Dialectal Variation in German 3-Verb-Clusters

A surface-oriented OT account

Tanja Schmid (schmidta@ims.uni-stuttgart.de)
Institut für Linguistik:Germanistik, Universität Stuttgart

Ralf Vogel (rvogel@ling.uni-potsdam.de)
Institut für Linguistik, Universität Potsdam

Abstract. We present data from an empirical investigation on the dialectal variation in the syntax of German 3-verb clusters, consisting of a temporal auxiliary, a modal verb and a predicative verb. The ordering possibilities vary largely among the dialects. Some of the orders that we found only occur under particular stress assignments. We assume that these orders fulfil an information structural purpose, and that the reordering processes are only changes in the linear order of the elements which is represented exclusively at the surface syntactic level, PF (Phonetic Form). Our Optimality Theoretic account offers a multifactorial perspective on the phenomenon.

1. Introduction

German dialects vary in which permutations of the verb order in clause-final 3-verb clusters they allow. In an empirical investigation we have found that each of the six logically possible permutations of the 3-verb cluster in (1) appears in German dialects:[1]

(1) Maria glaubt, dass ...
Maria thinks that ...

a. Peter singen müssen wird
Peter sing must will
'... Peter will have to sing'
b. Peter müssen singen wird
c. Peter wird müssen singen
d. Peter wird singen müssen
e. Peter singen wird müssen
f. Peter müssen wird singen

We will use the abbreviations in (2) for the above patterns throughout the paper.

(2) A = verb3 – verb2 – verb1
B = verb2 – verb3 – verb1
C = verb1 – verb2 – verb3
D = verb1 – verb3 – verb2

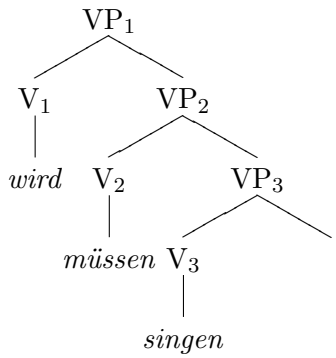
E = verb3 – verb1 – verb2
F = verb2 – verb1 – verb3

The verb clusters that we are exploring are exclusively three-verb clusters of the form abbreviated in (3).

(3) verb 1 = auxiliary
verb 2 = modal verb
verb 3 = predicative verb

We assume that the basic syntactic relations between the three verbs are *always* represented as in (4).

(4)



We follow recent assumptions in generative syntax insofar as we assume that the tree in (4) only encodes dominance, but not precedence relations (cf. (Kayne, 1994), (Chomsky, 1995)). Hence, that the heads are on the left in (4) is only a representational convention without any implication for the actual linear order of the elements. However, we also assume that linearisation is subject to an Optimality theoretic competition that takes the *translation of asymmetric c-command into precedence* as the default case, but also as a requirement that can be overridden by other demands. The details of this model will be discussed in section 3. Section 2 introduces the general outline of our proposal. Section 4 concentrates on our implementation of constraints about focus, and how their interaction with syntactic constraints derives the observed patterns. Section 5 discusses some problems and challenges for the proposed analysis.

In the remainder of this section, we will take a closer look at the data. The method that we chose in collecting the data is not very special: we made up a list of example sentences and asked the native speakers of the dialects to do two things: translate the clause literally into their home dialect, and then give a grammaticality judgement for that clause. What was special about it, was that we checked the

possibility of the different orders with particular *stress assignments*. The list of examples that we used consisted of all logically possible permutations of the verb order in a subordinate[2] clause with a 3-verb cluster (six possible orders), varying the place of the main stress (on subject, object, each of the three verbs). This gave us altogether 30 example sentences to test.

The variation with respect to verb order that we found includes two dimensions: variation in the default orders across dialect families and variation with respect to the possibility of additional optional orders within these dialect families. The following subsections present the details.

1.1. Macro-variation – variation across dialect 'families'

According to our findings, German dialects can be grouped into two different families that differ in which of the possible orders they take as 'default order'. Standard German dialects thereby differ from Swiss German dialects:

(5) Default orders
   a. Standard German (dialects):
      [A =] verb3 – verb2 – verb1 and
      [D =] verb1 – verb3 – verb2
   b. Swiss German (dialects):
      [C =] verb1 – verb2 – verb3

Swiss German dialects in this respect pattern together with West Germanic languages like Dutch, Afrikaans and West Flemish. The orders A (standard German) and C (Swiss German) are mirror images of each other that presumably reflect opposite settings of some parameter. The discussion of order D, which seems to be as unmarked as order A in standard German, is postponed till section 5.3. The finding illustrated in (5) confirms standard assumptions about the differences between standard and Swiss German.

1.2. Micro-variation – variation within dialect 'families'

Besides the default orders, dialects within these 'families' vary in which additional orders they allow for under certain circumstances. An interesting contrast that we observed, and which we want to discuss in more detail here, is the following one: The Swiss German dialect of St. Gallen (StG), and the Low German dialect 'Rheiderländer Platt' (RP, located in East Frisia) have the same additional patterns, namely the orders

312 (E) and 213 (F), but differ in which of the verbs receives main stress:

(6) St. Gallen (Swiss German dialect)
  a. stress on V: E = **V** Aux Mod
  b. stress on Mod: F = **Mod** Aux V

(7) 'Rheiderländer Platt' (Low German, standard German family)
  a. stress on Mod: E = V Aux **Mod**
  b. stress on V: F = Mod Aux **V**

The additional orders have the first (StG) or the last (RP) verb in the verb cluster stressed, as indicated by boldfacing. The possibility of order F is a rather surprising result in itself, as it is often said to be impossible in the verb cluster formation of Germanic languages (cf. (IJbema, 1997), (Wurmbrand, 2001b)). The discussion of this typology has to target three main issues: i) identify the 'parameter' that determines the default orders and is responsible for the division into two dialect families; ii) identify the factors that license the additional orders; iii) integrate 'extra-syntactic' factors like, e.g., stress assignment. In trying to achieve these three goals we developed a model within Optimality Theory (OT) that is introduced in the next section.

2. The model

The first important assumption that we make is that the reordering operations we are observing are PF operations. This is not a particularly new idea in the domain of verb clusters. The first such account that we know of has been developed by Haegeman and van Riemsdijk (1986). They propose a mechanism called 'PF inversion', the application of which is subject to particular syntactic conditions, and language specific parametrisation. Haegeman and van Riemsdijk assume that a Zurich German example like (8) is derived from an underlying standard German structure as exemplified by (9):

(8) Züri Tüütsch (Surface Structure):
das er en arie hät wele singe
that he an aria has want sing

(9) Züri Tüütsch (underlying structure, like standard German):
das er [[en arie singe] wele] hät
(cf. (Haegeman and van Riemsdijk, 1986), 428.)

Such a derivation has to proceed in two steps.[3] **Step 1** is a *Reanalysis* of the verb cluster: two adjacent verbal heads are syntactically reanalysed as being dominated by the same V^0 head:

(10) *Reanalysis* from a. to b.:
a. [$_{VP1}$ [$_{VP2}$ [$_{VP3}$ en arie singe] wele] hät]
b. [$_{VP1}$ [$_{VP2}$ en arie [$_{V2}$ [$_{V\alpha}$ singe] [$_{V\beta}$ wele]]] hät]

This configuration now makes 'PF inversion' possible. Vα and Vβ change their order. The result is, however, ungrammatical (order B = Mod-2 V-3 Aux-1):

(11) 'PF Inversion' of modal and predicative verb:
[$_{VP1}$ [$_{VP2}$ en arie [$_{V2}$ wele singe]] hät]

Therefore, a **step 2** is necessary, which repeats the processes in step 1. This now yields the Zurich German default order:

(12) Reanalysis from (11) to a., followed by inversion to yield b.:
a. [$_{VP1}$ en arie [$_{V1}$ [[$_{V\alpha}$ wele singe] [$_{V\beta}$ hät]]
b. [$_{VP1}$ en arie [$_{V1}$ hät wele singe]]

Following the basic intuition behind such an approach, we want to elaborate in a more extensive way on the idea that the verb order in verb clusters is a matter of linearisation and not so much of standard syntactic movement. More recent work that goes into the same direction has been presented by Wurmbrand (2000).

These accounts share with other purely (abstract) syntactic approaches like, for instance, those of Zwart (1996) and Koopman and Szabolcsi (2000) (see (Wurmbrand, 2001b) for an overview) that verb orders are mostly derived in a purely mechanical sense. That 'step 1' in a Haegeman and van Riemsdijk (1986) style analysis must be followed by further steps, has to be stipulated, for example. Within Koopman and Szabolcsi's (2000) theory, where all orders have to be derived by remnant VP movement, the number of stipulations needed to get the exact patterns for a single dialect becomes quite large, as demonstrated in ((Vogel, 2002)).

What is needed, in general, in explaining the typology is a *systematic* account not just of *how* orders are derived, but *why* dialects choose *which* subsets of the possible orders under *which* circumstances, thereby using a minimum of stipulative assumptions. Establishing a connection between *how* orders are derived and *why* they are derived is the major concern of our analysis.

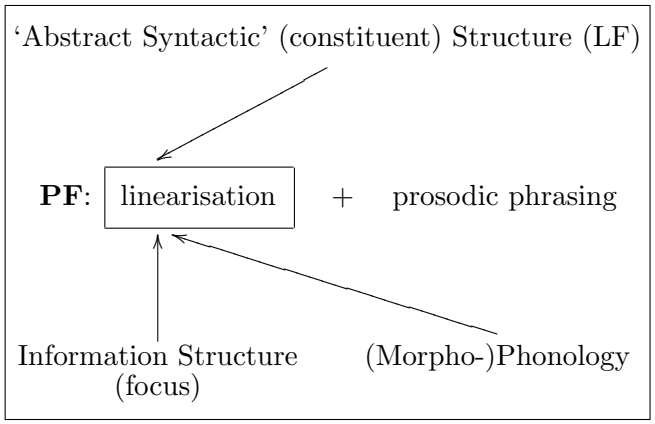
A second important idea that we make use of is the more traditional point of view that the syntax of verb clusters has multiple causes and is the result of the interaction of several independent factors. Predecessors of a multifactorial analysis are, for example, Lötscher (1978) and Maurer (1926):

> "'Dabei muß die zusätzliche Komplikation berücksichtigt werden, daß für die Erklärung der Wortstellung wohl kaum ein einziger Regeltyp vorausgesetzt werden kann. Vielmehr lassen sich mindestens drei interagierende, aber primär voneinander unabhängige Arten von Regeln ansetzen: Erstens **grammatisch bedingte Regeln** [...] die mehr oder minder willkürlich eine Abfolge bestimmen [...]; **performanzbedingte Regeln** [...]; endlich **funktional bedingte Regeln** [...] deren Zweck die Ermöglichung von bestimmten funktionalen Satzverhältnissen im Sinne der Thema-Rhema Unterscheidung ist."'[4] ((Lötscher, 1978, 11); boldfacing by us, TS/RV)

> "'[...]So liegt der Schluß nahe, daß die Wahl des Wortstellungstypus mit dem Tonfall der Rede, mit dem Akzent der betreffenden Mundart zusammenhängt. Vor allem fällt unser Augenmerk auch wieder auf den Rhythmus, der in Wortstellungsfragen eine ganz gewaltige Rolle spielt.[...]"'[5] ((Maurer, 1926, 72))

Abstract syntactic relations and properties (i.e., constituency, (asymmetric) c-command, selection, features etc.) constitute one class of the factors that are involved. Other important factors are (morpho-)phonological and information structural properties. These factors conflict, whenever they impose different requirements on the linear ordering of the verbs in a verb cluster.

(13) *The multifactorial model:*



For our OT grammar model, we take an abstract syntactic specification as syntactic part of the *input*. In what follows we call this structure LF

(Logical Form), using minimalist terminology. This might be somewhat misleading, insofar as LF is usually also assumed to be the input to the semantics component of the grammar, representing, for instance, covert movement. We are neutral about this, all we really need is a specification of the essential *abstract syntactic* relations, as listed in (14). The input also contains semantic information, in particular, what is important here, an information structural specification.

The *candidates* are PFs, i.e., linearised (inflected) words, prosodically and metrically structured. These are freely generated by the generation function *GEN*. This model certainly is only the *fragment* of a fully elaborated OT grammar. It contains only those aspects that are relevant for our discussion. It is a standard assumption among most OT practitioners that the structure of LFs themselves is also subject to optimisation.[6]

The *constraints* come from the areas discussed above. The architecture of this model is summarised in (14).

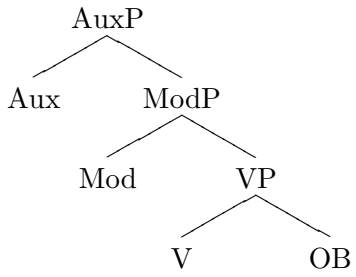
(14) The OT grammar model used here:
Input: ⟨*LF* (*constituent structure, abstract features*), *semantic representation (including focus)*⟩
Candidates: PFs, i.e., linearisation + phonological phrasing + stress assignment + morphology
Constraints: any constraints on PF formation, correspondence for LF-PF and semantics-PF, phonological and morphological restrictions.

The next step is the introduction of the constraints that we use. Section 3 discusses constraints on LF-PF correspondence, section 4 introduces the constraints imposed by information structure that we are using.

3. Syntactic Constraints on Linearisation

As already discussed above, we assume that the dominance and c-command relations encoded in an LF of a 3-verb cluster are always the same:[7]

(15) *Uniform abstract syntactic structure (LF) of the verb clusters*

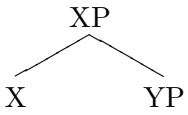


The linear order of auxiliary, modal and predicative verb is subject to an OT evaluation. Any conceivable order is a candidate. Under Kayne's (1994) "Linear Correspondence Axiom", rephrased in (16), the structure in (15) would be mapped into PF with the linearisation 'Aux Mod V OB'.

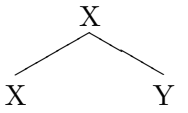
(16) **Kayne's Linear Correspondence Axiom (LCA) (rephrased)**
If α asymmetrically c-commands β at LF, then the PF-correspondent of α precedes the PF-correspondent of β at PF.

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 (17):

(17) string-ambiguous structures:

a. 

head-complement sisterhood

b. 

head-head sisterhood

Kayne's (1994) solution for (17-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, 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 (17-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. 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 (2000). However, our approach is even more radical in that it focuses on PF as the *central* level of representation in accounting for the phenomenon at issue.

The first constraint that we want to postulate is the one in (18). It does not talk about elements of LF in general, but only about elements of the same type, i.e., verbal heads that belong to the same extended projection (in the sense of (Grimshaw, 1991)):[8]

(18) **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.

At first sight, it might appear quite ad hoc to restrict the LCA to V-heads of the same extended projection, in the way this is done here. But we think that there is a deeper moral behind it: the need to translate asymmetric c-command into precedence is more urgent for

elements that could be held for each other, i.e., elements of the same kind within the same domain. Other constraints of this type could talk about DPs, e.g., the order of heads within a nominal extended projection, the relative order of subject and object etc.

The relative order of heads and their complements (in particular: verb and direct object, and, here, verb and VP complement) is not regulated by such a constraint, because here we are dealing with symmetric c-command. The violations of the constraint in (18) for each of the six possible orders of our 3-verb clusters are listed in (19).

(19) 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	*

The violations are counted pairwise. We have to consider three pairs of elements: (Aux,Mod), (Aux,V) and (Mod,V), and get at most three violations (order A).

The order of head and complement is, from the point of view of the LCA, string ambiguous, because we are dealing with a sisterhood relation, where two elements c-command each other. Our idea is here, as indicated above, that this string ambiguity is the source of the head parameter: because both possible modes of linearisation are equally (un)marked with respect to LF-PF mapping, we need to establish a linearisation convention. Parameters are typically expressed by conflicting constraints within OT. We assume the two complementary constraints in (20) and (21).

(20) **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.

(21) **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 violations of the three constraints introduced so far are listed in (22) for our six candidates. Violations of MAPch and MAPhc are again counted pairwise: We have to consider two pairs of elements: (Aux,ModP) and (Mod,VP), and get at most two violations (orders C,F, and A,E, respectively).[9]

(22) Violations of LF-PF mapping constraints:

	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	*	**	*

The interaction of these constraints derives the typological variation (without syntactic movement). VP complements cannot fulfil MAPlr(V^0) and MAPch simultaneously: As complements they should be on the left of their governing head to fulfil MAPch, but as co-heads of an extended projection of V their heads should be on its right to fulfil MAPlr(V^0). The relative ranking of these two constraints makes the difference between Swiss German and standard German verb clusters:

(23) Rankings:

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

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

Swiss German dialects, like Dutch and other West Germanic languages, make a difference between nominal and verbal complements of V: while VP complements occur to the right, DP complements occur to the left of V. That MAPch is ranked higher than MAPhc for Swiss German dialects, predicts exactly this: objects occur to the left of their governing verb. The default position of direct objects is left *adjacent* to the verb, as the Zurich German example in (24-a) shows. The object may move higher to the left, but it may not occur to the right:

(24) 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
The Joggel has the chop want eat
((Lötscher, 1978, 4))
c. *De Joggel hät welen-INF ässe-INF es gottlett
The Joggel has want eat the chop

Ranking MAPhc higher than MAPch would yield a VO language of the English type. Standard German treats both types of verbal complements alike – this is achieved by the high priority of MAPch.

4. Focus-dependent orders

As already discussed, stress placement on a particular verb may license a reordering of the verb cluster. StG allows the stressed verb to occur at the left edge of the verb cluster:

(25) e. ... dass sie das Lied SINGEN hat müssen
that she the song sing has must
f. ... das sie das Lied MÜSSEN hat singen
that she the song must has sing

RP allows the stressed verb at the right edge of the verb cluster:

(26) f. ... dass sie das Lied müssen hat SINGEN
e. ... dass sie das Lied singen hat MÜSSEN

Standard German, on the contrary, has no edge preferences, and, consequently, does not allow for the order F:

(27) e.' ... dass sie das Lied SINGEN hat müssen
e." ... dass sie das Lied singen hat MÜSSEN
f. *... dass sie das Lied müssen hat singen

We seem to find two different strategies that can be described as follows:

Strategy A: favouring one particular edge (RP: right edge; StG: left edge)

Strategy B: favouring the syntactically least marked configuration that serves the purpose (standard German)

Stress placement reflects focus interpretation in German. We interpret strategy A as a strategy that favours certain positions for focus

placement. In particular, we assume the following two constraints for StG and RP, respectively.[10]

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

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

The constraint rankings of StG and RP are then as follows:

(30) **StG:** FocL $\gg$ MAPlr(V^0) $\gg$ MAPch
RP: FocR $\gg$ MAPch $\gg$ MAPlr(V^0)

This illustrates the core idea of our analysis: abstract syntax and focus are two different 'forces' that compete in establishing the order in the verb clusters. While the highest syntactic constraints want the most prominent syntactic element, Aux, at the left or right edge, respectively, the focus constraints want the focus to occur at that edge. The rankings in (30) give higher priority to the focus constraints, but, as we will see below, the syntactic constraints still play a decisive role. Another interesting outcome is that the quite rare order F, which is possible in both dialects, occurs precisely under these circumstances: focus and abstract syntax compete for the same edge of the verb cluster for the element that they treat as most prominent.

Let us have a closer look at the predictions that are made by this system. We will first have a look at StG. In the following OT tableaux, the input is an abstract syntactic, semantic and information structural specification, but because everything is kept constant except for focus, we only specify this part of the input. The candidates are linearisations, i.e., PFs. We only look at the relevant parts of the candidates, i.e., the verb cluster internal linearisations.[11] We will start with the competitions for narrow scope on each of the three verbs.[12]

(31)

StG: Narrow Focus on V	FocL	MAPlrV^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	*!	*	**

For narrow focus on V, FocL selects the orders A and E, and the syntactic constraint MAPlrV^0 chooses among these two candidates, favouring order E (31). Thus, the LF-PF mapping is still obeyed as much as possible. This pattern also shows up with the other two competitions.

(32)

StG: Narrow Focus on Mod	FocL	MAPlrV^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, the orders B and F are selected by FocL, and order F is preferred by MAPlrV^0 (32).

(33)

StG: Narrow Focus on Aux	FocL	MAPlrV^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	*!	*	**

Narrow focus on Aux favours C and D, and then C emerges as optimal here Last. Thus, the three orders that we find in StG are already derived with these three competitions.

Strategy B, the standard German strategy, cannot simply be derived by ranking FocL and FocR equally high. This would yield order A for most foci, as in (34) for focus on V:

(34) *Wrong standard German ranking!*

Narrow Focus on V	FocR	FocL	MAPch	MAPlrV^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		*	*!*	*

The only candidates that are excluded by the focus placement constraints are those that have the focused verb at neither edge (B,D). For the determination of the winner among those candidates that survive, the ('syntactic') LF-PF mapping constraints are crucial. Hence, order A has a high chance to win in many competitions. This is indeed the case. The ranking in (34) yields the following winners for competitions with different foci:

(35) *Winners according to the ranking in (34)* :

Focus on V:	⟶ order A = V Mod Aux
Focus on Mod:	⟶ order D = Aux V Mod
Focus on Aux:	⟶ order A = V Mod Aux
Focus on V+Mod:	⟶ order A = V Mod Aux
Focus on Mod+Aux:	⟶ order A = V Mod Aux
Focus on V+Mod:	⟶ order A = V Mod Aux
Focus on V+Mod+Aux:	⟶ order A = V Mod Aux
Focus on no verb:	⟶ order A = V Mod Aux

If we want to know how strategy B works, we need to find out, what could be the advantage of the additional order E. We assume that it has to do with what we call *ideal* focus interpretation. If the most deeply

embedded constituent bears the main stress of the clause, and the words are in 'canonical order', then focus can be *maximally projected*. All the three indicated foci are possible in (36), which has the standard German default order A:

(36) ... weil Hans Maria (((SINGEN)$_{F1}$ hören)$_{F12}$ wird)$_{F123}$
because Hans Maria sing hear will

Thus, (36) is ambiguous with respect to focus. Furthermore, focus usually *tends* to be projected. Narrow focus on 'SINGEN' in (36) requires extra heavy stress. For narrow focus on V, order E is a better, because unambiguous, choice:

(37) e. ... weil Maria das Lied (SINGEN)$_{F3}$ wird müssen
because Maria the song sing will must

We assume that this is where the advantage of order E, and perhaps marked orders in general, lies. Although German does not have genuine focus positions, some configurations are better than others for the expression of a particular focus.[13]

We express this tendency as another violable OT constraint that evaluates the internal verb order and stress pattern in a phonological phrase formed by a verb cluster with respect to its ideal, i.e., maximal, focus interpretation.

(38) **Ideal Focus (IF)**
The intended focus interpretation given in the input matches the *ideal* focus interpretation of a candidate.

We define *Ideal Focus* in the following way:

(39) *Ideal focus*
The ideal focus is the set of elements that is constructed by the following procedure: start with the stressed element, project focus as far as possible *in one direction*, i.e., if the embedding verb is left (right) adjacent, then focus is projected, if the next embedding verb is again left (right) adjacent, focus is projected further again, etc.

For the six orders with stressed V, the ideal foci are as in (40):

(40) Ideal focus with stress on V:
a. [**V** Mod Aux]
b. [Mod **V**] Aux
c. [Aux Mod **V**]
d. Aux [**V** Mod]

e. [**V**] Aux Mod
f. Mod Aux [**V**]

(41) is a more schematic representation:

(41) a. [**V3** V2 V1]$_{F321}$
b. [V2 **V3**]$_{F32}$ V1
c. [V1 V2 **V3**]$_{F321}$
d. V1 [**V3** V2]$_{F32}$
e. [**V3**]$_{F3}$ V1 V2
f. V2 V1 [**V3**]$_{F3}$

The 'intended focus' is contained in the input. Hence, IF is another constraint on input-PF correspondence, here, it is semantics-PF correspondence. For standard German we assume that IF is ranked high.

For narrow focus on V, IF now chooses those candidates that have the focused verb *isolated* at one of the two edges. The difference to the failed implementation discussed above, with ranking FocL and FocR equally high, is that there the focus ambiguity of the evaluated configurations was not taken into account. What IF does, in a way, is determining the 'unmarked' focus for a candidate and comparing it with the focus specification given in the input.

(42) Standard German:

Narrow Focus on V	IF	MAPch	MAPlrV^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		**!	*

Now, the LF-PF mapping constraints only decide between the orders E and F, and MAPch decides for order E. Thus, it is correctly predicted that order E is possible with stress on V in standard German.

Narrow focus on Mod yields order D, by nearly the same procedure. Now the orders B and D are competing, and here, the lower ranked MAPlrV^0 makes the decision. Thus, we see that this constraint, which only seemed to be active in Swiss German, is also active in standard German.

(43) Standard German:

Narrow Focus on Mod	IF	MAPch	MAPlrV^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	*!	**	*

With narrow focus on Aux, we yield the default order, because for IF all candidates are equally good: Aux is the highest element, so no focus projection is possible and no ambiguity can arise. In many dialects that we explored, we observed this kind of freezing effect to the unmarked order, if Aux is stressed. One of our Swabian informants, for instance, rarely allowed for order A. But it suddenly was the only possible option with stress on Aux.

(44) Standard German:

Narrow Focus on Aux	IF	MAPch	MAPlrV^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 can now also discuss the narrow focus competitions for RP, where we will see that IF is also active, in addition to FocusRight:

(45)

RP: Narrow Focus on V	FocR	MAPch	IF	MAPlrV^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		**		*

The highest ranked FocR selects the orders C and F, which are equal at MAPch. And now IF makes the difference and decides for order F. The same happens with focus on Mod, where IF prefers order D over E:

(46)

RP: Narrow Focus on Mod	FocR	MAPch	IF	MAPlrV^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	*!	**	*	*

With narrow focus on Aux, we again yield the default order:

(47)

RP: Narrow Focus on Aux	FocR	MAPch	IF	MAPlrV^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	*!	**		*

In (48), we compare what we have derived so far with the empirical findings in the two dialects that we are comparing.

		stress on V	stress on Mod	stress on Aux
StG	Found:	C,E	C,F	C
	Predicted:	E	F	C
RP	Found:	A,D,F	A,D,E	A,D
	Predicted:	F	D	A

A number of issues still need to be resolved:

1. Order C is missing in StG for stress on V and Mod
2. Orders A and D are missing in RP for stress on V and Mod
3. Order E is missing in RP for stress on Mod
4. Order D as a second default pattern in RP is yet unexplained

These issues will be targeted in the next section, together with some further problems.

5. Some Problems

The general strategy for the solution of the problems of the missing orders, the first three problems listed at the end of the preceding section, is that there are more possible foci than narrow foci. We will see that most missing orders can be found as winners of such competitions. But we will also have to solve some harder problems. The subsequent subsections deal with more complex verb clusters (briefly) and the optional default order D in RP and other standard German dialects. We also briefly discuss another standard German dialect, Upper Hessian, which has some interesting properties.

5.1. Complex Foci

In section 4, we only looked at narrow focus competitions. When we consider competitions for more complex foci, we find some of the missing orders. Unproblematic cases in StG are the competitions for focus on Mod+Aux (stress on Mod), and V+Mod+Aux (stress on V). Both competitions are won by the default order C. Thus, the missing orders for StG are actually derived:

(48) Complex focus competitions for StG:

a.

Focus on Mod+Aux	FocL	MAPlrV^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		∗!	∗∗

b.

Focus on V+Mod+Aux	FocL	MAPlrV^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		∗!	∗∗

However, a problem occurs with a complex focus on V+Mod. Here, the ungrammatical order B is wrongly predicted to win:

c.

Focus on V+Mod	FocL	MAPlrV^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 truth is that the orders A and B are not possible in StG at all. Thus, in order to avoid the problem in (48), we only need a high ranked constraint that 'switches' these orders 'off'. What A and B have in common, is that they have the auxiliary in the final position. Obviously, this dialect tends to avoid functional and/or finite verbs at the right edge of the cluster. We can formulate this with a constraint

like MAPlrV^0 for functional verbs (where 'functional' first of all means finiteness and presumably tense):

(49) **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.

If the constraint in (49) was ranked high, we would never find an order, where V1 occurs last. However, a clause-final finite verb is possible in 2-verb clusters in StG ((Schönenberger, 1995, 366)):

(50) a. das t chatz fisch ässe mues
that the cat fish eat must
"that the cat must eat fish"
b. das t chatz fisch mues ässe
that the cat fish must eat

This order is even obligatory with an auxiliary ((Schönenberger, 1995, 66)):

(51) a. das t chatz fisch gässe hät
that the cat fish eaten has
b. *das t chatz fisch hät gässe
that the cat fish has eaten

The difference between (51) and (50) might be due to the difference between infinitival and participial verb forms, with the latter being required to be left to their governing verb more urgently than the former. For 3-verb-clusters, a final auxiliary is impossible, likewise for clusters with 4 (cf. (Schönenberger, 1995)) and presumably more verbs. Hence, the complexity of the verb cluster triggers the prohibition of verb final functional verbs. The method of constraint conjunction is a way to reflect this directly in OT:

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

This constraint is ranked high and thus blocks the orders A and B in 3-verb clusters:

(53) Violations of MAPlr(V^0_{func}) and MAPlr(V^0_{func})^2:

	MAPlr(V^0_{func})^2	MAPlr(V^0_{func})
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 only functional verb in the sense of the constraint is the auxiliary. When we determine the constraint violations, we need to consider two pairs, (Aux,Mod) and (Aux,V), and get at most two violations (A,B). The ranking for StG is as follows:

(54) StG ranking (revised):
MAPlr(V^0_{func})^2 $\gg$ FocL $\gg$ MAPlr(V^0) $\gg$ MAPch

This gives us the correct prediction for focus on V+Mod:

(55)

Focus on V+Mod	MAPlr(V_{func})^2	FocL	MAPlrV^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 findings for StG are now completely reconstructed:

(56) Predictions for StG:

		stress on V	stress on Mod	stress on Aux
StG	Found:	C,E	C,F	C
	Predicted:	C,E	C,F	C

The next subsection discusses some further evidence for the constraint that we just introduced.

5.2. Complexity: Another effect of MAPlr(V^0_{func})^2

For standard German, one can easily make the observation that the larger a verb cluster is, the stronger is the *pressure to give up the default order*:

(57) a. weil sie es sehen wird
because she it see will
Order: V2-V1

b. weil sie es sehen können wird
because she it see can will
Order: V3-V2-V1

c. ?weil er sie es sehen lassen können wird
because he her it see let can will
Order: V4-V3-V2-V1

d. ?*weil er sie die Kinder spielen sehen lassen können wird
because he her the children play see let can will
Order: V5-V4-V3-V2-V1

This is directly mirrored in our system of constraints by increasing violations of MAPlr(V^0_{func}). Swiss German dialects do not show such an effect, because their default order is already in accord with MAPlr(V^0_{func}). (58) shows improved standard German versions of (57-c,d):

(58) c'. weil er sie es wird sehen lassen können V1-V4-V3-V2

d'. ?weil er sie die Kinder wird können spielen sehen lassen V1-V2-V5-V4-V3

Our claim must thus be that some conjoined version of MAPlr(V^0_{func}) is ranked high enough in standard German to take effect – if not MAPlr(V^0_{func})^2, then perhaps MAPlr(V^0_{func})^3 or MAPlr(V^0_{func})^4.[14]

Schmid (2002) presents a more detailed discussion of this effect. It is also shown there, that MAPlr(V^0_{func}) might have to be seen as a family of constraints. The word order restrictions imposed by the different temporal auxiliaries are differently strong: perfect auxiliaries derived from *haben* ('to have') have the strongest tendency to occur in verb-cluster initial position, for the future auxiliary *werden* ('to become') and most finite modal verbs this is optional, and perfect future auxiliaries derived from *sein* ('to be') seem to cluster together with finite

predicative verbs in that they tend to occur in their default position independent of the size of the verb cluster.

5.3. Optionality of unmarked orders in RP

Thus far, we dealt with the optionality of word orders by treating it as what Müller (2000) calls 'pseudo-optionality', namely, that the optional orders are in fact the *only* winners of a *particular* competition that is defined by a particular information structural specification. This kind of strategy seems to be inapplicable in the case of the two apparently equally unmarked orders that we observe for many standard German varieties, including RP, the orders A and D. Here, we would like to see two winners for a competition within a neutral context. However, our LF-PF mapping constraints make a difference between the two orders we are talking about. Because the two candidates have different *constraint violation profiles*, they can never be winners within the same competition simultaneously, if all constraints are unambiguously ranked.

A second, in this case more promising, way of deriving 'real' optionality in OT is assuming that those constraints where the two optional candidates differ are not ranked with respect to each other. They are *tied*. This is the strategy we use in accounting for the problem of the two default orders in standard German and its dialects (like RP). In particular, we assume that MAPch and MAPlr(V_{func})^2 are *globally tied*, i.e., there are two *co-grammars* in that dialect, where the two constraints are ranked alternatingly:[15]

(59) Ranking for RP:
FocR $\gg$ MAPch $\circ$ MAPlr(V_{func})^2 $\gg$ IF $\gg$ MAPlr(V^0)

How the two co-grammars work is exemplified with the following two tableaux, which show how the two default orders are derived:

(60) a. First ranking (LF-PF constraints only, deriving default order):

	MAPch	MAPlr(V_{func})^2	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	*!*		*

b. Second ranking:

	MAPlr(V_{func})^2	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 in (60-b), why order D is an optimal candidate, as soon as order A is excluded: it performs quite well at both MAPch and MAPlr(V^0). With this global tie, we get the following winners for the different focus competitions:

(61) Outcomes in RP, winners only:

stress on V3	/F3/	⟶	[F213]
	/F32/	⟶	[D132]
	/F321/	⟶	[A321],[D132]
stress on V2	/F2/	⟶	[D132]
	/F21/	⟶	[A321],[E312]
stress on V1	/F1/	⟶	[A321]
no focus		⟶	[A321],[D132]

The predictions for RP are now nearly reconstructing our empirical findings, as illustrated in (62):

(62)

		stress on V	stress on Mod	stress on Aux
StG	Found:	A,D,F	A,D,E	A,D
	Predicted:	A,D,F	A,D,E	A

What is still missing, is the default order D for stress on Aux. If we want to include even this, we only need to assume a third co-grammar, where MAPlr$(V_{func})^2$ is ranked even higher than FocR. In that case, the only two orders that satisfy FocR for focus on Aux, A and B, are excluded from the start, and the syntactically least marked candidate, order D, is the winner, as it would also be for most other competitions.

We might abbreviate the three co-grammars in the following way:

(63) Ranking for RP:
(FocR $\gg$ MAPch) $\circ$ MAPlr$(V_{func})^2 \gg$ IF $\gg$ MAPlr(V^0)

This expresses that MAPlr$(V_{func})^2$ is tied with the subranking "FocR $\gg$ MAPch", and is ranked either below the two constraints, on top of them, or between them, while the relative order of FocR and MAPch remains constant.

A final issue that at least needs to be adressed briefly, is that, under the above ranking, we would predict that focused objects tend to be right dislocated in RP. This does not seem to be the case. On the contrary, direct objects have to strictly precede the predicative verb. To guarantee this, we would need a further constraint. Our suggestion in this case is that MAPch has to be obeyed more strictly for head-complement relations where the head assigns a thematic role to the complement. A constraint like the following would be appropriate:

(64) **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 natural 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:

(65) Final ranking for RP:
MAPch$\Theta \gg$ (FocR $\gg$ MAPch) $\circ$ MAPlr$(V_{func})^2 \gg$ IF $\gg$ MAPlr(V^0)

As FocusRight cannot be obeyed by a focused object, the system falls back to the default orders, A and D. Independent motivation might come from the fact that even verb clusters are sensitive for this restriction. If verb 1 is a causative verb, which presumably assigns a thematic role to its VP complement, then standard German only allows for the canonical order A, strictly obeying MAPch, as shown in (66).

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

This finding is predicted by the above ranking. From this perspective, it is no accident that the most flexible verb clusters are those that show no thematic relations between the verbs, as in our example clusters of predicative verb, modal and temporal auxiliary. However, a fully satisfactory account would have to establish the distinction between auxiliaries and modals on the one hand and causatives and other 'thematic' verbs on the other hand in semantic and θ-theoretic terms. This goes beyond the scope of this paper.

5.4. Upper Hessian

We finally want to introduce a dialect that has some further interesting properties. In Upper Hessian, only the D pattern wins sometimes. This is illustrated in (67):[16]

(67) a. *... dass sie es ihn singen gehört/hören hat/HAT
that she it him sing heard/hear has
d. ... dass sie es ihn hat/*HAT singen hören
that she it him has sing hear

The perfect auxiliary in this dialect cannot occur in final position and cannot be stressed either. To include this phenomenon, we assume the following constraint:

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

What counts as a weak element is obviously a parametrised difference among dialects. Most other standard German dialects do not seem to have such weak elements – all auxilaries can, for instance, be stressed in standard German.

Upper Hessian also has a very limited influence of stress marking on verb orders. In clusters with a weak Aux, only order D or E are possible. The outcomes are listed in (69):

(69) Outcomes with weak Aux in Upper Hessian, winners only:

stress on V3	/F3/	⟶	[E312]
	/F32/	⟶	[D132]
	/F321/	⟶	[D132]
stress on V2	/F2/	⟶	[D132]
	/F21/	⟶	[E312]
stress on V1	/F1/	⟶	impossible[17]
no focus	/NoF/	⟶	[D132]

The constraint ranking that we assume for Upper Hessian is the following one:

(70) Ranking for Upper Hessian:
*WkFin ≫ MAPch ≫ IF ≫ MAPlr(V^0) FocL FocR

We see that the focus constraints are ranked quite low. They only seem to have a marginal influence.

This kind of microvariation within standard German dialects can also be found within the Swiss German dialect family: For one *Bernese Swiss German* informant, only the default order C is possible, no matter which intonation is used. Here, we obviously have MAPlr(V^0) ranked on top, such that focus constraints take no effect.

6. Summary

We hope to have shown that OT is an ideal framework for the modelling of a multifactorial explanation of the word order facts in 3-verb clusters of German dialects, as well as its typological diversity. The factors that we looked at in particular are syntactic and information structural. Whether FocR and FocL are actually information structural, rather than phonological, is an open issue, however.

FocusRight could be an instance of the compound stress rule in RP. Northern German dialects show surprising stress patterns for compounded geographical names. The following examples are from Bremen (Northern Germany). The first one is the name of a federal state, the other two are street names in Bremen:

Niedersáchsen (Bremen) vs. *Níedersachsen* (standard German) (= 'Lower Saxony')
Sielwáll (Bremen) vs. *Síelwall* (SG) (≈ 'floodgate mound')

Buntentór (Bremen) vs. *Búntentor* (SG) (unclear, perhaps 'coloured gate', or 'union gate')

There might be a general tendency to favour the right edge of a prosodic domain for the main accent in Northern German dialects. This is an issue that needs further investigation. The same would have to be explored for Swiss German dialects. We might discover further factors, like metrical properties, and also morpho-phonological ones, as already found in Upper Hessian. The issue certainly requires more empirical investigation.

Zwart (1996) discusses the syntax of verb clusters in the light of the discussion, whether West Germanic varieties are underlyingly OV or VO structures. Our contribution to this discussion is perhaps an explanation, why this issue is so difficult to decide. Although the parameter setting "MAPch $\gg$ MAPhc" holds for all the varieties under discussion, there are a number of other factors that intervene in such a way that this 'underlying' parametrisation is very hard to recover.

Notes

[1] German dialects vary a lot in their morpho-phonology. As we are only concerned with word order facts here, we are abstracting away from these differences, and only give the examples, with a few exceptions, in their standard German 'translation'. Order (71-b) is extremely rare in German. We found only one native speaker of one Swabian dialect, who considered it as not ill-formed under a particular stress assignment. We are not discussing this dialect here, however. In Afrikaans, the order seems to have nearly unmarked status for some combinations (see (Schmid, 2002) for further details).

[2] Subordinate clauses are verb-final in German. While in main clauses the finite verb moves to second position, it remains within the clause-final verb cluster in standard subordinate clauses. What holds of clause-final 3-verb clusters in subordinate clauses, usually also holds of them in main clauses, too. Verb-Second is only an additional factor that we want to abstract away from in our discussion.

[3] This analysis cannot be found directly in ((Haegeman and van Riemsdijk, 1986)). However, we only make use of the mechanisms they propose.

[4] "The additional complication, that one single rule type is hardly sufficient to account for word order, must be taken into acount as well. Rather, there are at least three interacting but primarily independent kinds of rules: First, **grammatical rules** [...] that determine an order more or less arbitrarily [...]; **performance rules** [...]; at last, **functional rules** [...] that allow for certain functional relations in a sentence in the sense of the topic-comment distinction." (translation by us, TS/RV)

[5] "[...]So we may conclude that the choice of the word order type is connected to the intonation and to the accent of the dialect in question. Primarily, we look at the rhythm again that is hugely important for word order questions. [...]" (translation by us, TS/RV)

[6] One exception is the work of Pesetsky, Pesetsky (1997, 1998) who assumes an OT system for the LF-PF mapping, but not for 'syntax proper'. This kind of approach is fairly compatible to what we propose here.

[7] Wurmbrand (2000) assumes something along the same lines in arguing for a PF-oriented solution. She claims that the core semantic properties do not change with the order in the verb cluster. This is true for scope relations between the verbs, for instance. But information structural properties do change. However, these need not be encoded abstract syntactically.

[8] The notion of 'extended projection' takes V and N as basic syntactic categories on top of which several (semi-)functional projections can be stacked. An NP can, for instance, be projected up to the level of PP, and a VP, up to the level of CP. For our analysis, three assumptions are important. First, subordinate and matrix clauses are extended projections of different verbs. This is uncontroversial. Second, modals and auxiliaries do not constitute their own extended projections, at least not in German verb clusters. This is perhaps more controversial. As a rule of thumb might serve that an extended projection of V has to contain exactly one finite verb or infinitive marker (like *zu*, 'to'). A third assumption is about complementisers: Though Grimshaw treats complementisers and prepositions as the outmost heads of their extended projections, what is striking, at least in German, is that complementisers are totally different from verbs. The default complementiser, *dass*, 'that', in fact developed from the neuter d-pronoun, and is, thus, more nominal than verbal. It might be more conclusive to say that (German) complementisers are not part of extended projections, but, rather, that they only *embed* a verbal extended projection. For our discussion, we assume that this is the case. The syntax of complementisers is an independent issue that is not focused on in this paper. An alternative to Grimshaw's extended projections is the conception of 'M-Projection', developed by van Riemsdijk (1998). For our purposes, the two notions seem to be equivalent.

[9] Note that we assume the values of these constraints to be Boolean: in order to fulfil MAPch, *all parts* of a complement have to follow the head. But the constraint makes no difference in the grade of violation: whether all of the complement follows the head, or only a part of it, does not matter: MAPch is not fulfilled. Thus far, we see no reason to introduce gradiency into these constraints.

[10] Samek-Lodovici () proposed constraints which are similar, but he does not take phonological representations as basis, but rather requires focus to be aligned with the left or right edge of VP. In more recent work ((Samek-Lodovici, 2002)), he proposes constraints on prosodic phrasing that require the heads of prosodic representations to be aligned with their right edge. Because focus bearing elements have main stress, they are heads of their prosodic phrases and are thus required to occur at the edges. In this sense, the constraints we propose are comparable to those used by Samek-Lodovici. A difference might occur with complex foci: here it might occur that the focus is aligned with an edge, but that the stressed element is not necessarily also on the edge. This would be allowed by our constraints, but incur violations of Samek-Lodovici's constraints, unless he makes them more liberal. He furthermore only uses right-edge alignment. How to derive the St.Gallen findings, for example, in this approach, is not clear.

[11] Prosodic structuring is also left aside. For OT accounts of this problem, see ((Truckenbrodt, 1999), (Samek-Lodovici, 2002)). We assume that our approach is compatible with these. In the examples we discuss here, the relevant parts of the prosodic structure are kept constant, in particular, the verb clusters make up clause-final phonological phrases of their own.

[12] Narrow focus on the auxiliary has the effect of a kind of verum focus interpretation, emphasising that something is *indeed* the case, perhaps contrary to what has been claimed before.

[13] For a detailed discussion of the German focus facts see, for instance, (Uhmann, 1991).

[14] Multiple self-conjunctions of constraints establish what Legendre et al. (1998) call a *power hierarchy*. The idea is that multiple violations of a constraint can cumulate up to a point where they have a qualitative effect. The crucial scenario is the following: Assume that constraint A is ranked higher than constraint B. It's violations are more important, no matter how many violations of B we have. In order to establish cumulative effects, we need a constraint C that is ranked higher than A, and that is violated, if B is violated n times. As long as C is ranked immediately on top of B (or lower, which is usually excluded by convention), the system behaves as if C was not there at all.

[15] Several versions of constraint ties have been proposed in the literature. For an overview, see ((Müller, 2000)). A global tie is not really a tie, but actually a notational convention for the abbreviation of two existing co-grammars without tie.

[16] Upper Hessian is spoken in a region that is, roughly, between 30 and 100 kilometers north of Frankfurt/Main.

[17] This raises the issue of ineffability, which is a notorious problem in OT. In the case at hand, native speakers tend to use simple past instead of present perfect: "... singen HÖRte" ('sing HEARD'). It might be possible to include this structure as a (winning) candidate. Note that for Upper Hessian speakers there is no semantic difference between present perfect and simple past, and they have a strong preference for periphrastic tense forms, even in present tense, where 'tun'('do')-insertion is very frequent and has no stylistic or emphatic effects of any kind. It thus seems that synthetic tense forms are the marked case here, and might be a perfect candidate for a repair form.

References

Archangeli, D. and D. T. Langendoen (eds.): 1997, *Optimality Theory. An Overview*. Malden, Mass. and Oxford: Blackwell.

Barbosa, P., D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky (eds.): 1998, *Is the best good enough? Optimality and competition in syntax*. Cambridge, Massachusetts: MIT Press.

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

Chomsky, N.: 1995, *The Minimalist Program*. Cambridge, Massachusetts: MIT Press.

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

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

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

Grimshaw, J.: 1991, 'Extended Projection'. Manuscript, Brandeis Univesity.

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

Hayes, B., B. Tesar, and K. Zuraw: 1998, 'OTSoft. Software Package'. http://www.linguistics.ucla.edu/people/hayes/otsoft/.

IJbema, A.: 1997, 'Der IPP-Effekt im Deutschen und im Niederländischen'. *Groninger Arbeiten zur germanistischen Linguistik* **40**, 137–163.

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

Karttunen, L.: 1989, 'Radical Lexicalism'. In: M. Baltin and A. Kroch (eds.): *Alternative Conseptions of Phrase Structure*. Chicago: The University of Chicago Press, pp. 43–65.

Kayne, R.: 1994, *The Antisymmetry of Syntax*. Cambridge, Massachusetts: MIT Press.

Kleinhenz, U.: 1994, 'Focus and phrasing in German'. Technical Report 52, IBM Deutschland, Wiss. Zentrum, Inst. fr Wissensbasierte Systeme, Heidelberg.

Koopman, H. and A. Szabolcsi: 2000, *Verbal Complexes*. Cambridge, Massachusetts: MIT Press.

Legendre, G., P. Smolensky, and C. Wilson: 1998, 'When is less more? Faithfulness and Minimal Links in WH-Chains'. In (Barbosa et al., 1998), pp. 249–289.

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

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

Meurers, W. D.: 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, G.: 2000, *Elemente der optimalitätstheoretischen Syntax*. Tübingen: Stauffenburg.

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

Pesetsky, D.: 1997, 'Optimality Theory and Syntax: Movement and Pronunciation'. In (Archangeli and Langendoen, 1997), Chapt. 5, pp. 134–170.

Pesetsky, D.: 1998, 'Some Optimality Principles of Sentence Pronunciation'. In (Barbosa et al., 1998), pp. 337–383.

Samek-Lodovici, V., 'OT-interaction between Focus and Canonical Word Order: Deriving the crosslinguistic Typology of Structural Contrastive Focus'. Rutgers Optimality Archive, http://roa.rutgers.edu, ROA-257.

Samek-Lodovici, V.: 2002, 'Prosody-Syntax Interaction in the Expression of Focus'. Rutgers Optimality Archive, http://roa.rutgers.edu, ROA-524.

Schmid, T.: 2002, 'West-Germanic IPP-Constructions. An Optimality Theoretic Approach'. Ph.D. thesis, Institut für Linguistik, Universität Stuttgart.

Schönenberger, M.: 1995, 'Constituent Order in the VP: Verb Raising and Verb Projection Raising'. In: Z. Penner (ed.): *Topics in Swiss German Syntax*. Bern: Peter Lang, pp. 347–411.

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

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

Truckenbrodt, H.: 1999, 'On the Relation between Syntactic Phrases and Phonological Phrases'. *Linguistic Inquiry* **30**, 219–255.

Uhmann, S.: 1991, *Fokusphonologie*. Tbingen: Niemeyer.

van Riemsdijk, H.: 1998, 'Categorial Feature Magnetism: The Endocentricity and Distribution of Projections'. *Journal of Comparative Germanic Linguistics* **2**, 1–48.

Vogel, R.: 2002, 'Dialectal Variation in German 3-Verb Clusters. Looking for the best Analysis'. In: A. Mahajan (ed.): *Proceedings of the Workshop on Head Movement, UCLA, October 2001*, UCLA Working Papers in Linguistics. Los Angeles, UCLA: .

Wurmbrand, S.: 2000, 'Verb Clusters: Variation at the Right Periphery'. Handout, presented at the Workshop on Syntactic Microvariation, Amsterdam, August 2000.

Wurmbrand, S.: 2001a, *Infinitives: Restructuring and clause structure*. Berlin: Mouton de Gruyter.

Wurmbrand, S.: 2001b, 'Verb Clusters, Verb Raising, and Restructuring'. Online: http://www.arts.mcgill.ca/programs/linguistics/faculty/wurmbrand/research/files/verbclusters/Syncom.pdf. To be included in *The Syntax Companion*, Blackwell: http://www-uilots.let.uu.nl/syncom/.

Zwart, J.-W.: 1996, 'Verb clusters in Continental West Germanic dialects'. In: J. Black and V. Motapanyane (eds.): *Microparametric syntax and dialect variation*. Amsterdam: John Benjamins, pp. 229–258.