amsterdam: Benjamins.

Alexiadou, Artemis & Gisbert Fanselow. In press. On the Correlation between Morphology and Syntax: The Case of V to Infl. In: J.W. Zwart & W. Abraham, Proceedings of the Germanic Syntax Workshop. Groningen, 2000.

Baker, Mark. 1988. *Incorporation.* Chicago: The University of Chicago Press.

den Besten, Hans. 1989. *Studies in West Germanic Syntax.* Amsterdam: Atlanta.

Besten, Hans den, & Gert Webelhuth. 1987. Remnant topicalization and the constituent structure of VP in the Germanic SOV languages. Paper, presented at the X. GLOW Colloquium, Venice.

Besten, Hans den, & Gert Webelhuth. 1990. Stranding. In: *Scrambling and Barriers,* eds. Günter Grewendorf & Wolfgang Sternefeld. 77-92. Amsterdam: Benjamins.

Bhatt, Rakesh Mohan. 1999. *Verb Movement and the Syntax of Kashmiri.* Dordrecht: Kluwer.

Borer, Hagit. 1995. The Ups and Downs of Hebrew Verb Movement. *Natural Language and Linguistic Theory* 13: 527-606.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In: Ken Hale & Samuel J. Keyser (eds). *The View from Building 20.* Cambridge, Mass.: MIT-Press. 1-52.

Chomsky, Noam. 1995. *The minimalist program.* Cambridge, Mass.: MIT-Press.

Chomsky, Noam. 1999. Minimalist Inquiries: The framework. Ms., MIT.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads.* New York: Oxford UP.

Engels, Eva, in prep. Preverbal adverbs in English. Bidirectional optimization of adverb placement and interpretation. Ms., University of Potsdam.

Ernst, Thomas. 2001. *The Syntax of Adjuncts.* Cambridge: CUP.

Fanselow, Gisbert (2001). Features, q-roles, and free constituent order.*Linguistic Inquiry* 32,3.

Fanselow, Gisbert, in press a. Against Remnant VP Movement". In: Alexiadou, A., S. Barbiers, & H.-M. Gärtner. Proceedings of the Workshop on Remnant Movement.

Fanselow, Gisbert, in press b. Quirky Subjects and Other Specifiers.

Fanselow, Gisbert & Damir ?avar. 2001. Remarks on the Economy of Pronunciation. In: G. Müller & W. Sternefeld (eds.) *Competition in Syntax.* Amsterdam: Benjamins.

Frey, Werner 2001. Über die syntaktische Position der Satztopiks im Deutschen'. Unpublished Ms., Zentrum für Allgemeine Sprachwissenschaft, Berlin.

Gärtner, Hans Martin and Markus Steinbach (2001). Conditions on the Scrambling and Fronting of Reduced Pronominals in Dutch and German. Unpublished Ms., Potsdam and Mainz.

Grewendorf, Günther. 2001. Multiple Wh-Fronting. *Linguistic Inquiry* 32: 87-121.

Grewendorf, Günther & Joachim Sabel. 1994. Long scrambling and incorporation. *Linguistic Inquiry* 25: 263-308.

Grimshaw, Jane. 1991. Extended Projection. Ms., Brandeis.

Haider, Hubert. 1988. Matching Projections. In: *Constituent Structure. Papers from the 1987 GLOW Conferece.* Anna Cardinaletti, Guglielmo Cinque & Giovanni Guisti (eds). Venice: Annali di Ca'Foscari 27:101-121-

Haider, Hubert & Rosengren, Inger. 1998. Scrambling. *Sprache und Pragmatik 49,* Lund.

Holmberg, Anders. 2000. Scandinavian Stylistic Fronting: How any category can become an expletive. *Linguistic Inquiry* 31: 445-483.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT-Press.

Koeneman, Olaf. 2000. The flexible nature of verb movement. Utrecht: LOT.

Koopman, H. 1994. Licensing Heads, in N. Hornstein and D.Lightfoot, eds., *Verb Movemen*t, Cambridge University Press.

Koopman, Hilda. 1996. The Spec Head Configuration. Ms., UCLA.

Koopman, Hilda & Anna Szabolcsi. 2000. *Verbal Complexes.* Cambridge, Mass.: MIT-Press.

Koster, Jan 1978. *Locality Principles in Syntax.* Dordrecht: Foris.

Larson, Richard K. 1988. On the Double Object Construction. *Linguistic Inquiry* 19, 335-391.

Lenerz, Jürgen 1977. *Zur Abfolge nominaler Satzglieder im Deutschen.* Tübingen: Narr.

Mahajan, Anoop. 2001. Word Order and Remnant VP movement. Ms., UCLA.

Meinunger, André, 2001. Restrictions on Verb Raising. Ms., ZAS, Berlin.

Müller, Gereon. 1998. *Incomplete category fronting*. Dordrecht: Kluwer.

Müller, Gereon. 2003. Verb-Second as vP-First. In *Syntax at Sunset 3. Head Movement and Syntactic Theory,* Anoop Mahajan (ed.).UCLA/Potsdam Working Papers in Linguistics 10. 116-160. University of California, Los Angeles.

Müller, Gereon & Wolfgang Sternefeld. 1993. Improper movement and unambiguous binding. *Linguistic Inquiry* 24: 461-507.

Neeleman, Ad. 1994. Scrambling as a D-structure phenomenon. In: *Studies on scrambling,* ed. Norbert Corver and Henk van Riemsdijk, 387-429. Berlin: Mouton de Gruyter

Pesetsky, David & Esther Torrego, 2000. T-to-C Movement: Causes and Consequences. Ms.

van Riemsdijk, Henk. 1998. Head Movement and Adjacency. *Natural Language and Linguistic Theory* 16:633-678.

Rizzi, Luigi 1997. The Fine Structure of the Left Periphery. In *Elements of Grammar*, Liliane Haegeman (ed.) 281–337. Dordrecht: Kluwer.

Roberts, Ian. 1994.

Rudin, Catherine. 1988. On multiple questions and multiple wh-fronting. *Natural Language and Linguistic Theory* 6: 445-501.

Speas, Margaret. 1990. *Phrase structure in natural language*. Dordrecht: Kluwer.

Suñer, Margarita. 1994. V-movement and the licensing of argumental wh-phrases in Spanish. *Natural Language and Linguistic Theory* 12: 335-372.

Thiersch, Craig. 1985. VP and Scrambling in the German Mittelfeld. Ms., University of Tilburg.

Travis, Lisa 1984. Parameters and Effects of Word Order Variation. Unpublished doctoral dissertation, MIT, Cambridge, MA.

Tuller, Laurice. 1992. The syntax of postverbal focus constructions in Chadic. *Natural Language and Linguistic Theory* 12: 303-334

Vikner, S. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. New York, Oxford UP.

Wunderlich, Dieter 2001. The force of lexical case: German and Icelandic compared. In *The Wor*d, Sharon Inkelas and Kristin Khanson (eds.). Stanford: CSLI publi-cations. To appear.

Zwart, C. Jan-Wouter. 1993. *Dutch syntax*. Unpublished doctoral dissertation, University of Groningen.

Zwart, C. Jan-Wouter. 2001. Syntactic and Phonological Verb Movement. *Syntax* 4:36-62.