Optimality, Markedness, and Word Order in German

Gereon Müller (Universität Tübingen)[*]

Abstract

The main goal of this article is to develop an approach to free word order structures in German that reconciles the findings of two different lines of research: competition-based models that center around the interaction of factors like definiteness, animacy, Case, and focus on the one hand, and optimality theoretic syntax with its violable and ranked constraints on the other. The analysis relies on the existence of a syntactic scrambling operation. The major claims are: (i) Scrambling is triggered by a subhierarchy of violable and ranked linearization constraints. (ii) Optimality under at least one linearization constraint results in grammaticality, optimality under the whole subhierarchy results in an unmarked structure (unmarked structures do not correspond to D-structures, as is often assumed). (iii) The distinction between subhierarchies and matrix hierarchies in optimality theory parallels the traditional distinction between weak and strong rules. It accounts for the difference between weak pronoun fronting to a Wackernagel position, which results in a fixed order, and scrambling to VP, which does not. (iv) Language-specific variation with respect to scrambling options is due to constraint reranking.

1. Introduction

Recent research has accumulated substantial evidence in support of a scrambling operation (cf. Ross (1967)) as the basis of free word order structures in German (and other languages), and many of its properties have been unravelled. However, notwithstanding the progress that has been made, there are several problems that arise with free word order structures which have not received a convincing solution so far. I will address four of these in what follows. First, given economy constraints that block unforced movement (cf. Chomsky (1995)), scrambling cannot strictly speaking be an optional movement operation; rather, a trigger must be identified that forces scrambling. It is, however, not quite clear what this trigger might look like.[1] Second, the issue of markedness arises: Clause-internal word order in scrambling languages often exhibits degrees of markedness, rather than complete wellformedness or illformedness, and this fact is still in need of an explanation. Third, it must be clarified why a free word order language like German does in fact exhibit a fixed order domain in the Mittelfeld, viz., the Wackernagel position, to which weak pronouns are fronted. Finally, the question arises of how language-specific variation with respect to scrambling options is to be accounted for.

The goal of this article is to present solutions to these problems. To this end, I will try to combine insights of two different lines of research. First, it has repeatedly been suggested (especially in work with an orientation that is primarily empirical) that variable word order in the Mittelfeld of German clauses should be understood as the result of the interaction of factors which center around notions like definiteness, animacy, focus, Case, etc.; see Lenerz (1977, 27 & 62f), Hoberg (1981, 62), Lötscher (1981, 58-59), Uszkoreit (1984, 174ff; 1986, 896-899), Lerot (1985, 141), Reis (1986, 27ff), Jacobs (1988, 17ff), Stechow & Sternefeld (1988, 455), Siewierska (1993, 830ff), Dietrich (1994, 41), and

---

[1]Suggestions that have been advanced in the literature include (a) semantic factors (cf. Diesing (1992), de Hoop (1992), Meinunger (1995), Büring (1997), Kidwai (1997), Neeleman & Reinhart (1997), and much related work); (b) Case reasons (cf., e.g., van den Wyngaerd (1989), Zwart (1993)); (c) purely formal features (cf., e.g., Sauerland (1997), Müller (1998), Grewendorf & Sabel (1999)); and (d) combinations thereof (cf. some of the above). For space reasons, I cannot discuss these approaches here in any detail. However, regarding (a), it seems to me that while there can be little doubt that different S-structure orders of scope-bearing elements yield different (preferred) patterns of interpretation, it is both problematic and unnecessary to account for all attested scrambling operations by invoking semantic factors as a trigger for movement – it is problematic because non-scope-bearing items can be scrambled (e.g., proper names), and because even scrambling of scope-bearing items can sometimes lead to a non-distinct semantic interpretation (most obviously if only non-scope-bearing elements are crossed, but also in other cases); and it is unnecessary because the semantic effects of scrambling can usually be accounted for by a model of semantic interpretation where S-structure representations are highly relevant for relative scope relations; cf. Jackendoff (1972), Kroch (1974), among many others. (In line with this, topicalization often has the same effects on relative scope as scrambling). Turning next to (b), it seems clear that an analysis of scrambling as Case-driven movement faces severe problems of under-generation – for one thing, categories other than NPs can be scrambled in German (PPs, CPs, and even, to some extent, APs and VPs); and for another, whereas Case-driven movement typically targets fixed positions (resulting in an invariant NP order), NP scrambling in German creates permutation effects (all six orders are basically possible with three NP arguments). Finally (c), assuming a scrambling feature without inherent content may lead to a descriptively adequate analysis, but still has the flavour of an ad hoc-assumption, motivated mainly by the existence of economy constraints.

Primus (1994, 40-48), among others. It is often held that these factors are by themselves "weak" and not equally important. Competition-based models of this type have proven successful from an empirical point of view, but they have always raised a conceptual problem: The interplay of weak factors of varying importance did not otherwise appear to play a role in grammar. Hence, to the extent that competition-based models of free word order were theoretically implemented, the resulting system was very different from that which was otherwise employed in grammatical theory.[2]

However, with optimality theory (cf. Prince & Smolensky (1993)), a second line of research has recently come into existence that views grammatical constraints as systematically violable and ranked. For this reason, a unified approach does not seem to be impossible anymore. Thus, in what follows I would like to suggest that an approach to scrambling that reconciles some of the leading ideas of competition-based models with the basic tenets of optimality theory can provide solutions to the problems noted above. However, we will see that there is a price to be paid: It is a crucial optimality theoretic assumption that a suboptimal candidate is invariably ill formed, but in order to maintain this assumption in light of the empirical evidence from word order variation in the German Mittelfeld, a modification of standard optimality theory must be made somewhere. I will suggest that the constraint hierarchy must be split up into a matrix hierarchy and a subhierarcy. This modification will accomodate the scrambling evidence in German and leave previous reasonings in optimality theoretic syntax intact.

2. Basic Assumptions

2.1. Clause Structure and Scrambling

I will presuppose the following clause structure of German:

(1) $[_{CP} - C [_{TP} - [_{\pi P} - [_{VP} \text{ Adj } [_{VP} \text{ SUBJ } [_{V'} \text{ DO } [_{V'} \text{ IO } [_{V'} \text{ OBL V }]]]]] \pi] T]]$

Here, SpecC is the landing site for <u>wh</u>-movement. SpecT(ense) is the landing site for subject raising (cf. Chomsky (1995)); this movement is optional in German (cf. Grewendorf (1989) and Diesing (1992), among others). Specπ is the position that I assume to be the landing site for weak pronoun movement, i.e., the 'Wackernagel' position (also cf. Johnson's (1991) μP); evidence for this projection will be presented in section 5. Next, (certain) adverbs are base-adjoined to VP. Finally, the D-structure order of subject (SUBJ), direct object (DO), and indirect object (IO) is always (i.e., with all types of verbs) as in (1). In particular, the direct object (typically, the THEME) must asymmetrically c-command the indirect object (typically, the GOAL) in the base (cf. Larson (1988)); I will show that this assumption is supported by evidence from anaphoric binding and weak pronoun fronting.

---

[2]It is in principle possible to assume an interplay of weak factors and at the same time do without an explicit notion of competition. See, e.g., the coalition model proposed in Dietrich (1994; 1999), which assigns numerical precedence indices to constituents that are determined by factors like definiteness, agentivity, and focus. On this view, a "good" word order is one in which precedence indices decrease from left to right. Still, the conceptual problem remains: This mechanism is not needed in other domains of grammar.

Closest to the base position of V are oblique arguments; these are realized either by NPs bearing lexical Case (e.g., genitive), or by PPs. The word order determined by (1) can be changed by movement to a specifier position (SpecC, SpecT, Specπ), or by scrambling.

Throughout this article, I adopt the scrambling approach to free word order phenomena in German, and not an approach in terms of base-generation (as pursued by Haider (1989), Bayer & Kornfilt (1994), and Fanselow (1997), among others). It is impossible to review the evidence in support of a scrambling operation in any detail here; so I will confine myself to pointing out one phenomenon that is quite suggestive, and notoriously difficult to account for under a base-generation approach. As observed in Koster (1987), Fanselow (1988), and Müller (1995), scrambling does not only affect co-arguments; rather, the target position of extraction from PP (postposition stranding) and of extraction from NP can also be Mittelfeld-internal:

(2) a. daß der Fritz [$_{NP}$ da]$_1$ wahrscheinlich nicht [$_{PP}$ t$_1$ mit] gerechnet hat
         that ART Fritz   there probably   not   with   counted has
     b. daß Maria [$_{PP}$ über die Liebe]$_1$ gestern   abend [$_{NP}$ ein Buch t$_1$]
         that Maria   about the love   yesterday evening   a   book
         gelesen hat
         read   has

Given that there is strong evidence that extractions from PP and NP in German involve the creation of a trace, the examples in (2) must involve a TP-internal movement operation, i.e., scrambling. But if (2) is derived by scrambling, there is every reason to also hold this operation responsible for the creation of other word order variation effects, such as the permutation of two argument NPs. Accordingly, a language like English, which does not exhibit free word order phenomena, also lacks movement operations as in (2).

Furthermore, I will assume that scrambling in German can only be adjunction to VP (not to πP, TP, or CP, and not to NP or PP either);[3] and that it is a property of adjunction operations to be iterable (in contrast to specifier movements – note that this is in conflict with the view in Chomsky (1995)). The moved item is typically an NP, PP, or CP, with AP and VP being marginal options that will be ignored in what follows.

Finally, the question has been much discussed whether scrambling is to be viewed as an A- or an A-bar movement (-like) operation (see, e.g., the contributions in Corver & van Riemsdijk (1994), and references cited there). Here, I will not take a firm stand on this issue. However, all that follows is directly compatible with the A-bar movement approach, whereas an A-movement approach may necessitate additional assumptions.

2.2. Binding of Anaphors

A first consequence of the structure in (1) and the existence of scrambling as adjunction to VP concerns binding theory. Suppose that principle A of the binding theory is formulated

---

[3]Evidence against adjunction to TP will be presented in section 5. Other free word order languages may have more landing sites for scrambling. Thus, it seems that Russian permits scrambling to TP, CP, and NP (but not to PP); see Müller (1995) for discussion.

in a purely structural way, as requiring c-command of the anaphor by a co-indexed item in an A-position (cf. Chomsky (1981)). Then we predict that binding of an IO anaphor by a preceding DO should be possible in German, whereas binding of a DO anaphor by a preceding IO should not be. The reason is that the order IO $\succ$ DO can only be derived by scrambling under present assumptions. However, if scrambling is an A-bar movement operation, it cannot create new A-binding options.[4] As shown in (3) for reciprocals and in (4) for reflexive pronouns, these predictions are borne out (cf. Grewendorf (1988) for the original observation, and also Webelhuth (1992) and Müller & Sternefeld (1994)).[5]

(3)   a.   daß man die Gäste$_1$ einander$_1$ vorstellte
            that one the guests$_{acc}$ each other$_{dat}$ introduced

    b. *daß man den Gästen$_1$ einander$_1$ t$_1$ vorstellte
            that one the guests$_{dat}$ each other$_{acc}$ introduced

    c.   daß ich diese Leute$_1$ einander$_1$ von Herzen gönne
            that I these people$_{acc}$ each other$_{dat}$ from heart not begrudge

    d. *daß ich diesen Leuten$_1$ einander$_1$ t$_1$ von Herzen gönne
            that I these people$_{dat}$ each other$_{acc}$ from heart not begrudge

    e.   daß die UNO die Feinde$_1$ einander$_1$ nicht ausliefern darf
            that the UNO the enemies$_{acc}$ each other$_{dat}$ not extradite may

    f. *daß die UNO den Feinden$_1$ einander$_1$ t$_1$ nicht ausliefern darf
            that the UNO the enemies$_{dat}$ each other$_{acc}$ not extradite may

(4)   a.   daß der Arzt den Patienten$_1$ sich$_1$ im Spiegel zeigte
            that the doctor the patient$_{acc}$ himself$_{dat}$ in the mirror showed

    b. ?*daß der Arzt dem Patienten$_1$ sich$_1$ t$_1$ im Spiegel zeigte
            that the doctor the patient$_{dat}$ himself$_{acc}$ in the mirror showed

---

[4]If scrambling is A-movement, we can end up with the same consequence if we adopt Rizzi's (1986) constraint on A-chain formation – on this view, A-movement of the IO across the DO would result in an illicit crossover configuration.

[5]There is some variability in the judgements of the reflexive pronoun cases in (4); however, the general tendency is clear enough. Also note that it does not help to assume (as is sometimes done) that binding of a DO reciprocal is in general impossible in the presence of a dative NP, and hence, that the ill-formed data in (3) do not tell us anything about the argument order at D-structure. As shown in (i-ab), a DO reciprocal can be bound by the subject NP despite the presence of the IO.

(i)   a.   daß die Gastgeber$_1$ dem Besucher$_2$ einander$_1$ t$_2$ vorstellten
            that the hosts$_{nom}$ the visitor$_{dat}$ each other$_{acc}$ introduced

    b.   daß die Gastgeber$_1$ einander$_1$ dem Besucher$_2$ vorstellten
            that the hosts$_{nom}$ each other$_{acc}$ the visitor$_{dat}$ introduced

Finally, it has sometimes been claimed that an IO can in fact bind into a DO and thereby license a DO-internal anaphor. While examples of this type are indeed radical improvements over the ungrammatical sentences in (3) and (4), these NP-internal anaphors in German behave in every respect like their counterparts in English, for which Reinhart & Reuland (1993, 681-685) have shown that they are logophoric and do not obey structural binding conditions.

  c. daß der Vater die Kinder$_1$ sich$_1$ (selbst) überließ
    that the father the children$_{acc}$ themselves left
  d. *daß der Vater den Kindern$_1$ sich$_1$ (selbst) t$_1$ überließ
    that the father the children$_{dat}$ themselves   left

Furthermore, the prediction is that an anaphor in an oblique PP argument can be A-bound by both a DO and an IO; this is the case (cf. Grewendorf (1988)).

(5) a. daß man die Gäste$_1$ [$_{PP}$ an einander$_1$] verwies
    that one the guests$_{acc}$ to each other referred
  b. daß man der Frau$_1$ die Augen t$_1$ [$_{PP}$ über sich$_1$] öffnete
    that one the woman$_{dat}$ the eyes about herself opened

On the basis of these background assumptions, let me now turn to the problems noted above, and begin with the issue of markedness.

3. Markedness

3.1. The Empirical Determination of Markedness

The grammaticality of a given sentence is usually verified by invoking the intuitions of native speakers. An obvious question that arises with a graded notion like that of relative markedness is whether it can or must be determined in another way. Several suggestions have been made in the literature. A straightforward way of determining relative markedness relies on the concept of statistical frequency (in texts, discourses, etc.). This means of measuring markedness is now commonly believed to be highly unreliable (cf, e.g., Lenerz (1977, 28)), and I will not adopt it here. Second, one could do the same thing as with the notion of grammaticality and simply rely on speakers' intuitions. Indeed, it seems to me that speakers usually have fairly clear intuitions about, e.g., which of two competing grammatical linear orders represents the more marked one. Finally, Höhle (1982, 102 & 141) (based on earlier work by Lenerz (1977, 27)) has come up with a way of determining markedness that can to some extent be viewed as a formal account of speakers' intuitions, and that strikes me as most reliable. The basic idea is simple: The more context types a given sentence can occur in, the less marked it is. This idea can be formalized as in (6) (irrelevant terminological differences are tacitly adjusted):

(6) A candidate α is less marked than a candidate β if α, β are in the same candidate set, and α can have more foci than β (i.e., can occur in more context types), given appropriate stress assignment.

Following Höhle (1982), two candidates are in the same candidate set if they differ only with respect to word order and/or stress assignment. Note that if, as a special case of (6), a candidate α permits maximal focus projection (or maximal focus spreading, i.e., the whole sentence can be the focus), it is always unmarked. In what follows, I will presuppose that relative degrees of markedness can be empirically determined in one of the two ways envisaged here, i.e., either by directly invoking speakers' judgements, or by adhering to the number of context types in which the candidate is possible.

3.2. The Grammatical Implementation of Markedness

3.2.1. The DTC-Based Approach

Next, the question arises of how the empirically verifiable degrees of markedness of free word order structures can be predicted by the grammar. A commonly adopted approach builds on a concept that plays a crucial role in the classic Derivational Theory of Complexity (henceforth DTC, cf. Watt (1970)). According to the DTC, syntactic transformations are costly for the parser; simplifying somewhat, the more transformations have applied to a sentence, the longer parsing takes on this view. This idea also shows up in what I would like to call the DTC-based theory of markedness, which can be formulated as in (7), and which is adopted (in some form, though perhaps not always with this rigour) by Lenerz (1977, 85), Frey & Tappe (1991, 6ff), Haider (1992, 7), Fortmann & Frey (1997, 145), Wunderlich (1997, 45), and Haider & Rosengren (1998, 14), among others.

(7)   The D-structure order of arguments is the unmarked one; clause-internal order-changing movement operations create markedness.

An immediate consequence of (7) is that different D-structure orders must be assumed for different classes of double object verbs, and this point is indeed emphasized by Haider (1992). On this view, the unmarked sentences in (8), (9), and (10) exhibit the base-generated argument order; the marked sentences are derived by scrambling. Consider first the sentences in (8):

(8)   a.   ?daß man diesem Einfluß   die Kinder   ausgesetzt/ ausgeliefert/
             that one this   influence$_{dat}$ the children$_{acc}$ exposed/   extradited/
             entzogen   hat
             taken away from has
       b.   daß man die Kinder   diesem Einfluß   ausgesetzt/ ausgeliefert/
             that one the children$_{acc}$ this   influence$_{dat}$ exposed/   extradited/
             entzogen   hat
             taken away from has

Here, the unmarked order is DO $\succ$ IO, and accordingly the DTC-based approach views (8-b) as exhibiting base-generated argument order, and (8-a) as being derived by IO scrambling. The situation is reversed in (9):

(9)   a.   daß er der Maria   den Erfolg   gönnt
             that he ART Maria$_{dat}$ the success$_{acc}$ not begrudges
       b.   ?daß er den Erfolg   der Maria   gönnt
             that he the success$_{acc}$ ART Maria$_{dat}$ not begrudges
       c.   daß er dem Fritz   die Zeitung   überließ
             that he ART Fritz$_{dat}$ the newspaper$_{acc}$ left
       d.   ?daß er die Zeitung   dem Fritz   überließ
             that he the newspaper$_{acc}$ ART Fritz$_{dat}$ left

6

Now, different verbs are involved, and it seems that here, the unmarked order is IO ≻ DO. Thus, Haider (1992) assumes that with these verbs, IO ≻ DO is the base-generated argument order, and DO ≻ IO is derived by DO scrambling. Finally, consider (10):

(10)   a.   daß man dem Fritz   das Geld   gegeben hat
            that one   ART Fritz$_{dat}$ the money$_{acc}$ given   has
      b.(?)daß   man das Geld   dem Fritz   gegeben hat
            that one   the money$_{acc}$ ART Fritz$_{dat}$ given   has
      c.   daß ich der Maria   den Peter   vorgestellt habe
            that I   ART Maria$_{dat}$ ART Peter$_{acc}$ introduced have
      d.(?)daß   ich den Peter   der Maria   vorgestellt habe
            that I   ART Peter$_{acc}$ ART Maria$_{dat}$ introduced have

There is disagreement about the data in this case. On the one hand, Haider (1992), Höhle (1982), Reis (1986), Fanselow (1997), and Fortmann & Frey (1997) postulate that both orders are equally unmarked with verbs like geben ('give') or vorstellen ('introduce'). On the other hand, Lenerz (1977), Haftka (1981), Stechow & Sternefeld (1988), Webelhuth (1992), and Meinunger (1995) assume that only the order IO ≻ DO can be unmarked here. I will adopt the latter view. Then, given the DTC-based approach, (10-a) and (10-c) could be base-generated, whereas (10-b) and (10-d) would have to involve a scrambling operation, just as in (9). Thus, notwithstanding the empirical uncertainty surrounding the sentences in (10), the DTC-based approach offers a straightforward explanation of the markedness properties associated with the sentences of this subsection. Still, closer inspection reveals that this approach is untenable.

3.2.2.   *Arguments Against the DTC-Based Approach*

First, it is worth bearing in mind that, despite some suggestions to the contrary that can be found in the literature, (7) does not follow from anything. Thus, as pointed out by Höhle (1982, 137-138), there is no reason why relative degrees of markedness should reflect the degree of deviation from the D-structure order of arguments.[6]

     Second, the evidence involving binding of anaphors that was discussed in subsection 2.2 above clearly suggests that the D-structure order is uniformly DO ≻ IO, irrespective of the type of verb involved. Thus, a verb like ausliefern ('extradite') in (3) typically induces DO ≻ IO as the unmarked order (cf. (8)), whereas the unmarked order with verbs like gönnen ('not begrudge') and vorstellen ('introduce') in (3) appears to be IO ≻ DO (cf. (9) and (10)) – however, the evidence from anaphoric binding strongly suggests a D-structure order DO ≻ IO in all cases. Similarly, the verbs zeigen ('show') and überlassen ('leave to') are double object verbs that typically trigger the unmarked order IO ≻ DO (cf., e.g., (9)); but again, with respect to binding theory, they unambiguously show that the D-structure order must be DO ≻ IO; cf. (4). Thus, given the clear evidence from anaphoric licensing, it seems impossible to assume that verbs that behave differently with respect to unmarked argument order at S-structure can have different argument orders at D-structure.

---

[6]The situation might be different with parsing; cf. Schlesewsky, Fanselow, & Kliegl (1997, 16).

A third argument against the DTC-based approach rests on data involving weak pronoun movement to the Wackernagel position. If both DO and IO are weak pronouns that undergo movement, they have to reassemble in the order DO $\succ$ IO (i.e., in exactly the order that I have argued to be the D-structure one). Strikingly, this order is again invariant across verb classes; verbs like <u>gönnen</u> ('not begrudge') and <u>geben</u> ('give') (which otherwise appear to induce an unmarked argument order IO $\succ$ DO; cf. (9) and (10), respectively) must obey it in the same way as verbs like <u>entziehen</u> ('take away from') (which usually have DO $\succ$ IO as the unmarked order; cf. (8)). As noted by Reis (1986, 20), all the sentences exhibiting fixed pronoun order in (11) allow maximal focus projection; hence, the only order that is permitted is also unmarked, given Höhle's (1982) notion of markedness in (6).

(11)   a.   daß es$_1$ ihm$_2$/*ihm$_2$ es$_1$ der   Fritz t$_1$ t$_2$ gegeben hat
              that it   him/him   it   ART Fritz   given   has
         b.   daß es$_1$ ihm$_2$/*ihm$_2$ es$_1$ der   Fritz t$_1$ t$_2$ gegönnt   hat
              that it   him/him   it   ART Fritz   not begrudged has
         c.   daß es$_1$ ihm$_2$/*ihm$_2$ es$_1$ der   Fritz t$_1$ t$_2$ entzogen   hat
              that it   him/him   it   ART Fritz   taken away from has

Clearly, there is no obvious way to reconcile the evidence in (11) with the idea that markedness is related to argument order at D-structure.

Fourth, basically the same point can be made with sentences in which only one of the two objects is a weak pronoun. The pronoun always precedes the non-pronominal object NP, no matter whether the pronoun is the DO or the IO, and whether the verb otherwise induces IO $\succ$ DO or DO $\succ$ IO as the unmarked argument order:

(12)   a.   daß man es$_1$ Maria$_2$ gegeben hat
              that one   it$_{acc}$ Maria$_{dat}$ given   has
         b.   daß man es$_1$ Maria$_2$ gegönnt   hat
              that one   it$_{acc}$ Maria$_{dat}$ not begrudged has
         c.   daß man ihm$_2$ die Kinder$_1$ entzogen   hat
              that one   him$_{dat}$ the children$_{acc}$ taken away from has

Thus, (12-a) and (12-b) have DO $\succ$ IO as the only possible order (because DO is a weak pronoun, and IO is a lexical NP), even though <u>geben</u> and <u>gönnen</u> otherwise seem to induce the order IO $\succ$ DO. Conversely, (12-c) has the order IO $\succ$ DO, but the verb <u>entziehen</u> otherwise induces the unmarked order DO $\succ$ IO. Still, maximal focus projection appears possible in all these sentences, and they are clearly unmarked (again, cf. Reis (1986, 20)).

Finally, it turns out that, depending on the animacy status of its objects, one and the same verb can exhibit two different unmarked orders; cf. Fanselow (1995, 30) and Vogel & Steinbach (1995, 106; 1998, 70). Consider, e.g., the verb <u>entziehen</u>. Thus far, we have seen that with two non-prononimal object NPs, the unmarked order seems to be DO $\succ$ IO; the relevant data (cf. (8)) are repeated here in (13-ab). However, note that here the IO is [–animate], whereas the DO is [+animate]. Interestingly, if both IO and DO are [+animate], the unmarked order is reversed; cf. (13-cd).

(13) a. ?daß man diesem Einfluß$_2$ die Kinder$_1$ t$_2$ entzogen   hat
          that one this   influence$_{dat}$ the children$_{acc}$ taken away from has
      b. daß man die Kinder$_1$ diesem Einfluß$_2$ entzogen   hat
          that one the children$_{acc}$ this   influence$_{dat}$ taken away from has
      c. daß man der Maria$_2$ die Kinder$_1$ t$_2$ entzogen   hat
          that one ART Maria$_{dat}$ the children$_{acc}$ taken away from has
      d. ?daß man die Kinder$_1$ der Maria$_2$ entzogen   hat
          that one the children$_{acc}$ ART Maria$_{dat}$ taken away from has

To make this evidence compatible with the DTC-based approach to markedness, one would have to assume that the difference in animacy is reflected in the D-structure order of arguments already, such that the D-structure order would be DO $\succ$ IO with <u>entziehen</u> if IO is [–animate], and IO $\succ$ DO if IO is [+animate], clearly an undesirable result if this implies a proliferation of lexical entries.

To sum up so far, it seems that the DTC-based approach to markedness is both conceptually unmotivated and empirically problematic. I would like to conclude from this that another grammatical implementation of markedness is called for, one that does not determine relative markedness of a sentence by measuring the degree to which it deviates from D-structure. In the following sections, I will argue that the markedness problem can indeed receive a straightforward solution under an optimality theoretic approach, which also covers the other problems noted at the beginning of this article.

4. Optimality Theory

4.1. Basic Assumptions

An optimality theoretic syntax takes the following form (cf. Grimshaw (1997) and Pesetsky (1998), among others): One part of the grammar (GEN) consists entirely of constraints that are neither violable nor ranked, just as in standard approaches. GEN generates the set of candidates {K$_1$, K$_2$, ...}. These candidates are then subjected to the second, optimality theoretic part of the grammar (H-EVAL, 'Harmony Evaluation')) that determines the optimal candidate(s) K$_i$ in {K$_1$, K$_2$, ...}. The constraints in this latter part of the grammar are assumed to be universal, violable, and ranked, and it is these constraints that I will focus on in what follows. Central to the theory is the notion of competition: The grammaticality of a candidate K cannot be determined by solely looking at internal properties of K; rather, external factors (the competition of K and other candidates) are decisive. More specifically, grammaticality is viewed as synonymous with optimality, and can be defined as in (14) (cf. Grimshaw (1997)).

(14) Grammaticality:
      A candidate K$_i$ is grammatical (optimal) iff, for every candidate K$_j$ in the same candidate set, K$_i$ satisfies the highest-ranking constraint B$_i$ of the constraint hierarchy <B$_1$, B$_2$, ... B$_n$> on which K$_i$ and K$_j$ differ better than K$_j$.

It remains to be clarified what kinds of formal objects the candidates created by GEN are, and how the candidate set is defined. I will assume here that candidates are <D-

9

structure,S-structure> pairs, and that the candidate set is defined more or less as in Chomsky (1995): Two candidates are in the same candidate set iff they have the same numeration, i.e., simplifying somewhat, the same lexical material.[7]

4.2.   Constraints

Let us turn next to the constraints of the H-Eval system. Two types of constraints will become relevant in what follows, viz., (a) markedness constraints that trigger movement (X-Criteria); and (b) faithfulness constraints that prohibit (or minimize the effects of) movement. Among the first class of constraints I assume to be a constraint that forces weak pronouns to show up in the domain of the functional head π at S-structure: the Pronoun Criterion, or Pron-Crit. This constraint is inspired by Johnson (1991), but it can be traced back in its essentials at least to Thiersch's (1978, 84) rule C_1.[8]

(15)   Pron-Crit:
        Weak pronouns must be in the domain of π at S-structure.

Here, the domain of a head X comprises SpecX and XP adjuncts, perhaps even X adjuncts. I will assume that only one pronoun can be substituted in Specπ so as to fulfill Pron-Crit, and further pronouns must be adjoined to πP. Next, the Extended Projection Principle (EPP, cf. Chomsky (1982; 1995)) demands that NPs with a nominative Case feature are in SpecT at S-structure.[9]

(16)   EPP:
        NP_{nom} must be in SpecT at S-structure.

Turning now to faithfulness constraints, the first constraint that will play a role is Economy of Derivation, or Stay, which prohibits overt movement (cf. Chomsky (1991; 1995) for the

---

[7]Both assumptions may eventually turn out to be too simplistic, but they may suffice for present purposes. See Grimshaw (1997), Müller (1997), Ackema & Neeleman (1998), and Legendre, Smolensky, & Wilson (1998) for general discussion and more elaborate concepts that, i.a., incorporate the notion of (intended) identity of meaning/LF in candidate sets.

[8]Pron-Crit is analogous to other criteria as they have been proposed in the literature; cf. in particular the Wh-Criterion that is assumed by May (1985), Lasnik & Saito (1992), and Rizzi (1996), among others, and adopted within optimality theory in one form or another by Grimshaw (1997), Müller (1997), Ackema & Neeleman (1998), and Legendre, Smolensky, & Wilson (1998). The standard formulation of the Wh-Criterion is composed of two conjunctive statements, one that forces <u>wh</u>-phrases to move, and one stating that a C node bearing a [+wh] feature requires a lexical element in its domain. Indeed, it seems to me that the criteria adopted here should yield the same effect as the conjunctive formulation of the Wh-Criterion. In the case at hand, I will presuppose that indeed <u>only</u> weak pronouns can show up in the domain of π at S-structure (and not, e.g., lexical NPs). This could be achieved by adding a statement to this effect in (15). Alternatively, we may assume that there is a general and undominated (i.e., inviolable) constraint (that we may call F-Match) which demands that α can be moved to a position β only if α is equipped with features that match those of position β; F-Match ensures that there is no movement of [–wh] phrases to $SpecC_{[+wh]}$; and it also implies that there can be no movement of non-pronominal NPs to the π domain.

[9]I will leave open the question of whether Pron-Crit and EPP can also be satisfied by traces.

10

concept, and Grimshaw (1997) and Legendre, Smolensky, & Wilson (1998) for optimality theoretic applications).

(17)   STAY:
        S-structure movement is not allowed.

Finally, I will presuppose another faithfulness constraint that has the effect of minimizing the effect of syntactic movement operations, viz., Parallel Movement (PAR-MOVE), which is extensively motivated in Müller (1999) on the basis of fixed word order properties of, e.g., Scandinavian object shift (cf. Holmberg (1986), Vikner (1990), Collins & Thráinsson (1996)), multiple <u>wh</u>-movement in Bulgarian (cf. Rudin (1985)), simple <u>wh</u>-movement in English (superiority), and quantifier raising in German.

(18)   PAR-MOVE:
        If α c-commands β at level L_n, then α c-commands β at level L_{n+1}
        (where α, β are arguments).

Markedness constraints and faithfulness constraints impose conflicting requirements on candidates, which must be resolved by constraint ranking. For German, I will assume the ranking in (19), with PRON-CRIT dominating PAR-MOVE and STAY, and STAY and EPP being equally ranked, i.e., tied (cf. Prince & Smolensky (1993) for this concept):

(19)   PRON-CRIT $\gg$ EPP, STAY $\gg$ PAR-MOVE

This does not yet account for scrambling; however, it correctly predicts the distribution of weak pronouns in German.

5.   Weak Pronouns

5.1.   Empirical Evidence

As observed by Lenerz (1977; 1992), Haftka (1981), Hoberg (1981), Reis (1986), Haider & Rosengren (1998), and many others, weak (i.e., unstressed) pronouns in German show up in the left periphery of the Mittelfeld, and their order is fixed.[10] A subject pronoun must precede an object pronoun:

(20)   a.   daß sie$_1$ es$_2$ wahrscheinlich nicht t$_1$ t$_2$ lesen wollte
              that she$_{nom}$ it$_{acc}$ probably   not   read wanted
        b.   *daß es$_2$ sie$_1$ wahrscheinlich nicht t$_1$ t$_2$ lesen wollte

Similarly, as noted above, a DO pronoun precedes an IO pronoun:

---

[10]The situation is different with strong (stressed) and clitic pronouns, neither of which obey PRON-CRIT. Essentially, strong pronouns behave like non-pronominal NPs and can be scrambled, whereas clitic pronouns must cliticize onto a (phonologically) appropriate host. Accordingly, all ungrammatical sentences in this section can be made grammatical by stressing or cliticizing the pronouns. In what follows, I will disregard non-weak pronouns.

(21) a. daß es$_1$ ihm$_2$ der Fritz   t$_1$ t$_2$ gegeben hat
          that it$_{acc}$ him$_{dat}$ ART Fritz$_{nom}$ given   has
      b. *daß ihm$_2$ es$_1$ der Fritz t$_1$ t$_2$ gegeben hat

In the case of three pronominal arguments, the fixed order is SUBJ $\succ$ DO $\succ$ IO:

(22) a. daß sie$_1$ es$_2$ ihm$_3$ wahrscheinlich zum Geburtstag t$_1$ t$_2$ t$_3$ schenken
          that she$_{nom}$ it$_{acc}$ him$_{dat}$ probably   for   birthday   give
          wird
          will
      b. *daß sie$_1$ ihm$_3$ es$_2$ wahrscheinlich zum Geburtstag t$_1$ t$_2$ t$_3$ schenken wird
      c. *daß es$_2$ sie$_1$ ihm$_3$ wahrscheinlich zum Geburtstag t$_1$ t$_2$ t$_3$ schenken wird
      d. *daß es$_2$ ihm$_3$ sie$_1$ wahrscheinlich zum Geburtstag t$_1$ t$_2$ t$_3$ schenken wird
      e. *daß ihm$_3$ sie$_1$ es$_2$ wahrscheinlich zum Geburtstag t$_1$ t$_2$ t$_3$ schenken wird
      f. *daß ihm$_3$ es$_2$ sie$_1$ wahrscheinlich zum Geburtstag t$_1$ t$_2$ t$_3$ schenken wird
      g. *daß sie$_1$ es$_2$ wahrscheinlich zum Geburtstag t$_1$ t$_2$ ihm$_3$ schenken wird
      h. *daß wahrscheinlich zum Geburtstag sie$_1$ es$_2$ ihm$_3$ schenken wird

5.2. Analysis

These data follow from (19). The partial ranking Pron-Crit ≫ Stay implies that
weak pronouns must undergo raising to the domain of π at S-structure, either to Specπ
or, in cases of multiple pronoun fronting, to a πP-(left-)adjoined position. Hence, they
must precede VP adjuncts and VP-internal objects (on subjects, see subsection 5.3). The
fixed order property of fronted pronouns is an immediate consequence of Par-Move.
Under the partial ranking Pron-Crit ≫ Par-Move, weak pronouns can cross non-
pronominal NPs, thereby selectively violating Par-Move in order to fulfill the higher-
ranked Pron-Crit. However, if two (or more) pronouns undergo fronting to the domain
of π, all orders fulfill Pron-Crit equally well, and the lower-ranked Par-Move becomes
relevant. A low-ranked Par-Move predicts that, if possible, the D-structure order of
arguments should be preserved at S-structure, and this is exactly what the data in the
previous subsection show: The order among weak pronouns is identical to the D-structure
order postulated above for all German verbs, on the basis of evidence involving anaphoric
binding. Thus, given Par-Move, we now have a second strong argument for postulating
the clause structure in (1), with DO uniformly preceding IO at D-structure (also recall
that the order of weak pronouns is independent of the choice of verb).

Abstracting away from optional subject raising triggered by the EPP for the mo-
ment, the competition underlying (21), from which K_1 (= (21-a)) emerges as the optimal
candidate, is illustrated in tableau T_1.

The optimal candidate violates Par-Move (and Stay) twice (both DO and IO are c-
commanded by the subject at D-structure and c-command the subject at S-structure),
but makes up for this by respecting the higher-ranked Pron-Crit, in contrast to the
suboptimal candidates $K_3 - K_5$ which incur a fatal violation of Pron-Crit by leaving
either one or both pronouns in situ. Thus, Par-Move can be violated, but the violation
must be kept minimal, which it is not in K_2: Here, an additional, hence fatal, Par-Move
violation occurs because the fronted DO and IO pronouns do not show up in their D-

Tableau T$_1$: Pronoun Movement of DO and IO:

Candidates	Pron-Crit	Stay	Par-M
⇒K$_1$: es$_1$ ihm$_2$ der Fritz t$_1$ t$_2$ ...		**	**
*K$_2$: ihm$_2$ es$_1$ der Fritz t$_1$ t$_2$ ...		**	***!
*K$_3$: es$_1$ der Fritz t$_1$ ihm$_2$ ...	*!	*	*
*K$_4$: ihm$_2$ der Fritz es$_1$ t$_2$ ...	*!	*	**
*K$_5$: der Fritz es$_1$ ihm$_2$ ...	*!*		

structure order. Next, tableau T$_2$ shows that in the case of three weak pronouns (cf. (22)), the candidate emerges as a winner that respects both Pron-Crit (by moving all three pronouns) and Par-Move (by reassembling the pronouns in the π domain in their D-structure order), and violates Stay three times.

Tableau T$_2$: Pronoun Movement of SUBJ, DO, and IO

Candidates	Pron-Crit	Stay	Par-M
⇒K$_1$: sie$_1$ es$_2$ ihm$_3$ ... t$_1$ t$_2$ t$_3$		***	
*K$_2$: sie$_1$ es$_2$ ... t$_1$ t$_2$ ihm$_3$	*!	**	
*K$_3$: sie$_1$ ihm$_3$ ... t$_1$ es$_2$ t$_3$	*!	**	*
*K$_4$: sie$_1$ ihm$_3$ es$_2$ ... t$_1$ t$_2$ t$_3$		***	*!
*K$_5$: es$_2$ sie$_1$ ihm$_3$ ... t$_1$ t$_2$ t$_3$		***	*!
*K$_6$: es$_2$ ihm$_3$ sie$_1$ ... t$_1$ t$_2$ t$_3$		***	*!*
*K$_7$: ihm$_3$ sie$_1$ es$_2$ ... t$_1$ t$_2$ t$_3$		***	*!*
*K$_8$: ihm$_3$ es$_2$ sie$_1$ ... t$_1$ t$_2$ t$_3$		***	*!**
*K$_9$: ... sie$_1$ es$_2$ ihm$_3$	*!**		

5.3.   Pronoun Fronting and Subjects

Assuming that scrambling applies only VP-internally, it follows that VP adjuncts and non-pronominal object NPs cannot precede weak pronouns clause-internally in German. However, SpecT asymmetrically c-commands the π domain, and this implies that subjects (and only these) should be able to precede weak pronouns in the Mittelfeld, provided that they have undergone optional NP raising. As shown in (23), both predictions are borne out:

(23)   a.   daß Fritz   es$_1$ der Frau$_2$ wahrscheinlich zum Geburtstag t$_1$ t$_2$
             that Fritz$_{nom}$ it$_{acc}$ the woman$_{dat}$ probably   for   birthday
             schenken wird
             give   will
       b.   *daß Fritz der Frau$_2$ es$_1$ wahrscheinlich zum Geburtstag t$_1$ t$_2$ schenken wird
       c.   *daß Fritz wahrscheinlich zum Geburtstag es$_1$ der Frau$_2$ schenken wird

Even with a subject NP in SpecT, the fixed order among weak pronouns must be respected. This is shown in (24), which minimally contrasts with (21).

(24)   a.   daß der   Fritz$_3$ es$_1$ ihm$_2$ t$_3$ t$_1$ t$_2$ gegeben hat
             that ART Fritz$_{nom}$ it$_{acc}$ him$_{dat}$ given   has
       b.   *daß der Fritz$_3$ ihm$_2$ es$_1$ t$_3$ t$_1$ t$_2$ gegeben hat

13

At this point, the tie of EPP and STAY becomes relevant (ties being a standard optimality theoretic device to derive optionality). I will here assume a global concept of tie (cf. Prince & Smolensky (1993, 50)), according to which a candidate is grammatical if it is optimal under one ranking of tied constraints.[11] It now turns out that, even though (21-a) and (24-a) are part of the same candidate set, they are both optimal, depending on whether the tie is resolved into a ranking STAY ≫ EPP (under which (21-a), with the subject in situ, is optimal), or into a ranking EPP ≫ STAY (under which (24-a), with the subject in SpecT, is optimal). This is shown in tableau T_3, an extension of T_1 ($K_2 - K_5$ are not repeated again; K_6, K_7 underlie the sentences in (24-a) and (24-b), respectively):[12]

Tableau T_3: Pronoun Movement of DO and IO, plus EPP Effects:

Candidates	PRON-CRIT	EPP	STAY	PAR-M
⇒K_1: es$_1$ ihm$_2$ der Fritz t$_1$ t$_2$ ...		*	**	**
⇒K_6: der Fritz$_3$ es$_1$ ihm$_2$ t$_3$ t$_1$ t$_2$ ...			***	
*K_7: der Fritz$_3$ ihm$_2$ es$_1$ t$_3$ t$_1$ t$_2$ ...			***	*!

Thus, by assuming a global tie of EPP and STAY, the optionality of subject raising to SpecT in German can be accounted for. I will henceforth tacitly presuppose this tie, but since I will be exclusively concerned with VP-internal scrambling in what follows, I will abstract away from EPP-driven movement, and from candidates that are optimal under the ranking EPP ≫ STAY.[13]

5.4. Conclusion

To sum up this section, it has turned out that the rigid order of weak pronouns in German follows from the interaction of PRON-CRIT and PAR-MOVE.[14] Furthermore, in addition to the data on anaphoric binding, the rigid order of weak pronouns (in interaction with

---

[11]Prince & Smolensky also envisage a local concept of tie (via "crucial non-ranking") that would not be compatible with the analysis below, due to the two PAR-MOVE violations of K_1 in T_3 that K_6 does without. See Müller (1999a) for a general discussion of ties in syntax.

[12]Other candidates that fulfill EPP by raising the subject NP, but fail to move one of the weak pronouns (or both), are blocked as involving fatal PRON-CRIT violations, much as $K_3 - K_5$.

[13]A reviewer points out that object infinitives may precede weak pronouns clause-internally; cf.:

(i)   daß [den Hund zu füttern] sie keinem befohlen hat
       that   the dog   to feed   she no-one ordered has

The literature does not agree on whether infinitive fronting in (i) qualifies as an instance of scrambling (cf. the discussion and references in Müller (1995, ch. 3)). This movement operation shares properties with both scrambling and topicalization. It targets a TP-adjoined position, and it requires an I-topicalization (rise/fall) contour and topic interpretation. I will here assume that this movement is not triggered by the same constraint that triggers regular NP scrambling (SCR-CRIT, see below), but I will leave the exact nature of the constraint in question open.

[14]A bit more must be said about complex constructions, though. Coherent infinitives exhibit the same fixed order among fronted pronouns:

PAR-MOVE) offers a second argument for the base order DO ≻ IO in German (cf. (1)). On the basis of the system as it has emerged so far, let me now turn to scrambling.

6. Scrambling

6.1. Illformedness vs. Markedness

Recall that scrambling structures pose an economy problem and a markedness problem. In present terms, the economy problem can be rephrased as follows: Because of STAY and PAR-MOVE, each scrambling operation must be triggered by a higher-ranked constraint (a "criterion"), the nature of which remains to be clarified. The markedness problem arises straightforwardly under the standard optimality theoretic notion of grammaticality (i.e., optimality) in (14), according to which suboptimal candidates are invariably ungrammatical. It seems that this result does not reflect the actual data situation: Mittelfeld-internal argument orders often have an intermediate status that calls for a treatment in terms of relative markedness, and not complete well- or illformedness.[15]

A solution of the markedness problem that, at first sight, looks appealing is proposed by Keller (1996, 50). The basic idea is to give up the assumption that suboptimal candidates are invariably ungrammatical. More specifically, the notion of grammaticality in (14) is replaced by the notion of suboptimality in (25) (which, however, still uses (14) as

---

(i)   a.   daß es$_2$ ihm$_3$ keiner   t$_3$ [$_\alpha$ t$_2$ zu lesen] empfohlen   hat
            that it$_{acc}$ him$_{dat}$ no-one$_{nom}$ to read   recommended has
      b. ?*daß ihm$_3$ es$_2$ keiner   t$_3$ [$_\alpha$ t$_2$ zu lesen] empfohlen   hat
            that him$_{dat}$ it$_{acc}$ no-one$_{nom}$ to read   recommended has

If coherent constructions have a bisentential structure, as in (i), it seems that PAR-MOVE would wrongly predict the ill-formed S-structure order in (i-b) rather than the well-formed S-structure order in (i-a). Such a problem would not necessarily arise under a monosential approach to coherent infinitives, as in Haider (1993). Alternatively, the data in (i) could be taken to indicate the activity of a high-ranked output/output faithfulness constraint in the sense of McCarthy & Prince (1995), such that the pronoun order in complex clauses must reflect the pronoun order in simple clauses, which is in turn derived from PAR-MOVE; essentially, this would be an optimality theoretic implementation of the classical concept of analogy. For reasons of space and coherence, I will not pursue these matters here.

[15]Choi (1996), Büring (1997), and Costa (1998) devise optimality theoretic approaches to free word order according to which all clause-internal word orders are perfect realizations with respect to a given context. On this view, the notion of grammaticality in (14) is unproblematic. However, I would like to contend that this is the wrong move in view of the empirical evidence. Recall first that markedness is understood here in terms of context types, and is empirically verifiable by invoking native speaker intuitions. If all grammatical word orders are equally perfect, there is no simple grammar-internal way to derive that one sentence is less marked than a competing sentence. Hence, either additional assumptions have to be made grammar-internally (Costa suggests that unmarked sentences are optimal candidates that satisfy discourse-oriented constraints vacuously), or the concept of relative markedness of sentences as such must be given up, and be replaced by a concept of relative markedness of contexts (cf. Choi, Büring). Second, there are sentences which are not perfect realizations in any context, but nevertheless appear to be grammatical (see below); and this cannot be accounted for under the approaches of Choi (1996), Büring (1997), and Costa (1998). As we will see, the approach to be developed here evades both these problems.

an auxiliary definition that covers the notion of optimality in (25)).

(25) Suboptimality:
      A structure S_i is suboptimal with respect to a structure S_j if there are subsets R_i and R_j of the candidate set such that S_i is optimal for R_i and S_j is optimal for R_j and $R_i \subset R_j$ holds. A structure S_i is less grammatical than a structure S_j if S_i is suboptimal with respect to S_j.

Reinterpreting "less grammatical" in (25) as "more marked," it follows that the optimal candidate according to (14) is now viewed as the least marked, the second-best candidate is more marked than the optimal candidate, but less marked than the third-best candidate, and so forth. However, this graded approach to grammaticality faces a fundamental problem: It is a characeristic of most (if not all) syntactic analyses that have been developed within optimality theory that the "second-best" candidate is not less marked than the "third-best." In fact, in most cases, all suboptimal candidates are decidedly ungrammatical, with no variability involved. An arbitrarily chosen example from Grimshaw (1997, 378) may illustrate this (the same conclusion would apply in the case of pronoun fronting in German that was discussed above). Grimshaw adopts constraints that force <u>wh</u>-movement (Op-Spec), prohibit empty heads (Ob-Hd), and prohibit movement in general (Stay); the English ranking of these constraints is assumed to reflect the order in which I have presented them. From these assumptions it follows that (26-a) is optimal, and (26-bcd) are blocked as suboptimal (as matrix <u>wh</u>-questions), with (26-b) emerging as the second-best candidate (Op-Spec is satisfied, but Ob-Hd is not), (26-c) as the third-best (Op-Spec is violated, Ob-Hd and Stay are respected), and (26-d) as the worst (Op-Spec and Stay are violated).

(26) a. What will John read ?
      b. *What John will read ?
      c. *John will read what ?
      d. *Will John read what ?

Clearly, by adopting Keller's (1996) concept of suboptimality, we should expect that the sentences in (26-bcd) are not completely ill formed, and that their degree of markedness should increase from top to bottom. Neither prediction seems correct, though. Thus, if the optimality theoretic concept of grammaticality in (14) is replaced by the concept of suboptimality in (25), this implies abandoning the main bulk of analyses that have been proposed in optimality theoretic syntax.

How, then, can we maintain the standard assumption that suboptimal candidates are ungrammatical, and at the same time permit markedness differences among grammatical candidates? I would like to suggest that this dilemma can be solved by distinguishing two constraint levels, a matrix hierarchy and a subhierarchy. Fatal violations on the matrix hierarchy induce strict ungrammaticality. The constraints discussed up to now all belong to this hierarchy, and accordingly, suboptimal candidates in the competitions discussed above are correctly predicted to be ill formed. In contrast, fatal violations on the

subhierarchy only induce markedness.[16] The constraints that trigger scrambling belong to this latter hierarchy, and accordingly, we find degrees of markedness with (grammatical) candidates in this domain.[17]

6.2. A Subhierarchy

Suppose that scrambling is triggered by a Scrambling Criterion (Scr-Crit) that outranks Stay, just like Pron-Crit does. However, unlike Pron-Crit, Scr-Crit is in itself a subhierarchy of (potentially conflicting) linearization constraints. Among the linearization constraints are those listed in (27).[18]

(27) Scr-Crit: In the VP domain,
    a. Nom ('Nominative constraint'): [+nom] precedes [–nom] >
    b. Def ('Definiteness constraint'): [+def] precedes [–def] >
    c. An ('Animacy constraint'): [+animate] precedes [–animate] >
    d. Foc ('Focus constraint'): [–focus] precedes [+focus] >
    e. Dat ('Dative constraint'): [+dat] precedes [+acc] >
    f. Adv ('Adverb constraint'): [+NP] precedes [+adv] >
    g. Per ('Permutation constraint', 'Anti-Par-Move'): If α c-commands β at level L_n, then α does not c-commands β at level L_{n+1}.

I continue to indicate ranking on the matrix hierachy by using the symbol $\gg$; in contrast, ranking on the subhierarchy is indicated by $>$. As for the ranking on the matrix hierarchy, it is clear that Scr-Crit must dominate Stay and Par-Move in German (since scrambling exists and may change the order of NPs). Moreover, Pron-Crit must dom-

---

[16]The notion of subhierarchy here must not be confused with that used by Baković (1998) and Legendre, Smolensky, & Wilson (1998). In these latter approaches, constraints in subhierarchies behave like other constraints with respect to violability (i.e., a fatal violation creates ungrammaticality, not markedness).

[17]Note that this closely corresponds to the distinction between "weak" rules and what may be called "strong" rules that has been introduced by Daneš (1967); also cf. Lenerz (1977) and Reis (1986).

[18]The formulation of these constraints should by and large be self-explanatory; note, however, that [±focus] encodes a syntactic feature – it characterizes the focus exponent, but not necessarily the semantic focus (which can be enlarged by focus projection). Similarly, in the other linearization constraints, semantic/pragmatic properties are encoded as syntactic features; thus, all these subconstraints are strictly syntactic, and not semantic/pragmatic (this evades the pertinent objections in Reis (1986)). All of these constraints have, in one form or another, been proposed in the literature. (To choose an example that may be somewhat less obvious: Per has often been viewed as the ultimate rationale behind scrambling in German; see, e.g., Frey & Tappe (1991) and Haider (1992; 1993), where it is explicitly argued that scrambling in German is licensed only if it reverses the D-structure order of arguments. In the present framework, Per is the "weakest" possible trigger for scrambling.) However, the list in (27) is not exhaustive; in fact, several aspects of Mittelfeld-internal word order variation in German will be ignored in what follows (e.g., those concerning psych verb and unaccusative constructions, and those concerning scrambling of non-NPs). More generally, what follows should be viewed as the sketch of a model in which to tackle word order in German, rather than a comprehensive account. In line with this, even though I think that the specific rankings in this subhierarchy may not be too far off the track, they certainly demand further empirical (ideally, experimental) verification.

inate Scr-Crit because, rather than showing up in the VP domain in order to satisfy (some subconstraint of) Scr-Crit, weak pronouns must move to the π domain (where the linearization constraints in (27) do not apply):

(28)   Pron-Crit $\gg$ Scr-Crit $\gg$ EPP, Stay $\gg$ Par-Move

In accordance with the distinction between matrix hierarchy and subhierarchy, we can now define two different notions of optimality, viz., optimality as grammaticality and optimality as unmarkedness. The revised concept of grammaticality that takes into account the existence of subhierarchies is given in (29) (compare (14)):

(29)   Grammaticality (revised):
        A candidate K_i is grammatical iff, for every candidate K_j in the same candidate set, K_i satisfies the highest-ranking constraint B_k of the matrix hierarchy $<B_1$, B_2, ... $B_n>$ on which K_i and K_j differ better than K_j, where B_l is replaced by some C_i in $<C_1$, C_2, ... $C_n>$ if B_l is a subhierarchy $<C_1$, C_2, ... $C_n>$.

Basically, the determination of grammaticality works as before; however, in the case of a constraint that is a subhierarchy, a winning candidate must be optimal under some replacement of the subhierarchy itself by a constraint belonging to that subhierarchy. In other words, with respect to the notion of grammaticality, the subhierarchy is interpreted via logical disjunction of the subhierarchy constraints; the internal ranking of the subhierarchy is irrelevant. In the case at hand, suppose that Scr-Crit is the highest-ranking constraint on which competing candidates differ. Then a candidate will be grammatical if it is optimal under a ranking in which Scr-Crit is replaced by one of the linearizaton constraints of this subhierarchy, e.g., under a ranking in which Scr-Crit is replaced by Nom, or under a ranking in which Scr-Crit is replaced by Def, and so on. This means that the trigger for scrambling, and hence for VP-internal word order variation, is fairly weak. This theoretical prediction will be shown to be corroborated by empirical evidence – it is indeed the exception rather than the rule for a VP-internal word order in German not to be grammatical at all.[19]
    The notion of unmarkedness can be defined similarly:

(30)   Unmarkedness:
        A candidate K_i is unmarked iff, for every candidate K_j in the same candidate set, K_i satisfies the highest-ranking constraint B_k of the matrix hierarchy $<B_1$, B_2, ... $B_n>$ on which K_i and K_j differ better than K_j, where B_l is replaced by $<C_1$, C_2, ... $C_n>$ if B_l is a subhierarchy $<C_1$, C_2, ... $C_n>$.

The crucial difference here is that the unmarked candidate is determined not by taking one constraint of the subhierarchy and substituting it for the subhierarchy itself (here, Scr-Crit), but by substituting all of the ranked subhierarchy constraints for the subhierarchy (Scr-Crit). Thus, to determine unmarkedness, the distinction between matrix hierarchy and subhierarchy is ignored. It follows that an unmarked candidate must be grammatical,

---

[19]A similar result holds in certain competition-based approaches; cf. especially Uszkoreit (1984; 1986).

but not vice versa. What remains to be accounted for is the notion of degree of markedness, and here I will employ Keller's (1996) concept of suboptimality in (25): Among the grammatical candidates of a candidate set (determined according to the definition of optimality in (29)), a candidate K_j is more marked than another candidate K_i if K_j is suboptimal with respect to K_i according to the definition of optimality in (30) – i.e., the worse the constraint profile of a grammatical candidate, the more marked it is.

With this theory as background, let me now turn to the empirical evidence. For reasons of space, I will not extensively motivate each of the partial rankings in (27); the outlines of the approach should be become clear, though.

6.3. Empirical Evidence

6.3.1. Nominative vs. Definiteness

First, consider the pair of sentences in (31):

(31) a. daß eine Frau   den Fritz   geküßt hat
          that a   woman$_{nom}$ ART Fritz$_{acc}$ kissed has
      b. ?daß den Fritz$_2$ eine Frau$_1$ t$_2$ geküßt hat
          that ART Fritz$_{acc}$ a   woman$_{nom}$ kissed has

Both orders are grammatical, but (31-a) is generally assumed to be less marked than (31-b). There is a constraint conflict here between NOM (which requires the unmarked order in (31-a)) and DEF (which demands the marked order in (31-b)), and we can thus conclude that NOM dominates DEF in the subhierarchy. The competition is shown in tableau T_4.[20]

Tableau T_4: NOM > DEF:

Candidates	SCR-CRIT							STAY	PAR-M
	NOM	DEF	AN	FOC	DAT	ADV	PER		
⇒→K$_1$: eine Frau$_1$ den Fritz$_2$		*					*		
⇒ K$_2$: den Fritz$_2$ eine Frau$_1$ t$_1$	*?							*	*

K_1 is grammatical under substitution not only of NOM, but of any of the subconstraints of SCR-CRIT except for DEF and PER. The reason is that substitution of any of the constraints on which the two candidates do not differ (such as, e.g., AN) will pass on the decision to the lower-ranked constraints STAY and PAR-MOVE, and since K_1 does not involve scrambling (whereas K_2 does), it has an "inherent advantage" and will be optimal in this case. K_2, on the other hand, is grammatical under substitution of either DEF or PER, which can thus be seen as triggering the scrambling operation. However, only K_1 is unmarked: If the ordered constraints of the subhierarchy replace SCR-CRIT as a whole, only this candidate emerges as optimal.

---

[20]A remark on notation: A fatal violation of a grammatical candidate that renders it marked is indicated by "?" instead of "!", which I continue to assume (in line with the usual practice) to indicate a fatal violation that makes a candidate ungrammatical. In addition, whereas ⇒ continues to identify a grammatical candidate, unmarkedness will be indicated by →.

6.3.2. Definiteness vs. Animacy

The examples in (32) exhibit a conflict of DEF and AN and suggest the partial ranking
DEF > AN:

(32)   a. daß der Verkäufer   den Wein$_1$ einem Kunden$_2$ empfahl
           that the shop assistant$_{nom}$ the wine$_{acc}$ a   customer$_{dat}$ recommended
      b. ?daß der Verkäufer   einem Kunden$_2$ den Wein$_1$ t$_2$ empfahl
           that the shop assistant$_{nom}$ a   customer$_{dat}$ the wine$_{acc}$ recommended

Given the uniform D-structure order DO $\succ$ IO argued for above, K$_1$ in tableau T$_5$ ((32-a))
does without scrambling and is therefore grammatical not only under DEF substitution,
but also under substitution of NOM, FOC, and ADV (where the two candidates do not
differ). In contrast, K$_2$ ((32-b)), which involves IO scrambling, is grammatical under
substitution of AN, DAT, and PER, which thus act as triggers for scrambling in this case.
Given that DEF dominates AN, K$_1$ is predicted to be unmarked, and K$_2$ as marked with
respect to K$_1$, in line with what is observed in Stechow & Sternefeld (1988, 453).[21]

Tableau T$_5$: DEF > AN:

Candidates	SCR-CRIT							STAY	PAR-M
	NOM	DEF	AN	FOC	DAT	ADV	PER		
⇒→K$_1$: d. Wein$_1$ e. Kunden$_2$			*		*		*		
⇒ K$_2$: e. Kunden$_2$ d. Wein$_1$ t$_2$		*?						*	*

6.3.3. Animacy vs. Dative

In determining the ranking of AN and DAT, we can now return to the examples in (13) that
turned out to be problematic for a DTC-based approach to markedness (cf. subsection
3.2.2 above). First, the data in (33) clearly suggest that if IO and DO differ with respect
to animacy, the object that is [+animate] always precedes the one that is [–animate]
in the unmarked candidate, irrespective of the issue of Case (other things being equal).
Thus, in (33-ab) and (33-cd), the IO is [+animate], and in (33-ef), the DO is [+animate].
What distinguishes verbs like <u>geben</u> ('give') from verbs like <u>entziehen</u> ('take away from')
or <u>aussetzen</u> ('expose') on this view is not a different D-structure order of arguments
(which we have seen evidence against), but rather the fact that the first class of verbs
tend to have an IO that is [+animate] and a DO that is [–animate], whereas the second
class often selects an IO that is [–animate] and a DO that is [+animate].

(33)   a. daß man dem Fritz$_2$ das Buch$_1$ t$_2$ gegeben/geschickt hat
           that one ART Fritz$_{dat}$ the book$_{acc}$ given/sent   has

---

[21]Note that T$_5$ shows very clearly that the absolute number of SCR-CRIT violations of a candidate
is irrelevant, as predicted under (30) – the unmarked candidate K$_1$ has more SCR-CRIT violations than
the marked candidate K$_2$. The irrelevance of cumulative violations of lower-ranked subconstraints in the
present analysis can also be verified by considering tableau T$_7$ below.

  b. ?daß man das Buch$_1$ dem Fritz$_2$ gegeben/geschickt hat
      that one the book$_{acc}$ ART Fritz$_{dat}$ given/sent   has

  c. daß man der Claudia$_2$ den Wagen$_1$ t$_2$ zeigen/gönnen/verweigern muß
      that one ART Claudia$_{dat}$ the car$_{acc}$ show/not begrudge/deny must

  d. ?daß man den Wagen$_1$ der Claudia$_2$ zeigen/gönnen/verweigern muß
      that one the car$_{acc}$ ART Claudia$_{dat}$ show/not begrudge/deny must

  e. ?daß man diesem Einfluß$_2$ die Kinder$_1$ t$_2$ entzogen/
      that one this   influence$_{dat}$ the children$_{acc}$ taken away from/
      ausgeliefert/ ausgesetzt hat
      extradited/ exposed   has

  f. daß man die Kinder$_1$ diesem Einfluß$_2$ entzogen/   ausgeliefert/
      that one the children$_{acc}$ this   influence$_{dat}$ taken away from/ extradited/
      ausgesetzt hat
      exposed   has

If the demands imposed by AN and DAT coincide (and higher-ranked linearization constraints do not differentiate between the candidates), as in (33-a-d), the unmarked candidate has IO $\succ$ DO (the marked one is rendered grammatical under substitution of any constraint on which the two candidates do not differ, e.g., of NOM); cf. tableau T$_6$. If, however, AN and DAT conflict, as in (33-ef), the unmarked candidate respects the former constraint and violates the latter (cf. Fanselow (1995)); cf. tableau T$_7$ (note that the marked candidate K$_1$ emerges as grammatical only under substitution of DAT or PER).

Tableau T$_6$: AN > DAT – 'gönnen', 'überlassen' etc.:

Candidates	SCR-CRIT							STAY	PAR-M
	NOM	DEF	AN	FOC	DAT	ADV	PER		
⇒→K$_1$: dem Fritz$_2$ das Buch$_1$ t$_2$								*	*
⇒ K$_2$: das Buch$_1$ dem Fritz$_2$			*?		*		*		

Tableau T$_7$: AN > DAT – 'entziehen' etc.:

Candidates	SCR-CRIT							STAY	PAR-M
	NOM	DEF	AN	FOC	DAT	ADV	PER		
⇒ K$_1$: dem Einfluß$_2$ die Kinder$_1$ t$_2$			*?					*	*
⇒→K$_2$: die Kinder$_1$ dem Einfluß$_2$					*		*		

Tableau T$_6$ also highlights a general property of the present approach to markedness that is worth emphasizing: Scrambling does not automatically lead to markedness, as presupposed under the DTC-based approach, but can in fact apply so as to create the unmarked candidate (K$_1$, in this case) – this holds if the application of the movement operation results in an optimal constraint profile (with respect to (30)).

    Consider now the case where the animacy factor is excluded, i.e., where the candidates exhibiting the two object orders do not differ with respect to AN because both objects are [+animate]. All other things being equal (in particular, assuming that focus is not involved, see below), we expect that the decision (concerning unmarkedness) is passed on

to the lower-ranked constraint DAT, and that the unmarked candidate will always be the one in which IO scrambling across the DO has taken place. As shown by the data in (34), this prediction is borne out:

(34) a. daß er dem Fritz$_2$ die Maria$_1$ t$_2$ empfahl/vorstellte
          that he ART Fritz$_{dat}$ ART Maria$_{acc}$ recommended/introduced
      b. ?daß er die Maria$_1$ dem Fritz$_2$ empfahl/vorstellte
          that he ART Maria$_{acc}$ ART Fritz$_{dat}$ recommended/introduced
      c. daß man der Claudia$_2$ diesen Mann$_1$ t$_2$ zeigen/gönnen/verweigern muß
          that one ART Claudia$_{dat}$ this man$_{acc}$ show/not begrudge/deny must
      d. ?daß man diesen Mann$_1$ der Claudia$_2$ zeigen/gönnen/verweigern muß
          that one this man$_{acc}$ ART Claudia$_{dat}$ show/not begrudge/deny must
      e. daß man der Maria$_2$ die Kinder$_1$ t$_2$ entzogen/ zugeführt/
          that one ART Maria$_{dat}$ the children$_{acc}$ taken away from/ brought to/
          ausgeliefert hat
          extradited has
      f. ?daß man die Kinder$_1$ der Maria$_2$ entzogen/ zugeführt/
          that one the children$_{acc}$ ART Maria$_{dat}$ taken away from/ brought to/
          ausgeliefert hat
          extradited has

The sentence pairs (34-ab) and (34-cd) exhibit the same pattern as their counterparts in (33): IO precedes DO in the unmarked case. The underlying competition is illustrated in tableau T$_8$. More interesting in the present context are the data in (34-ef), with a verb like <u>entziehen</u>. Here, the unmarked order is not DO $\succ$ IO (as in the counterparts in (33)), but rather IO $\succ$ DO. This is shown in tableau T$_9$.

Tableau T$_8$: DAT – 'gönnen' etc.:

Candidates	SCR-CRIT							STAY	PAR-M
	NOM	DEF	AN	FOC	DAT	ADV	PER		
$\Rightarrow \rightarrow$K$_1$: d. Claudia$_2$ d. Mann$_1$ t$_2$								*	*
$\Rightarrow$ K$_2$: d. Mann$_1$ d. Claudia$_2$					*?		*		

Tableau T$_9$: DAT – 'entziehen' etc.:

Candidates	SCR-CRIT							STAY	PAR-M
	NOM	DEF	AN	FOC	DAT	ADV	PER		
$\Rightarrow \rightarrow$K$_1$: der Maria$_2$ die Kinder$_1$ t$_2$								*	*
$\Rightarrow$ K$_2$: die Kinder$_1$ der Maria$_2$					*?		*		

If a verb like <u>entziehen</u> selects an IO that is [+animate], and a DO that is [–animate], as is prototypically the case with a verb like <u>geben</u>, it behaves as expected – both AN and DAT favour the order IO $\succ$ DO:

(35) a. daß man dem Fritz$_2$ das Buch$_1$ t$_2$ entzogen hat
          that one ART Fritz$_{dat}$ the book$_{acc}$ taken away from has

b. ?daß man das Buch$_1$ dem Fritz$_2$ entzogen hat
   that one the book$_{acc}$ ART Fritz$_{dat}$ taken away from has

Thus, the ranking AN > DAT appears well motivated. More generally, we end up with the following picture concerning alleged verb classes in German: Instead of different base orders for different verbs, there is a uniform base order DO ≻ IO; IO ≻ DO orders are derived by scrambling. IO scrambling leads to the unmarked candidate in what is probably the majority of cases, and this is so for two reasons: First, DAT directly favours IO ≻ DO orders. And second, since verbs of giving and taking in a broader sense form the largest set of double object verbs in German (cf. Eisenberg (1986, 288)), and since these verbs usually realize the DO as [–animate], and the IO as [+animate], AN tends to favour IO ≻ DO orders as well, albeit indirectly.[22]

6.3.4. Focus vs. Animacy and Dative

In (27), FOC intervenes between AN and DAT. Consequently, it must be shown that AN > FOC, and that FOC > DAT. First, evidence for the ranking AN > FOC can be gained from the data in (36), which are versions of (33-f) ((36-ab)) and (33-e) ((36-cd)), with DO, IO as focus exponents (focus is indicated by capital letters). The competition is illustrated in tableau T$_{10}$.

---

[22]In principle, we also expect unmarked IO ≻ DO order to arise if both IO and DO are [–animate]. In many cases, DAT does indeed correctly single out the unmarked candidate (assuming that higher-ranked constraints are irrelevant for the decision); cf., e.g., (i-ab):

(i)   a.   daß man dem Film$_2$ den Oscar$_1$ verliehen hat
            that one the film$_{dat}$ the Oscar$_{acc}$ given has
      b.   ?daß man den Oscar$_1$ dem Film$_2$ verliehen hat
            that one the Oscar$_{acc}$ the film$_{dat}$ given has

However, sentence pairs like that in (ii) potentially pose a problem:

(ii)   a.   daß man das Buch der Prüfung unterzogen hat
            that one the book the test subjected to has
      b.   ??daß man der Prüfung das Buch unterzogen hat
            that one the test the book subjected to has

Here, DO precedes IO in the unmarked case, even though DAT is violated by this order (and AN is irrelevant). A possible explanation of this observation might rely on the fact that many of the pertinent examples involve light verb constructions (der Prüfung unterziehen ('examine') in the case at hand, similarly for der Gefahr aussetzen ('expose to danger')). As idioms, these constructions tend to resist syntactic split. This could be integrated into an optimality-theoretic system as a violable constraint that outranks linearization constraints like DAT. Alternatively, we might assume that the dative assigned in complex idiomatic expressions is always lexical, rather than structural, and DAT affects only NPs that bear structural Case (Gisbert Fanselow (p.c.)).

(36) a. daß man die Kinder$_1$ diesem EINFLUSS$_2$ entzogen/ ausgesetzt
     that one the children$_{acc}$ this influence$_{dat}$ taken away from/ exposed
     hat
     has

b. ?daß man die KINDER$_1$ diesem Einfluß$_2$ entzogen/ausgesetzt hat

c. ??daß man diesem Einfluß$_2$ die KINDER$_1$ t$_2$ entzogen/ausgesetzt hat

d. ?*daß man diesem EINFLUSS$_2$ die Kinder$_1$ t$_2$ entzogen/ausgesetzt hat

Tableau T$_{10}$: AN > FOC:

Candidates	SCR-CRIT							STAY	PAR-M
	NOM	DEF	AN	FOC	DAT	ADV	PER		
⇒→K$_1$: d. Kinder$_1$ d. EINFL.$_2$					*		*		
⇒ K$_2$: d. KINDER$_1$ d. Einfl.$_2$				*?	*		*		
⇒ K$_3$: d. Einfl.$_2$ d. KINDER$_1$ t$_2$			*?					*	*
⇒ K$_4$: d. EINFL.$_2$ d. Kinder$_1$ t$_2$			*?	*?				*	*

Second, evidence for the ranking FOC > DAT comes from data like those in (37), which have often been discussed in the literature (cf. Lenerz (1977, 39ff) and Uszkoreit (1984, 182ff; 1986, 895-898), among others).

(37) a. daß man der Maria$_2$ den FRITZ$_1$ t$_2$ vorstellte/ zeigte/ zum
     that one ART Maria$_{dat}$ ART Fritz$_{acc}$ introduced/ showed/ as
     Mann gab
     husband gave

b. ?daß man den Fritz$_1$ der MARIA$_2$ vorstellte/ zeigte/ zum Mann gab

c. ??daß man der MARIA$_2$ den Fritz$_1$ t$_2$ vorstellte/ zeigte/ zum Mann gab

d. ?*daß man den FRITZ$_1$ der Maria$_2$ vorstellte/ zeigte/ zum Mann gab

Tableau T$_{11}$ shows that (37-a) (K$_1$) is grammatical and unmarked; that (37-b) (K$_2$) (which respects FOC but violates DAT) is grammatical (under, e.g., FOC substitution), but more marked; that (37-c) (K$_3$) is grammatical (under, e.g., DAT substitution), but, crucially, yet more marked than (37-b); and that (37-d) (K$_4$) is grammatical (under, e.g., NOM substitution – the candidate does not involve scrambling), but clearly more marked than all its grammatical competitors (since it has the worst constraint profile, violating both FOC and DAT).[23]

---

[23]Note that there does not seem to be a context type in which (37-d) could be the optimal realization (not even with contrastive focus, as claimed by Choi (1996)). Consequently, this sentence must be viewed as ungrammatical in the optimality theoretic approaches to word order that are developed by Choi (1996), Büring (1997), and Costa (1998), which define candidate sets with respect to identity of discourse functions (cf. footnote 15). However, this result appears to be at variance with the facts. As noted by Uszkoreit (1986, 896), sentences of this type "still appear more acceptable than sentences in which fixed-order constraints are violated" (where, in the present framework, fixed-order constraints are those that belong to the matrix hierarchy, like PRON-CRIT).

Tableau T_{11}: FOC > DAT:

Candidates	SCR-CRIT							STAY	PAR-M
	NOM	DEF	AN	FOC	DAT	ADV	PER		
⇒→K$_1$: d. Maria$_2$ d. FRITZ$_1$ t$_2$								*	*
⇒ K$_2$: d. Fritz$_1$ d. MARIA$_2$					*?		*		
⇒ K$_3$: d. MARIA$_2$ d. Fritz$_1$ t$_2$				*?				*	*
⇒ K$_4$: d. FRITZ$_1$ d. Maria$_2$				*?	*?		*		

6.3.5. Adverbs

Finally, let me present some evidence for the linearization constraint ADV. As is well known, certain types of adverbs (e.g., locative adverbs) prefer to show up to the right of definite NP arguments (cf. Lenerz (1977), Haftka (1981), and Suchsland (1993), among many others). Thus, whereas all the examples in (38) (involving a nominative NP and an accusative NP) seem to be perfectly grammatical as such, only (38-a), with order-preserving VP-adjunction of both NPs above the VP-adjoined adverb, is unmarked; (38-b) is more marked than (38-a), and so forth:

(38) a. daß der Fritz$_1$ die Maria$_2$ in der Kneipe t$_1$ t$_2$ getroffen hat
          that ART Fritz$_{nom}$ ART Maria$_{acc}$ in the pub   met   has
      b. ?daß der Fritz$_1$ in der Kneipe t$_1$ die Maria$_2$ getroffen hat
      c. ??daß in der Kneipe der Fritz$_1$ die Maria$_2$ getroffen hat
      d. ?*daß die Maria$_2$ der Fritz$_1$ in der Kneipe t$_1$ t$_2$ getroffen hat

Tableau T_{12} illustrates the competition:[24]

Tableau T_{12}: ADV:

Candidates	SCR-CRIT							STAY	PAR-M
	NOM	DEF	AN	FOC	DAT	ADV	PER		
⇒→K$_1$: NP$_1$ NP$_2$ Adv t$_1$ t$_2$							*	**	
⇒ K$_2$: NP$_1$ Adv t$_1$ NP$_2$						*?	*	*	
⇒ K$_3$: Adv NP$_1$ NP$_2$						*?*?	*		
⇒ K$_4$: NP$_2$ NP$_1$ Adv t$_1$ t$_2$	*?							**	*

---

[24]This is only a rough indication of the direction that an articulated analysis might take. Other adverb types, e.g., may behave differently from locatives; similarly, the picture changes if indefinite NPs are taken into consideration. Cf. Müller (1998a) for a more comprehensive account. Furthermore, an interesting problem arises here because, in contrast to deriving the correct markedness degrees, the task of making sure that all candidates are grammatical in the first place is somewhat involved. K$_1$, K$_3$, and K$_4$ are grammatical under ADV, NOM, and PER substitution, respectively. However, it seems as though K$_2$ is not grammatical under substitution of any of the linearization constraints. This problem is solved if ADV is interpreted via reflexive local conjunction (cf. Legendre Smolensky, & Wilson (1998, 262)). Local conjunction of ADV with itself yields a new constraint ADV' which K$_3$ violates (because of two simple ADV violations) whereas K$_2$ does not. Thus, K$_2$ is grammatical under ADV' substitution.

6.4. String-Vacuous Scrambling

Thus far, we have seen that it is quite easy for candidates with different VP-internal word orders in German to be grammatical (and much harder to be unmarked); i.e., the trigger for scrambling as such is fairly weak. However, if scrambling does not lead to an improved behaviour towards any of the linearization constraints of SCR-CRIT, we expect ungrammaticality to arise. A classic case in point is the issue of string-vacuous scrambling. As has often been noted (cf., e.g., Ross (1967) and Frey & Tappe (1991)), one wants to rule out iterated scrambling to VP that does not have any effect; cf. (39-a) vs. (39-b):

(39) a. daß [$_{VP}$ die   Maria$_1$ [$_{V'}$ den   Fritz$_2$ geküßt hat]]
          that   ART Maria$_{nom}$ ART Fritz$_{acc}$ kissed has
      b. *daß [$_{VP}$ die Maria$_1$ [$_{VP}$ den Fritz$_2$ [$_{VP}$ t$_1$ [$_{V'}$ t$_2$ geküßt hat]]]]

As shown in tableau T$_{13}$, string-vacuous scrambling is straightforwardly blocked under present assumptions. Except for STAY, the two derivations have exactly the same constraint profile, and so STAY becomes decisive and rules out candidate K$_2$ in favour of the more economic K$_1$ – there is no "motivation" for string-vacuous scrambling, and hence it is excluded by economy.[25]

Tableau T$_{13}$: String-vacuous scrambling:

Candidates	Scr-Crit							Stay	Par-M
	Nom	Def	An	Foc	Dat	Adv	Per		
⇒K$_1$: NP$_1$ NP$_2$							*		
*K$_2$: NP$_1$ NP$_2$ t$_1$ t$_2$							*	*!*	

7. Cross-Linguistic Variation

Having provided solutions to the first three problems mentioned in section 1 (economy, markedness, and fixed pronoun order), let me now indicate how the last problem (parametrization) can be addressed. First, it is worth pointing out that it seems wrong to assume that there is a parameter ±scrambling: Languages like Dutch and Icelandic exhibit more freedom of clause-internal word order than English, but considerably less than German (which in turn has less word order variability than Russian, see footnote 3). For instance, Dutch and Icelandic permit IO and DO to precede or follow adjuncts as long as the D-structure order of arguments is preserved (Icelandic only under typical conditions for object shift, i.e., if the main verb has left the VP); however, abstracting away from a few intervening factors (like unaccusative and so-called "focus-scrambling" constructions in Dutch), a VP-internal permutation of arguments is impossible (cf., e.g., Koster (1986)

---

[25]In this context, it might also be worth pointing out that the present approach does not predict the co-existence of various grammatical candidates with different VP-internal NP orders to be always possible. As soon as there is a SCR-CRIT-independent constraint that is violated by one candidate and satisfied by another, optionality will break down. Relevant interfering constraints include those on weak crossover and binding theory – in (3) and (4), e.g., the order of DO and IO is strict, due to principle A.

and Neeleman (1994) on Dutch, and Collins & Thráinsson (1996) on Icelandic). Thus, scrambling seems to exist in Dutch and Icelandic as it exists in German, but it must be order-preserving.

Ideally, we would hope that cross-linguistic variation in this domain results from a reranking of constraints, which Prince & Smolensky (1993) identify as the major source of parametrization. Indeed, if SCR-CRIT dominates STAY (as in German) but is dominated by PAR-MOVE, languages of the Dutch/Icelandic type are predicted: Now all candidates that satisfy SCR-CRIT by changing the argument order are filtered out as ungrammatical by PAR-MOVE, whereas candidates that satisfy SCR-CRIT and at the same time maintain the argument order (e.g., by moving an NP across an adverb) are not.[26] Next, if the ranking differs from Dutch/Icelandic only in that SCR-CRIT is also dominated by STAY, languages of the English type are predicted, which prohibit all kinds of scrambling. Thus, on this approach, SCR-CRIT as a subhierarchy of linearization constraints is present as a constraint in English as it is in German; however, its effects are blurred by a higher-ranked STAY.[27]

A direct consequence of this approach to variation in scrambling options is that parametrization is not tied to other, independently observed properties of a given language: The extent to which a language exhibits clause-internal free word order depends on the relative ranking of SCR-CRIT, PAR-MOVE, and STAY; it does not depend on, e.g., the presence of an articulated system of morphological Case (cf., e.g., Fanselow (1993)), or on the availability of V government in the canonical direction (cf. Webelhuth (1987), Reuland & Kosmeijer (1993), Haider (1993), and Deprez (1994)).[28] To integrate these latter ideas into the analysis presented here, one might introduce meta-constraints that restrict possible rankings. For instance, one might stipulate that SCR-CRIT ≫ STAY only in languages in which canonical government by V is available for scrambling positions; or that SCR-CRIT ≫ PAR-MOVE only if a rich system of Case morphology is present in the language. However, modifications of this type introduce further machinery that one would ideally want to do without; and they also threaten to undermine the standard optimality theoretic concept of parametrization. Fortunately, such a move is empirically inadequate to begin with: There are languages that allow scrambling of the German type but exhibit neither OV order, nor a rich system of Case morphology. One such language

---

[26]It turns out that such an analysis of Dutch and Icelandic in addition requires certain assumptions about the D-structure order of Dutch and Icelandic as opposed to German; cf. Müller (1999).

[27]It is unclear to me whether reranking can also apply within the SCR-CRIT subhierarchy, thus yielding language types which could, e.g., minimally differ from German by ranking DAT higher than AN. This would yield the consequence that a verb like entziehen ('take away from') has IO ≻ DO as the unmarked order in both T_7 and T_9. Interestingly, if such reranking turned out not to be possible, this would correspond to what is the case with the concept of subhierarchy adopted by Baković (1998) and Legendre, Smolensky, & Wilson (1998) (which is otherwise radically different; cf. footnote 16).

[28]Assuming that scrambling can only apply if the scrambled NP ends up in a position that is governed by V in the canonical direction, it would follow that OV languages generally permit scrambling, at least to some degree (German, Dutch vs. English, Danish), and that VO languages permit scrambling if they have V-to-I(T) movement (Icelandic vs. English, Danish) – if V is in I(T), scrambled objects are still canonically governed.

is Bulgarian, as discussed in Molxova (1970, 27), Rudin (1985, 13-39), and Müller (1995, 110). Bulgarian is a VO language whose Case system is impoverished (in fact, it closely resembles the English Case system); however, its scrambling options are very much like those of German (e.g., the D-structure order of arguments does not have to be preserved). Thus, it seems fair to conclude that, at present, there is no theory that correctly predicts the emergence of free word order phenomena in a given language on a purely synchronic basis, and an approach in terms of reranking is not called into question by the fact that it does not rely on a correlation of scrambling and other (morphological) properties.[29]

8.   Concluding Remarks

To end this article, let me summarize the main findings and point out some further consequences. I have tried to show that optimality theory makes it possible to develop a precise and testable account of free word order in German that respects the main insights behind traditional competition-based models: The interaction of factors like definiteness, animacy, focus, and Case is re-interpreted as the interaction of linearization constraints of a subhierarchy that are violable and ranked. This re-interpretation has made it necessary to modify optimality theory by adding a system of subhierarchies, with a specific semantics that differs from that required for the matrix hierarchy. However, this step leaves all previous reasonings in standard optimality theory intact – except for subhierarchies, everything works as in Prince & Smolensky (1993). Thus, if what precedes is tenable in its essentials, there is no reason to assume that clause-internal word order variation must be addressed by theoretical means that are largely unrelated to what is otherwise documented in grammatical theory, or that it cannot be addressed by grammatical theory at all.

More specifically, the approach developed here provides solutions to some fundamental problems that arise if we adopt a scrambling operation (but closer inspection reveals that these problems show up in some form under a base-generation approach as well): the economy problem ('What is the trigger for scrambling?'), the markedness problem ('Why do free word order structures often exhibit degrees of markedness rather than complete wellformedness or illformedness?'), the fixed order domain problem ('Why do weak pronouns in the Mittelfeld exhibit fixed order when other NPs don't?'), and the parametrization problem ('Why, and to what extent, do languages differ with respect to scrambling options?').

Needless to say, the present approach raises a lot of further issues, which I cannot possibly address here in any detail. First, the obvious question arises of whether there are constraints other than SCR-CRIT that take the form of subhiearchies. For conceptual reasons alone, it seems to me that there should be. At first sight, it looks as though a

---

[29]Perhaps this result is indicative of a more general problem with morphology-based approaches to cross-linguistic variation in syntax. In line with what is proposed here, Grimshaw & Samek-Lodovici (1995) and Samek-Lodovici (1996) argue that the pro-drop parameter should in fact not be reduced to the morphological richness of subject (or, for that matter, object) inflection, but to constraint reranking. Also see Vikner (1999) for an approach to the V-to-I(T) raising parameter that relies on constraint reranking rather than the strength of verbal inflection.

subhierarchy is in fact nothing but a different notion of constraint tie. Ties and subhierarchies both create optionality (so that more than one candidate per candidate set can be optimal). Indeed, subhierarchies as defined here closely correspond to one notion of tie discussed in Broihier (1995) (which he calls "logical or" tie); cf. Müller (1999a). Furthermore, it is often the case that two or more optimal candidates that emerge as the result of a tie are not equally unmarked (cf. in particular Sells, Rickford, & Wasow's (1996, 614) analysis of inversion in AAVE). However, in many other cases (like, e.g., in the case of the EPP/STAY tie proposed in section 5), there is no discernible markedness difference among optimal candidates resulting from constraint ties. More importantly, subhierarchies as employed here have the effect of completely removing subconstraints from the constraint ranking on the matrix hierarchy, as regards the determination of grammaticality. Thus, if we were to completely replace constraint ties by subhierarchies, this would yield potentially disastrous consequences. In the case at hand, STAY would be completely irrelevant (as far as grammaticality is concerned) for candidates that respect EPP in German, an untenable result. For these reasons, I conclude that subhierarchies cannot take over the role of ties in optimality-theoretic syntax. However, this does not mean that SCR-CRIT is the only possible instantiation of a subhierarchy. Another application that strikes me as worth exploring concerns extraction theory; in particular, argument extraction from weak islands often yields intermediate degrees of acceptability, which is not straightforwardly accountable for under standard optimality theory, but might be amenable to an analysis along the lines sketched here.

Second, empirical issues arise. Let me emphasize again that what precedes is mainly designed as a model in which the constraints on clause-internal word order variation can be reconciled with the constraints otherwise adopted in grammar; as noted, the specific proposals would demand verification by further empirical, psycholinguistic studies.[30] Also recall that I have systematically neglected clause-internal word order variation in certain constructions.[31] What is more, it appears as though the linearization effects investigated here are not strictly confined to the VP domain (i.e., scrambling). They do not show up with pronoun movement to Specπ, but topicalization sometimes seems to obey similar markedness restrictions, and related issues arise with (multiple) extraposition of NPs. Again, an extension of the present approach seems promising, but it is beyond the scope of this article.

Finally, it has sometimes been proposed in the competition-based literature mentioned in section 1 that, in addition to the quality of linearization constraint violations, the quantity of these violations might be relevant. At least for the data discussed here, it seems to me that such a cumulative concept is not necessary (in fact, if naively applied, not even possible – recall footnote 21). Of course, the concept of cumulative violation could in

---

[30]See, e.g., McDonald, Bock, & Kelly (1993) for a study that concentrates on AN. Also, relevant work is currently being carried out by research groups of Hans Jürgen Uszkoreit and Rainer Dietrich, at Universität des Saarlandes and Humboldt-Universität Berlin, respectively.

[31]This holds, e.g., for constructions that involve psych verbs or unaccusative verbs. These domains are covered by the optimality-theoretic approach to German word order in Heck (1999, 3.4), which locates the optimization of unmarked orders at D-structure (not at S-structure, as I am doing here).

principle be integrated into the definition of markedness adopted here (cf. (30)) without too much ado (and without simultaneously affecting the determination of grammaticality). However, this would reduce the convergence of constraints on the matrix hierarchy and constraints on the subhierarchy which, as it stands, strikes me as significant. Thus, in the absence of clear empirical support for such a move, I take it that matrix/subhierarchy differences (i.e., differences between "weak" and "strong" rules) should be minimized, so as to permit a unified approach.

9.   References

Ackema, Peter, & Ad Neeleman (1998). Optimal Questions. Natural Language and Linguistic Theory 16, 443-490.

Baković, Eric (1998). Optimality and Inversion in Spanish. In Is the Best Good Enough?, eds. Pilar Barbosa et al., 35-58. Cambridge, Mass.: MIT Press & MITWPL.

Bayer, Josef, & Jaklin Kornfilt (1994). Against Scrambling as an Instance of Move-α. In Studies on Scrambling, ed. Norbert Corver & Henk van Riemsdijk, 17-60. Berlin: Mouton de Gruyter.

Broihier, Kevin (1995). Optimality Theoretic Rankings with Tied Constraints: Slavic Relatives, Resumptive Pronouns and Learnability. Ms., MIT, Cambridge, Mass.

Büring, Daniel (1997). Towards an OT Account of German Mittelfeld Word Order. Ms., Universität Köln.

Choi, Hye-Won (1996). Optimizing Structure in Context: Scrambling and Information Structure. Doctoral dissertation, Stanford University.

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

Chomsky, Noam (1982). Some Concepts and Consequences of the Theory of Government and Binding. Cambridge, Mass.: MIT Press.

Chomsky, Noam (1991). Some Notes on Economy of Derivation and Representation. In Principles and Parameters in Comparative Grammar, ed. Robert Freidin, 417-454. Cambridge, Mass.: MIT Press.

Chomsky, Noam (1995). Categories and Transformations. Chapter 4 of The Minimalist Program, 219-394. Cambridge, Mass.: MIT Press.

Collins, Chris, & Höskuldur Thráinsson (1996). VP-Internal Structure and Object Shift in Icelandic. Linguistic Inquiry 27, 391-444.

Corver, Norbert, & Henk van Riemsdijk (eds) (1994). Studies on Scrambling. Berlin: Mouton de Gruyter.

Costa, João (1998). Word Order Typology in Optimality Theory. Ms., Universidade de Lisboa & University of Leiden. To appear in Optimality Theoretic Syntax, eds. Jane Grimshaw, Géraldine Legendre, & Sten Vikner. Cambridge, Mass.: MIT Press.

Daneš, František (1967). Order of Elements and Sentence Intonation. In To Honor Roman Jakobson, vol (1, 499-512. The Hague/Paris: Mouton.

Deprez, Viviane (1994). Parameters of Object Movement. In Studies on Scrambling, ed. Norbert Corver & Henk van Riemsdijk, 101-152. Berlin: Mouton de Gruyter.

Diesing, Molly (1992). Indefinites. Cambridge, Mass.: MIT Press.

Dietrich, Rainer (1994). Wettbewerb – aber wie? Skizze einer Theorie der freien Wortstellung. In Was determiniert die Wortstellungsvariation?, ed. Brigitta Haftka, 33-47. Opladen: Westdeutscher Verlag.

Dietrich, Rainer (1999). On the Production of of Word Order and the Origin of Incrementality. In Representations and Processes in Language Production, ed. Ralf Klabunde & Christiane von Stutterheim. Amsterdam: Benjamins.

Eisenberg, Peter (1986). Grundzüge der deutschen Grammatik. Stuttgart: Metzler.

Fanselow, Gisbert (1988). German Word Order and Universal Grammar. In Natural Language Parsing and Linguistic Theories, ed. Uwe Reyle & Christian Rohrer, 317-355. Dordrecht: Kluwer.

Fanselow, Gisbert (1993). The Return of the Base Generators. Ms., Universität Passau.

(Also in Groninger Arbeiten zur Germanistischen Linguistik 36, 1-74.)

Fanselow, Gisbert (1995). A Minimalist Approach to Free Constituent Order. Ms., Universität Potsdam.

Fanselow, Gisbert (1997). Features, Θ-Roles, and Free Constituent Order. Ms., Universität Potsdam.

Fortmann, Christian, & Werner Frey (1997). Konzeptuelle Struktur und Grundabfolge der Argumente im Deutschen. In Zur Satzstruktur im Deutschen, ed. Franz-Josef d'Avis & Uli Lutz, 143-170. Tübingen: Arbeitspapiere des SFB 340.

Frey, Werner, & Thilo Tappe (1991). Zur Interpretation der X-bar-Theorie und zur Syntax des Mittelfeldes. Grundlagen eines GB-Fragmentes. Ms., Universität Stuttgart.

Grewendorf, Günther (1988). Aspekte der deutschen Syntax. Tübingen: Narr.

Grewendorf, Günther (1989). Ergativity in German. Dordrecht: Foris.

Grewendorf, Günther, & Joachim Sabel (1999). Scrambling in German and Japanese: Adjunction vs. Multiple Specifiers. Natural Language and Linguistic Theory 17, 1-65.

Grimshaw, Jane (1997). Projection, Heads, and Optimality. Linguistic Inquiry 28, 373-422.

Grimshaw, Jane, & Vieri Samek-Lodovici (1995). Optimal Subjects. In Papers in Optimality Theory, ed. Jill Beckmann et al., 589-605. University of Massachusetts, Amherst: UMass Occasional Papers in Linguistics 18.

Haftka, Brigitta (1981). Reihenfolgebeziehungen im Satz. In Grundzüge einer deutschen Grammatik, ed. Karl-Erich Heidolph et al., 702-764. Berlin: Akademieverlag.

Haider, Hubert (1989). Θ-Tracking Systems – Evidence from German. In Configurationality, ed. Laszlo Maracz & Pieter Muysken, 185-206. Dordrecht: Foris.

Haider, Hubert (1992). Branching and Discharge. Ms., Universität Stuttgart.

Haider, Hubert (1993). Deutsche Syntax, generativ. Tübingen: Narr.

Haider, Hubert, & Inger Rosengren (1998). Scrambling. Sprache und Pragmatik 49. Lund: Germanistisches Institut.

Heck, Fabian (1999). Tiefenoptimierung. Ms., Universität Stuttgart.

Hoberg, Ursula (1981). Die Wortstellung in der geschriebenen deutschen Gegenwartssprache. München: Hueber.

Höhle, Tilman (1982). Explikation für "normale Betonung" und "normale Wortstellung". In Satzglieder im Deutschen, ed. Werner Abraham, 75-153. Tübingen: Narr.

Holmberg, Anders (1986). Word Order and Syntactic Features in the Scandinavian Languages and English. Doctoral dissertation, Stockholm: Department of General Linguistics.

Hoop, Helen de (1992). Case Configuration and Noun Phrase Interpretation. Doctoral dissertation, University of Groningen.

Jackendoff, Ray (1972). Semantic Interpretation in Generative Grammar. Cambridge, Mass.: MIT Press.

Jacobs, Joachim (1988). Probleme der freien Wortstellung im Deutschen. Sprache und Pragmatik 5, 8-37.

Johnson, Kyle (1991). Object Positions. Natural Language and Linguistic Theory 9, 577-636.

Keller, Frank (1996). Extraction from Complex Noun Phrases. A Case Study in Graded

Grammaticality. Diplomarbeit, Universität Stuttgart.

Kidwai, Ayesha (1997). Scrambling and Binding in Hindi-Urdu. Doctoral dissertation, University of New Delhi.

Koster, Jan (1986). The Relation between pro-drop, Scrambling, and Verb Movements. Ms., Rijksuniversiteit Groningen.

Koster, Jan (1987). Domains and Dynasties. Dordrecht: Foris.

Kroch, Anthony (1974). The Semantics of Scope in English. Doctoral dissertation, MIT, Cambridge, Mass.

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

Lasnik, Howard, & Mamoru Saito (1992). Move α. Cambridge, Mass.: The MIT Press.

Legendre, Géraldine, Paul Smolensky, & Colin Wilson (1998). When is Less More? Faithfulness and Minimal Links in Wh-Chains. In Is the Best Good Enough?, eds. Pilar Barbosa et al., 249-289. Cambridge, Mass.: MIT Press & MITWPL.

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

Lenerz, Jürgen (1992). Zur Syntax der Pronomina im Deutschen. Sprache und Pragmatik 29, University of Lund.

Lerot, Jacques (1985). Zur Wortstellungsnorm im Deutschen. Deutsche Sprache, 137-142.

Lötscher, Andreas (1981). Abfolgeregeln für Ergänzungen im Mittelfeld. Deutsche Sprache, 44-60.

McCarthy, John, & Alan Prince (1995). Faithfulness and Reduplicative Identity. In Papers in Optimality Theory, ed. Jill Beckman, Laura Walsh-Dickie & Suzanne Urbanczyk, 249-384. Amherst, Massachussetts: UMass Occasional Papers in Linguistics 18.

McDonald, Janet, Kathryn Bock, & Michael H. Kelly (1993). Word and World Order: Semantic, Phonological, and Metrical Determinants of Serial Position. Cognitive Psychology 25, 188-230.

May, Robert (1985). Logical Form. Cambridge, Mass.: MIT-Press.

Meinunger, André (1995). Discourse-Dependent DP (De-)Placement. Doctoral dissertation, Universität Potsdam.

Molxova, Zana (1970). Xarakter i upotreba na člena v bylgarskija i anglijskija ezik. Sofia: NI.

Müller, Gereon (1995). A-bar Syntax. Berlin: Mouton de Gruyter.

Müller, Gereon (1997). Partial Wh-Movement and Optimality Theory. The Linguistic Review 14, 249-306.

Müller, Gereon (1998). Incomplete Category Fronting. Dordrecht: Kluwer.

Müller, Gereon (1998a). Zur Ableitung der NP-Adv-V-Stellung im Deutschen. Ms., Universität Stuttgart.

Müller, Gereon (1999). Order Preservation, Parallel Movement, and the Emergence of the Unmarked. Ms., Universität Stuttgart. To appear in Optimality-Theoretic Syntax, eds. Géraldine Legendre, Jane Grimshaw, & Sten Vikner. Cambridge, Mass.: MIT Press.

Müller, Gereon (1999a). Optionality in Optimality-Theoretic Syntax. GLOT International 4.5, 3-8.

Müller, Gereon, & Wolfgang Sternefeld (1994). Scrambling as A-bar Movement. In

Studies on Scrambling, ed. Norbert Corver & Henk van Riemsdijk, 331-385. Berlin: Mouton de Gruyter.

Neeleman, Ad (1994). Scrambling as a D-structure Phenomenon. In Studies on Scrambling, ed. Norbert Corver & Henk van Riemsdijk, 387-429. Berlin: Mouton de Gruyter.

Neeleman, Ad, & Tanya Reinhart (1997). The Syntax and Interpretation of Scrambling in Dutch. Ms., University of Utrecht.

Pesetsky, David (1998). Some Optimality Principles of Sentence Pronunciation. In Is the Best Good Enough?, eds. Pilar Barbosa et al., 337-383. Cambridge, Mass.: MIT Press & MITWPL.

Primus, Beatrice (1994). Grammatik und Performanz: Faktoren der Wortstellungsvariation im Mittelfeld. Sprache und Pragmatik 32, 39-86.

Prince, Alan & Paul Smolensky (1993). Optimality Theory. Constraint Interaction in Generative Grammar. Ms., Rutgers University.

Reinhart, Tanya, & Eric Reuland (1993). Reflexivity. Linguistic Inquiry 24, 657-720.

Reis, Marga (1986). Die Stellung der Verbargumente im Deutschen. Stilübungen zum Grammatik:Pragmatik-Verhältnis. Ms., Universität Tübingen. Appeared 1987 in Sprache und Pragmatik, ed. Inger Rosengren, 139-178. Stockholm: Almqvist & Wiksell International.

Reuland, Eric, & Wim Kosmeijer (1993). Projecting Inflected Verbs. In The Parametrization of Universal Grammar, ed. Gisbert Fanselow, 37-71. Amsterdam: Benjamins.

Rizzi, Luigi (1986). On Chain Formation. In The Syntax of Pronominal Clitics, ed. Hagit Borer, 65-95. Orlando: Academic Press.

Rizzi, Luigi (1996). Residual V/2 and the Wh-Criterion. In Parameters and Functional Heads, ed. Adriana Belletti & Luigi Rizzi, 63-90. New York & Oxford: Oxford University Press.

Ross, John (1967). Constraints on Variables in Syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Rudin, Catherine (1985). Aspects of Bulgarian Syntax: Complementizers and Wh-Constructions. Columbus, Ohio: Slavica.

Samek-Lodovici, Vieri (1996). Constraints on Subjects. An Optimality-Theoretic Analysis. Doctoral dissertation, Rutgers University: New Brunswick, New Jersey.

Sauerland, Uli (1997). Scrambling and Interpretability. Ms., MIT, Cambridge, Mass.

Schlesewsky, Matthias, Gisbert Fanselow, & Reinhold Kliegl (1997). The Costs of Wh-Movement in German. Ms., Universität Potsdam.

Sells, Peter, John Rickford, & Thomas Wasow (1996). An Optimality Theoretic Approach to Variation in Negative Inversion in AAVE. Natural Language and Linguistic Theory 14, 591-627.

Siewierska, Anna (1993). On the Interplay of Factors in the Determination of Word Order. In Syntax, vol I, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann, 826-846. Berlin: DeGruyter.

Suchsland, Peter (1993). The Structure of German Verb Projections. In The Parametrization of Universal Grammar, ed. Gisbert Fanselow, 123-143. Amsterdam: Benjamins.

Stechow, Arnim von, & Wolfgang Sternefeld (1988). Bausteine syntaktischen Wissens. Opladen: Westdeutscher Verlag.

Thiersch, Craig (1978). Topics in German Syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Uszkoreit, Hans Jürgen (1984). Word Order and Constituent Structure in German. Doctoral dissertation, University of Austin, Texas. Appeared 1987, Stanford: CSLI.

Uszkoreit, Hans Jürgen (1986). Constraints on Order. Linguistics 24, 883-906.

Vikner, Sten (1990). Verb Movement and the Licensing of NP Positions in the Germanic Languages. Doctoral dissertation, Université de Genève. Revised version appeared 1995 as Verb Movement and Expletive Subjects in the Germanic Languages. Oxford: Oxford University Press.

Vikner, Sten (1999). V-to-I Movement and 'Do'-Insertion in Optimality Theory. To appear in Optimality-Theoretic Syntax, eds. Géraldine Legendre, Jane Grimshaw, & Sten Vikner. Cambridge, Mass.: MIT Press.

Vogel, Ralf, & Markus Steinbach (1995). On the (Absence of a) Base Position for Dative Objects in German. FAS Papers in Linguistics 4, 99-131.

Vogel, Ralf, & Markus Steinbach (1998). The Dative – an Oblique Case. Linguistische Berichte 173, 65-90.

Watt, W (1970). On Two Hypotheses Concerning Psycholinguistics. In Cognition and the Development of Language, ed. J. Hayes, 137-220. New York: Wiley.

Webelhuth, Gert (1987). Eine universale Scrambling-Theorie. Paper presented at the annual meeting of the DGfS, Augsburg.

Webelhuth, Gert (1992). Principles and Parameters of Syntactic Saturation. Oxford: Oxford University Press.

Wunderlich, Dieter (1997). Cause and the Structure of Verbs. Linguistic Inquiry 28, 27-68.

Wyngaerd, Guido vanden (1989). Object Shift as an A-Movement Rule. MIT Working Papers in Linguistics 11, 256-271.

Zwart, Jan Wouter (1993). Dutch Syntax. A Minimalist Approach. Doctoral dissertation, University of Groningen.