OPTIONALITY IN OPTIMALITY-THEORETIC SYNTAX

by *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 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 SpecCP position, accompanied by *wh*-scope marker *was* inserted in the target SpecCP position:

(3)
a. Wen$_1$ glaubst du [$_{CP}$ t'$_1$ daß man t$_1$ einladen sollte]?
whom think you that one invite should
b. Was$_1$ glaubst du [$_{CP}$ wen$_1$ (daß) man t$_1$ 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$_1$ as-tu vu t$_1$?
who have-you seen
b. Tu as vu qui?
you have seen who
BOTH: 'Who did you see?'

Fifth, PP extraposition from NP is optional in English:

(5)
a. [$_{NP}$ A review [$_{PP}$ of this article]] came out yesterday
b. [$_{NP}$ A review t$_1$] came out yesterday [$_{PP}$ of this article]$_1$

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

(6)
a. daß [$_{NP1}$ keiner] [$_{NP2}$ den Fritz] gesehen hat
that no-one-NOM ART-ACC Fritz seen has
b. daß [$_{NP2}$ den Fritz] [$_{NP1}$ keiner] t$_2$ gesehen hat
that ART-ACC Fritz no-one-NOM 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 α, applies optionally throughout, subject to the requirement that the resulting syntactic object does not violate any constraints of grammar, creating 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 α to a given item α, and by leaving a unaffected.

The Minimalist Program (Chomsky 1995) takes the opposite direction: 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. One obvious and simple way out is to reanalyze optional movement transformations as obligatory instances of Move triggered by optional features. Concerning, for instance, 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.) 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. Clearly, (4a) involves more applications of Move than (4b), 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 defining candidate sets in terms of identity of meaning (or LF; see e.g., Fox 1995 and Reinhart 1998), 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).

Below I turn to Optimality theory, which 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.

One issue that is important in the context of optionality, and that I have not mentioned yet is the fact that optionality is often limited in the sense that in some contexts it breaks down ("syntactic alternation"); I discuss this in section 3.

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 1999):

(7) OPTIMALITY
A candidate C_i is optimal with respect to a constraint ranking $<Con_1>> \dots Con_i \dots >> Con_n>$ iff there is no candidate C_j in the same candidate set such that:
a. There is a constraint Con_k that C_j satisfies better than C_i; and
b. There is no constraint Con_l outranking 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 specifications, 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), 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), Choi (1996), Samek-Lodovici (1996), Büring (1997), Costa (1998, 1999), Legendre (1998) and Heck (1999), 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), 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 (8a) 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*). Still, this does not yet account for (4)–(6), where the lexical items are identical. 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 (8b).

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 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 (8b) is much harder. Any attempt at attributing a standard question interpretation to (4a), and an echo question interpretation to (4b), is disconfirmed by the empirical facts, according to which (4b) can receive a genuine question interpretation; but it might not be entirely futile to look for functional differences between the two candidates; cf. Chang (1997) for discussion.

However, pseudo-optionality raises an important problem: In all the cases discussed so far, optionality of two constructions may break down in certain contexts. I will refer to this phenomenon of restricted optionality (i.e., co-occurrence of optionality and its breakdown) as **syntactic alternation**. Thus, for most speakers of English, the complementizer 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]$_1$ she gave a book t$_1$
b. * I think — [$_{PP}$ to John]$_1$ she gave a book t$_1$
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. Höhle 1990, Rizzi 1992) and subject islands (cf. McDaniel 1989):

(11)
a. ? Wen$_1$ glaubst du nicht [$_{CP}$ t'$_1$ daß man t$_1$ einladen sollte]?
whom think you not that one invite should
'Who don't you think one should invite?'
b. * Was$_1$ glaubst du nicht [$_{CP}$ wen$_1$ (daß) man t$_1$ einladen
[+wh] think you not whom (that) one invite
sollte]?
should
c. ? [$_{PP}$ Mit wem]$_1$ ist es schade [$_{CP}$ t'$_1$ daß sie t$_1$ gesprochen
with whom is it a pity that she talked
hat]?
has
'Who is it a pity that she talked to?'
d. * Was$_1$ ist es schade [$_{CP}$ [$_{PP}$ mit wem]$_1$ sie t$_1$ gesprochen
[+wh] is it a pity with whom she talked
hat]?
has

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$_1$ es-tu venu t$_1$?
why have-you come
'Why did you come?'
b. * Tu es venu pourquoi$_1$?
you have come why
c. Je me demande [$_{CP}$ qui$_1$ C tu as vu t$_1$]
I ask myself who you have seen
'I wonder who you saw.'
d. * Je me demande [$_{CP}$ — (que) tu as vu qui$_1$]
I ask myself (that) you have seen who

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 presupposes that the two candidates in the example pairs in (9)–(12) do in fact compete, and thus belong to one and the same candidate set, an assumption that must be abandoned in a pseudo-optionality approach. Thus, we end up with the dilemma that it is not at all clear what candidate can block an ungrammatical sentence like (9d) if (9c) is not part of the same candidate set. In a nutshell, if syntactic optionality is reanalyzed as pseudo-optionality, with different winners of different competitions, the question arises of why optionality may turn into obligatoriness in certain contexts.

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. Such an approach has been pursued by Grimshaw (1997) and Vikner (1999).

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 actually 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.

Assuming that a finite IP can be the sister of V, (1a) and (1b) can now have an identical constraint profile (in fact, neither of the candidates violates any of the constraints in (13)); cf. tableau T1.

However, recall that the optionality of complementizer drop breaks down in certain contexts; cf. (9). And indeed, Grimshaw (1997) notes that *that* becomes obligatory in embedded topicalization contexts because the complementizer drop construction will have to fatally violate either OB-HD (with CP embedding), TOP-SCOPE (with the topic in situ), or PURE-EP (with topic adjunction to an IP that is the sister of the matrix V); cf. (9a) vs. (9b) and tableau T2.

As a second example, consider Vikner's (1999) extension of this proposal that captures an interesting cross-linguistic correlation: It seems that only languages with complementizer drop 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 bu the French data in (14):

(14)
a. Je crois que l'acteur voit vraiment le film
I believe that the actor sees really the film
'I think that the actor really sees the movie'
b. * Je crois — l'acteur voit vraiment le film
I believe the actor sees really the film

Two additional constraints become necessary: one that punishes movement of a lexical head (LX–MV, 'Lexical Movement'), and one that prohibits lowering of I features to V (PR-BD, 'Proper Binding'); the ranking is LX-MV >> PR-BD in English, and PR-BD >> 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**. In addition, it is theoretically attractive since it does not require any additional assumptions, given a notion of optimality like (7) that permits more than one winner in a single competition.

Still, there is a serious 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 T1; 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 T1, 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.

T1. OPTIONALITY OF COMPLEMENTIZER DROP IN ENGLISH (Grimshaw)

Candidates	PURE-EP	TOP-SCOPE	OB-HD	STAY
⇒ C_1: think [$_{CP}$ that [$_{IP}$ John is a fool]]				
⇒ C_2: think [$_{IP}$ John is a fool]]				
*C_3: think [$_{CP}$ – [$_{IP}$ John is a fool]]			*!	
*C_4: think [$_{CP}$ is$_2$ [$_{IP}$ John t$_2$ a fool]]	*!			*

T2. COMPLEMENTIZER DROP AND EMBEDDED TOPICALIZATION IN ENGLISH

Candidates	PURE-EP	TOP-SCOPE	OB-HD	STAY
⇒ C_1: think [$_{CP}$ that [$_{IP}$ to John [$_{IP}$ she gave ... t]]]				*
*C_2: think [$_{CP}$ – [$_{IP}$ to John [$_{IP}$ she gave ... t]]]			*!	*
*C_3: think [$_{IP}$ she gave ... to John]]		*!		
*C_4: think [$_{IP}$ to John [$_{IP}$ she gave ... t]]	*!			*

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. In what follows, I will note a tie of A and B as "A ∞ B" (in tableaux, ties will be represented by vertical dotted lines). Various concepts of tie have been proposed, which differ both conceptually, and empirically. However in the abstract tableau T3, with the ranking A >> B ∞ C >> 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.

T3. TIES

Candidates	A	B	C	D
⇒ C_1			*	
⇒ 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.

5.1. Ordered global ties

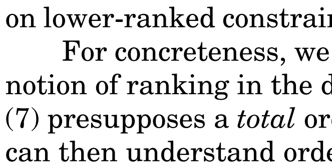
Ordered global ties have been adopted by Sells, Rickford & Wasow (1996), Ackema & Neeleman (1998), Schmid (1999) and others. They correspond to a notion of tie that is suggested in Prince & Smolensky (1993). The basic idea is that a constraint ranking that exhibits a tie of two constraints $C_1 \infty C_2$ is underspecified; it is an abbreviation that encodes the simultaneous presence of two hierarchies that exhibit the rankings $C_1 >> C_2$ and $C_2 >> C_1$. 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 Γ = <Con_1 >> … Con_i … >> Con_n> is a partial constraint order in language L, and Con_i ($1 \leq i \leq n$) ∈ Γ is a tie Con_{i1} o…o Con_{in}. Then, for every suborder O of the constraints in Con_i, Γ_O is a constraint order of language L, where Γ_O differs from Γ only in that 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. A diagram that illustrates how ordered global ties work is given in (16):

(16) DIAGRAM OF AN ORDERED GLOBAL TIE B ∞ 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 a different 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 constitutent 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 ∞ Q-MARKING >> 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. In the case when SPC outranks Q-MARKING, the candidate (4b) that dispenses with *wh*-movement (and thereby minimizes SPC violations) is optimal; if, however, Q-MARKING outranks SPC, the candidate (4a) that respects Q-MARKING by moving the *wh*-phrase emerges as the winner. This is shown in tableau T4, which should be viewed as an underspecified representation that simultaneously encodes two subtableaux — one in which C_1 is optimal, and one in which C_2 is optimal. (In line with this, (!) should be taken to indicate a violation that is fatal only under one resolution of the tie.)

T4. OPTIONAL *WH*-MOVEMENT IN FRENCH ROOT CLAUSES

Candidates	SPC	Q-MARKING	Q-SCOPE
⇒ C_1: qui_1 Aux_2 NP t_2 V t_1	*(!)*****		
⇒ C_2: – – NP Aux_2 V qui_1		*(!)	*
*C_3: qui_1 – NP Aux_2 V t_1	*(!)**	*(!)	
*C_4: – Aux_2 NP t_2 V qui_1	*(!)**	*(!)	*

Two remarks are in order concerning constraint satisfaction in this tableau. First, since the SPC counts nodes crossed in the course of movement, the exact number of SPC violations in tableau T4 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 number of *'s under SPC should be viewed as rough estimations. However, it seems uncontroversial that C_1, which has both *wh*-movement and I-to-C movement, involves the largest number of nodes crossed; that C_2, which leaves both items in situ, involves the smallest number of nodes crossed; and that C_3 and C_4, which exhibit *wh*-movement and V-to-I raising, respectively, are somewhere in between C_1 and C_2 with respect to the SPC.

Second, whereas Q-SCOPE is satsified 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^0 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 SpecCP 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 constraints 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 SPC ∞ Q-MARKING that results in a total order SPC >> Q-MARKING >> 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 an often noted 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 simulteneous 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 and alternation in every natural language, this result sheds doubt on the applicability of ordered global ties as a general mechanism to account for optionality.

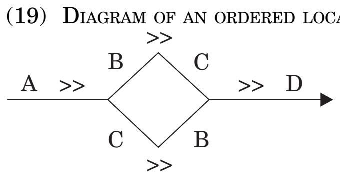
5.2. Ordered local ties

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∘B is resolved into its subrankings A >> B, B >> 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 Γ = <Con_1>> … Con_i … >> Con_n> is a total constraint order in language L, Con_i ($1 \leq i \leq n$) ∈ Γ is a tie Con_{i1} o…o Con_{in}, and O_1,…,O_n are the possible suborders of the constraints in Con_i. Then, 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 ∞ C



As an example, consider again English complementizer drop. 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 REC >> LE(CP) ∞ TEL >> 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: (1a) (= C_1 in T5) satisfies LE(CP) and violates TEL; therefore, it satisfies the tie under the resolution LE(CP) >> TEL. On the other hand, (1b) (= C_2 in T5) satisfies TEL and violates LE(CP); hence, it does not violate the tie under the resolution TEL >> LE(CP).

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 T6.

T5. Optionality of complementizer drop in English (Pesetsky)

Candidates	Rec	LE(CP)	Tel	DCP
⇒ C_1: ... think [$_{CP}$ that John ...			*	
⇒ C_2: ... think [$_{CP}$ – John ...		*		*!

T6. Complementizer drop and subject clauses in English

Candidates	Rec	LE(CP)	Tel	DCP
⇒ C_1: ... surprised me [$_{CP}$ that the earth ...			*	
*C_2: ... 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, so as to minimize Silent-t violations); cf. (21a–b). All other combinations, some of which are given in (21c–f), are ungrammatical.

(21)
a. On spotkał studenta [$_{CP}$ któremu$_1$ —on dał t_1 piątkę]
he met student whom-DAT he gave good mark
'he met the student who he gave a good mark.'
b. On spotkał studenta [$_{CP}$ –$_1$ co mu_1 on dał piątkę]
he met student that him-DAT he gave good mark
c. * On spotkał studenta [$_{CP}$ któremu$_1$ – on mu_1 dał piątkę]
he met student whom-DAT he him-DAT gave good mark
d. * On spotkał studenta [$_{CP}$ –$_1$ co on dał t_1 piątkę]
he met student that he gave good mark
e. * On spotkał studenta [$_{CP}$ któremu$_1$ co mu_1 on dał piątkę]
he met student whom-DAT that him-DAT he gave good mark
f. * On spotkał studenta [$_{CP}$ któremu$_1$ co on dał t_1 piątkę]
he met student whom-DAT that he gave good mark

Broihier and Pesetsky propose a three-way tie Silent-t ∞ LE(CP) ∞ Tel:

T7. Optionality of relativization strategies in Polish

Candidates	Rec	Silent-t	LE(CP)	Tel	DCP
⇒ C_1: wh_1 – ... t_1			*		*
⇒ C_2: –$_1$ co ... mu_1		*		*	*
*C_3: wh_1 – ... mu_1		*	*		*
*C_4: –$_1$ co ... t_1	*!			*	*
*C_5: wh_1 co ... mu_1		*	*	*	
*C_6: wh_1 co ... t_1			*	*	

A high-ranked Rec strictly prohibits simultaneous deletion of the oblique relative pronoun and its trace (as in C_4). 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_5, C_6). It must be shown that C_1 and C_2 in T7 do not violate the three-way tie, whereas C_3, C_5, and C_6 do. If the tie is resolved into a suborder in which Tel is highest, C_1 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_1 satisfies the tie; and under a resolution Silent-t >> LE(CP) >> Tel, C_4 satisfies the tie, but irrelevantly so because C_4 fatally violates Rec. Finally, if LE(CP) is highest, we hope that C_2 optimally satisfies the tie. However, unfortunately, this is not the case: Under any ranking of Silent-t and Tel, C_4 satisfies the tie better than C_2. Hence, it is wrongly predicted that only C_1 is grammatical.

Intuitively, the problem is that C_4 can be optimal within the tie, thereby creating an unwanted violation of the tie by C_2, even though C_4 violates a higher-ranked constraint. To remedy this situation, we have to ensure that a candidate like C_4 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 T1 that optimality-theoretic learning strategies like Tesar & Smolensky's 1993 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 Γ = <Con_1>> ... Con_i ... >> Con_n> is a total constraint order in language L, Con_i ($1 \leq i \leq n$) ∈ Γ is a tie Con_{i1} o...o Con_{in}, and O_1,...,O_n are the possible suborders of the constraints in Con_i. Then, Con_i is violated by a candidate C iff there is no suborder O_i such that C is optimal according to <Con_1>> ... Con_{i-1}>> O_i>

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_2 in T6 (complementizer drop in subject clauses) is wrongly predicted to be optimal under the global definition of tie in 5.1; and C_2 in T4 (*wh*-in situ in French) is wrongly predicted to be suboptimal under the local definition of tie in 5.2.

5.3. Conjunctive local ties

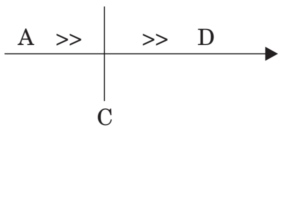
In addition to ordered global ties, Prince & Smolensky (1993) envisage a second concept of tie that relies on a "crucial nonranking" of constraints. This concept has been adopted by Legendre, Wilson, Smolensky, Homer & Raymond (1995), Legendre, Smolensky & Wilson (1998) and Müller (1997). In the first 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 approach 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 Γ = <Con_1>> ... Con_i ...>> Con_n> is a total constraint order in language L, and Con_i ($1 \leq i \leq n$) ∈ Γ is a tie Con_{i1} o...o Con_{in}. Then, Con_i is violated by a candidate C iff there is a constraint Con_{ij} that is violated by C.

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

(24) Diagram of a conjunctive local tie B ∞ 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 SpecCP, and $C_{[+wh]}$ requires filling of C or SpecCP.
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 postulated for Standard German is Wh-Crit >> Bar-Con >> Der-Econ ∞ Full-Int. Assuming that the *wh*-scope marker *was* ('what') is an expletive, each instance of *was* 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 (3a) violates Der-Econ twice, partial *wh*-movement accompanied by *was* insertion as in (3b) 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; see T8.

T8. Optionality of *wh*-scope marking in German

Candidates	Wh-Crit	Bar-Con	Der-Econ	Full-Int
⇒ C_1: [$_{CP}$ wen_1 ... [$_{CP}$ t'_1 daß ... t_1]]			**	
⇒ C_2: [$_{CP}$ was_1 ... [$_{CP}$ wen_1 (daß) ... t_1]]			*	*
*C_3: [$_{CP}$ was_1 ... [$_{CP}$ daß ... wen_1]]	*!			*
*C_4: [$_{CP}$ – ... [$_{CP}$ wen_1 ... t_1]]	*!		*	
*C_5: [$_{CP}$ was1 ... [$_{CP}$ was_1 ... wen_1]]	*!			**
*C_6: [$_{CP}$ – ... [$_{CP}$ daß ... wen_1]]	*!*			

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 >> Der-Econ; and *wh*-scope marking is correctly predicted to block long-distance *wh*-movement in Northern German varieties under a ranking Der-Econ >> 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 the 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 (11c) and *wh*-scope marking as in (11d) will invariably violate Bar-Con at S-structure. However, if *wh*-in situ must undergo LF raising to its scope position, (11d) incurs an additional, fatal Bar-Con violation at LF that (11c) doesn't incur. This is shown in tableau T9. The argument is identical in the case of the negative island asymmetry in (11a–b).

T9. *Wh*-scope marking and subject islands in German

Candidates	Wh-Crit	Bar-Con	Der-Econ	Full-Int
⇒ C_1: [$_{CP}$ mit wem_1 ... [$_{CP}$ t'_1 daß ... t_1]]		*	**	
*C_2: [$_{CP}$ was_1 ... [$_{CP}$ mit wem_1 (daß) ... t_1]]		**!	*	*

Thus far, the *wh*-scope marking alternation in German may 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. $Wann_1$ meinst du [$_{CP}$ t''_1 daß sie gesagt hat [$_{CP}$ t'_1 daß sie t_1 kommen würde]]?
when think you that she said has that she come would
b. Was_1 meinst du [$_{CP}$ $wann_1$ (daß) sie gesagt hat [$_{CP}$ t'_1 daß sie t_1 kommen würde]]?
[+wh] think you when that she said has that she come would
c. Was_1 meinst du [$_{CP}$ was_1 sie gesagt hat [$_{CP}$ $wann_1$ (daß) sie t_1 kommen würde]]?
[+wh] think you [+wh] she said has when that

she come would
ALL: 'When do you think she said she would come?'

Candidates (26a)–(26c) 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 T10:

T10. OPTIONALITY OF *WH*-SCOPE MARKING IN LONGER DEPENDENCIES IN GERMAN

Candidates	WH-CRIT	BAR-CON	DER-ECON	FULL-INT
⇒ C_1: $[_{CP}$ $wann_1$ .. $[_{CP}$ t''_1 .. $[_{CP}$ t'_1 .. t_1]]			***	
⇒ C_2: $[_{CP}$ was_1 .. $[_{CP}$ $wann_1$ .. $[_{CP}$ t'_1 .. t_1]]			**	*
⇒ C_3: $[_{CP}$ was_1 .. $[_{CP}$ was_1 .. $[_{CP}$ $wann_1$ .. t_1]]			*	**

This result cannot be obtained with ordered global ties (C_1 would be optimal under the order FULL-INT >> DER-ECON, C_3 would be optimal under the order DER-ECON >> FULL-INT, but C_2 would not be optimal under any resolution of the tie), or with ordered local ties (for essentially the same reason — there is no suborder of the tie under which C_2 could be optimal within the tie).

However, just like ordered local ties, conjunctive local ties make wrong predictions for the competition in T4 (optional *wh*-movement in French root clauses) that motivated an ordered global tie: however, this time it is not C_2 that would wrongly be predicted to be ill-formed, but C_1 (due to the large number of SPC violations). Moreover, even though conjunctive local ties make the same predictions as ordered local ties in T5 and T6, they are incompatible with the account in T7 (relativization in Polish): C_2 should be expected to be blocked as suboptimal by C_1 (due to an additional violation of the merged constraints).

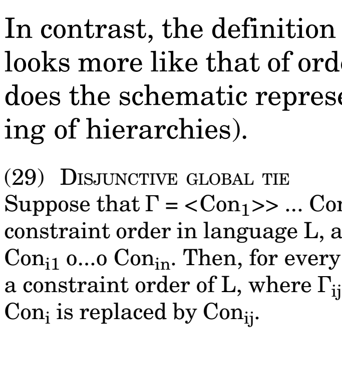
On a more general note, it is clear that conjunctive local ties face the same weaker version of the true optionality problem as ordered local ties: It must be ensured that lower-ranked constraints do not inadvertently destroy optionality. Furthermore, since there is no ranking whatsoever of the tied 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 be advanced (cf. Tesar & Smolensky 1993, Müller 1997 and Tesar 1998), 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 briefly introduce disjunctive ties, which are discussed in Broihier (1995) (in a local version) and Müller (1998) (in a global version, for scrambling in German). 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 can be illustrated by the diagram in (28).

(27) DISJUNCTIVE LOCAL TIE
Suppose that $\Gamma = <Con_1>> ... Con_i ...>> Con_n>$ is a total constraint order in language L, and Con_i ($1 \leq i \leq n$) $\in \Gamma$ is a tie Con_{i1} o...o Con_{in}. Then, Con_i is violated by a candidate C iff there is no constraint Con_{ij} that is satisfied by C.

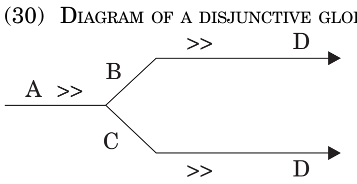
(28) DIAGRAM OF A DISJUNCTIVE LOCAL TIE B ∞ C



In contrast, 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 $\Gamma = <Con_1>> ... Con_i ...>> Con_n>$ is a partial constraint order in language L, and Con_i ($1 \leq i \leq n$) $\in \Gamma$ is a tie Con_{i1} o...o Con_{in}. Then, for every constraint Con_{ij} in Con_i, Γ_{ij} is a constraint order of L, where Γ_{ij} differs from Γ only in that Con_i is replaced by Con_{ij}.

(30) DIAGRAM OF A DISJUNCTIVE GLOBAL TIE B ∞ C



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 T7 under a disjunctive local tie; and we would expect that C_{1-3}, and C_6 can all be optimal in T7 under a disjunctive global tie. Apart from that, disjunctive concepts of tie encounter the standard problems of unwanted interference by lower-ranked constraints (local version) and complexity (global version). In addition, 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 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 in 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. The basic idea is that a difference in the input that normally creates optionality via two distinct candidate sets may under certain circumstances be "neutralized" in the output, thereby creating a breakdown in optionality. Neutralization approaches of various types have been proposed by Baković (1997), Keer & Baković (1997), Baković & Keer (1999), 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 (see especially Baković 1997).

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

(31) FAITH[SUB]:
A candidate must realize the [±Sub] specification of the input.

The feature [±Sub] specifies the subordination requirement of a predicate that embeds a clause. By assumption, [+Sub] requires embedding of CP (which, due to a high-ranked OB-HD, forces the presence of a complementizer, exactly as in Grimshaw's 1997 approach); and [–Sub] requires embedding of IP. Given that bridge verbs like *think* can be equipped with a [+Sub] or [–Sub] feature, and that this difference creates two candidate sets, both candidates in (1) emerge as optimal, each in a separate candidate set with different [±Sub] specification on *think*. This is shown in T11 and T12.

T11. *THINK* WITH [+SUB]

Candidates	PURE-EP	FAITH[SUB]
⇒ C_1: ... $think_{[+Sub]}$ $[_{CP}$ that $[_{IP}$ ...		
*C_2: ... $think_{[+Sub]}$ $[_{IP}$ ...		*!

T12. *THINK* WITH [–SUB]

Candidates	PURE-EP	FAITH[SUB]
*C_1: ... $think_{[-Sub]}$ $[_{CP}$ that $[_{IP}$ ...		*!
⇒ C_2: ... $think_{[-Sub]}$ $[_{IP}$ ...		

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[SUB] and ignore the [±Sub] specification of the input in order to satisfy the higher-ranked PURE-EP. This way, the difference in [±Sub] specification in the input is neutralized in topicalization contexts, and CP embedding becomes optimal in both candidate sets (cf. (9a) vs. (9b)):

T13. COMPLEMENTIZER DROP AND EMBEDDED TOPICALIZATION IN ENGLISH: FAITHFUL WINNER

Candidates	PURE-EP	FAITH[SUB]
⇒ C_1: ... $think_{[+Sub]}$ $[_{CP}$ that $[_{IP}$ to John $[_{IP}$ ...		
*C_2: ... $think_{[+Sub]}$ $[_{IP}$ to John $[_{IP}$ ...	*!	*

T14. COMPLEMENTIZER DROP AND EMBEDDED TOPICALIZATION IN ENGLISH: UNFAITHFUL WINNER

Candidates	PURE-EP	FAITH[SUB]
⇒ C_1: ... think$[_{-Sub]}$ $[_{CP}$ that $[_{IP}$ to John $[_{IP}$ ...		*
*C_2: ... think$[_{-Sub]}$ $[_{IP}$ 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. 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. 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 to cases where optionality breaks down 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 T14 that has an unfaithful winner in favour of the competition in T13 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, 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 serious 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 to be 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 the editors of Glot International for their comments.

An Optionality in Optimality 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*. Manuscript, Rutgers University.
Baković, E. & E. Keer (1999). Complementizers and *wh*-elements. [To appear in] *Optimality-Theoretic syntax*, edited by G. Legendre, J. Grimshaw & S. Vikner. Cambridge, Mass.: MIT Press.
Broihier, K. (1995). *Optimality theoretic rankings with tied constraints: Slavic relatives, resumptive pronouns and learnability*. Manuscript, MIT.
Büring, D. (1997). *Towards an OT account of German Mittelfeld word order*. Manuscript, Universität Köln.
Chang, L. (1997). *Wh-in-situ phenomena in French*. Doctoral dissertation University of British Columbia.
Choi, H.-W. (1996). *Optimizing structure in context: scrambling and information structure*. Doctoral dissertation, Stanford University.
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. *Principles and parameters in comparative grammar*, edited by R. Freidin, 417–454. Cambridge, Mass.: MIT Press.
Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.
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, HIL University of Leiden. [LOT International Series 14.] The Hague: Holland Academic Graphics.
Costa, J. (1999). Word order typology in Optimality Theory. [To appear in] *Optimality-Theoretic syntax*, edited by G. Legendre, J. Grimshaw & S. Vikner. 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*. Manuscript, Rutgers University.
Grimshaw, J. (1997). Projection, heads, and optimality. *Linguistic Inquiry* **28**, 373–422.
Grimshaw, J. (1999). *Clause structure and competition in syntax*. Manuscript, Rutgers University.
Grimshaw, J. & V. Samek-Lodovici (1995). Optimal subjects. *Papers in Optimality Theory*, edited by J. Beckmann, L. Walsh-Dickie & S. Urbanczyk, 589–605. [UMass Occasional Papers in Linguistics 18.]
Heck, F. (1998). *Relativer Quantorenskopus im Deutschen — Optimalitätstheorie und die Syntax der Logischen Form*. M.A. thesis, Universität Tübingen.
Heck, F. (1999). Quantifier scope in German and cyclic optimization. [To appear in] *Competition in Syntax*, edited by G. Müller & W. Sternefeld. Berlin: Mouton de Gruyter.
Höhle, T. (1990). *Die w ... w-Konstruktion im Deutschen*. Handout of paper presented at DGfS Saarbrücken.
Keer, E. & E. Baković (1997). Have faith in syntax. Paper presented at *WCCFL 16*.
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 α*. Cambridge, Mass.: MIT Press.
Legendre, G. (1998). *Why French stylistic inversion is optimal*. Manuscript, Johns Hopkins University.
Legendre G., C. Wilson, P. Smolensky, K. Homer & W. Raymond (1995). Optimality and *wh*- Extraction. *Papers in Optimality Theory*, edited by J. Beckman, L. Walsh-Dickie & S. Urbanczyk, 607–635. [UMass Occasional Papers in Linguistics 18.]
Legendre, G., P. Smolensky & C. Wilson (1998). When is less more? Faithfulness and minimal links in *wh*-chains. *Is the best good enough?*, edited by P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, 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. (1998). *German word order and Optimality Theory*. Manuscript, Universität Stuttgart. [Working paper no. 126, SFB 340, Stuttgart/Tübingen.]
Müller, G. (1999). *Elemente der optimalitätstheoretischen Syntax*. Manuscript, Universität Stuttgart. [To be published by Stauffenburg, Tübingen.]
Müller, G. & W. 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. *Is the best good enough?*, edited by P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, 291–313. Cambridge, Mass.: MIT Press & MITWPL.
Pesetsky, D. (1997). Optimality Theory and syntax: movement and pronunciation. *Optimality Theory. An overview*, edited by D. Archangeli & D. T. Langendoen, 134–170. Oxford: Blackwell.
Pesetsky, D. (1998). Some optimality principles of sentence pronunciation. *Is the best good enough?*, edited by P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, 337–383. Cambridge, Mass.: MIT Press & MITWPL.
Prince, A. & P. Smolensky (1993). *Optimality Theory. Constraint interaction in generative grammar*. Manuscript, Rutgers University.
Reinhart, T. *Wh*-in-situ in the framework of the Minimalist Program. *Natural Language Semantics* **6**, 29–56.
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*. Doctoral dissertation, Rutgers University.
Schmid, T. (1998). West Germanic 'Infinitivus Pro Participio' (IPP) constructions in Optimality Theory. [To appear in] *Proceedings of Console 1998*.
Schmid, T. (1999). *Ersatzinfinitiv und Optimalität*. Manuscript, Universität Stuttgart.
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. *Economy in linguistic theory*, edited by M. Bierwisch, H.-M. Gärtner & C. Wilder, 81–114. Berlin: Akademieverlag.
Stowell, T. (1981). *Origins of phrase structure*. Doctoral dissertation, MIT.
Tesar, B. (1998). Error-driven learning in Optimality Theory via the efficient computation of optimal forms. *Is the best good enough?*, edited by P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, 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.
Vikner, S. (1999). V-to-I movement and 'do'-insertion in Optimality Theory. [To appear in] *Optimality-Theoretic syntax*, edited by G. Legendre, J. Grimshaw & S. Vikner. Cambridge, Mass.: MIT Press.