Derivational Optimization of Wh-Movement[†]

Fabian Heck & Gereon Müller Universität Leipzig

1. Introduction

An idea that was originally proposed by Bresnan (1971; 1972) has been resurrected in recent syntactic theory: The PF- and LF-interfaces access syntactic information in a cyclic fashion ('cyclic spell-out'; see, e.g., Groat & O'Neill (1996), Uriagereka (1999), Chomsky (2001), Grohmann (2003), Fox & Pesetsky (2005)). Under this view, each application of cyclic spell-out takes into account only a part of the overall syntactic structure (the current spell-out domain). This can be interpreted as indicating that only part of the overall structure exists at each application of spell-out, i.e., that syntactic structure is built up derivationally.

In this article, we provide both spell-out independent evidence for a derivational view of syntax as well as evidence for a cyclic interaction of syntax and the PF-interface. The evidence suggests that during each of the proper subparts of the derivation only a proper subset of the overall syntactic information is available for the evaluation of syntactic principles. The analysis will be couched in terms of a theory of local optimization that reconciles the idea of a derivational syntax with an optimality-theoretic approach to syntactic repair phenomena. By 'repair' we mean an operation which is usually not legitimate as such but can exceptionally apply to avoid greater damage.[1] In optimality theory, a repair phenomenon is a competition in which the optimal candidate incurs a (normally fatal) violation of a high-ranked constraint C_i in order to respect an even higher-ranked constraint C_j (see Prince & Smolensky (2004)). However, standard global optimization procedures induce complexity of a type that more recent versions of the minimalist program (that do without transderivational constraints) manage to avoid (see Collins (1997) and Frampton & Gutman (1999; 2002), among others). Local optimization combines the two approaches by giving principled accounts of repair phenomena in a way that minimizes complexity.

Empirically, we focus on instances of what we call "repair-driven movement." By this we mean movement operations that are normally impossible in a language, but become possible and, in fact, obligatory if this is the only way to satisfy a high-ranked syntactic constraint.[2]

Let us begin by specifying the main features of the local optimization approach. Following Chomsky (1995; 2000; 2001; 2005), we assume a derivational approach according

[†] The present paper elaborates on and further develops an idea that was first introduced in Heck & Müller (2001). For comments and discussion, we would like to thank Artemis Alexiadou, Judith Aissen, Daniel Büring, Sandra Chung, Gisbert Fanselow, Silke Fischer, Jane Grimshaw, Kleanthes Grohmann, Norbert Hornstein, Shin-Sook Kim, Jason Merchant, Uli Sauerland, Peter Sells, Paul Smolensky, Wolfgang Sternefeld, Bruce Tesar, Sten Vikner, Ralf Vogel, an anonymous reviewer, and the audiences of presentations at Universität Stuttgart, Universität Potsdam, UCLA, UCSC, Rutgers University, Johns Hopkins University, University of Vitoria-Gasteiz, and Universität Leipzig. This research was supported by DFG grants MU 1444/1-1,2-1.

[1] See Speas (1995) and Grimshaw (1997) (both based on Chomsky (1957; 1991)) on optimality-theoretic approaches to *do*-support, and Pesetsky (1998) and Legendre, Smolensky & Wilson (1998) on resumptive pronouns.

[2] Throughout this article, we are only concerned with repair-driven movement that is triggered by constraints that are genuinely syntactic in nature; we leave open the question of whether repair-driven movement can also be triggered by (semantic or prosodic) interface constraints, or whether the pertinent effects are epiphenomena of operations that are motivated syntax-internally.

to which syntactic structure (which basically comprises CP, TP, vP, and VP in the clausal domain) grows incrementally from bottom to top as a result of operations like Merge, Move, and Delete (and, possibly, others). These operations belong to the Gen (generator) part of the grammar, which also contains inviolable constraints, among them the STRICT CYCLE CONDITION (SCC) in (1). The SCC precludes further modification of embedded structures; arguably, a constraint of this type is indispensable in any incremental-derivational system.[3]

(1) STRICT CYCLE CONDITION (SCC):
Within the current cyclic XP α, a syntactic operation may not target a position that is included within another XP β that is dominated by α.

Under a standard, global approach to optimization, we would then expect that there is a single, global H-Eval (harmony evaluation) procedure that affects either the complete derivation, or one (or more) complete representation(s) generated by the derivation.[4] In contrast, we would like to suggest that certain derivational units act as local optimization domains Σ, and that Gen and H-Eval alternate as many times as there are Σs in the derivation. More specifically, suppose that on the basis of a given input, syntactic operations of Gen can apply in accordance with inviolable constraints (like the SCC) in different ways, yielding different outputs at stage Σ_i. These outputs are then subject to optimization with respect to a set of ranked and violable constraints, and the optimal output is determined. Only an optimal output shows up in the input of subsequent derivational steps, together with items taken from the numeration and other optimal outputs. The derivation proceeds in various Gen-compatible ways, producing different outputs at the next optimization domain Σ_{i+1}. Here, optimization starts anew, yielding an optimal output that acts as part of the new input, and so on, until all material of the numeration is used up, the derivation reaches an end, and the optimal root clause is determined. Importantly, all locally suboptimal outputs are disregarded in subsequent derivational steps.[5] We assume that Σ = XP; i.e., every phrase is an optimization domain.[6]

In the following sections, we will argue on the basis of repair-driven movement in German and English that, independently of any complexity issue, local optimization turns out to

[3] The present formulation is based on Chomsky (1973) and Perlmutter & Soames (1979). The SCC is arguably derivable from more basic assumptions; see Chomsky (1995), Watanabe (1995), Bošković & Lasnik (1999), and Freidin (1999) for some recent attempts. Most of the existing attempts to derive the SCC would also be compatible with the main bulk of what follows.

[4] See Pesetsky (1998) and Broekhuis (2000) for the latter option, Müller (1997) for the former, all based on a derivational Gen.

[5] By being iterative, the present approach qualifies as an instance of what Prince & Smolensky (2004) call *harmonic serialism* (as opposed to the standard *harmonic parallelism*, according to which optimization applies only once). The idea of iterated optimization in syntax is also pursued in Heck (2000) and Wilson (2001); for phonology, see McCarthy (2000), Rubach (2000), and the contributions in Hermans & van Oostendorp (2000). However, in all these cases, optimization is global rather than local, in the sense that complete structures are being optimized (repeatedly). Iterated local optimization of the type advanced here is alluded to as a possibility in Archangeli & Langendoen (1997, 214) and Ackema & Neeleman (1998, 478); and, based on the approach that underlies the present article, it is also pursued in Fanselow & Ćavar (2000), Müller (2002), Fischer (2004), Heck (2004), and Heck & Müller (2002).

[6] In Heck & Müller (2005) we argue for a more radical approach where every derivational step is subject to input/output optimization. We believe that the arguments presented here are straightforwardly translatable into that system.

be empirically superior to global optimization in syntax. The cases of repair-driven movement that we discuss involve successive-cyclic *wh*-movement (section 2) and multiple *wh*-movement in sluicing constructions (section 3). The structure of the argument will be similar in the two cases. It takes the form in (i)–(iv).

(i) Movement is triggered by features. For concreteness, suppose that a FEATURE CONDITION (FC) demands that certain designated [∗F∗] features on a lexical item (the probe, see Chomsky (2000; 2001)) must be checked by movement of an item that bears a corresponding [F] feature (the goal).[7]

(2) FEATURE CONDITION (FC):
An [∗F∗] feature on X requires an item bearing [F] at edgeX.

We assume that edgeX is a specifier position of X that has been created by a Move operation; thus, FC can never be satisfied by (pure) Merge. FC is complemented by a constraint that blocks movement which is not feature-driven; following Chomsky (1995), Collins (1997), Kitahara (1997), and many others, we will refer to this constraint as LAST RESORT (LR):

(3) LAST RESORT (LR):
Movement requires matching [F] and [∗F∗] at an edge.

(ii) In certain contexts, it looks as though movement does in fact apply without being feature-driven. Apparently, the movement operation has taken place so as to fulfill another syntactic constraint Con; i.e., it is repair-driven. (iii) This presupposes ranking (Con $\gg$ LAST RESORT) and violability (of LAST RESORT), and thus supports an optimality-theoretic analysis. (iv) Repair-driven movement does not apply in all contexts in which the required constraint ranking would seem to force it. (v) The contexts in which it does not apply even though the constraint ranking as such seems to demand application are those which are filtered out as non-optimal sub-parts of the complete structure under local optimization, but which must be considered under global optimization. Thus, the empirical evidence shows that repair takes place instantaneously, not at some earlier or later stage in the derivation – backtracking and look-ahead are impossible. This suggests that optimization is local, not global.[8]

[7] A few remarks are due on FC. First, the [∗F∗]/[F] notation is taken from Sternefeld (2005); essentially, the asterisk notation ([∗F∗]) encodes the familiar concept of feature strength. Thus, FC implies that movement proceeds to satisfy the needs of the landing site ("Attract"), not the needs of the moved item ("Greed"). Furthermore, note that there is no recourse to specific generalized EPP features as the sole triggers for movement (as in Chomsky (2000; 2001)) – EPP features and contentful features collapse into a single [∗F∗] feature. Third, the question arises of whether the checking operation brought about by FC results in deletion of uninterpretable features. This may well be the case; however, for reasons of perspicuity, all [∗F∗]/[F] features will be retained in representations. Finally, FC is supposed to restrict only overt movement. We will generally remain uncommitted as far as the nature of covert movement is concerned.

[8] A terminological remark is in order here. Pre-theoretically, it makes sense to refer to Con-driven movement that does not check a feature as a "last resort" (= repair) operation. Thus, what we have here is a violation of LAST RESORT (the constraint) as a "last resort" (the meta-grammatical description). Although this terminological clash may be considered somewhat unfortunate, we will stick to the constraint name LAST RESORT to be compatible with the existing literature, and avoid the meta-grammatical use of the notion in favor of the notion "repair."

2. Repair-Driven Successive-Cyclic Wh-Movement

2.1. The Problem

A standard assumption is that simple *wh*-movement is triggered by a feature on C that we will refer to as [∗wh∗]; this feature attracts a *wh*-phrase with a matching [wh] feature, in accordance with LR and FC.[9,10] Under this assumption, we would a priori expect all instances of *wh*-movement to apply in a single step, and not successive-cyclically, via intermediate landing sites. However, there is evidence that *wh*-movement from an embedded CP, as in the English example in (4), applies successive-cyclically, via the embedded SpecC position.[11] Moreover, there is a growing body of evidence to the effect that successive-cyclic *wh*-movement as in (4) in addition uses Specv positions as intermediate landing sites, as originally proposed in Chomsky (1986).[12] How can these intermediate movement steps be reconciled with the assumption that *wh*-movement is driven by a feature like [∗wh∗]?

(4) What$_1$ do [$_{vP}$ t$_1'''$ you think [$_{CP}$ t$_1''$ that [$_{vP}$ t$_1'$ Mary likes t$_1$]]] ?

A common solution is to postulate optional pseudo-[∗wh∗] features on non *wh*-heads that trigger the intermediate movement of a *wh*-phrase to specifiers of v and declarative C (see Collins (1997), Chomsky (2000; 2001), Sabel (1998), and Fanselow & Mahajan (2000), among others). This assumption faces both conceptual and empirical problems. A conceptual problem is that even if we are willing to permit pseudo-[∗wh∗] features on non-*wh*-C heads, there does not seem to be a way to correlate the presence of the [∗wh∗] feature in question with the presence of a *wh*-phrase that "needs" it for further movement; hence, this approach induces a proliferation of derivations that are doomed to fail from the very beginning (e.g., if the relevant feature is present on a C node but there is no *wh*-phrase that might check it, or vice versa). An empirical problem is that given the availability of a feature on declarative C that triggers *wh*-movement, it is unclear what precludes partial *wh*-movement of an embedded *wh*-phrase to the embedded SpecC position in a multiple question (see also Bošković (2002)). This problem is especially pressing in a language like German, which ex-

---

[9] All features that play this role in this article are assumed to be privative rather than binary, but this is mainly to simplify exposition; in the present context, not much hinges on this issue.

[10] On this view, multiple *wh*-movement as in Bulgarian is probably best analyzed as a heterogeneous phenomenon, where the additional *wh*-movement operations are triggered by other features; see Bošković (2000).

[11] Theory-independent evidence for the intermediate movement step is provided by syntactic reflexes in the embedded C domain that can be found in a number of languages. These reflexes include the choice of complementizer in Modern Irish (see McCloskey (1979; 2002)), partial *wh*-movement in Ancash Quechua (see Cole (1982)), Iraqi Arabic (see Wahba (1982)), and German (on the assumption that the *wh*-scope marker *was* is actually the realization of a moved *wh*-feature; see Cheng (2000) and Sabel (2000)), *wh*-copying in German (see Fanselow & Mahajan (2000)), obligatory V raising to C with (certain types of) *wh*-phrases in Spanish (see Torrego (1984) and Baković (1998)) and Basque (see Ortiz de Urbina (1989)), obligatory CP extraposition in German (see Müller (1999), and the selection of subject pronouns by C in Ewe (see Collins (1994))).

[12] A widespread phenomenon that can be taken to support this view is the change of verbal (i.e., v+V) morphology as a result of the *wh*-movement operation, as it can be found, e.g., with tonal down-step in Kikuyu (see Clements, McCloskey, Maling & Zaenen (1983)), *wh*-agreement in Chamorro (see Chung (1994)), *meN* deletion in colloquial Singapore Malay (see Cole & Hermon (2000)), and participle agreement in Passamaquoddy (see Bruening (2001)). Additional evidence for the use of Specv in successive-cyclic movement can be gained on the basis of remnant stranding in Dutch (see Barbiers (2002)), and scope reconstruction in English (see Fox (2000)).

hibits such partial *wh*-movement in the presence of a *wh*-scope marker *was* ('what') instead of a real *wh*-phrase in the matrix clause (we assume that the scope marker is the realization of a moved bare *wh*-feature, i.e., it signals successive-cyclic *wh*-movement in the same way that (4) does; see note 11). Thus, compare (5-a) (successive-cyclic *wh*-movement, as in the English example in (4)), (5-b) (partial *wh*-movement with scope marking, an instance of successive-cyclic *wh*-movement), and (5-c) (*wh*-in situ) with the ill-formed (5-d) (partial *wh*-movement without scope marking).

(5) a. Wen$_1$ hat [$_{vP}$ t$_1'''$ sie gedacht [$_{CP}$ t$_1''$ dass [$_{vP}$ t$_1'$ Maria t$_1$ liebt]]] ?
whom has she thought that Maria loves

b. Was$_1$ hat sie gedacht [$_{CP}$ wen$_1$ Maria t$_1$ liebt] ?
what has she thought whom Maria loves

c. Wer$_1$ hat t$_1$ gedacht [$_{CP}$ – dass Maria wen$_2$ liebt] ?
who has thought that Maria whom loves

d. *Wer$_1$ hat t$_1$ gedacht [$_{CP}$ wen$_2$ Maria t$_2$ liebt] ?
who has thought whom Maria loves

The contrast between legitimate *wh*-movement to a declarative SpecC in (5-a) and (5-b) and illegitimate *wh*-movement to a declarative SpecC in (5-d) is striking. Note that at the stage of the derivation where the embedded TP and declarative C are merged, these two categories are nearly identical in all four cases under consideration.[13] This strongly suggests that it cannot be an internal property of either the embedded TP or the C head that is merged with it that forces or disallows *wh*-movement. The same kind of problem shows up in general form with Specv: If v can freely have a feature that attracts a *wh*-phrase, it is a priori unclear how multiple questions like (6) can be excluded in a language like English:

(6) *Who$_1$ t$_1$ thinks [$_{CP}$ that Mary [$_{vP}$ what$_2$ likes t$_2$]] ?

Chomsky (2000, 109; 2001, 34) suggests that the presence of features on v and C that trigger intermediate movement steps is not completely optional. By assumption, such features are not yet present in the numeration; they can be inserted on v and C heads, but only if they "have an effect on outcome." This solution avoids the conceptual problem of proliferation of unsuccessful derivations, and it also offers a handle on the ill-formedness of examples like (5-d) and (6). However, such an approach presupposes massive look-ahead: Whether or not an effect on outcome is produced by intermediate movement steps can only be verified when the moved XP reaches its target position.

In view of this, we would like to develop an alternative approach to successive-cyclic *wh*-movement according to which an intermediate movement step to Specv or SpecC is not feature-driven movement, but repair-driven movement that incurs a non-fatal violation of LR.[14] This approach arguably captures the same idea that underlies Chomsky's restriction

---

[13] To be sure, the embedded C is not exactly identical in all sentences, being *dass* ('that') in some cases, and phonologically empty in others. However, this difference is irrelevant in the present context: Varieties of German that permit Doubly-Filled Comp Filter configurations can have a *dass* in all four sentences, without any consequence for well-formedness.

[14] Other approaches to successive-cyclic movement which do not rely on features for intermediate steps (like the present approach), but which do not exploit the idea of constraint violability (unlike the present approach), can be found in Chomsky (1993), Takahashi (1994*a*), Bošković (2002), and Boeckx (2003).

on optional feature insertion to contexts where it has an effect on outcome; but it does not involve look-ahead capacity.[15] The analysis we suggest makes crucial use of Chomsky's (1995; 2000; 2001) concept of numeration (and the concept of workspace that is based on it, see below). Given that the numeration is accessible throughout a derivation, we suggest that there is a constraint that relies on information about the current make-up of the numeration, and that may trigger successive-cyclic *wh*-movement in violation of LR.[16]

2.2. Phase Balance and Phase Impenetrability

We propose that the constraint in question is a balance requirement imposed on phases (i.e., the propositional categories vP and CP; see Chomsky (2000; 2001; 2005)), as in (7).

(7) PHASE BALANCE (PB):
For every feature [∗F∗] in the numeration there must be an accessible feature [F] at the phase level.

Accessibility of a feature is defined in (8):

(8) *Accessibility*:
A feature [F] is accessible if (i) or (ii) holds:
(i) [F] is on X or edgeX of the present root of the derivation.
(ii) [F] is part of the workspace of the derivation.

The concept of workspace is based on the notion of a numeration: The workspace of a derivation D comprises the numeration N and material in trees that have been created earlier (with material from N) and have not yet been used in D. Thus, a feature is part of the workspace of a derivation if it is outside the present tree (see Frampton & Gutman (1999) for related discussion). PB forces material from the current phase P containing a feature [F] that will eventually be needed for a probe bearing [∗F∗] in a higher (though as yet non-existent) phase P′ (because of FC) to move to the edge of P, in violation of LR. Given a ranking FC, PB $\gg$ LR, successive-cyclic movement now emerges as a repair strategy.

As an illustration, let us consider the derivation that underlies the German example in (5-a), focusing on optimization of the embedded vP phase first. The competing vP outputs are generated by merging v first with VP, and then with the subject NP, both of which have

[15] An alternative approach that is even closer to Chomsky's system would consist in assuming that [∗F∗] features that are absent on v/C in the numeration can in principle be inserted at phase edges; but such insertion violates a (dependency) faithfulness constraint, call it Dep-[∗F∗]. This constraint would then play the role of LR in the approach to be developed below. Mainly for conceptual reasons (related to the unclear nature of the [∗F∗] features involved), we will not pursue such an approach in what follows, but it should be kept in mind that (at least for the data discussed in this article, but see Heck (2004)) it is empirically equivalent to the approach that we will develop.

[16] Isn't this another form of look-ahead? The question is primarily terminological, since there can be little doubt that this kind of procedure is much more restricted than genuine look-ahead – it utilizes a concept that has been proposed for independent reasons, and it does not have access to structural information provided by later parts of the derivation. Note that if the numeration is present, and accessible, anyway throughout the derivation, the simplest assumption is that the derivation is not blind to its properties – if we want to ensure that the derivation cannot look into the numeration, an additional stipulation to this effect is needed.

emerged as optimal outputs from (trivial) earlier optimization procedures.[17] The competition is shown in table T_1. One possible vP output, O_1, leaves the *wh*-phrase *wen* in situ. This, however, violates PB: For the *wh*-attracting feature [∗wh∗] of the $C_{[*wh*]}$ head that is still part of the numeration (and that will eventually become the matrix head), there is no accessible item that might check it; this is so because (a) there is no other *wh*-phrase left in the workspace, and (b) the *wh*-phrase *wen* is not at the left edge of vP. In contrast, in output O_2, *wh*-movement of *wen* has applied, in violation of LR. This output has the best constraint profile; hence, it qualifies as optimal. Consequently, [wh] of *wen* is now accessible for [∗wh∗] of $C_{[*wh*]}$ in the numeration, vP is balanced, and PB is respected.

T_1: Successive-cyclic movement: local optimization of embedded vP

Input: $[_{\mathrm{VP}}$ V $\mathrm{NP}_{[wh]1}$], $\mathrm{NP}_{[nom]2}$, v Numeration: $\{C_{[*wh*]}, ...\}$	FC	PB	EOC	LR
O_1: $[_{\mathrm{vP}}$ $\mathrm{NP}_{[nom]2}$ $[_{\mathrm{v'}}$ v $[_{\mathrm{VP}}$ V $\mathrm{NP}_{[wh]1}$]]]		*!		
☞O_2: $[_{\mathrm{vP}}$ $\mathrm{NP}_{[wh]1}$ $[_{\mathrm{vP}}$ $\mathrm{NP}_{[nom]2}$ $[_{\mathrm{v'}}$ v $[_{\mathrm{VP}}$ V t_1]]]				*
O_3: Ø			*!	

As shown in T_1, there is another relevant competing vP candidate: O_3, an empty output. Following Prince & Smolensky (2004), we assume that an empty output (the "null parse") is part of all competitions, i.e., it is generated as a possible option throughout. O_3 violates only the EMPTY OUTPUT CONSTRAINT (EOC) in (9), which also belongs to the H-Eval system.

(9) EMPTY OUTPUT CONSTRAINT (EOC):
Avoid an empty output (Ø).

The EOC defines a dividing line, in the sense that higher-ranked constraints in effect become inviolable in optimal candidates. The main task of the EOC is to ensure absolute ungrammaticality ("ineffability") in cases like the one at hand, where there does not appear to be any well-formed output. If the empty output Ø is optimal, the derivation cannot proceed; it crashes (see Chomsky (1995)). However, given a ranking FC, PB $\gg$ EOC $\gg$ LR, the derivation does not crash in T_1; it continues.[18]

Only O_2 in T_1 can show up in the input for the next steps in the derivation. In the next optimization cycle, which determines the optimal TP, nothing happens except for the Merge operation connecting T and vP (raising of the subject NP to SpecT is optional in German, and we can assume that the relevant feature is not present in the numeration; but see below on the situation in English). In particular, there is no repair-driven *wh*-movement because PB is vacuously fulfilled, and low-ranked LR thus excludes any output that involves

[17] The order of the two Merge operations, and their obligatoriness, can be taken to follow from constraints in Gen, with no room for optimization (but cf. Heck & Müller (2005)); but this question is orthogonal to the issue at hand. Furthermore, we will assume that all (pure) Merge operations must apply before all Move operations in a cycle (see Chomsky (1995; 2000; 2001)). Whether this constraint belongs to Gen or is high-ranked in H-Eval may ultimately prove to be an interesting question; but it is immaterial in the present context.

[18] An EOC-based approach to ineffability is also essentially adopted in Ackema & Neeleman (1998) and Wunderlich (2006)); alternative optimality-theoretic approaches (involving concepts like feature neutralization and interface incompatibility) are developed by Legendre, Smolensky & Wilson (1998), Pesetsky (1998), Baković & Keer (2001), and Fanselow & Féry (2002), among others.

unforced movement. Since the following optimization domain affects CP, i.e., a phase, and since the C head that is involved does not bear the feature [$*$wh$*$], the ensuing optimization procedure is in all relevant respects identical to that for vP in T_1: The optimal output CP has repair-driven *wh*-movement, in violation of LR. Optimization of the matrix VP, vP, and TP proceeds as one might expect (with repair-driven *wh*-movement confined to vP, and blocked in VP and TP by LR). Finally, consider optimization of the matrix CP, as shown in table T_2. The optimal CP employs regular, feature-driven *wh*-movement from the matrix Specv position (see O_{22}) rather than leaving the *wh*-phrase in this position (see O_{21}). O_{22} does not violate any of the constraints introduced so far, whereas the suboptimal derivations O_{21} and O_{23} fatally violate the FC and the EOC, respectively. Note that since nothing is left in the numeration at this point, PB is now respected by definition.[19]

T_2: Successive-cyclic movement: local optimization of matrix CP

Input: $[_{\text{TP}}$ ... $[_{\text{vP}}$ $\text{NP}_{[wh]1}$ v ... $[_{\text{CP}}$ t_1'' C ... $[_{\text{vP}}$ t_2' ... t_1 ...]]]], $C_{[*wh*]}$ Numeration: { ... }	FC	PB	EOC	LR
O_{21}: $[_{\text{CP}}$ $C_{[*wh*]}$ $[_{\text{TP}}$ ... $[_{\text{vP}}$ $\text{NP}_{[wh]1}$ v ... $[_{\text{CP}}$ t_1'' C ... $[_{\text{vP}}$ t_2' ... t_1 ...]]]]]	*!			
☞O_{22}: $[_{\text{CP}}$ $\text{NP}_{[wh]1}$ $C_{[*wh*]}$ $[_{\text{TP}}$ ... $[_{\text{vP}}$ t_1''' v ... $[_{\text{CP}}$ t_1'' C ... $[_{\text{vP}}$ t_2' ... t_1 ...]]]]]				
O_{23}: Ø			*!	

To sum up, an intermediate step in successive-cyclic *wh*-movement violates LR in order to provide an accessible item for a [$*$wh$*$]-feature in the numeration and ensure that the embedded phases respect PB.[20,21] The prediction is that no such repair-driven *wh*-movement

[19] The outputs are numbered O_{21}, O_{22}, ... so as to indicate that they are all descendants of O_2 in T_1.

[20] If intermediate and final steps of *wh*-movement are triggered differently, one might expect that reflexes of movement (like those listed in footnotes 11 and 12) are sensitive to this difference. McCloskey (2002, 5) assumes that they are not, and takes this as an argument against an approach like the present one, and in favor of a uniform classification of intermediate and final movement steps as feature-driven. However, it is by no means clear that the reflexes of successive-cyclic movement must be tied to feature-driven operations; they may also be brought about by certain structural configurations (in addition, or alternatively). Furthermore, it turns out that there are in fact reflexes of movement that are sensitive to the difference between intermediate and final movement steps. Thus, a *no* marker is a reflex of *wh*-movement in Duala, but only in the clause in which the moved item finds its final landing site (see Epée (1976)). The opposite situation can also be found: The marker *n-* in Kitharaka exclusively appears on the verbs of intermediate clauses on the path of long *wh*-movement (see Muriungi (2003)). Something similar holds for the complementizer *u-* in Wolof that appears in the context of successive-cyclic movement of overt *wh*-phrases (see Torrence (2005)). Also, obligatory CP extraposition can be viewed as a reflex of *wh*-movement in German, but only in intermediate clauses (see Müller (1999)); and the case is similar in Basque (see Ormazabal, Uriagereka & Uribe-Etxebarria (1994)). Under present assumptions, this implies that the reflex arises only with feature-driven *wh*-movement in Duala, and only with repair-driven *wh*-movement in Kitharaka, Wolof, German, and Basque. Another reflex that is sensitive to this general difference is the change of subject pronouns in Ewe (see Collins (1994)), which is obligatory in the clause containing the final landing site of *wh*-movement, and optional in intermediate clauses crossed by the movement operation. Without going into the details of what a theory of reflexes of movement looks like, it seems fair to conclude that the asymmetry between feature-driven and repair-driven movement that is inherent in the present approach can indeed play a role in this domain.

[21] A remark is due on partial *wh*-movement as in (5-b) in German. Such movement can be assumed to differ minimally from "full" *wh*-movement as in (5-a) in that a repair-driven movement operation to SpecC is fol-

to an intermediate position can be possible if there is another *wh*-phrase left in the workspace at this point that qualifies as accessible for the feature [∗wh∗] on C in the numeration. Consequently, in such a context, movement can only be non-local, across a phase boundary, and is predicted to violate the Phase Impenetrability Condition (PIC) proposed in Chomsky (2000; 2001).[22] Given that the PIC is ranked high, such violations will invariably be fatal.

(10) PHASE IMPENETRABILITY CONDITION (PIC):
The domain of a head X of a phase XP is not accessible to operations outside XP; only X and its edge are accessible to such operations.

In what follows, we will show that the interaction of PB and PIC offers a simple account of two types of *wh*-intervention effects: superiority-effects involving c-command and superiority-like effects that do not involve c-command.

2.3. Wh-Intervention by C-Command

2.3.1. Superiority Effects in English

Typical examples that illustrate superiority effects in English are given in (11), involving a subject and an object *wh*-phrase: Of two *wh*-phrases that are in a c-command relation, only the higher one can move to SpecC.

(11) a. (I wonder) who_1 bought what_2
b. *(I wonder) what_2 who_1 bought t_2

This contrast is often accounted for by assuming a Minimal Link Condition (MLC; see Chomsky (1995; 2000; 2001)), which can be viewed as a feature-based version of the original Superiority Condition (see Chomsky (1973)). However, it is worth noting that standard superiority effects as in (11) follow from the interaction of PB and PIC, without recourse to a constraint like the MLC. The crucial optimization procedure is the one affecting vP; see table T_3.

T_3: Superiority: local optimization of vP

Input: $[_{\text{VP}}$ V $\text{NP}_{[wh]2}$], $\text{NP}_{[nom],[wh]1}$, v Numeration: $\{\text{C}_{[*wh*]}, \text{T}_{[*nom*]}, ...\}$	FC	PB	PIC	EOC	LR
O_1: $[_{\text{vP}}$ $\text{NP}_{[nom],[wh]1}$ $[_{\text{v}'}$ v $[_{\text{VP}}$ V $\text{NP}_{[wh]2}$]]]		*!*			
☞O_2: $[_{\text{vP}}$ $\text{NP}_{[nom],[wh]1}$ $[_{\text{v}'}$ t_1 $[_{\text{v}'}$ v $[_{\text{VP}}$ V $\text{NP}_{[wh]2}$]]]]					*
O_3: $[_{\text{vP}}$ $\text{NP}_{[wh]2}$ $[_{\text{v}'}$ $\text{NP}_{[nom],[wh]1}$ $[_{\text{v}'}$ v $[_{\text{VP}}$ V t_2]]]]		*!			*
O_4: $[_{\text{vP}}$ $\text{NP}_{[nom],[wh]1}$ $[_{\text{v}'}$ $\text{NP}_{[wh]2}$ $[_{\text{v}'}$ t_1 $[_{\text{v}'}$ v $[_{\text{VP}}$ V t_2]]]]]					**!
O_5: Ø				*!	

lowed not by movement of the complete *wh*-phrase, but by movement of the bare [wh]-feature; this option is not available in English. A similar analysis can be given for languages like Bahasa Malay and Kikuyu, where "partial" *wh*-movement occurs without scope markers, if we assume that the bare [wh]-feature does not have to be phonologically realized after movement here. See Cole & Hermon (2000) and Sabel (1998), respectively. Additionally, it must be ensured in cases of partial *wh*-movement that the *wh*-phrase can be stranded in SpecC, but (typically) not in Specv; see Fanselow & Ćavar (2000) for a proposal that could be integrated into the present approach.

[22] Chomsky (2001) discusses two versions of the PIC; we adopt the more restrictive one.

O_1 does not involve any Move operation. Assuming that a category can be in edgeX only if it has been *moved* to a specifier of X, this output fatally violates PB for [∗wh∗] even though the subject NP is merged in Specv. As a matter of fact, it turns out that O_1 violates PB twice. The reason is that there must be a [∗F∗] feature on T that attracts the subject NP, and PB is violated for this feature too if the subject NP is not at the edge of v to provide an accessible [F] counterpart. For the purposes of this paper, we will assume that the feature in question is [∗nom∗], which is matched by [nom] on the subject NP. O_2 involves string-vacuous fronting of the subject *wh*-phrase to an outer Specv position, which satisfies PB for both [∗wh∗] and [∗nom∗], at the cost of a LR violation. However, this output is optimal because all competitors violate higher-ranked constraints. In particular, O_3, in which the object *wh*-phrase is moved to Specv, incurs a fatal violation of PB for [∗nom∗]. If both *wh*-phrases move, as in O_4, LR is violated twice, the second violation being unforced, hence fatal. Finally, the empty output O_5 is blocked by EOC: The derivation does not crash because outputs that respect all higher-ranked constraints are available.

Only O_2 can show up in the input for TP generation; the optimal TP output satisfies all of the above constraints by moving the subject NP to SpecT. Based on this, competing CP outputs are generated. At this point, FC becomes relevant for [∗wh∗], whereas PB is vacuously respected by all candidates (for [∗wh∗]). As shown in table T_4, O_{21} has no *wh*-movement, which fatally violates FC; O_{24} is excluded by EOC, as before; the decision between O_{22} (= (11-a)), with *wh*-movement of the subject, and O_{23} (= (11-b)), with *wh*-movement of the object, is then made by PIC: O_{23} must violate this constraint since NP_2 has not reached the edge of v. Thus, the superiority effect in (11) is derived.

T_4: Superiority: local optimization of CP

Input: $[_{TP}\ NP_{[nom],[wh]1}\ T\ [_{vP}\ t_1'\ [_{v'}\ t_1\ [_{v'}\ v\ [_{VP}\ V\ NP_{[wh]2}\ ]]]]]$, $C_{[*wh*]}$ Numeration: { ... }	FC	PB	PIC	EOC	LR
O_{21}: $[_{CP}\ C_{[*wh*]}\ [_{TP}\ NP_{[nom],[wh]1}\ T\ ...\ [_{vP}\ t_1'\ [_{v'}\ t_1\ [_{v'}\ v\ [_{VP}\ V\ NP_{[wh]2}\ ]]]]]]$	*!				
☞O_{22}: $[_{CP}\ NP_{[nom],[wh]1}\ C_{[*wh*]}\ [_{TP}\ t_1''\ T\ ...\ [_{vP}\ t_1'\ [_{v'}\ t_1\ [_{v'}\ v\ [_{VP}\ V\ NP_{[wh]2}\ ]]]]]]$					
O_{23}: $[_{CP}\ NP_{[wh]2}\ C_{[*wh*]}\ [_{TP}\ NP_{[nom],[wh]1}\ T\ ...\ [_{vP}\ t_1'\ [_{v'}\ t_1\ [_{v'}\ v\ [_{VP}\ V\ t_2\ ]]]]]]$			*!		
O_{24}: Ø				*!	

The superiority effect in (12-a) vs. (12-b) can be accounted for in essentially the same way. Here, two objects are involved that are not clause-mates. The lower *wh*-phrase NP_2 cannot move to Specv during vP optimization since PB can be satisfied without a violation of LR; consequently, any movement applying to NP_2 at a later step of the derivation will have to fatally violate PIC.

(12) a. Who_1 did you persuade t_1 $[_{CP}$ to read $what_2$] ?
b. *$What_2$ did you persuade who_1 $[_{CP}$ to read t_2] ?

2.3.2. Superiority Effects in German

Consider now the situation in German. It has often been noted that German does not exhibit superiority effects with *wh*-phrases that are clause-mates, as in (13-ab); see, e.g., Haider

(1983; 1993; 2004), Grewendorf (1988), and Bayer (1990).

(13) a. (Ich weiß nicht) wer$_1$ C t$_1$ was$_2$ gesagt hat
I know not who$_{nom}$ what$_{acc}$ said has

b. (Ich weiß nicht) was$_2$ C wer$_1$ t$_2$ gesagt hat
I know not what$_{acc}$ who$_{nom}$ said has

Interestingly, German differs from English in that subject raising to SpecT is optional (see den Besten (1981), Diesing (1992), and many others); in present terms, this means that T may or may not bear a feature [∗nom∗].[23] This implies that outputs that are analogous to O_2 and O_3 in T_3 can both be optimal in German if T does not have the feature [∗nom∗]: In this case, PB is satisfied if a [wh] feature shows up at the edge of v, which can be ensured by either subject movement or object movement (but if both NPs move, LR will be fatally violated). At this point, the derivation can continue in two directions, yielding eventually two optimal CP outputs (13-a) and (13-b).[24]

This analysis ties classic superiority effects (as they arise with a *wh*-subject and a *wh*-object) to the obligatoriness of subject raising. Consequently, we expect that German does have superiority effects if subject raising has applied. Unfortunately, evidence for subject raising in German that can be found in the literature is typically not quite as clear as one might wish (see, e.g., Haiders (1993) criticism of Diesing's (1992) arguments based on particle placement); in most cases, subject-initial orders that would seem to argue for subject raising can straightforwardly be derived by subject scrambling. However, there is at least one context where the assumption of subject raising to SpecT seems unavoidable (see Müller (2001)): A solid empirical generalization of German syntax is that nothing can intervene between C and unstressed pronouns, except for subject NPs. Given that German has scrambling of NPs, this may initially look surprising. The evidence can be accounted for if we assume that, whereas scrambling in German can target specifiers of v and V, unstressed pronouns must move to the phonological border of v (in the sense of Chomsky (2001)); i.e., they cannot be preceded by scrambled items. This excludes PPs and object NPs in front of an unstressed pronoun. Consequently, we can conclude that if a subject NP precedes unstressed pronouns but follows C, it must have undergone optional movement to SpecT. And indeed, a clear superiority effect arises in this context; compare (14-ab) (where the *wh*-subject follows the unstressed object pronoun) with (14-cd) (where the *wh*-subject precedes the unstressed object pronoun).[25]

---

[23] A separate issue that we will have nothing new to say about here is how [nom] that is not matched by [∗nom∗], and objective case features on NPs in general (in languages like English and German), can enter Agree relations; recall note 7. An obvious way to proceed here is to formally distinguish probe features with an EPP property ([∗F∗]) from probe features that trigger Agree without movement (this could be rendered as [∗F]); but there are other ways to implement the distinction.

[24] The analysis thus recasts the idea that superiority effects can be voided in languages that exhibit clause-internal movement of wh-objects (which, for the same reason, has been suggested by Grewendorf (1988) to be responsible for the lack of weak cross-over effects in German), see Takahashi (1993), Fanselow (1996), Bošković (1997), Grohmann (1997). However, in the present form, it does not tie the lack of superiority effects in German to the availability of scrambling.

[25] Haider (2004) argues for a similar asymmetry in Icelandic: There is a superiority effect if the *wh*-subject is in SpecT, but not in contexts where it can be assumed to occupy Specv. This effect can be derived in the same way.

(14) a. Wem$_2$ hat [$_{\text{vP}}$ es t$'_2$ wer$_1$ t$_2$ gegeben] ?
whom$_{dat}$ has it$_{acc}$ who$_{nom}$ given

b. Was$_2$ hat [$_{\text{vP}}$ ihm t$'_2$ wer$_1$ t$_2$ gegeben] ?
what$_{acc}$ has him$_{dat}$ who$_{nom}$ given

c. ?*Wem$_2$ hat wer$_1$ [$_{\text{vP}}$ es t$'_2$ t$_1$ t$_2$ gegeben] ?
whom$_{dat}$ has who$_{nom}$ it$_{acc}$ given

d. ?*Was$_2$ hat wer$_1$ [$_{\text{vP}}$ ihm t$'_2$ t$_1$ t$_2$ gegeben] ?
what$_{acc}$ has who$_{nom}$ him$_{dat}$ given

This superiority effect in German has the same source as its English counterpart in (11); the underlying optimization procedures are similar to those in T_3, T_4.[26] Note that the reasoning above presupposes that feature-driven scrambling of $\text{NP}_{[wh]2}$ in (14-c,d) to an outer specifier of vP must be blocked. If such *wh*-scrambling were possible, then the *wh*-phrases in question would be able to reach edgev without incurring an LR violation, which in turn would enable them to move on to SpecC. If one follows Müller (1998), Sauerland (1999), and Grewendorf & Sabel (1999) in assuming that scrambling is triggered by a feature pair $[*\Sigma*]/[\Sigma]$[27], then there are at least two ways to ensure this. One is to say that *wh*-phrases can never bear $[\Sigma]$ in the first place.[28] Another is to assume that $[\Sigma]$-driven scrambling disables a genuine *wh*-phrase from checking a [$*$wh$*$] probe. This alternative assumption is weaker as it principally allows for scrambling of an $\text{NP}_{[wh]}$ (see note 28) as long as this operation is not followed by feature-driven *wh*-movement at some later step. We will make the weaker assumption here.

We also expect a superiority effect to arise in German if the two *wh*-phrases are merged in different phases, as in (12) in English. This prediction is borne out (see Frey (1993), Büring & Hartmann (1994), Pesetsky (2000), and Fanselow (2004), among others).[29] Compare (15-a) (= (5-c)), in which the matrix *wh*-phrase is moved to the matrix $\text{SpecC}_{[*wh*]}$ position, with (15-b), in which the embedded *wh*-phrase undergoes such movement.

(15) a. Wer$_1$ hat t$_1$ gedacht [$_{\text{CP}}$ dass Maria wen$_2$ liebt] ?
who has thought that Maria whom loves

b. *Wen$_2$ hat wer$_1$ gedacht [$_{\text{CP}}$ dass Maria t$_2$ liebt] ?
whom has who thought that Maria loves

---

[26] Note that replacing the *wh*-phrase *wer* in (14-c) and (14-d) with a non-*wh*-subject NP like *der Fritz* ('the Fritz') yields complete well-formedness – in this case, PB forces movement of both the subject NP (for [$*$nom$*$] on T) and the object NP (for [$*$wh$*$] on C) to an outer specifier of v. Note also that replacing the unstressed object pronouns *es* ('it') and *ihm* ('him') with non-pronominal counterparts like *das Buch* ('the book') and *dem Fritz* ('the Fritz'), respectively, produces well-formed versions of (14-c) and (14-d).

[27] While it is fairly uncontroversial that different word orders (induced by scrambling) can have different (preferred) interpretations, it is not clear whether interpretational differences should be encoded by directly assigning an interpretation to $[\Sigma]$, or whether these effects should be made to follow from constraints that determine the mapping from surface structure to LF. We will remain uncommited as far as this question is concerned; all that follows is compatible with either view.

[28] An assumption we actually made in Heck & Müller (2001) (then following Fanselow (1990), Müller & Sternefeld (1993), and Rizzi (1996)). It predicts that *wh*-phrases are never able to undergo feature-driven scrambling. It has recently been argued that this is empirically incorrect, see Fanselow (2001) and references therein.

[29] Also see Takahashi (1993) and Pesetsky (2000) on a similar phenomenon in Japanese.

(15-b) is impossible because NP_2 can never make the first movement step in an optimal embedded vP: At this point, there is a [$*$wh$*$] feature on C and a [wh] feature on NP_1 in the numeration, and vP is balanced without a violation of LR, which therefore becomes fatal.[30]

On a more general note, given that the interaction of PB and PIC derives typical MLC effects in the domain of *wh*-movement, we would like to go one step further and suggest that the MLC can and should be dispensed with as a separate constraint. We would like to contend that there are both conceptual and empirical arguments for this (see Müller (2004) for an explicit discussion of these issues, couched in terms of a theory that assumes an impenetrability condition on phrases instead of phases). A conceptual argument is that dispensing with a constraint that chooses between two a priori possible Move operations in a sufficiently large search space (MLC) in favor of a constraint that permits only Move operations that originate in a small search space (PIC) is clearly preferable given a derivational organization of grammar, in which the search space for operations, i.e., representational residue, should be minimized (see Brody (2002)). The PIC must rely on another constraint to do its work, viz., PB; however, as noted above, PB can be viewed as a fairly straightforward explication of an independently motivated assumption (viz., Chomsky's stipulation concerning optional feature insertion at phase edges). Empirically, the PIC/PB-based approach to superiority effects is superior to an MLC-based approach because it does not need additional assumptions (involving concepts like, e.g., equidistance) to ensure optionality in highly local configurations (i.e., when two *wh*-phrases are merged in the same phase).[31] A further interesting consequence that arises under the PIC/PB-based approach pursued here (but not under an MLC-based approach) is that *wh*-intervention effects are predicted to also arise in contexts where the intervener does not exert c-command.

---

[30] *Wh*-movement from infinitives embedded by restructuring verbs do not give rise to superiority effects in German (see, e.g., Fanselow (1991) and Kim & Sternefeld (1997)), modulo a non-identity requirement on *wh*-phrases (see Haider (2000)). However, infinitives embedded by non-restructuring verbs yield deviance for many speakers in this case:

(i) a. Wer$_1$ hat t$_1$ versucht/gezögert [$_\beta$ dem Fritz was$_2$ zu klauen] ?
who has tried/hesitated the Fritz$_{dat}$ what to steal
b. Was$_2$ hat wer$_1$ versucht/?*gezögert [$_\beta$ dem Fritz t$_2$ zu klauen] ?
what has who tried/hesitated the Fritz$_{dat}$ to steal

One way to capture the restructuring/non-restructuring distinction is in terms of the concept of phase. On this view, the complement of a non-restructuring verb like *zögern* ('hesitate') qualifies as a CP phase; on the other hand, the the complement of a restructuring verb like *versuchen* ('try') is not a CP phase but rather a TP or a vP phase. Under this assumption, and provided that scrambling in German can target VP and vP but not TP or CP, the PIC blocks *wh*-movement in (i-b) in the former case because the lower *wh*-phrase is merged in a different CP phase (whose edge cannot be the target of scrambling), but not in the latter (because the two *wh*-phrases are only separated by a vP phase, whose edge can be targeted by scrambling).

[31] A question arises concerning *wh*-intervention in double object constructions. In German, either one of two *wh*-objects can move. The same goes for two *wh*-objects in prepositional object constructions in English, as in (i-ab) (see Chomsky (1973) and Fiengo (1980)). However, it appears that there is a discernable superiority effect in dative shift constructions in English, as in (i-c) vs. (i-d) (see Barss & Lasnik (1986); note that (i-c) is deviant because of an independent constraint against *wh*-movement of a dative-shifted object in English; see Stowell (1981)).

(i) a. What$_1$ did you give t$_1$ to whom$_2$?

2.4. Wh-Intervention without C-Command

A conspicuous property of the present approach to *wh*-intervention is that the actual hierarchical position of the intervening *wh*-phrase is irrelevant; what is relevant is the fact that it enters the derivation at a later stage. Hence, we expect superiority-like *wh*-intervention effects to arise that block successive-cyclic movement of a given *wh*-phrase $XP_{[wh]1}$ if there is another *wh*-phrase $XP_{[wh]2}$ that is merged outside of $XP_{[wh]1}$'s minimal phase even if $XP_{[wh]2}$ does not end up in a position that c-commands $XP_{[wh]1}$. Furthermore, the prediction is that if $XP_{[wh]2}$ is contained in a phase that enters a tree in which $XP_{[wh]1}$ has already been merged, $XP_{[wh]2}$ cannot move either, and the derivation will have to crash (i.e., the empty output violating only the EOC becomes optimal).[32] These consequences are quite far-reaching. They could in principle be avoided by imposing additional restrictions on the concept of accessibility.[33] However, it turns out that the far-reaching predictions for generalized *wh*-intervention effects are indeed confirmed by the empirical evidence.

2.4.1. Wh-Intervention without C-Command across Phases

Recall that German exhibits superiority effects with *wh*-movement if the intervening *wh*-phrase is merged in a higher phase, or if it is merged in the same phase but raised to SpecT. The prediction is that superiority-like effects will occur if the intervening *wh*-phrase is merged within a category that is in turn merged in a higher phase, or is merged within the same phase but raised to SpecT. This prediction is borne out. Consider first long-distance *wh*-movement. Suppose that a *wh*-phrase has been merged in an embedded clause, and that there is another *wh*-item outside the current tree that eventually ends up in a more deeply embedded position, e.g., in an island. In this case, the lower *wh*-phrase cannot move to

---

b. To whom$_3$ did you give what$_1$ t$_3$?
c. ?Who$_2$ did you give t$_2$ what$_1$?
d. *What$_1$ did you give who$_2$ t$_1$?

As it stands, an MLC-based analysis would seem to predict both (i-b) and (i-d) to be ill formed, whereas the present, PIC/PB-based analysis classifies both examples as well formed. Still, (i-d) can be excluded if we assume that the two objects in a double object construction are merged in a small clause-like, propositional constituent β that qualifies as a phase (see Kayne (1981), among many others); and that dative shift in English (but not in German) involves (Case-driven) movement of the indirect object to a β external specifier. Then, (i-d) receives the same kind of analysis as (11-b), whereas (i-abc) are correctly predicted not to incur a *wh*-intervention effect.

[32] Note in passing that this result holds independently of whether or not we assume that Merge operations in unconnected trees are intrinsically ordered or not. Still, throughout this article we will assume the null hypothesis that there is no extrinsic order of Merge operations (of course, Merge operations are intrinsically ordered, by the SCC, by linking constraints that relate argument structure and syntax, and so forth). Thus, suppose that α is merged with β, yielding δ, and that δ is then merged with a complex category γ. If there is no extrinsic order on Merge, the Merge operations that create γ may take place at any point in the derivation – after Merge applying to α and β, before Merge applying to α and β, or before the Merge operations that have created α or β (if these are complex).

[33] Thus, suppose that requirement (ii) of the definition of accessibility in (8) that plays a role for PB were revised as in in (ii)′.

(ii)′ [F] is on X or edgeX of a root in the workspace of the derivation.

SpecC$_{[*wh*]}$.[34] (16) illustrates this *wh*-intervention effect for long-distance movement with *wh*-in situ inside an adverbial CP.

(16) a. Wen$_1$ hat Fritz [$_{\text{CP}}$ nachdem er was$_2$ gemacht hat] t$_1$ getroffen ?
whom has Fritz after he what done has met

b. *Wen$_1$ hat Fritz [$_{\text{CP}}$ nachdem er was$_2$ gemacht hat] gesagt [$_{\text{CP}}$ dass Maria t$_1$
whom has Fritz after he what done has said that Maria
liebt] ?
loves

(16-a) involves clause-bound *wh*-movement of *wen*$_1$ ('whom'). This case is unproblematic: Given that the adverbial CP is merged in the same phase as *wen*$_1$ (i.e., within vP), the optimal vP is one that respects all higher-ranked constraints (FC, PB, PIC, and EOC) by applying repair-driven *wh*-movement of *wen*$_1$ to Specv, in violation of LR. Things are different in (16-b), though. This sentence is as ill formed as the long-distance superiority example in (15-b), even though the "intervening" *wh*-phrase *was*$_2$ does not c-command the base position of *wen*$_1$.[35] This is unexpected under MLC-based approaches to superiority. In the present approach, however, this intervention effect without c-command follows in more or less the same way as the long-distance superiority effect: Repair-driven successive-cyclic movement of the *wh*-phrase *wen*$_1$ in the lower vP is blocked by LR since PB is fulfilled by the accessibility of *was*$_2$ that is part of the workspace (either because it is still in the numeration, or because it is part of a tree that has already been created but not yet merged in the vP phase that is currently subject to optimization; see note 32). This is shown in table T$_5$.[36]

T$_5$: Wh-Intervention without c-command: local optimization of embedded vP

Input: [$_{\text{VP}}$ V NP$_{[wh]1}$], NP$_{[nom]0}$ v Numeration: $\{$C$_{[*wh*]}$, T$_{[*nom*]}$, N$_{[wh]2}$, ...$\}$	FC PB PIC	EOC	LR
☞O$_1$: [$_{\text{vP}}$ NP$_{[nom]0}$ [$_{\text{v}'}$ t$_0$ [$_{\text{v}'}$ v [$_{\text{VP}}$ V NP$_{[wh]1}$]]]]			*
O$_2$: [$_{\text{vP}}$ NP$_{[nom]0}$ [$_{\text{v}'}$ NP$_{[wh]1}$ [$_{\text{v}'}$ t$_0$ [$_{\text{v}'}$ v [$_{\text{VP}}$ V t$_1$]]]]]			**!
O$_3$: Ø		*!	

Consequently, feature-driven movement of *wen*$_1$ in the matrix CP fatally violates the PIC. As noted, there is a difference to the superiority case: In the superiority context, the higher *wh*-phrase moves to the matrix SpecC$_{[*wh*]}$ position. In the present context, *was*$_2$ cannot move either; compare (16-b) with (17).

(17) *Was$_2$ hat Fritz [$_{\text{CP}}$ nachdem er t$_2$ gemacht hat] gesagt [$_{\text{CP}}$ dass Maria wen$_1$ liebt] ?
what has Fritz after he done has said that Maria whom loves

Given that lexical items are trivial roots, the resulting approach would confine *wh*-intervention effects to c-command environments, and would thereby predict only *wh*-intervention effects that can be subsumed under superiority (i.e., a proper subset of the effects predicted by a strict MLC approach).

[34] Aoun & Li (2003, 20f.) report that this configuration also shows superiority-like effects in Lebanese Arabic.

[35] The example improves drastically if *was*$_2$ is replaced by a non-*wh*-item like *das* ('this').

[36] To simplify the exposition, we assume that *was*$_2$ ('what') is still part of the numeration at this point, but it might just as well be already in a tree in the workspace. In addition, only the most relevant candidates are given here.

Table T_6 shows that the derivation crashes in the very last step, with Ø emerging as the optimal output of the optimization procedure affecting the root CP.

T_6: Wh-Intervention without c-command: local optimization of matrix CP

Input: $[_{\text{TP}}$ .. $[_{\text{CP}}$ .. $\text{NP}_{[wh]2}$ ..] $[_{\text{CP}}$ C .. $\text{NP}_{[wh]1}$ ..]], $\text{C}_{[*wh*]}$ Numeration: { ... }	FC	PB	PIC	EOC	LR
O_1: $[_{\text{CP}}$ – $\text{C}_{[*wh*]}$ $[_{\text{CP}}$ ... $\text{NP}_{[wh]2}$ ...] ... $[_{\text{CP}}$ C ... $\text{NP}_{[wh]1}$ ...]]	*!				
O_2: $[_{\text{CP}}$ $\text{NP}_{[wh]1}$ $\text{C}_{[*wh*]}$ $[_{\text{CP}}$ ... $\text{NP}_{[wh]2}$ ...] ... $[_{\text{CP}}$ C ... t_1 ...]]			*!		
O_3: $[_{\text{CP}}$ $\text{NP}_{[wh]2}$ $\text{C}_{[*wh*]}$ $[_{\text{CP}}$ ... t_2 ...] ... $[_{\text{CP}}$ C ... $\text{NP}_{[wh]1}$ ...]]			*!		
☞O_4: Ø				*	

In O_1, the [∗wh∗] of C is not matched by [wh] at edgeC, which incurs a fatal FC violation. O_2 and O_3 both employ *wh*-movement. In O_2, the embedded *wh*-phrase *wen*$_1$ moves to $\text{SpecC}_{[*wh*]}$, which fatally violates PIC (see (16-b)). In O_3, the *wh*-phrase *was*$_2$ moves from the adverbial CP to $\text{SpecC}_{[*wh*]}$. Again, a fatal violation of PIC is unavoidable (see (17)). Note that the fatal PIC violation incurred by O_3 on the CP cycle cannot be avoided by applying repair-driven movement within the adverbial CP: Just as the presence of *was*$_2$ in the workspace of the part of the derivation creating the minimal phase above *wen*$_1$ blocks repair-driven movement of *wen*$_1$, the presence of *wen*$_1$ in the workspace of the part of the derivation creating the minimal phase above *was*$_1$ blocks movement of *was*$_1$. Hence, O_4 emerges as optimal.[37]

The same type of long-distance intervention effect as in (16) shows up in (18) and (19). Here, the *wh*-in situ item shows up within a complex NP, in a relative clause (see (18)) or in a PP (see (19)).

(18) a. ?Wen$_1$ hat Fritz $[_{\text{NP}}$ einem Mann $[_{\text{CP}}$ der was$_2$ kennt]] t$_1$ vorgestellt ?
whom$_{acc}$ has Fritz a man$_{dat}$ that what knows introduced

b. *Wen$_1$ hat Fritz $[_{\text{NP}}$ einem Mann $[_{\text{CP}}$ der was$_2$ kennt]] gesagt $[_{\text{CP}}$ dass er t$_1$ einladen soll] ?
whom$_{acc}$ has Fritz a man$_{dat}$ that what knows said that he invite should

c. *Was$_2$ hat Fritz $[_{\text{NP}}$ einem Mann $[_{\text{CP}}$ der t$_2$ kennt]] gesagt $[_{\text{CP}}$ dass er wen$_1$ einladen soll] ?
what$_{acc}$ has Fritz a man$_{dat}$ that knows said that he whom$_{acc}$ invite should

---

[37] Gisbert Fanselow (p.c.) points out to us that examples like (16-b) can be improved if the adverbial CP is extraposed to a position that follows the object clause, as in (i).

(i) ?Wen$_1$ hat Fritz gesagt $[_{\text{CP}}$ dass Maria t$_1$ liebt] $[_{\text{CP}}$ nachdem er was$_2$ gemacht hat] ?
whom has Fritz said that Maria loves after he what done has

Assuming with Haider (1994) that an extraposed adverbial CP is merged before an argument CP that precedes it, this fact falls into place. Similar considerations may also apply with respect to related constructions in English; compare Pesetsky's (1995) notion of cascade structure.

(19) a. Wen$_1$ hat Fritz [$_{NP}$ einem Freund von wem$_2$] t$_1$ vorgestellt ?
whom$_{acc}$ has Fritz a friend$_{dat}$ of whom introduced

b. ?*Wen$_1$ hat Fritz [$_{NP}$ einem Freund von wem$_2$] gesagt [$_{CP}$ (t$'_1$) dass Maria t$_1$ liebt] ?
whom$_{acc}$ has Fritz a friend$_{dat}$ of whom said that Maria loves

c. *[$_{PP_2}$ Von wem] hat Fritz [$_{NP}$ einem Freund t$_2$] gesagt [$_{CP}$ dass Maria wen$_1$ liebt] ?
of whom$_{acc}$ has Fritz a friend$_{dat}$ said that Maria whom$_{acc}$ loves

The ill-formedness of (18-b) and (19-b) can be can be accounted for as shown in table T_6. Movement of the intervening *wh*-phrase in (18-c) and (19-c) also leads to ungrammaticality, as before. Whereas this is unproblematic in the case of (18-c), something extra needs to be said about (19-c) (as long as we do not postulate that NPs are phases): As it stands, (19-c) corresponds to a version of O_3 in T_6 that does not violate PIC. However, this output (just like versions of O_3 that underlie (17) and (18-c)) violates the Condition on Extraction Domain (CED), according to which XPs in non-complement positions are barriers for movement (see Huang (1982), Chomsky (1986), and Cinque (1990), among many others); and it must do so fatally if one assumes a ranking CED $\gg$ EOC.

2.4.2. *Wh-Intervention without C-Command across Raised Subjects*

The question arises of whether there is a similar superiority-like effect without c-command in those cases where a *wh*-subject is raised to SpecT (compare the superiority effects in (11-b) in English, (14-c) and (14-d) in German)). Before we take a look at the data, let us see what the analysis developed so far predicts. Thus, suppose that V is first merged with an object NP$_{[wh]1}$, yielding a complex VP; and that v is then merged first with VP, and second with a subject NP$_3$ that is not marked [wh] itself but in turn dominates a PP$_{[wh]2}$. By assumption, T is marked [$*$nom$*$], and C is marked [$*$wh$*$]; both are still in the numeration. PB requires the optimal vP to be balanced, which implies string-vacuous movement of NP$_3$ to an outer Specv position. The question is: Does movement of NP$_3$ containing PP$_{[wh]2}$ suffice to balance vP for [$*$wh$*$] of C (which would preclude movement of NP$_{[wh]1}$ and, given that the subject NP is an island for extraction of PP$_{[wh]2}$ via the CED, lead to a crash of the derivation), or does additional movement of an item bearing [wh] have to occur to an outer Specv position (which would force movement of NP$_{[wh]1}$)? As it turns out, the answer to this question is not entirely obvious, due to a vagueness in the definition of accessibility in (8). At the point where the legitimacy of moving NP$_{[wh]1}$ to Specv is to be decided, [wh] on PP$_2$ is certainly not part of the workspace anymore; so [wh] of PP$_2$ does not qualify as accessible to [$*$wh$*$] of C in the numeration via clause (ii) of (8). The other possibility is that [wh] of PP$_2$ within NP$_3$ is “on edgeX of the present root of the derivation”, thereby satisfying clause (i) of (8) and blocking movement of NP$_{[wh]1}$. Note that “for a feature [F] to be on edgeX” can be understood in various ways: It may mean that [F] is a feature on the category that occupies an edge position of X; or that [F] is a feature on a category that is (reflexively) dominated by a category that occupies an edge position of X.

The choice between these two options can be made on the basis of empirical evidence; and even though the data situation may not be crystal-clear, it suggests that the second, more liberal interpretation of "on edgeX" is the correct one: There does indeed appear to be a discernible *wh*-intervention effect without c-command in subject raising contexts. Consider first the situation in English. In (20-a), $NP_{[wh]2}$ is embedded in a complex object NP_3, and $NP_{[wh]1}$ is a subject. $NP_{[wh]2}$ can never move (it is caught within its minimal vP phase because [wh] on NP_1 is accessible in the workspace for [$*$wh$*$] of C in the numeration), whereas $NP_{[wh]1}$ can and must move. However, in (20-b), $NP_{[wh]2}$ is embedded in a complex subject NP that is raised to SpecT, whereas $NP_{[wh]1}$ is an object. Provided that [wh] of NP_2 counts as accessible to [$*$wh$*$] of C in the numeration when the subject NP_3 is moved to an outer Specv (because of [$*$nom$*$] on T), it succeeds in blocking movement of $NP_{[wh]1}$ on the vP cycle. Since $NP_{[wh]2}$ cannot move itself either, Ø is correctly predicted to be the optimal vP (i.e., the derivation crashes); see (20-c).[38]

(20) a. Who$_1$ t$_1$ saw [$_{NP_3}$ the man that bought what$_2$] ?
b.?*Who$_1$ did [$_{NP_3}$ the man that bought what$_2$] see t$_1$?
c. *What$_2$ did [$_{NP_3}$ the man that bought t$_2$] see who$_1$?

We expect the same kind of superiority-like *wh*-intervention effect to arise in subject raising constructions in German, i.e., in constructions where *wh*-movement takes place across a subject NP that dominates a *wh*-phrase and precedes an unstressed object pronoun. The contrast between (21-a) (where the subject NP can be in situ) and (21-b) (where the subject NP is raised to SpecT) may not be one of perfect well-formedness vs. absolute ungrammaticality; but the tendency is clear enough, and conforms to expectations. Again, (21-c) is impossible.[39]

(21) a. ?Wem$_1$ hat [$_{NP_3}$ die Frau [$_{PP_2}$ mit welchem Mantel]] t$_1$ ein Buch gegeben ?
whom$_{dat}$ has the woman with which coat a book$_{acc}$ given
b.?*Wem$_1$ hat [$_{NP_3}$ die Frau [$_{PP_2}$ mit welchem Mantel]] es t$_1$ gegeben ?
whom$_{dat}$ has the woman with which coat it$_{acc}$ given

---

[38] The data in (20) (and several others of the same basic type) were checked with various native speakers, who unanimously provided the judgements given here. In particular, the contrast between examples of the type in (20-a) and those of the type in (20-b) was perceived as striking throughout. As far as we can tell, data such as (20-b) do not figure prominently in the literature (e.g., the studies of *wh*-in situ in Chomsky (1981), Huang (1982; 1995), Lasnik & Saito (1992), and Hornstein (1995) do not seem to mention the construction); but they should be expected to be well formed in most approaches that we are aware of. There is, however, an exception in the literature: Date of the type in (20-b) are are discussed in Fiengo et al. (1988) and, following them, in Fitzpatrick (2002), and judged grammatical. We have nothing to say here about the source of the diverging judgements, except for the observation that Fiengo et al. (1988) are primarily concerned with contrasting a construction of the type in (20-b) with a construction of the type in (20-c). It seems that (20-c) is indeed typically perceived as even worse than (20-b). Note in this context that since the CP output (20-b) (fatally) violates PIC, whereas the CP output in (20-c) (fatally) violates both PIC and CED, the constraint profile of (20-c) is clearly worse; and this fact might prove important in approaches that set out to reconcile optimality with graded (un-)grammaticality judgements (see Keller (2000)). For the time being, we will draw the dividing line between (20-a) on the one side and (20-b)/(20-c) on the other.

[39] We have to leave open the question as to why examples like (21-a) are in fact slightly deviant, as indicated; one might speculate that heavy subject NPs tend to leave the vP domain. Similarly for (21-b) vs. (21-c); see the last footnote.

c. *[$_{PP_2}$ Mit welchem Mantel] hat [$_{NP_3}$ die Frau t_2] es wem$_1$ gegeben ?
with which coat has the woman it$_{acc}$ whom$_{dat}$ given

To sum up sections 2.3 and 2.4: We have argued that superiority effects where an intervening *wh*-phrase c-commands a second *wh*-phrase, and superiority-like effects where an intervening *wh*-phrase does not c-command a second *wh*-phrase, can be given a uniform analysis that centers around the concept of repair-driven movement, based on the interaction of PB and PIC, and the violability of LR. However, there is a crucial theoretical difference between the two kinds of *wh*-intervention: In the first case, the intervening *wh*-phrase can itself move; in the second case, no *wh*-phrase can move, and ungrammaticality arises (the optimal output violates EOC non-fatally). It is this latter property that allows us to compare the local approach to optimization pursued here with a standard, global approach to optimization, and to develop an argument against the latter.

2.5. Wh-Intervention and Local vs. Global Optimization

In a local optimization procedure like the one sketched in T_6, there is no output that involves successive-cyclic *wh*-movement of the lower *wh*-phrase. Such an output is not available at this late stage of the derivation because it would either have to be a descendant of an output that is filtered out earlier in the derivation, due to a fatal LR violation (as shown in T_5), or its generation would imply a violation of the SCC. However, in the global approach, there is nothing that would preclude such a candidate from participating in and winning the competition. This would undermine the account of *wh*-intervention effects without c-command. In a nutshell, the problem with global optimization here is that an early LR violation (a LR violation in a low position) can eventually pay off later (in a higher position) when otherwise a higher-ranked constraint (the EOC) must be violated (note that the ranking EOC $\gg$ LR is unavoidable to permit successive-cyclic movement in the first place). This is shown in table T_7, in which complete structures (which may be derivations or representations) are subjected to a standard optimization procedure; here and henceforth, a wrong choice of optimal output in the global approach is signalled by ★.

T_7: Long-distance intervention: global optimization

Input: C, $C_{[*wh*]}$, $NP_{[wh]1}$, $NP_{[wh]2}$, ... Numeration: { ... }	FC	PB	PIC	EOC	LR
O_1: [$_{CP}$ $C_{[*wh*]}$ [$_{CP}$ ... $NP_{[wh]2}$ ...] ... [$_{CP}$ C ... $NP_{[wh]1}$ ...]]	*!				
O_2: [$_{CP}$ $NP_{[wh]2}$ $C_{[*wh*]}$ [$_{CP}$ ... t_2 ...] ... [$_{CP}$ C ... $NP_{[wh]1}$ ...]]			*!		
O_3: [$_{CP}$ wh_1 $C_{[*wh*]}$ [$_{CP}$ ... $NP_{[wh]2}$ ...] ... [$_{CP}$ C ... t_1 ...]]			*!		
★O_4: [$_{CP}$ wh_1 $C_{[*wh*]}$ [$_{CP}$ ... $NP_{[wh]2}$ ...]] ... [$_{CP}$ t_1' C ... t_1 ...]]					*
O_5: Ø				*!	

All other things being equal, the global approach thus makes the wrong prediction simply because it takes into account information from the complete structure. The empirical evidence, however, suggests that optimization remains caught in a local optimum, due to the fact that there is only a reduced amount of information available.

2.6. Relativized Wh-Intervention

In all the examples discussed so far, there was only one [∗wh∗] feature present in the initial numeration; we have not yet considered cases in which there is more than one [∗wh∗] feature. A consideration of these cases shows that the present approach is slightly too restrictive as it stands. To see this, consider an English example like (22).

(22) [$_{CP}$ Who$_1$ C$_{[*wh*]}$ t$_1$ asked whom$_2$ [$_{CP}$ what$_3$ C$_{[*wh*]}$ John [$_{vP}$ t$'_3$ likes t$_3$]]] ?

The well-formedness of (22) is unexpected under present assumptions: How can movement of NP$_{[wh]3}$ to Specv be forced in the embedded vP? At the point where it has to be decided whether NP$_{[wh]3}$ moves or not, there are two [∗wh∗] features on two C items in the numeration that need accessible [wh] features; and the two [∗wh∗] features do in fact find two [wh] features in the workspace (more specifically, in the numeration), on the two remaining *wh*-items N$_{[wh]1}$ and N$_{[wh]2}$. Consequently, *wh*-movement of NP$_{[wh]3}$ to the embedded Specv should be blocked, the optimal embedded CP should be Ø, and the derivation should crash early, yielding ill-formedness of (22), contrary to fact. The same problem shows up in the German examples in (23), which are both fully well formed.

(23) a. [$_{CP}$ Wer$_1$ C$_{[*wh*]}$-hat t$_1$ wen$_2$ gefragt [$_{CP}$ was$_3$ C$_{[*wh*]}$ [$_{vP}$ t$'_3$ Fritz t$_3$ mag]] ?
who$_{nom}$ has whom$_{acc}$ asked what$_{acc}$ Fritz$_{nom}$ likes

b. [$_{NP}$ Die Frage [$_{CP}$ wer$_1$ C$_{[*wh*]}$ t$_1$ was$_2$ mitbringt]]] ist relevant für die Frage [$_{CP}$ wie$_3$ Fritz denkt [$_{CP}$ t$''_3$ dass [$_{vP}$ t$'_3$ die Party t$_3$ wird]]]
the question who what brings is relevant to the question how Fritz thinks that the party will be

The difference between the unwanted instances of *wh*-intervention in (22) and (23) and the genuine instances of *wh*-intervention as they were discussed in the preceding sections is that in the former cases (but not in the latter), the intervening *wh*-item does not compete for the same SpecC$_{[*wh*]}$ position that the first-merged *wh*-phrase whose movement it blocks is supposed to end up in. Thus, the approach must be modified in such a way that *wh*-intervention is relativized with respect to the SpecC$_{[*wh*]}$ target position: An item bearing [wh] can block movement of another item bearing [wh] via PB only if the two items target the same [∗wh∗] C head. To execute this idea, suppose that [wh] features are accompanied by scope indices in the numeration. Then, an item that bears the feature ^{i}[wh] does not count as accessible for a contra-indexed feature ^{j}[∗wh∗] on some C, due to a simple feature mismatch. Hence, repair-driven movement of NP$_{^{i}[wh]3}$ to the embedded Specv position is the only way to fulfill PB for C$_{^{i}[*wh*]}$ in the numeration during vP optimization in (22) and (23): The remaining two *wh*-items N$_{^{j}[wh]1}$ and N$_{^{j}[wh]2}$ are not accessible for C$_{^{i}[*wh*]}$; they are accessible only for C$_{^{j}[*wh*]}$ (which then implies a possible superiority effect if N$_{^{j}[wh]1}$ is merged in Specv and subsequently undergoes subject raising to SpecT as NP$_{^{j}[wh]1}$).

Examples like those in (22) and (23) do not pose comparable problems in standard, MLC-based approaches to *wh*-intervention. Hence, at first sight it seems as though the additional assumption motivated by these data (viz., scope indices that accompany *wh*-features) might provide an argument against the present approach. However, such a conclusion would be premature. Closer inspection reveals that the same kind of problem shows up for MLC-based approaches in slightly different constructions like the one in (24).

(24) Who$_1$ C$_{[*wh*]}$ wonders [$_{\text{CP}}$ what$_3$ C$_{[*wh*]}$ who$_2$ bought t$_3$]

As noted by Lasnik & Saito (1992), (24) is ill formed (on a par with (11-b)) only if *who*$_2$ takes embedded scope; but it is well formed if *who*$_2$ takes matrix scope. This is a long-standing problem for MLC-based approaches because it is by no means clear how it can be ensured that a constraint like the MLC "should choose between two *wh*-phrases only when they are to wind up in the same Spec of CP at LF", as Lasnik & Saito (1992, 120) put it. In view of this problem, approaches that set out to reconcile the superiority effect with embedded scope of *who*$_2$ with the absence of such an effect with matrix scope of *who*$_2$ in cases like (24) either rely on the assumption that the constraint is not checked locally in the derivation, but globally, by taking into account the semantic interpretation of the whole sentence (see Kitahara (1993) and Grimshaw (1994) for suggestions of this kind); or they dispense with an MLC-type constraint in toto (see Lasnik & Saito (1992), who suggest a constraint on *wh*-absorption that also presupposes global look-ahead). In contrast, the present approach in terms of scope indices on *wh*-features accounts for (24) straightforwardly, and in a strictly local manner: *who*$_2$ does not block movement of *what*$_3$ if the two *wh*-items have different scope indices, but *who*$_2$ blocks movement of *what*$_3$ via PB if the indices are identical. Thus, it turns out that the solution to what initially looks like a problem that is confined to the present approach actually provides the solution to a long-standing problem with standard approaches to *wh*-intervention.

3. Repair-Driven Multiple Wh-Movement

3.1. The Ban on Multiple Wh-Movement in German

Only one *wh*-phrase moves overtly in German multiple questions; compare (25-a) with (25-b):[40]

(25) a. Wer$_1$ hat t$_1$ wen$_2$ getroffen ?
who has whom met
b. *Wer$_1$ wen$_2$ hat t$_1$ t$_2$ getroffen ?
who whom has met

As noted in section 2.1, this follows from LR if C$_{[*wh*]}$ can attract only one *wh*-phrase, and the features that are responsible for multiple *wh*-movement in languages like Bulgarian (see note 10) are absent in German. However, the preceding section has shown that *wh*-phrases can violate LR if this is the only possibility to satisfy a higher-ranked constraint – PB, in the case at hand. Thus far, we have been concerned with the intermediate steps in successive-cyclic *wh*-movement. However, violability of LR also opens up the possibility of repair-driven multiple *wh*-movement. In this section, we argue that German exhibits multiple overt *wh*-movement in a context where this is required by a higher-ranked constraint, viz., in

[40] Grewendorf (2001) argues for overt multiple *wh*-movement, accompanied by PF realization of exactly one *wh*-phrase in SpecC. On this view, the phenomenon to be discussed below would involve repair-driven multiple *wh*-realization instead of repair-driven multiple *wh*-movement.

multiple sluicing constructions. Before we address this issue, let us introduce an analysis of simple sluicing.

3.2. Simple Sluicing

(26) is a simple German sluicing construction of the type discussed in Ross (1969). In sluicing constructions, parts of an embedded *wh*-question are deleted (licensed by appropriate antecedent material in the matrix clause), with only the embedded *wh*-phrase retained (which has the same grammatical function as a quantified XP of the matrix clause if it is an argument).[41]

(26) Der Fritz hat irgendwen gesehen, aber ich weiß nicht [$_{CP}$ wen$_1$ C [$_{TP}$ ~~der Fritz~~ t$_1$
the Fritz has someone seen but I know not whom the Fritz
~~gesehen hat~~]]
seen has

Sluicing has been analyzed in terms of TP deletion at PF by Lasnik (1999) and Merchant (2001), and in terms of TP insertion at LF by Chung, Ladusaw, & McCloskey (1995). We will adopt a version of the deletion approach here that shares with Lasnik's and Merchant's analyses the basic tenet that only complete constituents can be deleted but deviates from these approaches in relying on the concept of phase.[42] For concreteness, suppose that C can optionally bear a deletion feature [Δ] and that a [Δ]-marked phase must be deleted. Chomsky (2001) proposes that PF spell-out proceeds cyclically, in a phase-by-phase manner; however, given that the top domain of a phase must in principle remain accessible for further operations, PF spell-out must be restricted to the non-edge domain of a phase. We would like to suggest that PF deletion in sluicing constructions operates in essentially this way: All material of a [Δ]-marked phase is deleted at PF as soon as the optimal phase output is determined, except for those items in the accessible domain (edge or head) of a phase that are required in their position by PB. In other words: For each item in a phase, it can be determined whether its presence in the position that it shows up in is required by a probe feature outside the current derivation or not; and if it is, the item must not undergo spell-out of the phase (see Müller (2006) for an elaboration of such a PB-based system of cyclic spell-out).

An immediate consequence of this approach is that it follows without further ado that the only items that can survive PF deletion in sluicing constructions are *wh*-phrases moved to SpecC (see Merchant's (2001, 62) "Sluicing-Comp-Generalization"). Thus, varieties of German that do not respect the Doubly-Filled Comp Filter, as in (27-a), still do not permit

[41] Material that is marked for deletion is crossed out here and in what follows.

[42] Ross's (1969) original approach permits discontinuous deletion. English examples like *John went to London with someone, but I don't know who with* seem to call for discontinuous deletion in sluicing, given that *who* precedes the preposition that selects it. An alternative that has been pursued in the literature is to treat such cases in terms of *wh*-pied piping plus internal *wh*-raising, which is compatible with the idea that only constituents can be deleted (see Merchant (2002) for such an analysis in terms of PF-movement). Culicover (1999) observes that both the idiosyncratic properties of internal *wh*-raising and its very existence are initially unexpected. It seems to us that both issues could be properly addressed in an optimality-theoretic approach. Similar considerations apply with respect to Ross's (1969) observation that (simple) sluicing can ameliorate island effects; see Merchant (2001) and Fox & Lasnik (2001) for recent approaches.

sluicing constructions like (27-b).[43]

(27) a. Der Fritz hat irgendwen gesehen, aber ich weiß nicht [$_{CP}$ wen$_1$ dass [$_{TP}$ er t$_1$
the Fritz has someone seen but I know not whom that he
gesehen hat]]
seen has

b. *Der Fritz hat irgendwen gesehen, aber ich weiß nicht [$_{CP}$ wen$_1$ dass [$_{TP}$ ~~der~~
the Fritz has someone seen but I know not whom that the
~~Fritz~~ t$_1$ ~~gesehen~~ ~~hat~~]]
Fritz seen has

In contrast, approaches that rely on TP deletion must resort to additional assumptions to ensure that C can never be filled in sluicing constructions, whether it is filled by Merge or by Move (i.e., by V/2); see Merchant (2001, ch.2).

With this approach to sluicing as background, we can note that simple sluicing as in (26) does not yet involve repair-driven movement of a new kind – the *wh*-phrase moves to the embedded SpecC$_{[*wh*]}$ position because of FC. The case is different with sluicing in multiple questions in German.

3.3. Multiple Sluicing

If the embedded *wh*-clause is a multiple question in a sluicing construction, something interesting happens: Although German normally does not exhibit multiple overt *wh*-movement, it seems that such multiple *wh*-movement becomes possible and, in fact, obligatory in cases of multiple sluicing. Compare (28-a), which is well formed, with (28-b), which is not a possible realization of an embedded multiple question:[44]

(28) a. Irgendjemand hat irgendetwas geerbt, aber der Fritz weiß nicht mehr
someone has something inherited but the Fritz knows not more
[$_{CP}$ wer$_1$ was$_2$ ~~C$_{[\Delta]}$ t$_1$ t$_2$~~ ~~geerbt~~ ~~hat~~]
who what inherited has

b. *Irgendjemand hat irgendetwas geerbt, aber der Fritz weiß nicht mehr
someone has something inherited but the Fritz knows not more
[$_{CP}$ wer$_1$ ~~C$_{[\Delta]}$ t$_1$~~ ~~was$_2$~~ ~~geerbt~~ ~~hat~~]
who what inherited has

---

[43] More generally, it follows in the present approach that a phase head π will undergo cyclic spell-out only if there is no other lexical item in the workspace that has a probe feature which requires it to be accessible, via PB; if there is such a lexical item, π will not be able to take part in spell-out. This correctly distinguishes between, say, V+v in cases of v-to-T(-to-C) movement (which is required in position by a probe feature on T, and thus does not participate in vP spell-out), and V+v+T+C in verb-second contexts, which typically does not undergo further movement (at least in finite contexts), and which therefore is predicted to undergo spell-out on the CP level. The latter consequence accounts for illegitimate matrix sluicing that maintains the verb in C, as in German **Wen schlug*? (‘Whom hitted’) vs. *Wen* (‘whom’) (context: I know that she hit someone; the question is: –).

[44] The PF-string in (28-b) is well formed as such – it can be interpreted as a simple embedded question, with an interpretation ‘I don’t know who inherited something’ instead of the multiple question interpretation ‘I don’t know who inherited what.’ A derivation with this interpretation does not have a [wh] feature on the embedded object. Hence, such a derivation does not compete with the derivations that generate the sentences in (28).

Obligatory multiple *wh*-movement in (28-a) strongly suggests an analysis in terms of repair-driven movement in violation of LR.[45] The question then is, what is the nature of the higher-ranked constraint that forces the LR violation? *Wh*-phrases differ from other categories in that their content can never be recovered after deletion (see Pesetsky (1998)). Accordingly, we suggest that the constraint in question is (29), which ensures the recoverability of *wh*-phrases from [Δ]-marked phases.

(29) WH-RECOVERABILITY (WH-R):
A *wh*-phrase must be at the edge of a [Δ]-marked phase.

Thus, the *wh*-phrases in (28-a) must first move to the edge of v and from there to the edge of the [Δ]-marked C. If they did not, WH-R would be fatally violated during CP optimization.

Before entering the optimization process, consider the following. There are two cases: either subject raising will take place on the TP cycle, or not. If there is subject raising, then it has to be prepared by PB-driven movement to an outer Specv on the vP cycle (due to an [$*$nom$*$] probe in the numeration; PB with respect to [$*$wh$*$] is satisfied as a free rider). This will render the subject $\mathrm{NP}_{[wh]2}$ accessible and thus enable it to move to edgeC later (in order to satisfy both FC and WH-R). But then the object $\mathrm{NP}_{[wh]1}$ cannot undergo repair-driven movement to an outer Specv. PB is already satisfied by $\mathrm{NP}_{[wh]2}$ and WH-R is irrelevant on the vP cycle (the [Δ] feature will enter the derivation on the C-head). Unmotivated movement of $\mathrm{NP}_{[wh]1}$ would fatally violate LR. $\mathrm{NP}_{[wh]1}$ must therefore reach the edge of v by feature-driven scrambling: $\mathrm{NP}_{[wh]1}$ must bear [Σ] and v must bear [$*\Sigma*$]. If there is no subject raising on the TP cycle, one of $\mathrm{NP}_{[wh]1}$ and $\mathrm{NP}_{[wh]2}$ must reach the edge of v via [Σ]-driven scrambling (because PB can only justify one LR violation by moving a *wh*-phrase to the edge of vP). Note, however, that if one of the *wh*-phrases undergoes [Σ]-driven scrambling, then the derivation will not succeed, because [Σ]-driven scrambling bleeds repair-driven scrambling (one *wh*-phrase in Specv is enough to satisfy PB). Thus, one of the *wh*-phrases is stranded within the inaccessible domain of vP, causing either a fatal WH-R or PIC violation on the CP cycle. (Bleeding is not an issue if subject raising takes place, because only the subject can bear [nom] for independent reasons).

The local optimization procedures that underly the data in (28) are shown in table T_8 and table T_9. (We only consider the successful derivation with subject raising, for convenience.) First consider the vP cycle.

As discussed above, both $\mathrm{NP}_{[wh]1}$ and $\mathrm{NP}_{[wh]2}$ in T_8 move to the edge of vP, albeit for different reasons: one to satisfy PB, thereby violating LR, the other to satisfy FC (see O_4). And as both *wh*-phrases have moved to outer specifiers of vP in T_8, they are both accessible on the CP cycle in T_9.

After subject raising to SpecT, $\mathrm{NP}_{[wh]2}$ will move to SpecC in order to satisfy FC: there is a [$*$wh$*$] probe on C.[46] As PB is trivially fulfilled, it cannot motivate movement of $\mathrm{NP}_{[wh]1}$. Rather, this movement is exclusively driven by WH-R, because C introduces the [Δ]-feature (WH-R, of course, is also a trigger for movement of $\mathrm{NP}_{[wh]2}$). This derives overt multiple

[45] Merchant (2001, 112) independently sketches an analysis along the same lines.

[46] Recall that the object $\mathrm{NP}_{[wh]1}$ cannot check [$*$wh$*$] on C, because it has already undergone [Σ]-driven scrambling. This assumption was a consequence of the superiority facts mentioned in section 2.3.2.

T_8: Multiple sluicing and clause bound wh-movement: local optimization of vP

Input: $[_{\text{VP}}\ \text{NP}_{[wh],[\Sigma]1}\ \text{V}\ ]$, $\text{NP}_{[wh]2}$, $\text{v}_{[*\Sigma*]}$ Numeration: $\{\text{C}_{[*wh*],[\Delta]}, \text{T}_{[*nom*]} ...\}$	FC	WH-R	PB	EOC	LR
O_1: $[_{\text{vP}}\ \text{NP}_{[wh],[nom]2}\ [_{\text{v}'}\ [_{\text{VP}}\ \text{NP}_{[wh],[\Sigma]1}\ \text{V}\ ]\ \text{v}_{[*\Sigma*]}\ ]]$	*(!)		*(!)*		
O_2: $[_{\text{vP}}\ \text{NP}_{[wh],[nom]2}$ ... $[_{\text{v}'}\ \text{t}_2\ [_{\text{v}'}\ [_{\text{VP}}\ \text{NP}_{[wh],[\Sigma]1}\ \text{V}\ ]\ \text{v}_{[*\Sigma*]}\ ]]]$	*!				*
O_3: $[_{\text{vP}}\ \text{NP}_{[wh],[\Sigma]1}$ ... $[_{\text{v}'}\ \text{NP}_{[wh],[nom]2}\ [_{\text{v}'}\ [_{\text{VP}}\ \text{t}_1\ \text{V}\ ]\ \text{v}_{[*\Sigma*]}\ ]]$			*!		
☞O_4: $[_{\text{vP}}\ \text{NP}_{[wh],[\Sigma]1}\ [_{\text{v}'}\ \text{NP}_{[wh],[nom]2}$ ... $[_{\text{v}'}\ \text{t}_2\ [_{\text{v}'}\ [_{\text{VP}}\ \text{t}_1\ \text{V}\ ]\ \text{v}_{[*\Sigma*]}\ ]]]]$					*
O_5: Ø				*!	

T_9: Multiple sluicing and clause bound wh-movement: local optimization of CP

Input: $[_{\text{TP}}\ \text{NP}_{[wh]2}\ [_{\text{vP}}\ \text{NP}_{[wh]1}\ \text{t}'_2\ [_{\text{v}'}\ \text{t}_2 ... \text{t}_1 ...\ ]]$, $\text{C}_{[\Delta],[*wh*]}$ Numeration: {...}	FC	WH-R	PB	EOC	LR
O_{41}: $[_{\text{CP}}\ \text{C}_{[\Delta],[*wh*]}\ [_{\text{TP}}\ \text{NP}_{[wh]2}$ ... $[_{\text{vP}}\ \text{NP}_{[wh]1}\ \text{t}'_2\ [_{\text{v}'}\ \text{t}_2 ... \text{t}_1 ...\ ]]]]$	*(!)	*(!)*			
O_{42}: $[_{\text{CP}}\ \text{NP}_{[wh]2}\ \text{C}_{[\Delta],[*wh*]}\ [_{\text{TP}}\ \text{t}''_2$ ... $[_{\text{vP}}\ \text{NP}_{[wh]1}\ \text{t}'_2\ [_{\text{v}'}\ \text{t}_2 ... \text{t}_1 ...\ ]]]]$		*!			
☞O_{43}: $[_{\text{CP}}\ \text{NP}_{[wh]2}\ \text{NP}_{[wh]1}\ \text{C}_{[\Delta],[*wh*]}\ [_{\text{TP}}\ \text{t}''_2$ ... $[_{\text{vP}}\ \text{t}'_1\ \text{t}'_2\ [_{\text{v}'}\ \text{t}_2 ... \text{t}_1 ...\ ]]]]$					*
O_{44}: Ø				*!	

wh-movement.[47]

Repair-driven successive-cyclic *wh*-movement differs substantially from repair-driven multiple *wh*-movement.[48] Still, the roles played by PB in the first case, and by WH-R in

---

[47] Nothing is said about the order in which the *wh*-phrases appear on edgeC; thus, in the example at hand both orders are predicted to be well-formed and no superiority-like effect is expected, as it can be observed in languages with regular overt multiple *wh*-movement like Bulgarian (see Rudin (1988), Pesetsky (2000), among others, ignoring that we assumed that not all *wh*-phrases in Bulgarian target edgeC to begin with, see note 10). However, only the order in (28-a) is possible:

(i) *Irgendjemand hat irgendetwas geerbt, aber der Fritz weiß nicht mehr was_2 wer_1
someone has something inherited but the Fritz knows not more what who

As we analyze superiority effects as an interaction of PB and PIC, the ungrammaticality of (i) cannot be due to superiority: Both *wh*-phrases are accessible at the relevant points. We therefore assume that the ill-formedness of (i) is due to some parallelism requirement that holds between the order of the *wh*-phrases in the sluiced clause on the one hand and the order of the antecedents in the superordinate clause on the other hand (similar to the parallelism requirements that can be observed in other ellipsis environments, see Fox 2000). This view is corroborated by the fact that a similar (though slightly weaker) effect can be observed with two object *wh*-phrases, a case for which superiority does not arise in German if no sluicing is involved (see note 31):

(ii) a. Der Fritz hat jedem irgendetwas gegeben, aber ich weiß nicht mehr wem was
the Fritz has everyone something given but I know not more whom what
b. ?*Der Fritz hat jedem irgendetwas gegeben, aber ich weiß nicht mehr was wem
the Fritz has everyone something given but I know not more what whom

[48] In fact, an alternative would be to assume that multiple sluicing involves *wh*-scrambling to TP rather than *wh*-movement to SpecC. We see at least one problem with such an approach, though: There is good evidence that scrambling in German cannot target TP. Recall the generalization from section 2.3.2. that in German full

the second, are similar. Accordingly, it does not come as a surprise that an argument for local (as opposed to global) optimization can be given on the basis of sluicing constructions that resembles the argument given in the previous section on the basis of long-distance intervention effects.

3.4. Multiple Sluicing and Local vs. Global Optimization

In cases of multiple sluicing, movement of the second *wh*-phrase across a sentence boundary is impossible; the two *wh*-phrases have to be clause-mates (see Takahashi (1994*b*) on Japanese, Sauerland (1999b) on German). This is initially surprising since a *wh*-phrase can move across a sentence boundary in simple sluicing constructions. As we will see, the facts fall out directly under the present system of local optimization, but must remain a mystery under global optimization.

(30) is a well-formed instance of simple long-distance sluicing in German:

(30) Maria hat behauptet dass sie irgendetwas geerbt hat aber Fritz weiß nicht mehr
Maria has claimed that she something inherited has but Fritz knows not more
[$_{CP}$ was$_1$ ~~C$_{[\Delta]}$ Maria t$'''_1$ behauptet hat~~ [$_{CP}$ ~~t$''_1$ dass sie t$'_1$ t$_1$ geerbt hat~~]]
what Maria claimed has that she inherited has

This construction involves repair-driven *wh*-movement. However, the LR violation is not forced by WH-R (in the final step); rather, it is forced by PB (in the intermediary steps), just like the cases of successive-cyclic *wh*-movement that were discussed in section 3. The optimization procedure affecting the most deeply embedded CP is shown in table T_{10}. (Here and in what follows, intermediate steps that target the edge of vP are not explicitely discussed, but only indicated by traces; note that this does not affect the argument.)

---

object NPs cannot precede unstressed pronouns within TP, only subject NPs can. This observation received a straightforward explanation by assuming that SpecT, the position optionally targeted by subject NPs, is above unstressed pronouns, whereas scrambling in German exclusively targets positions below such pronouns (and hence below TP).

On the other hand, multiple sluicing cannot involve bare vP deletion, which would make an analysis of multiple sluicing in terms of repair-driven *wh*-scrambling to vP possible (in agreement with the observations about unstressed pronouns). If multiple sluicing could involve bare vP deletion in German, then a subject NP in SpecT should survive deletion. This, however, is not possible:

(i) *Jeder hat irgendjemandem geholfen, aber es ist unklar, wem der Fritz ~~geholfen hat~~
everyone has someone helped but it is unclear whom the Fritz helped has

Sauerland (1999b) advances an argument for the scrambling approach to multiple sluicing in German: Restructuring verbs permit scrambling and *wh*-movement from an infinitive, non-restructuring verbs permit only *wh*-movement. It turns out that multiple sluicing is much more acceptable with restructuring verbs; see (i).

(ii) Irgendjemand hat versucht/?*gezögert irgendetwas zu klauen aber ich weiß nicht [$_{CP}$ wer$_1$ was$_2$ t$_1$
someone has tried/hesitated something to steal but I know not who what
~~versucht/gezögert hat~~ [$_\beta$ t$_2$ ~~zu klauen~~]]
tried/hesitated has to steal

However, this also follows under the *wh*-movement approach to multiple sluicing adopted here. As remarked in note 30, evidence from superiority effects points to a CP phase status of infinitival complements of non-restructuring verbs, and a TP or a vP phase status of infinitives embedded by restructuring verbs. In what follows, we will see that PB and the PIC systematically predict *wh*-movement across a CP phase in multiple sluicing constructions to be impossible.

T_{10}: Simple sluicing, local optimization of embedded CP

Input: [$_{\text{TP}}$ Maria was$_1$ geerbt hat], C Numeration: {$C_{[*wh*],[\Delta]}$, ... }	FC	WH-R	PB	LR
O_1: [$_{\text{CP}_5}$ – C [$_{\text{TP}}$ ... [$_{\text{vP}}$ *wh*$_1$ ... t$_1$ ...]]]			*!	
☞O_2: [$_{\text{CP}_5}$ *wh*$_1$ C [$_{\text{TP}}$ ... [$_{\text{vP}}$ t$'_1$ ... t$_1$ ...]]]]				*

O_2 is then the sole input for subsequent optimization. The next CP phase involves merging a [Δ]-marked $C_{[*wh*]}$. As shown in table T_{11}, the optimal embedded question has *wh*-movement to Spec$C_{[*wh*]}$. This movement is in accordance with WH-R, but it is not repair-driven because it is independently required by the FC, and therefore does not violate LR.

T_{11}: Simple sluicing, local optimization of matrix CP

Input: [$_{\text{TP}}$ M. [$_{\text{vP}}$ was$_1$ behauptet hat ... [$_{\text{CP}_5}$ t$''_1$ M. [$_{\text{vP}}$ t$'_1$ t$_1$ geerbt hat]]], $C_{[*wh*],[\Delta]}$ Numeration: { ... }	FC	WH-R	PB	LR
O_{21}: [$_{\text{CP}_7}$ – ~~$C_{[*wh*],[\Delta]}$ [$_{\text{TP}}$ ... [$_{\text{vP}}$ *wh*$_1$ ...~~ ~~[$_{\text{CP}_5}$ t$''_1$ C [$_{\text{TP}}$ ... t$'_1$ ... t$_1$ ...]]]]~~]	*!	*		
☞O_{22}: [$_{\text{CP}_7}$ *wh*$_1$ ~~$C_{[*wh*],[\Delta]}$ [$_{\text{TP}}$ ... [$_{\text{vP}}$ t$'''_1$ ...~~ ~~[$_{\text{CP}_5}$ t$''_1$ C [$_{\text{TP}}$ ... t$'_1$ ... t$_1$ ...]]]]~~]				

Consider now the case of multiple long-distance sluicing in (31):

(31) *Irgendjemand hat behauptet, dass Maria irgendetwas geerbt hat, aber Fritz weiß
someone has claimed that Maria something inherited has but Fritz knows
nicht mehr [$_{\text{CP}}$ wer$_1$ was$_2$ ~~$C_{[\Delta]}$ t$_1$ behauptet hat~~ [$_{\text{CP}}$ ~~dass Maria t$_2$ geerbt hat~~]]
not more who what claimed has that Maria inherited has

In (31), the two *wh*-phrases are not clause-mates, and the result is ill formed. The present approach accounts for this as follows. First, optimization of the most deeply embedded CP ensures that the lower *wh*-phrase *was*$_2$ must stay in situ (modulo [Σ]-driven scrambling on the vP cycle, an option that is not available in the TP and CP domains) – PB can be fulfilled without repair-driven *wh*-movement (with the other *wh*-phrase *wer*$_1$ still part of the numeration), and the LR violation incurred by movement of *was*$_2$ is therefore fatal; see table T_{12}.

T_{12}: Multiple long-distance sluicing, local optimization of embedded CP

Input: [$_{\text{TP}}$ M. was$_2$ geerbt hat], C Numeration: {$C_{[*wh*]}$, *wh*$_1$, ... }	FC	WH-R	PB	EOC	LR
☞O_1: [$_{\text{CP}_5}$ – C [$_{\text{TP}}$ ... *wh*$_2$ ...]]					
O_2: [$_{\text{CP}_5}$ *wh*$_2$ C [$_{\text{TP}}$ ... t$_2$ ...]]					*!
O_3: Ø				*!	

The optimization of the next CP phase involves outputs that are generated by merging TP (a descendant of O_1) with $C_{[*wh*],[\Delta]}$. The situation is now similar to that in the well-formed multiple question (15-a). In the latter case, the *wh*-phrase in the lower clause stays in situ, and the *wh*-phrase in the higher clause moves to Spec$C_{[*wh*]}$. However, there is one important difference in the present case: Leaving the lower *wh*-phrase in situ during matrix CP optimization implies a violation of WH-R, and given the ranking WH-R $\gg$ EOC, the empty output will win. Since (31) is in fact ungrammatical, this is the right result. Thus, the effect here is comparable to that shown in table T_6: Just as moving the higher *wh*-phrase induces a fatal violation of the CED in T_6, leaving the lower *wh*-phrase in situ induces a fatal violation of WH-R; ineffability results in both cases. The competition is sketched in table T_{13}.[49]

T_{13}: Multiple long-distance sluicing, local optimization of matrix CP

Input: [$_{\text{TP}}$ wer$_1$ behauptet hat ... [$_{\text{CP}_5}$ C [$_{\text{TP}}$ M. was$_2$ geerbt hat]]], C$_{[*wh*],[\Delta]}$ Numeration: { ... }	FC	WH-R	PIC	EOC	LR
O$_{11}$: [$_{\text{CP}_7}$ – ~~C$_{[*wh*],[\Delta]}$ [$_{\text{TP}}$ wh_1 ... [$_{\text{CP}_5}$ – C [$_{\text{TP}}$ ... wh_2 ...]]]~~]	*!	**			
O$_{12}$: [$_{\text{CP}_7}$ wh_1 ~~C$_{[*wh*],[\Delta]}$ [$_{\text{TP}}$ t$_1$ ... [$_{\text{CP}_5}$ – C [$_{\text{TP}}$ ... wh_2 ...]]]~~]		*!			
O$_{13}$: [$_{\text{CP}_7}$ wh_1 wh_2 ~~C$_{[*wh*],[\Delta]}$ [$_{\text{TP}}$ t$_1$ ... [$_{\text{CP}_5}$ – C [$_{\text{TP}}$ ... t$_2$ ...]]]~~]			*!		*
O$_{14}$: [$_{\text{CP}_7}$ wh_2 ~~C$_{[*wh*],[\Delta]}$ [$_{\text{TP}}$ wh_1 ... [$_{\text{CP}_5}$ – C [$_{\text{TP}}$ ... t$_2$ ...]]]~~]		*!	*		
☞O$_{15}$: Ø				*	

If both *wh*-phrases stay in situ, as in O_{11}, the FC is fatally violated, and so is WH-R (twice). In O_{12}, the higher *wh*-phrase moves to SpecC$_{[*wh*]}$, but this leaves the lower *wh*-phrase in the [Δ]-marked TP domain, in fatal violation of WH-R. Evidently, moving only the lower *wh*-phrase, as in O_{14}, can only make things worse, because the WH-R violation is now accompanied by a PIC violation. Finally, moving both *wh*-phrases also fatally violates the PIC; see O_{13}. Hence, the empty output's EOC violation emerges as non-fatal, and ungrammaticality is derived.[50]

Suppose now that we were to adopt a global optimization approach to repair-driven *wh*-movement in sluicing constructions. This would accommodate the well-formed (i.e., clause-bound) cases of multiple *wh*-movement without problems. However, as with long-distance intervention effects (see T_7), the global approach overgenerates: A variant of O_{13} in T_{13} that establishes an intermediate trace (and thereby fulfills the PIC) cannot be excluded. This candidate incurs two LR violations (one by embedded *wh*-movement, and one by subsequent multiple *wh*-movement to SpecC$_{[*wh*]}$), but manages to avoid a violation of a higher-ranked constraint. This wrong outcome, which would render (31) well formed, is sketched in table T_{14}.

Again, for the local approach this is not an issue, because the information that accepting

[49] Note that PB is ignored here for space reasons, whereas the PIC becomes relevant again and is accordingly re-introduced into the table.

[50] One might think that a [Δ]-marking on the embedded and on the matrix C head plus [Σ]-driven scrambling to the intermediary vP edges could induce repair-driven *wh*-movement in both the embedded and the matrix CP, and thereby help to avoid a fatal PIC violation with multiple *wh*-movement in T_{13}. However, this strategy is not available since it would imply successive-cyclic deletion, and hence, deletion of a non-constituent after the first step.

T_{14}: Multiple long-distance sluicing, global optimization

Input: C, $C_{[*wh*]}$, $NP_{[wh]1}$, $NP_{[wh]2}$, T, $T_{[\Delta]}$, ... Numeration: { ... }	FC	WH-R	PIC	EOC	LR
O_{11}: $[_{CP_7}$ – $C_{[*wh*]}$ ~~$[_{TP_{[\Delta]}}$ wh_1 ... $[_{CP_5}$ – C $[_{TP}$ ... wh_2 ...]]]~~]	*!	**			
O_{12}: $[_{CP_7}$ wh_1 $C_{[*wh*]}$ ~~$[_{TP_{[\Delta]}}$ t_1 ... $[_{CP_5}$ – C $[_{TP}$ ... wh_2 ...]]]~~]		*!			
O_{13}: $[_{CP_7}$ wh_1 wh_2 $C_{[*wh*]}$ ~~$[_{TP_{[\Delta]}}$ t_1 ... $[_{CP_5}$ – C $[_{TP}$ ... t_2 ...]]]~~]			*!		*
O_{14}: $[_{CP_7}$ wh_2 $C_{[*wh*]}$ ~~$[_{TP_{[\Delta]}}$ wh_1 ... $[_{CP_5}$ – C $[_{TP}$ ... t_2 ...]]]~~]		*!	*		
O_{15}: Ø				*	
★O_{16}: $[_{CP_7}$ wh_1 wh_2 $C_{[*wh*]}$ ~~$[_{TP_{[\Delta]}}$ t_1 ... $[_{CP_5}$ t'_2 C $[_{TP}$ ... t_2 ...]]]~~]					**

an additional LR violation on an embedded cycle will pay off on a higher cycle is simply not available. The derivation thus remains in a local optimum.

4. Conclusion

In this article, we provided evidence for a derivational view of syntax. The arguments were based on both purely syntactic phenomena from the domain of *wh*-movement (successive-cyclic *wh*-movement, superiority) on the one hand, as well as on the interaction of *wh*-movement with the PF-interface (PF spell-out in sluicing and multiple sluicing constructions) on the other. The evidence suggested that at different sub-stages of the derivation only a proper subset of the overall syntactic information is available. Moreover, we opted for a system that involves violable ranked constraints. This enabled us to express the idea that there exists non-feature driven *wh*-movement, which can apply in violation of LR in contexts where this is required by higher ranked constraints.

References

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

Archangeli, Diana, & D. Terence Langendoen. 1997. Afterword. In *Optimality Theory. An Overview*, ed. Diana Archangeli & D. Terence Langendoen, 200-215. Oxford: Blackwell.

Aoun, Joseph & Yen-Hui Audrey Li. 2003. *The Diversity of Wh-Constructions*. Cambridge, Mass.: MIT Press.

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

Baković, Eric & Ed Keer. 2001. Optionality and Ineffability. In *Optimality-Theoretic Syntax*, ed. Géraldine Legendre, Jane Grimshaw & Sten Vikner, 97-112. Cambridge, Mass.: MIT Press.

Barbiers, Sjef. 2002. Remnant Stranding and the Theory of Movement. In *Dimensions of Movement*, ed. Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers & Hans-Martin Gärtner, 47-67. Amsterdam: Benjamins.

Barss, Andrew, & Howard Lasnik. 1986. A Note on Anaphora and Double Objects. *Linguistic Inquiry* 17:347-354.
Bayer, Josef. 1990. Notes on the ECP in English and German. *Groninger Arbeiten zur Germanistischen Linguistik* 30:1-51.
Besten, Hans den. 1981. Government, syntaktische Struktur und Kasus. In *Sprache: Formen und Strukturen*, ed. Manfred Kohrt & Jürgen Lenerz, 97-107. Tübingen: Niemeyer.
Boeckx, Cedric. 2003. *Islands and Chains*. Amsterdam: John Benjamins.
Bošković, Željko. 1997. On certain violations of the superiority condition, AgrO, and economy of derivation. *Journal of Linguistics* 33:227-254.
Bošković, Željko. 2000. What is Special about Multiple Wh-Fronting? In *Proceedings of NELS 30*, ed. M. Hirotani, A. Coetzee, N. Hall & J.-Y. Kim, 83-107. Amherst, Mass.: GLSA.
Bošković, Željko. 2002. A-Movement and the EPP. *Syntax* 5:167-218.
Bošković, Željko, & Howard Lasnik. 1999. How Strict is the Cycle? *Linguistic Inquiry* 20:691-703.
Bresnan, Joan. 1971. Sentence Stress and Syntactic Transformations. *Language* 47:257-281.
Bresnan, Joan. 1972. Stress and Syntax: a Reply. *Language*, 48:326-342.
Brody, Michael. 2002. On the Status of Representations and Derivations. In *Derivation and Explanation in the Minimalist Program*, ed. Samuel David Epstein & T. Daniel Seely, 19-41. Oxford: Blackwell.
Broekhuis, Hans. 2000. Against Feature Strength: The Case of Scandinavian Object Shift. *Natural Language and Linguistic Theory* 18:673-721.
Bruening, Benjamin. 2001. *Syntax at the Edge: Cross-Clausal Phenomena and the Syntax of Passamaquoddy*. Doctoral dissertation, MIT, Cambridge, Mass.
Büring, Daniel, & Katharina Hartmann. 1994. The Dark Side of Wh-Movement. *Linguistische Berichte* 149:56-74.
Cheng, Lisa. 2000. Moving Just the Feature. In *Wh-Scope Marking*, ed. Uli Lutz, Gereon Müller, & Arnim von Stechow, 77-99. Amsterdam: Benjamins.
Chomsky, Noam. 1973. Conditions on Transformations. In *A Festschrift for Morris Halle*, ed. Stephen Anderson & Paul Kiparsky, 232-286. New York: Academic Press.
Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
Chomsky, Noam. 1986. *Barriers*. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory. In *The View from Building 20*, eds. Kenneth Hale & Samuel Jay Keyser, 1-52. MIT, Cambridge, Mass.
Chomsky, Noam. 1995. Categories and Transformations. Chapter 4 of *The Minimalist Program*, 219-394. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In *Step by Step*, ed. Roger Martin, David Michaels & Juan Uriagereka, 89-155. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 2001. Derivation by Phase. In *Ken Hale. A Life in Language*, ed. Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 2005. On Phases. Ms., MIT.
Chung, Sandra. 1994. Wh-Agreement and 'Referentiality' in Chamorro. *Linguistic Inquiry* 25:1-44.
Chung, Sandra, William Ladusaw, & James McCloskey. 1995. Sluicing and Logical Form. *Natural Language Semantics* 3:239-282.
Cinque, Guglielmo. 1990. *Types of A′-Dependencies*. Cambridge, Mass.: MIT Press.

Clements, George, James McCloskey, Joan Maling, & Annie Zaenen. 1983. String-Vacuous Rule Application. *Linguistic Inquiry* 14:1-17.

Cole, Peter. 1982. Subjacency and Successive Cyclicity: Evidence from Ancash Quechua. *Journal of Linguistic Research* 2:35-58.

Cole, Peter, & Gabriella Hermon. 2000. Partial Wh-Movement: Evidence from Malay. In *Wh-Scope Marking*, ed. Uli Lutz, Gereon Müller, & Arnim von Stechow, 101-130. Amsterdam: Benjamins.

Collins, Chris. 1994. Economy of Derivation and the Generalized Proper Binding Condition. *Linguistic Inquiry* 25:45-61.

Collins, Chris. 1997. *Local Economy*. Cambridge, Mass.: MIT Press.

Culicover, Peter. 1999. *Syntactic Nuts*. Oxford: Oxford University Press.

Epée, Roger. 1976. On Some Rules that are Not Successive Cyclic in Duala. *Linguistic Inquiry* 7:193-198.

Fanselow, Gisbert. 1990. Scrambling as NP-movement. In *Scrambling and Barriers*, ed. Günther Grewendorf & Wolfgang Sternefeld, 113-140. Amsterdam: Benjamins.

Fanselow, Gisbert. 1991. *Minimale Syntax*. Habilitation thesis, Universität Passau.

Fanselow, Gisbert. 1996. The Proper Interpretation of the Minimal Link Condition. Ms., Universität Potsdam.

Fanselow, Gisbert. 2001. Features, Theta-Roles, and Free Constituent Order. *Linguistic Inquiry* 32:405-437.

Fanselow, Gisbert. 2004. The MLC and Interface Economy. In *Minimality Effects in Syntax*, ed. Arthur Stepanov, Gisbert Fanselow, & Ralf Vogel, 73-124. Berlin: Mouton de Gruyter.

Fanselow, Gisbert, & Damir Ćavar. 2000. Remarks on the Economy of Pronunciation. In *Competition in Syntax*, ed. Gereon Müller & Wolfgang Sternefeld, 107-150. Berlin: Mouton de Gruyter.

Fanselow, Gisbert, & Anoop Mahajan. 2000. Towards a Minimalist Theory of Wh-Expletives, Wh-Copying, and Successive Cyclicity. In *Wh-Scope Marking*, ed. Uli Lutz, Gereon Müller, & Arnim von Stechow, 195-230. Amsterdam: Benjamins.

Fanselow, Gisbert & Caroline Féry. 2002. Ineffability in Grammar. In *Resolving Conflicts in Grammars*, ed. Gisbert Fanselow & Caroline Féry, 265-307. Hamburg: Buske.

Fiengo, Robert. 1980. *Surface Structure*. Harvard: Harvard University Press.

Fiengo, Robert, James Huang, Howard Lasnik, & Tanya Reinhart. 1988. The Syntax of Wh-in-situ. In *Proceedings of WCCFL 7*, ed. Hagit Borer, 81-98. Stanford, Calif.: CSLI Publications.

Fischer, Silke. 2004. *Towards an Optimal Theory of Reflexivization.* Doctoral dissertation, Universität Tübingen.

Fitzpatrick, Justin. 2002. On Minimalist Approaches to the Locality of Movement. *Linguistic Inquiry* 33:443-463.

Fox, Danny. 2000. *Economy and Semantic Interpretation*. Cambridge, Mass.: MIT Press.

Fox, Danny, & Howard Lasnik. 2003. Successive Cyclic Movement and Island Repair: The Difference between Sluicing and VP Ellipsis. *Linguistic Inquiry* 34:143-154.

Fox, Danny & David Pesetsky. 2005. Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31:1-46.

Frampton, John, & Sam Gutman. 1999. Cyclic Computation. *Syntax* 2:1-27.

Framton, John, & Sam Gutman. 2002. Crash-Proofed Syntax. In: *Derivation and Explana-*

tion in the Minimalist Program, ed. Samuel David Epstein & T. Daniel Seely, 90-105. Oxford: Blackwell.
Freidin, Robert. 1999. Cyclicity and Minimalism. In *Working Minimalism*, ed. Samuel David Epstein & Norbert Hornstein, 95-126. Cambridge, Mass.: MIT Press.
Frey, Werner. 1993. *Syntaktische Bedingungen für die Interpretation*. Berlin: Akademieverlag.
Grewendorf, Günther. 1988. *Aspekte der deutschen Syntax*. Tübingen: Narr.
Grewendorf, Günther. 2001. Multiple Wh-Fronting. *Linguistic Inquiry* 32:87-122.
Grewendorf, Günther, & Joachim Sabel. 1999. Scrambling in German and Japanese. *Natural Language and Linguistic Theory* 17:1-65.
Grimshaw, Jane. 1994. Heads and Optimality. Handout of talk, Universität Stuttgart.
Grimshaw, Jane. 1997. Projection, Heads, and Optimality. *Linguistic Inquiry* 28:373-422.
Groat, Eric, & John O'Neill. 1996. Spellout at the LF-Interface. In *Minimal Ideas*, ed. Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson, & Jan-Wouter Zwart, 113-139. Amsterdam: Benjamins.
Grohmann, Kleanthes. 1997. German Superiority. *Groninger Arbeiten zur Germanistischen Linguistik* 40:97-107.
Grohmann, Kleanthes. 2003. Successive Cyclicity under (Anti-)Local Considerations. *Syntax* 6:260-312.
Haider, Hubert. 1983. Connectedness Effects in German. *Groninger Arbeiten zur Germanistischen Linguistik* 23:82-119.
Haider, Hubert. 1988. Theta-Tracking Systems. In *Configurationality*, ed. Laszlo Marácz & Pieter Muysken, 185-206. Amsterdam: Benjamins.
Haider, Hubert. 1993. *Deutsche Syntax – Generativ*. Tübingen: Narr.
Haider, Hubert. 1994. Detached Clauses – The Later the Deeper. Ms., Universität Stuttgart.
Haider, Hubert. 2000. Towards a Superior Account of Superiority. In *Wh-Scope Marking*, ed. Uli Lutz, Gereon Müller, & Arnim von Stechow, 231-248. Amsterdam: Benjamins.
Haider, Hubert. 2004. The Superiority Conspiracy. In *Minimality Effects in Syntax*, ed. Arthur Stepanov, Gisbert Fanselow, & Ralf Vogel, 147-176. Berlin: Mouton de Gruyter.
Heck, Fabian. 2000. Quantifier Scope in German and Cyclic Optimization. In *Competition in Syntax*, ed. Gereon Müller & Wolfgang Sternefeld, 175-209. Berlin: Mouton de Gruyter.
Heck, Fabian. 2004. *A Theory of Pied Piping*. Doctoral dissertation, Universität Tübingen.
Heck, Fabian, & Gereon Müller. 2001. Repair-Driven Movement and the Local Optimization of Derivations. Ms., Universität Stuttgart & IDS Mannheim.
Heck, Fabian, & Gereon Müller. 2002. Quantifier Raising in VP Ellipsis Constructions Revisited. Ms., Universität Stuttgart & IDS Mannheim.
Heck, Fabian, & Gereon Müller. 2005. Extremely Local Optimization. Ms., Universität Leipzig.
Hermans, Ben, & Mark van Oostendorp (ed.). 2000. *The Derivational Residue in Phonological Optimality Theory.* Amsterdam: Benjamins.
Hornstein, Norbert. 1995. *Logical Form*. Oxford: Blackwell.
Huang, Cheng-Teh James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Doctoral dissertation, MIT, Cambridge, Mass.
Huang, Cheng-Teh James. 1995. Logical Form. In *Government and Binding Theory and the Minimalist Program*, ed. Gert Webelhuth, 125-175. Oxford: Blackwell.
Kayne, Richard. 1981. Unambiguous Paths. In *Levels of Syntactic Representation*, ed.

Robert May & Jan Koster, 143-183. Dordrecht: Foris.
Keller, Frank. 2000. *Gradience in Grammar. Experimental and Computational Aspects of Degrees of Grammaticality*. Doctoral dissertation, University of Edinburgh.
Kim, Shin-Sook, & Wolfgang Sternefeld. 1997. Superiority vs. Crossover. Ms., Universität Tübingen.
Kitahara, Hisatsugu. 1993. Deducing 'Superiority' Effects from the Shortest Chain Requirement. *Harvard Working Papers in Linguistics* 3:109-119.
Kitahara, Hisatsugu. 1997. *Elementary Operations and Optimal Derivations*. Cambridge, Mass.: MIT Press.
Lasnik, Howard, & Mamoru Saito. 1992. *Move* α. Cambridge, Mass.: MIT Press.
Lasnik, Howard. 1999. On Feature Strength: Three Minimalist Approaches to Overt Movement. *Linguistic Inquiry* 30:197-217.
Legendre, Géraldine, Paul Smolensky & Colin Wilson. 1998. When is Less More? Faithfulness and Minimal Links in Wh-Chains. In *Is the Best Good Enough?*, ed. Pilar Barbosa et al., 249-289. Cambridge, Mass.: MIT Press & MITWPL.
McCarthy, John. 2000. Harmonic Serialism and Parallelism. *Proceedings of NELS* 30:501-524.
McCloskey, James. 1979. *Transformational Syntax and Model Theoretic Semantics. A Case Study in Modern Irish*. Dordrecht: Reidel.
McCloskey, James. 2002. Resumption, Successive Cyclicity, and the Locality of Operations. In: *Derivation and Explanation in the Minimalist Program*, ed. Samuel David Epstein & T. Daniel Seely, 184-226. Oxford: Blackwell.
Merchant, Jason. 2001. *The Syntax of Silence*. Oxford: Oxford University Press.
Merchant, Jason. 2002. Swiping in Germanic. In: *Studies in comparative Germanic syntax*, ed. C. Jan-Wouter Zwart and Werner Abraham, 295-321. Amsterdam: John Benjamins.
Müller, Gereon. 1997. Partial Wh-Movement and Optimality Theory. *The Linguistic Review* 14:249-306.
Müller, Gereon. 1998. *Incomplete Category Fronting*. Dordrecht: Kluwer.
Müller, Gereon. 1999. Imperfect Checking. *The Linguistic Review* 16:359-404.
Müller, Gereon. 2002. Two Types of Remnant Movement. In *Dimensions of Movement*, ed. Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers & Hans-Martin Gärtner, 209-241. Amsterdam: Benjamins.
Müller, Gereon. 2004. Phrase Impenetrability and Wh-Intervention. In *Minimality Effects in Syntax*, ed. Arthur Stepanov, Gisbert Fanselow, & Ralf Vogel, 289-325. Berlin: Mouton de Gruyter.
Müller, Gereon. 2006. Towards a Relativized Concept of Cyclic Linearization. Ms., Universität Leipzig.
Müller, Gereon, & Wolfgang Sternefeld. 1993. Improper Movement and Unambiguous Binding. *Linguistic Inquiry* 24:461-507.
Muriungi, Peter Kinyua. 2003. *Wh-questions in Kitharaka*. M.A. thesis, University of the Witwatersrand, Johannesburg.
Ormazabal, Javier, Juan Uriagereka & Miriam Uribe-Etxebarria. 1994. Word Order and *Wh*-Movement: Towards a Parametric Account. Ms., University of Connecticut & University of Maryland.
Ortiz de Urbina, Jon. 1989. *Parameters in the Grammar of Basque*. Dordrecht: Foris.
Perlmutter, David, & Scott Soames. 1979. *Syntactic Argumentation and the Structure of*

English. Berkeley: The University of California Press.
Pesetsky, David. 1995. *Zero Syntax*. Cambridge, Mass.: MIT Press.
Pesetsky, David. 1998. Some Optimality Principles of Sentence Pronunciation. In *Is the Best Good Enough?*, ed. Pilar Barbosa et al., 337-383. Cambridge, Mass.: MIT Press & MITWPL.
Pesetsky, David. 2000. *Phrasal Movement and Its Kin*. Cambridge, Mass.: MIT Press.
Prince, Alan & Paul Smolensky. 2004. Optimality Theory. Constraint Interaction in Generative Grammar. In *Optimality Theory in Phonology – A Reader*, ed. John J. McCarthy, 3-71. Oxford: Blackwell.
Rizzi, Luigi. 1996. Residual Verb Second and the *Wh*-Criterion. In *Parameters and Functional Heads*, ed. Adriana Belletti & Luigi Rizzi, 63-90. New York & Oxford: Oxford University Press.
Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Doctoral dissertation, MIT, Cambridge, Mass.
Ross, John Robert. 1969. Guess Who. *Papers from the 5th Regional Meeting of the Chicago Linguistic Society*, 252-286. University of Chicago.
Rubach, Jerzy. 2000. Glide and Glottal Stop Insertion in Slavic Languages: A DOT Analysis. *Linguistic Inquiry* 31:271-317.
Rudin, Catherine. 1988. On Multiple Questions and Multiple Fronting. *Natural Language and Linguistic Theory* 6:445-501.
Sabel, Joachim. 1998. *Principles and Parameters of Wh-Movement*. Habilitation thesis, Universität Frankfurt/Main.
Sabel, Joachim. 2000. Expletives as Features. In *Proceedings of WCCFL 19*, ed. Roger Billerey & Brook Danielle Lillehaugen, 411-424. Somervile, MA: Cascadilla Press.
Sauerland, Uli. 1999a. Erasability and Interpretation. *Syntax* 3:161-188.
Sauerland, Uli. 1999b. Locality in Ellipsis: Sluicing vs. Multiple Sluicing. Ms., Universität Tübingen.
Schmid, Tanja. 2005. *Infinitival Syntax. Infinitivus pro participio as a repair strategy.* Amsterdam: John Benjamins.
Speas, Margaret. 1995. Generalized Control and Null Objects in Optimality Theory. In *Papers in Optimality Theory*, ed. Jill Beckman, Laura Walsh-Dickie & Suzanne Urbanczyk, 637-653. Amherst, Massachussetts: UMass Occasional Papers in Linguistics 18.
Stechow, Arnim von. 2000. Partial Wh-Movement, Scope Marking, and Transparent Logical Form. In: *Wh-Scope Marking*, ed. Uli Lutz, Gereon Müller, & Arnim von Stechow, 448-478. Amsterdam: Benjamins.
Sternefeld, Wolfgang. 2005. *Syntax. Eine merkmalsbasierte Analyse*. Book ms., Universität Tübingen.
Stowell, Tim. 1981. *Origins of Phrase Structure*. Doctoral dissertation, MIT, Cambridge, Mass.
Takahashi, Daiko. 1993. Movement of Wh-Phrases in Japanese. *Natural Language and Linguistic Theory* 11:655-678.
Takahashi, Daiko. 1994*a*. *Minimality of Movement*. PhD thesis. University of Connecticut, Storrs.
Takahashi, Daiko. 1994*b*. Sluicing in Japanese. *Journal of East Asian Linguistics* 3:265-300.

Torrence, Harold. 2005. Asking Questions Quietly: The Syntax of Silent Wh-Words in Wolof. Ms., University of Kansas.

Torrego, Esther. 1984. On Inversion in Spanish and Some of Its Effects. *Linguistic Inquiry* 15:103-129.

Uriagereka, Juan. 1999. Multiple Spell-Out. In: *Working Minimalism*, ed. Samuel David Epstein and Norbert Hornstein, 251-282. Cambrigde, Mass.: MIT Press.

Wahba, Wafaa Abdel-Faheem Batran. 1992. LF Movement in Iraqi Arabic. In *Logical Structure and Linguistic Structure*, ed. C.-T. James Huang & Robert May, 253-276. Dordrecht: Kluwer.

Watanabe, Akira. 1995. Conceptual Basis of Cyclicity. *MITWPL* 27:269-291. Cambridge, Mass.: MIT.

Wilson, Colin. 2001. Bidirectional Optimization and the Theory of Anaphora. In *Optimality Theoretic Syntax*, ed. Géraldine Legendre, Jane Grimshaw & Sten Vikner, 465-508. Cambridge, Mass.: MIT Press.

Wunderlich, Dieter. 2006. The Force of Lexical Case: German and Icelandic Compared. In *The Nature of the Word: Essays in Honor of Paul Kiparsky*, ed. Kristin Hanson & Sharon Inkelas. Cambridge, Mass.: MIT Press.