

# Optionality in Optimality-Theoretic Syntax

Gereon Müller

## 1. Introduction

Pre-theoretically, we can conceive of *syntactic optionality* as a name for a situation in which different ways of saying what seems to be the same thing show a clear correspondence in form. Such a situation may or may not be problematic for a given syntactic theory. Classic transformational grammar of the sixties acknowledges syntactic optionality by introducing a distinction between obligatory and optional transformations. Instances of syntactic optionality can be traced back to transformations that apply optionally. Some typical instances of syntactic optionality are given in (1)–(6), together with an account in terms of classic transformational grammar. First, English complementizer drop is optional in declarative object clauses (embedded by bridge verbs):

- (1) a. I think that John is a fool  
b. I think – John is a fool

Second, English dative shift (that removes the preposition *to* and reorders direct and indirect object) is optional:

- (2) a. John gave [NP a book ] [PP to Mary ]  
b. John gave [NP Mary ] [NP a book ]

Third, German *wh*-scope marker insertion is optional; i.e., in long-distance *wh*-dependencies, successive-cyclic *wh*-movement co-exists with partial movement of the *wh*-phrase to an embedded SpecC position, accompanied by *wh*-scope marker insertion in the target SpecC position:

- (3) a. Wen<sub>1</sub> glaubst du [CP t'<sub>1</sub> dass man t<sub>1</sub> einladen sollte ] ?  
whom think you that one invite should  
b. Was<sub>1</sub> glaubst du [CP wen<sub>1</sub> (dass) man t<sub>1</sub> einladen sollte ] ?  
[+wh] think you whom that one invite should  
BOTH: 'Who do you think we should invite?'

Fourth, French *wh*-movement of argument XPs is optional in root clauses:

- (4) a. Qui<sub>1</sub> as-tu vu t<sub>1</sub> ?  
who have-you seen  
b. Tu as vu qui<sub>1</sub> ?  
you have seen who  
BOTH: 'Who did you see?'

Fifth, PP extraposition from NP is optional in English:

- (5) a. [NP A review [PP<sub>1</sub> of this article ]] came out yesterday  
b. [NP A review t<sub>1</sub> ] came out yesterday [PP<sub>1</sub> of this article ]

And finally, scrambling is optional in free word order languages like German:

- (6) a. dass [NP<sub>1</sub> keiner ] [NP<sub>2</sub> den Fritz ] gesehen hat  
that no-one<sub>nom</sub> ART Fritz<sub>acc</sub> seen has  
b. dass [NP<sub>2</sub> den Fritz ] [NP<sub>1</sub> keiner ] t<sub>2</sub> gesehen hat  
that ART Fritz<sub>acc</sub> no-one<sub>nom</sub> seen has  
BOTH: ‘that no-one saw Fritz’

In standard GB theory (Chomsky 1981, Lasnik & Saito 1992), there is but one transformation left. This transformation, Affect  $\alpha$ , applies optionally throughout, subject to the requirement that the resulting syntactic object does not violate any constraints of grammar – of course, this requirement often creates the effect of obligatory rule application. Still, in this system, syntactic optionality is not problematic; it simply arises whenever the constraints of grammar can all be fulfilled both by applying Affect  $\alpha$  to a given item  $\alpha$ , and by leaving  $\alpha$  unaffected.

The Minimalist Program (Chomsky 1995) takes the opposite direction: Here, all syntactic operations are obligatory. The operation Move in particular applies if and only if it is triggered by features with certain properties. Hence, syntactic optionality at first sight poses a problem. However, there is an obvious and simple way out: Those cases in which classic transformational grammar posited an optional movement transformation can be reanalyzed as obligatory instances of Move triggered by optional features. Concerning, e.g., French root *wh*-movement in (4), we can say that the feature that triggers *wh*-movement is optional in root contexts. Thus, it looks as though syntactic optionality is unproblematic under minimalist assumptions. However, closer inspection reveals that this conclusion is in need of a qualification.

There are versions of the Minimalist Program that employ transderivational economy constraints like Chomsky’s (1991) Fewest Steps; cf., e.g., Epstein (1992), Kitahara (1993; 1997), Collins (1994), Chomsky (1995), Fox (1995), and Nakamura (1998). (Transderivational economy is a recurrent theme in all chapters of Chomsky (1995), even though it becomes clear towards the end of the book that Chomsky would ideally want to dispense with this notion in favour of local economy, on which see also Collins (1997). Still, in the guise of the Merge before Move constraint, a remnant of transderivational economy has survived even in Chomsky (1999).) In a minimalist syntax that incorporates Fewest Steps, a derivation is well formed only if it violates none of the standard (local) constraints of grammar, and, in addition, involves the fewest instances of syntactic operations in its

candidate (or reference) set. Depending on how candidate sets are defined, syntactic optionality may now become problematic again. For instance, (4-a) involves more applications of Move than (4-b), and the situation is similar in at least some of the other examples given above.

To maintain transderivational economy in view of this situation, it seems that we have to ensure that two derivations generating sentences that exhibit syntactic optionality are not in the same candidate set. This means that sufficiently many conditions must be imposed on the definition of candidate set, so that candidate sets get smaller, and syntactic optionality turns out to be only apparent – each “optional” derivation is the most economical candidate in its candidate set. Thus, instead of, e.g., defining candidate sets in terms of identity of meaning (or LF; recall the above notion of “saying the same thing”), we can define them in terms of identity of lexical material (or numeration, cf. Chomsky 1995). This would suffice to ensure that both sentences can survive in (1)–(3); and assuming that different feature specifications on lexical items must result in non-identity of lexical material, (4)–(6) could be accounted for in the same way: Derivations that differ only with respect to some optional feature triggering movement are not part of the same candidate set. However, as noted in Sternefeld (1997), such a step straightforwardly leads to a dilemma: A significant reduction of competition in candidate sets may indeed reconcile optionality with transderivational economy; but as an unwanted side effect, it also threatens to undermine the notion of transderivational economy itself, because many ill-formed derivations that could be shown to be blocked via, e.g., Fewest Steps, will now survive because the more economical derivation is not part of the same candidate set anymore. Finding a suitable definition of candidate set that is weak enough to permit optionality and strong enough to actually do some work may not be an impossible task, but the attempts that I am aware of suggest that it is a fairly difficult one (cf. Müller & Sternefeld 1996).

Optimality theory recognizes the notion of competition in candidate sets as the fundamental concept of grammar. It is therefore not surprising that optionality is inherently difficult to account for in optimality-theoretic syntax.

## 2. Optimality-theoretic syntax

The structure of an optimality-theoretic syntax looks as follows (cf. Prince & Smolensky 1993). Based on a given input, a first part of the grammar that contains only inviolable and unranked constraints (GEN) generates candidates, which are then subjected to a competition in terms of harmony evaluation in a second part of the grammar that employs violable and ranked constraints (H-EVAL). Harmony evaluation selects the optimal (= grammatical) candidate(s) in a candidate set. Optimality can be defined as follows. ((7)

is a modification of the definition in Grimshaw (1997), which closer scrutiny reveals to be incompatible with the existence of more than one optimal candidate in a candidate set; see Heck 1998 and Müller 2000.)

(7) *Optimality:*

A candidate  $C_i$  is optimal with respect to a constraint ranking  $\langle \text{CON}_1 \gg \dots \text{CON}_i \dots \gg \text{CON}_n \rangle$  iff there is no candidate  $C_j$  in the same candidate set such that:

- a. There is a constraint  $\text{CON}_k$  that  $C_j$  satisfies better than  $C_i$ ; and
- b. There is no constraint  $\text{CON}_l$  outranking  $\text{CON}_k$  on which  $C_i$  and  $C_j$  differ.

The notions of input, candidates, and candidate set remain to be clarified. For the time being, it will suffice to adopt some simplified definitions (based on Grimshaw 1997). Suppose that inputs are predicate/argument structures with tense and aspect specification; that candidates are S-structure representations (although assuming candidates to be full derivations, as suggested in Müller (1997) and envisaged in Legendre, Smolensky & Wilson (1998), would not make a difference in what follows); and that candidate sets are defined as in (8):

(8) *Candidate Set:*

Two candidates  $C_i, C_j$  belong to the same candidate set iff:

- a.  $C_i, C_j$  are realizations of identical predicate/argument structures.
- b.  $C_i, C_j$  have identical LFs.

Now we are in a position to formulate the problem that optionality as in (1)–(6) raises for optimality-theoretic syntax: Two (or more) candidates are optimal even though they seem to belong to the same candidate set. In view of this state of affairs, several proposals have been made, all of which belong to one of the four following types:

- I. Pseudo-Optionality:  $C_i, C_j$  belong to different candidate sets and do not interact.
- II. True Optionality:  $C_i, C_j$  have an identical constraint profile.
- III. Ties:  $C_i, C_j$  differ only on two (or more) constraints that are tied. Ties can be interpreted in various ways: globally or locally; ordered, conjunctive, or disjunctive.
- IV. Neutralization:  $C_i, C_j$  belong to different candidate sets, but interact nevertheless.

In what follows, I will address these proposals in turn.

### 3. Pseudo-optionality

Pseudo-optionality has been pursued by Grimshaw & Samek-Lodovici (1995; 1998), Choi (1996; 1999), Samek-Lodovici (1996), Costa (1998; 2001), Legendre (1998), Buring

(2001), and Heck (2001), among others. The basic strategy is similar to the one that is standardly pursued in minimalist systems with transderivational economy constraints: The observed optionality in sentences pairs like those in (1)–(6) is only apparent, the sentences belong to different candidate sets after all, and they are the sole optimal candidates in their respective candidate sets. Of course, a precondition for such an analysis is that candidate sets are sufficiently small. Under the notion of candidate set in (8), e.g., we would expect the candidates in each of the above examples to belong to the same candidate set and, hence, to compete for optimality. How, then, can candidate sets be reduced appropriately to avoid competition in optionality contexts?

As a first step, we might replace the notion of “identical predicate/argument structures” in (8-a) with the stricter notion of “identical numeration,” just as shown above for transderivational minimalism. This would create two candidate sets in the example pairs (1) (presence vs. absence of the complementizer *that*), (2) (presence vs. absence of the Case-marking preposition *to*), and (3) (presence vs. absence of the *wh*-scope marker *was*). Thus, in each case, both sentences can emerge as optional. Still, this does not yet account for (4)–(6), where the lexical items are identical. In principle, one could resort to optional features here, again in analogy to what has been said above about transderivational minimalism. In practice, proponents of pseudo-optionality approaches have often tried to argue for subtle differences in meaning in cases like these, which create two candidate sets for each example pair via (8-b).

These subtle differences in meaning are often not relevant for truth conditions; they typically relate to pragmatic or functional notions like topic, focus, and the like. Something along these lines has independently been proposed for NP scrambling vs. NP in situ in German and other free word order languages (cf. (6)). In addition, it has been suggested that extraposition from NP in English (cf. (5)) has minor semantic effects. Similarly (and, given the difference in numerations, redundantly), it has often been claimed that dative shift can somehow create “affectedness” of the indirect object (see, e.g., Speas 1990 and literature cited there). In the case of optional root *wh*-movement in French (cf. (4)), the identification of a difference related to (8-b) is much harder. Any attempt at attributing a standard question interpretation to (4-a), and an echo question interpretation to (4-b), is disconfirmed by the empirical facts, according to which (4-b) can receive a genuine question interpretation; but it might not be entirely futile to look for functional differences between the two candidates; see Chang (1997) for discussion. Also see Cheng & Rooryck (2000), who argue that French *wh*-in situ shows a different intonation pattern.

However, pseudo-optionality raises an important problem: In all the cases discussed so far, optionality of two constructions may break down in certain contexts; here and henceforth, I will refer to this phenomenon of restricted optionality (i.e., co-occurrence of

optionality and its break-down) as *syntactic alternation*. Thus, for most speakers, complementizer drop (cf. (1)) becomes obligatory with embedded topicalization (cf. Rochemont 1989, Grimshaw 1997), and in subject clauses (cf. Stowell 1981):

- (9) a. I think that [PP to John ]<sub>1</sub> she gave a book t<sub>1</sub>  
 b. \*I think – [PP to John ]<sub>1</sub> she gave a book t<sub>1</sub>  
 c. It surprised me [CP that the earth is round ]  
 d. \*It surprised me [CP – the earth is round ]

Similarly, dative shift (cf. (2)) can be obligatory in certain contexts, and blocked in others (cf. Baker 1988, Larson 1988, and references cited there):

- (10) a. \*The orange socks cost [NP two dollars ] [PP to/for Linda ]  
 b. The orange socks cost [NP Linda ] [NP two dollars ]  
 c. I donated [NP money ] [PP to charity ]  
 d. \*I donated [NP charity ] [NP money ]

Moreover, partial *wh*-movement in German (cf. (3)) is blocked in certain contexts, like negative islands (cf. Rizzi 1992, Höhle 2000) and subject islands (cf. McDaniel 1989):

- (11) a. ?Wen<sub>1</sub> glaubst du nicht [CP t'<sub>1</sub> dass man t<sub>1</sub> einladen sollte ] ?  
           whom think you not that one invite should  
 b. \*Was<sub>1</sub> glaubst du nicht [CP wen<sub>1</sub> (dass) man t<sub>1</sub> einladen sollte ] ?  
           [+wh] think you not whom that one invite should  
           BOTH: 'Who don't you think one should invite?'  
 c. ?[PP Mit wem ]<sub>1</sub> ist es schade [CP t'<sub>1</sub> dass sie t<sub>1</sub> gesprochen hat ] ?  
           with whom is it a pity that she talked has  
 d. \*Was<sub>1</sub> ist es schade [CP [PP mit wem ]<sub>1</sub> sie t<sub>1</sub> gesprochen hat ] ?  
           [+wh] is it a pity with whom she talked has  
           BOTH: 'Who is it a pity that she talked to?'

Optionality of *wh*-movement in French (cf. (4)) holds only in root contexts, and only for *wh*-arguments. As soon as a *wh*-adjunct or an embedded context is involved, *wh*-movement becomes obligatory, as in English:

- (12) a. Pourquoi<sub>1</sub> es-tu venu t<sub>1</sub> ?  
           why have-you come  
 b. \*Tu es venu pourquoi<sub>1</sub> ?  
           you have come why  
           BOTH: 'Why did you come?'  
 c. Je me demande [CP qui<sub>1</sub> C tu as vu t<sub>1</sub> ]  
           I ask myself who you have seen

- d. \*Je me demande [<sub>CP</sub> – (que) tu as vu qui<sub>1</sub> ]  
 I ask myself that you have seen who  
 BOTH: ‘I wonder who you saw.’

Analogous instances of alternations can be observed with extraposition in English (cf. (5)) and scrambling in German (cf. (6)).

Syntactic alternation can be captured in optimality-theoretic syntax in a straightforward way, by assuming a competition from which only one of the candidates emerges as optimal in certain contexts. But this, of course, presupposes that the two candidates in the example pairs in (9)–(12) do in fact compete, i.e., belong to one and the same candidate set – an assumption that must be abandoned in a pseudo-optimality approach. Thus, we end up with the dilemma that it is not at all clear what candidate an ungrammatical sentence like, e.g., (9-d) can be blocked by if (9-c) is not part of the same candidate set. In a nutshell, if syntactic optionality is reanalyzed as pseudo-optimality, with different winners of different competitions, the question arises of why optionality may turn into obligatoriness in certain contexts.

For the time being, I will leave it at that (I will return to a more promising version of the pseudo-optimality approach at the end). Let us continue to assume that candidate sets are defined as proposed in (8) – i.e., without recourse to a notion like “identical numeration,” and without paying attention to subtle semantic (pragmatic, functional) factors that do not affect truth conditions –, and turn to a second approach to optionality in optimality-theoretic syntax.

#### 4. True optionality

Given the definition of optimality in (7), a candidate can be optimal without having a better constraint profile than all competitors; it suffices if there is no competitor that has a better constraint profile. Hence, if two candidates have an identical constraint profile, they can both be optimal; true optionality can arise within a single candidate set. Furthermore, in contrast to what we have seen with pseudo-optimality, alternation effects can be accounted for: In certain contexts, two alternating constructions may cease to have an identical constraint profile, and one construction will block the other one. Such an approach has been pursued by Grimshaw (1997) and Vikner (2001).

Consider first Grimshaw’s (1997) analysis of English complementizer drop. The violable constraints that are relevant are given in (13), together with their ranking. (TOP-SCOPE is not actually proposed by Grimshaw, but tacitly assumed; this constraint could be formulated in a simpler, but arguably less transparent way.)

- (13) a. PURE-EP (‘Purity of Extended Projection’):

There is no adjunction to the highest XP of an extended projection or its head.

b. »TOP-SCOPE (‘Topic-Scope’):

Topics are adjoined to IP.

c. »OB-HD (‘Obligatory Heads’):

A projection has a head.

d. »STAY (‘Derivational Economy’):

Trace is not allowed.

The optionality of complementizer drop in (1) now simply follows from the fact that both candidates have an identical constraint profile; as a matter of fact, neither of the candidates violates any of the constraints in (13). This is shown in tableau  $T_1$ . (Note that the analysis here forces us to assume that complementizer drop constructions can be instances of IP embedding; a candidate with an embedded CP and an empty C head fatally violates OB-HD, and a candidate that tries to fill this C head by Subject/Aux inversion violates the highest-ranked PURE-EP (in addition to the low-ranked STAY).)

$T_1$ : *Optionality of complementizer drop in English* (Grimshaw)

Candidates	PURE-EP	TOP-SCOPE	OB-HD	STAY
☞C <sub>1</sub> : think [CP that [IP John is a fool ]]				
☞C <sub>2</sub> : think [IP John is a fool ]]				
C <sub>3</sub> : think [CP – [IP John is a fool ]]			*!	
C <sub>4</sub> : think [CP is <sub>2</sub> [IP John t <sub>2</sub> a fool ]]	*!			*

However, complementizer drop is an alternation, in the technical sense of optionality that breaks down in certain contexts; cf. (9). And indeed, Grimshaw (1997) notes that the two constructions cease to exhibit an identical constraint profile in embedded topicalization contexts. Here *that* becomes obligatory because the complementizer drop construction will have to fatally violate either OB-HD (with CP embedding), TOP-SCOPE (with the topic in situ – of course, the string is optimal in another candidate set in which *to John* is not a topic), or PURE-EP (with topic adjunction to an IP that is the sister of the matrix V); cf. (9-a) vs. (9-b) and tableau  $T_2$ .

$T_2$ : *Complementizer drop and embedded topicalization in English*

Candidates	PURE-EP	TOP-SCOPE	OB-HD	STAY
☞C <sub>1</sub> : think [CP that [IP to John [IP she gave ... t ]]]]				*
C <sub>2</sub> : think [CP – [IP to John [IP she gave ... t ]]]]			*!	*
C <sub>3</sub> : think [IP she gave ... to John ]]		*!		
C <sub>4</sub> : think [IP to John [IP she gave ... t ]]	*!			*

As a second example, consider Vikner’s (2001) extension of this proposal that captures an interesting cross-linguistic correlation: It seems that only those languages can have complementizer drop that lack overt V-to-I raising (English, Danish), not those that have V-to-I raising (French, Icelandic; following the work of Emonds, Pollock, and others, the evidence for V-to-I raising at S-structure can be gained from the relative order of VP adverbials and finite V). The obligatoriness of complementizers in the latter class of languages is exemplified by the French data in (14):

- (14) a. Je crois que l’acteur voit vraiment le film  
           I think that the actor sees really the film  
       b. \*Je crois – l’acteur voit vraiment le film  
           I think the actor sees really the film  
           БОТН: ‘I think that the actor really sees the film.’

Two additional constraints become necessary: one that punishes movement of a lexical head (LX-MV, ‘No Lexical Movement’), and one that prohibits lowering of I features to V (PR-BD, ‘Proper Binding’); the ranking is LX-MV  $\gg$  PR-BD in English, and PR-BD  $\gg$  LX-MV in French. Consequently, the candidate with V-in situ will be optimal in English, and Grimshaw’s account of complementizer drop optionality can be maintained essentially unchanged. In contrast, the candidate with V-to-I raising will be optimal in French. But since V-to-I raising incurs a fatal PURE-EP violation if IP is the highest extended projection, and since bare IP embedding has turned out to be a precondition for complementizer drop optionality in this system, it follows that complementizers are obligatory in French, for basically the same reasons as in embedded topicalization structures in English.

Thus, the true optionality approach seems empirically superior to the pseudo-optionality approach because it can straightforwardly account for syntactic alternation, i.e., the breakdown of optionality in certain contexts. In addition, it is theoretically attractive since it does not require any specific additional assumptions, given a notion of optimality like (7) that permits more than one winner in a single competition.

Still, there is a severe problem that has been noted in Baković (1997). On closer inspection, it turns out to be extremely difficult to maintain an identical constraint profile of two candidates. For instance, there must not be a single constraint in grammar that may distinguish between the two optimal candidates in  $T_1$ ; in particular, the presence of a complementizer, and of the additional structure it requires, must be completely costless. This would be incompatible with the existence of a constraint like Grimshaw’s (1993) MIN-PROJ (‘Minimal Projection’) that demands economy of projection, which is however not adopted in Grimshaw (1997). More importantly, as Grimshaw herself notes, it is in fact incompatible with the existence of a low-ranked constraint HD-RT (‘Head-

Right’) that is adopted in Grimshaw (1997). HD-RT requires heads (including C) to be rightmost in their projections. In English, the effects of HD-RT are usually blurred by a higher-ranked HD-LFT (‘Head-Left’), but even an extremely low-ranked HD-RT can differentiate between the optional variants  $C_1$ ,  $C_2$  in tableau  $T_1$ , and thereby wrongly predict complementizer drop to be obligatory in this context. And indeed, for reasons like this one, Grimshaw (1999) abandons this approach to complementizer drop.

## 5. Ties

The basic idea behind ties is that two (or more) constraints are equally important, i.e., “tied.” If two candidates differ only with respect to a tie of constraints, they can both be optimal, even if their constraint profile is not completely identical. As with true optionality, alternation phenomena can be traced back to the impact of another constraint that may differentiate between candidates that otherwise differ only on ties. Various notations of ties of two constraints A, B can be found in the literature, e.g., “ $A \ll B$ ”, “ $A \approx B$ ”, and “A/B;” in what follows, I will note a tie of A and B as “ $A \circ B$ ” (in tableaux, ties will be represented by vertical dotted lines). More importantly, various concepts of tie have been proposed. I am aware of at least five distinct notions of tie that can be shown to be both conceptually different, and empirically incompatible. Still, in the abstract tableau  $T_3$ , with the ranking  $A \gg B \circ C \gg D$ , they turn out to make identical predictions:  $C_1$  and  $C_2$  are both optimal, whereas  $C_3$  and  $C_4$  are blocked as suboptimal.

$T_3$ : *Ties*

Candidates	A	B   C	D
$\Leftarrow C_1$		*	
$\Leftarrow C_2$		*	
$C_3$		*   *!	
$C_4$	*!		

A basic distinction can be made between what I will call global and local ties. Global ties can be viewed as abbreviations for multiple constraint rankings co-existing in a single language; local ties can be viewed as special constraint types. (Legendre 2001 reserves the term “tie” for what I call “local tie,” and refers to what I call “global tie” as “floating constraints” or “partial orderings,” for reasons that will become clear immediately.)

### 5.1. Ordered global ties

Ordered global ties have been adopted by Sells, Rickford & Wasow (1996), Ackema & Neeleman (1998), Schmid (2000), and others. They correspond to a notion of tie that

is suggested in Prince & Smolensky (1993). The basic idea is that a constraint ranking that exhibits a tie of two constraints  $C_1 \circ C_2$  is underspecified; it is an abbreviation that encodes the simultaneous presence of two hierarchies that exhibit the rankings  $C_1 \gg C_2$  and  $C_2 \gg C_1$ . On this view, a candidate is grammatical if it is optimal under one of the possible resolutions of a tie; a different behaviour on lower-ranked constraints is irrelevant.

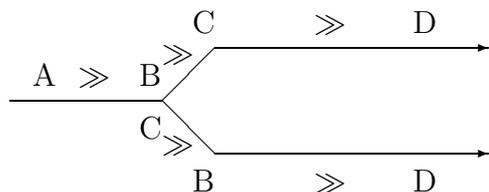
For concreteness, we can assume that the notion of ranking in the definition of optimality in (7) presupposes a *total* order of constraints. We can then understand ordered global ties as underspecified representations with *partial* orders:

(15) *Ordered Global Tie:*

Suppose that  $\Gamma = \langle \text{CON}_1 \gg \dots \text{CON}_i \dots \gg \text{CON}_n \rangle$  is a partial constraint order in language L, and  $\text{CON}_i$  ( $1 \leq i \leq n$ )  $\in \Gamma$  is a tie  $\text{CON}_{i_1} \circ \dots \circ \text{CON}_{i_n}$ . Then, for every suborder O of the constraints in  $\text{CON}_i$ ,  $\Gamma_O$  is a constraint order of language L, where  $\Gamma_O$  differs from  $\Gamma$  only in that  $\text{CON}_i$  is replaced by O.

Given that optimality is only defined with respect to total constraint orders, and not with respect to partial constraint orders, (15) implies that a candidate is grammatical in a language iff it is optimal under at least one total order that is generated by resolving all ties that exist in the language. The following diagram illustrates ordered global ties:

(16) *Diagram of an ordered global tie  $B \circ C$*



As an example, let us consider the analysis of optional *wh*-movement in French root clauses (cf. (4)) that is developed (but ultimately rejected) by Ackema & Neeleman (1998). The relevant constraints are given in (17). (SPC is called STAY by Ackema & Neeleman. Since this constraint is understood in a rather different way from the constraint STAY in Grimshaw (1997), I have chosen another name here.)

(17) a. SPC ('Shortest Paths Condition'):

Minimize movement paths: Every node crossed by movement creates a \*.

b. Q-MARKING:

In a question, assign a [+Q] feature to the constituent corresponding to the proposition.

c. Q-SCOPE:

[+Q] elements must c-command the constituent corresponding to the proposition.

Ackema & Neeleman then suggest the following constraint ranking for French: SPC $\circ$ Q-MARKING  $\gg$  Q-SCOPE. Since this ranking involves an ordered global tie, and hence, a partial order of constraints, optimality cannot yet be checked – first, total orders must be created. Then, if SPC outranks Q-MARKING, the candidate (4-b) that dispenses with *wh*-movement (and thereby minimizes SPC violations) is optimal; if, however, Q-MARKING outranks SPC, the candidate (4-a) that respects Q-MARKING by moving the *wh*-phrase emerges as the winner. This is shown in tableau T<sub>4</sub>, which should be viewed as an underspecified representation that simultaneously encodes two subtableaux – one in which C<sub>1</sub> is optimal, and one in which C<sub>2</sub> is optimal. (In line with this, (!) should be taken to indicate a violation that is fatal only under one resolution of the tie.)

T<sub>4</sub>: *Optional wh-movement in French root clauses*

Candidates	SPC	Q-MARKING	Q-SCOPE
C <sub>1</sub> : qui <sub>1</sub> Aux <sub>2</sub> NP t <sub>2</sub> V t <sub>1</sub>	*(!)*****		
C <sub>2</sub> : – – NP Aux <sub>2</sub> V qui <sub>1</sub>		*(!)	*
*C <sub>3</sub> : qui <sub>1</sub> – NP Aux <sub>2</sub> V t <sub>1</sub>	*(!)**	*(!)	
*C <sub>4</sub> : – Aux <sub>2</sub> NP t <sub>2</sub> V qui <sub>1</sub>	*(!)**	*(!)	*

Two remarks are in order concerning constraint satisfaction in this tableau. First, since the SPC does not count movement operations (like STAY), but nodes crossed in the course of movement, the exact number of SPC violations in tableau T<sub>4</sub> depends on specific assumptions about the number of functional projections between VP and CP, the inclusion/exclusion distinction, etc., which are not really relevant in the present context. Thus, the numbers of \*'s under SPC should be viewed as rough estimations. However, it seems uncontroversial that C<sub>1</sub>, which has both *wh*-movement and I-to-C movement, involves the largest number of nodes crossed; that C<sub>2</sub>, which leaves both items in situ, involves the smallest number of nodes crossed; and that C<sub>3</sub> and C<sub>4</sub>, which exhibit *wh*-movement and V-to-I raising, respectively, are somewhere in between C<sub>1</sub> and C<sub>2</sub> with respect to the SPC.

Second, whereas Q-SCOPE is satisfied by *wh*-movement and violated otherwise, matters are a bit more intricate with Q-MARKING. Ackema & Neeleman (1998) assume that Q-MARKING requires sisterhood of the *wh*-clause to a lexical X<sup>0</sup> head that bears a [+Q] feature. In root clauses, this constraint can only be fulfilled by a combination of *wh*-movement and I-to-C movement: The verb in C derives its [+Q] feature from a *wh*-phrase in SpecC and can then Q-mark its clausal sister. Neither *wh*-movement nor I-to-C movement can avoid a Q-MARKING violation alone.

An important difference to the true optionality approach is that the case of non-identical behaviour of two optimal candidates in one candidate set on lower-ranked con-

straints does not pose a problem. Thus, the *wh*-in situ candidate  $C_2$  incurs a violation of the lowest-ranked Q-SCOPE constraint that the *wh*-movement candidate  $C_1$  manages to avoid. Still,  $C_2$  is grammatical because it is optimal under the resolution of the tie  $\text{SPC} \circ \text{Q-MARKING}$  that results in a total order  $\text{SPC} \gg \text{Q-MARKING} \gg \text{Q-SCOPE}$ : The fatal violation of  $C_1$  on SPC under this ranking renders all differences on lower-ranked constraints (Q-MARKING and Q-SCOPE) irrelevant.

Despite this advantage, there is a well-known problem with ordered global ties that is related to complexity considerations. A tie of two constraints creates an underspecified representation, the resolution of which generates two constraint orders, and hence, two grammars that are simultaneously present in a single speaker. Two ordered global ties of two constraints therefore result in four grammars, three ties yield eight grammars, and so on. Furthermore, there is nothing that would preclude assuming an ordered global tie of three constraints; indeed, this is suggested by Sells, Rickford & Wasow (1996) to account for the existence of three types of negative inversion in African American Vernacular English. Such a three-way tie generates six grammars; and if there are two ties of this type, we end up with a simultaneous presence of thirty-six grammars, all within a single language. It is not quite clear how serious this consequence is from an empirical point of view; but it seems fair to conclude that ordered global ties are conceptually somewhat problematic, especially if more than two constraints participate in the tie, and if there are more than a few ties in a given grammar. Since it is likely that there are more than a few instances of syntactic optionality in every natural language, this result sheds doubt on the applicability of ordered global ties as a general mechanism to account for optionality, and makes it worthwhile to look for another concept of ties.

## 5.2. Ordered local ties

Pesetsky (1997; 1998) and Broihier (1995) propose a local concept of ordered ties that avoids the problem of grammar proliferation. As before, a tie  $A \circ B$  is resolved into its subrankings  $A \gg B$ ,  $B \gg A$ . However, this time the tie is not an abbreviation for the simultaneous presence of two (or more) constraint orders in a language; rather, the tie itself acts as a violable and ranked constraint (a “meta-constraint,” in the terminology of Broihier 1995). Pesetsky’s (1998) suggestion reads as follows: “The output of a set of tied constraints is the union of the outputs of every possible ranking of those constraints.” As a first approximation, we can postulate the following definition of ordered local ties, still presupposing the notion of optimality in (7):

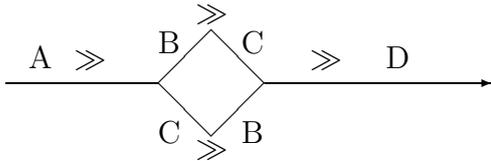
(18) *Ordered Local Tie:*

Suppose that  $\Gamma = \langle \text{CON}_1 \gg \dots \text{CON}_i \dots \gg \text{CON}_n \rangle$  is a total constraint order in language L,  $\text{CON}_i$  ( $1 \leq i \leq n$ )  $\in \Gamma$  is a tie  $\text{CON}_{i_1} \circ \dots \circ \text{CON}_{i_n}$ , and  $O_1, \dots, O_n$  are the

possible suborders of the constraints in  $\text{CON}_i$ . Then,  $\text{CON}_i$  is violated by a candidate  $C$  iff there is no suborder  $O_i$  such that  $C$  is optimal with respect to  $O_i$ .

Thus, a tie is treated as a constraint that can be violated or satisfied by a candidate (unlike “ordinary” constraints, it cannot be violated multiply, though). Optionality can arise because there is more than one possibility to satisfy the tie, depending on the ranking into which the tie is resolved. Schematically, ordered local ties look as in (19).

(19) *Diagram of an ordered local tie  $B \circ C$*



As an example, consider again English complementizer drop. Recall that complementizer drop is optional in declarative object clauses embedded by bridge verbs (cf. (1-a) and (1-b)), but impossible in subject clauses (cf. (9-c) vs. (9-d)). The following constraints are proposed in Pesetsky (1998).

- (20) a. REC (‘Recoverability’):  
A syntactic unit with semantic content must be pronounced unless it has a sufficiently local antecedent.
- b. LE(CP) (‘Left Edge(CP)’):  
The first pronounced word in CP is the complementizer that heads it.
- c. TEL (‘Telegraph’):  
Do not pronounce function words.
- d. DCP (‘Deletion in CP’):  
The head or specifier of a CP may be deleted only if that CP is a complement.
- e. SILENT-T (‘Silent Trace’):  
Minimize pronunciation of traces.

In English, the ranking of the first four constraints is held to be  $\text{REC} \gg \text{LE}(\text{CP}) \circ \text{TEL} \gg \text{DCP}$ . Since REC and DCP are always satisfied in declarative object clauses (*that* is semantically empty, and the CP is a complement), optionality in (1) follows from the tie of LE(CP) and TEL: (1-a) (=  $C_1$  in  $T_5$ ) satisfies LE(CP) and violates TEL; therefore, it satisfies the tie under the resolution  $\text{LE}(\text{CP}) \gg \text{TEL}$ . On the other hand, (1-b) (=  $C_2$  in  $T_5$ ) satisfies TEL and violates LE(CP); hence, it does not violate the tie under the resolution  $\text{TEL} \gg \text{LE}(\text{CP})$ . (Note that unlike Grimshaw (1997), Pesetsky (1998) excludes bare IP-embedding in GEN; in his system, such a candidate would always block the presence of a complementizer because, with no CP present, it would vacuously satisfy

both TEL and LE(CP).)

*T<sub>5</sub>: Optionality of complementizer drop in English (Pesetsky)*

Candidates	REC	LE(CP)   TEL	DCP
☞ C <sub>1</sub> : ... think [CP that John ...		*	
☞ C <sub>2</sub> : ... think [CP – John ...		*	

The breakdown of optionality in subject clauses follows from the constraint that discriminates between subject and object CPs, viz., DCP. Like LE(CP), DCP prohibits complementizer deletion, but it is more specific since it does so only in a certain context. That DCP is ranked low does not make it irrelevant here; since an ordered local tie (almost) functions just like any other constraint in the hierarchy (it can be satisfied or violated), the behaviour of two otherwise optimal candidates that both satisfy the tie on a lower-ranked constraint does make a difference; see T<sub>6</sub>.

*T<sub>6</sub>: Complementizer drop and subject clauses in English*

Candidates	REC	LE(CP)   TEL	DCP
☞ C <sub>1</sub> : ... surprised me [CP that the earth ...		*	
C <sub>2</sub> : ... surprised me [CP – the earth ...		*	*!

Closer scrutiny reveals that (18) is not yet fully adequate to capture Pesetsky's intentions. This can be shown by looking at relativization in Polish, as analysed in Broihier (1995) and Pesetsky (1998). If the relative pronoun is oblique, relativization can either proceed by pronouncing the relative pronoun and deleting both the complementizer and the relativization trace, or by deleting the relative pronoun and pronouncing both the complementizer and the trace (the latter as a resumptive pronoun, not as a full copy, so as to minimize SILENT-T violations.); cf. (21-ab). All other combinations, some of which are given in (21-c-f), are ungrammatical.

- (21) a. On spotkał studenta [CP któremu<sub>1</sub> – on dał t<sub>1</sub> piątkę ]  
 he met student whom<sub>dat</sub> he gave good mark
- b. On spotkał studenta [CP –<sub>1</sub> co mu<sub>1</sub> on dał piątkę ]  
 he met student that him<sub>dat</sub> he gave good mark
- c. \*On spotkał studenta [CP któremu<sub>1</sub> – on mu<sub>1</sub> dał piątkę ]  
 he met student whom<sub>dat</sub> he him<sub>dat</sub> gave good mark
- d. \*On spotkał studenta [CP –<sub>1</sub> co on dał t<sub>1</sub> piątkę ]  
 he met student that he gave good mark
- e. \*On spotkał studenta [CP któremu<sub>1</sub> co mu<sub>1</sub> on dał piątkę ]  
 he met student whom<sub>dat</sub> that him<sub>dat</sub> he gave good mark

- f. \*On spotkał studenta [<sub>CP</sub> któremu<sub>1</sub> co on dał t<sub>1</sub> piątke ]  
 he met student whom<sub>dat</sub> that he gave good mark  
 ALL: ‘He met the student who he gave a good mark.’

Broihier and Pesetsky propose a three-way tie SILENT-T ◊ LE(CP) ◊ TEL:

*T<sub>7</sub>: Optionality of relativization strategies in Polish*

Candidates	REC	SILENT-T   LE(CP)   TEL	DCP
☞ C <sub>1</sub> : wh <sub>1</sub> - ... t <sub>1</sub>		*	*
☞ C <sub>2</sub> : - <sub>1</sub> co ... mu <sub>1</sub>		*     *	*
C <sub>3</sub> : wh <sub>1</sub> - ... mu <sub>1</sub>		*   *	*
C <sub>4</sub> : - <sub>1</sub> co ... t <sub>1</sub>	*!	*	*
C <sub>5</sub> : wh <sub>1</sub> co ... mu <sub>1</sub>		*   *   *	
C <sub>6</sub> : wh <sub>1</sub> co ... t <sub>1</sub>		*   *	

A high-ranked REC strictly prohibits simultaneous deletion of the oblique relative pronoun and its trace (as in C<sub>4</sub>). A low-ranked DCP irrelevantly prohibits deletion in the domain of the non-complement CP because candidates that respect DCP will have to violate the higher-ranked constraints LE(CP) and TEL, and will incur a fatal violation of the tie (cf. C<sub>5</sub>, C<sub>6</sub>). It must be shown that C<sub>1</sub> and C<sub>2</sub> in T<sub>7</sub> do not violate the three-way tie, whereas C<sub>3</sub>, C<sub>5</sub>, and C<sub>6</sub> do. If the tie is resolved into a suborder in which TEL is highest, C<sub>1</sub> will be optimal within this tie, and will therefore not incur a violation here. If SILENT-T is highest, the ranking of the other two constraints becomes relevant: Under a resolution SILENT-T ≫ TEL ≫ LE(CP), C<sub>1</sub> satisfies the tie; and under a resolution SILENT-T ≫ LE(CP) ≫ TEL, C<sub>4</sub> satisfies the tie, but irrelevantly so because C<sub>4</sub> fatally violates REC. Finally, if LE(CP) is highest, we hope that C<sub>2</sub> optimally satisfies the tie. However, unfortunately, this is not the case: Under any ranking of SILENT-T and TEL, C<sub>4</sub> satisfies the tie better than C<sub>2</sub>. Hence, it is wrongly predicted that only C<sub>1</sub> is grammatical.

Intuitively, the problem is that C<sub>4</sub> can be optimal within the tie, thereby creating an unwanted violation of the tie by C<sub>2</sub>, even though C<sub>4</sub> violates a higher-ranked constraint. To remedy this situation, we have to ensure that a candidate like C<sub>4</sub> that violates a higher-ranked constraint cannot participate in a tie-internal optimization procedure. This can be done as in (22); but it seems that ordered local ties are then not quite as local anymore – to find out whether a candidate violates a tie, all the constraints that dominate the tie must also be considered. (Note incidentally that it is exactly in competitions like T<sub>7</sub> that optimality-theoretic learning strategies like Tesar & Smolensky’s (1993; 2000) constraint demotion algorithms encounter severe problems; cf. Broihier 1995. This may or may not be taken to suggest a deeper conceptual problem.)

(22) *Ordered Local Tie* (revised):

Suppose that  $\Gamma = \langle \text{CON}_1 \gg \dots \text{CON}_i \dots \gg \text{CON}_n \rangle$  is a total constraint order in language L,  $\text{CON}_i$  ( $1 \leq i \leq n$ )  $\in \Gamma$  is a tie  $\text{CON}_{i_1} \circ \dots \circ \text{CON}_{i_n}$ , and  $O_1, \dots, O_n$  are the possible suborders of the constraints in  $\text{CON}_i$ . Then,  $\text{CON}_i$  is violated by a candidate C iff there is no suborder  $O_i$  such that C is optimal according to  $\langle \text{CON}_1 \gg \dots \text{CON}_{i-1} \gg O_i \rangle$

This complication notwithstanding, ordered local ties evade the the complexity problem raised by ordered global ties. However, they reintroduce the problem that we have encountered with the true optionality approach: To derive optionality, it must be ensured that there are no lower-ranked constraints that may undermine the co-existence of two optimal candidates in one candidate set. Still, this problem is arguably less pressing than with true optionality. The reason is that the ordered local tie approach makes it possible to postulate constraints that refer to the pertinent items in alternating candidates, as long as these constraints are tied. For instance, one constraint may prohibit complementizer realization (TEL), another one may force it (LE(CP)); this is a situation that is not possible in the true optionality approach.

Finally, it is worth bearing in mind that ordered global ties and ordered local ties do not only differ conceptually, but also empirically. Thus, C<sub>2</sub> in T<sub>6</sub> (complementizer drop in subject clauses) is wrongly predicted to be optimal under the global definition of tie in (15); and C<sub>2</sub> in T<sub>4</sub> (*wh*-in situ in French) is wrongly predicted to be suboptimal under the local definition of tie in (18).

### 5.3. Conjunctive local ties

In addition to ordered global ties, Prince & Smolensky (1993) envisage a second concept of tie that relies on a “crucial nonranking” of constraints. This concept has been adopted by Legendre, Wilson, Smolensky, Homer & Raymond (1995), Legendre, Smolensky & Wilson (1998), and Müller (1997). In the first two studies, the optionality of resumptive pronouns and traces with argument topicalization in Chinese is derived from a tie of the constraints STAY, which blocks traces, and FILL, which prohibits the use of resumptive pronouns. I will not further discuss this analysis here because it is also compatible with approaches in terms of ordered (global or local) ties (in fact, only the reference to Tesar (1998) makes clear what concept of tie the authors have in mind). In contrast, we will see that conjunctive local ties are crucial for the analysis in Müller (1997).

The basic idea is that ties are treated as ordinary constraints, but there is no resolution of the tie into suborders. Rather, the two constraints are merged into a single constraint that is interpreted via logical conjunction: A candidate violates a tie if it violates a

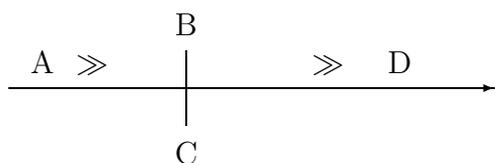
constraint that is part of this tie, and multiple violations add up.

(23) *Conjunctive Local Tie:*

Suppose that  $\Gamma = \langle \text{CON}_1 \gg \dots \text{CON}_i \dots \gg \text{CON}_n \rangle$  is a total constraint order in language L, and  $\text{CON}_i$  ( $1 \leq i \leq n$ )  $\in \Gamma$  is a tie  $\text{CON}_{i_1} \circ \dots \circ \text{CON}_{i_n}$ . Then,  $\text{CON}_i$  is violated by a candidate C iff there is a constraint  $\text{CON}_{i_j}$  that is violated by C.

A diagram illustrating conjunctive local ties is given in (24):

(24) *Diagram of a conjunctive local tie B $\circ$ C*



The relevant example in Müller (1997) is the treatment of the alternation of *wh*-scope marking and long-distance *wh*-movement in German (cf. (3)), and its breakdown in weak island contexts (cf. (11)). The constraints are given in (25).

- (25) a. WH-CRIT ('Wh-Criterion'; holds at S-structure):  
*Wh*-operators are in SpecC, and  $C_{[+wh]}$  requires filling of C or SpecC.
- b. BAR-CON ('Barriers Condition'):  
Chain formation must not cross a barrier.
- c. DER-ECON ('Derivational Economy,' a version of STAY):  
S-structure movement is not allowed.
- d. FULL-INT ('Full Interpretation'):  
A lexical item must have an interpretation.

The constraint ranking assumed for Standard German is WH-CRIT  $\gg$  BAR-CON  $\gg$  DER-ECON  $\circ$  FULL-INT. By assumption, the *wh*-scope marker *was* ('what') is an expletive; hence, each instance of *wh*-scope marker insertion incurs a violation of FULL-INT. On the other hand, each instance of overt *wh*-movement incurs a violation of DER-ECON. Given that these two constraints are tied, optionality arises: Successive-cyclic *wh*-movement as in (3-a) violates DER-ECON twice, partial *wh*-movement accompanied by scope marker insertion as in (3-b) violates both DER-ECON and FULL-INT once, and since this adds up to two violations of the tie in each case, the two candidates can both be optimal. They do indeed turn out to be optimal because all competing candidates that try to improve their behaviour on the tie violate the higher-ranked WH-CRIT fatally; cf. T<sub>8</sub>.

The optionality disappears in languages that exhibit a similar ranking but resolve the tie. Thus, *wh*-scope marking will always be blocked by long-distance *wh*-movement in English under a ranking FULL-INT  $\gg$  DER-ECON; and *wh*-scope marking is correctly

*T<sub>8</sub>: Optionality of wh-scope marking in German*

Candidates	WH-CRIT	BAR-CON	DER-ECON   FULL-INT
☞ C <sub>1</sub> : [CP wen <sub>1</sub> ... [CP t' <sub>1</sub> dass ... t <sub>1</sub> ]]			**
☞ C <sub>2</sub> : [CP was <sub>1</sub> ... [CP wen <sub>1</sub> (dass) ... t <sub>1</sub> ]]			*   *
C <sub>3</sub> : [CP was <sub>1</sub> ... [CP dass ... wen <sub>1</sub> ]]	*!		*
C <sub>4</sub> : [CP - ... [CP wen <sub>1</sub> ... t <sub>1</sub> ]]	*!		*
C <sub>5</sub> : [CP was <sub>1</sub> ... [CP was <sub>1</sub> ... wen <sub>1</sub> ]]	*!		**
C <sub>6</sub> : [CP - ... [CP dass ... wen <sub>1</sub> ]]	*!*		

predicted to block long-distance *wh*-movement in Northern German varieties under a ranking DER-ECON  $\gg$  FULL-INT. More importantly in the present context, optionality is also predicted to break down in Standard German as soon as another constraint becomes active that differentiates between the alternating candidates. As argued in Müller (1997), this is the case in weak island contexts. Given that subject clauses are barriers due to a lack of L-marking (cf. Chomsky (1986)), and given that a scope marker forms an S-structure chain with its associate, both long-distance *wh*-movement as in (11-c) and *wh*-scope marking as in (11-d) will invariably violate BAR-CON at S-structure. However, if *wh*-in situ must undergo LF raising to its scope position, (11-d) incurs an additional, fatal BAR-CON violation at LF that (11-c) does not incur. This is shown in tableau T<sub>9</sub>, where suboptimal candidates that fatally violate WH-CRIT are not represented again. (The argument is identical in the case of the negative island asymmetry in (11-a), (11-b).)

*T<sub>9</sub>: Wh-scope marking and subject islands in German*

Candidates	WH-CRIT	BAR-CON	DER-ECON   FULL-INT
☞ C <sub>1</sub> : [CP mit wem <sub>1</sub> ... [CP t' <sub>1</sub> dass ... t <sub>1</sub> ]]		*	**
C <sub>2</sub> : [CP was <sub>1</sub> ... [CP mit wem <sub>1</sub> (dass) ... t <sub>1</sub> ]]		**!	*   *

Thus far, the *wh*-scope marking alternation in German could equally well be captured with ordered (global or local) ties. Things begin to differ in the case of longer *wh*-dependencies that cross two CPs; here, only conjunctive local ties make correct predictions:

- (26) a. Wann<sub>1</sub> meinst du [CP t'<sub>1</sub> dass sie gesagt hat [CP t'<sub>1</sub> dass sie t<sub>1</sub> kommen  
when think you that she said has that she come  
würde ]] ?  
would
- b. Was<sub>1</sub> meinst du [CP wann<sub>1</sub> (dass) sie gesagt hat [CP t'<sub>1</sub> dass sie t<sub>1</sub> kommen  
[+wh] think you when that she said has that she come  
würde ]] ?  
would

- c. Was<sub>1</sub> meinst du [CP was<sub>1</sub> sie gesagt hat [CP wann<sub>1</sub> (dass) sie t<sub>1</sub> kommen [+wh] think you [+wh] she said has when that she come würde ]] ?  
 would  
 ALL: ‘When do you think she said she would come?’

Candidates (26-a)–(26-c) employ either *wh*-movement, or *was* insertion, or both. All lead to three violations of the tie, and hence they can all be optimal given (23); see T<sub>10</sub>.

T<sub>10</sub>: *Optionality of wh-scope marking in longer dependencies in German*

Candidates	WH-CRIT	BAR-CON	DER-ECON   FULL-INT
☞ C <sub>1</sub> : [CP wann <sub>1</sub> .. [CP t'' <sub>1</sub> .. [CP t' <sub>1</sub> .. t <sub>1</sub> ]]			***
☞ C <sub>2</sub> : [CP was <sub>1</sub> .. [CP wann <sub>1</sub> .. [CP t' <sub>1</sub> .. t <sub>1</sub> ]]			**   *
☞ C <sub>3</sub> : [CP was <sub>1</sub> .. [CP was <sub>1</sub> .. [CP wann <sub>1</sub> .. t <sub>1</sub> ]]			*   **

This result cannot be obtained with ordered global ties (C<sub>1</sub> would be optimal under the order FULL-INT ≫ DER-ECON, C<sub>3</sub> would be optimal under the order DER-ECON ≫ FULL-INT, but C<sub>2</sub> would not be optimal under any resolution of the tie), or with ordered local ties (for essentially the same reasons – there is no suborder of the tie under which C<sub>2</sub> could be optimal within the tie). However, just like ordered local ties, conjunctive local ties make wrong predictions for the competition in T<sub>4</sub> (optional *wh*-movement in French root clauses) that motivated an ordered global tie; however, this time it is not C<sub>2</sub> that would wrongly be predicted to be ill formed, but C<sub>1</sub> (due to the large number of SPC violations). Moreover, even though conjunctive local ties make the same predictions as ordered local ties in T<sub>5</sub> and T<sub>6</sub>, they are incompatible with the account in T<sub>7</sub> (relativization in Polish): C<sub>2</sub> should be expected to be blocked as suboptimal by C<sub>1</sub> (due to an additional violation of the merged constraints).

On a more general note, it is clear that conjunctive local ties face the same weaker version of the true optionality problem as ordered local ties: It must be ensured that lower-ranked constraints do not inadvertently destroy optionality. Furthermore, since there is no ranking whatsoever of the tied constraints in this approach, questions concerning learnability and language change arise: How does the child find out that two constraints in its UG endowment literally count as one, and how can a language ever create or get rid of a tie if the tie does not have any inherent structure? Speculative answers to these kinds of questions can of course been advanced (cf. Tesar & Smolensky 1993; 2000, Tesar 1998, and Müller 1997), but these answers are far from obvious.

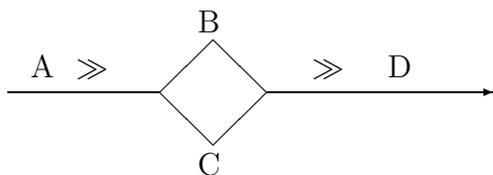
## 5.4. Disjunctive ties

The three concepts of tie discussed so far arguably predominate within optimality-theoretic syntax. Still, they do not exhaust the concepts that are logically possible and linguistically plausible. As a final example, I will now briefly introduce disjunctive ties, which are discussed in Broihier (1995) (in a local version) and Müller (1999) (in a global version). The gist of both approaches is that a candidate can be optimal on a tie of constraints if it satisfies at least one of the constraints. Disjunctive local ties can be defined as in (27) (cf. the definition of conjunctive local tie in (23)), and they can be illustrated by the diagram in (28).

(27) *Disjunctive Local Tie:*

Suppose that  $\Gamma = \langle \text{CON}_1 \gg \dots \text{CON}_i \dots \gg \text{CON}_n \rangle$  is a total constraint order in language L, and  $\text{CON}_i$  ( $1 \leq i \leq n$ )  $\in \Gamma$  is a tie  $\text{CON}_{i_1} \circ \dots \circ \text{CON}_{i_n}$ . Then,  $\text{CON}_i$  is violated by a candidate C iff there is no constraint  $\text{CON}_{i_j}$  that is satisfied by C.

(28) *Diagram of a disjunctive local tie B  $\circ$  C*



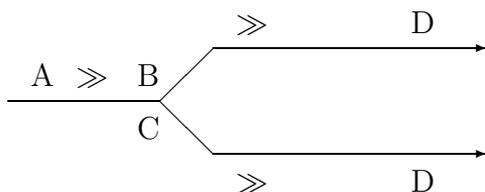
Brohier (1995) mentions this concept of tie but does not actually use it. As a matter of fact, I am not aware of any analysis in which disjunctive local ties play a role. (However, there is a strong similarity to the concept of local constraint conjunction that is employed, e.g., in Legendre, Smolensky & Wilson (1998); here, a conjoined constraint A&B is violated by a candidate if the candidate satisfies neither A nor B.)

The definition of disjunctive global ties looks more like that of ordered global ties, and so does the schematic representation (with a branching of hierarchies).

(29) *Disjunctive Global Tie:*

Suppose that  $\Gamma = \langle \text{CON}_1 \gg \dots \text{CON}_i \dots \gg \text{CON}_n \rangle$  is a partial constraint order in language L, and  $\text{CON}_i$  ( $1 \leq i \leq n$ )  $\in \Gamma$  is a tie  $\text{CON}_{i_1} \circ \dots \circ \text{CON}_{i_n}$ . Then, for every constraint  $\text{CON}_{i_j}$  in  $\text{CON}_i$ ,  $\Gamma_{i_j}$  is a constraint order of L, where  $\Gamma_{i_j}$  differs from  $\Gamma$  only in that  $\text{CON}_i$  is replaced by  $\text{CON}_{i_j}$ .

(30) *Diagram of a disjunctive global tie B  $\circ$  C*



Disjunctive global ties are proposed in Müller (1999) to account for the optionality of scrambling in German; cf. (6). On this view, scrambling is possible in a candidate (in violation of a lower-ranked STAY) if this leads to a satisfaction of at least one of a number of tied linearization constraints. (In addition, these tied constraints are assumed to be subject to a superimposed ranking that is eventually important for an account of degrees of markedness, but irrelevant for grammaticality considerations, and that can be ignored here.) Tied constraints that play a role in (6) could include the following:

- (31) a. DEF (‘Definiteness’):  
 Definite NPs precede indefinite NPs.  
 b. CASE:  
 Nominative NPs precede accusative NPs.

The optionality of scrambling in (6) now follows from a disjunctive global tie of DEF and CASE, as illustrated in tableau T<sub>11</sub>: C<sub>1</sub> is optimal if DEF is ignored, and C<sub>2</sub> if CASE is ignored.

T<sub>11</sub>: *Optionality of scrambling in German*

Candidates	DEF   CASE	STAY
☞ C <sub>1</sub> : ... keiner <sub>1</sub> den Fritz <sub>2</sub> ...	*	
☞ C <sub>2</sub> : ... den Fritz <sub>2</sub> keiner <sub>1</sub> t <sub>2</sub> ...	*	*

Note that C<sub>1</sub> and C<sub>2</sub> differ on the low-ranked constraint STAY because C<sub>2</sub> involves scrambling. This is unproblematic under a disjunctive global tie, but, as is typical with local notions of tie, a problem would arise under a disjunctive local tie, as in (27). However, it seems that we would end up with the same outcome in T<sub>11</sub> if we were to adopt the more standard ordered global tie. While this conclusion is indeed correct in the case at hand, it does not generalize to all instances of optionality discussed in Müller (1999). For instance, suppose that there is yet another tied linearization constraint FOCUS that requires that non-focussed NPs precede focussed NPs (i.e., de-focussing is viewed as another trigger for scrambling). This constraint is respected in a first version of C<sub>1</sub> (C<sub>1a</sub>) in which focus is on the object NP<sub>2</sub>, and violated in a second version of C<sub>1</sub> (C<sub>1b</sub>) in which focus is on the subject NP<sub>1</sub>. Assume furthermore that there is no other constraint on which C<sub>1a</sub> and C<sub>1b</sub> differ. Since C<sub>1b</sub> now violates all of the constraints that C<sub>1a</sub> violates, but not vice versa (i.e., since C<sub>1b</sub> is “harmonically bounded” by C<sub>1a</sub>), we wrongly expect C<sub>1b</sub> to be blocked as suboptimal by C<sub>1a</sub> under an ordered global tie (in fact, under all non-disjunctive notions of tie). No such problem arises under a disjunctive global tie, because C<sub>1a</sub> and C<sub>1b</sub> have exactly the same constraint profile in a hierarchy in which the whole tie of constraints is replaced by CASE alone.

As should be clear by now, local and global versions of disjunctive ties differ with respect to the question of whether lower-ranked constraints can destroy optionality. Interestingly, they can also be shown to make predictions that differ significantly from those we have seen so far. For example, we would expect that only  $C_6$  can be optimal in the Polish relativization tableau  $T_7$  under a disjunctive local tie; and we would expect that  $C_{1-3}$  and  $C_6$  can all be optimal in  $T_7$  under a disjunctive global tie. Apart from that, both versions of disjunctive ties have a peculiar property that renders them unsuitable as the single means to account for optionality: For a candidate that satisfies one constraint of such a tie, any other constraint in the tie becomes completely irrelevant. Thus, if, e.g., STAY and some other constraint X were to participate in a disjunctive tie, a candidate that satisfies X could violate STAY arbitrarily often, and in completely independent domains. This would hardly be an acceptable result.

## 5.5. Conclusion

All the concepts of tie discussed in this section seem to work fairly well for the subcases of optionality that they are designed for; but at the same time, they can all be shown to raise serious conceptual problems – related to complexity in the case of global ties, and related to a possible unwanted interference from lower-ranked constraints in the case of local ties. Furthermore, it has turned out that the different concepts of tie make empirical predictions that are often strikingly different and incompatible. However, a unified concept of tie seems out of reach. Therefore, I take it that there is every reason to continue looking for another way to handle optionality and alternation in optimality-theoretic syntax. One possibility is a radical rethinking of the pseudo-optionality approach.

## 6. Neutralization

Recall that the main problem with the pseudo-optionality approach is that it does not cover alternation: To account for optionality, alternating optimal candidates must belong to different candidate sets; but to account for the breakdown of optionality in certain contexts, it seems that they must belong to the same candidate set. This dilemma may look unsolvable at first sight, but such a conclusion turns out to be premature, provided that we alter some basic assumptions. Suppose that, e.g., (1-a) and (1-b) are indeed the sole optimal candidates in their respective candidate sets; but suppose furthermore that each of these candidates is also included in the other’s candidate set, where it is filtered out as suboptimal due to a fatal faithfulness violation (i.e., it deviates too much from the input). Then, a breakdown of optionality (cf. (9), in the case at hand) can be traced

back to a situation in which a candidate, say, (9-a), emerges as optimal not only in the candidate set in which it is faithful to the input, but also in the candidate set in which it is unfaithful to the input; in the latter case, (9-a) blocks the faithful (9-b) as suboptimal because (9-b) violates an even higher-ranked constraint. This means that a difference in the input can be “neutralized” in the output. Neutralization approaches of various types have been proposed by Baković (1997), Keer & Baković (1997), Baković & Keer (2001), Legendre, Smolensky & Wilson (1998), and Schmid (1998). I will briefly introduce the neutralization approach to English complementizer drop alternation that is given in the system of Baković and Keer (2001).

In addition to Grimshaw’s (1997) PURE-EP (cf. (13-a)), this approach relies on a faithfulness constraint FAITH[COMP] that is lower-ranked in English:

(32) FAITH[COMP]:

The output value of [Comp] is the same as the input value.

Simplifying slightly, we can assume that the feature [ $\pm$ Comp] specifies the subordination requirement of a predicate that embeds a clause. By assumption, [+Comp] requires embedding of CP (which, due to a high-ranked OB-HD, then forces the presence of a complementizer, exactly as in Grimshaw’s (1997) approach); and [–Comp] requires embedding of IP. FAITH[COMP] can then be understood as requiring faithfulness to the subordination requirement. Given that bridge verbs like *think* can be equipped with a [+Comp] or [–Comp] feature, and that this difference creates two candidate sets, both candidates in (1) emerge as optimal – each in the candidate set where it is faithful to the [ $\pm$ Comp] specification in the input. This is shown in T<sub>12</sub> and T<sub>13</sub>.

T<sub>12</sub>: *Optionality of complementizer drop in English: [+Comp] in the input*

Candidates	PURE-EP	FAITH[COMP]
☞ C <sub>1</sub> : ... think <sub>[+Comp]</sub> [CP that [IP ...		
C <sub>2</sub> : ... think <sub>[+Comp]</sub> [IP ...		*!

T<sub>13</sub>: *Optionality of complementizer drop in English: [–Comp] in the input*

Candidates	PURE-EP	FAITH[COMP]
C <sub>1</sub> : ... think <sub>[–Comp]</sub> [CP that [IP ...		*!
☞ C <sub>2</sub> : ... think <sub>[–Comp]</sub> [IP ...		

Thus far, the approach is similar to standard pseudo-optionality approaches. However, as soon as PURE-EP becomes relevant, a difference emerges: In the case of conflict, the optimal candidate will violate FAITH[COMP] and ignore the [ $\pm$ Comp] specification of the

input in order to satisfy the higher-ranked PURE-EP. This way, the difference in [ $\pm$ Comp] specification in the input is neutralized in topicalization contexts, and CP embedding becomes optimal in both candidate sets (cf. (9-a) vs. (9-b)). See T<sub>14</sub> and T<sub>15</sub>.

T<sub>14</sub>: *Complementizer drop and embedded topicalization in English: faithful winner*

Candidates	PURE-EP	FAITH[COMP]
☞C <sub>1</sub> : ... think <sub>[+Comp]</sub> [CP that [IP to John [IP ...		
C <sub>2</sub> : ... think <sub>[+Comp]</sub> [IP to John [IP ...	*!	*

T<sub>15</sub>: *Complementizer drop and embedded topicalization in English: unfaithful winner*

Candidates	PURE-EP	FAITH[COMP]
☞C <sub>1</sub> : ... think <sub>[-Comp]</sub> [CP that [IP to John [IP ...		*
C <sub>2</sub> : ... think <sub>[-Comp]</sub> [IP to John [IP ...	*!	

It seems likely that other instances of syntactic alternations can be treated equally well in a neutralization approach. Furthermore, it is worth noting that the concept of neutralization has been successfully employed to account for another recalcitrant problem of optimality-theoretic syntax, viz., that of deriving absolute ungrammaticality, i.e., handling situations in which there does not seem to be any well-formed candidate in a candidate set (cf. in particular Legendre, Smolensky & Wilson 1998).

However, the neutralization approach is not completely unproblematic either because it increases complexity in various domains. First, given that inputs that differ only with respect to an abstract selection feature create different candidate sets, it is clear that there will now have to be more candidate sets to be considered for one and the same surface form. Second, the candidate sets also get bigger: GEN must be somewhat more powerful than hitherto assumed; e.g., it must be allowed to selectively ignore all kinds of information present in the input. These pieces of information may not be confined to features (as in the case at hand); they may also be semantically contentful items, which then necessitates a more liberal definition of candidate set (cf. Legendre, Smolensky & Wilson 1998, who suggest that the notion of “having identical LFs” in (8) is to be replaced by the weaker notion of “targeting identical LFs”). Ultimately, on this view everything (including Chomsky’s famous *ba*) competes with everything else in each competition, but all the candidates that are unfaithful to the input to a significant degree will of course never get a chance to be optimal. Third, and perhaps most importantly, the neutralization approach systematically creates vacuous ambiguities in well-formed sentences: There is an optimal candidate that is faithful to the input, and there is another optimal candidate in a (minimally) different candidate set that is not. This ambiguity might be considered

problematic from the point of view of parsing and learnability. In view of this, Prince & Smolensky (1993) suggest a principle of “Input Optimization” (standardly referred to in phonology as “Lexical Optimization”) that removes vacuous ambiguities of this kind. Input Optimization essentially states that in cases where different inputs lead to optimal outputs that are identical, the language learner determines the optimal output that is most harmonic, and selects the input of this output as the underlying form. Thus, in the case at hand, Input Optimization filters out the competition in  $T_{15}$  that has an unfaithful winner in favour of the competition in  $T_{14}$  that has a faithful, more harmonic winner. This, however, is a non-trivial extension of the standard optimality-theoretic system: In addition to the standard mechanism of what one might call “first-order optimization,” which chooses the best candidate in a candidate set, we now also have a procedure of “second-order optimization,” which chooses the best competition in a set of competitions (that is defined by identity of the optimal candidate).

## 7. Conclusion

The existence of optionality, viewed as a situation in which there is more than one winner in a candidate set, poses a problem for optimality-theoretic syntax. This problem is solved in a brute force way in the pseudo-optionality approach by postulating that optionality involves winners of different candidate sets (i.e., no competition) after all. However, this solution is called into doubt by the fact that optionality is often to be analyzed as alternation, i.e., breakdown of optionality in certain contexts, because this strongly suggests competition. The approaches in terms of true optionality, (global and local) ties, and neutralization all acknowledge the existence of competition in cases of optionality, and they are all in principle capable of handling both optionality and its breakdown. However, they raise other problems: In the case of true optionality and, to a lesser extent, local ties, it must be ensured that the alternating optimal candidates exhibit a completely identical behaviour on all constraints, even ones that are ranked very low, which may turn out to be an extremely difficult task in many cases. With global ties, a substantial increase in complexity arises because a language with just a few instances of syntactic optionality will already require the simultaneous presence of many grammars (i.e., different rankings). Similarly, neutralization raises complexity issues because it postulates more and bigger candidate sets and vacuous ambiguities that can only be avoided by additional second-order optimization.

Thus, it seems fair to conclude that optionality continues to be a problem in optimality-theoretic syntax. The situation is far from being hopeless, though. It might be that interference problems with true optionality and local ties turn out not to be that severe after all once larger fragments of optimality-theoretic grammars are developed; it might

be that complexity problems with global ties and neutralization can be reduced or, in fact, shown to be spurious once our understanding of these issues broadens; and it might even be that pseudo-optionality will eventually prove the right approach for a significant subset of instances of syntactic optionality (especially those that do not involve alternation). It is also worth noting that an attempt is made in the recent literature to combine different approaches to optionality in optimality-theoretic syntax (e.g., pseudo-optionality and true optionality, pseudo-optionality and neutralization, local ties and neutralization, ordered global ties and disjunctive local ties, etc.). This might prove a viable strategy, even though it is incompatible with the famous German saying “In Gefahr und höchster Not bringt der Mittelweg den Tod.” (In greatest danger or in dread, take the middle road and you’re dead.)

### Acknowledgements

Work on this article was supported by DFG grants MU 1444/1-1,2-1. I would like to thank Sten Vikner, Tanja Schmid, Wolfgang Sternefeld and Lisa Cheng & Rint Sybesma for their comments.

### An Optionality in Optimality-Theoretic Syntax Bibliography

- Ackema, P. & A. Neeleman (1998). Optimal Questions. *Natural Language and Linguistic Theory* 16, 443-490.
- Baker, M. (1988). *Incorporation*. Chicago: The University of Chicago Press.
- Baković, E. (1997). Faithfulness & Optionality. Ms., Rutgers University.
- Baković, E. & E. Keer (2001). Optionality and Ineffability. In *Optimality-Theoretic Syntax*, eds. G. Legendre, J. Grimshaw & S. Vikner, 97-112. Cambridge, Mass.: MIT Press.
- Broihier, K. (1995). Optimality Theoretic Rankings with Tied Constraints: Slavic Relatives, Resumptive Pronouns and Learnability. Ms., MIT, Cambridge, Mass.
- Büring, D. (2001). Let’s Phrase It! Focus, Word Order, and Prosodic Phrasing in German Double Object Constructions. In *Competition in Syntax*, eds. G. Müller & W. Sternefeld, 69-105. Berlin: Mouton de Gruyter.
- Chang, L. (1997). *Wh-in-situ Phenomena in French*. Doctoral dissertation, University of British Columbia.
- Cheng, L. & J. Rooryck (2000). Licensing Wh-in-situ. *Syntax* 3, 1-19.
- Choi, H.-W. (1996). *Optimizing Structure in Context: Scrambling and Information Structure*. Dissertation, Stanford University.
- Choi, H.-W. (1999). *Optimizing Structure in Context. Scrambling and Information Structure*. Stanford: CSLI Publications.

- Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. (1986). *Barriers*. Cambridge, Mass.: MIT Press.
- Chomsky, N. (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, N. (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chomsky, N. (1999). Derivation by Phase. MIT Occasional Papers in Linguistics 18. MIT, Cambridge, Mass.
- Collins, C. (1994). Economy of Derivation and the Generalized Proper Binding Condition. *Linguistic Inquiry* 25, 45-61.
- Collins, C. (1997). *Local Economy*. Cambridge, Mass.: MIT Press.
- Costa, J. (1998). *Word Order Variation*. Doctoral dissertation, University of Leiden. (LOT dissertations, 14)
- Costa, J. (2001). The Emergence of Unmarked Word Order. In *Optimality-Theoretic Syntax*, eds. G. Legendre, J. Grimshaw & S. Vikner, 171-203. Cambridge, Mass.: MIT Press.
- Epstein, S. (1992). Derivational Constraints on A-bar Chain Formation. *Linguistic Inquiry* 23, 235-259.
- Fox, D. (1995). Economy and Scope. *Natural Language Semantics* 3, 283-341.
- Grimshaw, J. (1993). Minimal Projection, Heads, and Optimality. Ms., Rutgers University. New Brunswick, New Jersey.
- Grimshaw, J. (1997). Projection, Heads, and Optimality. *Linguistic Inquiry* 28, 373-422.
- Grimshaw, J. (1999). Clause Structure and Competition in Syntax. Ms., Rutgers University. New Brunswick, New Jersey.
- Grimshaw, J. & V. Samek-Lodovici (1995). Optimal Subjects. In *Papers in Optimality Theory*, eds. J. Beckmann, L. Walsh-Dickie & S. Urbanczyk, 589-605. University of Massachusetts, Amherst: UMass Occasional Papers in Linguistics 18.
- Grimshaw, J. & V. Samek-Lodovici (1998). Optimal Subjects and Subject Universals. In *Is the Best Good Enough?*, eds. P. Barbosa et al., 193-219. Cambridge, Mass.: MIT Press & MITWPL.
- Heck, F. (1998). *Relativer Quantorenskopis im Deutschen – Optimalitätstheorie und die Syntax der Logischen Form*. M.A. thesis, Universität Tübingen.
- Heck, F. (2001). Quantifier Scopepe in German and Cyclic Optimization. In *Competition in Syntax*, eds. G. Müller & W. Sternefeld, 175-209. Berlin: Mouton/de Gruyter.
- Höhle, T. (2000). The W-...W-Construction: Appositive or Scope Indicating? In *Wh-Scope Marking*, eds. U. Lutz, G. Müller & A. von Stechow, 249-270. Amsterdam: Benjamins.
- Keer, E. & E. Baković (1997). Have FAITH in Syntax. *Proceedings of WCCFL* 16, 255-269.
- Kitahara, H. (1993). Deducing ‘Superiority’ Effects from the Shortest Chain Requirement.

- Harvard Working Papers in Linguistics* 3, 109-119.
- Kitahara, H. (1997). *Elementary Operations and Optimal Derivations*. Cambridge, Mass.: MIT Press.
- Larson, R. (1988). On the Double Object Construction. *Linguistic Inquiry* 19, 335-391.
- Lasnik, H. & M. Saito (1992). *Move  $\alpha$* . Cambridge, Mass.: MIT Press.
- Legendre, G. (1998). Why French Stylistic Inversion is Optimal. Ms., Johns Hopkins University.
- Legendre, G. (2001). An Introduction to Optimality Theory in Syntax. In *Optimality-Theoretic Syntax*, eds. G. Legendre, J. Grimshaw & S. Vikner, 1-27. Cambridge, Mass.: MIT Press.
- Legendre G., C. Wilson, P. Smolensky, K. Homer & W. Raymond (1995). Optimality and Wh-Extraction. In *Papers in Optimality Theory*, eds. J. Beckman, L. Walsh-Dickie & S. Urbanczyk, 607-635. University of Massachusetts, Amherst: UMass Occasional Papers in Linguistics 18.
- Legendre, G., P. Smolensky & C. Wilson (1998). When is Less More? Faithfulness and Minimal Links in Wh-Chains. In *Is the Best Good Enough?*, eds. P. Barbosa et al., 249-289. Cambridge, Mass.: MIT Press & MITWPL.
- McDaniel, D. (1989). Partial and Multiple Wh-Movement. *Natural Language and Linguistic Theory* 7, 565-604.
- Müller, G. (1997). Partial Wh-Movement and Optimality Theory. *The Linguistic Review* 14, 249-306.
- Müller, G. (1999) Optimality, Markedness, and Word Order in German. *Linguistics* 37, 777-818.
- Müller, G. (2000). *Elemente der optimalitätstheoretischen Syntax*. Tübingen: Stauffenburg.
- Müller, G. & Wolfgang Sternefeld (1996). A-bar Chain Formation and Economy of Derivation. *Linguistic Inquiry* 27, 480-511.
- Nakamura, M. (1998). Reference Set, Minimal Link Condition, and Parametrization. In *Is the Best Good Enough?*, eds. P. Barbosa et al., 291-313. Cambridge, Mass.: MIT Press & MITWPL.
- Pesetsky, D. (1997). Optimality Theory and Syntax: Movement and Pronunciation. In *Optimality Theory. An Overview*, eds. D. Archangeli & D. T. Langendoen, 134-170. Oxford: Blackwell.
- Pesetsky, D. (1998). Some Optimality Principles of Sentence Pronunciation. In *Is the Best Good Enough?*, eds. P. Barbosa et al., 337-383. Cambridge, Mass.: MIT Press & MITWPL.
- Prince, A. & P. Smolensky (1993). *Optimality Theory. Constraint Interaction in Generative Grammar*. Ms., Rutgers University.

- Rizzi, L. (1992). Argument/Adjunct (A)symmetries. *Proceedings of NELS 22*, 365-381.
- Rochemont, M. (1989). Topic Islands and the Subjacency Parameter. *Canadian Journal of Linguistics* 34, 145-170.
- Samek-Lodovici, V. (1996). *Constraints on Subjects. An Optimality-Theoretic Analysis*. Dissertation, Rutgers University: New Brunswick, New Jersey.
- Schmid, T. (1998). West Germanic “Infinitivus Pro Participio” (IPP) Constructions in Optimality Theory. In *Proceedings of Console VII*, ed. Tina Cambier-Langeveld, Anikó Lipták, Michael Redford & Erik Jan van der Torre, 229-244. Leiden: SOLE.
- Schmid, T. (2000). Die Ersatzinfinitivkonstruktion im Deutschen. *Linguistische Berichte* 183, 325-351.
- Sells, P., J. Rickford & T. Wasow (1996). An Optimality Theoretic Approach to Variation in Negative Inversion in AAVE. *Natural Language and Linguistic Theory* 14, 591-627.
- Speas, M. (1990). *Phrase Structure in Natural Language*. Dordrecht: Kluwer.
- Sternefeld, W. (1997). Comparing Reference Sets. In *Economy in Linguistic Theory*, eds. M. Bierwisch, H.-M. Gärtner & C. Wilder, 81-114. Berlin: Akademie Verlag.
- Stowell, T. (1981). *Origins of Phrase Structure*. Doctoral dissertation, MIT, Cambridge, Mass.
- Tesar, B. (1998). Error-Driven Learning in Optimality Theory via the Efficient Computation of Optimal Forms. In *Is the Best Good Enough?*, eds. P. Barbosa et al., 421-435. Cambridge, Mass.: MIT Press & MITWPL.
- Tesar, B. & P. Smolensky (1993). The Learnability of Optimality Theory.: An Algorithm and Some Basic Complexity Results. Technical Report, Computer Science Department, University of Colorado, Boulder.
- Tesar, B. & P. Smolensky, (2000). *Learnability in Optimality Theory*. Cambridge, Mass.: MIT Press.
- Vikner, S. (2001). V-to-I Movement and ‘Do’-Insertion in Optimality Theory. In *Optimality-Theoretic Syntax*, eds. G. Legendre, J. Grimshaw & S. Vikner, 427-464. Cambridge, Mass.: MIT Press.