Ralf Vogel (ed.)

Three papers on German verb movement

Linguistics in Potsdam

Series Editors:

Susann Fischer
Institut für Linguistik Potsdam

Ruben van de Vijver
Institut für Linguistik Potsdam

Ralf Vogel
Institut für Linguistik Potsdam

Linguistics
in Potsdam
22

Ralf Vogel (ed.)

Three papers on German verb movement

Introduction

In recent generative syntactic debate, the view of verb movement as head movement "has gained a bad reputation " (Fanselow, this volume) on conceptual grounds. But it is still an open question to what extent the alternatives that have been proposed fare better. Those alternatives are first of all remnant XP movement and postponing head movement to the level of Phonological form (PF).

This volume offers new arguments and perspectives in the ongoing debate about the optimal analysis of verb movement, mainly, but not exclusively, in German. Fanselow and Meinunger deal with verb second (V2) movement in German main clauses. Fanselow argues that head movement of the substitution type follows the standard minimalist conceptions of MERGE and MOVE and is therefore not subject to the same objections as head movement as head adjunction which violates Chomsky's minimalist extension condition, operates countercyclically, and fails to let the moved head c-command its trace. Fanselow argues for V2 movement as head movement of the substitution type. Meinunger discusses a restriction on V2 movement imposed by phrases like "*mehr als*" ('more than'), as in "*Der Wert hat sich weit mehr als verdreifacht*" ('the value has far more than tripled') where V2 movement is ruled out (cf. *'Der Wert verdreifachte sich mehr als'). Meinunger claims that this restriction is best analysed in phonological terms: the preposition/complementiser "*als*" acts as a prefixal clitic to its host, the finite verb, which therefore may not move without it. With

respect to the V2 debate, Meinunger argues for an interface perspective. He shows that V2 is restricted from both the conceptual and the phonological interface. Vogel, finally, discusses the syntax of clause-final verbal complexes and their dialectal variation in German. He compares three different syntactic analyses, a minimalist head movement analysis, a minimalist XP movement analysis, and an Optimality theoretic PF movement analysis. The three accounts are evaluated relative to the additional assumptions they have to make, the complications they face and how they fit the observations. Vogel argues in favour of the phonologically oriented OT analysis because of its ability to create a direct link between the coming about of a particular word order pattern and its basically phonological trigger. Each of the three papers recognises the relevance of surface forms in the analysis of German verb movement. They differ, however in the extent to which phonological aspects take part in the explanations they offer. The papers by Fanselow and Vogel are slightly modified versions of two papers by the same authors which are included under the same titles in: Anoop Mahajan (ed.), *Syntax at Sunset 3. Head Movement and Syntactic Theory*, UCLA/Potsdam Working Papers in Linguistics 10, Los Angeles: UCLA, 2003.

I hope you enjoy reading the papers in this volume,

Potsdam, February 11, 2004,

Ralf Vogel

Contents

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

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)   [$_{CP}$ α [$_C$[$_T$[$_V$ VERB] ∅] ∅] [$_{TP}$ β t$_T$ [$_{VP}$ t$_V$ γ]]]

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 tVP]   saare phal   tIP (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) $[_{YP}[_Y \text{ Y }][_{XP} \ldots X \ldots]]$ $\Longrightarrow$

                       $[_{YP}[_Y \text{ X Y }][_{XP} \ldots \cancel{X} \ldots]]$

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 $[\alpha\Sigma]$ 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~~$_+$-$[_{V'}$ ~~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 $[_\alpha H \alpha]$. Therefore, in $[_{YP} [_Y X Y] [_{XP} \ldots. \overline{X} \ldots]]$, 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 $[_\alpha H \alpha]$ 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 $[_\alpha \text{ H } \alpha]$ 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) $[_{XP} [_{YP} ... \text{ Y } ...] \text{ X } [_{ZP} .... \text{ t}_{YP} ....]$

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. $\alpha\beta\gamma$ [... ~~$\alpha\beta\gamma$~~ ...]

        b.   ~~$\alpha\beta\gamma$~~   [... $\alpha\beta\gamma$ ...]

        c.   ~~α~~$\beta\gamma$ [... α~~$\beta\gamma$~~ ...]

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) $[_{YP}[X[Y][_{XP} .... t_X ....]]]$

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[$_α$ … [$_β$ …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 $[_{kP} \ldots k \ldots]$ – 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) $[_{XP} \ldots X \ldots] \rightarrow$

        $X [_{XP} \ldots t_X \ldots]$

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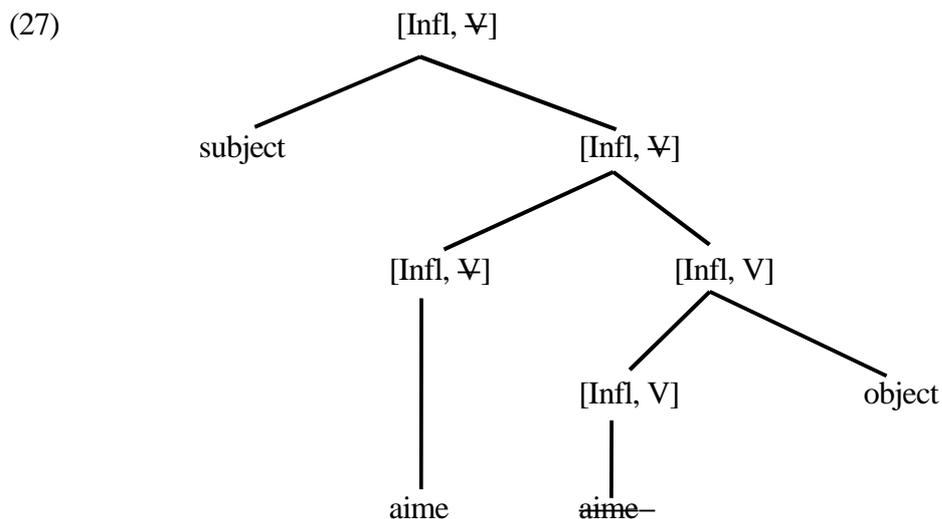
In a language like English, *love* is entered as a verb (and not as a V-Infl) into the syntactic representation. Therefore, an empty Infl element must be selected in the numeration if an IP is to be generated. The resulting structure (28) is quite classical.

(28)

```
                              [Infl]
                             /      \
                        subject      [Infl]
                                    /      \
                                [Infl]      [V]
                                  |        /   \
                                 |       [V]     object
                                 |        |
                                 Ø      loves
```
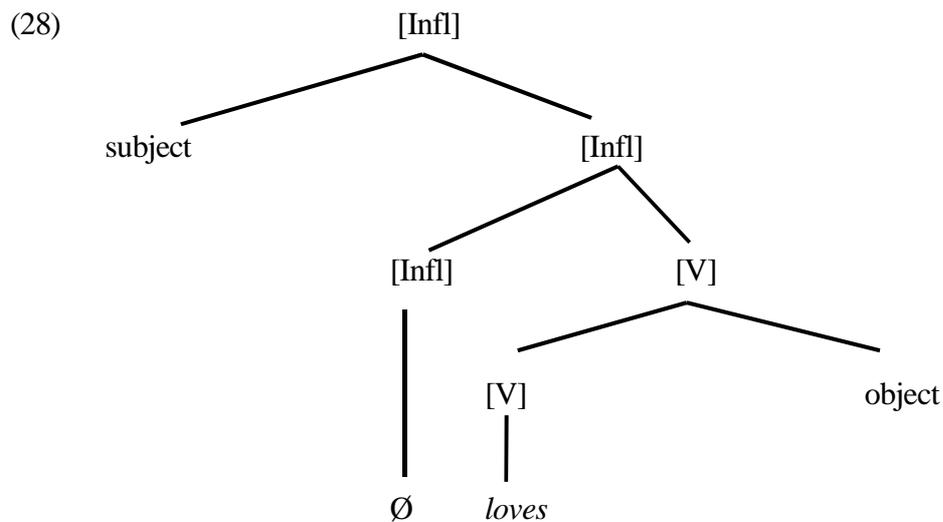
If nothing else is said (but see the final subsection), the V-Infl entities we assume for languages with verb movement just need to move *at some point* in order to get rid of the strong V feature they possess, but they may do so *at any point* in the derivation. What we expect to find in the empirical 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 $[\alpha\ V.\text{FIN}\ [_{\Sigma}\dots]]$, 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) [$_\alpha$ A [$_\beta$ B ... [$_\gamma$ [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 $[[[[\Sigma]\ f]\ g]\ h]$, 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 Movement*, 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 Word*, 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.

Interface restrictions on verb second

André Meinunger

ZAS Berlin

0   Introduction

Verb second (henceforth V2) in Germanic goes on being a mystery. It has been a focus of research for decades and still enjoys a lively and controversial debate. The fact that the German finite verb appears in the sentence final position ('rechte Satzklammer', right sentence bracket) in subordinate clauses and in the second position in main clauses ('linke Satzklammer', left s.b.) has all possible analyses: the canonical one is still standard and goes back to den Besten's analysis of Dutch (1977/1983). According to it the finite verb raises via head-to-head movement from its base position under V° to the head-initial C° node. Recently three different proposals are on the market: (i) Fanselow (2002, this volume) revives Holmberg's 'Münchhausen' idea, (ii) Müller (2002) proposes an analysis that does without head movement at all (see also Koopman and Szablolcsi (2000)), and Chomsky (2001) banishes verb second into the phonological component getting it outside the core area of syntax. A good overview can also be drawn from the volume 'Dimensions of movement' where all current sorts of movement approaches are presented (Alexiadou et al. 2002).

This article does not aim at defending the traditional approach or one of the new proposals, nor will it deliver yet another one. The intention of the present

work is to point at two relatively unrelated observations, the correct analysis of which might help to decide for the correct approach.

What will become clear is that the V2 rule is not blind with respect to the interfaces. It is not a merely syntactic movement rule whose output is (to be) interpreted by the phonological component on the one hand and by the conceptual interpretative system on the other. It is rather such that both systems restrict V2, they have a direct influence on the verb getting to its position. We will see that parsing strategies guide V2: phonological processing as well as interpretation related parsing.

The paper is organized as follows. In the first part a special construction is presented and discussed. This relatively neglected pattern is given an analysis that recurs to phonological wellformedness constraints. It is shown that a tree gets a syntactic and a phonological interpretation, sometimes parallel and corresponding to each other – sometimes diverging. If either one fails to meet wellformedness conditions the derivation crashes. The second part discusses the behavior of speech act adverbials. These may appear in a position before a canonical CP, verb third emerges (V3). Often the same adverbials, however, are also fine in the canonical 'Vorfeld' (prefield) giving rise to a regular V2 structure. Sometimes V2 is not permitted. The reason for this pattern is a parsing effect that has a similar impact as other better known garden path effects. Both constructions thus show that V2 is restricted from outside the core syntax.

1   Elements that block verb raising – a discussion

1.1 Haider's observation

The crucial observation that led to the present investigation, which I will discuss in the first part, was made in Haider (1997). The following story is very much a

recapitulation of Meinunger (2001). Haider considered the comparison constructions in (1) – (3) (slightly modified for our purposes). (1) consists of a periphrastic tense construal where the finite verb is an auxiliary that has moved to the V2 position. (2) is the unfelicitous attempt to move the full main verb of a simple tense into the V2 position.

(1)   Der   Wert hat sich weit <u>mehr als</u>   verdreifacht.
       the value has   self far   more than tripled
       'the value has far more than (only) tripled'

(2)   *Der Wert verdreifachte sich weit <u>mehr als</u>.

(3)   weil sich der Wert <u>mehr als</u> verdreifachte
       because 'self' the value more then tripled

Haider's argumentation goes as follows. The comparative element <u>mehr</u> (the German counterpart of <u>more</u>) must c-command its associate at s-structure. Since in the case at hand <u>mehr</u> quantifies over the verb <u>verdreifach(en)</u>, the verb must stay in its base position inside the VP. This is not given in (2). A subordinate structure, where any verbal form occurs rightperipherally, i.e. not only in periphrastic tenses, renders the sentence grammatical because the verb does not raise over <u>mehr</u> (3) – at least in a stucture without a head final I°, which Haider assumes.

Haider's explanation, however, is not completely compelling for several reasons: (i) under other circumstances verb movement does not care about scopal configurations depending on (so-called) d- or s-structure; (ii) other comparative constructions involving <u>mehr</u> allow the associate to appear to the left of it, and (iii) the same phenomenon of impossible verb movement can be observed in many other cases that do not involve comparison.

Ad (i): Haider himself agrees that negative polarity verbs (NPIs) do not obey an s-structure requirement on the licensor c-commanding the polarity verb. The well-known case is <u>brauchen</u>, (something like <u>need</u> or <u>must</u> under negation and question operators) which can appear in the V2 position leaving its licensor (<u>nicht</u> in (4)) behind.

(4)   Man braucht$_i$ es *(nicht) zu sehen e$_i$.
       one   need   it   (not)   to see _
       'one need not see it'

It would be a mystery if the 'mehr als' comparative construction required s-structure c-command whereas negative polarity did not. This is even more surprising for NPI licensing is known to be a very strong s-structure phenomenon. Comparative constructions on the other hand are less strict (ad (ii)). In certain examples the comparative associate may precede <u>mehr</u>:

(5)   *Als Jo   / ^{ok}im Vergleich zu Jo   / ^{ok}verglichen mit Jo
       *than Jo / in comparison with Jo / compared   to   Jo
       hat Eva mehr Aufgaben gelöst.
       has Eva more tasks   solved
       'Compared to Jo, Eva solved more problems.'

The interesting thing, however, is that verb movement of the sort discussed here can be observed in more environments that are not related to comparison at all (iii). Thus, it seems that there is no blind mechanism that raises the finite verb in main clauses to C° automatically and irrespectively of all factors.

1.2   The other constructions

In the following paragraph I list a number of constructions (hopefully almost exhaustive) that behave very much like <u>mehr als</u> above, that is, they contain verbal forms in which the main verb must be placed after a certain element in the middle-field. In these cases V2-movement of the full verb form into C° is ungrammatical. Some constructions below belong to informal spoken registers. The contrast, however, is very sharp:

(6)   a.   Der Angeklagte hat <u>so gut wie</u> gestanden.[1]

              the accused   has <u>so gut wie</u> confessed

       b.   weil der Angeklagte <u>so gut wie</u> gestanden hat / gestand

       c.   *Der Angeklagte gestand <u>so gut wie.</u>

              'The accused almost confessed (his crime).'

(7)   a.   Der Wert hat sich nun <u>soviel wie</u> verdoppelt.

              the value has self now <u>soviel wie</u> doubled

       b.   weil sich der Wert nun <u>soviel wie</u> verdoppelt hat / verdoppelte

       c.   *Der Wert verdopplete sich nun <u>soviel wie.</u>

              'In the meantime the value almost tripled.'

(8)   a.   Der Wert hat sich <u>mehr denn</u> verdoppelt.

              the value has self <u>mehr denn</u> doubled

       b.   weil sich der Wert <u>mehr denn</u> verdoppelt hat / verdoppelte

---

[1]   In order to prevent redundancy I outline the patterns displayed in examples (6) – (13). The a. example with perfect tense in main clause shape is glossed. The b. example in each block is not. It has subordinate shape with perfect and simple past tense separated by '/'. The c. example, which is always ungrammatical, is an attempt to raise the simple past tense verb to the V2 position. A tentative translation follows. Since in German simple past and perfect tense can be used synonymously, one translation is sufficient.

    c. *Der Wert verdopplete sich <u>mehr denn.</u>

        'The value more than merely doubled.'

(9)   a. Der Besuch hat <u>sowas von</u> geprahlt.

        the guest has <u>sowas von</u> boasted

    b. weil der Besuch <u>sowas von</u> geprahlt hat / prahlte

    c. *Der Besuch prahlte <u>sowas von.</u>

        'The   guests were boasting in such an incredible manner.'

(10) a. Der Hamster hat <u>so eine Art von</u> gehustet.

        the hamster   has <u>so eine Art von</u> coughed

    b. weil der Hamster <u>so eine Art von</u> gehustet hat / hustete

    c. *Der Hamster hustete <u>so eine Art von</u>.

        'The hamster sort of coughed.'

(11) a. Die Kommission   hat <u>nichts als</u> gemeckert.

        the commission   has <u>nichts als</u> grumbled

    b. weil die Kommission <u>nichts als</u> gemeckert hat / meckerte

    c. *Die Kommission meckerte <u>nichts als.</u>

        'The commission didn't do anything but grumble.'

(12) a. Die Kommission   hat <u>nichts außer</u> gemeckert.

        the commission   has <u>nichts außer</u> grumbled

    b. weil die Kommission <u>nichts außer</u> gemeckert hat / meckerte

    c. *Die Kommission meckerte <u>nichts außer.</u>

        'The commission didn't do anything but grumble.'

(13) a. Die Kommission   hat <u>weder</u> gemeckert,   ...noch...[2]

        the commission   has neither grumbled   ...nor...

    b. weil die Kommission <u>weder</u> gemeckert hat / meckerte,   ...noch...

    c. *Die Kommission meckerte <u>weder</u>,   ...noch...

---

[2] This example is a bit different and less clear. <u>weder</u> can induce V2 (13 d.) and save the construction.

    d. <u>weder</u> meckerte die Kommission,   ...noch...

        'The commission neither grumbled, nor...'

All examples become grammatical if something follows the underlined chunks. In some cases it even suffices to have a separable prefix of a morphological complex verb (17):

(14)   Der Angeklagte gestand <u>so gut wie</u> gar nicht.

(15)   Der Besuch prahlte <u>sowas von</u> {oft / unverschämt / peinlich...}.

(16)   Die Kommission meckerte <u>weder</u> davor, (noch danach).

(17)   Unser Besuch gibt <u>sowas von</u> an.

A similar pattern of verb movement blocking can be observed in other languages. French has a construction consisting of a negative clitic element which in most cases is morphologically attached to a finite verb in I° and semantically associated with a constituent that starts with the element <u>que</u>. Such a construction triggers a focus reading on the relevant phrases – similar to <u>only</u> in English. In perephrastic tenses and constructions movement of the full verb across <u>que</u> triggers ungrammaticality (19). Putting something after <u>que</u> makes the sentence grammatical (20) again. In case the verb is supposed to be the target of focus in a non-periphrastic construction, a dummy verb (<u>faire</u> ? do-support) has to be inserted (21).

(18)   Il n'a que bossé.   (French)

       he not-has 'que' hard-worked

       'He didn't do anything but work.'

(19)   *Il ne bosse que.

       he not hard-work   that

'He doesn't do anything but work.'

(20)   Il ne bosse que le dernier jour avant l'examen   / pour son frère...

       he not hard-work 'que' the last day befor the exam / for his brother

       'He works hard only the day before his exam / only for his brother.'

(21)   Il ne fait que bosser.

       he not does 'que' work-hard

        'He doesn't anything but work.'

Similar constructions also occur in English. The situation is different insofar as that there is no verb movement of full verbs. However, some quantifying elements that target the verb must precede it, whereas others may also follow the verb.

(22)   He <u>somehow</u> wheezed

(23)   He <u>sort of</u> wheezed.

(24)   He wheezed <u>somehow.</u>

(25)   *He wheezed <u>sort of</u>.   (under unmarked flat intonation)

Here one can also add examples inspired by Kajita, who discovered the peculiarities of such constructions as early as 1977. He is, however, not concerned with the contrast of the examples in (26) to (29).

(26)   ?This amount <u>far from</u> suffices.

(27)   *This amount suffices <u>far from</u>.

(28)   The number of visitors to London <u>close to</u> tripled.

(29)   *The number of visitors to London tripled <u>close to</u>.

1.3   A possible explanation

A look at the elements that may block verb movement suggests two possible directions for an explanation: one is rather semantic, the other one is lexical/phonological. In the end neither will be completely satisfying in isolation and a mixed approach will be correct. It will be argued, however, that the latter one is much more important. Let us start with the first option. All the underlined elements in examples (1) – (29) somehow indeed operate on the verb, i.e. they focus it, measure or compare it. The explanation might thus be something along Haider's lines: the operator must c-command the operandum at s-structure. To put it differently, the quantifying element seems to have to c-command and therefore linearly precede the verb. A similar idea has been proposed in the 'fast'-generalization for German advocated by Rapp and von Stechow (1999). fast roughly means almost and as such it is close in meaning to other expressions blocking V2, e.g. soviel wie, so gut wie (see above). Rapp and von Stechow give the judgements in (31), (32) and claim:

(30)   'fast-generalization: fast cannot attach to a phrase with a phonetically
         empty head.
(31)   Gestern hätte sie mich fast getroffen.
         yesterday had she me almost met.
         'Yesterday she almost met me.'
(32)   *Gestern traf sie mich fast.'

'fast' is also an element which operates on the verbal meaning. However, a rule like (30) would be too strong. There are many other constructions in which the operator element need not c-command the verb. That means although the verbal action is quantified over, the full verb may raise over its operator and leave the

latter behind. Also <u>fast</u> – contrary to what Rapp and v. Stechow claim – can certainly be stranded by the verb (even (32) is grammatical for most speakers).

(33)   Genug! Der Reifen platzt fast.
       enough! the tire bursts almost
       'Enough! The tire is likely to burst soon.'

(34)   Sie erkannte mich fast.
       she recognized me <u>fast</u> (=almost)
       'She almost recognized me.'

The same is true for operator words like <u>nicht, nur, wieder, kaum, mehrmals</u> and so on. Each of them can be used to quantify over the verbal action and yet is completely compatible with a finite verb to their left and moreover, all of these expressions can stand right-peripherally, i.e. they do not require a continuation. Almost all of these elements are semantically very similar, or even synonymous to some of the blocking elements. For example strandable <u>nur</u> is synonymous to <u>nichts als,</u> which does not allow for stranding (see above (11)):

(35)   Der Wert verdoppelte sich fast _. / wieder_.   / kaum_. / nicht_. / nur_.
       the value doubled self   almost / again   / hardly / not   / only

Given this, the facts seem to call for another option, possibly the one with a phonological background. It appears that all the blocking elements – even across languages – and irrespective of the verb's final landing site (no or very little verb movement in English, I° in French and C° in German) end in or even only consist of an element that CAN be analyzed as either a C° element, or as some sort of functional preposition (P°, K°).

I would like to go here through all elements mentioned so far that are able to block verb movement. First there <u>wie</u> from <u>so gut wie</u>, and <u>soviel wie</u>. Apart from its status as an interrogative or relative manner pronoun, 'wie' can act as a C° (Zimmermann 1991, also Hahnemann 1999 see below). It does so with a special semantics in subordinate clauses selected by verbs of perception as in (36). In contrast to the canonical complementizer <u>dass</u> (=that), the use of <u>wie</u> implies that the matrix subject was a witness of the hair cutting process, a reading which is not obligatorily triggered by the use of <u>dass</u>, hence the English translation with the acc-ing construction.

(36)   Ich habe gesehen, wie du ihm die Haare geschnitten hast.
         I have seen, 'wie' you him the hair cut have
         'I saw you cutting his hair.'

<u>wie</u> also appears in comparative constructions of equality. Its corresponding element of inequality in Standard German (not necessarily in substandard dialects) is <u>als</u>. <u>als</u> is another element able to block verb movement, see above. I suggest that it is possible to analyze many of its uses as instances of C° as well (also proposed in Hahnemann 1999). Of course there are problems with such an analysis:

(37)   Er tut so, als wüsste er alles _.   vs.   *Er tut so, als _ er alles wüsste
         he does so, as knew he everything
(38)   Er tut so, als ob er alles wüsste.
(39)   Er tut so, als wenn er alles wüsste.
         'He pretends to knew everything.'

(37) shows that <u>als</u> cannot block V2 in subordinate sentences, which makes it unlikely to act as a regular C°. (38) and (39) show that only if CP-recursion is admitted, could <u>als</u> possibly be a C°. In a different analysis it must be something else because <u>ob</u> (=if) is the relevant C°. In one construction, however, <u>als</u> seems to act a relative complementizer to the modal adjunct <u>insofern</u> (or <u>insoweit</u>). Here most speakers can use it in complementary distribution with dass.

(40)   Das ist insofern günstig, als ich dann noch in Berlin sein werde.
        This is insofar opportune 'als' I then still in Berlin be will
        'This is opportune insofar as that at that time I will still be in Berlin.'

(41)   (?) Das ist insofern günstig, dass ich dann noch in Berlin sein werde.

In southern varieties, however, doubling emerges again and (42) seems to be structurally close to (38) and (39).

(42)   Das ist insofern günstig, als dass ich dann noch in München sein werde.

A similar point can be made with respect to <u>außer</u>[3], which is the next element making verb movement crash. <u>außer</u> behaves almost exactly as <u>als</u>, only that instead of <u>ob</u> it combines with <u>wenn</u> (as in (44) or <u>falls</u>, which both translate into English with <u>if</u>, or marginally again with <u>dass</u>:

(43)   Er tut es, außer du tötest den Hund.
        he does it except you kill the dog
        vs. *Er tut es, außer du den Hund tötest

(44)   Er tut es, außer wenn du den Hund tötest

---

[3] Note that there is a crucial difference, however. <u>außer</u> cannot trigger V2, itself 'representing' the 'Vorfeld', whereas <u>als</u> can ((40) vs. (43)).

(45) Er tut es, außer falls du den Hund tötest

      'He'll do it, unless you kill the dog.'

(46) ?Er tut es, außer dass du den Hund tötest

The solution to this 'problem' is not far fetched: all these constructions call for an analysis in terms of either CP-recursion or the assumption of a split CP layer: als can then be considered a head element in a split left periphery (Rizzi 1997). At any rate the claim made here – contrary to Haider – that it is the als, and not the mehr that triggers the failure of verb movement in (2). denn – the next element to be considered – is just a synonym of als in comparative constructions. It means the same, sounding just a bit old-fashioned. In other contexts denn behaves as another sort of C° that blocks V-to-C, but obligatorily embeds a V2 structure. In yet other constructions it seems to be an ellipsis of es sei denn (= unless), in which case it is synonymous with außer[4]:

(47) Er tut nichts (es sei) denn faulenzen.

      he does nothing (it be) 'denn' faulenzen

(48) Er tut nichts außer faulenzen.

      'He doesen't do anything but laze around.'

The final blocker in German is von, which cannot be analyzed as a C°. von, like English 'of', is a semantically empty preposition, thus a P°, or K°. So are 'to' and arguably 'from'. French que is unproblematic. It is the complementizer par

---

[4] In other positions different from the middle-field es sei cannot be left out. The parallelism with außer remains, however:

    (i) Er tut es, es sei denn du tötest den Hamster _.   vs.

        *Er tut es, es sei denn du den Hamster tötest

    (ii) Er tut es, es sei denn falls du den Hamster tötest

    (iii) Er tut es, es sei denn wenn du den Hamster tötest

excellence. Another regular French complementizer is <u>si</u> (if, whether). Here the data are less clear, the facts seem to be similar, however[5]. There is definitely a contrast.

(49)   (?)Il n'a pas si bossé *que ça.*
       He neg-has not si hard-worked que this
       Something like 'He didn't work THAT hard...'

(50)   *Il ne bosse pas si *que ça.*

(51)   ?Il n'a pas si bossé *qu'il soit fatigué.*
       He neg-has not si hard-worked que'he is-subj tired
       'He didn't work so hard that he should be tired.'

(52)   *Il ne bossait pas si *qu'il soit fatigué.*

<u>weder</u> is a more complicated case. Its categorization as a C°-element is not evident. However, it does not seem impossible to analyze it as such. In any case it shares some striking similarities with its French counterpart <u>ni</u>. <u>ni</u> is another element which – similar to <u>que</u> – associates with clitic negation. Periphrastic tense constructions are acceptable (53). Ordinary verb movement in simple tenses is impossible (54). Auxiliaryless constructions must resort to other strategies. Either like in German, where <u>weder</u> must be placed into Spec;CP or some related left peripheral position to precede the verb (55), '<u>ni'</u> can be fronted and replace the negation clitic <u>ne</u> and hence precede the finite full verb (56), or again pleonastic 'faire' (do-support) can save the construction (57). So it seems

---

[5]   The structure here is more complicated for a constituent beginning with que is necessary. The relevant que-phrase (an NP or DP in (46) or a CP in (48)), however, opens a new phonological constituent and should therefore be considered as *not being there*, i.e. *not counting*, leaving the element <u>si</u> stranded at some abstract phonological level (for the spirit of this reasoning see below).

that 'ni' (like weder) is very flexible in position and categorial status, what matters is its location in front of something.

(53)   Il n'a ni travaillé, ni dormi.

      he neg-has nor worked, nor slept.

(54)   *Il ne travaille ni, (il ne) dort ni.

(55)   Ni il travaille, ni il dort.   or

(56)   Il ne travaille pas, ni ne dort.

(57)   Il ne fait ni travailler, ni dormir.   (do support)

      'He neither works, nor sleeps.'

None of the other quantifying elements that do NOT trigger ungrammaticality in case they stand right peripherally (e.g. nicht, nur, fast an so on, see (35) above) can be analyzed as either a complementizer or an empty preposition.

It thus seems that the impossibility of verb raising in the constructions considered here is due to the nature of the (last) elements that occur between the relevant verb positions, immediately before the right sentence bracket. I have argued that the blocking elements are canonical occupants of either C° or P° (K°). A proposal that suggests itself is the following. According to Grimshaw's theory of extended projections (1991), these elements, i.e. C°s and P°s have something in common. They both close off the projection, i.e. they are the highest functional elements: C° of sentences, P° (K°) of noun phrases. Having this status implies being a functional category, i.e. the relevant elements act as function(al) words in their canonical use. Many authors have argued that function words are special and distinct from lexical words in that they are not phonological words whereas lexical words are. According to Hall (1999) such a view is widely accepted by phonologists. Being not a phonological word means that the relevant chunk of phonological material has to undergo phonological

processes in order to survive at PF. Such processes can be attachment to a phonological word (cliticization) or stressing/focussing for example[6].

The decisive factor in the case at hand seems to be a special variant of cliticization. Given the phonological deficiency of C°/P°-categories and the C°/P°-analysis from above, the expectation is the following. The relevant functional elements can survive only if they can attach to a host to be a part of a phonological word. The host necessarily finds itself within the same phonological constituent. The most reasonable assumption is to let the beginning of a sentence, i.e. a CP in syntactic terms, coincide with the beginning of a phonological constituent that is or contains a phonological word. This phonological entity will then automatically serve as the host material for the phonologically deficient C°-element. The phenomenon is not different with extended nominal projections. This explains the proclitic-like behavior of the elements under discussion. Orphanage thus leads to un-grammaticality. Putting phonological material after the clitic-like element saves the structure (14)–(17), (20) (even though with semantic consequences). In case quantifying over the verb is intended, only some form of do-support (or its language specific counterparts) can achieve the desired result.

The peculiarity of the whole phenomenon is that in contrast to canonical cliticization the clitics discussed here need NOT move in syntax, but they get hold of a host 'with less effort'. They need not look for an attachment site, they just require something to follow them. If this requirement is not met, the structure is ungrammatical.

---

[6] I am aware that some of the crucial items ('wie', 'von' etc.) can appear in isolation as independent elements (minimal words), however, not in the discussed examples where they are parts of larger expressions, which makes them very hybrid in character. Here they must integrate twice. Being a part of a larger unit makes it impossible for them to achieve the status of a phonologically (and morphologically) independent expression.

1.4 Riemsdijk's grafting approach as a possible alternative?

Another promising way of coping with the data is to apply Riemsdijk's very appealing theory of grafting (1998, 2000). Riemsdijk in his work is concerned with true or apparent mismatches between syntax proper and other linguistic components (morpho-phonology, pragmatics). His theory seems to work well for a number of grammatical phenomena like transparent free relatives. Look at his analysis of (58).

(58)
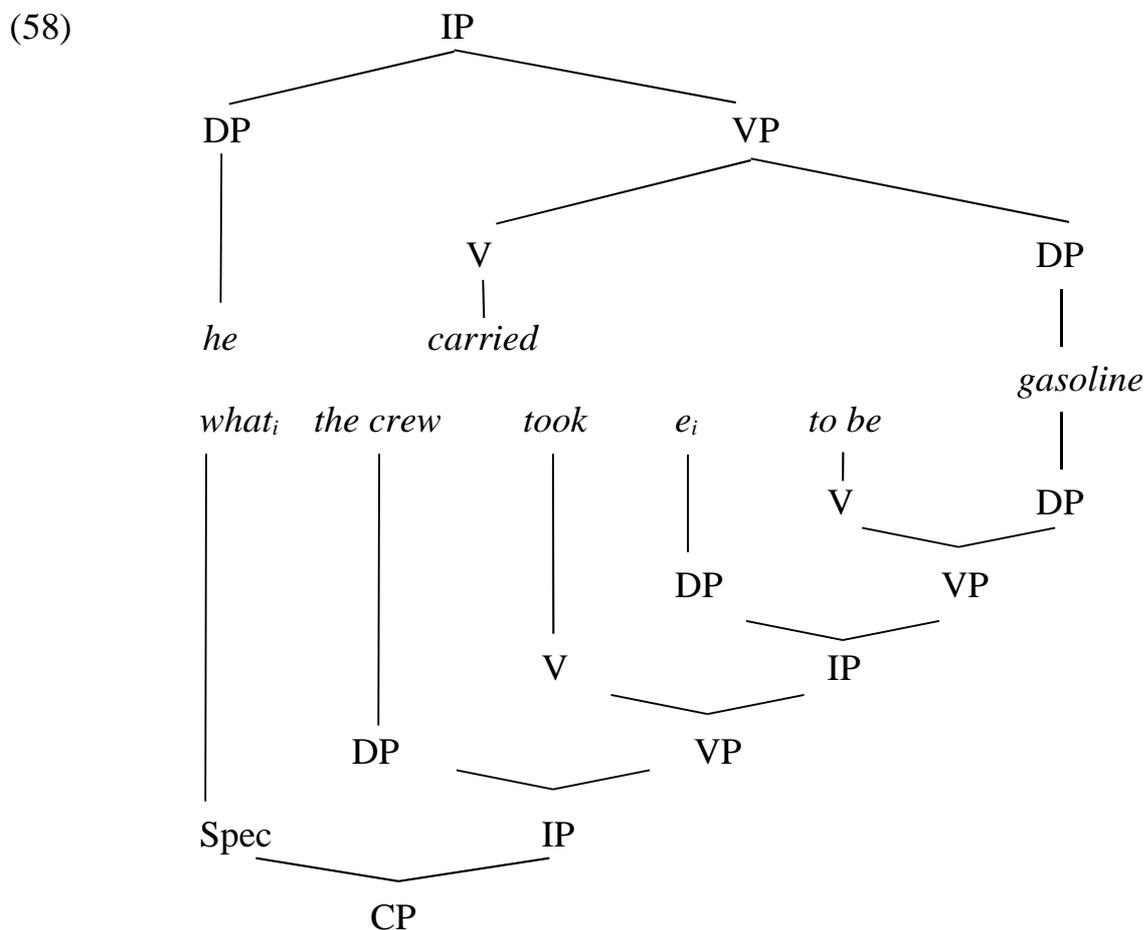
The same structure he proposes for cases like 'a far from simple matter' (59) first brought up in Kajita (1977 see above).

Riemsdijk's analysis:

(59)

```
                              DP
                 ┌─────────────┴──────────┐
               D°                         N'
               │                   ┌──────┴──────┐
               a                  AP             N°
                                   │              │
                                   A              │
                                   │              │
                                 simple         matter
                                   │
                    from           A
                     │             │
              far    P            AP
               │      └─────┬─────┘
               A           PP
                └─────┬─────┘
                     AP
```
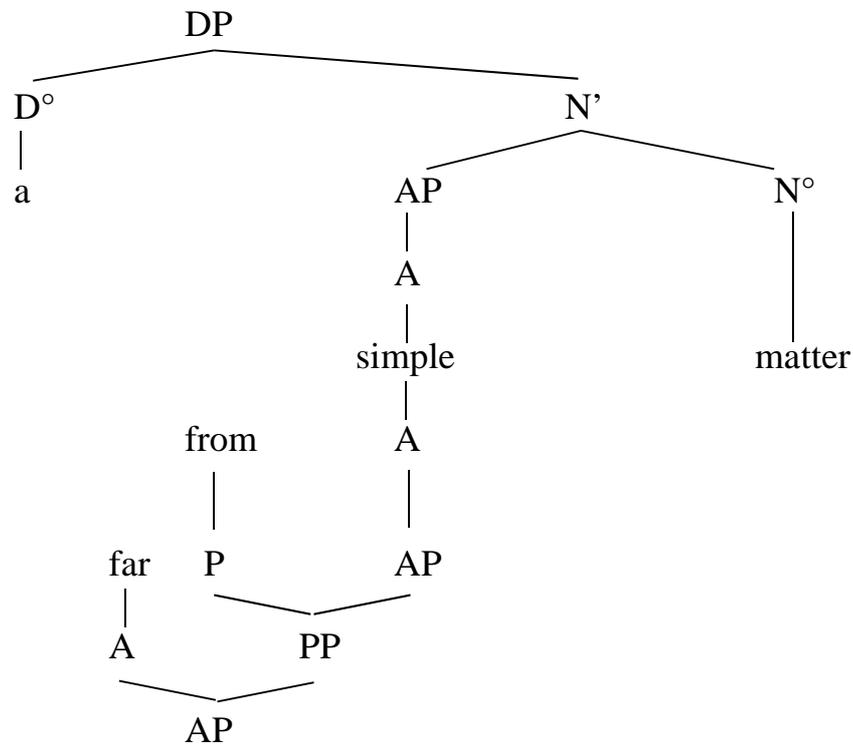
There seems to me to be an important difference between the two constructions, however. In (58) both trees can be generated independently and can potentially appear in different environments. The deepmost constituent 'gasoline' is a good argument of both verbs involved. This is not so with the adjective 'simple'. From a compositional, semantically driven point of view the bottom tree in (59) makes not much sense. From the labeling it does not fall out that 'simple' is the head of the construction. Instead it triggers a reading where 'simple' is the complement of the preposition 'from', which is not very uncontroversial[7]. On the contrary, the fact that almost any category can appear after P in such a construction makes it unlikely that P acts here as a regular preposition, all the more that an otherwise canonical DP complement seems to be the most marked option.

---

[7] The given argumentation still holds if 'far' is to be interpreted as the grafted construction's head with 'from' only having the status of an empty P° introducing the argument of the higher adjective.

(60)	She is far from out of the woods	PP
(61)	This amount far from suffices	VP (or even V°)
(62)	She is far from being an ideal partner/ a beauty	IP or possibly VP or CP
(63)	This is far from what I expected	CP
(64)	She solved the problem far from elegantly	AdvP
(65)	??/*She is far from a beauty	DP
(66)	??/*She is far from an ideal partner	DP
(67)	better: She is a far from ideal partner	AP like:
(68)	a far from simple matter	AP

This suggests that there is no selectional restriction along the path in the Riemsdijk's bottom tree, but that 'far from' is to be analyzed as a quantificational adjunct, semantically and syntactically parallel to 'hardly'. This is already done in Kajita (1977 p. 50), who proposes a reanalysis.

Note also that Riemsdijk discusses transparent free relatives like (69) and (70).

(69)   (?)John is what I'd call snoring.

(70)   *John what I'd call snores.

Here Riemsdijk argues convincingly for string identity as decisive criterion for grafting, categorial affiliation being of minor importance. (69) is good because 'snoring' can have (i) a verbal and (ii) a nominal reading (as predicative NP, see Riemsdijk (1998) and Wilder (1998). (70) is out because 'snores' is unambiguously verbal and as such not licit in a (sub-) structure '*I'd call (that) snores.' So far Riemsdijk's argumentation is reasonable and seems to speak in favor of his grafting approach to transparent free relatives. But it reveals itself as problematic for the cases discussed here, for verbal forms CAN appear adjacent

to elements like 'close to', 'far from' etc, see Kajita's example (71) and those given above.

(71)   It far from exhausts the relevant considerations.

(72)   it   exhausts   the relevant ...

                                                          |
                                 from   V

                        far | |
                                        | |
                         | P   (?)
                         | _____   _____/
                                                  \ /
                        A   PP
                         _____   _____/
                                   \ /
                              AP
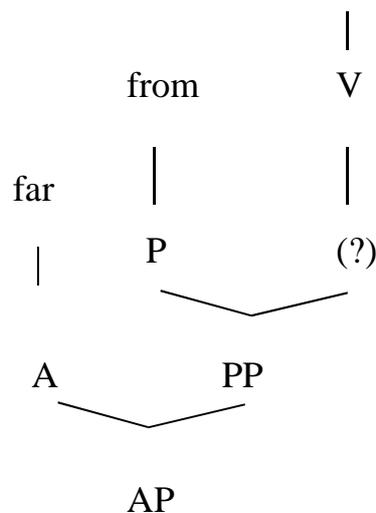
The same applies to all the German examples given at the beginning of the article, which are the main subject of this investigation. It seems odd to argue for a tree like (72), where a finite verb(al element) is the ultimate argument of a preposition. Such an approach is even less convincing under an analysis where the German finite verb in subordinate clauses is argued to occupy a head final I° node.

Instead the proposal advocated here is different, but in some sense similar. The idea is that there are indeed two distinct trees, but one is syntactic, the other one is phonological – a situation which is always given.

(73)

```
                              VP
              _____
             /                                   \
          AdvP                                    VP
         /\                                        |
        /  \                                       V°
       /____\                                     /\
      /      \                                   /  \
     /_____\                                 /____\

    [so-viel          (  wie   ]       ge –  stan –  den    )

    ?       ?          cl=?            ?      ?       ?
      \    /             \    /          \    /
       \  /               \  /            \  /
        \/                 \/              \/
        F                  F               F
        |                    \           /
        |                     \         /
        ?                      \       /
                                \     /
        ?                        \   /
         \                        \ /
          _____ ?
                      ?
```

[ ] = lexical, semantic word boundary

(...) = phonological (word) boundary

Nothing is strange with the analysis in (73). The upper part is classical syntax, the lower tree is classical phonology. Often there is a correspondence between syntactic and phonological trees, but it is well known that in some cases things pattern differently and the parallelism breaks down. It seems to me that in certain cases a morpho-syntactic: phonological mismatch is even systematic and motivated. Very often in derivational morphology (often considered to be a part of generative syntax) morpheme and syllable boundaries diverge systematically (in the case of suffixation).

(74)   find-2nd/sg   dove-infinitive   horse-pl   forest-genitive

[(fin) (d][est)]   [(gra)(b][en]]   [(Pfer)(d][e)]   [(Wal)(d][es)]
     ?   ?   ?   ?   ?   ?   ?   ?

          ?   ?   ?   ?

          [ ] = morpheme word boundary
          (...) = syllable boundary

The functional explanation is that these mismatches create larger units, they concatenate pieces like a zipper, with some sort of interlocking principle. Sometimes this dovetail device is the regular pattern, sometimes it comes as some sort of parasitic construction.

1.4   Intermediate Summary

The picture that emerges is that in the relevant constructions (see (1) – (17) below), verb movement is blocked because it creates a structure that violates phonological wellformedness constraints. The parser of phonological constituents cannot assign a legitimate structure to the syntactic object when a proclitic-like element ends up in a right peripheral position. As soon as some phonological material follows the phonological parser finds a host and the structure is saved.

2   Parsing problems with speech act adverbials in the pre-field

In this section I would like to present another restriction on V2. The observation can be laid down as follows. Some expressions, which semantically act as speech act adverbials or some sort of adverbials commenting on the mode of uttering (see the examples below), can occupy a position before the regular 'Vorfeld'.

(75)   Ehrlich gesagt, ich bin von dir total enttäuscht.
        Honestly said, I am of you totally disappointed.
        'To be honest/ honestly, I am completely disappointed with you.'
(76)   Ehrlich, ich bin von dir total enttäuscht.
(77)   Ganz offen gestanden, ich bin von dir total enttäuscht
(78)   Ganz offen, ich bin von dir total enttäuscht.

From these examples it is not obvious that the pre-comma string is indeed an integral part of the sentence and should therefore be integrated into the core tree of the respective clause. However, a few arguments can be given. In a theory of adverbials like Cinque (1999), these elements are taken to pattern like other adverbials and occupy a specific (base) position in the tree of sentence – universally a very high position in the tree of a sentence. However, there is also strong evidence from German clause structure itself that these elements can occupy a sentence internal position. They are fine both in the 'Vorfeld' directly preceding the finite verb or in the upper 'Mittelfeld' (middle field)[8].

(79)   Ehrlich gesagt bin ich von dir total enttäuscht.

---

[8] To argue for a clause internal reading one has to make sure that the intonation is different from a paranthetical structure, which is also possible with this word order. (80), however is also fine without intonational setting off, i.e. (80) can be read / pronounced without pauses integrating 'ehrlich gesagt' like any other adverbial forming an intonational phrase with following clause internal material.

(80)   Ich bin ehrlich gesagt von dir total enttäuscht.

This is a strong enough piece of evidence for the integration of this sort of adverbials into the sentence (CP) they modify. (81) – (86) bring some more frequently used examples.

(81)   Ganz nebenbei bemerkt, ich habe mir die Sache ganz anders vorgestellt.
       Wholly nearby remarked, I have the story completely differently pictured

(82)   Ganz nebenbei, ich habe mir die Sache ganz anders vorgestellt.

(83)   Ganz nebenbei bemerkt habe ich mir die Sache ganz anders vorgestellt.
       'By the way, I've had a completely different idea about it.'

(84)   Im Vertrauen gesagt, ich hab die Schnauze voll.
       In confidence said I have the moth full.

(85)   Im Vertrauen, ich hab die Schnauze voll.

(86)   Im Vertrauen gesagt habe ich die Schnauze voll.
       'Confidentially, I've got enough!'

(87) gives a list of pretty frequent adverbials that pattern exactly alike (some examples are inspired from Pittner (1999).

(87) offen gestanden   frankly
     offen gesagt
     (ganz) im Vertrauen gesagt   confidentially
     hinter vorgehaltener Hand gesagt   ??? confidentially
     zugegeben   admittedly
     ernst(haft)gesagt   seriously speaking
     ohne Scheiß gesagt   no kidding
     nebenbei bemerkt / gesagt   as a marginal remark, by the way

am Rande <u>bemerkt</u>	as a marginal remark, by the way
kurz <u>gesagt</u>	briefly, in brief
ohne zu <u>übertreiben</u>	without exaggerating
ohne Übertreibung <u>gesagt</u>	without exaggerating
überspitzt <u>formuliert</u>	with (a bit of) exaggeration
ohne Umschweife <u>gesagt</u>	to say it straight to the point
mit anderen Worten <u>gesagt / gesprochen</u>	to put it differently
anders <u>ausgedrückt / gesagt</u>	to put it differently
wenn ich ehrlich <u>bin / sein soll</u>	honestly
mit Verlaub gesagt / zu sagen	with all due respect

All these expressions come in a certain pattern, there is something – mostly some adverbial like phrase and then a verbal form, mostly a participle of a verbum dicendi, i.e. of a verb of communication, a verb of saying (always underlined in the examples.) However, as illustrated above, it is also possible to drop the performative verb (76), (78), (82), (85). One can obtain the same effect if one leaves out the verbal part (short form[9]). Semantically the sentences either with the verbal element or with just the short form are equal. Now comes the observation. (75) – (78) are V3 structures, in classical terms: the adverbials seem to be CP-adjoined. If the short form is placed within the regular pre-field giving rise to V2, the sentences are bad (88)–(90); whereas the long form is a good occupant of the 'Vorfeld', see (79), (83), (86).

(88)   *Ehrlich bin ich total enttäuscht von dir.

(89)   *Nebenbei habe ich mir die Sache anders vorgestellt.

(90)   *Im Vertrauen hab ich die Schnauze voll.

---

[9] Often the short form sounds better if the adverbial element 'ganz' (completely, wholly) precedes it. This is, however, just an improvement, it is not obligatory.

The question is: why is V2 blocked in these cases? It cannot be that the XPs in Spec,CP are no legitime occupants of the 'Vorfeld'. Under other conditions strings like 'im Vertrauen' or 'ehrlich' are fine in sentence initial position. The following data seem to suggest an explanation for the pattern. In some V2 cases dropping of the verbal part does not necessarily lead to ungrammaticality as in the above examples (88)–(90). Consider the following data in (91) to (96). The reading, however, is such that the speech act oriented interpretation is impossible.

(91)   Am Rande bemerkt, steht da auch eine Telefonzelle (für den Fall...)
       'By the way, there's a phone booth, ... just in case...'

(92)   Am Rande steht da auch eine Telefonzelle.
       'On the edge of it, there's also phone booth.'   (local reading)
       /##'By the way...'   (speech act reading)

(93)   Nebenbei bemerkt, ist so ein Job gar nicht zu schaffen.
       'By the way, such a job is undoable.'

(94)   Nebenbei ist so ein Job gar nicht zu schaffen.
       'In addition to what you are already doing, such a job is too much.'
       (manner reading)
       /##'By the way...'

(95)   Offen gestanden, war der Tresor ein Kinderspiel für den Dieb.
       'Frankly (speaking), the safe was a child's play for the thief.'

(96)   Offen war der Tresor ein Kinderspiel für den Dieb.
       'Being open, the safe was a child's play for the thief.'
       /##'Frankly,...'

Thus, if the adverbial form is not unambiguously specified for a speech act reading, this reading will not emerge. In case a reasonable manner reading (or

something similar) is possible, the sentence is grammatical, but only with that reading.

Second, related observation: German has 'bare' adverbials that are inherently speech act oriented, i.e. they cannot have a reading where they can possibly modify or affect the interpretation of the proposition in any way. These obligatorily speech act referring elements are: 'übrigens', 'erstens', 'zweitens' (by the way, first(ly), second(ly) – respectively – and so on[10] and marginally 'ungelogen' (literally: 'un-lied', meaning truely). Being unable to get a proposition internal reading these expressions can appear in the Vorfeld without triggering an unwanted interpretation or leading to ungrammaticality (97). Unsurprisingly they are also fine in the pre-prefield (V3) (98) or in the upper middle field (99). The semantics is always the same. They are bad, however, if they are put in a position close to the right sentence bracket (100).

(97)   Übrigens bin ich vorige Woche in München gewesen.

(98)   Übrigens, ich bin vorige Woche in München gewesen.

(99)   Ich bin übrigens vorige Woche in München gewesen.

(100) *Ich bin vorige Woche in München übrigens gewesen.

       'By the way, I was in Munich last week.'

And thirdly: more research is required concerning the following findings. Frey and Pittner (1998) mention in a footnote that scrambling of some (short, adjective-like) manner adverbials is bad (101) vs. (102).

---

[10] But interestingly not: 'letztens' ('finally, as the last'), which has a non speech act reading, meaning 'recently', 'some time ago'. This lexical feature seems to block the use of 'letztens' as speech act adverbial.

(101) Sie hat jedes Hemd sorgfältig gebügelt.

(102) *Sie hat sorgfältig jedes Hemd gebügelt

      'She ironed every shirt carefully.'

Whatever the reason for this behavior is, the topmost position in the upper middle field seems to be forbidden for adverbials with a manner interpretation. Nevertheless one can find manner adverbials there. The interesting thing is that HERE we observe the opposite to what happens in the 'Vorfeld'. The relevant sentences are grammatical, but the reading of the adverbials must be speech act oriented. A manner reading is impossible[11].

(103) Ich bin (ganz) ehrlich von dir total enttäuscht.

(104) Ich bin (ganz) nebenbei erst seit gestern von dieser Sache überzeugt.

(105) Ich habe (ganz) im Vertrauen erst gestern von dieser Sache erfahren.

Thus, a sentence corresponding in word order to (102) is not necessarily ungrammtical, consider (106) and (107).

(106) Sie hat (ganz) ehrlich alle Fragen beantwortet.

      Honestly, she answered all questions.

(107) Sie hat (ganz) ernsthaft alle Aufgaben gemacht.

      Really/Seriously, she did all tasks.

However, '(ganz) ehrlich' cannot get a reading where the answering is done in a honest way, i.e. without cheating; but – as the translation suggests, the only

---

[11] This finding should be considered crucial for the discussion whether V2 is a blind role that puts the finite verb between the first constituent and the rest, or whether there is a decisive semantic and categorial difference between the 'Vorfeld' and the highest position in the middle field.

interpretation possible is that the speaker of (106) wants to convey explicitly that the statement is taken to be true. The same for (107): 'ernsthaft' cannot mean that she worked on every task with seriousness.

The picture that emerges can be summarized in a table (108).

(108)

	'Vor-vor-Feld' (V3)	Canonical 'Vorfeld' (V2)	Upper middle-field position (TopP)
Complex, i.e. unambiguous speech act adverbial	ok	ok	ok
Bare adverbial, i.e. short form	ok	* / other reading	ok
Bare, but inherently speech act related adverbial (e.g. 'übrigens')	ok	ok	ok
Adverbial with a reading that is not speech act related, (mostly manner)	*	ok	*

This suggests very much that the speech act reading must be made explicit. This can be done in two ways: (i) either a long form is used, which by its very meaning must be interpreted as speech related (upper line), or (ii) by putting the adverbial in an unambiguous position (V3)(or some TopP-like position in the upper middle field) (first and third column). If the (bare) adverbial is put into the regular first position in a V2 clause, the parser wants to assign a sentence internal reading (mostly manner, but also local etc.). If such a reading is available, the sentence is grammatical, but there is no way to get a speech act reading. If no such reading can be triggered, the interpretative parser crashes. Thus, again we have a parsing related restriction for verb second. This time not from the PF path, but from the CI side.

References

Abney, Steve (1991) Syntactic affixation and performance structures. In: *Views on Phrase Structure*, K. Leffel and D. Bouchard (eds.), 215–227. Dordrecht/Boston/London: Kluwer.

Alexiadou, Artemis, E. Anagnoastopoulou, S. Barbiers and H.-M. Gärtner (2002) Dimensions of Movement – Introduction. *Linguistik Aktuell – Linguistics Today*, 3–16. Amsterdam/Philadelphia: John Benjamins.

Chomsky, Noam A. (2001) Derivation by Phase. In: *A Life in Language,* M. Kenstowicz and K. Hale, 1–52. Cambridge, Mass: MIT Press.

den Besten, Hans (1977/1983) On the interaction of root transformations and lexical deletive rules. Published in: *On the Formal Syntax of the Westgermania,* W. Abraham (ed.), 47–131. Amsterdam: John Benjamins.

Fanselow, Gisbert (2002) Münchausen-Style Head Movement and the Analysis of Verb Second. Ms. University of Potsdam.

Frey, Werner, and Karin Pittner (1998) Zur Positionierung der Adverbiale im Deutschen Mittelfeld. *Linguistische Berichte 176*, 489–534. Westdeutscher Verlag.

Grimshaw, Jane (1991) Extended projection. Ms., Brandeis University, Brandeis, Mass.

Hahnemann, Suzan (1999) *Vergleiche im Vergleich.* Tübingen: Niemeyer.

Haider, Hubert (1997) Typological Implications of a directionality constraint on projections. In: *Studies on Universal Grammar and Typological Variation*, Artemis Alexiadou and Tracy Hall (ed.), 17–33. Amsterdam/Philadelphia: John Benjamins.

Kajita, Masaru (1977) Towards a dynamic model of syntax. In: *Studies in English Linguistics,* Kajita, M., N. Toshio and M. Ukaji, 44–76. Asahi Press.

Koopman, Hilda, and A. Szablolcsi (2000) *Verbal Complexes*, Cambridge MA: MIT Press.

Meinunger, André (2001) Restrictions on Verb Raising, *Linguistic Inquiry* [Squib] 32.4: 732–740.

Müller, Gereon (to appear) Verb-Second as vP-First In: *Journal of Comparative Germanic Linguistics.*

Pittner, Karin (1999) *Adverbiale im Deutschen. Untersuchungen zu ihrer Stellung und Interpretation.* (Studien zur deutschen Grammatik 60). Stauffenberg. Tübingen.

Rapp, Irene, and Arnim v. Stechow (1999) Fast-Hand-out, presented at ZAS/Berlin.

Riemsdjik, Henk van (1998) Trees and Scions – Science and Trees: Fest-Web-Page for Noam Chomsky (32pp) URL:
   http://mitpress.mit.edu/chomskydisc/riemsdyk.html. – apparently not available anymore.

Riemsdjik, Henk van (2000) A far from simple matter: Syntactic reflexes of syntax-pragmatics misalignments. In: *Semantics, Pragmatics and Discourse. Perspectives and Connections. A Festschrift for Ferenc Kiefer*, R. M. Harnish and I. Kenesei (eds.). Amsterdam: John Benjamins.

Rizzi, Luigi (1997) The fine structure of the left periphery. In: *Elements of Grammar: handbook in generative syntax.* L. Haegemann (ed.) Dordrecht: Kluwer, 281–337.

Wilder, Chris (1998) Transparent free relatives *ZAS Papers in Linguistics* 10: 191–199. Berlin ZAS.

Zimmermann, Ilse (1991) Die subordinierende Konjunktion wie. In: *Fragesätze und Fragen*, Marga Reis and Inger Rosengren (eds.), 113–122. Tübingen: Niemeyer.

Dialectal variation in German 3-verb clusters.

Looking for the best analysis*

Ralf Vogel

University of Potsdam

German dialects vary in which of the possible orders of the verbs in a 3-verb cluster they allow. In a still ongoing empirical investigation that I am undertaking together with Tanja Schmid, University of Stuttgart (Schmid and Vogel (2004)) we already found that each of the six logically possible permutations of the 3-verb cluster in (1) can be found in German dialects.[1]

(1) *Maria glaubt, daß* ...
     Maria believes that ...

    a.   *Peter die Arie singen müssen wird*   '... she will hear Peter sing the aria'
         Peter the aria sing hear   will
    b.   (Peter die Arie müssen singen wird)
    c.   Peter die Arie wird müssen singen
    d.   Peter die Arie wird singen müssen
    e.   Peter die Arie singen wird müssen
    f.   Peter die Arie müssen wird singen

The type of cluster exemplified in (1) is the most flexible one, consisting of auxiliary, modal and predicative verb. A perception verb in place of modal

---

[1] German dialects vary a lot in their morpho-phonology. As I am only concerned with word order facts in this paper, I am abstracting away from these differences, and only give the examples, with a few exceptions, in their Standard German "translation".

yields has a by and large equivalent flexibility. Throughout the paper, I use the following abbreviations for the above patterns:

(2)   A =   main verb – modal – auxiliary
        B =   modal – main verb – auxiliary
        C =   auxiliary – modal – main verb
        D =   auxiliary – main verb – modal
        E =   main verb – auxiliary – modal
        F =   modal – auxiliary – main verb

In this paper, I want to compare three different ways of accounting for the observed typology:[2]

- An LF derivation with head movement (minimalist)

  This is a standard minimalist approach.

- An LF derivation without head movement (minimalist)

  This is a 'Kaynean' approach, deriving the effects of head movement by remnant movement.[3]

- A PF-oriented solution (OT-style)

  This approach is radically different from the other two in that it assumes that the LF-to-PF matching is subject to an optimality theoretic competition.

We will see that the hardest problem comes with some unexpected optional orders. The advantage of the OT-account over the minimalist ones might be that

---

[2] The notion 'LF' might be a bit misleading here. What I have in mind is not an object that is specifically designed to be an input to the semantics component of the grammar, but rather simply a syntactic constituent structure. As the two have become quite indistinguishable from a representational perspective in recent minimalist work, I use the term 'LF' throughout the paper for something that should better simply be called 'syntactic structure', or 'constituent structure', as in LFG.

[3] The most actual approach of this kind on verbal complexes is (Koopman and Szabolcsi (2000)). This approach is much more complex and sophisticated than the 'toy grammars' I want to discuss here. Nevertheless, as we are concerned with the conceptual implications of such approaches, what is said here about accounts without head movement *in general*, should also hold of the account of Koopman and Szabolcsi (2000).

it integrates the *triggers* for these orders in a more direct manner than purely syntactic accounts would be able to do.

1   The Typology

I will take a look at two rather extreme cases: the Swiss German spoken in St. Gallen and a Low German dialect called "Rheiderländer Platt". The dialects vary in two ways: a) They have different *default orders* – these are possible with varying stress assignments. b) They have the same *additional orders* – but these are possible only with specific stress assignments, and these differ between the two dialects.

(3)   **St. Gallen Swiss German (StG)**

      a.   Default order:
              C = Aux Mod V
      b.   Additional orders:
            (i)   stress on Mod: F = **Mod** Aux V
            (ii)   stress on V: E = **V** Aux Mod

(4)   **Rheiderländer Platt (RP)**

      a.   Default orders:
              A = V Mod Aux
              D = Aux V Mod
      b.   Additional orders:
            (i)   stress on Mod: E = V Aux **Mod**
            (ii)   stress on V: F = Mod Aux **V**

As we see, the additional orders have the first (StG) or the last (RP) verb in the verb cluster stressed, as indicated by boldfacing. Note that *the possibility of order F is a rather surprising result* that has rarely been noticed in the literature (if at all). This order is a syntactically very interesting case, as we will see below. Standard German, which will not be discussed in detail, but might be used as a 'control dialect', observes the following patterns:

(5)   **Standard German**

    a.   Default orders:
        A = V Mod Aux
        D = Aux V Mod

    b.   Additional order:
        (i)   stress on Mod or/and V: E = **V** Aux **Mod**

A straightforward way of describing the differences between Standard German and the two dialects might be that there is a requirement to place the stressed verb at an edge of the verb cluster. While in Standard German, this could be the left or the right edge, in RP, it must be the right edge, and in StG, it must be the left edge. Hence, in Standard German, the highly marked order F can be avoided, while in RP and StG it cannot.

1.1   Object Placement

The dialects also have slightly different possibilities of accusative object placement. In the default orders, the most natural position for the direct object is left adjacent to the verb:

(6)   Default orders:

    a.   St. Gallen (StG):
        Order C: Aux Mod <u>OB</u> V

    b.   Rheiderländer Platt (RP):
        Order A: <u>OB</u> V Mod Aux   **and!:**
        Order D: <u>OB</u> Aux V Mod

The exception to this generalisation is (6-b) with order D in RP, where the object occurs in front of the whole verb cluster, although it is no more adjacent to the main verb.

(7)   All possible Object orders:

    a.   StG:
        Order C: (<u>OB</u>) (…) Aux (<u>OB</u>) Mod (<u>OB</u>) V
        Order E: (<u>OB</u>) (…) V Aux Mod
        Order F: (<u>OB</u>) (…) Mod Aux V

    b.   RP:
           Order A: (OB) (. . .) V Mod Aux
           Order D: (OB) (. . .) Aux (OB)V Mod
           Order E: (OB) (. . .) V Aux Mod
           Order F: (OB) (. . .) Mod Aux V

2   Treatment of StG in terms of LF Movement – with and without Head Movement

For the comparison of the two minimalist accounts, I assume the following 'scenario':

- Cyclicity is obeyed. Merge/Move have to extend their target, and target the tree's top.

- Head movement does not count as violation of cyclicity (although it does not literally extend the target) for the head movement approach.

- Subjacency is obeyed, in particular, extraction out of islands (XPs in specifiers, adjoined XPs) is impossible.

I will use a **rightward branching**, **binary** structure, with a **vP** for transitive verbs, as assumed in most work based on Chomsky's recent writings. The branching direction is only a notational convention here. Syntactic trees only encode *dominance relations*. Linearisation follows from some version of Kayne's (1994) Linear Correspondence Axiom – though we might allow for multiple specifiers and/or multiple adjuncts.

The default order of Swiss German can be derived with nearly no movement:

(8)   Default order C = Aux Mod OB V:

```
              AuxP
             /    \
          Aux      ModP
                  /    \
               Mod      vP
                       /   \
                     OB     vP
                           /  \
                        tSU    vP
                              /  \
                             v    VP
                                 /   \
                                V     tOB
```
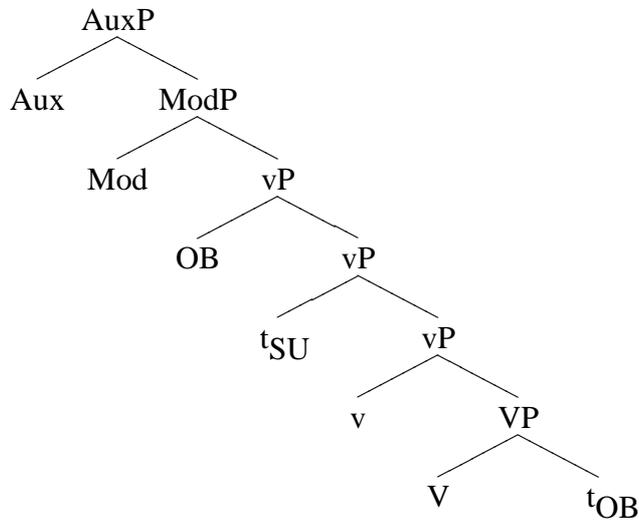
The subject has already moved to its position higher in the tree, it is only represented by its trace here. For the German dialects, strictly observing OV order, we must assume that the direct object has a strong case feature that it needs to check, and therefore obligatorily moves to its case position. Chomsky (1999) treats this on a par with *object shift*. Chomsky's (1999) way of representing this is the one indicated here, namely, adjunction to an outer specifier of vP. The additional orders cause bigger problems. Let us first take a look at order E:

(9)   Additional Order E = OB V Aux Mod – with head movement:

```
              AuxP
             /    \
           OB      AuxP
                  /    \
               Aux      ModP
              /  \      /   \
             V   Aux  Mod    vP
                            /   \
                         tOB     vP
                                /  \
                             tSU    vP
                                   /  \
                                  v    VP
                                      /   \
                                    tV     tOB
```
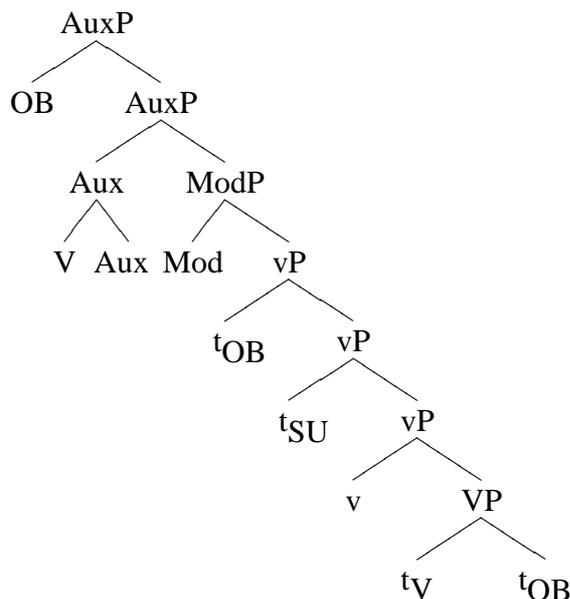
Two operations are necessary to derive this order:

1. head movement of V to Aux

   ⟶ If it is performed in a single step, then the 'head movement constraint' (Travis (1984)) is violated: V skips *v* and Mod on its way to Aux.

   ⟶ short successive head movement, however, requires *excorporation* of V, at least after adjunction to Mod. V must be able to "jump" from adjunction site to adjunction site.

2. Scrambling of the direct object: it may not occur on the right edge of the clause

The first operation discussed above is impossible under the standard assumptions for head movement. We would either have to allow for long head movement, or for excorporation. An additional problem is that, although the two operations have to apply both, they seem to be independent of each other. The structure in (10) shows that it is quite easy to derive order E without head movement:

(10)   Additional order E = OB V Aux Mod –
        without head movement:
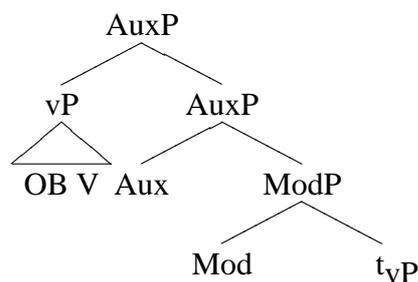
We only need vP-movement to AuxP to derive this order. So, in this case, XP-movement is clearly preferred over head movement. However, the derivation of order F is the harder problem. Let us again consider a head movement analysis first:

(11)   Additional Order F = OB Mod Aux V –
        With head movement:
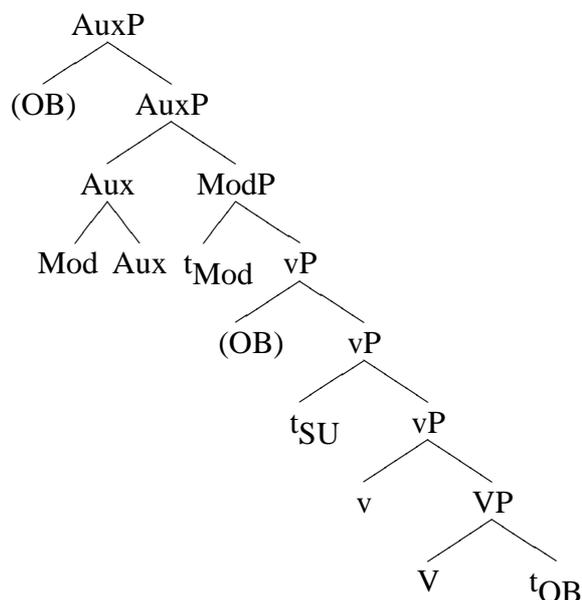
We only need a single step: head movement of Mod to Aux. Given that order F is extremely rare and highly 'marked', one could suspect that, if it was so simple to derive this order, why is it so special? On the other hand, markedness and economy, in the minimalist sense, do not go hand in hand anyway, so such worries are not helpful at all for our discussion.

Without head movement, it is much harder to derive this order. It is certainly impossible to do it within one step: VP is contained within ModP, but ModP must be raised and VP must be left behind. One option might be a split spell-out:

(12)   Without head movement: ModP→AuxP with split spell-out:
        [$_{AuxP}$ [$_{ModP}$ Mod ~~[$_{vP}$ OB v [$_{VP}$ V]]~~] Aux [$_{ModP}$ ~~Mod~~ [$_{vP}$ OB v [$_{VP}$ V]]]]

ModP is adjoined to AuxP, but the vP contained within ModP is spelled out in the position of the trace of ModP. This looks very ad hoc. What could be the trigger for such an operation? Note that spelling out vP within the moved ModP would yield the ungrammatical order B (= Mod V Aux).

A true movement solution requires additional projections: if vP must be left behind by ModP-movement, then it must extract before that movement, but only to a position lower than AuxP, hence, an additional (functional) projection is needed, call it FP:

(13)   (OB) [$_\text{AuxP}$ [$_\text{ModP}$ Mod t$_\text{vP}$] Aux [$_\text{FP}$ [$_\text{vP}$ (OB) v [$_\text{VP}$ V t$_\text{OB}$]] F t$_\text{ModP}$]]

Such an account faces a number of problems, among which are the following:

- The only 'evidence' for such an FP, as I see it, is that the analysis would not work otherwise. An analysis along the lines of (13) says that 3-verb clusters are in fact 4-verb clusters, only that one verb is invisible, and has no other function than providing a landing site.

- VP may not scramble to FP: this would wrongly bring OB to clause-final position.

- vP-movement to FP is obligatory, if ModP moves to AuxP, but what is the connection between the two steps?

The 'optimal' minimalist account of the St. Gallen German pattern seems thus to be a strategy that uses head movement to derive order F, and XP movement to derive order E. This is summarised in table (14).

(14)   Derivation of StG verb clusters with and without head movement (HM):

	Order C: Aux Mod V	Order E: OB V Aux Mod	Order F: (OB) Mod Aux ... ... (OB) V
with HM	default	V→Aux OB→AuxP	Mod→Aux
without HM	default	vP→AuxP	additional: FP vP→FP ModP→AuxP

Table (15) lists the set of operations needed for the obligatory and optional orders, assuming that the mixed strategy described above is the most promising one.

(15)   Possible operations in verb clusters in StG:

Obligatory operations:	Object Shift
	Subject movement
Optional operations:	Object Scrambling (→ModP,AuxP)
	vP→AuxP
	Mod→Aux

We thus have four different optional operations:

(16)   **Operation 1**: Object scrambling to ModP (⟶ Aux OB Mod V)
          Operation 2: Object scrambling to AuxP (⟶ OB Aux Mod V)
          Operation 3: vP→AuxP (⟶ OB V Aux Mod)
          Operation 4: Mod→Aux (⟶ Mod Aux OB V)

We now need to establish triggers for these optional operations and verify that they do not combine in the wrong way. Some combinations of the operations lead to orders that are not possible in StG. This is listed in (17):

(17)   a.   Op3+Op4 ⟶ Order A: V Mod Aux (impossible in StG)
          b.   Op1+Op3 ⟶ V Aux OB Mod (ill-formed because of OB position)
          c.   Op2+Op3 ⟶ V OB Aux Mod (ill-formed because of OB position)

All combinations involving operations 3 and 4 simultaneously yield the standard German default order A='V Mod Aux', which is impossible in StG. Operation 1 or 2 combined with operation 3 bring OB to the right of V. As we saw, there is a general ban on objects occuring to the right of V in all German dialects.

How can these combinations be avoided?

ad (17-a): Operation 3 and 4 could be triggered by a *strong* 'V-EPP' feature in Aux (this might be reminiscent of Koopman and Szabolcsi's (2000) 'VP+'). This feature can either be satisfied by head movement (of Mod, yielding order

F) or XP movement (of vP, yielding order E). Why does not the whole ModP move? Here, one could assume that head movement is the 'cheaper' version, because ModP movement would involve pied-piping of vP which does not check anything. Such an assumption makes sense in an approach that uses head movement. The bigger problem is that vP is lower than ModP, hence, why should it move at all, given that ModP is closer to Aux? We should observe a violation of the Minimal Link condition (MLC): ModP is the closer potential checker and should thus block VP from entering a checking relation with Aux. This problem might be unsolvable without a relaxation of the MLC.

To avoid this, one could assume that Aux has actually two different verbal EPP-features: a Mod-EPP feature and a V-EPP feature. If they are strong, then the movement of the respective elements is triggered. But now we have the same problem as before, because we have to prevent that both of these features are strong at the same time. This could, however, be stipulated in the functional lexicon of StG: It does not contain auxiliaries with the feature combination "$[s\text{Mod}][s\text{V}]$", but only those in (18):[4]

(18)   Aux-$[w\text{Mod-EPP}][w\text{V-EPP}]$
         Aux-$[s\text{Mod-EPP}][w\text{V-EPP}]$
         Aux-$[w\text{Mod-EPP}][s\text{V-EPP}]$

ad (17-b): Operation 1 might be triggered by a scrambling feature either in OB or in ModP. But now the triggers for operation 1 and 3 are again *independent of each other*. So we need an additional assumption, namely, that Aux only selects a ModP with a scrambling feature, if Aux itself has a weak V-EPP feature.

---

[4] 's' and 'w' stand for 'strong' and 'weak', respectively. This kind of solution has been pointed out to me by Jens Michaelis (p.c.), who I had the pleasure to discuss these problems with.

Problem (17-c) can be solved by a similar lexical stipulation: Aux only can have a scrambling feature, if it has a weak V-EPP feature. A better solution would, of course, be restricting the number of possible adjunctions to a single XP node to one, as usual in the Kaynean framework.[5]

A number of very specific lexical stipulations need to be postulated to make the correct predictions. These are not only about the feature strength of some element, but also about the feature strength of some other element that it selects. The content of these features is rather meaningless, EPP- or scrambling features are only there to yield correct orders. Nothing is said yet about the connection between these somehow derived orders and their information structural interpretations.

An alternative to these treatments would be attractive, if it was able to directly relate the additional orders to their information structural properties, and on the other hand still had enough flexibility to capture the typological variation. A second weakness of the minimalist accounts are the lexical stipulations that we had to make in order to rule out unwanted combinations of optional operations. It would be nice, if this could be derived in a less arbitrary, ad hoc fashion. The optimality theoretic treatment developed by Schmid and Vogel (2004) that I will present in the next subsection, tries to fulfil both of these requirements.

3   An OT-solution in terms of linearisation ('LF-to-PF-Mapping')

What follows is derived from the account developed in (Schmid and Vogel (2004)). We assume a uniform underlying LF for all cases we are exploring:

---

[5] It might be important to note that a treatment without head movement needs even more stipulations. To derive order F, we need an additional projection, FP in (13). The optional operations we then need in addition are vP-to-FP movement and ModP-to-AuxP movement. The number of optional operations is five in this account, one more than with head movement, and this also increases the number of combinations that need to be ruled out.

(19)   *Uniform structure of the verb clusters*

```
              AuxP
             /    \
          Aux      ModP
                   /   \
                Mod     VP
                        / \
                       V   OB
```
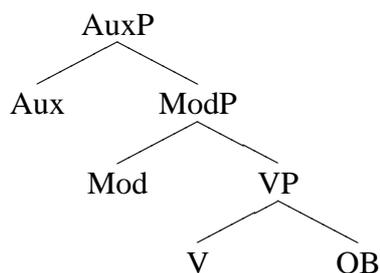
Note that this structure is simpler than the minimalist one we used before, in that it has no vP. OT encodes in constraints what is very often expressed in terms of structure in minimalism. So such a simplification is expected, but does not really say much about the conceptual complexity of the frameworks. Under Kayne's (1994) "Linear Correspondence Axiom", rephrased in (20), the structure in (19) would be mapped into PF with the linearisation 'Aux Mod V OB'.

(20)   **Kayne's Linear Correspondence Axiom (LCA) (rephrased)**
         If α asymmetrically c-commands β at LF, then the PF-correspondent of α precedes the PF-correspondent of β at PF.

The basic idea of our Optimality theoretic account is that constraints like the LCA indeed have their place in the grammar, but they are *violable*. The LCA is only one among a number of factors that determine linearisation. The constraints on linear correspondence that we use are in some respect different from the Kaynean version. For Kayne, the LCA is an inviolable constraint, and it is the only one that determines linearisation. Therefore, he has to take care that there are no LCA-ambiguous structures. This imposes some interesting restrictions on what a possible syntactic structure is. Problematic cases are those, where two elements c-command each other symmetrically, i.e., sisterhood relations, like those illustrated in (21):

(21)   string-ambiguous structures:

a.   XP   head-complement sisterhood

        X   YP

b.   X   head-head sisterhood

        X   Y
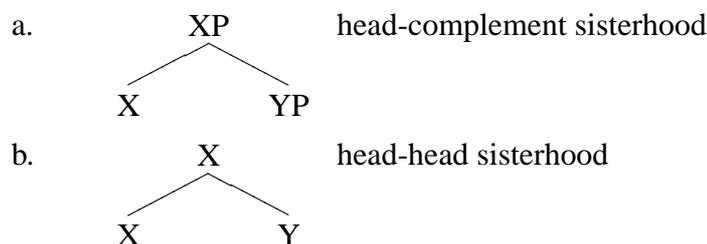
Kayne's (1994) solution for (21-a) is the decision that the LCA only talks about the relative order of heads (and the terminal nodes they dominate), not that of heads and maximal projections. It is, in fact, sufficient to do so, since maximal projections are built of heads. But it is somewhat counterintuitive that we cannot directly talk about the linear order of DPs with respect to each other. The main argument against such a way of formulating the LCA is that we would not get a total ordering of the terminals. In an OT setting, this might no longer be problematic. The string ambiguity of the head-complement sisterhood relation can be interpreted as the *source* of the 'head parameter': there is typological variation in the relative order of heads and complements (in particular: verb and direct object), precisely *because* this relation is string ambiguous, and hence needs to be fixed by a language particular convention. This convention might come into conflict with the LCA, and it thus becomes crucial which principle has the higher priority – we get an optimality theoretic setting.

To solve the problem in (21-b), string ambiguity of head adjunction, Kayne defines c-command in such a way that adjoined elements asymmetrically c-command the category they are adjoined to. But, intuitively speaking, adjuncts are still parts of their host categories, under standard assumption, and a category usually does not c-command something it is part of. So, while technically accurate, this is also somewhat counterintuitive. Recent work in the Kaynean framework tries to get rid of head movement at all. Koopman and Szabolcsi (2000), e.g., develop a theory of verb complex formation which is fully based

on remnant movement. Another way of getting rid of head movement, is, however, attributing it to the LF-PF interface, as first suggested for Germanic verb clusters by Haegeman and van Riemsdijk (1986), and, more recently, by Wurmbrand (2001). This is the kind of solution that we also prefer in (Schmid and Vogel (2004)). However, our approach is more radical in that it focuses on PF as the central representation in accounting for the phenomenon at issue.

The residue of the LCA that we make use of is restricted to relations between heads of the same *extended projection*, in the sense of Grimshaw (1991). The heads within an extended projection, e.g., C, INFL, V, or: P, D, N, usually asymmetrically c-command each other. Asymmetric c-command is to be translated into left-to-right ordering.[6]

(22)   **MAP–left-right(V^0) (MAPlr(V^0))**
         The heads of an extended projection of V are linearised in a left-to-right fashion, i.e.,
         if head A asymmetrically c-commands head B at LF, then the PF correspondent of A
         precedes the one of B at PF.

The violations of MAPlr(V^0) are counted pairwise, i.e. if Aux c-commands Mod, and both c-command V, asymmetrically!, then the following violations occur: [7]

---

[6] The definition in (22) only talks about extended projections of V, not about the heads of any extended projection. Hence, there might be another constraint talking about the extended projection of N. Whether these two can be collapsed under one general constraint, cannot be discussed within the limited range of this paper.

[7] Note again that the candidates that we are talking about here and below are PFs, i.e., linearisations of terminal elements of syntactic structures, and their prosodic phrasing.

(23)   Violations of MAPlr(V^0):

	MAPlr(V^0)
A: V Mod Aux	∗∗∗
B: Mod V Aux	∗∗
C: Aux Mod V	
D: Aux V Mod	∗
E: V Aux Mod	∗∗
F: Mod Aux V	∗

As already indicated above, we also re-establish the head parameter as a linearisation convention:

(24)   **MAP(complement before head) (MAPch)**
        If A and B are sister nodes at LF, and A is a head and B is a complement, then the correspondent of B precedes the one of A at PF.

(25)   **MAP(head before complement) (MAPhc)**
        If A and B are sister nodes at LF, and A is a head and B is a complement, then the correspondent of A precedes the one of B at PF.

The relative ranking of these two constraints instantiates the 'head parameter'. The violations for the six possible verb cluster linearisations of the tree in (19) are given in (26) (object and subject ignored):

(26)

	MAPlr(V^0)	MAPch	MAPhc
A: V Mod Aux	∗∗∗		∗∗
B: Mod V Aux	∗∗	∗	∗
C: Aux Mod V		∗∗	
D: Aux V Mod	∗	∗	∗
E: V Aux Mod	∗∗	∗	∗∗
F: Mod Aux V	∗	∗∗	∗

We observe a crucial *constraint conflict* here: VP complements cannot simultaneously fulfil MAPlr(V^0) and MAPch: as complements their head should be on the left of their governing head to fulfil MAPch, but as co-heads of an extended projection of V, they should be on its right to fulfil MAPlr(V^0). The relative ranking of these two constraints makes the difference between Swiss German (including StG) and Standard German (including RP) verb clusters:

(27) Rankings:

    a. StG, Swiss German:
        MAPlr(V^0) $\gg$ MAPch $\gg$ MAPhc $\rightarrow$ order C (= Aux Mod V)

    b. RP, Standard German:
        MAPch $\gg$ MAPlr(V^0) $\gg$ MAPhc $\rightarrow$ order A (= V Mod Aux)

That MAPch is ranked higher than MAPhc for Swiss German dialects, predicts that objects occur to the left of their governing verb. The default position of direct objects is indeed left adjacent to the verb, as the Zurich German example in (28-a) shows. The object may move higher to the left, but it may not occur to the right:

(28) a. *De Joggel hät welen   es gottlett ässe*
         The Joggel has want-INF the chop eat-INF

      b. *De Joggel hät es gottlett welen*-INF *ässe*-INF ((Lötscher, 1978, 4))
           The Joggel has the chop   want   eat

         c. **De Joggel hät welen*-INF *ässe*-INF *es gottlett*
              The Joggel has want   eat   the chop

Ranking MAPhc over MAPch would yield a language of the English type.

3.1 The trigger for additional orders: Focus

The next step is the implementation of triggers for the additional orders. StG prefers the left edge of the verb cluster for focused verbs:

(29) **St. Gallen Swiss German** – additional orders

    a. stress on Mod: F = **Mod** Aux V
    b. stress on V: E = **V** Aux Mod

RP prefers the right edge:

(30) **RP** – additional orders

    a. stress on Mod: E = V Aux **Mod**
    b. stress on V: F = Mod Aux **V**

We can capture this by assuming two symmetric constraints that directly express these tendencies:

(31) **FocusLeft(FocL)**
      Focused material occurs at the left edge of its phonological phrase.

(32) **FocusRight(FocR)**
      Focused material occurs at the right edge of its phonological phrase.

(33) FocL $\gg$ MAPlr(V^0) $\gg$ MAPch (StG)
      FocR $\gg$ MAPch $\gg$ MAPlr(V^0) (RP)

3.2 Competitions

I will now briefly show, how the orders that we find in StG are predicted with this system of constraints. We assume that focus information is part of the input, just as any semantic information is. This is a standard assumption in OT syntax. The six different candidates are also already optimised with respect to prosodic phrasing. This needs to be accounted for independently.[8] For the beginning, we leave out objects, and only look at narrow focus assignments to one of the three verbs. Let us start with narrow focus on the predicate verb, V. The table in (34) illustrates this competition:

(34)

	Narrow Focus on V	FocL	MAPlr(V^0)	MAPch
	A V Mod Aux		***!	
	B Mod V Aux	*!	**	*
	C Aux Mod V	*!		**
	D Aux V Mod	*!	*	*
☞	E V Aux Mod		**	*
	F Mod Aux V	*!	*	**

The highest constraint, FocusLeft, only leaves the candidates A and E within the competition. The next lower ranked constraint, MAPlr(V^0), favors E over A, and we have a winner, namely order E, which is now predicted to occur under narrow focus on V in StG. This fits to our findings.

---

[8] For an Optimality Theoretic approach on this issue see (Truckenbrodt (1999)).

(35)

Narrow Focus on Mod	FocL	MAPlr(V^0)	MAPch
A V Mod Aux	*!	***	
B Mod V Aux		**!	*
C Aux Mod V	*!		**
D Aux V Mod	*!	*	*
E V Aux Mod	*!	**	*
☞ F Mod Aux V		*	**

With narrow focus on Mod, FocusLeft again reduces the set of competitors to two, this time to the candidates B and F, the ones that have Mod on the left of the verb cluster. B performs worse than F in the next lower ranked MAPlr(V^0), and again we have a correctly predicted winner for narrow focus on Mod, namely, order F.

(36)

Narrow Focus on Aux	FocL	MAPlr(V^0)	MAPch
A V Mod Aux	*!	***	
B Mod V Aux	*!	**	*
☞ C Aux Mod V			**
D Aux V Mod		*!	*
E V Aux Mod	*!	**	*
F Mod Aux V	*!	*	**

With narrow focus on Aux, the syntactically least marked candidate C is undefeatable, because it has the auxiliary already in the right position at the left edge of the verb cluster.

The three occurring orders are thus derived. The orders E and F are the optimal orders for narrow focus on Mod and V, respectively, because they *preserve* the syntactic information in the mapping from LF to PF as much as possible, under the premise to obey FocusLeft.

In this system, direct objects are either placed left adjacent to the verb: obeying MAPch, yielding default order. Or, when they are focused, they occur at the left edge of the verb cluster. There is, thus, one order missing, namely, 'Aux OB Mod V'. Our answer to this problem would be along the following lines: Ob-

ject placement is governed by many factors in addition to focus. Definiteness, Givenness, animacy and others have influence on NP placement in general. A full picture of the word order problem would include all those factors, and then hopefully derive this order as an order with a specific and unique information structural implication.

A harder problem is the following one: With complex focus on [Mod V], FocusLeft favours the orders A and B, neither of which occurs in StG. Here, we cannot hope for an external solution. We need an additional constraint. What is special about these two orders, is that the the finite verb, the auxiliary, is in final position:

(37)   A =   main verb – modal – auxiliary
         B =   modal – main verb – auxiliary

While MAPlr(V^0) requires left to right order for verbs in general, this requirement might be even stronger for functional verbs or verbs carrying functional features like finiteness, agreement, a.o. Let us assume that this tendency is reflected in a more specific constraint, MAPlr(V^0_{func}:

(38)   **MAPlr(V^0_{func}):**
         If A is a functional verb (or a verb containing functional features) that asymmetrically c-commands at LF another verb B that belongs to the same extended projection, then the correspondent of A precedes that of B at PF.

MAPlr(V^0_{func}) is violable by winners in StG. E.g., the orders E and F have one violation of MAPlr(V^0_{func}). The following table shows all violations of MAPlr(V^0_{func}) for the six possible verb orders of a 3-verb cluster:[9]

---

[9] Note, that we here assume that Aux is a functional verb in the sense of MAPlr(V^0_{func}), but not Mod.

(39)   Violations of MAPlr(V^0_{func}):

	MAPlr(V^0)
A: V Mod Aux	**
B: Mod V Aux	**
C: Aux Mod V	
D: Aux V Mod	
E: V Aux Mod	*
F: Mod Aux V	*

Ranking this constraint high would be too restrictive. What is crucial, it seems, is double violation of this constraint. This is also expressible by constraint conjunction, a mode of *constraint composition* that has been established by Smolensky (1995). It is necessary to integrate effects of cumulativity into OT. Usually a lower ranked constraint A cannot supersede a higher constraint B, no matter how often A is violated. Constraint conjunction offers a means to implement this for cases where it is needed. Thus, we can formulate a new constraint, A&A, that is ranked higher than B. In our case, the conjoined constraint is sensitive to double violation of MAPlr(V^0_{func}).

(40)   **MAPlr(V^0_{func})^2:**
        No double violation of (V^0_{func}) by the same V^0_{func}.

It is usually necessary to specify conjoined constraints for particular domains. We do not want the constraint to count violations of V^0_{func} by different verb clusters, e.g., in main clause and subordinate clause. Rather, we are interested in those violations that are incurred by the same element. This is the reason for the restriction "by the same V^0_{func}" in the definition above. Only candidates A and B violate this constraint because of their double violation of MAPlr(V^0_{func}) (see table (39)). The constraint ranking for StG is now as in (41):

(41)   StG ranking (revised):
        MAPlr(V^0_{func})^2 ≫ FocL ≫ MAPlr(V^0) ≫ MAPch

This ranking has the effect of blocking candidates A and B in toto in StG.

The problematic competition with the complex focus [Mod V] can no longer have one of these two as winner. As the remaining candidates, C, D, E, F, all violate FocusLeft for this input (because Mod and V are not together at the left edge of the verb cluster), this constraint cannot be decisive either, and the next lower constraint makes the difference – MAPlr(V^0), which prefers the syntactic default order C.

For the existence of a constraint like MAPlr(V^0_{func}), there is independent evidence from Finnish, as described by Dowty (1996), following Karttunen (1989):

(42)   a. ***En** minä **ole** **aikonut ruveta pelaamaan** näissä tennistä* 'I did not intend to
              not I   have intend start play   these-in tennis
              start to play tennis in these (clothes)'
   b. **En** minä näissä **ole** tennistä **aikonut ruveta pelaamaan**
   c. **En** minä tennistä näissä **ole aikonut ruveta pelaamaan**
   d. **En** minä **ole** tennistä **aikonut** näissä **ruveta pelaamaan**

Karttunen (1989) claims that the NPs can permute freely in (42). The only restriction is that the relative order of the functional verbs ('*En*','*ole*') remains constant. Finnish is a language that strictly obeys the constraint MAPlr(V^0_{func}).

3.3 Summary

By taking into account external factors directly, LF-PF mapping yields the correct results without stipulating additional structure, features or their (in)compatibility. The price that has to be paid is the inclusion of syntax-external factors within the constraint set. They require their own motivations and explanations. We seem to be in a situation where we reinvent Chomsky's (1973) 'Move α' as a PF device: everything can be moved and displaced at PF. But this is not really a problem, because the necessary constraints and restrictions on this powerful device are already there in the form of optimality theoretic constraints. Furthermore, because some of these constraints are concerned with the LF-PF mapping,

it is ensured that LF information is preserved at PF as much as possible. In the competitions discussed above, it is always a syntactic constraint, MAPlr(V^0), that makes the final decision for the winning candidate.

4   RP

I will now more briefly discuss the RP dialect, and make the same comparisons as we did before – again starting with the minimalist treatments.

4.1   LF Movement – with and without Head Movement

The additional orders that RP has are the same ones as those in StG. So we do not need to make any additions here, but can rather take over the analyses we developed for StG.

Deriving the default orders

The default orders in RP are A = 'OB V Mod Aux' and D = 'OB Aux V Mod'. Order A can be derived by successive-cyclic head movement of V to v, v to Mod and Mod to Aux, followed by OB scrambling to AuxP, for order D the last step in the successive-cyclic head movement is simply skipped:

(43)   Order A = 'OB V Mod Aux' – with HM:

    a.   [$_{ModP}$ V-v-Mod [$_{vP}$ OB t$_v$ [$_{VP}$ t$_V$ t$_{OB}$]]]
    b.   (OB) [$_{AuxP}$ V-v-Mod-Aux …]

(44)   Order D = 'OB Aux V Mod' – with HM:

    a.   [$_{ModP}$ V-v-Mod [$_{vP}$ OB t$_v$ [$_{VP}$ t$_V$ t$_{OB}$]]]
    b.   (OB) [$_{AuxP}$ OB Aux [$_{ModP}$ V-v-Mod] …]

Order A can be derived by XP-movement in the same manner as before with head movement, i.e., successive-cyclically: vP adjoins to ModP, and ModP then adjoins to AuxP. Order D is derived by first adjoining OB to ModP, and then

adjoining vP to ModP, and then adjoining OB to AuxP, again skipping ModP-to-AuxP movement:

(45)   Order A = 'OB V Mod Aux' – without HM:

  a. $[_{ModP}$ $[_{vP}$ OB $[_{vP}$ v $[_{vP}$ V]]] $[_{ModP}$ Mod t_{VP}]]
  b. $[_{AuxP}$ $[_{ModP}$ $[_{vP}$ OB $[_{vP}$ v $[_{vP}$ V]]] $[_{ModP}$ Mod t_{VP}]] $[_{AuxP}$ Aux t_{ModP}]]

(46)   Order D = 'OB Aux V Mod' – without HM:

  a. $[_{ModP}$ OB $[_{ModP}$ Mod $[_{vP}$ t_{OB} v $[_{vP}$ V t_{OB}]]]]
  b. $[_{ModP}$ $[_{vP}$ . . . V . . .] $[_{ModP}$ OB $[_{ModP}$ Mod t_{vP}]]]
  c. $[_{AuxP}$ OB $[_{AuxP}$ Aux $[_{ModP}$ $[_{vP}$ . . . V . . .] $[_{ModP}$ t_{OB} $[_{ModP}$ Mod t_{vP}]]]]]

XP-movement takes fewer steps than head movement in both cases and should therefore be preferred. A general, but perhaps less serious problem is that it takes more steps to derive the unmarked orders (A, and D with OB outside the cluster) than it takes to derive the marked ones: Economy and markedness do not go hand in hand – but they need not necessarily do so.

Table (47) lists the options that we have for deriving RP verb clusters with and without head movement. The derivations that need fewer derivational steps are again underlined. The "optimal system" uses head movement only in the case of order F. This is very much parallel to StG.

(47)   Derivation of RP verb clusters with and without head movement (HM):

	Order A: V Mod Aux	Order D: Aux V Mod	Order E: V Aux Mod	Order F: Mod Aux V
with HM	V→v v→Mod Mod→Aux OB→AuxP	V→v v→Mod OB→AuxP	V→Aux OB→AuxP	<u>Mod→Aux</u>
without HM	<u>vP→ModP</u> <u>ModP→AuxP</u>	(OB→ModP) <u>vP→ModP</u> (OB→AuxP)	vP→AuxP	additional: FP vP→FP ModP→AuxP

The list of operations that we need for RP is given in (48):

(48)   Operations needed in RP:

  Operation 1: Object scrambling to ModP ($\longrightarrow$ Aux OB Mod V)

Operation 2: Object scrambling to AuxP ($\longrightarrow$ OB Aux Mod V)
Operation 3: vP→AuxP ($\longrightarrow$ OB V Aux Mod)
Operation 4: Mod→Aux ($\longrightarrow$ Mod Aux OB V)
Operation 5: vP→ModP ($\longrightarrow$ Aux OB V Mod)
Operation 6: ModP→AuxP ($\longrightarrow$ OB V Mod Aux, only after op.5!)

A number of problems have to be solved, first of all again the exclusion of unwanted combinations of optional operations. The discussion that follows is perhaps not exhaustive.

All six operations in (48) are optional, but one out of the operations 3 to 5 always has to apply. A straightforward solution could be that either ModP or AuxP has a strong V-EPP feature, or both of them do, but not neither.

Operation 6 only applies after operation 5. Why is that so? It might be the case that the V-EPP feature of Aux can only be checked by vP. Operation 6 would then be ModP-pied-piping to check a strong V-EPP-feature in AuxP. Operation 4 must then be triggered by a different feature, perhaps a strong Mod-EPP-feature, as already proposed for StG.

If operation 4 applies, then operation 2 has to apply. This follows from nothing. One would have to assume that Aux always (and only then) has a strong NP-scrambling feature, if it has a strong Mod-EPP feature.

The operations 1 and 3 may not apply both at once. This is the same problem as in StG. We assumed there that Aux only selects a ModP with a strong scrambling feature, if Aux itself has a weak V-EPP feature. This will also help here.

A combination of operation 1, followed by operation 5 and then operation 4 would yield the order 'Mod Aux V OB'. This is also ill-formed. In fact, to get the right object placement, we have to assume that at least one of Mod and Aux always has a strong scrambling feature. The operations 1 and 5 are both adjunctions to ModP. Their co-occurrence could also be blocked by a prohibition

against multiple adjunction.

The list of problems is a little bit longer for RP than for StG, but perhaps they can still be solved with the correctly chosen stipulations for the functional lexicon of RP. However, this is also the weakness of such an account. This whole methodology looks like constructing the theory after the facts, and it does not do anything more than deriving particular syntactic structures. It still remains to be clarified what the connection is between particular word orders and their information structural implications. Let us see, whether the more complicated facts of RP can be accounted for within the OT approach, as those of StG can.

4.2   The OT-account for RP

The first problem that has to be solved is how to derive order D as one of the two default orders. The difference between order A and order D is the position of the auxiliary. In order A, it is at the right edge of the cluster, while in order D it is at the left edge. Modal and predicative verb can be assumed to remain in their positions:

(49)   A =   main verb – modal – <u>auxiliary</u>
          D =   <u>auxiliary</u> – main verb – modal

We saw that in StG, there is a total ban on the orders A and B, which have the auxiliary at the right edge of the verb cluster. In standard German dialects, the orders A and D usually are both default orders for 3-verb clusters with auxiliaries. For StG, we assumed the constraint $\mathrm{MAPlr}(\mathrm{V}^0_{func})^2$ to account for the total absence of the orders A and B in that dialect. However, this constraint cannot be held responsible for the optionality of order D in standard German dialects, because this option crucially depends on the kind of verb that bears the finite morphology. In Upper Hessian, a standard German dialect spoken in

a region about 70 kilometers north of Frankfurt/Main, 3-verb clusters with a perfect auxiliary cannot occur with order A at all:

(50)   Upper Hessian

    a.   … *dass sie singen gemusst/müssen hat*
                that   she sing   must-PART/INF has

    d.   … *dass sie* hat *singen   müssen*
                that she has sing-INF must-INF

Interestingly, this correlates with the impossibility of the perfect auxiliary to bear stress:

(51)   a.   … *dass sie singen gemusst/müssen* HAT
                that   she sing   must-PART/INF has

    d.   … *dass sie* HAT *singen   müssen*
                that   she has   sing-INF must-INF

This dialect also has another property that differentiates it from standard German, namely, it has weak pronouns. Non-subject pronouns may not occur in clause-initial position, and they cannot be stressed either (focal stress is again indicated by uppercase):[10]

(52)   Upper Hessian

    a.   *en/se   hu   ich gesehe*
              him-/her-ACC have ich seen

    b.   *ich hu   EN/SE   gesehe*
              I   have HIM-/HER-ACC seen

In such situations, Upper Hessian native speakers use d-pronouns:

(53)   Upper Hessian

    a.   *den/däi   hu   ich gesehe*
              him-/her-ACC have ich seen

    b.   *ich hu   DEN/DÄI   gesehe*
              I   have HIM-/HER-ACC seen

It thus seems that the perfect auxiliary in this dialect shares two properties

---

[10] For further discussion of this and related problems in Hessian syntax, see (Gärtner and Steinbach (2001)). The datum in (52-a) with the feminine pronoun is a counterexample to their analysis that allows for weak pronouns in clause-initial position under homonymy with the subject weak pronoun.

with weak pronouns: it cannot be stressed and it cannot occur in prominent position. The generalisation on word order that we need can be expressed with the following constraint:

(54)   ***WeakFinal (*WkFin)**:
         Weak elements may not occur in final position.

This constraint might actually describe only one subcase of a more general constraint banning prominence marking on weak elements. Note that "weakness" must be a lexical property of the perfect auxiliary in Upper Hessian. The future auxiliary, for instance, does not have the same restriction:

(55)   a.   *dass sie singen müssen* wird
                that she sing   must   will
         d.   dass sie *wird* singen müssen

Order D can even be blocked with contrastive focus accent on '*wird*':

(56)   a.   dass sie singen müssen *WIRD*
         d.   *dass sie WIRD singen müssen

Order D is, on the other hand, totally blocked, if the finite verb is thematically 'heavier', like, e.g., a causative verb. Here, order A is required in all standard German dialects:

(57)   a.   *dass sie die Kinder   spielen gehen* liess
                that she the children play   go   let
         d.   *dass sie die Kinder *liess* spielen gehen

To account for the *optionality* of the RP and standard German default orders A and D, we have essentially two options: we either might assume that there are two co-existing constraint rankings, or, that temporal auxiliaries exist in two versions, a weak and a 'normal' one. We have empirical evidence for the latter approach in Upper Hessian, where perfect auxiliaries have to be specified as 'weak' in the lexicon. For other auxilaries in Upper Hessian, or even the perfect

auxiliaries in other standard German dialects, I will assume that they exist in two versions in the lexicon, a 'weak' and a 'normal' one. I will indicate this optionality with brackets around the violations of the constraint *WeakFinal, as shown in table (58).

(58)   Violations of *WeakFinal

	*WkFin
A: V Mod Aux	(*)
B: Mod V Aux	(*)
C: Aux Mod V	
D: Aux V Mod	
E: V Aux Mod	
F: Mod Aux V	

The constraint ranking that we need for RP is the one in (59):

(59)   RP constraint ranking:
FocusRight $\gg$ *WkFin $\gg$ MAPch $\gg$ MAPlr(V^0)

4.3   Competitions

We again start with leaving out object placement. The first competition that we are looking at is narrow focus on V:

(60)

	Narrow Focus on V	FocR	*WkFin	Mapch	MAPlr(V^0)
	A V Mod Aux	*!	(*)		***
	B Mod V Aux	*!	(*)	*	**
☞↯	C Aux Mod V			**	
	D Aux V Mod	*!		*	*
	E V Aux Mod	*!		*	**
☹	F Mod Aux V			**	↯*!

We see that we are predicting the wrong winner, order C, which never occurs in RP. The winner that we would like to get, is order F. This order performs as well as order C in FocusRight. In fact, the two candidates perform equally well till

they reach MAPlr(V^0). Here, order C is optimal. The problem, thus, seems to be that a very low ranked syntactic constraint becomes decisive. As this is the only constraint, where the two candidates differ, there is no way to make order F the winner by reranking. So we need an additional constraint.

Such a constraint can in fact be motivated. The difference between the two candidates is that order F is indeed the better order for the intended narrow focus on V. The reason for this is that there is a general tendency to *project* a focus as far as possible:

(61)   *Focus Projection* – General observation about focus interpretation (cf., e.g., Uhmann (1991)):
        If a focussed element A is adjacent to the element B that selects it directly, then the focus can be 'projected' to [A B].

The idea for the formulation of the constraint that we need is that, in the ideal case, the focus *is* projected:

(62)   *Ideal Focus* (definition):
        The ideal focus of a clause is the maximally projectable focus.

We now can formulate the following constraint:

(63)   **IdealFocus (IF)**:
        The intended focus interpretation given in the input matches the ideal focus of a candidate.

Order C is a perfect candidate for global focus projection, if V bears nuclear stress. Mod, which directly embeds V, is right adjacent to V, and Aux, which directly embeds Mod, is right adjacent to Mod. Hence, the ideal focus for order C with stressed V is focus on all three verbs together. The 'ideal foci' with stressed V for all six different verb orders are listed in (64):

(64)   Ideal focus with stress on the predicative verb:
        a.   *weil   Maria [SINGEN müssen wird*]$_{\text{F:V-Mod-Aux}}$
              because Maria sing   must   will
        b.   weil Maria [müssen SINGEN]$_{\text{F:V-Mod}}$ wird

    c.   weil Maria [wird müssen SINGEN]$_{F:V\text{-}Mod\text{-}Aux}$
    d.   weil Maria wird [SINGEN müssen]$_{F:V\text{-}Mod}$
    e.   weil Maria [SINGEN]$_{F:V}$ wird müssen
    f.   weil Maria müssen wird [SINGEN]$_{F:V}$

Order F has the focused V at the right edge, with a left adjacent Aux, which does not directly embed it. Thus, focus cannot project, and order F is 'ideal' for narrow focus on V. The same holds for order E, except that here V is isolated at the right edge of the verb cluster. IdealFocus is usually fulfilled by two candidates which are *mirror images* of each other. For each 'edge' of the verb cluster, there is one ideal order for a given focus. However, IF is not a constraint that simply sums up FocusLeft and FocusRight. One difference comes with a broad focus on all three verbs: while IF says that the orders A and C are best here, FocusLeft and FocusRight cannot be violated in such a competition, because all we are looking at here is the verb cluster, and as all three verbs are focused, there is no way to violate FocL or FocR within the verb cluster. IF is, thus, much more sensitive to the total order of the verbs. FocL and FocR only look at the edges of the verb cluster.[11]

We rank IdealFocus (IF) immediately above MAPlr(V^0). We then get the following competition for narrow focus on V:

(65)

Narrow Focus on V	FocR	*WkFin	MAPch	IF	MAPlr(V^0)
A V Mod Aux	∗!	(∗)		∗	∗∗∗
B Mod V Aux	∗!	(∗)	∗	∗	∗∗
C Aux Mod V			∗∗	∗!	
D Aux V Mod	∗!		∗	∗	∗
E V Aux Mod	∗!		∗		∗∗
☞ F Mod Aux V			∗∗		∗

---

[11] For this reason, we have the impression (in Schmid and Vogel (2004)) that FocL and FocR are not truly about focus itself, but rather about stress, i.e., they reflect phonological constraints. In particular, a good candidate for such a constraint is the compound stress rule. There is some evidence that Northern German dialects prefer the right edge of compounds as the default location for nuclear stress, while in standard and southern German dialects, including Swiss German dialects, it is the left edge. For further details, see (Schmid and Vogel (2004)).

Now order F is the correctly predicted winner. With narrow focus on Mod, we have a competition between the orders D and E, after the evaluation of Focus-Right. IdealFocus is again the decisive constraint, favouring order D:

(66)

	Narrow Focus on Mod	FocR	*WkFin	MAPch	IF	MAPlr(V^0)
	A V Mod Aux	*!	(*)		*	***
	B Mod V Aux	*!	(*)	*		**
	C Aux Mod V	*!		**	*	
☞	D Aux V Mod			*		*
	E V Aux Mod			*	*!	**
	F Mod Aux V	*!		**	*	*

Narrow focus on Aux favours the default order, A:[12]

(67)

	Narrow Focus on Aux	FocR	*WkFin	MAPch	IF	MAPlr(V^0)
☞	A V Mod Aux					***
	B Mod V Aux			*!		**
	C Aux Mod V	*!		**		
	D Aux V Mod	*!		*		*
	E V Aux Mod	*!		*		**
	F Mod Aux V	*!		**		*

Thus far, we have not derived order E. But we have not yet considered all possibilities. Let us have a look at a more complex focus, Aux+Mod:

(68)

	Narrow Focus on Aux+Mod	FocR	*WkFin	MAPch	IF	MAPlr(V^0)
☞	A V Mod Aux		(*)(!)	☺		***
	B Mod V Aux	*!	(*)	*	*	**
	C Aux Mod V	*!		**		
	D Aux V Mod	*!		*	*	*
☞	E V Aux Mod		☺	*(!)		**
	F Mod Aux V	*!		**		*

Here, we have two different winners for weak and 'normal' auxiliary, the orders E and A, respectively.

---

[12] Stress on Aux requires non-weak auxiliaries, so *WeakFinal is not violable here.

Objects are usually placed left adjacent to the verb: obeying MAPch, yielding default order. But when they are focused, they are wrongly predicted to occur at the right edge. We again need another constraint. The idea here is that MAPch must more urgently be obeyed, if the head-complement relation is thematic. ModP is a complement of Aux, but Aux assigns no thematic role to ModP. Much of the observed word order freedom with 3-verb clusters is due to this factor. Remember example (57), where the syntactically highest verb of a 3-verb cluster was a causative verb: in such verb clusters the order is fixed to the default order A in standard German, obeying MAPch, and this correlates with the fact that the highest verb, the causative verb, assigns a thematic role to the VP that it embeds. The constraint that reflects this is the one in (69):

(69)   **MAP(complement before headΘ) (MAPchΘ):**
        If A and B are sister nodes at LF, and A is a head and B is a thematically dependent complement, then the correspondent of B precedes the one of A at PF.

A usual optimality theoretic assumption would be that MAPchΘ universally outranks the simple MAPch – the same holds for the mirror image constraints MAPhcΘ and MAPhc. For RP, we need a ranking where MAPchΘ is ranked higher than FocusRight, while the simple constraint MAPch is ranked lower. We thus get the following ranking:

(70)   Final ranking for RP:
         MAPchΘ ≫ FocR ≫ *WkFin ≫ MAPch ≫ IF ≫ MAPlr(V^0)

As FocusRight cannot be obeyed by a focused object, the system falls back to the default order, yielding A and D order. Object placement can be influenced by a number of additional factors, a discussion of which is beyond the scope of this paper.

5   Summary

RP is a more complex case and this is mirrored in the more complex accounts. The OT-PF-mapping account might have the advantage that each of the three additional constraints that we introduced can be justified independently. For some of the stipulations necessary for the minimalist accounts, such independent justifications might be harder to find. The OT-account follows a fundamentally different strategy: it focuses on linearisation in a direct manner, and assumes that the underlying syntactic structure (LF) is only one among several factors constraining the linear order of verb clusters at PF: IF is a semantic constraint on PF, and FocL and FocR might best be viewed as phonological restrictions. *WkFin refers to morphological properties of lexical items.

If one wants to do without head movement, however, an account in terms of LF-PF-mapping might be a better replacement than remnant XP movement – if one accepts the line of reasoning that I followed in this paper, namely, that an evaluation has to be made in terms of *qualitative criteria*, i.e., that not only the *number* of additional assumptions is of interest, but first of all the degree to which they can be motivated independently. If one adopts an OT version of 'PF movement', the apparent unrestrictedness of such an operation is also no longer a problem.

References

Büring, Daniel (2001). 'Lets's Phrase it. Focus, Word Order, and Prosodic Phrasing in German Double Object Constructions.' *In* Müller and Sternefeld (2001), pp. 69–106.

Chomsky, Noam (1973). 'Conditions on Transformations.' *In* Anderson, Stephen and Paul Kiparsky, eds., *A Festschrift for Morris Halle*, pp. 232–286. New York: Academic Press.

— (1999). 'Minimalist Inquiries: the Framework.' Manuscript, MIT.

— (2001). 'Beyond Explanatory Adequacy.' Manuscript, MIT.

Dowty, David R. (1996). 'Toward a minimalist theory of syntactic structure.' *In* Bunt, Harry and Arthur van Horck, eds., *Discontinuous Constituency*, pp. 11–62. Berlin: Mouton de Gruyter.

Fitzpatrick-Cole, Jennifer (1999). 'The Alpine Intonation of Bern Swiss German.' *In Proceedings of ICPhS 1999, San Francisco*, pp. 941–944.

Gärtner, Hans-Martin and Markus Steinbach (2001). 'What do reduced pronominals reveal about the syntax of Dutch and German?' *Linguistics in Potsdam* **9**:7–62.

Grimshaw, Jane (1991). 'Extended Projection.' Manuscript, Brandeis Univesity.

Haegeman, Liliane and Henk van Riemsdijk (1986). 'Verb Projection Raising, Scope and the Typology of Verb Movement Rules.' *Linguistic Inquiry* **17**:417–466.

Jacobs, Joachim (1993). 'Integration.' *In* Reis, Marga, ed., *Wortstellung und Informationsstruktur*, pp. 63–116. Tübingen: Niemeyer.

Karttunen, Lauri (1989). 'Radical Lexicalism.' *In* Baltin, Mark and Anthony Kroch, eds., *Alternative Conceptions of Phrase Structure*, pp. 43–65. Chicago: The University of Chicago Press.

Kayne, Richard (1994). *The Antisymmetry of Syntax*. Cambridge, Massachusetts: MIT Press.

Kleinhenz, Ursula (1994). 'Focus and phrasing in German.' *Tech. rep. 52*, IBM Deutschland, Wiss. Zentrum, Inst. für Wissensbasierte Systeme, Heidelberg.

Koopman, Hilda and Anna Szabolcsi (2000). *Verbal Complexes*. Cambridge, Massachusetts: MIT Press.

Lötscher, Andreas (1978). 'Zur Verbstellung im Zürichdeutschen und in anderen Varianten des Deutschen.' *Zeitschrift für Dialektologie und Linguistik* **XLV**(1):1–29.

Maurer, Friedrich (1926). *Untersuchungen über die deutsche Verbstellung in ihrer geschichtlichen Entwicklung*. Heidelberg: Carl Winter's Universitätsbuchhandlung.

Meurers, Walt Detmar (1999). *Lexical Generalizations in the Syntax of German Non-Finite Constructions*. Ph.D. thesis, Seminar für Sprachwissenschaft, Universität Tübingen, Tübingen, Germany. Published 2000 as Volume 145 in Arbeitspapiere des SFB 340, ISSN 0947-6954/00.

Müller, Gereon and Wolfgang Sternefeld, eds. (2001). *Competition in Syntax*. Berlin: Mouton de Gruyter.

Schmid, Tanja and Ralf Vogel (2004). 'Dialectal Variation in German 3-Verb Clusters. A Surface-Oriented Optimality Theoretic Account.' *Journal of Comparative Germanic Linguistics* pp. 235–274.

Smolensky, Paul (1995). 'On the structure of the constraint component Con of UG.' Handout of talk at UCLA, 4/7/95. Rutgers Optimality Archive, http://roa.rutgers.edu, ROA-86.

Sternefeld, Wolfgang (1994). 'Subjects, Adjuncts, and SOV-Order in Asymmetric Syntax.' *Groninger Arbeiten zur germanistischen Linguistik* **37**:227–246.

Travis, Lisa (1984). *Parameters and Effects of Word Order Variation*. Ph.D. thesis, MIT, Cambridge, Massachusetts.

Truckenbrodt, Hubert (1995). 'Extraposition from NP and Prosodic Structure.' *In* Beckman, Jill N., ed., *Proceedings of the North Eastern Linguistic Society*, vol. 25, pp. 503–518.

— (1999). 'On the Relation between Syntactic Phrases and Phonological Phrases.' *Linguistic Inquiry* **30**:219–255.

Uhmann, Susanne (1991). *Fokusphonologie*. Tübingen: Niemeyer.

Wurmbrand, Susi (2001). *Infinitives: Restructuring and clause structure*. Berlin: Mouton de Gruyter.