

Partial *wh*-movement and Optimality Theory¹

GEREON MÜLLER

Abstract

Partial wh-movement constructions in German and Hungarian exhibit a number of properties that are unexpected under standard approaches to movement. In contrast, I will show that these properties follow directly under an optimality-theoretic approach to wh-dependencies. This approach is primarily devised so as to account for languages like English, Korean, and Bulgarian, and centers around six general and commonly accepted constraints (PROJ-PRIN, DER-ECON, WH-CRIT, FULL-INT, MIN-CHAIN, and BAR-CON. However, it turns out that the properties of partial wh-movement constructions of both the German and the Hungarian type correspond exactly to one specific ranking of these constraints. On a more general note, the analysis presented here provides arguments for postulating complete derivations as members of the reference set, and for constructing the reference set via LF identity.

-
1. For helpful comments and discussion, I would like to thank Veneeta Dayal, Gisbert Fanselow, Caroline Féry, Hans-Martin Gärtner, Gabriela Hermon, Junko Itô, Shin-Sook Kim, Anoop Mahajan, Armin Mester, Henk van Riemsdijk, Ian Roberts, Bernhard Rohrbacher, Joachim Sabel, Ken Safir, Vieri Samek-Lodovici, Peter Staudacher, Arnim von Stechow, Wolfgang Sternefeld, Imre Szigeti, Sten Vikner, two careful reviewers for *The Linguistic Review*, and the audiences of the 1995 GGS meeting at the University of Jena, the 11th Comparative Germanic Syntax Workshop at Rutgers University, the Linguistics Colloquium at Brown University, the Workshop on *Wh*-Scope Marking at Universität Tübingen, and the Interdisciplinary Colloquium at the University of Potsdam. A special thank you goes to Jane Grimshaw for initial motivation; her approach to expletive *do* in English inspired the present approach to expletive *was* in German.

1. Introduction

The main goal of this article is to develop an optimality-theoretic account of several long-standing problems with *wh*-scope marking and partial *wh*-movement in German that fits into a general theory of *wh*-dependencies in languages like English, Korean, Bulgarian, and Hungarian. The resulting analysis will in turn have interesting consequences for the definition of reference sets (see Chomsky 1995); in particular, I will argue that reference sets are defined in terms of LF-identity.

I will proceed as follows. In section 2, I address several striking properties of partial *wh*-movement in German that have so far resisted an explanatorily adequate analysis. Section 3 introduces background assumptions of Optimality Theory (see Prince and Smolensky 1993) and develops an approach to *wh*-movement that relies on the ranking of six general constraints (PROJ-PRIN, DER-ECON, WH-CRIT, FULL-INT, MIN-CHAIN, and BAR-CON). This theory is designed to derive the basic facts related to the “Wh-Movement Parameter” — that is, it differentiates between *wh*-in situ (as in Korean and Japanese), simple *wh*-movement at S-structure (as in English), and multiple *wh*-movement at S-structure (as in Bulgarian). In section 4, then, I show that a specific ranking of the above constraints successfully accounts for all the problems highlighted in section 2. Section 5 shows that a re-ranking of two constraints (PROJ-PRIN and BAR-CON) suffices to capture the basic properties of partial and long-distance *wh*-movement in Hungarian, where *wh*-dependencies share properties both with the Bulgarian and the German type. Finally, section 6 draws a conclusion.

2. Some properties of partial *wh*-movement in German

2.1. Alternation

In partial *wh*-movement constructions in German (cf. [1a]), the target [+wh] position of a *wh*-phrase is filled by a scope marker *was* ‘what’, and the *wh*-phrase is moved to an intermediate [-wh] SpecC position. Perhaps the most conspicuous property of this construction is its alternation with successive-cyclic long-distance movement (cf. [1b]):²

2. Here and henceforth, I assume that the complementizer *daß* may optionally co-occur together with a *wh*-phrase to its left, as in (1a). These optional violations of the Doubly-Filled Filter (cf. Chomsky and Lasnik 1977) regularly show up in colloquial German. They are not generally permitted in Standard German, but I take this to be theoretically insignificant; cf. Müller (1993, chapt. 2) for further discussion.

- (1) a. *Was*₁ *glaubst du* [_{CP} *wann*₁ (*daß*) *sie* *t*₁ *gekommen ist*]?
 [+wh] think you when that she come is
 b. *Wann*₁ *glaubst du* [_{CP} *t*₁ *daß* *sie* *t*₁ *gekommen ist*]?
 when think you that she come is

Such an alternation is *a priori* unexpected under the following two assumptions: First, (1a) and (1b) have identical LFs. And second, exactly those derivations compete with respect to Economy constraints that have identical LFs. If these two assumptions are correct, it seems that we should incorrectly expect one of the two sentences in (1) to be impossible (given the notions of Economy in Chomsky 1993, 1995). Before proceeding, I will try to substantiate these two assumptions.

2.1.1. *LF identity.* According to the first assumption, the LF of (1a) and (1b) looks roughly as in (2).

- (2) [_{CP} *wann*₁ *glaubst du* [_{CP} *sie* *t*₁ *gekommen ist*]]?
 when think you she come is

To derive (2) from (1a), the partially moved *wh*-phrase in (1a) must undergo substitution in the matrix SpecC-position bearing the feature [+wh], thereby overwriting the expletive scope marker. Moreover, I assume (following Lasnik and Saito 1984, 1992; and Chomsky 1986a, among others) that an embedded complementizer *daß* (if present in the first place) is deleted on the way to LF, as an item that does not contribute to semantic interpretation. Turning next to (1b), a derivation of the LF in (2) is straightforward — apart from complementizer deletion, nothing happens.³

The assumption that (1a) and (1b) yield identical LF-representations used to be a common one. Assuming that *was* in (1a) is an expletive, and that the embedded *wann* ‘when’ is a *wh*-operator, LF-movement of *wann* in (1a) (hence, LF identity of [1a] and [1b]) virtually follows under standard approaches to question semantics in a framework with LF. Thus, in Heim’s (1992) and von Stechow’s (1996a) adaptation of Karttunen’s (1977) theory, *wh*-operators are interpreted as existential quantifiers which must be located immediately to the left of the C node bearing the [+wh] feature, to obtain a compositional semantics. This implies that all *wh*-phrases must be moved to the specifier position of a C_[+wh] head at LF.

However, this assumption is called into question by Dayal (1994), who argues that the German scope marker *was* (like its Hindi counterpart *kyaa*) is not

3. For ease of exposition, I ignore here the possibility of deletion of intermediate traces (cf. Lasnik and Saito 1984 and Chomsky and Lasnik 1993, among others). It seems reasonable to postulate that presence or absence of intermediate traces can never create LF non-identity (in this sense) of two otherwise identical derivations.

actually an expletive base-generated in SpecC, but rather a genuine *wh*-object that has undergone movement out of the direct object position. More precisely, she assumes that the scope marker *was* in (1a) is a *wh*-operator that quantifies over propositions, that is, it denotes a relation between sets of propositions. The embedded CP in (1a), on this view, acts as the restriction of the *wh*-operator *was*. Thus, in this approach, (1a) and (1b) do not have identical LF-representations (*wann* ‘when’ in [1a] is interpreted in the embedded SpecC-position). However, it is pointed out in Höhle (1990) and Müller and Sternefeld (1996: 507–509) that such an “indirect *wh*-dependency” approach to partial *wh*-movement constructions in German faces a number of problems. Basically, these counter-arguments against Dayal’s (1994) approach are syntactic in nature; they all show that the scope marker *was* in German does not behave like an object base-generated in argument position, but rather like an expletive that is overwritten by LF-raising of the *wh*-phrase it is co-indexed with. I will summarize three of the arguments here (cf. Beck and Berman 1996 for semantic problems with Dayal’s analysis).

2.1.1.1 *Multiple questions* A scope marker *was* may in principle show up in multiple questions; cf. (3a). However, as observed by Höhle (1990), it must occupy the SpecC position; as shown in (3b), it may never occur *in situ*. This is straightforwardly explained if the scope marker *was* is an expletive base-generated in SpecC; it must remain a mystery if the scope marker originates as an ordinary object NP within the VP — all the more so since a true object *was* can show up *in situ* without problems; cf. (3c).

- (3) a. *Was meint wer* [_{CP} *wann₁ sie t₁ gekommen ist*]?
 [+wh] thinks who when she come is
 b. **Wer meint was* [_{CP} *wann₁ sie t₁ gekommen ist*]?
 who thinks [+wh] when she come is
 c. *Wer meint was?*
 who thinks what

2.1.1.2 *Echo questions* A similar counter-argument (again going back to Höhle 1990) is based on echo questions. It is shown in (4a) that a genuine object *was* can occur *in situ* in echo questions. In contrast to that, a scope marker *was* in partial *wh*-movement constructions can never occur *in situ* in an echo question (cf. [4b]) — if an echo question is to be formed via partial *wh*-movement and scope marking, the scope marker must occupy SpecC; cf. (4c):

- (4) a. *Fritz hat WAS gesagt?*
 Fritz has WHAT said
 b. **Fritz hat WAS gesagt* [_{CP} *wann₁ sie t₁ gekommen ist*]?
 Fritz has [+wh] said when she come is
 c. *WAS hat Fritz gesagt* [_{CP} *wann₁ sie t₁ gekommen ist*]?
 [+wh] has Fritz said when she come is

Again, this state of affairs suggests that the German scope marker *was* is not a *wh*-operator base-generated in an object position.

2.1.1.3 *Intervening daß* For many speakers, sentences like (5) are well formed in German:

- (5) *Was meinst du* [_{CP1} *daß sie gesagt hat* [_{CP2} *wann₁ sie t₁ kommen*
 [+*wh*] think you that she said has when she come
würde]]?
 would

Here, the scope marker *was* and the partially moved *wh*-phrase *wann* are separated by an intervening CP with an — apparently — empty SpecC-position. I will come back to this construction in section 2.7; for the time being, it may suffice to note that (5) is problematic for the indirect *wh*-dependency analysis, according to which the “appositive” clause in partial *wh*-movement constructions must be a *wh*-clause. The reason is that a direct interpretation without LF-movement of *wann* does not appear to be possible because of semantic type mismatch: The embedded CP₁ is a declarative clause, and not a question. Consequently, it denotes a *proposition*, but not a *set of propositions*, as would be required if *was* is to quantify over propositions (recall that the scope marker *was* is analyzed as a relation between sets of propositions in Dayal’s indirect *wh*-dependency analysis).

For these reasons (and others discussed in Müller and Sternefeld 1996), I conclude that Dayal’s (1994) approach to partial *wh*-movement in German is not on the right track, and that partial and long-distance *wh*-movement (as in [1a] and [1b]) do indeed create identical LF-representations.

2.1.2. *Reference sets.* Recall that LF-identity of derivations with partial and long-distance *wh*-movement is problematic under Economy considerations only if yet another assumption is made, namely, that the notion of reference set is defined via LF-identity of derivations, as in (6).

- (6) Reference set:
 Two derivations D₁ and D₂ are in the same reference set iff they yield the same LF-output.

Such a notion of reference set is argued for by Epstein (1992), Kitahara (1993), Fox (1994), and Reinhart (1994), among others.⁴ In contrast, Chomsky (1995) adopts a notion of reference set that is not based on LF-identity; rather, it

4. Fox and Reinhart actually assume that the relevant factor is identity of meaning, rather than LF-identity. There are certain differences between the two concepts; nevertheless, with respect to the case at hand (that is, partial *wh*-movement), the prediction would be the same — (1a) and (1b) are in the same reference set. See Sternefeld (1997) for a detailed comparison.

basically states that those derivations compete that have the same lexical material (the same numeration). The notion of reference set in Chomsky (1995: 227) can be formulated as follows:

(7) Reference set (Chomsky 1995):

Two derivations D_1 and D_2 are in the same reference set iff they start with the same numeration and converge at LF and PF.

Clearly, the derivations D_1 and D_2 generating (1a) (partial *wh*-movement) and (1b) (long-distance *wh*-movement) respectively, would then not end up in the same reference set — irrespective of the issue of the complementizer *daß*, D_1 would differ from D_2 in that a scope marker *was* is present in D_1 , but not in D_2 . Hence, there is no competition of (1a) and (1b) if the definition of reference set in Chomsky (1995) is adopted. However, notwithstanding the substantial evidence in support of the notion of reference set in (6) provided in the literature just cited, and notwithstanding several problems that emerge with the notion of reference set in (7) (see Sternefeld 1997), I will later argue that there is indeed good evidence that partial *wh*-movement and long-distance *wh*-movement do compete, that is, that (6) is to be preferred over (7). But this means that, for the time being, free alternation as in (1a) and (1b) potentially poses a problem — we have to ensure that neither sentence (or, more precisely, neither of the two derivations generating the two S-structure representations) blocks the other one. That said, let me turn to the next peculiar property of partial *wh*-movement.

2.2. Parametrization

Closely related languages like English and Dutch do not exhibit partial *wh*-movement constructions as in (1a). Here, only successive-cyclic long-distance movement of the *wh*-phrase is possible. This is shown for English in (8), and for Dutch in (9):

- (8) a. *Why₁ do you think* [_{CP} *t₁' that she came t₁']?
 b. **What₁ do you think* [_{CP} *why₁ she came t₁']?**
- (9) a. *Wie₁ denk je* [_{CP} *t₁' dat t₁ het gedaan heeft']?
 b. **Wat₁ denk je* [_{CP} *wie₁ t₁ het gedaan heeft']?**

Although this fact is well known, it does not seem to have received a convincing explanation so far. Standardly (see, for example, Koster 1987), it is suggested that English and Dutch lack partial *wh*-movement constructions because they lack a scope marker. However, a cross-linguistic survey reveals that languages typically use the equivalent of *what* if they are in need of a scope marker for *wh*-phrases. This holds for *was* 'what' in German, for *kyaa* 'what' in Hindi (cf. Mahajan 1990: 129; Dayal 1994: 138), for *mit* 'what' in Hungarian (cf., for

example, Marácz 1990: 325), for *so* ‘what’ in Romani (cf. McDaniel 1989: 568), for *sheno* ‘what’ and its contracted version *sh-* in Iraqi Arabic (cf. Wahba 1992: 264), etc. Thus, it seems that there is no *a priori* ‘lexical’ reason why *what* in English and *wat* in Dutch should not be able to function as scope markers. I conclude that a syntactic explanation for the illformedness of (8b) and (9b) should be sought.⁵

2.3. *Anti-locality*

A surprising property of *wh*-scope marking in German is that it exhibits an “anti-locality” effect, quite in contrast to what is the case with other dependencies: a scope marker and the *wh*-phrase it is co-indexed with may never be clause-mates (cf. Stechow and Sternefeld 1988). This is shown in (10a), where *warum* ‘why’ shows up *in situ*, and a scope marker *was* is located in SpecC. Note that the very combination of *warum in situ* and *was* in SpecC is fully grammatical if the latter acts as an object, and not as a scope marker; cf. (10b).⁶

- (10) a. **Was*₁ *ist* [_{IP} *sie warum*₁ *gekommen*]?
 [+wh] is she why come
 b. *Was*₁ *hat* [_{IP} *sie warum*₂ *t*₁ *getan*]?
 what has she why done

2.4. *Partiality*

Another property of scope marking constructions in German that calls for an explanation concerns the obligatoriness of partial movement. Thus, it is not quite obvious why the *wh*-phrase has to undergo partial movement even though its scope is indicated by a scope marker; that is, why (11b) is not permitted in addition to (or instead of) (11a) (= [1a]).⁷

-
5. As noted by a reviewer, this view is reinforced by the fact that Dutch *wat* can indeed function as an expletive in exclamative sentences of the type in (i):
 (i) *Wat heb ik lekker geslapen!*
 what have I well slept
 See Corver (1990) for discussion.
6. The wellformedness of (10b) shows that there is no general prohibition against *wh*-adjuncts *in situ* at S-structure in German, in contrast to what is the case in English and French (cf. Aoun, Hornstein and Sportiche 1981; Lasnik and Saito 1984, among others). See Haider (1986), Bayer (1990), and Müller (1993) for accounts of this fact.
7. In order to account for both anti-locality and partiality, Stechow and Sternefeld (1988: 355) postulate that “each *wh*-element *in situ* is scope-bound by a *wh*-operator” in an A-bar position. What is relevant here is that scope markers are not *wh*-operators, hence (in contrast to a genuine

- (11) a. *Was₁ glaubst du* [_{CP} *wann₁ (daß) sie t₁ gekommen ist*]?
 [+wh] think you when that she come is
- b. **Was₁ glaubst du* [_{CP} *daß sie wann₁ gekommen ist*]?
 [+wh] think you that she when come is

2.5. Scope marking

Conversely, a scope marker must show up even though the *wh*-phrase has undergone movement. This is shown for matrix *wh*-clauses in (12a) (= [1a]) versus (12b). Also note that partial *wh*-movement is not possible if the eventual scope position of the partially moved *wh*-phrase is occupied by another *wh*-phrase, as in (12c) (cf. Epstein 1992; Müller and Sternefeld 1996).

- (12) a. *Was₁ glaubst du* [_{CP} *wann₁ (daß) sie t₁ gekommen ist*]?
 [+wh] think you when that she come is
- b. *_{[CP —} *glaubst du* [_{CP} *wann₁ (daß) sie t₁ gekommen ist*]]?
 think you when that she come is
- c. *_{[CP} *Wer₂ glaubt t₂* [_{CP} *wann₁ (daß) sie t₁ gekommen ist*]]?
 who thinks when that she come is

Partial *wh*-movement in German is not confined to contexts where the ultimate [+wh] target position is in a matrix clause; as noted by Höhle (1990), it can also be the SpecC-position of an embedded clause. But again the same picture arises: partial *wh*-movement is possible if the [+wh] target position is occupied by a scope marker (cf. [13a]), but not if it remains empty (cf. [13b]), or if it is occupied by another *wh*-phrase at S-structure (cf. [13c]):

- (13) a. *Ich weiß nicht* [_{CP} *was₁ sie gemeint hat* [_{CP} *wann₁ sie t₁*
 I know not [+wh] she said has when she
kommen würde]]
 come would
- b. **Ich weiß nicht* [_{CP} *daß sie gemeint hat* [_{CP} *wann₁ sie t₁ kommen*
 I know not that she said has when she come
würde]]
 would

wh-phrase in SpecC) they fail to license *wh in situ* phrases. Since *wh in situ* phrases systematically show up in languages like Korean and Japanese without being scope-bound by a *wh*-operator, the condition must be language-specific. Although this principle does indeed correctly predict the anti-locality and partiality properties of partial *wh*-movement in German, it seems to me that ideally, it should be derivable from deeper assumptions, and there should be a principled reason why its effects are absent in *wh in situ* languages.

- c. **Ich weiß nicht* [_{CP} *wer*₂ (*daß*) *t*₂ *gemeint hat* [_{CP} *wann*₁ *sie* *t*₁
 I know not who that said has when she
kommen würde]]
 come would

2.6. Late was

Another problem with partial *wh*-movement and scope marking in German shows up when we consider somewhat longer *wh*-dependencies. Thus, in (14), the [+*wh*] SpecC-position is in the matrix clause (CP₂), and the *wh*-phrase originates in the most deeply embedded clause (CP₄). As one might expect, there is free alternation between long-distance movement of the *wh*-phrase to the SpecC-position of CP₂ (cf. [14a]), multiple scope marking and partial *wh*-movement to the SpecC-position of CP₄ (cf. [14b]), and an “intermediate” version that exhibits both scope marking in the matrix SpecC-position and successive-cyclic *wh*-movement to the SpecC-position of CP₃ (cf. [14c]):

- (14) a. [_{CP2} *Wann*₁ *meinst du* [_{CP3} *t*'₁ *daß sie gesagt hat* [_{CP4} *t*'₁ *daß*
 when think you that she said has that
sie *t*₁ *kommen würde*]]]?
 she come would
- b. [_{CP2} *Was*₁ *meinst du* [_{CP3} *was*₁ *sie gesagt hat* [_{CP4} *wann*₁ *sie* *t*₁
 [+*wh*] think you [+*wh*] she said has when she
kommen würde]]]?
 come would
- c. [_{CP2} *Was*₁ *meinst du* [_{CP3} *wann*₁ *sie gesagt hat* [_{CP4} *t*'₁ *daß sie* *t*₁
 [+*wh*] think you when she said has that she
kommen würde]]]?
 come would

The data in (14) suggest that syntactic theory has to incorporate a concept of generalized *wh*-dependencies (or generalized *wh*-CHAINS, in Chomsky's 1981 sense) that may include (a) one *wh*-phrase, (b) any number of traces, and (c) any number of scope markers. Given this state of affairs, the question arises why the order of scope markers, *wh*-phrase, and its traces in generalized *wh*-dependencies is strict. As observed by McDaniel (1989), the scope marker may never occur to the right of the *wh*-phrase, that is, there is a prohibition against “late” *was*. This is shown in (15):

- (15) * [_{CP2} *Was*₁ *meinst du* [_{CP3} *wann*₁ *sie gesagt hat* [_{CP4} *was*₁ *sie* *t*₁
 [+*wh*] think you when she said has [+*wh*] she
kommen würde]]]?
 come would

It seems *a priori* unclear how to derive this ban on late *was*, given that scope markers are as legitimate as members of generalized *wh*-dependencies in German as traces.⁸

2.7. Long-distance scope marking

In addition to the sentences in (14), another derivation with the same LF is possible, namely, one that generates the S-structure in (16) (= [5]):

- (16) [_{CP2} *Was*₁ *meinst du* [_{CP3} *daß sie gesagt hat* [_{CP4} *wann*₁ *sie* *t*₁
 [+wh] think you that she said has when she
kommen würde]]]?
 come would

Example (16) differs from (14b) only in that the scope marker *was* in the SpecC-position of CP₂ is replaced by a complementizer *daß* in the C position of this clause. This gives rise to what looks like a case of long-distance scope marking — scope marker and partially moved *wh*-phrase are separated by an intervening CP (namely, CP₃) with an apparently empty SpecC-position. However, long-distance *wh*-dependencies of this kind are otherwise unattested in German. Hence, the wellformedness of (16) poses a problem.⁹ The same kind of problem arises with sentences like (17), where again an apparently empty SpecC-position intervenes between a scope marker *was* and the partially moved *wh*-phrase.

- (17) [_{CP2} *Was*₁ *meinst du* [_{CP3} *hat sie gesagt* [_{CP4} *wann*₁ *sie* *t*₁
 [+wh] think you has she said when she
kommen würde]]]?
 come would

-
8. This problem has not gone unnoticed in the literature, and attempts have been made to solve it. Thus, McDaniel (1989: 580) postulates that wellformed generalized *wh*-CHAINS must have the property that “for any scope marker α_i , $1 \leq i < n$, $\langle \alpha_{i+1} \dots \alpha_{n-1} \rangle$ contains a true *wh*-phrase”. This is observationally adequate, but it clearly is a restatement of the facts, and should be derived from deeper assumptions. Another approach is suggested by Brandner (1994: 204), who assumes that the embedded scope marker *was* in a sentence like (15) creates a *wh*-island that blocks movement of the *wh*-phrase *wann*. This solution is called into question by the standard assumption (for which I will present additional evidence later) that scope markers and partially moved *wh*-phrases are co-indexed — if this is so, the “offending” late scope marker in (15) is co-indexed with the *wh*-phrase, and actually part of the generalized *wh*-CHAIN. Hence, it should not erect a *wh*-island, but should rather be able to function as a (CHAIN-internal) antecedent-governor (cf. Chomsky 1986a).
9. Note that there is no general agreement as to whether sentences like (16) are completely grammatical; van Riemsdijk (1982: 12), McDaniel (1989: 576), and Haider (1993: 98), for example, assume that they are not. I will here follow the judgement in Höhle (1990: 2) and Müller and Sternefeld (1996: 509), according to which (16) is wellformed to the same degree as, for example, (14a) (cf. also Stechow and Sternefeld 1988: 358 for remarks on this issue).

The only difference between (17) and (16) is that instead of the complementizer *daß*, the C node of CP₃ is filled by a V2 head *hat* 'has'.

2.8. Weak islands

The final problem raised by partial *wh*-movement which is to be discussed here concerns weak islands. The crucial fact is that partial *wh*-movement of an argument is much more sensitive to weak islands than successive-cyclic long-distance movement of an argument (or *wh in situ*, that is, LF-movement of an argument). I will show this for negative islands, CNPC (complex noun phrase constraint) islands, and subject islands.

2.8.1. *Negative islands*. It has often been noted (see Höhle 1990; Rizzi 1991; and Beck 1996, chapt. 2) that partial *wh*-movement is completely impossible if subsequent LF-raising of the partially moved *wh*-phrase crosses a negative island; cf. (18a). In contrast, successive-cyclic long-distance movement of an argument across a negative island is at worst slightly marginal in German (cf. [18b]), and *wh in situ* seems to be even more acceptable (cf. [18c]).¹⁰

- (18) a. **Was*₁ *glaubst du nicht* [_{CP} *wen*₁ (*daß*) *Hans t*₁ *getroffen* [+wh] think you not whom(ACC) that Hans met *hat*]?
has
- b. ??*Wen*₁ *glaubst du nicht* [_{CP} *t*₁ *daß* *Hans t*₁ *getroffen hat*]?
whom(ACC) think you not that Hans met has
- c. *Wer* *glaubt nicht* [_{CP} *daß* *Hans wen*₁ *getroffen hat*]?
who(NOM) thinks not that Hans whom(ACC) met has

2.8.2. *CNPC islands*. As another instance of weak islands, consider complex noun phrases.¹¹ It is noted in van Riemsdijk (1982) that partial *wh*-movement is strictly prohibited in CNPC contexts; cf. (19a). Stechow and Sternefeld (1988: 358) observe that there is a clear contrast to analogous cases of long-distance *wh*-movement out of a CNPC context — although the resulting sentence (19b) is not fully acceptable (it exhibits a standard Subjacency effect), it is clearly

10. A similar asymmetry occurs in Bahasa Indonesia; see Saddy (1991: 188–194).

11. Complex noun phrases are classified as strong islands in Cinque (1990). However, the standard view is that strong islands are those that indiscriminately block adjunct and argument extraction, whereas weak islands strictly block adjunct extraction, but typically yield more variable results with argument extraction. Clearly, then, complex noun phrases in German (at least those that contain complement CPs) belong to the latter group.

much better than (19a), and this fact calls for an explanation. As before, an *in situ* occurrence of a *wh*-argument within a complex NP is unproblematic; cf. (19c).¹²

- (19) a. **Was*₁ *hast du* [_{NP} *ein Gerücht* t₃] *gehört* [_{CP3} *wen*₁
 [+wh] have you a rumour heard whom(ACC)
Ede t₁ *mag*]?
 Ede likes
- b. ??*Wen*₁ *hast du* [_{NP} *ein Gerücht* t₃] *gehört* [_{CP3} t'₁ *daß*
 whom(ACC) have you a rumour heard that
Ede t₁ *mag*]?
 Ede likes
- c. *Wer*₂ t₂ *hat* [_{NP} *ein Gerücht* t₃] *gehört* [_{CP3} *daß* *Ede*
 who(NOM) has a rumour heard that *Ede*
*wen*₁ *mag*]?
 whom(ACC) likes

2.8.3. *Subject islands*. Finally, let us consider the case of subject islands in German. Again, the same picture arises (cf. McDaniel 1989: 571): whereas partial *wh*-movement within subject islands (with the scope marker outside of the island) is strictly ungrammatical even in the case of arguments, overt extraction of an argument from a subject island is only mildly deviant; compare (20a) with (20b). Given what has been said so far, it does not come as a surprise that *wh in situ* in this context is close to perfect (cf. [20c]).

- (20) a. **Was*₁ *ist es schade* [_{CP} [_{PP} *mit wem*]₁ *Hans* t₁ *gesprachen hat*]?
 [+wh] is it too bad with whom Hans spoken has
- b. ??[_{PP} *Mit wem*]₁ *ist es schade* [_{CP} t'₁ *daß* *Hans* t₁ *gesprachen hat*]?
 with whom is it too bad that Hans spoken has
- c. *Für wen*₂ *ist es* t₂ *schade* [_{CP} *daß* *Hans* [_{PP} *mit wem*]₁
 for whom is it too bad that Hans with whom
gesprachen hat]?
 spoken has

To sum up, partial *wh*-movement turns out to be much more sensitive to weak islands in German than both long-distance *wh*-movement and *wh in situ*. The asymmetry is particularly conspicuous in the case of negative islands, but it is also pervasive in the case of CNPC islands and subject islands. It seems to me

12. The contrast between (19b) and (19c) is accounted for if we follow Huang (1982) and Lasnik and Saito (1984, 1992), among others, in assuming that Subadjacency does not constrain LF-movement.

that a comprehensive account of these phenomena that does not resort to construction-specific assumptions is still outstanding.¹³

2.9. Conclusion

In the preceding subsections, I have discussed eight properties of partial *wh*-movement that call for an explanation. In what follows, I will argue that the key to a solution of a number of these problems is actually based on the very first problem — viz., that partial *wh*-movement and successive-cyclic long-distance *wh*-movement alternate, that is, compete with each other. The basic idea is that whereas “normally,” the two constructions may co-exist, one may block the other under special circumstances. To execute this idea, I will first develop an optimality-theoretic approach to *wh*-movement in section 3. I will return to partial *wh*-movement in German in section 4.

3. Optimality Theory and *wh*-movement

3.1. Premises

The basic tenets of Optimality Theory (see Prince and Smolensky 1993; among many others) are the following:

-
13. An approach to the asymmetries between partial and long-distance *wh*-movement that show up with weak islands in German has been developed by Rizzi (1991: 5ff.) on the basis of the negative island facts. Rizzi observes that partial *wh*-movement patterns with long-distance *wh*-movement of an adjunct in this respect. He accounts for this by assuming that scope markers like *was* in German do not carry a referential index — a property that (simplifying somewhat) they share with adjuncts, but not with arguments. Hence, a chain link of a generalized *wh*-CHAIN that contains a scope marker will depend on antecedent-government, and cannot resort to binding; and, in the framework developed in Rizzi (1990), antecedent-government is blocked by weak islands, in contrast to binding. Rizzi’s (1991) approach works well for the cases at hand; however, there is no denying that it involves a construction-specific assumption, namely, that a chain link that contains a scope marker and a *wh*-argument behaves just like an adjunct chain link, and not like an argument chain link. Another approach is developed by Beck (1996, chapt. 2). She accounts for the negative island facts by assuming that scope-bearing elements (like negation) invariably erect barriers at LF (but not at S-structure). Thus, overt movement of an argument across a negation is fine, but LF-movement is not. This analysis might be empirically problematic for two reasons. First, it does not generalize to CNPC islands and subject islands (in these cases, there is no obvious candidate for a scope-bearing element that creates an LF-barrier). And second, it automatically excludes (18c), where a *wh*-argument *in situ* shows up that must undergo LF-movement to its scope position across an intervening negation.

- (21) a. Constraints are universal.
 b. Constraints can be violated.
 c. Constraints are ranked.
 d. An optimal candidate in a reference set is grammatical, all non-optimal candidates are ungrammatical.

The candidates of a given reference set are generated by a part of the grammar (GEN) that contains only inviolable constraints. For example, I will assume (contra Chomsky 1995) that there are principles of X-bar theory, and that these principles are inviolable and belong to GEN. The candidates built up by GEN are then subjected to a process of harmony evaluation that is based on (21), with the result that only “optimal” candidates are grammatical. The central notion of optimality can be understood as follows (see, for example, Prince and Smolensky 1993 and Grimshaw 1995):

- (22) Optimality:

A derivation D_i is optimal iff, for every derivation D_j in the same reference set, D_i satisfies the highest-ranking constraint on which D_i and D_j conflict better than D_j .

There are two ways in which one derivation D_i can satisfy a constraint C better than another derivation D_j . First, and obviously, this is the case if D_i fulfills C and D_j does not. Second, however, D_i may violate C, but D_j violates C more often; in such a situation, D_i 's satisfaction of C is still better than that of D_j . With this in mind, let me now first turn to the constraints that I will adopt in what follows, and then focus on their ranking.¹⁴

3.1.1. *The Projection Principle.* The Projection Principle is formulated as follows:

- (23) PROJ-PRIN:

Adjunction to CP, C^0 , and SpecC is prohibited at S-structure.

Clearly, this formulation of PROJ-PRIN differs substantially from the one that can be found in Chomsky (1981); it goes back to Rizzi and Roberts (1989, sect. 3) and Grimshaw (1995). I do not want to propose that (23) covers all the effects that have been accounted for by invoking the Projection Principle of Chomsky (1981). Rather, what is meant here is that the statement in (23) can ideally be made to follow from something like the classic Projection Principle; indeed, this has been proposed in the case of adjunction to CP in Chomsky

14. Note that all of the following constraints have — in some version — been proposed as inviolable in the literature. Moreover, several of these constraints already show up as violable and ranked in Grimshaw's (1995) optimality-theoretic approach to projection and inversion, albeit in slightly different formulation, and with slightly different consequences.

(1986a), and in the case of movement to C^0 in Rizzi and Roberts (1989).¹⁵ Crucially (and in contrast to the formulation of PROJ-PRIN in Grimshaw 1995), (23) also rules out adjunction to SpecC; this will become relevant.

3.1.2. *Derivational Economy*. Derivational Economy is defined as follows:¹⁶

- (24) DER-ECON:
S-structure movement is not allowed.

This derivational constraint can be viewed as an amalgamation of Chomsky's (1993, 1995) transderivational constraints Procrastinate (which prefers LF-movement over S-structure movement) and Fewest Steps (which prefers no movement over movement). According to (24), any S-structure movement step creates a DER-ECON violation. If we do not want to resort to the concept of Form Chain (cf. Chomsky 1993), this means that overt successive-cyclic *wh*-movement involving, for example, three SpecC-positions will invariably incur three violations of DER-ECON. Note, however, that LF-movement can never violate (24), be it successive-cyclic or not.

3.1.3. *The wh-criterion*. I will adopt the following version of the *wh*-criterion (see, for example, May 1985: 17 and Lasnik and Saito 1992: 7–8).¹⁷

- (25) WH-CRIT:
a. *Wh*-operators are in SpecC.
b. $C_{[+wh]}$ requires lexicalization of C or SpecC.

Clearly, the “simplest” way to satisfy WH-CRIT is to move a *wh*-phrase to a SpecC-position that agrees with a $C_{[+wh]}$ head; however, WH-CRIT can also be satisfied if the *wh*-phrase moves to some other SpecC-position (thereby fulfilling the (a)-part of [25]), and the (b)-part of (25) is fulfilled by lexicalizing $C_{[+wh]}$ or by inserting a *wh*-expletive (that is, a scope marker) into the specifier position

15. The important point in Rizzi and Roberts (1989, sect. 3) is that I-to-C movement is permitted in root CPs, but not in embedded CPs. This follows under the definition of PROJ-PRIN in (23) if we assume that matrix C is radically empty (so that I-to-C movement can be substitution), whereas an embedded C, being selected, is never radically empty (so that I-to-C movement would have to resort to adjunction, which is blocked by 23). See Rizzi (1991) and Müller and Sternefeld (1993) for further discussion of this issue.

16. Under the assumption that all movement leaves traces, this constraint would be identical to Grimshaw's (1995) STAY.

17. The first part of (25) is a strengthened version of Grimshaw's (1995) OP-SPEC constraint. Note also that WH-CRIT differs from May's and Lasnik and Saito's formulation in one important respect — it is not required that *wh*-operators be in a SpecC-position with a feature [+wh].

of the $C_{[+wh]}$ head. For the time being, I will assume that WH-CRIT holds at S-structure.¹⁸

3.1.4. *Full Interpretation.* The principle of Full Interpretation goes back to Chomsky (1986b). Here, I will adopt Grimshaw's (1995) reformulation:

(26) FULL-INT:

A lexical item has an interpretation.

A consequence of FULL-INT is that it prohibits the insertion of expletives like *do* in English. Of course, another operation that will inevitably lead to a violation of FULL-INT is scope marker insertion; this violation can be read off S-structure representations.

3.1.5. *Minimize chain links.* This condition is taken from Chomsky (1993, 1995); it incorporates the basic insight behind Rizzi's (1990) Relativized Minimality.¹⁹

(27) MIN-CHAIN:

Chain formation is minimal at each step of the derivation.

Note that this constraint, in contrast to PROJ-PRIN, WH-CRIT, and FULL-INT, is not a simple S-structure filter; rather, it is formulated as a derivational constraint (cf. Chomsky 1993, 1995), just like DER-ECON. This implies that a violation of (27) occurs if a step in the derivation is not compatible with it — even if, for instance, the S-structure representation itself has only minimal chain links.

3.1.6. *The barriers condition.* Finally, I assume that many weak island effects follow from the prohibition against crossing a barrier, as in (28) (cf. Chomsky 1986a; Cinque 1990; and Chomsky and Lasnik (1993); among many others):

(28) BAR-CON:

Chain formation must not cross a barrier.

BAR-CON is a derivational constraint, just like MIN-CHAIN. As regards the central notion of barrier, I will assume (following Chomsky and Cinque) that an

18. As far as LF is concerned, I follow Huang (1982) and Lasnik and Saito (1984, 1992), among others, in postulating that all *wh*-phrases must be in $\text{Spec}C_{[+wh]}$ at this level, without qualification. As noted above, under Heim's (1992) and Stechow's (1996a) approach to *wh*-semantics adopted here, this requirement is actually a semantic one, and does not have to be stipulated by a syntactic constraint. Thus, WH-CRIT, which is violable, triggers overt *wh*-movement, and semantic requirements, which are inviolable, trigger LF *wh*-movement of all *wh*-phrases that are still *in situ* at S-structure.

19. The notion of "chain formation" used here is distinct from Chomsky's (1993) concept of Form Chain; see section 4.4 below.

XP γ is a barrier between α and β iff (i)–(iii) hold: (i) γ includes β ; (ii) γ excludes α ; and (iii) γ is not directly selected. Direct selection of an XP is to be understood as either semantic selection under sisterhood (that is, direct θ -marking), or morphological selection of the head of XP (for example, I morphologically selects V, and C morphologically selects I, as evidenced by the presence of inflectional features on complementizers in many languages). Under these assumptions, a NegP (which I assume to intervene between V and I, see Pollock 1989) is automatically a barrier because it is neither morphologically selected, nor semantically selected by the embedding head. Similarly, a CP embedded by N is invariably a barrier due to a lack of direct selection if we assume (following Stowell 1981 and others) that not only relative clause CPs, but even CP “complements” fail to be θ -marked NP-internally. Finally, a subject XP is a barrier because, not being a sister of a lexical head, it can never be directly selected. Thus, negative islands, CNPC islands and subject islands can all be reduced to BAR-CON.

Summarizing this section, I have introduced six constraints, three of them representational (PROJ-PRIN, WH-CRIT and FULL-INT) and three derivational (DER-ECON, MIN-CHAIN and BAR-CON). Clearly, these constraints partly impose conflicting requirements on candidates (that is, derivations). For instance, WH-CRIT forces overt movement, whereas DER-ECON bans such movement. This tension is a fundamental characteristic of Optimality Theory; the resolution of such conflicts is accomplished by ranking the constraints — that is, the demands imposed by one constraint may be overridden by the demands of another constraint that is ranked higher. It is the ranking of the constraints introduced above that I will focus on now.²⁰ I will begin by considering data from English, Korean, and Bulgarian related to the “Wh-movement Parameter,” and I will show that cross-linguistic variation in this area can straightforwardly be accounted for by manipulating the ranking of the three constraints PROJ-PRIN, DER-ECON and WH-CRIT.²¹

20. A remark is in order concerning two differences to standard optimality-theoretic assumptions. First, as just noted, I postulate that there are derivational constraints, in addition to representational constraints (S-structure filters). Second, and related to that in an obvious way, I assume throughout (basically following Chomsky 1993, 1995) that the competing candidates (that is, the members of the reference set) are *derivations*. This view deviates from the standard position taken in Optimality Theory (see Prince and Smolensky 1993), according to which the members of a reference set are *output representations*, that is, (in syntax) S-structure representations (cf. Grimshaw 1995). The position taken here is closer in spirit to a proposal by McCarthy and Prince (1995) according to which the competing candidates are not simply output representations, but rather (Input,Output) pairs. Under this approach, then, there does not seem to be a principled reason not to allow *n*-tuples as candidates, or even complete derivations; cf. Smolensky (1995) and Legendre *et al.* (1996, sect. 9).

21. Also see Ackema & Neeleman (1995) and Legendre *et al.* (1996) for optimality-theoretic approaches to the Wh-Movement Parameter. These approaches share certain assumptions with the one developed here, and differ in others. A careful comparison is beyond the scope of this article, though.

3.2. *Wh-movement at S-structure and at LF*

It is well known that languages like English, Korean (Japanese, Chinese), and Bulgarian (Romanian) differ with respect to the options of *wh*-movement. In what follows, I address these three language types in turn.

3.2.1. *English*. If one considers only simple questions like (29), it seems as though it might suffice to assume that English has obligatory *wh*-movement at S-structure:

- (29) a. **(I wonder)* [_{CP} — [_{IP} *he gave what*₁ *to John*]]
 b. *(I wonder)* [_{CP} [_{SpecC} *what*₁] [_{IP} *he gave t*₁ *to John*]]

However, multiple questions as in (30) show that the issue is somewhat more involved.

- (30) a. **(I wonder)* [_{CP} — [_{IP} *he gave what*₁ *to whom*₂]]
 b. *(I wonder)* [_{CP} [_{SpecC} *what*₁] [_{IP} *he gave t*₁ *to whom*₂]]
 c. **(I wonder)* [_{CP} [_{SpecC} [_{SpecC} *what*₁] *to whom*₂] *he gave t*₁ *t*₂]]

Thus, not only is overt *wh*-movement of the second *wh*-phrase *to whom* in (30) not required; it is actually prohibited (see [30b] versus [30c]). The correct generalization seems to be that, for each C_[+wh] node, there must be exactly one *wh*-phrase that undergoes S-structure movement to its specifier in English; if other *wh*-phrases are present, they must remain *in situ* at S-structure, raising to SpecC_[+wh] at LF.

3.2.2. *Korean*. The situation in Korean (Japanese, Chinese) can be characterized as follows: no *wh*-phrase may undergo movement to SpecC at S-structure. That *wh*-phrases do not have to move to SpecC in Korean is shown by (31), where the *wh*-object *muös-ül* ‘what’ and the *wh*-adjunct *wae* ‘why’ both remain in their respective base positions.

- (31) [_{CP} — [_{IP} *Nö-nün muös-ül*₁ *wae sa-ss-*]-*ni*?]
 you(TOP) what(ACC) why buy-PAST-O

To show that *wh*-phrases also cannot undergo overt *wh*-movement in Korean is a bit more complicated, given the wellformedness of strings like (31).

- (32) *Muös-ül*₁ [_{IP} *nö-nün t*₁ *wae sa-ss-*]-*ni*?
 what(ACC) you(TOP) why buy-PAST-O

Since the *wh*-phrase occupies a CP-initial position at S-structure, one might assume that it has optionally been moved to SpecC. However, it is argued in some detail in Saito (1985) (on the basis of the analogous construction in Japanese) and Müller and Sternefeld (1993) that (32) involves scrambling in

front of the subject (presumably to IP), rather than optional *wh*-movement to SpecC. One piece of evidence is based on the restrictions on long-distance movement. It is well known that adjuncts cannot undergo long-distance scrambling in Korean (and Japanese), in contrast to arguments. Interestingly, *wh*-fronting behaves exactly like scrambling in this respect — as shown in (33a), *wh*-adjuncts cannot be moved out of a CP. This corresponds exactly to our expectations if overt *wh*-fronting in Korean (Japanese) must be scrambling, but it remains a mystery if such *wh*-fronting is analyzed as optional *wh*-movement: not only is it the case that overt *wh*-movement to SpecC in languages like English and German can extract adjuncts from an embedded CP in successive-cyclic fashion without problems; what is more, the wellformedness of an example like (33b) in Korean (which differs from [33a] only in that the *wh*-adjunct is *in situ*) shows that *wh*-adjuncts in *wh-in situ* languages like Korean and Japanese can undergo successive-cyclic *wh*-movement from an embedded clause, but only at LF (see Lasnik and Saito 1984, 1992).

- (33) a. *_{CP} *Wae*₁ *nŏ-nŏn* [_{CP} *t*₁ *Ch'ŏlsu-ka* *t*₁ *o-ass-ta-ko*]
 why you-TOP Ch'ŏlsu-nom come-PAST-DECL-C
saengkakha-ni?
 think-Q
 'Why₁ do you think that Ch'ŏlsu came *t*₁?'
 b. [_{CP} *Nŏ-nŏn* [_{CP} *Ch'ŏlsu-ka* *wae*₁ *o-ass-ta-ko*]
 you-TOP Ch'ŏlsu-nom why come-PAST-DECL-C
saengkakha-ni?
 think-Q
 'Why₁ do you think that Ch'ŏlsu came *t*₁?'

For reasons like this, I conclude that Korean and Japanese are “true” *wh in situ* languages: overt *wh*-movement to SpecC is not only unnecessary, it is also prohibited.

3.2.3. *Bulgarian*. Turning finally to Bulgarian, it has been argued by Rudin (1988) that all *wh*-phrases must be moved to SpecC at S-structure. Thus, consider (34), where both *wh*-phrases (*koj* ‘who’ and *kogo* ‘whom’) have undergone overt movement to SpecC. Here, the *wh*-subject undergoes substitution in SpecC, the *wh*-object is right-adjoined to the *wh*-subject, and the result is grammatical.²²

22. If the order of *wh*-subject and *wh*-object is reversed, ungrammaticality arises. This, among other things, motivates the assumption that adjunction to a filled SpecC-position always has to be to the right; see Rudin (1988) for extensive discussion. Note incidentally that there is no strict order of subject and object in Bulgarian if only [-wh] phrases are involved — Bulgarian in principle permits scrambling in front of the subject (see, for example, Molxova 1970: 27 and Müller and Sternefeld 1993: 474).

- (34) $[_{CP} [_{SpecC} [_{SpecC} Koj_1] \quad kogo_2] [_{IP} t_1 \text{ vižda } t_2]]?$
 who(NOM) whom(ACC) sees

Just as in Korean and English, there is no optionality involved in Bulgarian *wh*-movement. An English-like version of (34) where *kogo* remains *in situ* at S-structure is illicit (cf. [35a]), as is a Korean-like version with strict *wh-in situ* (cf. [35b], which is of course string-identical to [35a] in the case at hand):

- (35) a. $*[_{CP} [_{SpecC} Koj_1] [_{IP} t_1 \text{ vižda } kogo_2]]?$
 who(NOM) sees whom(ACC)
 b. $*[_{CP} \text{ — } [_{IP} Koj_1 \quad \text{ vižda } kogo_2]]?$
 who(NOM) sees whom(ACC)

3.2.4. *The parameter.* Recall that one of the basic premisses of Optimality Theory is that constraints are universal (cf. [21]); therefore, resort to standard concepts of parametrization (“the constraint comes equipped with an open parameter that is settled differently from language to language,” “the constraint holds in some languages, but not in others”) is not possible. Rather, cross-linguistic variation is traced back to constraint re-ranking in Optimality Theory. I would like to suggest the following ranking of the constraints PROJ-PRIN, DER-ECON and WH-CRIT in English (on the ranking of the other constraints introduced in section 3.1, see below):

- (36) Ranking in English:
 PROJ-PRIN > WH-CRIT > DER-ECON

As noted before, DER-ECON and WH-CRIT impose conflicting requirements on *wh*-phrases: WH-CRIT triggers movement, DER-ECON prohibits movement. However, since DER-ECON is ranked lower than WH-CRIT, DER-ECON can be violated by a derivation if such a violation leads to an improved behavior on WH-CRIT. In other words, the ranking in (36) implies that it is better to have *wh*-movement than to have *wh in situ* in English. Still, by assumption, PROJ-PRIN is ranked higher than WH-CRIT in this language. Again, these two constraints are in conflict: WH-CRIT induces multiple overt *wh*-movement, since it requires every *wh*-phrase to occupy a SpecC (substitution or adjunction) position at S-structure, whereas PROJ-PRIN prohibits adjunction to SpecC. Since PROJ-PRIN is ranked higher than WH-CRIT in English, it is obvious how this conflict is resolved — there is no overt adjunction to SpecC in English, that is, all *wh*-phrases but one remain *in situ* in multiple questions, at the cost of WH-CRIT violations. All this is shown schematically for the sentences in (30) in Table 1.²³

23. A few remarks on notation: D_1, D_2 etc. are the derivations that compete (the candidates from the reference set); following the colon is an abbreviated version of the S-structure generated by the respective derivation. Thus, D_1 is a derivation that generates an S-structure with both

Table 1. Wh-movement in English

Candidates	PROJ-PRIN	WH-CRIT	DER-ECON
*D ₁ : [_{CP} — [_{IP} ... <i>what</i> ₁ <i>to whom</i> ₂]]		**!*	
→D ₂ : [_{CP} <i>what</i> ₁ [_{IP} ... t ₁ <i>to whom</i> ₂]]		*	*
*D ₃ : [_{CP} <i>what</i> ₁ <i>to whom</i> ₂ [_{IP} ... t ₁ t ₂	*!		**

Here, the optimal candidate is derivation D₂, with one violation of WH-CRIT (because *to whom* is *in situ* at S-structure), and one violation of DER-ECON (because *what* has undergone movement to SpecC at S-structure). In contrast, D₁ exhibits three violations of WH-CRIT: the (a)-part of (25) is violated twice (with both *wh*-phrases *in situ*), and the (b)-part of (25) is violated once (C_[+wh] is not lexicalized by filling of C or SpecC). However, the second WH-CRIT violation of D₁ is already fatal, because this is the highest-ranking constraint where D₁ and D₂ differ, and D₂ violates WH-CRIT only once, which implies that D₂ satisfies WH-CRIT better than D₁. Derivation D₃, on the other hand, does not contain any violation of WH-CRIT; but it violates the highest-ranked constraint PROJ-PRIN (in addition to a double violation of the lowest-ranked constraint DER-ECON), and the PROJ-PRIN violation is decisive. Thus, D₂ is far from perfect from an absolute point of view, but it is the best that one can get, given the constraint ranking in (36), and therefore it is well formed.²⁴

Given the definition of optimality in (22), no number of violations of a lower-ranked constraint can outweigh a single violation of a higher-ranked con-

wh-phrases *in situ*; D₂ is a competing derivation (that is, one with the same LF) that generates an S-structure with one *wh*-phrase in SpecC, and one *wh*-phrase *in situ*; and D₃ is a competing derivation that generates an S-structure with both *wh*-phrases in SpecC. Every violation of a constraint by a derivation is indicated by a star. If a violation is fatal for a candidate, the star that is assigned as a consequence of this violation is followed by an exclamation mark. An optimal (grammatical) derivation (according to the definition of optimality in (22) above) is assigned an arrow →, non-optimal (ungrammatical) derivations are assigned a star. This latter star (which is not to be confused with the star assigned if a constraint is violated), the arrow →, and the exclamation mark are strictly speaking redundant; their only purpose is to simplify exposition.

24. Note that the present system does not say anything about which *wh*-phrase has to undergo movement to SpecC in a multiple question. Thus, if nothing more is said, a derivation in which *to whom* has undergone overt movement to SpecC, and *what* stays *in situ*, is also optimal. In what follows, I neglect the issue of Superiority effects (cf. Chomsky 1973), and their derivation by invoking a concept like the Shortest Path Condition (cf. Chomsky 1993; 1995); see Ackema and Neeleman (1995) and Müller (1996) for optimality-theoretic approaches.

straint.²⁵ However, it might seem at first sight as though the optimal candidate in Table 1 could also be determined by simply counting the number of violations, without recourse to constraint ranking (D_1 and D_3 exhibit three violations, D_2 involves only two). That this is but a mere coincidence will become apparent when we consider *wh*-movement in Korean-type languages. I would like to suggest that PROJ-PRIN, DER-ECON and WH-CRIT are ranked as in (37) in Korean:

(37) Ranking in Korean:

PROJ-PRIN > DER-ECON > WH-CRIT

The only difference from English is that DER-ECON and WH-CRIT have changed places. This implies that derivations with *wh in situ* will typically win against competing derivations with overt *wh*-movement — it is now more important to satisfy DER-ECON than to satisfy WH-CRIT. Table 2 shows that the derivation generating the S-structure representation in (31) (namely, D_1) satisfies the ordered set of constraints in Korean better than all competing candidates:

Table 2. *Wh-movement in Korean*

Candidates	PROJ-PRIN	DER-ECON	WH-CRIT
$\sim D_1$: [_{CP} — [_{IP} ... <i>muös-ül</i> ₁ <i>wae</i> ₂ ...]]			**
* D_2 : [_{CP} <i>muös-ül</i> ₁ [_{IP} ... <i>wae</i> ₂ ...]]		*!	*
* D_3 : [_{CP} <i>muös-ül</i> ₁ <i>wae</i> ₂ [_{IP} ... <i>t</i> ₁ <i>t</i> ₂ ...]]	*!	**	

Evidently, D_2 involves just as many violations as D_1 , so simple counting does not help here. Rather, what is relevant is that D_1 satisfies the highest-ranked constraint on which D_1 and D_2 differ (namely, DER-ECON) better than D_2 — no number of violations of WH-CRIT by D_1 could turn D_2 into a winning candidate.²⁶ Thus, D_1 is optimal, and D_2 and D_3 (which has a fatal violation of PROJ-PRIN) are not.²⁷

25. This is one of the main conceptual differences between Optimality Theory and earlier connectionist approaches; cf. Prince and Smolensky (1993, chapt. 10).

26. Note that D_1 violates WH-CRIT only twice (unlike D_1 in Table 1). This is so because the (b)-part of (25) is fulfilled by the Q-morpheme *ni*. I turn to this issue immediately, in section 3.2.5.

27. Whereas the ranking of all three constraints is crucial in English, the actual position of PROJ-PRIN is irrelevant for the Korean data under discussion. If PROJ-PRIN were ranked between DER-ECON and WH-CRIT, or even below WH-CRIT, D_1 would still come out as optimal. The only difference would be that D_3 's fatal violation would be one of DER-ECON. The exact position of PROJ-PRIN in Korean can only be determined by taking into account other constructions (multiple topicalization being an obvious candidate). Since nothing depends on this issue for the data currently under discussion, I will stick to the ranking in (37), mainly for expository reasons.

Suppose next that the Bulgarian ranking differs from the English one only in that WH-CRIT and PROJ-PRIN have changed places:

(38) Ranking in Bulgarian:

WH-CRIT > PROJ-PRIN > DER-ECON

This implies that it is more important in this language to have all *wh*-phrases in SpecC at S-structure, than to respect the prohibition against adjunction to SpecC (that is, PROJ-PRIN), or to respect the prohibition against overt movement in general (that is, DER-ECON). Accordingly, derivation D_3 generating the S-structure in (34) emerges as optimal and thereby blocks derivations D_2 (generating the S-structure in [35a]) and D_1 (generating the S-structure in [35b]). This is shown in Table 3.²⁸

Table 3. *Wh-movement in Bulgarian*

Candidates	WH-CRIT	PROJ-PRIN	DER-ECON
* D_1 : [_{CP} — [_{IP} <i>koj</i> ₁ <i>vižda kogo</i> ₂]]	*! **		
* D_2 : [_{CP} <i>koj</i> ₁ [_{IP} <i>t</i> ₁ <i>vižda kogo</i> ₂]]	*!		*
→ D_3 : [_{CP} <i>koj</i> ₁ <i>kogo</i> ₂ [_{IP} <i>t</i> ₁ <i>vižda t</i> ₂]]		*	**

Summarizing so far, it seems that an optimality-theoretic account in terms of the three constraints PROJ-PRIN, WH-CRIT and DER-ECON offers a straightforward explanation of the effects standardly derived by invoking some version of the *Wh*-movement Parameter. Before turning to an analysis of partial *wh*-movement, I will integrate a fourth constraint into the present framework, namely, FULL-INT, and discuss some of its effects on what has become known as “clausal typing” (cf. Cheng 1991).

3.2.5. *Clausal typing.* Cheng (1991) observes an interesting correlation: languages with overt *wh*-movement generally do not employ lexical question markers in C; the existence of a Q-morpheme in C_[+wh] positions seems to be confined to *wh in situ* languages. Furthermore, the Q-morpheme is often obligatory in *wh in situ* languages. This holds for Korean:

28. A consequence arises that is similar to the one mentioned for Korean in the last footnote. Whereas it is important that WH-CRIT dominates both PROJ-PRIN and DER-ECON in Bulgarian, the relative order of PROJ-PRIN and DER-ECON is not fixed by the data discussed here. Again, for reasons of simplicity, I postulate the ranking that differs minimally from the English one.

- (39) a. [_{CP} — [_{IP} *Nŏ-nŭn muŏs-ŭl₁ wae sa-ss-*] [_C *ni*]]?
 you(TOP) what(ACC) why buy-PAST Q
 b. * [_{CP} — [_{IP} *Nŏ-nŭn muŏs-ŭl₁ wae sa-ss-*] [_C —]]?
 you(TOP) what(ACC) why buy-PAST

The illformed sentence in (39b) differs from the wellformed sentence in (39a) (= [31]) only insofar as (39a) does, and (39b) does not, contain a Q-morpheme in C. The presence of a Q-morpheme is one way to satisfy the (b)-clause of WH-CRIT. However, given that the lexical Q-morpheme is an expletive that does not receive an interpretation at LF, it violates FULL-INT.²⁹ Thus, there is an inherent conflict between WH-CRIT and FULL-INT as regards Q-morpheme insertion. I would like to suggest that in *wh in situ* languages like Korean, WH-CRIT, although by itself ranked fairly low, is still ranked higher than FULL-INT:

- (40) Ranking in Korean (extended):

PROJ-PRIN > DER-ECON > WH-CRIT > FULL-INT

This implies that in Korean *wh*-clauses, which inevitably violate WH-CRIT (because DER-ECON is ranked higher than WH-CRIT), violations of WH-CRIT are nevertheless kept minimal by violating the lower ranked constraint FULL-INT. To see this, consider Table 4, which is an extension of Table 2.

Table 4. *Wh-movement in Korean, extended*

Candidates	PROJ-PRIN	DER-ECON	WH-CRIT	FULL-INT
→D ₁ : [_{CP} — [_{IP} ... <i>muŏs-ŭl₁ wae₂ ...</i>] <i>ni</i>]			**	*
*D ₂ : [_{CP} <i>muŏs-ŭl₁</i> [_{IP} ... <i>t₁ wae₂ ...</i>] <i>ni</i>]		*!	*	*
*D ₃ : [_{CP} <i>muŏs-ŭl₁ wae₂</i> [_{IP} ... <i>t₁ t₂ ...</i>] <i>ni</i>]	*!	**		*
*D ₄ : [_{CP} — [_{IP} ... <i>muŏs-ŭl₁ wae₂ ...</i>] —]			***!	
*D ₅ : [_{CP} <i>muŏs-ŭl₁</i> [_{IP} ... <i>t₁ wae₂ ...</i>] —]		*!	*	
*D ₆ : [_{CP} <i>muŏs-ŭl₁ wae₂</i> [_{IP} ... <i>t₁ t₂ ...</i>] —]	*!	**		

D₁–D₃ are identical to the derivations in Table 2; D₄–D₆ are counterparts of D₁–D₃ that lack a Q-morpheme *ni* in C. As can be read off Table 4, D₄–D₆

29. This is not to say that C_{+wh} does not receive an interpretation. On the contrary, given the theory of question semantics presupposed here (cf. Heim 1992; Stechow 1996a), such a C-node *does* contribute to interpretation. The point is that it is the C-node that carries semantic information, and not the lexical Q-morpheme inserted in that position — the interpretation of C is identical in languages with Q-morphemes and languages without (such as English).

satisfy FULL-INT better than D_1 - D_3 , which in turn fare better with respect to WH-CRIT. Since, by assumption, WH-CRIT is ranked higher than FULL-INT in Korean, D_1 is still the best derivation in its reference set — the next best candidate (D_4) has an additional, hence fatal, violation of WH-CRIT. Thus, the fact that *wh in situ* languages like Korean must resort to a Q-morpheme in $C_{[+wh]}$ is explained by a conspiracy of WH-CRIT, DER-ECON and FULL-INT: WH-CRIT is ranked between DER-ECON and FULL-INT, and therefore it can be violated to satisfy DER-ECON (that is, *wh*-phrases can stay *in situ*), but it cannot be violated to satisfy FULL-INT (that is, $C_{[+wh]}$ must be lexicalized by inserting an expletive Q-morpheme).³⁰

Let me now turn to the observation that languages with obligatory overt *wh*-movement (of either the English or the Bulgarian type) never seem to permit overt Q-morphemes in *wh*-clauses. This restriction becomes apparent when we consider a minimally different version of the English example (29b) where C is lexicalized by a Q-morpheme; compare (41a) (= [29b]) with (41b):

- (41) a. (*I wonder*) [_{CP} [_{SpecC} *what*₁] [_C —] [_{IP} *he gave t₁ to John*]]
 b. *(*I wonder*) [_{CP} [_{SpecC} *what*₁] [_C *if*] [_{IP} *he gave t₁ to John*]]

Recall that WH-CRIT is ranked higher than DER-ECON in English. Hence, *wh*-movement of *what* to SpecC must apply in (41) to satisfy the (a)-requirement of WH-CRIT. But this movement simultaneously satisfies the (b)-requirement of WH-CRIT — the [+wh] feature on C is lexicalized by filling of SpecC. Therefore, nothing is to be gained by an additional Q-morpheme insertion operation; and since such an operation violates FULL-INT, it is barred.³¹ Interestingly, this

30. There are *wh in situ* language where Q-morphemes are not obligatory, or even not present at all; cf., for example, Mandarin Chinese and Hindi. This poses a problem for Cheng's (1991) theory of clausal typing, where additional assumptions (concerning, inter alia, the existence of phonologically empty Q-morphemes) are called for. In the approach adopted here, a language where *wh in situ* occurs without a Q-morpheme (like Hindi; see, for example, Mahajan 1990: 128) has a ranking that differs from the Korean one in that FULL-INT dominates WH-CRIT. This would imply that the analogue to D_4 in Table 4 would be the optimal derivation. Optionality of a Q-morpheme in *wh in situ* languages could be derived by assuming equal ranking of WH-CRIT and FULL-INT. This would work for Mandarin Chinese matrix questions, but more would have to be said about embedded clauses (where a Q-morpheme does not show up, see Cheng 1991: 35). I will return to the concept of equal ranking in the following section.

31. A reviewer notes that Dutch permits constructions like (i-a), which at first sight seem to contradict the generalization that Q-morphemes (*of(dat)* in this case) cannot show up in C if SpecC is occupied by a *wh*-phrase.

- (i) a. *Ik vraag* [_{CP} [_{SpecC} *wat*₁] [_C *of(dat)*] [_{IP} *Jan gedaan heeft*]]
 I ask what if(that) Jan done has
 b. *Wie*₁ *denk je* [_{CP} *t₁ dat/of(dat) Jan gekust heeft*?]
 who(ACC) think you that/if(that) Jan kissed has

However, as Zwart (1993: 267) observes, *of(dat)* can also optionally show up in what must be a [-wh] C node if *wh*-extraction has taken place via its specifier, as in (i-b). This strongly

reasoning holds irrespective of where FULL-INT is ranked in English.³²

With this approach to the Wh-movement Parameter and clausal typing as a background, I will now address the peculiar properties of partial *wh*-movement highlighted in section 2, and show how they fall out of the theory of *wh*-dependencies adopted here.

4. Partial *wh*-movement and Optimality Theory

4.1. *Alternation, partiality, and scope marking*

Let me begin with the first problem discussed in section 2: partial *wh*-movement and successive-cyclic long-distance *wh*-movement co-exist in German; neither version seems to block the other one (cf. section 2.1). The relevant pair of examples is repeated in (42) (cf. [1]):

- (42) a. *Was₁ glaubst du* [_{CP} *wann₁ (daß) sie t₁ gekommen ist*]?
 [+wh] think you when that she come is
- b. *Wann₁ glaubst du* [_{CP} *t'₁ daß sie t₁ gekommen ist*]?
 when think you that she come is

A standard optimality-theoretic approach to free alternation of competing candidates is to postulate that two constraints are equally ranked (tied). This means that a violation of one constraint is just as important as a violation of the other one.³³ I would like to propose that the very existence of partial *wh*-movement in German does indeed result from a tie of two constraints, namely DER-ECON and FULL-INT. I will assume that the four constraints PROJ-PRIN, DER-ECON, WH-CRIT and FULL-INT are ranked as follows in German:

- (43) Ranking in German:
 PROJ-PRIN > WH-CRIT > DER-ECON/FULL-INT

Thus, German is basically like English (cf. [36]): WH-CRIT dominates DER-ECON, so that overt *wh*-movement is forced, and PROJ-PRIN in turn dominates WH-CRIT, so that only one *wh*-phrase undergoes S-structure movement in multiple questions. In addition, however, DER-ECON and FULL-INT are equally

suggests that *of(dat)* in Dutch can act as a [-wh] reflex on C of *wh*-movement to SpecC, just like *daß* in German; and it plays this role in both (i-a) and (i-b). Similar considerations apply to complementizers in the Celtic languages that indicate successive-cyclic *wh*-movement; see, for example, McCloskey (1979) for Irish and Hendrick (1988) for Breton and Welsh.

32. Note that the situation is slightly more involved if the target position of a *wh*-phrase is in a higher clause. In that case, ranking of FULL-INT will become important. I return to this issue below.

33. See, for example, Pesetsky (1994: 11) for an approach to optional complementizer deletion in English that is based on a tie of two constraints.

ranked. This implies that it is just as “costly” (from an optimality-theoretic point of view) to insert a scope marker *was* (thereby violating FULL-INT) as it is to move the *wh*-phrase (thereby violating DER-ECON). Hence, the alternation problem is solved: WH-CRIT requires that the *wh*-phrase is in SpecC at S-structure, and that the matrix C_[+wh] node is lexicalized. This can be accomplished either by long-distance *wh*-movement as in (42b) (with two violations of DER-ECON), or by partial *wh*-movement as in (42a) (with one violation of DER-ECON, and one violation of FULL-INT, the scope marker not being interpretable, that is, being an expletive) — both derivations are optimal.³⁴ Along the same lines, the problems addressed in sections 2.4 (partiality) and 2.5 (scope marking) are solved. The relevant examples are repeated here (cf. [11b] and [12b]):

- (44) a. *Was₁ glaubst du [CP daß sie wann₁ gekommen ist]?
 [+wh] think you that she when come is
 b. *[CP — glaubst du [CP wann₁ (daß) sie t₁ gekommen ist]]?
 think you when that she come is

The derivations generating the S-structure representations in (44a) and (44b) are blocked as non-optimal by those derivations that generate (42a) and (42b). Example (44a) has one violation of DER-ECON/FULL-INT less than (42a) and (42b), but it incurs a violation of WH-CRIT (the [a]-part). Since WH-CRIT is ranked higher than DER-ECON/FULL-INT, (44a) is blocked. Similarly, (44b) fares better than (42a) and (42b) with respect to DER-ECON/FULL-INT, but it also violates the higher ranked constraint WH-CRIT (the [b]-part), and is thus excluded. Another illformed S-structure representation that is generated by a derivation in the same reference set looks as in (45):

- (45) *Was₁ wann₁ glaubst du [CP t'₁ daß sie t₁ gekommen ist]?
 [+wh] when think you that she come is

Here, the scope marker *was* and the *wh*-phrase show up in the same SpecC-position (one item being adjoined). Such a derivation inevitably involves an additional violation of DER-ECON/FULL-INT, and a (fatal) violation of PROJ-PRIN, which, as we have seen, dominates all other constraints addressed so far in German. All this is shown schematically in Table 5; the tie of FULL-INT and DER-ECON is indicated by the lack of a vertical line between the two constraints.³⁵

34. This analysis does not offer a principled reason for why the scope marker must be an XP in SpecC in German, and cannot be an X⁰-category in the matrix C node, that is, a Q-morpheme. I presume that the choice between Q-morpheme and scope marker is mainly governed by morphological properties of a given language.

35. Clearly, this approach necessitates the assumption that clauses where *daß* ‘that’ shows up in C are not more costly than clauses where C is empty. In particular, one has to ensure that the presence of *daß* does not violate FULL-INT (in contrast to the presence of a scope marker, a Q-morpheme, or the auxiliary *do*); cf. Grimshaw (1995) for a proposal (based on a more elaborate definition of FULL-INT) that has this effect.

Table 5. *Partial and long-distance wh-movement in German*

Candidates	PROJ- PRIN	WH- CRIT	DER- ECON	FULL- INT
→D ₁ : [_{CP} was ₁ ... [_{CP} wann ₁ (daß) ... t ₁ ...]]			*	*
→D ₂ : [_{CP} wann ₁ ... [_{CP} t' ₁ daß ... t ₁ ...]]			**	
*D ₃ : [_{CP} was ₁ ... [_{CP} — (daß) ... wann ₁ ...]]		*!		*
*D ₄ : [_{CP} — ... [_{CP} wann ₁ (daß) ... t ₁ ...]]		*!	*	
*D ₅ : [_{CP} was ₁ wann ₁ ... [_{CP} t' ₁ daß ... t ₁ ...]]	*!		**	*

4.2. Anti-locality

Recall next from section 2.3 that partial *wh*-movement is “anti-local,” in the sense that the scope marker and the *wh*-phrase cannot be clause-mates; compare the illformed sentence (46a) (= [10a]) with the wellformed sentence (46b), where short *wh*-movement has applied:

- (46) a. *Was₁ ist [_{IP} sie warum₁ gekommen]?
 [+wh] is she why come
 b. Warum₁ ist [_{IP} sie t₁ gekommen]?
 why is she come

It remains to be shown why the derivation generating (46b) blocks the competing derivation generating (45a) as non-optimal. Given what has been said so far, this is straightforward: scope marking as in (46a) involves one violation of WH-CRIT (more precisely, of the [a]-part of (25) — the *wh*-phrase is not in SpecC), whereas short movement as in (46b) involves no violation of WH-CRIT (the [a]- and the [b]-parts are simultaneously satisfied by *wh*-movement to SpecC). To be sure, (46a) satisfies DER-ECON better than (46b), but this has no effect because it violates FULL-INT instead, and because WH-CRIT dominates DER-ECON in German anyway.

Additional candidates that compete with the derivations generating the sentences in (46) are the derivations that yield (47a) and (47b):

- (47) a. *— ist [_{IP} sie warum₂ gekommen]?
 is she why come
 b. *Was₁ warum₂ ist [_{IP} sie t₂ gekommen]?
 [+wh] why is she come

However, it is evident that these derivations do not satisfy the ordered set of constraints in German better than the derivation generating (46b)–(47a) exhibits two violations of WH-CRIT, the first of which is already fatal, and (47b) involves a fatal violation of PROJ-PRIN, in addition to one violation of DER-ECON and one violation of FULL-INT. Thus, the optimal version of short *wh*-movement in German is always one where DER-ECON is violated once, and PROJ-PRIN, WH-CRIT, and FULL-INT are not violated at all, as in (46b). This is shown in Table 6:

Table 6. *Short wh-movement in German*

Candidates	PROJ-PRIN	WH-CRIT	DER-ECON	FULL-INT
*D ₁ : [_{CP} <i>was</i> ₁ ... [_{IP} ... <i>warum</i> ₁ ...]]		*!		*
→D ₂ : [_{CP} <i>warum</i> ₁ ... [_{IP} ... <i>t</i> ₁ ...]]			*	
*D ₃ : [_{CP} — ... [_{IP} ... <i>warum</i> ₁ ...]]		*!*		
*D ₄ : [_{CP} <i>was</i> ₁ <i>warum</i> ₁ ... [_{IP} ... <i>t</i> ₁ ...]]	*!		*	*

Note that Table 6 provides evidence for the above assumption (cf. section 2.1.2) that scope marking constructions compete with *wh*-movement constructions. If D₁ and D₂ were not in the same reference set, D₂ could not block D₁, in the approach adopted here; rather, with D₂ out of the way, D₁ would incorrectly be predicted to be optimal.

4.3. Parametrization

4.3.1. *English*. Thus far, I have assumed that PROJ-PRIN is ranked higher than WH-CRIT in English, which in turn is ranked higher than DER-ECON (cf. [36]). However, I have not yet integrated FULL-INT into the English order of constraints. Clearly, since the tie of DER-ECON and FULL-INT is responsible for the free alternation of partial *wh*-movement and long-distance *wh*-movement in German, FULL-INT and DER-ECON cannot be equally ranked in English; recall from section 2.2 that English does not exhibit partial *wh*-movement.³⁶ The relevant data are repeated in (48) (= [8]):

- (48) a. *Why*₁ *do you think* [_{CP} *t*₁ *that she came t*₁]?
 b. **What*₁ *do you think* [_{CP} *why*₁ *she came t*₁]?

36. The following remarks are supposed to carry over to Dutch and other Germanic languages, where partial *wh*-movement is also absent.

I would like to suggest that FULL-INT is integrated into the English ranking as in (49):

(49) Ranking in English (extended):

PROJ-PRIN > WH-CRIT > FULL-INT > DER-ECON

That is, English differs minimally from German in that FULL-INT and DER-ECON are not tied; rather, FULL-INT dominates DER-ECON. Hence, it is more important to satisfy FULL-INT than it is to satisfy DER-ECON in English.³⁷ This has the effect that the derivation generating (48a), with its two DER-ECON violations, is the only one that comes out as optimal; it successfully blocks the derivation that yields the S-structure representation in (48b) because that latter derivation involves a fatal FULL-INT violation, in addition to one violation of DER-ECON. At this point, it is worth pointing out that the non-existence of partial *wh*-movement as in (48b) supports the assumption that reference sets are defined in terms of LF-identity, just like the non-existence of “local” scope markers did. If the derivations generating (48a) and (48b) did not compete, there would be no way to block (48b) by invoking (48a) — partial *wh*-movement would incorrectly be predicted to be optimal, despite the FULL-INT violation.³⁸

Apart from the prohibition against partial *wh*-movement, long-distance

37. Interestingly, Grimshaw (1995) and Vikner (1995b) arrive at essentially the same conclusion, on the basis of data from an independent empirical domain (V movement and *do*-support).

38. As pointed out by a reviewer, this analysis implies that the derivations generating (i-a, b) in English do not compete — if they did, (i-b) should be blocked by (i-a).

- (i) a. *Someone₁ was t₁ in the room*
 b. *There was someone₁ in the room*

Examples (i-a) and (i-b) have the same LF output (that is, they compete) if an “expletive replacement” approach to *there* constructions along the lines of Chomsky (1986b, 1993) is adopted. However, the first thing to note is that expletive replacement is not externally motivated with *there* (in contrast to *was*, given a Karttunen-type approach to question semantics); it can only be forced by theory-internal considerations (for example, relating to checking of Case or Agr features). Second, as has been observed by den Dikken (1995: 348), the postulation of *there* replacement at LF is incompatible with the hypothesis that binding principles apply exclusively at LF; cf. the contrast in (ii):

- (ii) a. *Some applicants₁ seem to each other₁ to be t₁ eligible for the job*
 b. **There seem to each other₁ to be some applicants₁ eligible for the job*

If expletive replacement applies at LF in (ii-b), the impossibility of anaphoric binding is unexpected, given that principle A holds at LF. Finally, as is well known (cf., for example, Drubig 1992: 175), *there* replacement at LF is not compatible with facts about relative scope. Thus, whereas (iii-a) is ambiguous, (iii-b) only permits a reading in which *must* has scope over *someone*; but if *someone* were to replace *there* at LF, we would expect it to be able to have wide scope, contrary to fact.

- (iii) a. *Someone₁ must be t₁ in the house*
 b. *There must be someone₁ in the house*

For these reasons, I conclude that expletive replacement does not apply in the case of *there* constructions; accordingly, (i-b) is not blocked by (i-a) because it has a different LF output.

wh-movement in English resembles long-distance *wh*-movement in German, as depicted in Table 5 above. Thus, the derivations that generate the illformed S-structure representations in (50) are also correctly excluded as non-optimal:

- (50) a. **What₁ do you think* [_{CP} *that she came why₁*]?
- b. *— *you think* [_{CP} *why₁ she came t₁*]?
- c. **What₁ why₁ do you think* [_{CP} *t₁ that she came t₁*]?

Table 7 summarizes the relevant options and shows that only long-distance *wh*-movement can emerge as optimal in English questions where a *wh*-phrase originating in an embedded clause takes matrix clause scope:

Table 7. *Partial and long-distance wh-movement in English*

Candidates	PROJ-PRIN	WH-CRIT	FULL-INT	DER-ECON
*D ₁ : [_{CP} <i>what₁</i> ... [_{CP} <i>why₁</i> ... <i>t₁</i>]]			*!	*
→D ₂ : [_{CP} <i>why₁</i> ... [_{CP} <i>t₁' that</i> ... <i>t₁</i>]]				**
*D ₃ : [_{CP} <i>what₁</i> ... [_{CP} — <i>that</i> ... <i>why₁</i>]]		*!	*	
*D ₄ : [_{CP} — ... [_{CP} <i>why₁</i> ... <i>t₁</i>]]		*!		*
*D ₅ : [_{CP} <i>what₁ why₁</i> ... [_{CP} <i>t₁' that</i> ... <i>t₁</i>]]	*!		*	**

4.3.2. *Northern German dialects.* The question arises as to why the alternation of long-distance movement and partial *wh*-movement of the German (and Hungarian, see below) type is a somewhat marked and unusual phenomenon across languages. The reason could be that it crucially involves the existence of a tie of two constraints. One might speculate that ties are always marked options (perhaps arising only in the course of language change), and that languages strive to (re-)establish a fixed order between two constraints in the long run. In the previous section, I have argued that English and Dutch avoid a tie of DER-ECON and FULL-INT by ranking FULL-INT higher. Suppose now that a language chooses the reverse strategy, such that DER-ECON is ranked higher than FULL-INT, and that all the other constraints are ranked as in German (cf. [43]). In such a language, violations of FULL-INT would not be as grave as violations of DER-ECON, and consequently, long-distance *wh*-movement would invariably be blocked by partial *wh*-movement. This situation arguably obtains in northern German dialects, where all kinds of *wh*-extraction from *daß*-clauses are prohibited, and scope marking is chosen as a means to express long-distance *wh*-dependencies (see, for example, van Riemsdijk 1982: 13; Fanselow 1987: 50; Grewendorf

1988: 306–313; and Müller 1993, chapt. 8).³⁹ I think that such an approach to rigid German dialects can be viable, but there are further complications that have to be resolved. One obvious complication is that whereas northern German dialects block *wh*-extraction from *daß*-clauses, they do not block *wh*-extraction from V2 clauses; compare (51a) (= [42b], with a northern German judgement) with (51b):

- (51) a. ?**Wann*₁ *glaubst du* [_{CP} *t*'₁ *daß sie* *t*₁ *gekommen ist*]?
 when think you that she come is
 b. *Wann*₁ *glaubst du* [_{CP} *t*'₁ *ist*₂ *sie* *t*₁ *gekommen* *t*₂]?
 when think you is she come

One way out of this dilemma would be to assume that (51b) does not actually involve long-distance extraction from a V2 clause; on this view, *glaubst du* 'think you' is a parenthetical expression. And indeed, although Tappe (1981) and Grewendorf (1988: 83–87) (among others) have argued against analyzing (51b) in terms of parenthesis, Reis (1996) presents new evidence that this analysis might be correct after all, and that V2 clauses are always strict islands in German. Moreover, strict islandhood of embedded V2 clauses is what one would expect based on evidence from other Germanic languages (cf. Vikner 1995a). I will, however, not embark on this issue here.

4.4. Longer dependencies

4.4.1. *Long-distance scope marking*. Up to now, I have confined myself to considering *wh*-dependencies where at most two SpecC-positions are involved. Now I turn to longer dependencies, as in (14). These examples are repeated here in (52):

- (52) a. [_{CP2} *Wann*₁ *meinst du* [_{CP3} *t*'₁ *daß sie gesagt hat* [_{CP4} *t*'₁ *daß sie* *t*₁
 when think you that she said has that she
 kommen würde]]]?
 come would
 b. [_{CP2} *Was*₁ *meinst du* [_{CP3} *was*₁ *sie gesagt hat* [_{CP4} *wann*₁ *sie* *t*₁
 [+wh] think you [+wh] she said has when she
 kommen würde]]]?
 come would
 c. [_{CP2} *Was*₁ *meinst du* [_{CP3} *wann*₁ *sie gesagt hat* [_{CP4} *t*'₁ *daß sie* *t*₁
 [+wh] think you when she said has that she
 kommen würde]]]?
 come would

39. In fact, this analysis can be viewed as an implementation of van Riemsdijk's suggestion that "the existence of the scope marker *was* is related to the impossibility of long movement" in these dialects.

The three derivations generating these S-structure representations are all part of the same reference set. The task now is to show that they all turn out as optimal, under the assumptions made thus far. The derivation generating (52a) involves three violations of DER-ECON, and no violation of FULL-INT; the derivation that yields (52c) violates DER-ECON only twice, but incurs an additional FULL-INT violation; finally, the derivation with multiple scope marking as in (52b) violates DER-ECON once, but compensates for this by violating FULL-INT twice. Thus, given that in the case of ties, a violation of one constraint costs exactly as much as a violation of the other (so that, in effect, the two constraints count as one), the derivations generating (52a), (52b) and (52c) share the property of being optimal in their reference set, as intended.⁴⁰ However, closer scrutiny reveals a potential complication: it seems that long-distance scope marking, as in (53) (cf. [16] from section 2.7), poses a problem.

- (53) [_{CP2} Was₁ meinst du [_{CP3} daß sie gesagt hat [_{CP4} wann₁ sie t₁
 [+wh] think you that she said has when she
 kommen würde]]]]?
 come would

At first sight, it looks as though the derivation yielding (53) violates DER-ECON once (due to movement of *wann* in CP₄), and FULL-INT once (due to scope marker insertion in CP₂). Hence, with a total of two violations of DER-ECON/FULL-INT, it should block the three competing derivations yielding the S-structure representations in (52), an incorrect prediction. At this point, the constraint MIN-CHAIN (cf. [27]) comes into play, according to which chain formation must be minimal at each step of the derivation. Suppose that a scope marker must always be co-indexed with a *wh*-phrase; that is, scope marker insertion is an instance of chain formation, just like movement (see section 4.4.3 below). It then follows that the derivation just envisaged for (53) violates MIN-CHAIN: the intermediate SpecC₃ position is left empty, and chain formation applies long-distance. To ensure that this violation is fatal (so that this derivation does not block the derivations of the S-structure representations in [52]), we have to assume that MIN-CHAIN dominates DER-ECON/FULL-INT in German. For the time being, I will simply postulate that MIN-CHAIN also dominates WH-CRIT

40. Note, however, that Prince and Smolensky (1993: 51) envisage two definitions of ties. According to one notion, equal ranking of two constraints A, B actually means that a candidate can be successful on this tie only if it is successful under one of the two possible rankings of A and B. Under this notion of equal ranking, (52c) should be non-optimal — (52a) wins if FULL-INT dominates DER-ECON, (52b) wins if DER-ECON dominates FULL-INT, but (52c) does not win under either ranking of DER-ECON and FULL-INT. For this reason, I will assume the second notion of tie suggested by Prince and Smolensky (1993), according to which a tie implies a “crucial nonranking” of constraints — a definition of which Prince and Smolensky say that they “have not yet found evidence.” Thanks to Armin Mester for pointing this out to me.

and PROJ-PRIN (again, see section 4.4.3 for justification). Now we have derived that a true long-distance derivation of the string in (53) is impossible. Why, then, is (53) not ill formed? The answer is that there is another derivation of (53) that has exactly the same constraint profile as the derivations of the sentences in (52). In this derivation, the scope marker is inserted in the intermediate SpecC₃ position and then moved to the matrix SpecC₂ position, thereby inducing an additional violation of DER-ECON and circumventing a fatal MIN-CHAIN violation at the same time.

Table 8 shows that derivations D₁–D₄ (corresponding to the four S-structures in [52b], [52a], [52c], and [53], in that order) are all optimal since they satisfy the tied constraints FULL-INT and DER-ECON equally well.⁴¹ D₅ is the alternative derivation of (53) that does without scope marker movement; it is filtered out by MIN-CHAIN. Similarly, D₆ is a non-optimal derivation that violates MIN-CHAIN; it is a variant of D₂ that fails to establish the intermediate trace t'₁, by improving its behavior on DER-ECON/FULL-INT. Finally, D₇–D₉ are derivations that are ruled out in a by now familiar way, as involving a fatal violation of either PROJ-PRIN or WH-CRIT.

Table 8. *Multiple scope marking and wh-movement in German*

Candidates	MIN-CHAIN	PROJ-PRIN	WH-CRIT	DER-ECON	FULL-INT
→D ₁ : was ₁ ... was ₁ ... wann ₁ (daß) ... t ₁ ...				*	**
→D ₂ : wann ₁ ... t' ₁ daß ... t' ₁ daß ... t ₁ ...				***	
→D ₃ : was ₁ ... wann ₁ (daß) ... t' ₁ daß ... t ₁ ...				**	*
→D ₄ : was ₁ ... t ₁ daß ... wann ₁ (daß) ... t ₁ ...				**	*
*D ₅ : was ₁ ... — daß ... wann ₁ (daß) ... t ₁ ...	*!			*	*
*D ₆ : wann ₁ ... — daß ... t' ₁ daß ... t ₁ ...	*!			**	
*D ₇ : was ₁ wann ₁ ... t' ₁ daß ... t' ₁ daß ... t ₁ ...		*!		***	*
*D ₈ : was ₁ ... was ₁ ... was ₁ ... wann ₁ ...			*!		***
*D ₉ : — C _[+wh] ... — daß ... wann ₁ (daß) ... t ₁			*!	*	

There are many more derivations in the same reference set that must be excluded. For most of these, it seems clear that they satisfy the ordered set of constraints

41. Under the ranking DER-ECON > FULL-INT of northern German dialects, D₁ is predicted to be the sole optimal candidate. This result seems to be tenable; cf. van Riemsdijk (1982: 13).

worse than D_1 - D_4 . An interesting competing derivation, however, is one that involves a “late was”.

4.4.2. *Late was*. As noted in section 2.6, a scope marker *was* may never follow a *wh*-phrase it is co-indexed with; recall (15), which is repeated here as (54a), and contrasted with the minimally different (and wellformed) (52c), repeated as (54b):

- (54) a. $*[{}_{CP_2} Was_1 \text{ meinst du } [{}_{CP_3} wann_1 \text{ sie gesagt hat } [{}_{CP_4} was_1 \text{ sie } t_1$
 [+wh] think you when she said has [+wh] she
kommen würde]]]?
 come would
- b. $[{}_{CP_2} Was_1 \text{ meinst du } [{}_{CP_3} wann_1 \text{ sie gesagt hat } [{}_{CP_4} t'_1 \text{ daß sie } t_1$
 [+wh] think you when she said has that she
kommen würde]]]?
 come would

The relevant difference is that there is a complementizer *daß* in CP_4 in (54b), and a scope marker *was* in CP_4 in (54a). The derivations generating these two S-structures yield identical LFs, so it remains to be shown that any derivation that produces (54a) is blocked as non-optimal by derivation D_3 of Table 8 that produces (54b) (as well as by D_1 , D_2 , and D_4 of Table 8). Two derivations of (54a) must be considered: one in which *wann* ‘when’ moves successive-cyclically, via the SpecC-position of CP_4 , and the scope marker *was* is inserted later, thereby overwriting the intermediate trace; and one in which *wann* moves in one swoop to the SpecC-position of CP_3 , and *was* is inserted without deleting an intermediate trace. In Table 9 these two derivations are shown as D_2 and D_3 , respectively, together with D_1 (= D_3 from from Table 8).

Table 9. *The ban on late was*

Candidates	MIN-CHAIN	PROJ-PRIN	WH-CRIT	DER-ECON	FULL-INT
$\rightarrow D_1: was_1 \dots wann_1 (da\beta) \dots t'_1 da\beta \dots t_1 \dots$				**	*
$*D_2: was_1 \dots wann_1 (da\beta) \dots (t'_1) was_1 \dots t_1 \dots$				**	**!
$*D_3: was_1 \dots wann_1 (da\beta) \dots was_1 \dots t_1 \dots$	*!			*	**

MIN-CHAIN requires that chain formation operations be minimal at each step of the derivation. Movement is a chain formation operation, but it cannot be minimal at each step of the derivation in D_3 . No matter whether movement precedes scope marker insertion or vice versa, the movement operation by itself will skip a potential landing site in this derivation, thereby inducing a fatal MIN-CHAIN violation. Note that since MIN-CHAIN is formulated derivationally, it does not help

that the S-structure representation generated by D_3 remedies this violation (all chain links are indeed minimal at S-structure). As far as D_2 is concerned, it does not involve a severe violation: the only thing that distinguishes D_2 from the optimal derivation D_1 is that D_2 has an additional violation of the tied constraints DER-ECON and FULL-INT, which, per se, are ranked very low. This violation is induced by late insertion of the scope marker, which is therefore blocked. Intuitively, this means that, although scope marker insertion does not cost much from the point of view of optimality, it still costs something, and if there is nothing to be gained by the presence of a scope marker (that is, if its FULL-INT violation does not help to circumvent a violation of some other constraint), it is blocked by competing derivations as superfluous.

4.4.3. *Two assumptions.* To end this section, I will try to substantiate two assumptions that I have made in section 4.4.1, namely, (a) scope marker insertion involves chain formation, and (b) MIN-CHAIN is ranked very high. The first assumption seems to be a standard one; it follows immediately if we adopt the constraint on LF-movement in (55) (see Lasnik and Saito 1984, 1992, and Epstein 1992, among others):

- (55) LF *wh*-movement of a *wh*-phrase *X* that occupies SpecC at S-structure to another SpecC-position that is filled by some item *Y* at S-structure is possible only if *X* and *Y* are co-indexed.

Given (55), LF raising of a partially moved *wh*-phrase to the position of the scope marker is possible only if the two items are co-indexed. This constraint also ensures that partial *wh*-movement in German cannot apply if the target position is filled by another *wh*-phrase (rather than by a scope marker), as in (12c) (repeated here as [56a]):

- (56) a. * $[_{CP} \text{Wer}_2 \text{ glaubt } t_2 [_{CP} \text{ wann}_1 (\text{da\ss}) \text{ sie } t_1 \text{ gekommen ist}]]?$
 who thinks when that she come is
 b. $[_{CP} \text{ Wer}_2 \text{ glaubt } t_2 [_{CP} \text{ — da\ss sie wann}_1 \text{ gekommen ist}]]?$
 who thinks that she when come is

Without something like (55), we would wrongly predict that (56a), which avoids a WH-CRIT violation, blocks (56b), which minimizes DER-ECON violations. Clearly, to exclude (56a), we have to assume that (55) is either ranked very high (that is, above WH-CRIT) or in fact located in GEN (that is, inviolable).⁴²

42. It is argued in Müller and Sternefeld (1996) that (55) does not have to be stated as a primitive of grammar; it follows from the Principle of Unambiguous Binding (PUB) introduced in Müller and Sternefeld (1993), which demands that A-bar movement must be uniform in the sense that only positions of the same formal type can be used in the course of successive-cyclic movement. On this view, the PUB permits partial *wh*-movement in the presence of a scope marker because S-structure substitution in SpecC is followed by the same operation at LF (namely,

Turning now to the ranking of MIN-CHAIN, there is evidence that MIN-CHAIN at least dominates WH-CRIT in German. Violations of the *wh*-island condition by *wh*-movement are strictly prohibited in German, even with arguments (see Bayer 1990; Müller and Sternefeld 1993):

- (57) **Wen*₁ fragst du dich [_{CP} wann₂ sie t₂ t₁ getroffen hat]?
 whom(ACC) ask you REFL when she met has

If MIN-CHAIN is ranked higher than WH-CRIT, the derivation generating (57) is blocked by one in which *wen* ‘whom’ stays *in situ* at S-structure. Recall now that I have assumed that MIN-CHAIN is a derivational constraint; unlike DER-ECON, it is not confined to overt movement, but also applies between S-structure and LF. This produces the result that the optimal candidate is one that has *wen in situ* at LF, too. However, *wh in situ* is not interpretable under the semantics of questions along the lines of Karttunen (1977), which I assume here. Hence, we end up with a situation where there is no good output because the optimal candidate is uninterpretable; see Prince and Smolensky (1993: 50).⁴³

On a more general note, this analysis clearly presupposes that a derivation D_i with a *wh*-phrase *in situ* at LF can compete with another derivation D_j in which the *wh*-phrase is in SpecC_[+wh]. Thus, the notion of “identical LF output” in the definition of reference set (cf. [6]) must be understood in a sufficiently abstract sense, such that D_i and D_j can both be members of the same reference set even though the *wh*-phrase occupies different positions in D_i and D_j, and even though D_j is not interpretable. Although this task does not strike me as trivial, I assume that it can be accomplished in a fairly straightforward way (also see Kitahara 1993; Grimshaw 1995; and Ackema and Neeleman 1995 for related discussion).

4.5. Weak islands

Let me finally turn to the last of the eight problems with partial *wh*-movement in

substitution in the SpecC-position occupied by *was* at S-structure); but the PUB excludes (56a) because S-structure substitution of *wann* in SpecC is followed by a formally different operation at LF (namely, adjunction to *wer*). Just like (55), the PUB must be either inviolable or ranked very high.

43. Two further derivations must be blocked that proceed via intermediate adjunction to *wann*, at either S-structure or LF. Assuming that intermediate adjunction avoids a MIN-CHAIN violation (see Chomsky 1995), these derivations respect MIN-CHAIN. However, if intermediate adjunction of *wen* to *wann* takes place at S-structure, a violation of PROJ-PRIN occurs that is fatal because PROJ-PRIN dominates WH-CRIT in German. On the other hand, if *wen* stays *in situ* at S-structure and moves to the matrix SpecC position via intermediate adjunction at LF, the S-structure constraint PROJ-PRIN is respected. However, this derivation would violate the PUB (cf. the previous footnote), because adjunction to SpecC is followed by substitution in SpecC, and is thus excluded.

German raised in section 2, namely, the asymmetry between partial and long-distance *wh*-movement that shows up in weak island contexts. Thus, recall from section 2.8 that long-distance *wh*-movement of an argument (at S-structure or at LF) may cross weak islands with only weak ungrammaticality (or no ungrammaticality at all) arising; partial *wh*-movement across a weak island, in contrast, is strictly prohibited in German. I have discussed three instances of this weak island asymmetry: (a) negative islands; (b) CNPC islands; and (c) subject islands. The negative island data are repeated here:

- (58) a. **Was*₁ *glaubst du nicht* [_{CP} *wen*₁ (*daß*) *Hans t*₁ *getroffen hat*]?
 [+wh] think you not whom(ACC) that Hans met has
 b. ?*Wen*₁ *glaubst du nicht* [_{CP} *t*₁ *daß* *Hans t*₁ *getroffen hat*]?
 whom(ACC) think you not that Hans met has

4.5.1. *Deriving the asymmetry.* The guiding idea will be that the derivation D_1 that yields (58b) blocks the competing derivation D_2 that yields (58a) as non-optimal. Both D_1 and D_2 violate BAR-CON (cf. [28]), according to which chain formation must not cross a barrier. But recall that this does not yet imply that D_1 and D_2 satisfy BAR-CON equally well. It might be that one of the two derivations violates BAR-CON less often than the other one, and, given that they satisfy all other constraints equally well, this means that one blocks the other as non-optimal. Indeed, it turns out that this situation arises in the case at hand. Since movement and scope marker insertion both involve chain formation, both D_1 (generating [58b]) and D_2 (generating [58a]) exhibit a BAR-CON violation at S-structure. However, BAR-CON is a derivational constraint that *inter alia* applies to the mapping from S-structure to LF. Since partial *wh*-movement at S-structure must be followed by LF raising to the position of the scope marker, this implies that D_2 involves a second, fatal violation of BAR-CON (induced by LF-movement) that D_1 does not exhibit (there is no additional LF *wh*-movement operation). Intuitively, if the best derivation in a reference set must violate BAR-CON, then it will come out as optimal, as long as violations of BAR-CON are minimal.

In order to derive this result from the ranking of constraints in German, we have to ensure that BAR-CON is dominated by WH-CRIT in this language, in contrast to, for example, MIN-CHAIN. The revised ranking in German is given in (59) (compare [43]):

- (59) Ranking in German (extended):

MIN-CHAIN > PROJ-PRIN > WH-CRIT > BAR-CON > DER-ECON/FULL-INT

If BAR-CON were ranked higher than WH-CRIT, this would mean that weak islands would strictly block extraction of any kind in German, in complete analogy to what has been said about MIN-CHAIN in section 4.4.3 — the optimal derivation would be an uninterpretable one with *wh in situ* at LF. But if BAR-CON is ranked lower than WH-CRIT, it can be violated by *wh*-movement to fulfill WH-CRIT at S-structure,

as required.⁴⁴ The competition between D_1 (cf. [58b]), D_2 (cf. [58a]), and other non-optimal derivations in the same reference set is shown in Table 10.

Table 10. *Weak islands and partial wh-movement in German*

Candidates	MIN-CHAIN	PROJ-PRIN	WH-CRIT	BAR-CON	DER-ECON	FULL-INT
$\neg D_1$: <i>wen</i> ₁ ... Neg ... <i>t</i> ' ₁ <i>daß</i> ... <i>t</i> ₁ ...				*	**	
* D_2 : <i>was</i> ₁ ... Neg ... <i>wen</i> ₁ (<i>daß</i>) ... <i>t</i> ₁ ...				**!	*	*
* D_3 : — ... Neg ... <i>wen</i> ₁ (<i>daß</i>) ... <i>t</i> ₁ ...			*!	*	*	
* D_4 : <i>was</i> ₁ <i>wen</i> ₁ ... Neg ... <i>t</i> ' ₁ <i>daß</i> ... <i>t</i> ₁ ...		*!		*	**	*
* D_5 : <i>was</i> ₁ ... Neg ... <i>daß</i> ... <i>wen</i> ₁ ...	*!		*	**		*
* D_6 : — ... Neg ... <i>daß</i> ... <i>wen</i> ₁			*!*	*		

Thus, although normally partial *wh*-movement and successive-cyclic *wh*-movement co-exist as optimal in German, things are different in a weak island context; here, partial *wh*-movement is more costly than long-distance *wh*-movement, and therefore blocked. More generally, this means that we have again evidence that derivations with and without a scope marker can compete, as postulated in section 2 above.⁴⁵

So far, so good. Still, upon closer scrutiny it turns out that this approach to weak island asymmetries between partial and long-distance movement raises at least two problems, which I will address in the following section.

4.5.2. *Problems and extensions.* A first potential problem with Table 10 is that D_1 , with overt *wh*-movement of an argument NP across a weak island, is predicted to be fully well formed, since it is optimal in its reference set. Whereas this result might perhaps be tenable in the case of negative islands, it is dubious in the case of CNPC islands and subject islands. The relevant data are repeated here:

44. Note that although I have postulated the ranking BAR-CON > DER-ECON/FULL-INT in (59), the data are compatible with the reverse ranking. Nothing hinges on this in the present context.

45. Note that this approach correctly predicts that true *wh in situ*, as in (i) (= [18c]) behaves like overt *wh*-movement, and not like partial *wh*-movement, with respect to weak islands:

(i) *Wer glaubt nicht* [_{CP} *daß* *Hans wen*₁ *getroffen hat*]?
 who(NOM) thinks not that Hans whom(ACC) met has

The derivation generating (i) violates BAR-CON only once (because of LF raising of *wen*), and there is no competing derivation that makes do with fewer BAR-CON violations (or, more generally, has a better constraint profile).

- (60) a. **Was*₁ *hast du* [_{NP} *ein Gerücht* t₃] *gehört* [_{CP3} *wen*₁ *Ede* t₁
 [+wh] have you a rumour heard whom(ACC) Ede
mag]?
 likes
- b. ??*Wen*₁ *hast du* [_{NP} *ein Gerücht* t₃] *gehört* [_{CP3} t'₁ *daß* *Ede* t₁
 whom(ACC) have you a rumour heard that Ede
mag]?
 likes
- (61) a. **Was*₁ *ist es schade* [_{CP} [_{PP} *mit* *wem*]₁ *Hans* t₁ *gesprachen hat*]?
 [+wh] is it too bad with whom Hans spoken has
- b. ??[_{PP} *Mit* *wem*]₁ *ist es schade* [_{CP} t'₁ *daß* *Hans* t₁ *gesprachen hat*]?
 with whom is it too bad that Hans spoken has

There is a clear contrast between, for example, (61a) and (61b), and this contrast is correctly predicted. However, (61b) itself is not fully grammatical; rather, it has a variable, intermediate status that is standardly classified as being due to a Subjacency(like) violation (see, for example, Chomsky 1986a and Chomsky and Lasnik 1993). The same goes for (60b), which correctly blocks (60a) but nevertheless is not fully well formed. At first sight, it seems that deriving an intermediate status of a derivation is not possible in an optimality-theoretic approach — the best derivation should be perfect. Towards a solution of this problem, suppose that BAR-CON is in fact duplicated in the grammar. First, it is a violable constraint which is part of the optimality-theoretic evaluation part of the grammar; second, it shows up as an inviolable constraint in GEN (in a minimally different form, such that LF-movement is not affected here; cf. note 12). Suppose next that a derivation that violates BAR-CON in GEN has an intermediate (that is, Subjacency-like) status which — in contrast to derivations with genuine, “strong” violations in GEN — does not prevent it from being passed over to the optimality-theoretic system of constraints that are universal, violable and ranked. Then, an optimal derivation involving a BAR-CON violation (such as D₁ in Table 10) still comes equipped with a ?(?) received in GEN, and the result is that, although it satisfies the ordered set of constraints better than all its competitors, it is still not perfect, due to an “early” GEN violation.

Other approaches to reduced wellformedness in Optimality Theory are conceivable which might avoid the redundancy that results from duplicating a constraint in two different parts of the grammar. I will, however, not pursue this issue here, confining myself to noting that the problem at hand also shows up in more or less the same form in other theories of movement such as those developed by Cinque (1990, chapt. 1) and Rizzi (1990, chapt. 1); see Sternefeld (1991, 169) for discussion.

That said, let me turn to a second problem that concerns the well-known fact that weak islands show argument/adjunct asymmetries. Consider adjunct extraction from a negative island in (62a), from a CNPC island in (62b), and from a subject

island in (62c) (and compare the analogous cases of argument extraction in [58b], [60b], and [61b], respectively):

- (62) a. **Warum₁ glaubst du nicht* [_{CP} *t'₁ daß Fritz t₁ geschlafen hat*]?
 why think you not that Fritz slept has
 b. **Wie₁ hast du* [_{NP} *ein Gerücht t₃*] *gehört* [_{CP₃} *t'₁ daß Ede t₁ schnarcht*]?
 how have you a rumour heard that Ede snores
 c. **Wie₁ ist es schade* [_{CP} *t'₁ daß Fritz t₁ mit Ede gesprochen hat*]?
 how is it too bad that Fritz with Ede spoken has

In the optimality-theoretic approach adopted here, an analysis of this fact will rely on the idea that the derivations generating the S-structure representations in (62) are blocked by other derivations in the same reference set as non-optimal. These optimal derivations can in turn either be well formed (in which case they yield acceptable S-structure representations), or they can involve a “no good output” situation. I will assume the latter — thus, the derivation generating, for example, (62a) is blocked by an optimal derivation in the same reference set that does not produce an interpretable output. I have argued for such a case of no good output in section 4.4.3 already — if MIN-CHAIN dominates WH-CRIT, *wh*-islands will always be strict in a language, blocking *wh*-extraction of any kind (unless there is a way to extract from a *wh*-island without violating MIN-CHAIN; see section 5.5). But note now that whereas the ranking MIN-CHAIN > WH-CRIT correctly predicts the facts in German, it does not in a language like English. Here, *wh*-islands are somewhat less strict in the case of argument extraction, and typically give rise to an intermediate status (see, for example, Chomsky 1986a). This suggests that MIN-CHAIN is dominated by WH-CRIT in English. However, if this is so, the very same problem arises in English with respect to MIN-CHAIN that we have encountered in German with respect to BAR-CON — adjunct extraction from *wh*-islands is completely ungrammatical in English.

Thus, what seems to be called for is a means to account for argument/adjunct asymmetries. Standardly, this role is played by the Empty Category Principle (ECP). In what follows, I will briefly sketch an analysis that incorporates the main insights behind the approaches developed in Lasnik and Saito (1984, 1992), Chomsky (1986a), and Chomsky and Lasnik (1993). First, suppose that the ECP is an LF filter, as in (63):

- (63) ECP:
 Offending traces must not be present at LF.

A trace is an offending trace iff its generation violates a locality constraint like BAR-CON or MIN-CHAIN (this is the analogue to [-γ]-marked traces in Lasnik and Saito 1984 and to *-marked traces in Chomsky and Lasnik 1993). Offending traces at LF can be avoided by trace deletion; but the only kinds of traces that can be deleted during the derivation are intermediate traces of arguments; adjunct traces

and initial argument traces are undeletable. Under these assumptions, it follows that the (intermediate) adjunct traces in (62) violate the ECP at LF (in addition to the BAR-CON violation that occurs in the derivation), whereas no such additional ECP violation occurs in the case of argument extraction across a weak island (given that the offending intermediate trace is deletable).⁴⁶ What remains to be clarified at this point is the ranking of the ECP. If the ECP is ranked lower than MIN-CHAIN or BAR-CON, its effects are inevitably overridden; however, if the ECP dominates MIN-CHAIN or BAR-CON, it may induce argument/adjunct asymmetries. In particular, this situation arises with *wh*-movement if, in a given language, the ECP is ranked higher than WH-CRIT, and MIN-CHAIN or BAR-CON is ranked lower than WH-CRIT. As far as German is concerned, I would like to propose that indeed, the ECP dominates WH-CRIT, which in turn dominates BAR-CON. Hence, the BAR-CON violation in (62a), (62b), and (62c), which per se is not fatal for *wh*-movement, automatically triggers a violation of the higher-ranked ECP, which is fatal — a competing candidate with *wh in situ* (hence, without an ECP violation) is optimal, but uninterpretable. Similarly, whereas the effects of the ECP cannot be seen in the case of MIN-CHAIN violations in German (cf. section 4.4.3), they can be seen in English, assuming that here, the ECP dominates WH-CRIT, which in turn dominates MIN-CHAIN.

To summarize, the approach to weak islands adopted here shares many of the properties of standard analyses as developed in, for example, Chomsky (1986a), Rizzi (1990), Cinque (1990), and Chomsky and Lasnik (1993). Recurrent features of all these approaches, as different as they may seem upon first glance, are (a) a statement to the effect that weak islands prohibit movement or chain formation, and (b) a distinct statement that this prohibition is stronger in the case of adjuncts than in the case of arguments. Moreover, as noted above, the problem of ensuring an intermediate status of weak island violations with arguments that arose in the present analysis also exists in the approaches developed in Rizzi (1990) and Cinque (1990). However, notwithstanding the fact that the present analysis can to a large extent be viewed as a version of standard approaches to locality, it has two notable properties: First, *wh*-islands on the one hand and negative islands, CNPC islands, and subject islands on the other hand are not derived by a single constraint, but

46. Two remarks are in order. First, it might actually be more appropriate to assume that trace deletion can only affect intermediate traces of items that are *referential* in a sense to be made precise, rather than intermediate argument traces; cf. Aoun (1986), Rizzi (1990), and Cinque (1990), among others. To simplify matters, I abstract away from this qualification here. And second, note that in this approach, an initial argument trace can create an ECP violation if its generation implies a violation of, for example, BAR-CON. If we were to assume (in contrast to what is stated in section 3.1.6) that direct selection of IP by C fails to take place if C is filled by a complementizer without agreement features, IP would turn into a barrier for a subject trace in this context and would thereby induce an ECP violation, in addition to a BAR-CON violation; cf. Koster (1987) and Müller (1993) for extensive discussion.

rather by two distinct constraints (MIN-CHAIN and BAR-CON, respectively); this move is empirically motivated by the fact that the two types of islands do indeed not form a homogeneous class in German — *wh*-islands are strict for *wh*-extraction of arguments in this language, whereas the other kinds of islands are not (that is, they are “weak”). And second (and most importantly), the peculiar behavior of partial *wh*-movement with respect to weak islands in German follows without additional assumptions.

It is clear that there is much more that could (ultimately, should) be said about the role of locality constraints in the present approach, but I will leave it at that, and turn to an extension of the analysis to another language that employs *wh*-scope marking.

5. Partial *wh*-movement in Hungarian

5.1. Premisses

Wh-movement strategies in Hungarian are interesting in the present context because this language shares properties with both the Bulgarian type and the German type: like Bulgarian, Hungarian exhibits multiple *wh*-movement at S-structure, and like German, partial *wh*-movement is possible. In this section, I will examine how Hungarian fits into the system developed so far.⁴⁷ Simplifying somewhat, Hungarian clauses are typically divided into three parts — the topic position(s), followed by the focus position(s), followed by the finite verb and what Kiss (1987) calls the “propositional component.” I will assume, following Marácz (1990), that the topic and focus positions arise as a consequence of the possibility of CP recursion, that the finite verb is in C at S-structure, and that the propositional component is identified as IP. *Wh*-movement in Hungarian is focus movement; it ends up in the lowest SpecC-position, immediately in front of the finite verb in C, and following the complementizer in embedded clauses. In multiple questions, one *wh*-phrase is substituted in SpecC, and we may assume that the remaining *wh*-phrases are adjoined to SpecC; cf. (64) (from Kiss 1987: 56–57):

- (64) *Marinak* [_{SpecC} [_{SpecC} *ki*₁] *mit*₂] *adott* *t*₁ *t*₂ *nászajándékba?*
 Mary(DAT) who what(ACC) gave wedding-present-to

Since S-structure movement of both *ki*₁ and *mit*₂ in (64) is obligatory, it is crucial that WH-CRIT is ranked higher than both PROJ-PRIN and DER-ECON in Hungarian,

47. The following analysis should be viewed as tentative. Unless quoted otherwise, the data are due to Imre Szigeti (personal communication).

as in Bulgarian (see Table 3). For the time being, suppose that this is the only difference to German as regards the constraints discussed up to now.⁴⁸

5.2. Partial versus long-distance *wh*-movement

Long-distance *wh*-movement alternates with partial *wh*-movement in Hungarian; in the latter case, a scope marker *mit* 'what' shows up in the matrix SpecC-position (cf. Marác 1990: 297–298, 325).⁴⁹

- (65) a. *Kit*₁ *gondolsz* [_{CP} — *hogy* [_{CP} *t*'₁ *látta t*₁ *Jánost*]]?
 who(ACC) you think that saw John(ACC)
 b. *Mit*₁ *gondolsz* [_{CP} — *hogy* [_{CP} *ki*₁ *látta Jánost t*₁]]?
 [+wh] you think that who saw John(ACC)

As in German, this follows if DER-ECON and FULL-INT are tied. Indeed, the parallelism does not stop here. Like German, Hungarian does not permit scope marking without partial *wh*-movement (cf. [66a]), and it does not permit partial *wh*-movement without scope marking either (cf. [66b]); see Marác (1990: 336–337) and Brody (1995: 107–108). Moreover, an “unmotivated” scope marker *mit* co-occurring with long-distance *wh*-movement is prohibited (cf. [66c]).

- (66) a. **Mit*₁ *gondolsz* [_{CP} — *hogy* [_{CP} — *látta ki*₁ *Jánost*]]?
 [+wh] you think that saw who John(ACC)
 b. *— *gondolsz* [_{CP} — *hogy* [_{CP} *ki*₁ *látta t*₁ *Jánost*]]?
 you think that who saw John(ACC)
 c. **Mit kit*₁ *gondolsz* [_{CP} — *hogy* [_{CP} *t*'₁ *látta t*₁ *Jánost*]]?
 [+wh] who(ACC) you think that saw John(ACC)

The illformedness of (66a) and (66b) is due to a fatal violation of WH-CRIT, whereas the illformedness of (66c) results from a fatal violation of PROJ-PRIN. This is shown in Table 11, which differs from its German counterpart Table 5 only in that PROJ-PRIN and WH-CRIT have changed places — but since the optimal candidates satisfy both these constraints, this difference in ranking does not affect the competition.

48. The question arises of why WH-CRIT can only be fulfilled by movement to the lowest, V-adjacent SpecC-position, and not by movement to a higher one. Several possibilities arise; for example, we might revise WH-CRIT in such a way that it requires *wh*-operators to show up in the *lowest* SpecC-position of a CP “cluster,” where CP clusters can be trivial (one-membered), as in English or German, or multi-membered, as in Hungarian.

49. For reasons which do not concern me here, the embedded *wh*-subject acquires accusative Case assigned by the matrix verb on its way to the target position in (65a).

Table 11. *Partial and long-distance wh-movement in Hungarian*

Candidates	WH-CRIT	PROJ-PRIN	DER-ECON	FULL-INT
→D ₁ : [_{CP} <i>kit</i> ₁ ... [_{CP} ... <i>t</i> ' ₁ <i>látta</i> ... <i>t</i> ₁ ...]]			**	
→D ₂ : [_{CP} <i>mit</i> ₁ [_{CP} ... <i>ki</i> ₁ <i>látta</i> ... <i>t</i> ₁ ...]]			*	*
*D ₃ : [_{CP} <i>mit</i> ₁ ... [_{CP} ... — <i>látta</i> ... <i>ki</i> ₁ ...]]	*!			*
*D ₄ : [_{CP} — ... [_{CP} ... <i>ki</i> ₁ <i>látta</i> ... <i>t</i> ₁ ...]]	*!		*	
*D ₅ : [_{CP} <i>mit</i> ₁ <i>kit</i> ₁ ... [_{CP} ... <i>t</i> ' ₁ <i>látta</i> ... <i>t</i> ₁]]		*!	**	*

A further striking piece of evidence for treating Hungarian and German in a similar way is the fact that free alternation of partial and long-distance *wh*-movement does not exist in all Hungarian dialects. As noted by Marácz (1990: 339), there are dialects in which only partial *wh*-movement is permitted; these varieties are amenable to the same account as the northern German dialects discussed in section 4.3.2 (that is, the tie is resolved by ranking DER-ECON higher than FULL-INT).

5.3. Short *wh*-movement

Next, note that a scope marker *mit* in Hungarian can never exempt a *wh*-phrase in the same clause from undergoing local *wh*-movement; that is, just as in German, scope marker insertion is subject to an anti-locality requirement; this is shown for a multiple question in (67a). Co-occurrence of a scope marker and one or more locally moved *wh*-phrases is also precluded; cf. (67b) and (67c).

- (67) a. **Marinak mit adott ki₁ mit₂ nászajándékba?*
 Mary(DAT) [+wh] gave who what(ACC) wedding-present-to
 b. **Marinak mit ki₁ adott t₁ mit₂ nászajándékba?*
 Mary(DAT) [+wh] who gave what(ACC) wedding-present-to
 c. **Marinak mit ki₁ mit₂ adott t₁ t₂ nászajándékba?*
 Mary(DAT) [+wh] who what(ACC) gave wedding-present-to

The sentences in (67) are generated by derivations that compete with the derivation that creates the wellformed multiple question in (64), and they are successfully blocked by that derivation under present assumptions: (67a) and (67b) have a fatal WH-CRIT violation which is due to the fact that a *wh*-phrase has remained *in situ*; (67c), on the other hand, involves an additional, hence fatal, violation of PROJ-PRIN. Table 12 shows why derivation D₁, with multiple S-structure *wh*-movement and without scope marking (cf. [64]), comes out as an optimal candidate that blocks the competing derivations D₃–D₅ which generate the sentences in (67). However, derivation D₂ is also optimal. It differs from D₁ only in that the order of

the two *wh*-phrases is reversed — a possibility in Hungarian, where Superiority effects do not show up; cf. Kiss (1987).⁵⁰

Table 12. *Short wh-movement in Hungarian*

Candidates	WH-CRIT	PROJ-PRIN	DER-ECON	FULL-INT
$\rightarrow D_1$: [_{CP} <i>ki</i> ₁ <i>mit</i> ₂ ... [_{IP} ... <i>t</i> ₁ <i>t</i> ₂ ...]]		*	**	
$\rightarrow D_2$: [_{CP} <i>mit</i> ₂ <i>ki</i> ₁ ... [_{IP} ... <i>t</i> ₁ <i>t</i> ₂ ...]]		*	**	
* D_3 : [_{CP} [+wh] ... [_{IP} ... <i>ki</i> ₁ <i>mit</i> ₂ ...]]	*!*			*
* D_4 : [_{CP} [+wh] <i>ki</i> ₁ ... [_{IP} ... <i>t</i> ₁ <i>mit</i> ₂ ...]]	*!	*	*	*
* D_5 : [_{CP} [+wh] <i>ki</i> ₁ <i>mit</i> ₂ ... [_{IP} ... <i>t</i> ₁ <i>t</i> ₂ ...]]		*!*	**	*

The optimal candidates D_1 and D_2 involve violations of PROJ-PRIN, but given that this constraint is ranked lower than WH-CRIT in Hungarian, this cannot be avoided in a multiple question. Any candidate that tries to circumvent a PROJ-PRIN violation will automatically incur a fatal WH-CRIT violation. In sum, local scope marking is blocked as non-optimal in Hungarian as it is in German, simply because nothing is to be gained by a scope marker in this case.

5.4. Multiple scope marking, multiple movement

Let me now turn to longer *wh*-dependencies. Again, as in German, a *wh*-dependency that crosses two CPs can be created in various ways. As shown in (68a), one option is a combination of multiple scope marker insertion and short *wh*-movement (cf. Marácz 1990: 330). Another option is successive-cyclic long-distance movement of the *wh*-phrase, as in (68b). Third, a scope marker can be inserted in the matrix SpecC-position, with successive-cyclic *wh*-movement to the intermediate SpecC-position, as in (68c).

- (68) a. *Mit*₁ *gondolsz* [_{CP}₁ *hogy* *Mari* *mit*₁ *mondott* [_{CP}₂ *hogy* *ki*₁ [+wh] *you* *think* *that* *Mary* [+wh] *said* *that* *who látta* *t*₁ *Jánost*]]?
 saw *John*(ACC)

50. Throughout, I note the scope marker *mit* as [+wh] if ambiguity could arise in a table.

- b. *Kit*₁ *gondolsz* [_{CP1} *hogy Mari t*'₁ *mondott* [_{CP2} *hogy t*'₁ *látta t*₁
 who(ACC) you think that Mary said that saw
Jánost]]?
 John(ACC)
- c. *Mit*₁ *gondolsz* [_{CP1} *hogy Mari kit*₁ *mondott* [_{CP2} *hogy t*'₁
 [+wh] you think that Mary who(ACC) said that
*látta t*₁ *Jánost*]]?
 saw John(ACC)

This is expected: (68a) involves two FULL-INT violations plus one DER-ECON violation; (68b) involves three DER-ECON violations and no FULL-INT violation; and (68c) involves two DER-ECON violations and one FULL-INT violation. Given that FULL-INT and DER-ECON are tied, all these derivations come out as optimal. In analogy to German, we also expect to find cases of apparent long-distance *wh*-scope marking (given that these cases can be analyzed in terms of *wh*-movement of the scope marker), and this does indeed seem to be the case.⁵¹

- (69) (?)*Mit*₁ *gondolsz* [_{CP1} *hogy Mari — mondott* [_{CP2} *hogy ki*₁ *látta t*₁
 [+wh] you think that Mary said that who saw
Jánost]]?
 John(ACC)

Finally, Hungarian exhibits the ban on “late” scope marker insertion:

- (70) **Mit*₁ *gondolsz* [_{CP1} *hogy Mari kit*₁ *mondott* [_{CP2} *hogy mit*₁
 [+wh] you think that Mary who(ACC) said that [+wh]
*látta t*₁ *Jánost*]]?
 saw John(ACC)

As before, this follows because a derivation of (70) will either violate MIN-CHAIN (if *kit*₁ undergoes one-swoop movement to SpecC₁), or will incur a fatal violation of FULL-INT (if movement of *kit*₁ proceeds successive-cyclically, and the scope marker *mit* is inserted in the position of the intermediate trace in SpecC₂).

All this is summarized in Table 13 on page 296 (cf. Table 8, Table 9 in German). Here, D₁–D₄ are the derivations generating (68a), (68b), (68c), and (69); D₅ and D₆ are two derivations of (70); and D₇–D₉ are some non-optimal derivations in the same reference set that violate high-ranked constraints (PROJ-PRIN or WH-CRIT).

51. Example (69) is not assumed to be grammatical by Marác (1990: 330); I follow the judgement of Imre Szigeti (personal communication) here. Also recall note 9.

Table 13. *Multiple scope marking and wh-movement in Hungarian*

Candidates	MIN-CHAIN	WH-CRIT	PROJ-PRIN	DER-ECON	FULL-INT
$\rightarrow D_1: mit_1 \dots mit_1 \dots ki_1 \dots t_1 \dots$				*	**
$\rightarrow D_2: kit_1 \dots t'_1 \dots t'_1 \dots t_1 \dots$				***	
$\rightarrow D_3: mit_1 \dots kit_1 \dots t'_1 \dots t_1 \dots$				**	*
$\rightarrow D_4: mit_1 \dots t_1 \dots ki_1 \dots t_1 \dots$				**	*
$*D_5: mit_1 \dots kit_1 \dots (t'_1) mit_1 \dots t_1 \dots$				**	**!
$*D_6: mit_1 \dots kit_1 \dots mit_1 \dots t_1 \dots$	*!			*	**
$*D_7: mit_1 kit_1 \dots t'_1 \dots t'_1 \dots t_1 \dots$			*!	***	*
$*D_8: mit_1 \dots mit_1 \dots mit_1 \dots ki_1 \dots$		*!			***
$*D_9: - C_{[+wh]} \dots \dots ki_1 \dots t_1$		*!		*	

5.5. Multiple long-distance questions

Despite the fact that Hungarian employs a multiple *wh*-movement strategy, the data involving partial *wh*-movement considered so far have turned out to lend themselves to basically the same analysis as their German counterparts. Things become slightly more involved, though, as soon as we consider multiple long-distance questions, that is, questions where two (or more) *wh*-phrases originating in a [-*wh*] embedded clause take matrix scope. One option is that a scope marker is inserted in the [+*wh*] target position, and the *wh*-phrases both undergo movement to the local SpecC-position; cf. (71a). As shown in (71b), multiple embedded *wh*-movement is obligatory here, which we expect if WH-CRIT outranks PROJ-PRIN in Hungarian. The status of alternative realizations is initially somewhat more surprising (cf. Horvath 1986: 227 and Marác 1990: 331). As shown in (71c), it is possible that one *wh*-phrase undergoes partial movement in the embedded clause, and the other one raises to the [+*wh*] target position in the matrix clause. However, it is not possible that both *wh*-phrases show up in the matrix SpecC-position at S-structure; cf. (71d):

- (71) a. *Mit gondolsz* [_{CP} — *hogy* [_{CP} *János* [_{CP} *kinek*₁ *mit*₂ adott *t*₁ *t*₂]]]]?
gave

- b. **Mit gondolsz* [_{CP} — *hogy* [_{CP} *János* [_{CP} *mit*₂ *adott* [_{+wh}] you think that John what(ACC) gave *kinek*₁ *t*₂]]]])?
who(DAT)
- c. *Kinek*₁ *gondolsz* [_{CP} — *hogy* [_{CP} *János* [_{CP} *t*'₁ *mit*₂ *adott* *t*₁ *t*₂]]]])?
who(DAT) you think that John what(ACC) gave
- d. **Kinek*₁ *mit*₂ *gondolsz* [_{CP} — *hogy* [_{CP} *János* [_{CP} *t*'₁ *t*'₂ *adott* *t*₁ *t*₂]]]])?
who(DAT) what(ACC) you think that John gave

Example (71c) is a legitimate instance of partial *wh*-movement in a context where the [+wh] LF target position is not occupied by a scope marker at S-structure, but by another *wh*-phrase; and (71d) is ungrammatical even though Hungarian in principle permits multiple S-structure *wh*-movement. However, a closer look at the candidates in (71) reveals that these data are not quite as unexpected as one might assume at first sight:

Table 14. Multiple long-distance questions in Hungarian

Candidates	MIN-CHAIN	WH-CRIT	PROJ-PRIN	DER-ECON	FULL-INT
→D ₁ : [+wh] ... <i>kinek</i> ₁ <i>mit</i> ₂ ... <i>t</i> ₁ <i>t</i> ₂ ...			*	**	*
→D ₂ : <i>kinek</i> ₁ ... <i>t</i> ' ₁ <i>mit</i> ₂ ... <i>t</i> ₁ <i>t</i> ₂ ...			*	***	
*D ₃ : [+wh] ... <i>mit</i> ₂ ... <i>kinek</i> ₁ <i>t</i> ₂ ...		*!		*	*
*D ₄ : <i>kinek</i> ₁ <i>mit</i> ₂ ... <i>t</i> ' ₁ <i>t</i> ' ₂ ... <i>t</i> ₁ <i>t</i> ₂ ...			**!	****	
*D ₅ : <i>kinek</i> ₁ ... <i>mit</i> ₂ ... <i>t</i> ₁ <i>t</i> ₂ ...	*!			**	

D₃ (which underlies [71b]) is blocked by D₁ (which generates [71a]) as non-optimal because it avoids D₁'s PROJ-PRIN violation at the cost of a fatal violation of WH-CRIT; this is just the reverse of the situation found in German, a consequence of the different ranking of PROJ-PRIN and WH-CRIT. In contrast, D₂ (generating [71c]) has the same status as D₁: It avoids D₁'s FULL-INT violation at the cost of an additional DER-ECON violation that results from moving the *wh*-phrase to the matrix [+wh] position in successive-cyclic fashion. D₂ violates PROJ-PRIN (due to intermediate adjunction to the embedded SpecC-position), but so does D₁. Note that this kind of derivation is in principle an option in Hungarian (as opposed to German) because ranking PROJ-PRIN lower than WH-CRIT yields

the result that PROJ-PRIN violations are possible if they help to circumvent WH-CRIT and MIN-CHAIN violations; recall that MIN-CHAIN can be fulfilled in the course of extraction from a clause with a filled SpecC-position if adjunction to SpecC takes place.⁵² Clearly, successive-cyclic movement as in D₂ (which creates a PROJ-PRIN violation) is the only way to generate (71c); if the PROJ-PRIN violation of D₂ is avoided by one-swoop movement, as in D₅, the higher-ranked MIN-CHAIN is violated. Finally, D₄ (which creates [71d]) is blocked by D₁ and D₂ because it involves an additional PROJ-PRIN violation: intermediate adjunction to SpecC as required by MIN-CHAIN is accompanied by adjunction to SpecC in the matrix.⁵³

Direct analogues to (71c) are ungrammatical in German, and this follows because PROJ-PRIN is ranked higher than WH-CRIT here, so that a derivation generating (72a) via intermediate adjunction to SpecC is blocked by a competing derivation generating (72b).⁵⁴

- (72) a. **Wann₂ denkst du* [_{CP} *wen₁ sie t₂ t₁ getroffen hat*]?
 when think you whom(ACC) she met has

-
52. More generally, a partial ranking MIN-CHAIN > WH-CRIT > PROJ-PRIN paves the way for systematic violations of the *wh*-island constraint, as they do indeed occur in Hungarian; see Horvath (1986: 221), Marácz (1990: 322), and Kiss (1992: 85).
53. Interestingly, derivations like D₄ are possible in Bulgarian, which does not exhibit *wh*-scope marking and requires all *wh*-phrases to be in their LF target position at S-structure; cf. Rudin (1988: 455) and references cited there. Note that the present analysis presupposes that the two *wh*-phrases in D₄ cannot undergo movement from the embedded SpecC-position into the matrix SpecC-position as a single unit; in that case, only one PROJ-PRIN violation would be expected. There are various possibilities to ensure this. Perhaps the most straightforward solution would be to dispense with the idea that multiple *wh*-movement involves adjunction to SpecC, as assumed thus far. Indeed, Comorovski (1989, chapt. 3), Dayal (1996, chapt. 4), Beck (1996, chapt. 2), and Stechow (1996b) argue that multiple *wh*-movement should be handled in terms of adjunction to CP (or C'), and not to SpecC. Assuming this to be correct, we would have to modify one constraint in the present framework: WH-CRIT would have to require *wh*-operators to be in the "domain" of C (where the domain of a head X comprises an X-adjunct, an XP-adjunct, and SpecX, but crucially not a SpecX-adjunct). Furthermore, PROJ-PRIN could be formulated more like in Grimshaw (1995), as prohibiting S-structure adjunction to CP and C⁰ only. It can easily be verified that such a revision would leave all previous reasonings intact; in addition, it would straightforwardly ensure that the two *wh*-phrases in D₄ cannot move into the matrix clause as a single item. Also see Müller (1996) for a proposal of how to derive the Bulgarian superiority effects (that Rudin 1988 takes to be a strong argument in favor of SpecC-adjunction) in an approach where multiple *wh*-movement involves CP-adjunction.
54. Recall from section 4.4.3 that partial *wh*-movement in German is also blocked if the matrix SpecC-position is occupied by another *wh*-phrase at S-structure that originates in the matrix clause, so that MIN-CHAIN and PROJ-PRIN cannot be involved. I have argued that this is excluded by (55)/the PUB. However, D₁ and D₂ in Table 14 do not involve violations of (55)/the PUB, and neither does a derivation of (72a) that exhibits a combination of substitution of *wann* in the embedded SpecC-position, followed by adjunction of *wen*, followed by substitution of *wann* in the matrix SpecC-position, followed finally by LF adjunction of *wen* in the matrix clause.

- b. *Was*₁ *denkst du* [_{CP} *wen*₁ *sie wann*₂ *t*₁ *getroffen hat*]?
 [+wh] think you whom(ACC) she when met has

5.6. *Weak islands*

Let me finally turn to the issue of weak islands. Recall that weak islands induce asymmetries between partial and long-distance *wh*-movement (of arguments) in German; this was accounted for by ranking BAR-CON lower than WH-CRIT (cf. section 4.6). Interestingly, no such asymmetries arise in Hungarian. Thus, as noted by Marácz (1990: 306, 307, 329, 330), CNPC islands and subject islands are strict for both long-distance *wh*-movement and scope marking.⁵⁵ This is shown for CNPC effects in (73):

- (73) a. **Kit*₁ *említett(e)* [_{NP} *azt* *a tényt* [_{CP} *hogy Mari*
 who(ACC) he mentioned that(ACC) the fact(ACC) that *Mary*
*megcsókolt t*₁]]?
 kissed
- b. **Mit*₁ *említett(e)* [_{NP} *azt* *a tényt* [_{CP} *hogy Mari*
 [+wh] he mentioned that(ACC) the fact(ACC) that *Mary*
*kit*₁ *megcsókolt t*₁]]?
 who(ACC) kissed

Under present assumptions, this implies that BAR-CON dominates WH-CRIT in Hungarian. It then follows that the derivations that generate (73a) and (73b) are blocked by a competing derivation in which the *wh*-phrase does not leave the complex NP at either S-structure or LF, and which is therefore uninterpretable, in complete analogy to what has been said about MIN-CHAIN in German (cf. section 4.4.3).

To sum up: abstracting away from a few intervening factors, it seems that the main properties of partial and long-distance *wh*-movement in Hungarian are accounted for under the approach developed here if we assume that Hungarian differs from German in that WH-CRIT dominates PROJ-PRIN (which triggers multiple S-structure *wh*-movement), and is in turn dominated by BAR-CON (so that weak islands are strict for all kinds of *wh*-dependencies). The following ranking is minimally different from the one suggested for German in (59) (BAR-CON and PROJ-PRIN have changed places) and covers the data of this section:

- (74) *Ranking in Hungarian*
 MIN-CHAIN > BAR-CON > WH-CRIT > PROJ-PRIN > DER-ECON/FULL-INT

55. See, however, Horvath (1995: 102) for a different view concerning subject islands.

6. Concluding remarks

I have set out to present an account of several long-standing problems related to partial *wh*-movement in German; and I have argued that an approach in terms of Optimality Theory offers a straightforward explanation of the — at first sight — peculiar properties of partial movement, given the proper formulation of six well-known and (more or less) commonly accepted constraints (PROJ-PRIN, DER-ECON, WH-CRIT, FULL-INT, MIN-CHAIN, and BAR-CON), and given their correct ranking. From the perspective assumed here, the properties of partial *wh*-movement in German are not surprising; they correspond exactly to what we would expect under a specific ranking of these constraints. Furthermore, it has turned out that a re-ranking of two constraints (PROJ-PRIN and BAR-CON) offers a promising approach to partial *wh*-movement in Hungarian, thereby lending further support to the optimality-theoretic analysis developed here. In addition, the relative markedness of the partial *wh*-movement alternation of the German/Hungarian type could be accounted for; it is due to a tie of the two constraints DER-ECON and FULL-INT (ties, by assumption, being a marked option in universal grammar).

In the course of the discussion, it has turned out several times that what at first sight looks like a problem raised by partial *wh*-movement (namely, its alternation with successive-cyclic long-distance movement) eventually offers the solution to other problems: scope marking and long-distance movement are different grammatical devices to achieve the same thing, and in certain configurations one option may indeed block the other as non-optimal; recall, for instance, the discussion of anti-locality (sections 4.2 and 5.3), parametrization (sections 4.3 and 5.2), late *was/mit* (sections 4.5 and 5.4), and weak islands (section 4.6). Evidently, such blocking is possible only if derivations with scope marking (partial *wh*-movement) and derivations with successive-cyclic long-distance *wh*-movement can be competing candidates, that is, can be part of the same reference set. This is the case if reference sets are defined via LF-identity. Hence, to the extent that the results presented here are tenable, the preceding discussion can be viewed as an argument in support of a notion of reference set that is based on the concept of LF output.

Let me conclude by briefly addressing two further issues that are raised by the preceding analysis, one concerning language acquisition, and one additional cross-linguistic variation.

First, the question arises of how constraint rankings such as those in (59) (German) and (74) (Hungarian) can be learned in the course of language acquisition. A possible scenario might look as follows. Suppose that all constraints are unranked in the initial state of the grammar, and that language acquisition implies instantiating a (partial) ranking. I will assume that such a ranking can be achieved by something like Tesar and Smolensky's (1993) Constraint Demotion Algorithm (CDA). Essentially, the CDA takes a wellformed candidate, constructs the reference set of this candidate, compares the violations incurred by the competing

candidates, and concludes that the constraints violated by the winning candidate must be lower-ranked than the constraints violated by the losing candidates. According to the CDA-based approach to acquisition, certain constructions that are prominent in the child's environment serve as triggers for demoting constraints. For example, by being exposed to wellformed *wh*-constructions, the English child can conclude that DER-ECON must be ranked lower than WH-CRIT; and upon entering multiple questions, it can conclude that WH-CRIT must be ranked lower than PROJ-PRIN.

It seems to me that the main bulk of the rankings proposed in this article can be learned along these lines; still, it turns out that a bit more must be said concerning the acquisition of the tie of DER-ECON and FULL-INT proposed for German and Hungarian. Given that German and Hungarian children are exposed to long-distance movement constructions and partial movement constructions, they have evidence for both ranking DER-ECON lower than FULL-INT, and ranking FULL-INT lower than DER-ECON; and this inconsistency could plausibly be taken to be the reason for the child's postulation of a crucial non-ranking of DER-ECON and FULL-INT. Interestingly, however, it seems to be the case that children acquiring English go through a phase in which they permit partial *wh*-movement, in addition to long-distance *wh*-movement (see McDaniel, Chiu and Maxfield 1995; and references cited there). This raises two questions: Why does the English child postulate a tie of DER-ECON and FULL-INT even though it is not exposed to partial *wh*-movement constructions? And why does it eventually abandon this tie even though it might only have negative evidence for doing so?

Towards an (albeit speculative) answer to the first question, one might put to use recent suggestions to the effect that the set of universal constraints is inherently richly structured, that is, that constraints can be grouped together and form "families" (see, for example, Smolensky 1995). As regards the case at hand, DER-ECON and FULL-INT can be viewed as belonging to a family of faithfulness constraints that minimize the effects of "Affect α " (movement, insertion, deletion, etc.). What I would like to suggest on this basis is that what superficially looks like the child's demotion of DER-ECON to a position below WH-CRIT (to account for the existence of simple English questions in the child's input) is actually to be analyzed as a demotion of the whole Affect α family, including FULL-INT. On this view, the internal structure of this constraint family is not yet accessible to the child (which, of course, does not imply that the constraints are not yet active); hence, FULL-INT is demoted even though there is no direct trigger for this demotion — there is a trigger only for the demotion of DER-ECON. This accounts for the alternation of partial and long-distance *wh*-movement in English child language.

Turning to the eventual loss of this alternation in the adult language, the crucial factor must be that the internal structure of the Affect α family becomes accessible to the child. Ideally, we might hope to find a trigger in the child's input that is incompatible with its present grammar (in particular, with the tie of DER-ECON and FULL-INT), and that forces it to resolve the tie by demoting DER-ECON. However,

given that negative evidence (that is, the illformedness of partial *wh*-movement in adult English) cannot per se force the child to alter its grammar, and given that the set of *wh*-constructions permitted by the child grammar is a proper superset of the set of *wh*-constructions permitted by the target grammar, it is far from obvious what this trigger might look like.⁵⁶ It seems to me that a viable alternative could be to correlate the instability of partial *wh*-movement in English with the very fact that it relies on a tie of two constraints. Thus far, I have simply postulated that ties are “marked” options that languages try to avoid; this markedness can now be located in the process of language acquisition. The ultimate rationale behind the CDA is to establish a ranking of as many constraints as possible; furthermore, there seems to be no reason not to assume that in cases where specific ranking of two constraints is irrelevant, ranking is nevertheless established arbitrarily. In line with this, I would like to suggest that language acquisition eventually results in a ranking of as many constraints as possible; this means that ties will automatically be resolved by the child as long as this resolution does not conflict with positive evidence that the child is exposed to. Thus, under present assumptions, as soon as the internal structure of the Affect α family becomes accessible to the child, the tie of DER-ECON and FULL-INT will normally be resolved in such a way as to ensure compatibility with the target grammar (that is, a child acquiring English is exposed only to long-distance *wh*-movement constructions, and resolves the initial tie by ranking FULL-INT higher; a child acquiring a northern German dialect is exposed only to partial *wh*-movement constructions, and accordingly resolves the tie by ranking DER-ECON higher) — however, if there is positive evidence in the child’s input for partial *and* long-distance *wh*-movement, it cannot but maintain the tie; and this is the case in (Standard) German and Hungarian. Thus, although this approach of course requires further empirical verification, for the time being we can conclude that the rankings postulated in this article are in principle learnable on the basis of an unranked initial setting of constraints (grouped together in constraint families), given the CDA plus the auxiliary assumption that non-rankings are eventually resolved by the child whenever possible.

Second, and finally, let me point out that the approach to *wh*-dependencies developed in this article has focused on languages where partial *wh*-movement either occurs together with a scope marker (German, Hungarian), or does not occur at all (English, Korean, Bulgarian). Interestingly, however, there are languages that permit partial *wh*-movement, but can dispense with scope markers. This holds, for example, for Iraqi Arabic (cf. Wahba 1992), Bahasa Indonesia (cf. Saddy 1991), and Ancash Quechua (cf. Cole 1982). Here, scope marker insertion ranges from being optional throughout to being obligatory in certain contexts; such obligatoriness arises, for example, with finite clauses in Iraqi Arabic. A further suggestive

56. Note incidentally that the same problem shows up in a Principles and Parameters approach to the acquisition of *wh*-dependencies; cf. McDaniel, Chiu and Maxwell (1995: 743–744).

property of these languages is that partial *wh*-movement alternates not only with long-distance *wh*-movement, but also with *wh in situ*. The obvious next step would be to try to provide an analysis of partial *wh*-movement constructions in languages like Iraqi Arabic, Bahasa Indonesia, and Ancash Quechua that is based on a re-ranking of the constraints assumed in this article. However, mainly for reasons of space and coherence, I refrain from attempting such a thing here.

Received 15 September 1995

SfS, Universität Tübingen

Revised 25 September 1996

References

- Ackema, Peter and Ad Neeleman (1995). Optimal Questions. Unpublished manuscript, University of Utrecht.
- Aoun, Joseph (1986). *Generalized Binding*. Dordrecht: Foris.
- Aoun, Joseph, Norbert Hornstein and Dominique Sportiche (1981). On some aspects of wide scope interpretation. *Journal of Linguistic Research* 1: 69–95.
- Bayer, Josef (1990). Notes on the ECP in English and German. *Groninger Arbeiten zur Germanistischen Linguistik* 30: 1–51.
- Beck, Sigrid (1996). *Wh*-constructions and transparent Logical Form. Unpublished Ph.D. dissertation, University of Tübingen.
- Beck, Sigrid and Stephen Berman (1996). *Wh*-scope marking: Direct versus indirect dependency. Unpublished manuscript, University of Tübingen.
- Brandner, Ellen (1994). Specifiers and functional projection. Unpublished Ph.D. dissertation, University of Stuttgart.
- Brody, Michael (1995). *Lexico-Logical Form*. Cambridge, MA: MIT Press.
- Cheng, Lisa (1991). On the typology of *wh*-questions. Unpublished Ph.D. dissertation. Massachusetts Institute of Technology, Cambridge.
- Chomsky, Noam (1973). Conditions on transformations. In *A Festschrift for Morris Halle*, Stephen Anderson and Paul Kiparsky (eds.), 232–286. New York: Academic Press.
- (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- (1986a). *Barriers*. Cambridge, MA: MIT Press.
- (1986b). *Knowledge of Language*. New York: Praeger.
- (1993). A minimalist program for linguistic theory. In *The View from Building 20*, Kenneth Hale and Samuel Jay Keyser (eds.), 1–52. Cambridge, MA: MIT Press.
- (1995). Categories and transformations. Chapter 4 of *The Minimalist Program*, 219–394. Cambridge, MA: MIT Press.
- Chomsky, Noam and Howard Lasnik (1977). Filters and control. *Linguistic Inquiry* 8: 425–504.
- (1993). Principles and parameters theory. In *Syntax*, Joachim Jacobs, Wolfgang Sternefeld, Arnim von Stechow and Theo Vennemann (eds.), 506–569. Berlin: Mouton de Gruyter.
- Cinque, Guglielmo (1990). *Types of A'-Dependencies*. Cambridge, MA: MIT Press.
- Cole, Peter (1982). Subjacency and successive cyclicity: Evidence from Ancash Quechua. *Journal of Linguistic Research* 2: 35–58.
- Comorovski, Ileana (1989). Discourse and the syntax of multiple constituent questions. Unpublished Ph.D. dissertation, Cornell University.
- Corver, Norbert (1990). The syntax of left branch extractions. Unpublished Ph.D. dissertation, University of Tilburg.

- Dayal, Veneeta (1994). Scope marking as indirect wh dependency. *Natural Language Semantics* 2: 137–170.
- (1996). *Locality in WH Quantification*. Dordrecht: Kluwer.
- Dikken, Marcel den (1995). Binding, expletives, and levels. *Linguistic Inquiry* 26: 347–354.
- Drubig, H. Bernhard (1992). Zur Frage der grammatischen Repräsentation thetischer und kategorischer Sätze. In *Informationsstruktur und Grammatik*, Joachim Jacobs (ed.), 142–195. Opladen: Westdeutscher Verlag.
- Epstein, Samuel David (1992). Derivational constraints on \bar{A} -chain formation. *Linguistic Inquiry* 23: 235–259.
- Fanselow, Gisbert (1987). *Konfigurationsalität*. Tübingen: Narr.
- Fox, Danny (1994). Quantifier scope in VP ellipsis. Unpublished manuscript, Massachusetts Institute of Technology, Cambridge.
- Grewendorf, Günther (1988). *Aspekte der deutschen Syntax*. Tübingen: Narr.
- Grimshaw, Jane (1995). Projection, heads, and optimality. Unpublished manuscript, Rutgers University. To appear in *Linguistic Inquiry*.
- Haider, Hubert (1986). Affect α : A reply to Lasnik and Saito, 'On the nature of proper government'. *Linguistic Inquiry* 17: 113–126.
- (1993). *Deutsche Syntax, generativ*. Tübingen: Narr.
- Heim, Irene (1992). The problem of pied-piping. Unpublished manuscript, Massachusetts Institute of Technology, Cambridge.
- Hendrick, Randall (1988). *Anaphora in Celtic and Universal Grammar*. Dordrecht: Kluwer.
- Höhle, Tilman (1990). Die w...w-Konstruktion im Deutschen. Unpublished manuscript, University of Tübingen.
- Horvath, Julia (1986). *Focus in the Theory of Grammar and the Syntax of Hungarian*. Dordrecht: Foris.
- (1995). Partial wh-movement and wh-'Scope-Markers'. In *Approaches to Hungarian 5*, István Kenesei (ed.), 91–124. Szeged: Jate.
- Huang, Cheng-Teh James (1982). Logical relations in Chinese and the theory of grammar. Unpublished Ph.D. dissertation. Massachusetts Institute of Technology, Cambridge.
- Karttunen, Lauri (1977). Syntax and semantics of questions. *Linguistics and Philosophy* 1: 3–44.
- Kiss, Katalin (1987). *Configurationality in Hungarian*. Dordrecht: Kluwer.
- (1992). Move-alpha and scrambling in Hungarian. In *Approaches to Hungarian, vol. 4*, István Kenesei and Csaba Pléh (eds.), 67–98. Szeged: Jate.
- Kitahara, Hisatsugu (1993). Deducing 'superiority' effects from the shortest chain requirement. *Harvard Working Papers in Linguistics* 3: 109–119.
- Koster, Jan (1987). *Domains and Dynasties*. Dordrecht: Foris.
- Lasnik, Howard and Mamoru Saito (1984). On the nature of proper government. *Linguistic Inquiry* 15: 235–289.
- (1992). *Move α* . Cambridge, MA: MIT Press.
- Legendre, Géraldine, Paul Smolensky and Colin Wilson (1996). When is less more? Faithfulness and minimal links in wh-chains. Unpublished manuscript, Johns Hopkins University. To appear in *Proceedings of the Workshop on Optimality in Syntax*. Cambridge, MA: MIT Press and MITWPL.
- McCarthy, John and Alan Prince (1995). Faithfulness and reduplicative identity. Unpublished manuscript, University of Massachusetts, Amherst and Rutgers University.
- McCloskey, James (1979). *Transformational Syntax and Model Theoretic Semantics*. Dordrecht: Reidel.
- McDaniel, Dana (1989). Partial and multiple wh-movement. *Natural Language and Linguistic Theory* 7: 565–604.
- McDaniel, Dana, Bonnie Chiu and Thomas Maxfield (1995). Parameters for wh-movement types: Evidence from Child English. *Natural Language and Linguistic Theory* 13: 709–753.
- Mahajan, Anoop (1990). The A/A-bar distinction and movement theory. Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

- Marácz, László (1990). Asymmetries in Hungarian. Unpublished Ph.D. dissertation, University of Groningen.
- May, Robert (1985). *Logical Form*. Cambridge, MA: MIT Press.
- Molxova, Zana (1970). *Xarakter i upotreba na členu bylgarskija i anglijskija ezik*. Sofia: NI.
- Müller, Gereon (1993). On deriving movement type asymmetries. Unpublished Ph.D. dissertation, University of Tübingen.
- (1996). Parallel movement. Unpublished manuscript, University of Tübingen. To appear in *The Pointing Finger*, Joost Dekkers *et al.* (ed.).
- Müller, Gereon and Wolfgang Sternefeld (1993). Improper movement and unambiguous binding. *Linguistic Inquiry* 24: 461–507.
- (1996). A-bar chain formation and economy of derivation. *Linguistic Inquiry* 27: 480–511.
- Pesetsky, David (1994). Optimality principles of sentence pronunciation. Unpublished manuscript, Massachusetts Institute of Technology, Cambridge.
- Pollock, Jean-Yves (1989). Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.
- Prince, Alan and Paul Smolensky (1993). Optimality theory: Constraint interaction in generative grammar. Unpublished manuscript, Rutgers University.
- Reinhart, Tanya (1994). *Wh-in situ* in the framework of the minimalist program. Unpublished manuscript, University of Utrecht.
- Reis, Marga (1996). Extractions from V2 clauses in German? In *On Extraction and Extraposition in German*, Uli Lutz and Jürgen Pafel (eds.), 45–88. Amsterdam: John Benjamins.
- Riemsdijk, Henk van (1982). Correspondence effects and the Empty Category Principle. *Tilburg Papers in Language and Literature* 12. University of Tilburg.
- Rizzi, Luigi (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.
- (1991). Argument/adjunct asymmetries. Unpublished manuscript, University of Geneva.
- Rizzi, Luigi and Ian Roberts (1989). Complex inversion in French. *Probus* 1: 1–30.
- Rudin, Catherine (1988). On multiple questions and multiple *wh* fronting. *Natural Language and Linguistic Theory* 6: 445–501.
- Saddy, Douglas (1991). *Wh*-scope mechanisms in Bahasa Indonesia. In *More Papers on Wh-movement* [= *MIT Working Papers in Linguistics*, 15], Lisa Cheng and Hamida Demirdash (eds.), 182–218. Massachusetts Institute of Technology, Cambridge.
- Saito, Mamoru (1985). Some asymmetries in Japanese and their theoretical implications. Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.
- Smolensky, Paul (1995). On the internal structure of the constraint component *Con* of UG. Unpublished manuscript, John Hopkins University.
- Stechow, Arnim von (1996a). Against LF pied-piping. *Natural Language Semantics* 4: 57–110.
- (1996b). Partial *wh*-movement and Logical Form. Unpublished manuscript, University of Tübingen.
- Stechow, Arnim von and Wolfgang Sternefeld. (1988). *Bausteine syntaktischen Wissens*. Opladen: Westdeutscher Verlag.
- Sternefeld, Wolfgang (1991). *Syntaktische Grenzen*. Opladen: Westdeutscher Verlag.
- (1997). Comparing reference sets. In *The Role of Economy Principles in Linguistic Theory*, Chris Wilder, Hans-Martin Gärtner and Manfred Bierwisch (eds.), 81–114. Berlin: Akademie Verlag.
- Stowell, Tim (1981). Origins of phrase structure. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge.
- Tappe, Thilo (1981). Wer glaubst du hat recht? In *Sprache: Formen und Strukturen*, Manfred Kohrt and Jürgen Lenerz (eds.), 203–212. Tübingen: Niemeyer.
- Tesar, Bruce and Paul Smolensky (1993). The learnability of optimality theory. Unpublished manuscript, University of Colorado, Boulder.
- Vikner, Sten (1995a). *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford: Oxford University Press.

- (1995b). An optimality-theoretical sketch of V2 and V-to-I movement in Germanic. Unpublished manuscript, University of Stuttgart.
- Wahba, Wafaa Abdel-Faheem Batran (1992). LF-movement in Iraqi Arabic. In *Logical Structure and Linguistic Structure*, C.-T. James Huang and Robert May (eds.), 253–276. Dordrecht: Kluwer.
- Zwart, Jan Wouter (1993). Dutch syntax: A minimalist approach. Unpublished Ph.D. dissertation, University of Groningen.