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

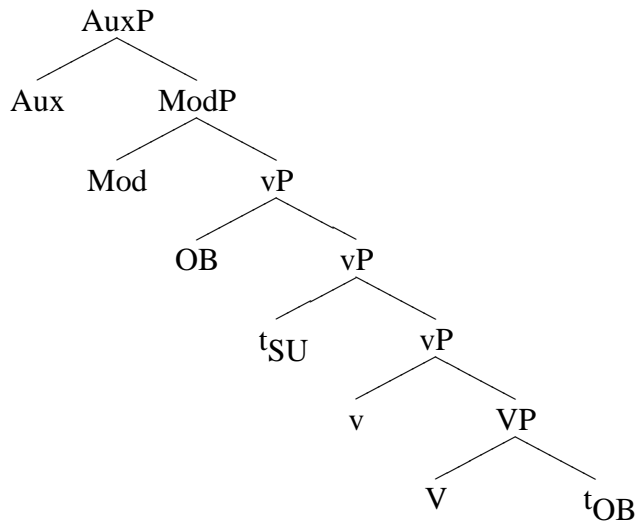## 2   Treatment of StG in terms of LF Movement – with and without Head Movement

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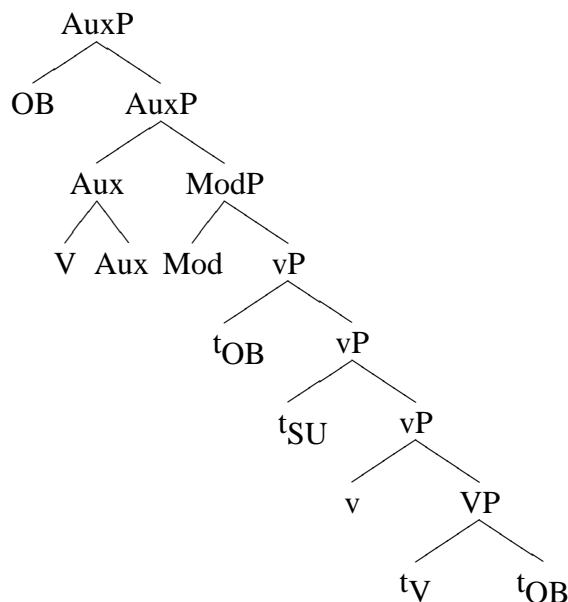
The default order of Swiss German can be derived with nearly no movement:

(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
```

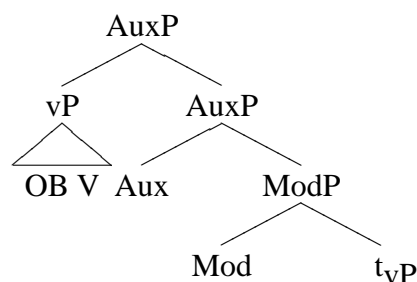Two operations are necessary to derive this order:

1. head movement of V to Aux

   $\longrightarrow$ 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.

   $\longrightarrow$ 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
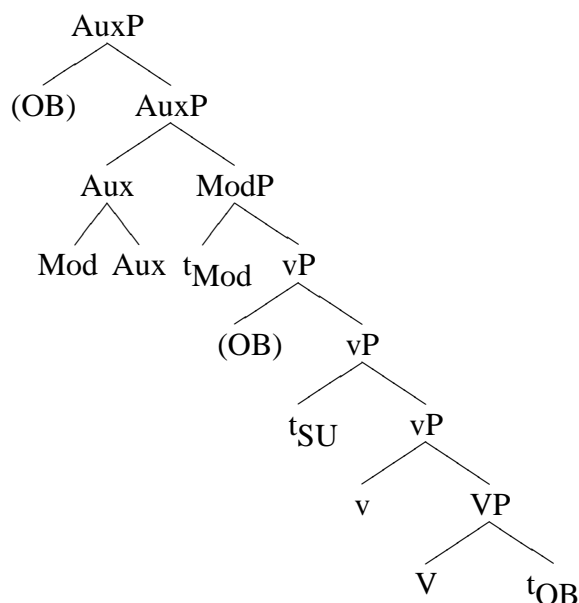
The first operation discussed above is impossible under the standard assumptions for head movement. We would either have to allow for long head movement, or for excorporation. An additional problem is that, although the two operations have to apply both, they seem to be independent of each other. The 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:

```
                AuxP
              /      \
            vP        AuxP
           / \       /    \
      OB V   Aux    ModP
                    /    \
                  Mod     t_vP
```

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:

```
                    AuxP
                   /    \
              (OB)      AuxP
                       /    \
                    Aux      ModP
                   /   \     /    \
               Mod  Aux  t_Mod    vP
                                 /    \
                             (OB)      vP
                                      /    \
                                  t_SU      vP
                                           /    \
                                          v      VP
                                                /    \
                                              V      t_OB
```

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

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 "[sMod][sV]", but only those in (18):[4]

(18)   Aux-[wMod-EPP][wV-EPP]
        Aux-[sMod-EPP][wV-EPP]
        Aux-[wMod-EPP][sV-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]
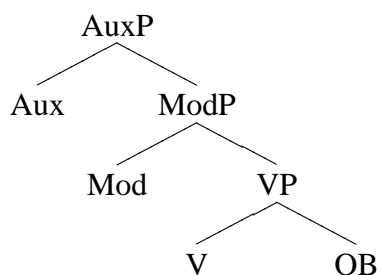
A number of very specific lexical stipulations need to be postulated to make the correct predictions. These are not only about the feature strength of some element, but also about the feature strength of some other element that it selects. The content of these features is rather meaningless, EPP- or 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
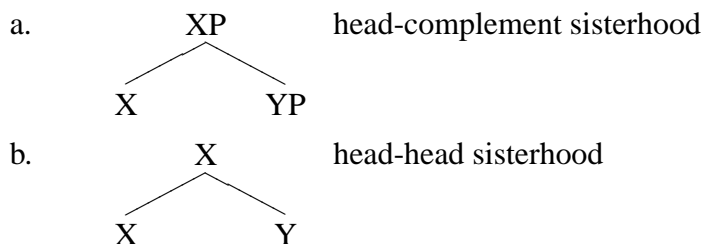
Kayne's (1994) solution for (21-a) is the decision that the LCA only talks about the relative order of heads (and the terminal nodes they dominate), not that of heads and maximal projections. It is, in fact, sufficient to do so, since maximal projections are built of heads. But it is somewhat counterintuitive 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})$^2 \gg$ FocL $\gg$ MAPlr(V^0) $\gg$ 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. $[_{ModP} [_{vP} \ldots$ V $\ldots]$ $[_{ModP}$ OB $[_{ModP}$ Mod t_{vP}]]]
    c. $[_{AuxP}$ OB $[_{AuxP}$ Aux $[_{ModP} [_{vP} \ldots$ V $\ldots]$ $[_{ModP} t_{OB} [_{ModP}$ 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 $\text{MAPlr}(V^0_{func})^2$ 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*]_{F:V-Mod-Aux}
               because Maria sing   must   will
         b.   weil Maria [müssen SINGEN]_{F:V-Mod} wird

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

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.