Dialectal Variation in German 3-Verb Clusters
A Surface-Oriented Optimality Theoretic Account*

TANJA SCHMID
Fachbereich Sprachwissenschaft, Universität Konstanz, Germany
(E-mail: tanja.schmid@uni-konstanz.de)

RALF VOGEL
Institut für Linguistik, Universität Potsdam, Postfach 601553, D-14415 Potsdam, Germany
(E-mail: rvogel@ling.uni-potsdam.de)

Key words:

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 greatly among the dialects. Some of the orders that we found occur only under particular stress assignments. We assume that these orders fulfil an information structural purpose and that the reordering processes are changes only 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 as to which permutations of the verb order in clause-final 3-verb clusters they allow. In an empirical investigation we have-

found that five of the six logically possible permutations of the 3-verb cluster in (1) are clearly acceptable in at least some German dialects:[1]

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

   a. sie das Lied singen müssen wird
      she the song sing   must   will

      she will have to sing the song

   b. %sie das Lied müssen singen wird
   c. sie das Lied wird müssen singen
   d. sie das Lied wird singen müssen
   e. sie das Lied singen wird müssen
   f. sie das Lied müssen wird singen

The verb clusters that we are exploring are exclusively 3-verb clusters of the form abbreviated in (2). We are well aware, however, that both verb class of V2 and construction type (modal construction, future tense, perfect tense with either past participle or Infinitivus Pro Participio, IPP) have a large influence on the verb order possibilities in 3-verb clusters (see e.g., Den Besten and Edmondson 1983, Wurmbrand 2001, Schmid 2002 for discussion).

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

We restrict ourselves to constructions as given in (2) for the following reason: we focus on the role of information structure in 3-verb clusters in this paper. To 'pin down' the pure effects of information structure, we try to keep our examples as constant as possible and deliberately exclude other factors like verb class and construction type. This does not mean, however, that these factors are not important. In an exhaustive account of verb clusters (which is beyond the scope of this paper) they are to be integrated as well.

   The reason why we chose verbal complexes like (2) is that modals as V2 are very 'flexible'. We find more variation in verb order than with other combinations. This may be connected to the fact that modals need

---

[1] German dialects vary greatly in their morphophonology. As we are concerned only 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 (1-b) is extremely rare in German. However, it is very common in other West Germanic languages, for instance, Afrikaans (see Donaldson 1993, Robbers 1997, Schmid 2002 for further discussion) and West Flemish.

not be realized as past participles in auxiliary-modal-verb constructions due to their IPP-property (see Schmid 2002). In most West Germanic languages, the order of a bare infinitive is less restricted than the order of a past participle.

We *always* represent the basic syntactic relations between the three verbs as in (3).

(3) $[_{VP1}$ V1 $[_{VP2}$ V2 $[_{VP3}$ V3]]]

We follow recent assumptions in generative syntax insofar as we assume that a structure like (3) encodes only dominance but not precedence relations (cf. Kayne 1994). Hence, that the heads are on the left in (3) is only a representational convention without any implication for the actual linear order of the elements. However, we also assume that linearization 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 them was as follows: we developed a questionnaire that contained five repetitions of the block of six permutations in (1). The blocks differed as to which element carried the main stress. Main stress was indicated by uppercase and varied over subject, object, V3, V2, V1 (in that order). Each of the blocks was accompanied by a context sentence that was supposed to help identify the focus interpretation that is correlated with the particular stress pattern – the context clause usually gave preference to narrow focus. Altogether there were 30 example sentences to test.[2]

The native speakers of the dialects were asked to do two things: translate the clause literally into their home dialect, and then give a grammaticality judgment for that clause. The number of informants was rather small – usually, but not always, there was only one person per dialect. Table I

---

[2]   Only subordinate clauses were tested. 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 is true of clause-final 3-verb clusters in subordinate clauses usually also holds true in main clauses as well. Verb-Second is only an additional factor that we wanted to abstract away from.

TABLE I
Possible orders of 3-verb-clusters in German dialects with differing location of main stress

	RP	UH	TSwa	Franc.	BavFor	StG	Bernese	Meran
SUBJECT	321 132	132	123 132	321 132	321 132 312	123	123	321 123 132 312 213
OBJECT	321 132	132	123 132 312	132	321 132 312	123	123	321 123 132 312 213
VERB	321 132 213	132 312	123 132 312	321 132	321 123 132 312	123 312	123	321 123 132 312 213
MODAL	321 132 312	132 312	132 312	321 132 312	321 123 132 312	123 213	123	321 123 132 312 213
AUX	321 132	132 (*haben)[a]	321	321 132 312	321 132 312	123	123	321 123 132

Abbreviations: RP = Rheiderländer Platt, East Frisia; UH = Upper Hessian; TSwa = Swabian (Tübingen); Franc. = Franconian; BavFor = Bavarian Forest; StG = St. Gallen Swiss German; Bernese = Bernese Swiss German; Meran = Meran, Southern Tirol.
[a] The auxiliary *haben* 'have' cannot be stressed in Upper Hessian, hence this stress pattern is impossible with this auxiliary.

lists the orders that our informants found clearly acceptable given the indicated location of main stress.

Given the observed variation, it is clear that there cannot be an uncontroversial result for Standard German. The general tendency suggests to us that three orders are possible, 321, 132, and 312. Order 132 appears to be possible with nearly any stress pattern. Order 321 is possible with most patterns, perhaps with the exception of main stress on the modal verb. Order 312 is more restricted. Only with stress on the modal or the predicative verb might it be uncontroversially accepted. Most German speakers might agree on our list in Table II for Standard German. It summarizes the impressions that we got from informants and from the literature. Orders in parentheses are dispreferred under a given stress pattern and are rejected by some speakers and accepted by others.

TABLE II
Possible orders in Standard German

Main stress on	Standard German
SUBJECT	132 321 (312)
OBJECT	132 321 (312)
VERB	132 321 312
MODAL	132 (321) 312
AUX	(132) 321 (312)

The variation in the verb orders 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'*

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

(4)   Default orders
        a. Standard German (dialects): 321 and 132
        b. Swiss German (dialects): 123

Two criteria are essential for the determination of the default orders: first, default orders should be the most frequent orders; second, under verum focus (main stress on the auxiliary) we often observe restricted variability,

often only the default order is possible. Swabian seems to have the default order only in the case of verum focus, and here it is the only order accepted. For most other variants, default orders can be detected by the first criterion.

With respect to the default order, Swiss German dialects pattern together with West Germanic languages like Dutch. The orders 321 (Standard German) and 123 (Swiss German) are mirror images of each other that presumably reflect opposite settings of a syntactic parameter. The discussion of order 132 which is as unmarked as order 321 in Standard German, is postponed until section 5.3. The finding illustrated in (4) 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 under certain circumstances. An interesting contrast that we observed, and which we want to discuss in more detail here, is the following: 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 and 213, but differ as to which of the verbs receives the main stress:

(5)　　St. Gallen (Swiss German dialect)
　　　　a. stress on V:　　312 = **V** Aux Mod
　　　　b. stress on Mod: 213 = **Mod** Aux V

(6)　　'Rheiderländer Platt' (Low German, Standard German family)
　　　　a. stress on Mod: 312 = V Aux **Mod**
　　　　b. stress on V:　　213 = Mod Aux **V**

The additional orders stress the first (StG) or the last (RP) verb in the verb cluster, as indicated by boldfacing. The possibility of order 213 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 2001). 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 develop 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 not instances of classical syntactic movement. 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 parameterization. Haegeman and van Riemsdijk (1986) assume that a Zürich German example like (7) is derived from an underlying Standard German structure as exemplified by (8):

(7)   Züri Tüütsch (Surface Structure):

       das   er   en   arie   hät   wele   singe
       that   he   an   aria   has   want   sing

(8)   Züri Tüütsch (underlying structure like Standard German):
       er [[en arie singe] wele] hät
                    (cf. Haegeman and van Riemsdijk 1986, p. 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 reanalyzed as being dominated by the same V^0 head:

(9)   *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. $V\alpha$ and $V\beta$ change their order. The result is, however, ungrammatical (order 231 = Mod-2 V-3 Aux-1):

(10)   '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 Zürich German default order:

(11)   Reanalysis from (10) 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]]

---

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

Following the basic intuition behind such an approach, we want to elaborate on the idea that the verb order in verb clusters is a matter of linearization and not so much of standard syntactic movement. More recent work that goes in 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 2001 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 by Vogel (2003).

What is needed 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):[4]

> "The additional complication, that one single rule type is hardly suffi-
> cient to account for word order, must be taken into account 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." (Lötscher 1978, p. 11; boldfacing by us, TS/RV)

"[. . .] So we may conclude that the choice of word order type is connected to intonation and to the accent of the dialect in question. Primarily, we again look at rhythm, which is extremely important for word order questions. [. . .] (Maurer 1926, p. 72)

Abstract syntactic relations and properties (i.e., constituency, 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.

(12)   *The multifactorial model:*
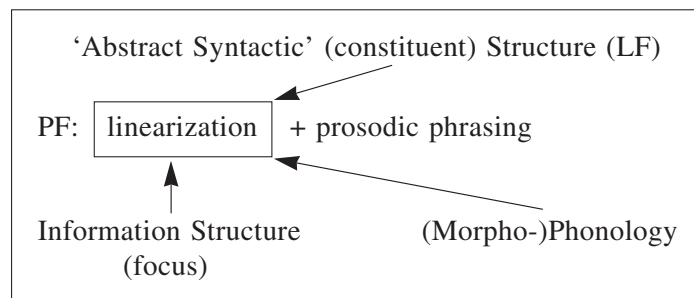
For our OT grammar model, we take an abstract syntactic structure as syntactic part of the *input*. In what follows we call this structure LF (Logical Form), using Minimalist terminology (see Chomsky 1995). This might be somewhat misleading, insofar as LF is usually also assumed to be the input to the 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 (13). The input also contains semantic information, in particular – what is important here – an information structural specification.

The *candidates* are PFs, i.e., linearized (inflected) words, prosodically and metrically structured. These are freely generated by the generation function *GEN*. This model is certainly only a *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 optimization.[5]

---

[5] One exception is the work of Pesetsky (1997, 1998), who assumes an OT system for the

The *constraints* come from the areas discussed above. The architecture of this model is summarized in (13).

(13)   The OT grammar model used here:
    a. **Input** $\langle$*LF* (*constituent structure*, *abstract features*), *semantic representation (including focus)*$\rangle$
    b. **Candidates**: PFs, i.e., linearization + phonological phrasing + stress assignment + morphology
    c. **Constraints**: any constraints on PF formation, correspondence for LF-PF and semantics-PF, phonological and morphological restrictions.

We will now introduce the constraints that we use. Section 3 discusses constraints on LF-PF correspondence; section 4 introduces the information structure constraints.

3. Syntactic constraints on linearization

We assume that the dominance and c-command relations in a 3-verb cluster with a direct object are always the same:[6]

(14)   *Uniform abstract syntactic structure (LF) of the verb clusters:*
$[_{AuxP}$ Aux $[_{ModP}$ Mod $[_{VP}$ V NP$_{Obj}]]]$

But we also assume that the linear order of auxiliary, modal, and predicative verb is subject to an OT evaluation. Any conceivable order is a candidate. What would be an optimal linearization of (14)?

One option for an OT constraint on linearization might be Kayne's (1994) "Linear Correspondence Axiom", rephrased in terms of our model in (15):

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

(15) prefers for the structure in (14) the linearization 'Aux Mod V NP$_{Obj}$'.

---

syntax-PF mapping, but not for 'syntax proper'. This kind of approach is fairly compatible with what we propose here.

[6] 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 abstract syntactically encoded.

Any deviation from that order would have to be derived by syntactic movement. As we are looking for an explanation that functions without syntactic movement, the LCA in the general version given in (15) seems to be too strict for our purpose. Another problem is that we want to be able to talk about particular departures from the order required by the LCA: the relative order of verbs is different in standard and Swiss German dialects, while the relative order of verbs and NPs is not. Instead of the one general constraint in (15) we need a set of less general constraints.

We nevertheless want to maintain the intuition that, as a default, asymmetric c-command is translated into precedence but restricted to cases where this seems to be most crucial, namely, asymmetric c-command relations between elements of the same syntactic category:

(16)   *Translation of asymmetric c-command into precedence*
        If α asymmetrically c-commands β at LF and both are of the same syntactic category, then the correspondent of α precedes the one of β at PF.

One motivation for this restriction is that this factor is important for minimal link phenomena. For instance, movement of an object *wh*-NP is possible in English, but it is blocked if the subject is also a *wh*-NP:

(17) a. What did John say?
     b.*What did who say?

(16) can thus also be seen as a mode of implementing a central aspect of Rizzi's (1990) theory of *relativized minimality*: movement of X across Y is blocked if Y is a potential antecedent for the trace of X. The most important criterion for being a potential antecedent is, of course, equivalence in syntactic features. (16) is not a constraint itself but describes a family of OT constraints. The constraint that will be crucial in our discussion is defined as follows:

(18)   **MAP–left-right (V^0) (MAPlr(V^0))**
        The heads of an extended projection of V are linearized 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.

This constraint does not talk about verbal heads in general but only about verbal heads that belong to the same extended projection (in the sense of Grimshaw 1991).[7] This restriction is mainly assumed for methodological

---

[7] The notion of 'extended projection' takes V and N as basic syntactic categories on top of

reasons and could presumably be left out. The interesting relations in our case are those between verbs of the same extended projection, and we simply want to take those into account. Verbs belonging to different extended projections are usually not ordered relative to each other – only the extended projections they belong to are, and this is presumably regulated by other constraints. There is some independent evidence for the constraint in (18). Finnish, as discussed by Dowty (1996), following Karttunen (1989), makes an interesting distinction between verbs and noun phrases: while the relative order of verbs is fixed, the order of noun phrases seems to be totally unconstrained.

(19) a. En   minä   ole   aikonut ruveta pelaamaan näissä
        not   I   have intend   start   play   these-in

        tennistä
        tennis

        I did not intend to 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 NP *tennistä* and the adverbial *näissä* can permute freely in (19). The only restriction is that the relative order of the verbs remains constant. Thus, Finnish seems to be a language that strictly obeys MAPlr(V^0) but perhaps not a parallel constraint on the relative order of NPs.

   The violations of MAPlr(V^0) for each of the six possible orders of our 3-verb clusters are listed in (20). According to the definition in (18), the violations are counted pairwise. We have to consider three pairs 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 an extended projection of V has to contain exactly one finite verb or infinitive marker (like *zu*, 'to') – one IP. A third assumption is about complementizers: though Grimshaw treats complementizers and prepositions as the outmost heads of their extended projections, what is striking, at least in German, is that complementizers are totally different from verbs. The default complementizer, *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) complementizers 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 complementizers 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 Riemsdijk (1998). For our purposes, the two notions seem to be equivalent.

elements: (Aux, Mod), (Aux, V), and (Mod, V), and therefore get at most three violations (order 321).

(20)   Violations of $\text{MAPlr}(V^0)$:

	$\text{MAPlr}(V^0)$
321: V Mod Aux	***
231: Mod V Aux	**
123: Aux Mod V	
132: Aux V Mod	*
312: V Aux Mod	**
213: Mod Aux V	*

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 symmetrically c-command each other. Contrary to Kayne (1994), we assume that this string ambiguity is the *source* of the head parameter: because both possible modes of linearization are equally (un)marked with respect to LF-PF mapping, the grammar needs to establish a *linearization convention*. Parameters are typically expressed by opposing constraints within OT. We assume the two complementary constraints in (21) and (22).

(21)   **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.

(22)   **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.

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 123, 213 and 321, 312, respectively).[8] The

---

[8] Note that we assume the values of these constraints to be Boolean: in order to fulfil MAPch, *all parts* of a complement have to precede 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 assume that these constraints can be partially fulfilled.

violations of the three constraints introduced so far are listed in (23) for our six candidates.

(23)   Violations of LF-PF mapping constraints:

	MAPlr(V^0)	MAPch	MAPhc
321: V Mod Aux	***		**
231: Mod V Aux	**	*	*
123: Aux Mod V		**	
132: Aux V Mod	*	*	*
312: V Aux Mod	**	*	**
213: Mod Aux V	*	**	*

The interaction of these constraints derives the typological variation in the unmarked orders of the two German dialect families. 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:

(24)   Rankings:
    a. Swiss German:
        MAPlr(V^0) ≫ MAPch ≫ MAPhc → order 123
    b. Standard German:
        MAPch ≫ MAPlr(V^0) ≫ MAPhc → order 321

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, NP complements occur to the left of V. Ranking MAPch on top of MAPhc for Swiss German dialects yields this pattern: objects occur to the left of their governing verb. The default position of direct objects is left *adjacent* to the verb, as the Zürich German example in (25-a) shows. The object may move higher to the left, but it may not occur to the right:

(25) 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, p. 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 complements alike – this is achieved by the high priority of MAPch.

To sum up, we propose that there are two types of syntactic mapping constraints. One type of constraint requires, for asymmetric c-command relations between elements of the same category, a mapping into precedence relations. We assume a family of such constraints for the different syntactic categories, mainly along the lines of the theory of extended projections. Second, for the symmetric c-command relation of head-complement sisterhood we assume a pair of complementary constraints that require head-complement or complement-head order.

4. Focus-dependent orders

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

(26) a. . . . dass sie das Lied SINGEN hat müssen (= 312)
          that she the song sing   has must

     b. . . . das sie das Lied MÜSSEN hat singen (= 213)
          that she the song must   has sing

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

(27) a. . . . dass sie das Lied müssen hat SINGEN (= 213)
     b. . . . dass sie das Lied singen hat MÜSSEN (= 312)

Standard German, on the contrary, has no edge preferences and does not allow for the order 213:

(28) a. . . . dass sie das Lied SINGEN hat müssen (= 312)
     b. . . . dass sie das Lied singen hat MÜSSEN (= 312)
     c.*. . . dass sie das Lied müssen hat singen (= 213)

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

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

Strategy B: favoring the syntactically least marked configuration that serves the purpose (Standard German)

The position of main stress indicates focus in German, as it does in many other languages. We interpret strategy A as a strategy that favors edge positions for focus placement within phonological phrases. We assume the following two constraints for StG and RP, respectively. (30) was introduced by Truckenbrodt (1999):[9]

(29) **FocusLeft (FocL)**
A focused constituent is aligned with the left edge of a phonological phrase.

(30) **FocusRight (FocR)**
A focused constituent is aligned with the right edge of a phonological phrase.

Samek-Lodovici (2001) uses constraints with the same name which require focus to be aligned with the left or right edge of VP.[10]

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

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

In our examples the verb clusters always constitute a phonological phrase of their own if one of the verbs is stressed.[11] (32) displays the prosodic phrasing for a cluster with main stress on the predicative verb in order 321:[12]

(32)
$$(\qquad\qquad\qquad X \qquad\qquad)_{\text{IntP}}$$
$$(\qquad X \;)_{\text{PhP}} (\; X \qquad\qquad)_{\text{PhP}}$$
$$\text{dass sie das Lied} \qquad \text{singen müssen wird}$$

Abstract syntax and focus compete in establishing the order in the verb-clusters. While the syntactic constraints want the syntactically most-prominent element, Aux, to occur at the left or right edge, the focus

---

[9] Truckenbrodt's exact formulation is the following:

(i) ALIGN-FOC = ALIGN(Foc, R; P, R)
"Each focused constituent is right-aligned with a p-boundary"

A p-boundary is the boundary of a phonological phrase. The syntax of this constraint follows the conception of generalized alignment as introduced by McCarthy and Prince (1993).

[10] In more recent work (Samek-Lodovici, 2002), he takes prosodic representations as a base and uses Truckenbrodt's (1999) constraints on prosodic phrasing.

[11] The possibility of intonational breaks is the most reliable indicator for phonological phrase boundaries in German (see, among others, Kleinhenz (1994) for detailed discussion).

[12] IntP stands for intonation phrase, PhP for phonological phrase. We follow the theory of prosodic structure as developed by Selkirk (1984), Nespor and Vogel (1986), Truckenbrodt (1999), and others.

constraints want the focused element to occur at that edge. The rankings in (31) give higher priority to the focus constraints, but, as we will see below, the syntactic constraints still play a decisive role. The quite rare order 213, 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 these rankings for the dialects under discussion. We will first examine 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 the focus, we only specify this part of the input. The candidates are linearizations, i.e., PFs. We only look at the relevant parts of the candidates, i.e., the verb-cluster-internal linearizations. We will start with the competitions for narrow focus on each of the three verbs.

For narrow focus on V, FocL selects the orders 321 and 312, and the syntactic constraint MAPlrV^0 chooses between these two candidates, favoring order 312. Thus, the LF-PF mapping is still obeyed as much as possible. This pattern also shows up with the other two competitions:

(33)

StG: narrow focus on V	FocL	MAPlrV^0	MAPch
321 V Mod Aux		***!	
231 Mod V Aux	*!	**	*
123 Aux Mod V	*!		**
132 Aux V Mod	*!	*	*
☞ 312 V Aux Mod		**	*
213 Mod Aux V	*!	*	**

With narrow focus on Mod, the orders 231 and 213 are selected by FocL, and order 213 is preferred by MAPlrV^0:

(34)

StG: narrow focus on Mod	FocL	MAPlrV^0	MAPch
321 V Mod Aux	*!	***	
231 Mod V Aux		**!	*
123 Aux Mod V	*!		**
132 Aux V Mod	*!	*	*
312 V Aux Mod	*!	**	*
☞ 213 Mod Aux V		*	**

Narrow focus on Aux[13] favors 123 and 132,and then 123 emerges as optimal:

(35)

StG: narrow focus on Aux	FocL	MAPlrV^0	MAPch
321 V Mod Aux	*!	***	
231 Mod V Aux	*!	**	*
☞ 123 Aux Mod V			**
132 Aux V Mod		*!	*
312 V Aux Mod	*!	**	*
213 Mod Aux V	*!	*	**

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 321 for most foci, as in (36) for focus on V:[14]

(36)   *Wrong Standard German ranking!*

SG: narrow focus on V	FocR	FocL	MAPch	MAPlrV^0
☞ 321 V Mod Aux	*			***
231 Mod V Aux	*	*!	*	**
123 Aux Mod V		*	*!*	
132 Aux V Mod	*	*!	*	*
312 V Aux Mod	*		*!	**
213 Mod Aux V		*	*!*	*

The only candidates that are excluded by the focus placement constraints are those that have the focused verb at neither edge (231, 132). For the determination of the winner among those candidates that survive, the LF-PF mapping constraints are crucial. Hence, the unmarked order 321 has a high chance to win in many competitions. This is indeed the case. The ranking in (36) yields the following winners for competitions with different foci:

---

[13]   Narrow focus on the auxiliary has the effect of a verum focus interpretation, emphasizing that something is *indeed* the case, perhaps contrary to what has been claimed before.
[14]   'SG' stands for 'Standard German' in the tables below.

(37)   *Winners according to the ranking in* (36):

Focus on V:	$\longrightarrow$ order 321 = V Mod Aux
Focus on Mod:	$\longrightarrow$ order 132 = Aux V Mod
Focus on Aux:	$\longrightarrow$ order 321 = V Mod Aux
Focus on V+Mod:	$\longrightarrow$ order 321 = V Mod Aux
Focus on Mod+Aux:	$\longrightarrow$ order 321 = V Mod Aux
Focus on V+Mod:	$\longrightarrow$ order 321 = V Mod Aux
Focus on V+Mod+Aux:	$\longrightarrow$ order 321 = V Mod Aux
Focus on no verb:	$\longrightarrow$ order 321 = 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 312 which is missing in (37). 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 if the words are in 'canonical order', then focus can be *maximally projected*.[15] All three indicated foci are possible in (38), which has the Standard German default order 321:

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

Thus, (38) is ambiguous with respect to focus. Furthermore, focus usually *tends* to be projected. Narrow focus on 'SINGEN' in (38) requires a contrastive stress that is often stronger than the normal main stress. For narrow focus on V, order 312 is a better, because unambiguous, choice:

(39)   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 312, 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.

We express this tendency as another violable OT constraint that eval-

---

[15]   For a detailed discussion of the German focus facts see, for instance, Höhle 1982, who invented the notion of focus projection, Uhmann 1991, and Büring 1997. See also Cinque 1993 and Reinhart 1995 for related proposals. The central idea behind the notion of focus projection is that "the focus of a clause is a(ny) constituent containing the main stress of the intonational phrase" (Reinhart 1995, p. 62). Focus projection is maximally variable in default orders while it is usually blocked for elements in their non-default positions or at the non-recursive side of the syntactic structure.

uates the internal word order and stress pattern in a phonological phrase formed by a verb cluster with respect to its ideal, i.e., maximal, focus interpretation.

(40)   **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:

(41)   *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 (42):

(42)   Ideal focus with stress on V:
       a. [**V** Mod Aux] (= 321)
       b. [Mod **V**] Aux (= 231)
       c. [Aux Mod **V**] (= 123)
       d. Aux [**V** Mod] (= 132)
       e. [**V**] Aux Mod (= 312)
       f. Mod Aux [**V**] (= 213)

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 from 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 determine the 'unmarked' focus of a candidate and compare it with the focus specification given in the input.

(43)

SG: narrow focus on V	IF	MAPch	MAPlrV^0
321 V Mod Aux	*!		***
231 Mod V Aux	*!	*	**
123 Aux Mod V	*!	**	
132 Aux V Mod	*!	*	*
☞ 312 V Aux Mod		*	**
213 Mod Aux V		**!	*

Now, the LF-PF mapping constraints decide only between the orders 312 and 213, and MAPch chooses order 312. Thus, it is correctly predicted that order 312 is possible with stress on V in Standard German.

Narrow focus on Mod yields order 312, by nearly the same procedure. Now the orders 231 and 132 are competing, and here the lower ranked MAPlrV^0 makes the decision. Thus, we see that this constraint, which seemed to be active only in Swiss German, is also active in Standard German. It is one cause for the frequent acceptability of the order 132.

(44)

SG: narrow focus on Mod	IF	MAPch	MAPlrV^0
321 V Mod Aux	*!		***
231 Mod V Aux		*	**!
123 Aux Mod V	*!	**	
☞ 132 Aux V Mod		*	*
312 V Aux Mod	*!	*	**
213 Mod Aux V	*!	**	*

With narrow focus on Aux, we yield the default order 321 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 order 321. But it suddenly was the only possible option with stress on Aux.

(45)

SG: narrow focus on Aux	IF	MAPch	MAPlrV⁰
☞ 321 V Mod Aux			***
231 Mod V Aux		*!	**
123 Aux Mod V		*!*	
132 Aux V Mod		*!	*
312 V Aux Mod		*!	**
213 Mod Aux V		*!*	*

We can now discuss the narrow focus competitions for RP, where we will see that IF is also active, in addition to FOCUSRIGHT:

(46)

RP: narrow focus on V	FocR	MAPch	IF	MAPlrV⁰
321 V Mod Aux	*!		*	***
231 Mod V Aux	*!	*	*	**
123 Aux Mod V		**	*!	
132 Aux V Mod	*!	*	*	*
312 V Aux Mod	*!	*		**
☞ 213 Mod Aux V		**		*

The highest ranked constraint FocR selects the orders 123 and 213, which are equal at MAPch. IF now makes the difference and chooses order 213. The same happens with focus on Mod, where IF prefers order 132 over 312:

(47)

RP: narrow focus on Mod	FocR	MAPch	IF	MAPlrV⁰
321 V Mod Aux	*!		*	***
231 Mod V Aux	*!	*		**!
123 Aux Mod V	*!	**	*	
☞ 132 Aux V Mod		*		*
312 V Aux Mod		*	*!	**
213 Mod Aux V	*!	**	*	*

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

(48)

RP: narrow focus on Aux	FocR	MAPch	IF	MAPlrV^0
☞ 321 V Mod Aux				***
231 Mod V Aux		*!		**
123 Aux Mod V	*!	**		
132 Aux V Mod	*!	*		*
312 V Aux Mod	*!	*		**
213 Mod Aux V	*!	**		*

In (49), we compare what we have derived so far with the empirical findings in the two dialects that we are looking at.

(49)

		Stress on V	Stress on Mod	Stress on Aux
StG	Found:	123, 312	123, 213	123
	Predicted:	312	213	123
RP	Found:	321, 132, 213	321, 132, 312	321, 132
	Predicted:	213	132	321

A number of problems still need to be resolved:

1. Order 123 is missing in StG for stress on V and Mod.
2. Orders 321 and 132 are missing in RP for stress on V and Mod.
3. Order 312 is missing in RP for stress on Mod.
4. Order 132 as a second default pattern in RP is yet unexplained.

These issues are addressed in the next section.

5. Some problems

In both StG and RP, the default orders allow stress on any of the three verbs. As already mentioned above, a strategy that can often be observed for the indication of narrow contrastive focus in the default order is the use of heavier stress.[16] Heavy stress and word reordering seem to be two

---

[16] This might in fact only be necessary where the context does not provide enough information for disambiguation.

alternative disambiguation strategies with respect to focus. Use of heavy stress avoids the need to reorder the verb cluster while reordering avoids the use of heavy stress. To capture the word order effects of this strategy we would need to assume that the two dialects have a second (co-)grammar where FOCUSLEFT and FOCUSRIGHT, respectively, are ranked below the crucial syntactic constraints. We thus might assume a constraint tie between the relevant focus constraint and the highest syntactic mapping constraint.

This would capture the presumably correct intuition that the influence of focus placement constraints licenses some additional orders, but it does not override the defaults set by the syntactic mapping constraints. This might be an appropriate treatment, given that the information structure is a 'soft factor': its effects are often expressed in terms of markedness rather than grammaticality. However, in OT any possible ranking is a possible grammar. While the grammar of StG might have the proposed constraint tie, the grammars of other German dialects might not. The data we elicited for Bernese German (see Table I) show that this dialect could be a variant that shows no information structural influence. But we still need to figure out what happens if FOCUSLEFT is ranked unambiguously high.

We therefore want to explore another strategy that has been adapted by several OT researchers in accounting for the optionality of scrambling in the middle field of German clauses. The main idea is that the optionality is only apparent and that each optional order is ideal with respect to the expression of particular information structural properties in particular contexts.[17] According to this strategy, the orders not yet predicted according to the table in (49) conform to contexts that have not been considered until now.

The following subsection will check whether these missing orders arise under consideration of foci that are more complex than narrow focus. The subsequent subsections deal with more complex verb clusters (briefly) and the optional default order 132 in RP and other Standard German dialects. We will also briefly discuss the dialect of Upper Hessian which introduces another interesting complication.

---

[17]   Noteworthy examples for this strategy applied to German NP scrambling are Choi 1996, Müller 1999, Büring 2001. See also Costa 1998, Costa 2001, Samek-Lodovici 2001 for equivalent accounts of focus placement in Romance.

5.1. *Complex foci*

In section 4, we looked only at narrow focus competitions. The task of this subsection is to explore whether or not the model predicts attested orders for complex foci only under unambiguously high rank of the focus constraints in StG and RP. 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 123.

(50)   Complex focus competitions for StG:

a.

Focus on Mod+Aux	FocL	MAPlrV^0	MAPch
321 V Mod Aux	*!	***	
231 Mod V Aux	*!	**	*
☞ 123 Aux Mod V			**
132 Aux V Mod	*!	*	*
312 V Aux Mod	*!	**	*
213 Mod Aux V		*!	**

b.

Focus on V+Mod+Aux	FocL	MAPlrV^0	MAPch
321 V Mod Aux		*!**	
231 Mod V Aux		*!*	*
☞ 123 Aux Mod V			**
132 Aux V Mod		*!	*
312 V Aux Mod		*!*	*
213 Mod Aux V		*!	**

These two competitions already give us the two orders that have been missing for StG according to (49). However, a problem occurs with a complex focus on V+Mod. Here, the ungrammatical order 231 is wrongly predicted to win:

(50)   c.

Focus on V+Mod	FocL	MAPlrV^0	MAPch
321 V Mod Aux		***!	
☺☞ 231 Mod V Aux		**	*
123 Aux Mod V	*!		**
132 Aux V Mod	*!	*	*
312 V Aux Mod	*!	**	*
213 Mod Aux V	*!	*	**

The orders 321 and 231 are not possible in StG at all. Our model seems to be too liberal up to now. Syntactic mapping can be disrespected for the purpose of focus expression but only to a certain extent. What is the correct description of this 'extent'? What these orders 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 special version of MAPlrV^0 for functional verbs (we restrict 'functional' to finiteness and tense here, i.e., those elements traditionally classified as 'INFL' within Government and Binding theory; Chomsky 1981):

(51)   **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 (51) 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, p. 366):

(52) 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 the perfect auxiliary (Schönenberger 1995, p. 366):

(53) 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 (53) and (52) might be due to the difference between infinitival and participial verb forms, with the participle being required to precede its governing verb more urgently than the infinitive is. In the 3-verb clusters that we examined, a final auxiliary is impossible. Schönenberger (1995) reports the same for clusters with four verbs. We assume that the complexity of the verb cluster triggers the prohibition of verb-final functional verbs. The method of constraint conjunction is a way to reflect cumulative effects in OT:

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

This constraint is ranked high and thus blocks the orders 321 and 231 in 3-verb clusters:

(55)   Violations of MAPlr(V_{func}^0) and MAPlr(V_{func}^0)^2:

	MAPlr(V_{func}^0)^2	MAPlr(V_{func}^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 functional verb in our verb clusters in the sense of the constraint is the auxiliary. In determining the constraint violations, we consider two pairs, (Aux, Mod) and (Aux, V), and get at most two violations (321, 231). The ranking for StG is as follows:

(56)   StG ranking (revised):
        MAPlr(V_{func}^0)^2 >> FocL >> MAPlr($_{func}^0$) >> MAPch

Order 123 now wins the competition for focus on V+Mod, a prediction that appears to be correct:

(57)

Focus on V+Mod	$MAPlr(V_{func}^0)^2$	FocL	$MAPlrV^0$	MAPch
321 V Mod Aux	*!		***	
231 Mod V Aux	*!		**	*
☞ 123 Aux Mod V		*		**
132 Aux V Mod		*	*!	*
312 V Aux Mod		*	*!*	*
213 Mod Aux V		*	*!	**

The findings for StG are now completely reconstructed:

(58)   Predictions for StG:

		Stress on V	Stress on Mod	Stress on Aux
StG	Found:	123, 312	123, 213	123
	Predicted:	123, 312	123, 213	123

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

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

We observe in Standard German that the larger a verb cluster is, the stronger is the *pressure to give up the default order*:

(59) a.   weil   sie   es   sehen   wird
          because she   it   see   will
          Order: 21

    b.   weil   sie   es   sehen   können   wird
          because she   it   see   can   will
          Order: 321

    c. ? weil   er   sie   es   sehen   lassen   können   wird
          because he   her   it   see   let   can   will
          Order: 4321

d.?*weil   er   sie die Kinder   spielen sehen lassen
   because he   her the children play   see   let

können wird
can   will
Order: 54321

This can be directly mirrored in our system of constraints by the increasing violations of $MAPlr(V^0_{func})$ that go along with larger verbal complexes in the Standard German default order. Swiss German dialects do not show such an effect because their default order is already in accord with $MAPlr(V^0_{func})$. (60) shows improved Standard German versions of (59-c, d):

(60) c′. weil er sie es wird sehen lassen können
          1432

     d′.?weil er sie die Kinder wird können spielen sehen lassen
          12543

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 $MAPlr(V^0_{func})^4$.[18]

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* 'have' have the strongest tendency to occur in verb-cluster initial position; for the future auxiliary *werden* 'become' and most finite modal verbs this is optional, and perfect auxiliaries derived from *sein* '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*

Section 5.1 dealt with the optionality of word orders by treating it as what Müller (1999b, 2000) calls 'pseudo-optionality', namely, that the optional

---

[18]   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 accumulate 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. Its violations are more important, no matter how many violations of B we have. In order to establish cumulative effects, we need a constraint B^n that is ranked higher than A and that is violated if B is violated at least n times. As long as B^n is ranked immediately on top of B (or lower, which is usually excluded by convention), the system behaves as if B^n was not there at all.

orders are in fact the *only* winners of *particular* competitions which are
defined by particular information structural specifications. This 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, namely the orders 321 and 132. Here, we would like to see
two winners for a competition within a neutral context. However, our
LF-PF mapping constraints distinguish between the two orders we are
talking about. Because the two candidates have different *constraint vio-
lation 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' option-
ality in OT is assuming that those constraints where the two optional
candidates differ are not ranked with respect to each other. They are *tied*.
We will use this strategy here. 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:[19]

(61) 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 by the following two
tableaux, which show how the two default orders win in a neutral context:

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

	MAPch	MAPlr(V_{func})^2	MAPlr(V^0)
☞ 321: V Mod Aux		*	***
231: Mod V Aux	*!	*	**
123: Aux Mod V	*!*		
132: Aux V Mod	*!		*
312: V Aux Mod	*!		**
213: Mod Aux V	*!*		*

---

[19] Several versions of constraint ties have been proposed in the literature. For an overview,
see Müller 1999b, 2000. A global tie is not really a tie but actually a notational convention for
the abbreviation of two existing co-grammars without a tie.

b. Second ranking:

	MAPlr(V_{func})^2	MAPch	MAPlr(V^0)
321: V Mod Aux	*!		***
231: Mod V Aux	*!	*	**
123: Aux Mod V		**!	
☞ 132: Aux V Mod		*	*
312: V Aux Mod		*	**!
213: Mod Aux V		**!	*

We see in (62-b) why order 132 is an optimal candidate as soon as order 321 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:

(63)   Outcomes in RP, winners only:

stress on V3   /F3/ $\longrightarrow$ [213]
                /F32/ $\longrightarrow$ [132]
                /F321/ $\longrightarrow$ [321], [132]

stress on V2   /F2/ $\longrightarrow$ [132]
                /F21/ $\longrightarrow$ [321], [312]

stress on V1   /F1/ $\longrightarrow$ [321]

no focus $\longrightarrow$ [321], [132]

These outcomes now nearly match our empirical findings, as illustrated in (64):

(64)

		Stress on V	Stress on Mod	Stress on Aux
RP	Found:	321, 132, 213	321, 132, 312	321, 132
	Predicted:	321, 132, 213	321, 132, 312	321

What is still missing is the default order 132 for stress on Aux. If we want to include this as well, we need to assume a third co-grammar, where MAPlr(V_{func})^0 is ranked even higher than FocR. In this case, the only two orders that satisfy FocR for focus on Aux, 321 and 231, are excluded from the start, and the syntactically least marked candidate, order 132, is the winner, as it also is for most other competitions.

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

(65)   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 sub-ranking "FocR $\gg$ MAPch" and is ranked either below the two constraints, on top of, or between them while the relative ranking of FocR and MAPch remains constant.

5.4. *Upper Hessian*

We will now take a look at a dialect that displays further syntactic and prosodic restrictions on auxiliaries. In Upper Hessian sometimes only the 132 pattern is acceptable. This is illustrated in (66):[20]

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

    b. . . . dass sie es ihn hat/*HAT singen hören
           that she it him has   sing   hear

The verb *haben* 'have' in its perfect auxiliary usage cannot occur in final position and cannot be stressed either. It seems to be characteristic of this dialect that some function words occur only in weak forms. The personal pronouns have the same 'defect'. Selkirk (1996) shows that English function words can occur both in a phonologically strong and a weak form but that the weak form cannot occur in clause-final position. In the following examples, the clause-final verbs *can* and *is* cannot be reduced while this is possible with the first occurrence of *can* in (67-a) (the vowel is reduced to schwa) and the second occurrence of *is* in (67-b) (reduced to *'s*):

(67) a. I can eat more than Sara *cán*
     b. Wherever Ray *ís*, he's having a good time

Selkirk assumes that weak function words do not project prosodic words. She further assumes a highly ranked constraint that requires the right edge of a phonological phrase to be aligned with the right edge of a prosodic word in English. This derives the observed restriction. Upper Hessian seems to show the same behavior – with the additional complication that no strong form of the auxiliary is available in the example in (66). Without

[20]   Upper Hessian is spoken in a region which is, roughly, between 30 and 100 kilometres north of Frankfurt/Main.

going into more detail here, we assume the constraint in (68) to integrate this phenomenon:

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

Whether (particular) function words occur in weak forms only or also in strong forms is obviously a (possibly parameterized) lexical difference among German dialects. Most other Standard German dialects do not seem to lack strong function words – all auxiliaries 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 132 or 312 are possible. The constraint ranking that we assume for Upper Hessian is the following one:

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

The outcomes are listed in (70):

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

	stress on V3	/F3/	⟶	[312]
		/F32/	⟶	[132]
		/F321/	⟶	[132]
	stress on V2	/F2/	⟶	[132]
		/F21/	⟶	[312]
	stress on V1	/F1/	⟶	impossible[21]
	no focus	/NoF/	⟶	[D132]

We see that the focus constraints are ranked quite low. They have only a marginal influence.[22]

---

[21]   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 extremely 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.

[22]   This kind of micro-variation within Standard German dialects can also be found within the Swiss German dialect family: For our *Bernese Swiss German* informant, only the default order 123 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.

5.5. *Focus on the object in RP*

A problem that requires closer examination is that, under the assumed ranking for RP, we would predict that focused objects tend to be right dislocated. This is, however, not the case. On the contrary, direct objects strictly have to precede the predicative verb. The same is true of adverbs and other constituents. In fact, it seems as if FOCUSRIGHT took effect only in verbal complexes.

A less appealing strategy for excluding such unwanted effects of FOCUSRIGHT would be the assumption of a number of special constraints that take care of all contexts where FOCUSRIGHT has no effect. One could, for instance, assume 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:

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

MAPchΘ is ranked higher than FOCUSRIGHT while the simple constraint MAPch is ranked lower. We thus get the following ranking:

(72)   Final ranking for RP:
        MAPchΘ $\gg$ (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, 321 and 132. As such, this kind of strategy is not implausible.[23]

A plausible alternative would be a treatment in terms of morphology

---

[23]   Independent motivation might come from the fact that even verb clusters are sensitive to this restriction. If verb 1 is a causative verb, which presumably assigns a thematic role to its VP complement, then Standard German allows for only the canonical order 321, strictly obeying MAPch, as shown in (i).

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

rather than thematic roles. Objects receive case from verbs. The force to obey MAPch more strictly might be correlated with morphological dependency. This can be correlated with the phenomenon of 'Infinitivus pro participio' in German. For many German speakers and dialects, the alternation between the orders 321 and 132 goes hand in hand with a shift in the morphology of verb 2, as in our example with a perception verb (cf. Schmid 2000, p. 344):

(73) a. . . . dass sie ihn das Lied hat   singen **hören**
            that she him the song has   sing   hear-INF

    b. . . . dass sie ihn das Lied singen **gehört**   hat
            that she him the song sing   heard-PART has

In order 321 (73-b), the perception verb preferably occurs in participial form while in order 132 (73-a) the infinitive is preferred. A constraint 'MAPch(morph)' could be formulated that requires a complement to occur before its head if it bears morphology that indicates dependency, like case for NPs or the participial form for verbs. Perhaps both MAPch(morph) and MAPch(θ) are part of the constraint hierarchy.

An interesting observation in this respect is that adjuncts, being morphologically independent, are easier to extrapose than arguments in Standard German:

(74) a.*Ich habe gestern   gelesen das Buch
         I   have yesterday read   the book

    b.?Ich habe das Buch gelesen gestern
         I   have the book read   yesterday

But note also that the extraposed constituent must not bear the main stress in (74-b). Focus on *gestern* would require it to occur left of the verb, ideally adjacent.

Our solution so far implies that the focus always occurs on the very right of the clause in RP although this appears to be an *exception* which can only be observed in verbal complexes. The model takes the exception to be the rule.

An alternative approach would be to try to restrict FOCUSRIGHT to verbal complexes from the very beginning. How can this be achieved? Verbal complexes, though they are syntactically construed, could be treated on a par with compound expressions at PF,[24] in particular, with respect to stress

---

[24] This intuition is also guiding the theory of 'reanalysis' put forward by Haegeman and van Riemsdijk (1986): a sequence of two adjacent, but distinct verbal heads is reanalysed as a single head consisting of two verbs.

assignment. Let us assume that in RP the *compound stress rule* (CSR) requires stress to be on the rightmost constituent. Usually, compounds are lexically fixed, and main stress is on their most embedded constituent. The CSR therefore is a default rule that applies very rarely, namely, in cases where either the deepest constituent cannot be determined or where narrow focus demands main stress to be on a different constituent. Verbal complexes have a syntactic source and are for this reason syntactically more flexible than lexical compounds. Given high priority, the nuclear stress rule can now induce a change in the word order of a verbal complex – preferring a narrowly focused constituent at the right edge.

What we call "FocusRight" could thus be an instance of the compound stress rule in RP. Some 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', 'union gate' or 'outer gate')

There might be a general tendency to favor the right edge of a prosodic domain for the main accent in Northwestern Low German dialects. This is an issue that needs further investigation.[25]

In Swiss German dialects, one can also observe an interesting departure from Standard German metrical preferences. Abbreviations like 'BMW' and 'EU', which are pronounced letter-wise (i.e., like 'USA' and

---

[25] One anonymous reviewer remarks that geographical names might be bad examples because they might be stored as simplexes by speakers in those areas where they are used frequently. We agree that this is a possible explanation for *Niedersáchsen*' as word stress usually goes to the penultimate syllable in German. However, we contacted some Low German speakers from more Eastern areas of Lower Saxony, and none of them favored the Bremen version of pronouncing *Niedersáchsen*. This word should be just as common to those speakers. Commonality cannot be the only responsible factor. Likewise, the word stress rule would still predict *Síelwall*, not *Sielwáll*. The reviewer further suggests that a good test would be how Bremeners would pronounce a less known 'X-Tor' in another city. They certainly do not pronounce it differently than other standard German speakers. The reason for this expectation is that compound stress usually falls on the most deeply embedded element. Common geographical names might be somewhat 'intermediate' words in that their compositionality has not been forgotten, but the embedding relation is no longer relevant for stress assignment. In such a case, a default stress assignment rule for compounds might apply, and this rule might differ from Standard German in the relevant dialect. But note that this is highly speculative.

not word-wise like 'NATO', or'AIDS'), have the main stress on the first letter in Swiss German dialects but on the final one in Standard German:

(75) a. ÉU (Swiss German), EÚ (Standard German)
     b. B́MW (Swiss German), BMẂ (Standard German)

Here, we find a preference for the left edge of a prosodic domain where Standard German prefers the right edge – again a parallelism to the edge preference that we found for focused verbs in StG verbal complexes. Both of these observations could have a connection to the data we discussed in this article. The ultimate source of the variation that we found in StG and RP, in particular, the surprising occurrence of the 213 order, might be due to a departure from Standard German metrical and prosodic phonology which, in the dialects at hand, takes syntactic 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 are syntactic, phonological, and information structural. Whether FocR and FocL are actually information structural, rather than phonological, is an open issue, however.

Our results suggest a view on dialectal variation in syntax that relates it to variation in prosodic and metrical phonology and morphology. Syntactic variation at a micro-syntactic level occurs if such non-syntactic factors overrule syntactic linearization constraints. Our work thus supports a view on the syntax-phonology interface where both interact at a level playing field, contrary to the feeding relation from syntax to phonology that is standardly assumed in generative syntax.[26]

Zwart (1996) addresses the syntax of verb clusters in 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 ≫ MAPhc" holds for all the varieties under discussion, there are a number of other factors that intervene in such a way that this 'underlying' parameterization is very hard to recover.

Most of the dialects displayed in Table I have not been discussed here.

---

[26] This view of the interface has also been argued for more recently by Samek-Lodovici (2002).

This is mainly due to reasons of space. The typology of the constraint system that we developed is rich enough to describe these and many more dialects. In many cases of optional orders we have to make use of constraint ties. The set of dialects we have collected is rather accidental; future work will have to explore in much more detail how individual dialects work and include many more dialects. As of this writing, the number of (not only) generative explorations into the syntax of German dialects in Germany has been extremely small, and there is little hope that this situation will change soon. But we hope to have shown that efforts in this direction are worth pursuing, especially from the perspective of the theory of grammar.

References

Büring Daniel: 1997, *The Meaning of Topic and Focus – The 59th Street Bridge Accent*, Routledge, London.

Büring Daniel: 2001, 'Let's Phrase It! – Focus, Word Order, and Prosodic Phrasing in German Double Object Constructions', in G. Müller and W. Sternefeld (eds.), *Competition in Syntax*, Mouton de Gruyter, Berlin & New York, pp. 101–137.

Choi, Hye-Won: 1996, *Optimizing Structure in Context: Scrambling and Information Structure*, Ph.D. dissertation, Stanford University.

Chomsky, Noam: 1981, *Lectures on Government and Binding*, Foris, Dordrecht.

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

Cinque, Guglielmo: 1993, 'A Null Theory of Phrase and Compound Stress', *Linguistic Inquiry* **24**, 239–298.

Costa, João: 1998, *Word Order Variation*, Ph.D. dissertation, HIL, University of Leiden.

Costa, João: 2001, 'The Emergence of Unmarked Word Order', in G. Legendre, J. Grimshaw, and S. Vikner (eds.), *Optimality Theoretic Syntax*, MIT Press, Cambridge.

Den Besten, Hans and Jerold A. Edmondson: 1983, 'The Verbal Complex in Continental West Germanic', in W. Abraham (ed.), *On the Formal Syntax of the Westgermania*, John Benjamins, Amsterdam, pp. 155–216.

Donaldson, Bruce C.: 1993, *A Grammar of Afrikaans*, Mouton de Gruyter, Berlin.

Dowty, David R.: 'Toward a Minimalist Theory of Syntactic Structure', in H. Bunt and A. van Horck (eds.), *Discontinuous Constituency*, Mouton de Gruyter, Berlin.

Grimshaw, Jane: 1991, 'Extended Projection', ms., Brandeis University.

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

Höhle, Tilman: 1982, 'Explikationen für "normale Betonung" und "normale Wortstellung"', in W. Abraham (ed.), *Satzglieder im Deutschen*, Narr, Tübingen, pp. 75–153.

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

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

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

Kleinhenz, Ursula: 1994, *Focus and Phrasing in German*, Technical Report 52, IBM Deutschland, Wiss. Zentrum, Inst. Für Wissensbasierte Systeme, Heidelberg.

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

Legendre, Géraldine, Paul, Smolensky, and Colin Wilson: 1998, 'When is Less More? Faithfulness and Minimal Links in WH-Chains', in P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky (eds.), *Is the Best Good Enough? Optimality and Competition in Syntax*, MIT Press, Cambridge, pp. 249–289.

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

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

McCarthy, John J. and Alan Prince: 1993, 'Generalized Alignment', in G. Booj and J. van Marle (eds.), *Yearbook of Morphology 1993*, Kluwer, Dordrecht, pp. 79–153.

Müller, Gereon: 1999a, 'Optimality, Markedness, and Word Order in German', *Linguistics* **37**, 777–818.

Müller, Gereon: 1999b, 'Optionality in Optimality Theoretic Syntax', *Glot International* **4**, 3–8.

Müller, Gereon: 2000, *Elemente der optimalitätstheoretischen Syntax*, Stauffenburg, Tübingen.

Nespor, Marina and Irene Vogel: 1986, *Prosodic Phonology*, Foris, Dordrecht.

Pesetsky, David: 1997, 'Optimality Theory and Syntax: Movement and Pronunciation', in D. Archangeli and D. T. Langendoen (eds.), *Optimality Theory, An Overview*, Blackwell, Malden MA and Oxford, pp. 134–170.

Pesetsky, David: 1998, 'Some Optimality Principles of Sentence Pronunciation', in P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky (eds.), *Is the Best Good Enough? Optimality ad Competition in Syntax*, MIT Press, Cambridge, pp. 337–383.

Reinhart, Tanya: 1995, 'Interface Strategies', OTS Working Papers, University of Utrecht.

Rizzi, Luigi: 1990, *Relativized Minimality*, MIT Press, Cambridge.

Robbers, Karin: 1997, *Non-Finite Verbal Complements in Afrikaans, A Comparative Approach*, Holland Academic Graphics, Den Haag.

Samek-Lodovici, Vieri: 2001, 'Crosslinguistic Typologies in Optimality Theory', in G. Legendre, J. Grimshaw, and S. Vikner (eds.), *Optimality Theoretic Syntax*, MIT Press, Cambridge, pp. 315–353.

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

Schmid, Tanja: 2000, 'Die Ersatzinfinitivkonstruktion im Deutschen', *Linguistische Berichte* **183**, 325–351.

Schmid, Tanja 2002, *West Germanic IPP-Constructions, An Optimality Theoretic Approach*, Ph.D. dissertation, Institut für Linguistik, Universität Stuttgart.

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

Selkirk, Elisabeth: 1984, *Phonology and Syntax, The Relation between Sound and Structure*, MIT Press, Cambridge.

Selkirk, Elisabeth: 1996, 'The Prosodic Structure of Function Words', in J. Morgan and K. Demuth (eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, Lawrence Erlbaum Associates, Mahwah, New Jersey.

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

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

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

Vogel, Ralf: 2003, 'Dialectal Variation in German 3-Verb Clusters, Looking for the Best Analysis', in A. Mahajan (ed.), *Head Movement and Syntactic Theory*, vol. 10 of *UCLA/Potsdam Working Papers in Linguistics*, Los Angeles, UCLA.

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

Wurmbrand, Susi: 2001, '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, Jan-Wouter: 1996, 'Verb Clusters in Continental West Germanic Dialects', in J. Black and V. Motapanyane (eds.), *Microparametric Syntax and Dialect Variation*, John Benjamins, Amsterdam, pp. 229–258.