6. German Word Order Variation

This chapter presents an analysis of word order variation in German, which will be formalized in the following chapter. There exist several accounts of analyzing word order variation, and many of its properties have been unraveled. Research has accumulated substantial evidence for assuming a “scrambling operation”; however, at least three elementary problems arise for such a scrambling operation for traditional grammar approaches. First, a trigger to enforce scrambling is needed, as “economy of movement”, a feature widely accepted in the Generative literature, would block unforced movement. It is not at all clear what such a trigger might be. Second, clause-internal word order variation in languages often exhibits degrees of markedness rather than complete illformedness or wellformedness. The issue of (un-)markedness is not accounted for in traditional approaches. Finally, languages show different reasons or options for changing word order; how can such language-specific variation be explained?

Optimality Theory has brought new possibilities into research to address these problems of scrambling approaches. The system of violable and ranked constraints provides a solution for the first and last of the above problems. The markedness issue remains a problem for pure OT; but proposals were made (cf. Keller 1996) that allow for the treatment of markedness under OT principles.

One of the fundamental strengths of Optimality Theory is its ability to allow for the interaction of different linguistic faculties, e.g. phonology and morphology, or phonology and syntax. For that reason it does not astonish that many analyses of word order variation take into account the interaction of phonology and syntax:
the trigger for clause-internal (forward) movement of constituents is steered by, e.g., accentuation for topicalised elements. This clearly is a striking account for spoken natural language analysis but does not present a valuable framework for computational approaches to syntactic phenomena.

A purely syntactic analysis of German word order variation is proposed in Müller (1998) and will thus be taken as a basis for the formalization. In the following I will present this approach.

### 6.1 Empirical Evidence for Scrambling and Markedness

We have already seen some examples for word order variation in section 3.3. Let us now have a closer look at such phenomena before dealing with Müllers analysis. Consider the following example (Müller 1998, 9):

6.1 a) dass er dem Fritz die Zeitung überliess.
   that he DET Fritz dat the newspaper _akk_ left.

6.1 b) dass er die Zeitung dem Fritz überliess.
   that he the newspaper _akk_ DET Fritz dat left.

Both sentences are grammatical, though (6.1b) is judged to be more marked by many native speakers of German. A possible reason for this judgement could be the unmarked (D-structure) order of the arguments of the verb: in (6.1a), the indirect object precedes the direct (IO > DO), whereas in b) we find the opposite order. We can now assume IO > DO to be the basic order of arguments, and explain the DO > IO order through scrambling. However, consider (6.2) (Müller 1998, 8–9):
Here, the DO > IO order seems to be the base-generated one, contradicting example (6.1). Thus, basic argument order cannot serve as an explanation for deriving markedness phenomena. A second and similar explanation, case, can be rejected with the same arguments as DO/IO above: E.g., accusative not always precedes dative nor does dative precede accusative in all cases. Another option immediately comes to mind when examining the above examples: the word order varies with different types of verbs, so we could ascribe also markedness to this. However, how can we describe semantic differences of verbs in syntax? Obviously, this would entail further unwanted problems we should try to avoid.

We can’t derive the markedness phenomena from clause structure or morphological features like case, nor can we gain it from semantic properties of the verb. But what about semantic features of the arguments? In (6.1) and (6.2), the unmarked sentence shows precedence of animate objects over inanimate; in the marked variation, inanimate objects precede animate. Thus we can identify animacy as a steering device for markedness (and also as a possible trigger for scrambling): *animate* > *inanimate*. Likewise, definiteness can be identified as a feature steering markedness: *definite* > *indefinite* as in (6.3).

Finally, focus also is such a feature, however the precedence relation here is inverse: focussed arguments align with the right edge of the VP, whereas
topicalized (i.e., accentuated or in any other way highlighted) elements align left. Therefore, [-focus] > [+focus].

We have seen that features like definiteness, animacy and focus can influence the word order of arguments in German. In the following these features will be used to construe an OT analysis of word order variation.

### 6.2 Overview and Basic Assumptions

Müller (1998) suggests an approach for motivating word order variation by postulating a “Scrambling Criterion” (SCR-CRIT) constraint ranked higher than the constraint that blocks movement (STAY/*t). SCR-CRIT does not encode a possible trigger for NP movement, but instead is composed of a constraint (sub)hierarchy that incorporates features like definiteness, animacy, focus, etc., driving the scrambling movement. He thus proposes to split up the overall constraint ranking into a matrix hierarchy and a subhierarchy that itself builds a constraint in the matrix hierarchy. This allows for a treatment of word order variation in OT, where principally an inferior candidate is ungrammatical, but in natural language, variation need not necessarily entail ungrammaticality but markedness. For Müller, ungrammaticality arises only with fatal violations of constraints in the matrix hierarchy; the subhierarchy serves for determining markedness. Language-specific parameterization can follow on the one hand from reranking of the matrix hierarchy (i.e. forbidding scrambling by ranking STAY/*t higher than SCR-CRIT) and on the other hand by differences in the scrambling subhierarchy. All of this, of course, presupposes that word order variation does not happen due to differences
in the base generation of the arguments of the verb, as is suggested in other analyses (see Müller 1998, 4 for references regarding base generation).

Müller assumes an underlying clause structure of German (simplified), showed in (6.4):

\[(6.4) [CP – C [TP – [VP Adj [VP SUBJ [V′ DO [V′ IO [V′ OBL V ]]]]]] π ] T]\]

SpecC is the landing site for *wh*-movement, SpecT(ense) for subject raising (optional in German). Spec\(\pi\) is the position assumed to be the landing site for weak pronoun movement, i.e. the so-called “Wackernagel” position. The D-structure order of subject (SUBJ), direct object (DO) and indirect object (IO) always (i.e. with all types of verbs!) is assumed to be as in (6.1). Evidence for this comes from the order of weak pronouns in German clauses.\(^1\) Closest to the base position of the verb are oblique arguments realized by NPs bearing lexical case (e.g., genitive) or by PPs. The word order determined by (6.1) can be varied through movement to a specifier position (SpecC, SpecT, Spec\(\pi\)), or by scrambling. Furthermore, Müller assumes that scrambling in German can only be analyzed as adjunction to VP and not to NP, PP, TP etc., and that it is a property of scrambling to be iterable (in contrast to movement to specifier positions).

---

\(^1\) In the case of three pronominal arguments, their order is fixed as SUBJ > DO > IO: “dass sie\(_{\text{SUBJ}}\) es\(_{\text{DO}}\) ihm\(_{\text{IO}}\) gegeben hat.” Subject pronouns thus obligatorily precede object pronouns, and similarly, a direct object pronoun always prececes a indirect one. (Cf. Müller 1998, 15ff.)
6.3 The Constraints

Let me now turn to the relevant constraints for an Optimality-theoretic approach to scrambling: Müller (1998) distinguished two types of constraints, viz., markedness constraints that trigger movement ("X-Criteria", Müller 1998, 13), and faithfulness constraints that prohibit or minimize the effects of movement. As we have seen according to the basic German clause structure in (6.4) that weak pronouns show up in the "Wackernagel" position πP, Müller (1998) assumes a constraint that forces movement of weak pronouns to the domain of the functional head π at S-structure: the Pronoun Criterion, or PRON-CRIT (Müller 1998, 14):

(6.5) Pronoun Criterion (PRON-CRIT):
Weak pronouns must be in the domain of π at S-structure.

The domain of a head X comprises SpecX and XP adjuncts, perhaps also X adjuncts. Müller (1998) assumes that only one pronoun can be substituted in Specπ to fulfill PRON-CRIT; further pronouns adjoin to πP (p. 14).

Next, the Extended Projection Principle, or EPP, (cf. Chomsky 1995) requires NPs with nominative case to be in SpecT (the subject position) at S-structure (Müller 1998, 14):

(6.6) Extended Projection Principle (EPP):
NP\text{nom} must be in SpecT at S-structure.

These two constraints are the markedness constraints triggering movement of pronouns and subjects. They are counteracted by the faithfulness constraint STAY:

(6.7) STAY:
S-structure movement is not allowed.
Müller presupposes another faithfulness constraint minimizing the effect of syntactic movement, demanding parallel movement of NPs (Müller 1998, 15):

(6.8) Parallel Movement (PAR-MOVE):
If \( \alpha \) c-commands \( \beta \) at level \( L_n \), then \( \alpha \) c-commands \( \beta \) also at level \( L_{n+1} \) (where \( \alpha, \beta \) are arguments). ²

The ranking of these four constraints for German is as following: PRON-CRIT necessarily dominates STAY (since weak pronouns move to \( \pi P \)) and PAR-MOVE, and EPP and STAY are tied:

(6.9) Constraint ranking for German:
PRON-CRIT ≫ STAY ≪ EPP ≫ PAR-MOVE.

A constraint tie means that a candidate may be optimal under any possible ranking of the tied constraints, i.e. either STAY or EPP may dominate the other in a competition while resulting in the same optimal output.

With these constraints we are able to explain weak pronoun movement in German, allowing us to verify the ranking in (6.9). In (6.10a), the two pronouns (the DO and IO) precede the subject NP as expected, whereas the sentence in b) is impossible:

(6.10) a) dass es₁ ihm₂ der Fritz t₁ t₂ gegeben hat.
that itₜₜ dat him₂ DET Fritz nom given has.
"that Fritz gave it to him."

b) *dass ihm₂ es₁ der Fritz t₁ t₂ geben hat.

² c-command: A node \( A \) c-commands a node \( B \) iff (i) \( A \) does not dominate \( B \) and \( B \) does not dominate \( A \); and (ii) the first branching node dominating \( B \) also dominates \( B \). (Haegeman 1994, 134).
In OT we can account for this in the following way: the partial ranking $\text{PRON-CRIT} \gg \text{STAY}$ implies that weak pronouns must undergo movement to the domain of $\pi$ at S-structure, either to Spec $\pi$ or to a $\pi P$-(left-)adjoined position.

The fixed order of the pronouns is ensured by $\text{PAR-MOVE}$. Under the partial ranking $\text{PRON-CRIT} \gg \text{PAR-MOVE}$ weak pronouns can cross subject NPs, thereby violating $\text{PAR-MOVE}$ in order to satisfy the higher-ranked $\text{PRON-CRIT}$. Judging only by $\text{PRON-CRIT}$, all orders of pronouns were equally well, so this is where $\text{PAR-MOVE}$ comes into action. A low-ranked $\text{PAR-MOVE}$ predicts that the D-structure order of arguments is preserved at S-structure (if possible), exactly what we need for the analysis. Furthermore, this behavior strengthens the assumption of the underlying clause structure in (6.4), with DO uniformly preceding IO.

With the constraints in (6.5) through (6.8) it is possible to analyze weak pronoun fronting in German, providing evidence for the underlying clause structure in (6.4). But so far, we are not able to account for scrambling, as no motivation for movement of real nouns (i.e., not just pronouns) has yet been presented. Because of $\text{STAY}$ and $\text{PAR-MOVE}$, another trigger is needed: Müller (1998) proposes a Scrambling Criterion $\text{SCR-CRIT}$, similar in nature to $\text{PRON-CRIT}$.

In contrast to pronoun movement, reordering of Mittelfeld-internal arguments does not necessarily invoke ungrammaticality. Thus, the treatment of such reordering must differ somehow from that of pronoun fronting, and $\text{SCR-CRIT}$ therefore needs to differ from $\text{PRON-CRIT}$. For this reason Müller (1998) assumes the Scrambling Criterion to consist of several “sub-constraints” which build up $\text{SCR-CRIT}$. This results in two constraint levels, a matrix hierarchy and a
subhierarchy. Fatal violations on the matrix hierarchy necessarily induce suboptimality in the sense of pure OT, i.e. strict ungrammaticality. All constraints discussed until now (besides SCR-CRIT) belong to that matrix hierarchy. In contrast, fatal violation on a subhierarchy only leads to markedness. The constraints that trigger scrambling belong to this latter hierarchy, and accordingly we find degrees of markedness with candidates in this domain.

SCR-CRIT consists of several conflicting linearization constraints; to distinguish the subhierarchy from the matrix hierarchy, Müller uses $>$ to indicate ranking on the subhierarchy (in contrast to $\gg$ on the matrix hierarchy). Thus, SCR-CRIT looks as in (6.11) (Müller 1998, 22):

(6.11) Scrambling Criterion (SCR-CRIT)\(^3\)

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 [-dat] >
f) ADV (‘Adverb constraint’): NP [+def] precedes a VP adverb >
g) PER (‘Permutation constraint’): If $\alpha$ c-commands $\beta$ at level L\(_n\), then $\alpha$ does not c-command $\beta$ at level L\(_{n+1}\).

As for the ranking on the matrix hierarchy, clearly SCR-CRIT must dominate STAY and PAR-MOVE, since scrambling exists in German and may change the word order of NPs in the VP. Furthermore, Pron-Crit must dominate SCR-CRIT because weak pronouns must move to the p domain (where SCR-CRIT does not apply), rather than showing up inside VP in order to satisfy any of the SCR-CRIT linearization sub-constraints. Hence, the ranking will be as in (6.12):

\(^3\) See Müller (1998, 22).
After having introduced subhierarchies to OT, we need to redefine the notion of optimality, since subhierarchies are not accounted for in traditional OT. As markedness also needs to be defined, a distinction between optimality as unmarkedness and optimality as grammaticality is required.

Basically, the definition of optimality as grammaticality should remain the same: we only need to incorporate subhierarchies. In the case of a constraint that is a subhierarchy, the winning candidate is optimal if the subhierarchy is replaced by a constraint that belongs to the subhierarchy (Müller 1998, 23). In logical terms: the subhierarchy is interpreted via disjunction of its constraints – the relative ranking of these internal constraints is irrelevant.

(6.13) **Grammaticality** (Müller 1998, 23)

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 $<B_1, B_2, \ldots, B_n>$ on which $K_i$ and $K_j$ conflict better than $K_j$, where $B_l$ is replaced by some $C_k$ in $<C_1, C_2, \ldots, C_n>$ if $B_l$ is a subhierarchy $<C_1, C_2, \ldots, C_n>$.

In the present case of SCR-CRIT, supposing that SCR-CRIT is the highest-ranking constraint, this means that a candidate will be optimal when SCR-CRIT is replaced by any of the linearization constraints in (6.11). Concretely, a candidate would be grammatical if it is optimal under a ranking where, e.g., SCR-CRIT is replaced by NOM (or DEF, and so on).

Turning now to markedness, we can record that the notion of unmarkedness can be defined similarly:
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 \( <B_1, B_2, \ldots, B_n> \) on which \( K_i \) and \( K_j \) conflict better than \( K_j \), where \( B_l \) is replaced by \( <C_1, C_2, \ldots, C_n> \) if \( B_l \) is a subhierarchy \( <C_1, C_2, \ldots, C_n> \).

The crucial difference is that the unmarked candidate is determined not by substituting any of the constraints of the subhierarchy for \( \text{SCR-CRIT} \), but by substituting \textit{all} of the constraints for it. Thus, for the determination of unmarkedness, the distinction of matrix hierarchy and subhierarchy is overridden. Clearly, it follows that an unmarked candidate is grammatical, but not the inversion. What is not yet incorporated is how to account for different degrees of markedness. Here, Müller adopts Keller’s (1996) concept of suboptimality in (6.15): “Among the grammatical candidates of a candidate set (determined according to the definition in (6.13)), 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 unmarkedness in (6.14)] – i.e., the worse the constraint profile of a grammatical candidate is, the more marked it is.” (Müller 1998, 24).

\[ \text{(6.15) Suboptimality (Keller 1998; cit. Müller 1998, 20)} \]

A structure \( S_i \) is suboptimal with respect to a structure \( S_j \) if there are subsets \( R_i \) and \( R_j \) of the reference 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 \( S_j \).

With these theoretical assumptions as background we will now return to empirical evidence to prove the rankings in (6.11) and (6.12), respectively, as well as the constraints involved. I will restrict analyses to only some representative examples,
just to be able to principally verify the subhierarchy. For a full proof see Müller (1998, 24–36). First, consider the pair of sentences in (6.16) (Müller 1998, 24):

\[(6.16) \quad \begin{align*}
    a) & \quad \text{dass eine Frau} \quad \text{den Fritz} \quad \text{geküsst hat.} \\
        & \quad \text{that a woman} \text{nom DET Fritz} \text{akk kissed has.} \\
    b) & \quad ?\text{dass den Fritz} \quad \text{eine Frau} \quad \text{geküsst hat.} \\
        & \quad \text{that DET Fritz} \text{akk a woman} \text{nom kissed has.}
\end{align*}\]

There is a conflict here between NOM (which requires the order in a) and DEF (which in turn demands the order in b). Both orders are grammatical, but generally it is assumed that b. is more marked than a. Thus we can conclude that in the ranking of the subhierarchy, the “Nominative Constraint“ NOM dominates the “Definiteness Constraint“ DEF:

\[
\begin{array}{|c|c|c|c|c|c|c|}
\hline
\text{Candidates} & \text{SCR-CRIT} & \text{STAY/*t} & \text{PAR-MOVE} \\
\hline
& \text{NOM} & \text{DEF} & \text{AN} & \text{FOC} & \text{DAT} & \text{ADV} & \text{PER} \\
\hline
a) & * & * & & & & & \\
b) & * & & & & & * & * \\
\hline
\end{array}
\]

Candidate a) is grammatical under substitution of SCR-CRIT by NOM (or any of the other linearization constraints but DEF and PER\textsuperscript{4}) but also unmarked (if the subhierarchy replaces SCR-CRIT as a whole, only this candidate emerges as optimal due to the violation of the highest-ranked NOM by candidate b). Candidate b) in turn is grammatical when SCR-CRIT is replaced by either DEF or PER, which can for this reason be identified as trigger for the scrambling operation, but is therefore necessarily marked.

\textsuperscript{4} Substitution of any of the constraints in the subhierarchy on which the two candidates do not differ will pass on the decision two the lower-ranked constraints STAY and PAR-MOVE. Candidate a) does not involve scrambling (whereas candidate b) does) and thus does not violate these constraints. Therefore it has an „inherent advantage“ (Müller 1998, 25) over candidate b) and will be optimal.
As a second example, let me illustrate the competition between animacy and focus. This will also serve as the basis of the implementation in chapter 7. Evidence for the ranking AN > FOC can be gained from the data in (6.17) (Müller 1998, 33):

\[
(6.17) \begin{align*}
\text{a) dass man die Kinder} & \text{ diesem EINFLUSS entzogen/} \\
& \text{ that one the children\textunderscore acc this influence\textunderscore dat deprived/} \\
& \text{ ausgesetzt/ausgeliefert hat.} \\
& \text{ exposed/ extradited has.}
\end{align*}
\]

b) ?dass man die KINDER diesem Einfluss entzogen/ausgesetzt/ ausgeliefert hat.

c) ??dass man diesem Einfluss die KINDER entzogen/ausgesetzt/ ausgeliefert hat.

d) ??*dass man diesem EINFLUSS die Kinder entzogen/ausgesetzt/ ausgeliefert hat.

<table>
<thead>
<tr>
<th>Candidates</th>
<th>SCR-CRIT</th>
<th>STAY/*</th>
<th>PAR-MOVE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NOM</td>
<td>DEF</td>
<td>AN</td>
</tr>
<tr>
<td>a)</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c)</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>d)</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Candidates a) and b) vary only with respect to the position of the focused element (both violate DAT and DEF): In candidate b), there is a violation of FOC since not the VP-final element is focused. This renders candidate b) more marked then candidate a). Candidates c) and d) both involve scrambling (of the IO) and thus both violate STAY and PAR-MOVE. Additionally, both violate AN (avoiding violations of DAT and PER), hence they lose to candidates a) and b). However, the animacy constraint acts as trigger for scrambling here (hence the violations of STAY and PAR-MOVE), so the violations only lead to (further and stronger) markedness –
all four candidates are (correctly) predicted to be grammatical, even if candidate
d) already appears close to the border of ungrammaticality.

I will confine myself to these examples as by now it should be clear how the
ranking of the subhierarchy can be deduced.

We have found that it is quite easy for candidates with different VP-internal word
order to be grammatical, how markedness arises and how this can be analyzed.
What remains to be done is to rule out instances of iterated scrambling: If
scrambling does not lead to any improved behaviour regarding one of the
linearization constraints of SCR-CRIT, then we want that case to be ungrammatical.

Compare the following sentences (Müller 1998, 35):

(6.21) a) dass [\text{VP die Maria} \ [\text{V' den Fritz geküsst hat}]].
    that DET Maria nom DET Fritz acc kissed has.

b) *dass [\text{VP die Maria} [\text{VP den Fritz} [\text{VP t1 [V' t2 geküsst hat]]}}].

So-called \text{string-vacuous scrambling} is straightforwardly blocked under the
discussed assumptions. Candidate b) accumulates a second violation of \text{STAY},
which necessarily becomes decisive in this context and rules out candidate b). in
favor of the more ecomomic candidate a). Put simply, there is just no motivation
for string-vacuous scrambling, as there is always a more ecomomic candidate.

Finally, I want to address the topic of cross-linguistic variation. Scrambling not
only exists in German, there is also substantial evidence for a scrambling
operation, e.g., in Dutch and Icelandic (see Müller 1998, 36 for references). In these
languages however, it must be order-preserving: Dutch and Icelandic allow DO
and IO to precede or follow adjunct, as long as the D-structure order of the
arguments is obeyed. We can account for this if, in the matrix hierarchy, SCR-CRIT dominates STAY but is in turn outranked by PAR-MOVE. All candidates that satisfy SCR-CRIT by changing the order of arguments are filtered out by PAR-MOVE and all that maintain argument order are not. Thus languages of this type (i.e. with the Dutch/Icelandic constraint ranking) show considerably less freedom of word order than German, which, in turn shows less than a language like Russian\(^5\). In turn, English prohibits all kinds of scrambling and thus exhibits considerably less freedom of word ordering than languages of the Dutch/Icelandic type. The ranking for English differs from that of Dutch/Icelandic by ranking STAY too over SCR-CRIT, thus blurring its effects.

The advantage of the approach of Müller (1998) is that scrambling is not tied to other, independently motivated properties of a language (e.g., morphological case) but solely depends on the relative ranking of the three constraints SCR-CRIT, STAY, and PAR-MOVE.

### 6.4 Summary

The goal of this chapter was to present the Optimality-theoretical approach of Müller (1998) to scrambling. We have seen that a Scrambling criterion SCR-CRIT is proposed that consists of multiple features hierarchically ranked themselves to form a constraint subhierarchy embedded in the overall (matrix) hierarchy. Among the features that make up SCR-CRIT are syntactic attributes like

---

\(^5\) Russian exhibits more landing sites for scrambling than German. According to (Müller 1998, 5, footnote 3), Russian permits scrambling to TP, CP, and NP.
definiteness, animacy, focus, and others. Formulated as linearization (precedence) constraints, these features can be seen as triggers for scrambling.

The distinction of a matrix hierarchy and a subhierarchy furthermore allows Müller to introduce a notion of markedness into Optimality Theory that is not accounted for in traditional OT. By incorporating subhierarchies into the definition of optimality, the grammaticality of all candidates involving scrambling is assured by replacing the subhierarchy constraint SCR-CRIT with any one of its internal linearization constraints (i.e. a linearization constraint does not play a role for grammaticality). Markedness on the other side is analyzed by substituting the entire subhierarchy for SCR-CRIT, and there the different violations in the subhierarchy become decisive: The worse the violations are, the more marked the candidate is.

The advantage of Müller’s (1998) account of scrambling is that it is not dependent of any non-syntactic properties. Also, it manages without any other, independently motivated constraint. Cross-linguistic variation can simply be derived from re-ranking of the three constraints SCR-CRIT, STAY, and PAR-MOVE. The model of Müller (1998) will serve as basis of the computational approach to scrambling in the following chapter.