Repair-Driven Movement and the Local Optimization of Derivations

Fabian Heck (Universität Stuttgart)
Gereon Müller (IDS Mannheim)

Correspondence address:
Gereon Müller
IDS Mannheim
P.O.Box 101621
D-68016 Mannheim
Germany

Email:
heck@ims.uni-stuttgart.de
gereon.mueller@ids-mannheim.de

Keywords:
optimality theory, minimalist program, scrambling, *wh*-movement, sluicing, quantifier raising

Abstract:
In this article, we justify a derivational model of local optimization in syntax that incorporates essential features of the minimalist program and optimality theory. Empirically, we focus on instances of repair-driven movement in German and English; i.e., exceptional movement that differs from standard instances of displacement in being triggered by the requirements of high-ranked constraints rather than by features, and that violates the ban on movement which is not feature-driven. Repair-driven movement is shown to underlie cases of *wh*-scrambling, successive-cyclic *wh*-movement, multiple *wh*-movement in sluicing contexts, and quantifier raising in VP ellipsis constructions. Finally, other syntactic repair phenomena (*do*-support, resumptive pronouns) are briefly addressed, and some general consequences of the local optimization approach are discussed.

Repair-Driven Movement and the Local Optimization of Derivations

Fabian Heck (Universität Stuttgart)
Gereon Müller (IDS Mannheim)

1. Introduction

Optimality-theoretic approaches have proven successful in the domain of syntactic repair phenomena like *do*-support and resumptive pronouns, where it looks as though a given operation which is not legitimate as such 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 (1993)). 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), among others). In this article, we want to justify a new, derivational model of local optimization that reconciles 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. Assume that syntactic structure is created derivationally by applying operations like Merge, Move, and Delete, with each XP a cyclic node (see Chomsky (1995)). Then, in classical optimality theory we would expect that there is a single, global optimization procedure that affects either the complete derivation, or one (or more) complete representation(s) generated by the derivation. In contrast, in the approach we pursue here, each subpart of the derivation from one cyclic XP to the next cyclic XP

[1]See Speas (1995) and Grimshaw (1997) (both based on Chomsky (1957; 1991)) on *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.

is subject to input/output optimization with respect to a set of violable and ranked constraints. The optimal subderivation that is the output of one optimization procedure is then used as the sole input for the next optimization step, and so on, until the numeration is empty and the root CP is reached. Since only an optimal output qualifies as a source for further optimization and all locally suboptimal outputs can be disregarded in subsequent derivational steps, it is clear that iterated local optimization reduces complexity, compared to global optimization. At any given stage of the derivation, the set of competing subderivations (the reference or candidate set) is further reduced by the presence of inviolable constraints. Among these, the STRICT CYCLE CONDITION (SCC) proves particularly important:[3]

(1) STRICT CYCLE CONDITION (SCC):
Within the current cyclic domain D, a syntactic operation may not target a position that is included within another cyclic domain D', such that D dominates D'.

Inviolable (and unranked) constraints belong to the Gen (generator) part of the grammar that creates a set of subderivation outputs for a given input; violable and ranked constraints belong to the H-Eval (harmony evaluation) part that selects the optimal output. The basic idea that we want to pursue here is that Gen and H-Eval continuously alternate in syntactic derivations, each affecting only a small portion of structure at any given stage of the derivation.[4]

[3]This formulation is based on Chomsky (1973) and Perlmutter & Soames (1979). The SCC is arguably derivable from more basic assumptions; see Chomsky (1995), Watanabe (1995), and Bošković & Lasnik (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]By being iterative, the present approach qualifies as an instance of what Prince & Smolensky (1993) 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 Wilson (1998) and Heck (2000); for phonology, see McCarthy (1999), 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) and Müller (2000).

In the following sections, we will argue on the basis of repair-driven movement in German and English that, independently of the complexity issue, local optimization turns out to be empirically superior to global optimization in syntax. The cases of repair-driven movement that we discuss involve *wh*-scrambling (section 2), successive-cyclic *wh*-movement (section 3), multiple *wh*-movement in sluicing constructions (section 4), and quantifier raising in VP ellipsis constructions (section 5). The structure of the argument will be similar in the four cases. It takes the following form. First, there is a constraint that blocks movement which is not feature-driven; following Chomsky (1995), Collins (1997), Kitahara (1997), and many others, we will call this constraint LAST RESORT:

(2) LAST RESORT (LR):
    Movement must result in feature checking.

LAST RESORT is complemented by the FEATURE CONDITION (see Chomsky (1995)), which forces overt movement in the presence of strong features.[5]

(3) FEATURE CONDITION (FC):
    Strong features must be checked by overt movement.

Second, 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. Third, this presupposes ranking (Con $\gg$ LAST RESORT) and violability (of LAST RESORT), and thus supports an optimality-theoretic analysis. Fourth and finally, the empirical evidence shows that repair takes place instantaneously, not at some earlier or later stage in the derivation. This suggests that optimization is local, not global.[6]

[5]For the moment, we are only concerned with overt movement. See section 5 on LF movement.

[6]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 favour of the notion "repair."

2. Repair-Driven Wh-Scrambling

2.1. The Ban on Wh-Scrambling in German

A standard assumption is that simple *wh*-movement is triggered by a strong *wh*-feature on C that attracts a *wh*-phrase with a matching feature, in accordance with LR and FC.[7] Following Grewendorf & Sabel (1999) and Sauerland (1999a), scrambling in German can be given a similar analysis. For concreteness, we will assume that a specific formal feature (which we can refer to as [Σ]) can optionally be inserted on functional heads of the extended projection of VP (i.e., on v, Neg, or T) in German, and matching [Σ] features can be instantiated on NP, PP, and CP; [Σ] is checked by movement to an (outer) specifier.[8]

As has often been noted (see, e.g., Fanselow (1990), Müller & Sternefeld (1993), and Rizzi (1996)), scrambling of *wh*-phrases normally results in ungrammaticality. This is shown by the contrast in (4): The second *wh*-phrase in a multiple question must stay in its base position; it cannot undergo scrambling to a SpecvP position.[9]

(4) a. Wie$_1$ hat [$_{vP}$ der Fritz t$_1$ was$_2$ repariert] ?
how has ART Fritz what fixed

b. *Wie$_1$ hat [$_{vP}$ was$_2$ [$_{vP}$ der Fritz t$_1$ t$_2$ repariert]] ?
how has what ART Fritz fixed

In a feature-based approach to scrambling, the ban on *wh*-scrambling in German can be expressed in terms of a feature co-occurrence restriction (cf. Gazdar et al. (1985)) that precludes a simultaneous instantiation of the features [+wh] and [Σ] on a single item. Under this assumption, the presence of the obligatory feature [+wh] on an NP blocks the presence of the optional feature [Σ]; consequently, *wh*-scrambling will always

[7] 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 (1999).

[8] 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 [Σ], 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.

[9] Note that subject NPs can stay in a vP-internal position; cf. Haider (1993). We assume that the EPP feature of T is optionally strong in German.

violate LR. Thus, at the stage of the derivation where v is merged with VP, and then with the subject NP, thereby creating the cyclic node vP, scrambling of a VP-internal *wh*-phrase to an outer specifier of vP is successfully blocked by LR, irrespectively of whether v is itself equipped with an optional [Σ] feature or not. Using optimality-theoretic notation, this reasoning can be illustrated as in table T_1; here O_1 and O_2 are the competing output vPs, the optimal vP O_1 is indicated by ☞, and the exclamation mark signals that O_2's LR violation is fatal.[10]

T_1: The ban on wh-scrambling in German

Input: $[_{\mathrm{VP}}\ wh_2\ \mathrm{V}\ ]$, v, NP	LR
☞O_1: $[_{\mathrm{vP}}$ NP ... $[_{\mathrm{VP}}\ wh_2$ ...]]	
O_2: $[_{\mathrm{vP}}\ wh_2\ [_{\mathrm{v'}}$ NP ... $[_{\mathrm{VP}}\ t_2$ ...]]]	*!

Interestingly, it has often been observed that there are certain systematic exceptions to the ban on *wh*-scrambling in German. Particularly interesting in the present context are those exceptions where it looks as though *wh*-scrambling can apply to fulfill an independently motivated syntactic constraint.[11] In the following three subsections, we consider three such cases.

2.2. Negative Intervention

As noted by Beck (1996) and Pesetsky (1999), *wh*-scrambling is not only possible, but in fact obligatory in contexts in which a *wh*-phrase is separated from its scope position by an intervening negative quantifier. Compare *wh*-scrambling with *wh*-in situ in (5).

(5) a. Wieso$_1$ hat sie $[_{\mathrm{NegP}}$ was$_2$ $[_{\mathrm{NegP}}$ niemals $[_{\mathrm{vP}}$ den Eltern t_2
why has she what$_{acc}$ never ART parents$_{dat}$
vorgeworfen]]] ?
accused of

b. *Wieso$_1$ hat sie $[_{\mathrm{NegP}}$ niemals $[_{\mathrm{vP}}$ den Eltern was$_2$ vorgeworfen]] ?
why has she never ART parents$_{dat}$ what$_{acc}$ accused of

This strongly suggests that *wh*-scrambling can and must apply in violation of LR, so as to respect another, higher-ranked constraint that precludes *wh*-items below negation.

[10] If v bears a [Σ] feature, FC is also violated, unless some other [Σ]-bearing item can be attracted.

[11] Other exceptions involve D-linked *wh*-phrases and connectedness effects.

Beck (1996) assumes that the relevant constraint (her Minimal Negative Structure Constraint; MNSC) applies only at LF; on this view, (5-b) violates the MNSC after *wh*-movement of was_2 to the domain of $C_{[+wh]}$ at LF. However, since we contend that the constraint has an immediate overt reflex – viz., forcing *wh*-scrambling in (5-a) as a repair operation –, we adopt a version that applies throughout the derivation:

(6) NEGATIVE INTERVENTION CONSTRAINT (NIC):
A negative item must not c-command a *wh*-phrase.

The ranking NIC $\gg$ LR implies that as soon as a Merge or Move operation creates a situation in which a negative item c-commands a *wh*-phrase, the NIC demands movement of the *wh*-phrase across the negative quantifier, as a repair operation that violates LR.[12] In the case at hand, suppose that the adverbial *niemals* is generated in the specifier of an empty head Neg. When Neg is merged with vP (containing *was*), the NIC becomes active. Given the ranking NIC $\gg$ LR, the optimal NegP will be one in which the *wh*-phrase has undergone scrambling to an outer specifier of Neg, in violation of LR, and not one in which the *wh*-phrase stays in situ, which fatally violates the NIC. The selection of the optimal NegP is shown in table T_2.[13]

T_2: Repair-driven wh-scrambling in negative contexts

Input: $[_{vP}$ ... wh_2 ... $]]$, Neg, niemals	NIC	LR
O_1: $[_{NegP}$ niemals Neg $[_{vP}$ ... wh_2 ... $]]$	*!	
☞O_2: $[_{NegP}$ wh_2 $[_{Neg'}$ niemals Neg $[_{vP}$ ... t_2 ... $]]]$		*

The optimal NegP O_2 forms part of the input for the next generation and optimization procedures that determine the optimal TP, and finally the optimal root CP (5-a). Crucially, the ungrammatical sentence (5-b) is never generated in the first place:

[12] At first sight, it may seem as though the NIC is too strict as it stands: A negative item that is merged in the matrix clause does not trigger repair-driven long-distance *wh*-scrambling from an embedded clause in German. This evidence could be reconciled with the NIC by adding a clause-mate requirement on the negative item and the *wh*-phrase in (6). However, we will later see that such a move is superfluous: As shown in section 6.3, the NIC itself is violable in favour of constraints that conspire to rule out long-distance scrambling of *wh*-phrases.

[13] We assume that the NIC (and other constraints of the same type that will be discussed below) applies only to output XPs, not to all derivational steps.

Its locally suboptimal NegP structure is filtered out early in the derivation and can never give rise to a complete CP candidate of this type.[14]

2.3. Weak Crossover

As a second example of repair-driven *wh*-scrambling, consider weak crossover contexts. For many speakers of German, scrambling can remedy weak crossover violations (cf. Haider (1988), Webelhuth (1992), and Frank, Lee & Rambow (1995), among others). Compare the ungrammatical example (7-b), where the pronoun *seinem*$_1$ ('his') is not bound by the quantified NP it is co-indexed with (*jeden Gast*$_1$ ('every guest')), with the well-formed example (7-a), where NP$_1$ is scrambled and binding is possible.

(7) a. Der Fritz hat [$_{NP}$ jeden Gast]$_1$ [$_{NP}$ seinem$_1$ Nachbarn] t$_1$ vorgestellt
ART Fritz has every guest$_{acc}$ his neighbour$_{dat}$ introduced

b. *Der Fritz hat [$_{NP}$ seinem$_1$ Nachbarn] [$_{NP}$ jeden Gast]$_1$ vorgstellt
ART Fritz has his neighbour$_{dat}$ every guest$_{acc}$ introduced

This, as such, does not yet argue for an optimality-theoretic approach: We can assume that [Σ] is instantiated on v and NP$_1$ in (7-a); so scrambling of NP$_1$ to the domain of v respects LR. However, the case is different with *wh*-phrases. As shown in (8-a) vs. (8-b), *wh*-scrambling may occur in this context so as to avoid a weak crossover violation, even though *wh*-phrases may not normally undergo scrambling, because of LR – as a matter of fact, *wh*-scrambling turns out to be impossible in the very same context if the pronoun bears a different index (e.g., that of the subject NP$_2$).

(8) a. Wann hat der Fritz$_2$ wen$_1$ [$_{NP}$ seinem$_{1/*2}$ Nachbarn] t$_1$ vorgestellt ?
when has ART Fritz whom$_{acc}$ his neighbour$_{dat}$ introduced

b. Wann hat der Fritz$_2$ [$_{NP}$ seinem$_{*1/2}$ Nachbarn] wen$_1$ vorgestellt ?
when has ART Fritz his neighbour$_{dat}$ whom$_{acc}$ introduced

For present purposes, it may suffice to adopt the following simplified Weak Crossover Constraint (see Reinhart (1983), Mahajan (1990), and Heim & Kratzer

[14] Note that this analysis also presupposes repair-driven intermediate *wh*-scrambling in simple questions like (i); but this has no effect on surface form, given subsequent FC-driven *wh*-movement.

(i) Was$_1$ hat sie [$_{NegP}$ t$'_1$ niemals den Eltern t$_1$ vorgeworfen] ?
what$_{acc}$ has she never the parents$_{dat}$ accused of

(1998) for related proposals).[15]

(9) WEAK CROSSOVER CONSTRAINT (WCC):
A bound-variable pronoun must be bound from an L-related position.

A pronoun that is co-indexed with a quantified NP must be interpreted as a bound variable, and is therefore subject to the WCC. Given a ranking WCC $\gg$ LR, the contrast between (8-a) and (8-b) is accounted for. (8-a) receives the following analysis. First, by principles that regulate the mapping from argument structure to syntax, the optimal VP is one in which the indirect (dative) object asymmetrically c-commands the direct (accusative) object. Let us assume that scrambling (be it feature- or repair-driven) is not available in the VP domain on general grounds. Then, the optimal VP must violate WCC.[16] Things are different in the subsequent vP cycle, though. Here, scrambling can in principle apply, and the ranking WCC $\gg$ LR ensures that it moves the direct *wh*-object across the indirect object that contains the co-indexed pronoun, to an outer specifier of vP. This is shown in table T_3 (where "x" designates the bound-variable pronoun).

T_3: Repair-driven wh-scrambling in weak crossover contexts

Input: $[_{\text{VP}}\ [_{\text{NP}}\ ...\ x_1\ ...\ ]\ wh_1\ \text{V}\ ]$, v, NP_{Subj}	WCC	LR
O_1: $[_{\text{vP}}\ \text{NP}_{Subj}\ [_{\text{NP}}\ ...\ x_1\ ...\ ]\ ...\ wh_1\ ...\ ]$	*!	
☞O_2: $[_{\text{vP}}\ wh_1\ [_{\text{v}'}\ \text{NP}_{Subj}\ [_{\text{NP}}\ ...\ x_1\ ...\ ]\ ...\ \text{t}_1\ ...\ ]]$		*

Thus, repair-driven movement applies as soon as possible to provide a binder for the bound-variable pronoun. Only O_1 is used as the input for the next generation and optimization procedures, which eventually lead to the optimal CP in (8-a) (with

[15]Among the L-related positions are the specifiers of T and v, but not of C. We assume that the distinction between L-related and non-L-related positions is not identical to the A- vs. A-bar distinction. Thus, we can maintain that German scrambling is movement to an L-related A-bar position; the latter will be relevant in the following section.

[16]Non-fatal WCC violations of this type are particularly obvious if a bound-variable pronoun is bound by a quantified NP that is merged much later in the derivation, e.g., in a matrix clause. That said, if we were to assume that VP qualifies as a scrambling site, the argument that follows would not be significantly different – repair-driven movement would apply during optimization of VP rather than vP. The main reason for not pursuing this option here is ease of exposition; otherwise, we would have to introduce constraints on the mapping from argument structure to syntax.

optional subject raising to SpecT). If the pronoun in T_3 bears a different index, *wh*-scrambling is blocked, due to a fatal LR violation – either because the pronoun is interpreted referentially (hence, not subject to the WCC), or because it is co-indexed with a different quantified NP that is merged later (so that a WCC is unavoidable).[17]

2.4. Parasitic Gaps

A third context that exhibits repair-driven *wh*-scrambling in German involves parasitic gaps. Parasitic gaps need a binder in an A-bar position; this requirement is captured by the following constraint.[18]

(10) PARASITIC GAP CONSTRAINT (PGC):
A parasitic gap is A-bar bound.

An A-bar binder for a parasitic gap can be provided by scrambling (see Felix (1983) and Bennis & Hoekstra (1984)). As shown by the data in (11), *wh*-scrambling is not just possible, but obligatory in a multiple question if it can provide an antecedent for a parasitic gap within an adjunct clause.[19]

(11) a. Wann hat die Maria was_1 [$_{CP}$ ohne e_1 zu lesen] dem Fritz t_1 zurückgegeben ?
when has ART Maria $what_{acc}$ without to read ART $Fritz_{dat}$ returned

[17]As with the NIC (see note 14), the analysis predicts intermediate *wh*-scrambling in cases like (i), which explains why the WCC can be fulfilled here even though the binder occupies a non-L-related position (see Grewendorf (1988)).

(i) Wen_1 hat der Fritz [$_{vP}$ t'_1 [$_{NP}$ $seinem_1$ Nachbarn] t_1 vorgestellt] ?
$whom_{acc}$ has ART Fritz his $neighbour_{dat}$ introduced

Intermediate repair-driven *wh*-scrambling without a visible output effect will also follow for simple questions from the analysis in the following section, but we will not specifically mention it there.

[18]The PGC might be a theorem that is derivable from more general assumptions (see, e.g., Chomsky (1982)). If so, it stands for a set of more primitive constraints that yield the effect it describes.

[19]As expected, *wh*-scrambling is impossible in this context if the parasitic gap e_1 is replaced by a resumptive pronoun es_1 ('it'). Note that a resumptive pronoun in an infinitival adjunct clause can escape the WCC. See Stowell (1991) and section 6.3 below.

b. *Wann hat die Maria [$_{CP}$ ohne e_1 zu lesen] dem Fritz was$_1$
when has ART Maria without to read ART Fritz$_{dat}$ what$_{acc}$
zurückgegeben ?
returned

This phenomenon is accounted for under the ranking PGC ≫ LR. The relevant parts of the derivation proceed as follows. First, the optimal VP (containing the potential binder *was*$_1$) and the optimal adjunct CP (containing the parasitic gap e_1) are independently determined by iterated optimization procedures; in the CP case, this involves iterated non-fatal violations of the PGC (there is no way how the parasitic gap can be A-bar bound CP-internally, assuming that an empty operator is either not available, or does not suffice to fulfill the PGC). However, at some point, CP and VP enter a common structure. Let us assume that this can take place at the vP-level.[20] Now a fulfillment of the PGC by *wh*-scrambling of the antecedent to an outer specifier of vP becomes an option. And indeed, as shown in table T_4, this strategy qualifies as optimal, given PGC ≫ LR.

T_4: Repair-driven wh-scrambling in parasitic gap contexts

Input: [$_{VP}$ NP wh_1 V], v, NP$_{Subj}$, [$_{CP}$ ... e_1 ...]	PGC	LR
O_1: [$_{vP}$ [$_{CP}$ ... e_1 ...] [$_{v'}$ NP$_{Subj}$ [$_{VP}$ ... wh_1 ...]]]	*!	
☞O_2: [$_{vP}$ wh_1 [$_{v'}$ [$_{CP}$ ... e_1 ...] [$_{v'}$ NP$_{Subj}$ [$_{VP}$... t_1 ...]]]]		*

Subsequent TP optimization involves fronting of the subject NP to SpecT (the EPP feature of T being optionally strong in the derivation that underlies (11-a)); finally, CP optimization implies movement of the second *wh*-phrase *wann* ('when') to SpecC, and of the auxiliary to C. However, as before, these later steps are not decisive for the contrast between (11-a) and (11-b) – the latter sentence will not be generated in the first place, its vP structure being filtered out as suboptimal earlier in the derivation.

The same type of repair-driven scrambling in parasitic gap constructions applies with morphologically simplex *wh*-phrases that are interpreted as indefinites. As shown in (12), indefinite *wh*-phrases normally cannot undergo scrambling (see Haider (1993)),

[20]The following reasoning would be compatible with an approach in which adjuncts are inserted in specifiers of designated functional heads; cf. Alexiadou (1997) and Cinque (1999). In that case, the functional head that hosts the adjunct CP defines the domain in which repair-driven movement applies.

just like regular *wh*-phrases.

(12) a. dass sie dem Fritz was zurückgegeben hat
that she ART Fritz_{dat} something_{acc} returned has

b. *dass sie was_1 dem Fritz t_1 zurückgegeben hat
that she something_{acc} ART Fritz_{dat} returned has

However, *wh*-indefinites can and must scramble so as to license parasitic gaps; see (13).

(13) a. dass die Maria was_1 [$_{\text{CP}}$ ohne e_1 zu lesen] dem Fritz t_1 zurückgegeben hat
that ART Maria something_{acc} without to read ART Fritz_{dat} returned has

b. *dass die Maria [$_{\text{CP}}$ ohne e_1 zu lesen] dem Fritz was_1 zurückgegeben hat
that ART Maria without to read ART Fritz_{dat} something_{acc} returned has

Given that indefinite *wh*-phrases cannot bear a [Σ] feature in German (like regular *wh*-phrases), the contrast in (13) follows as shown in table T_4.[21]

Thus far, we have presented evidence for repair-driven *wh*-scrambling, an operation that violates LR (because it is not feature-driven) in order to satisfy other constraints – NIC, WCC, and PGC. Since this implies a violability of LR, and a ranking NIC, WCC, PGC $\gg$ LR, the analysis supports an optimality-theoretic approach. However, thus far, the local approach that relies on iterated optimization of small portions of structure is empirically equivalent to the standard global approach according to which complete structures (complete derivations or representations, depending on the specific analysis) are optimized in a single step. We will now show that the local optimization approach is empirically superior because it successfully blocks certain unwanted instances of *wh*-scrambling that are predicted to be possible in the global approach. More specifically, it turns out that local optimization systematically precludes cases of chain interleaving in derived island contexts (see Browning (1991), Collins (1994), and Müller (1998)).

[21] Fanselow (1995) observes the same effect with other items that cannot normally undergo scrambling, like the NP *Gesindel* ('riff-raff'). It should be noted that Fanselow himself takes the wellformedness of examples like (13-a) as an argument against the parasitic gap analysis of this construction in general. Fanselow (1995; 1998) gives further arguments to this effect, at least some of which seem to us to be straightforwardly addressable within an optimality-theoretic approach. For reasons of space and coherence, we will not pursue these issues here.

2.5. Derived Islands

A robust generalization is that movement creates an island for extraction; this is a standard freezing effect (see Ross (1967), Wexler & Culicover (1980)). In the German passive example (14), (optional) NP-movement to SpecT co-occurs with *wh*-movement from NP; the result is ungrammatical.

(14) *[$_{CP}$ [$_{PP}$ Über wen]$_1$ ist [$_{TP}$ [$_{NP}$ ein Buch t$_1$]$_2$ [$_{vP}$ (t$'_1$) [$_{vP}$ [$_{PP}$ von keinem] t$_2$
about whom is a book by no-one
gelesen worden]]]] ?
read been

(14) can be generated by several derivations, and to ensure its ungrammaticality, each of these derivations must fatally violate some constraint. In derivation D_1, *wh*-movement to SpecC precedes NP-movement to SpecT; the second operation violates the SCC, which we have assumed to belong to the class of inviolable constraints. In derivation D_2, the two movement operations are reversed: *Wh*-movement applies after NP-movement. Here, the second movement step violates the CONDITION ON EXTRACTION DOMAIN (CED) which demands that movement must not cross a barrier (see Huang (1982), Chomsky (1986)). The reason is that, by (15-b), movement (to a non-complement position) invariably creates a barrier (see Cinque (1990)). Assuming that the CED is ranked high, it cannot be violated by a well-formed derivation.

(15) CONDITION ON EXTRACTION DOMAIN (CED):
a. Movement must not cross a barrier.
b. An XP is a barrier if it is a non-complement.

However, a third derivation D_3 must also be ruled out. Here, chain interleaving applies, in the sense that the *wh*-phrase is first scrambled out of the NP (while the latter is still in object position), and moved to an outer vP specifier; second, NP-movement to SpecT takes place; and finally, the *wh*-phrase is moved from the vP domain to SpecC, giving rise to t$'_1$ in (14). This derivation is not blocked by the SCC or the CED. Collins (1994) excludes this third derivation by invoking the transderivational economy condition Fewest Steps. Complex conditions of this type are abandoned in Collins (1997). In line with this, Collins suggests that chain interleaving can also be successfully blocked by local economy. Indeed, given that *wh*-phrases cannot instantiate a [Σ] feature, it is clear that *wh*-scrambling in D_3 must violate the local economy constraint LR in the present approach. Still, we have seen that LR can selectively be violated so as to fulfill higher-ranked constraints. Hence, it remains to be shown that the LR violation in D_3 is fatal, and repair-driven *wh*-scrambling is unavailable in this context. It turns out that this result can be achieved under local optimization, but not under global optimization, given an independently motivated ranking.

Let us first address the local optimization approach, and see how the illformedness of (14) can be ensured. Among the constraints of the H-Eval part of the grammar that are relevant here are the FC (which triggers *wh*-movement to SpecC, and NP-movement to SpecT if the EPP feature of T is strong), the CED (which blocks movement across a barrier), LR (which bans movement that is not feature-driven), and, finally, an EMPTY OUTPUT CONSTRAINT.[22]

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

Following Prince & Smolensky (1993), we assume that the empty output (the "null parse") is part of all competitions, i.e., it is generated as a possible option throughout. If the empty output Ø does not violate any other constraint, it 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)).

Suppose now that the ranking among these four constraints is CED, FC $\gg$ EOC $\gg$ LR (the ranking of CED and FC is not determined by the empirical evidence), and consider the outcome of local vP optimization in the derivation that underlies (14). The competition is illustrated in table T_5.

Among the vP candidates that are generated are outputs O_1 and O_2. In O_1, v is merged with VP, and the *by*-phrase is merged in a specifier of vP. In O_2, there is an additional scrambling operation that moves the *wh*-phrase to an outer specifier of

[22]Outputs that do not respect the SCC can be disregarded in what follows because they are not submitted to the optimization procedure in the first place.

T_5: Derived islands, local vP optimization

Input: [$_{VP}$ [$_{NP}$ ein Buch [$_{PP}$ über wen]$_1$]$_2$ V Aux], v, PP	CED	FC	EOC	LR
☞O_1: [$_{vP}$ PP [$_{VP}$ [$_{NP}$ ... [$_{PP}$ *wh*]$_1$]$_2$ ...] v]				
O_2: [$_{vP}$ [$_{PP}$ *wh*]$_1$ [$_{v'}$ PP [$_{VP}$ [$_{NP}$ ... t_1]$_2$ ...] v]]				*!
O_3: Ø			*!	

v. This movement step violates LR, and it does so fatally because the LR violation is locally unmotivated: At this point of the derivation, all higher-ranked constraints can be satisfied without a LR violation (the CED is vacuously fulfilled since there is no strong feature present that would require movement). O_3 fatally violates the EOC. O_1 is then used as the sole input for TP generation and optimization; assuming a strong EPP feature on T, the optimal TP has NP_2 movement to SpecT. Finally, various CP outputs are generated, and CP optimization applies. Clearly, since (14) is ungrammatical, there must be another CP output that is optimal; this turns out to be the empty output Ø. The competition is illustrated in table T_6 (the outputs are numbered O_{11}, O_{12}, ... so as to indicate that they are all descendants of O_1 in T_5).

T_6: Derived islands, local CP optimization

Input: [$_{TP}$ [$_{NP}$ ein Buch [$_{PP}$ über wen]$_1$]$_2$ [$_{vP}$ von keinem t_2 gelesen Aux]], C	CED	FC	EOC	LR
O_{11}: [$_{CP}$ [$_{PP}$ *wh*$_1$] $C_{[+wh]}$ [$_{TP}$ [$_{NP}$ ... t_1]$_2$ [$_{vP}$ PP t_2 ...]]]	*!			
O_{12}: [$_{CP}$ $C_{[+wh]}$ [$_{TP}$ [$_{NP}$ ... [$_{PP}$ *wh*$_1$]]$_2$ [$_{vP}$ PP t_2 ...]]]		*!		
☞O_{13}: Ø			*	

Here, O_{11} involves *wh*-movement from the NP in subject position, which fatally violates the CED. In O_{12}, the FC is fatally violated because the *wh*-phrase stays in situ, and the strong *wh*-feature of C remains unchecked in overt syntax. Finally, O_{13} is the empty output. Given the high ranking of the CED and the FC in German, the EOC violation incurred by Ø is non-fatal in this context. Ø becomes optimal, inducing a crash of the derivation, and hence the absence of any well-formed output, as desired for cases of derived islands.[23]

[23] As an alternative to this approach to ineffability in terms of the EOC and Ø (which is also essentially adopted in Ackema & Neeleman (1998) and Wunderlich (2000)), one might pursue an

2.6. Derived Islands and Local vs. Global Optimization

Let us now turn to the (standard) global optimization approach and see whether the derived island effect in (14) can be accounted for. The task is to show that (14) is successfully blocked by Ø. Note first that the ranking EOC $\gg$ LR must be assumed in the global approach just as it is required in the local approach. Otherwise, the evidence involving repair-driven scrambling in tables T_2–T_4 would be unaccountable (instead of repair-driven movement, we would expect a crash of the derivation, incurred by an optimal Ø). Evidently, the CED and the FC must outrank the EOC if (14) is to be blocked by Ø. Thus, the ranking must be the one determined for the local approach in T_5 and T_6 if the global approach is to succeed at all. However, it turns out that the wrong winner is selected under the global approach. This is shown in table T_7, in which complete structures (which may be derivations or representations) are subjected to an optimization procedure; here and henceforth, a wrong choice of optimal output in the global approach is signalled by ★.

T_7: Derived islands, global optimization

Input: V, Aux, v, PP, $T_{[EPP]}$, C, $[_{NP}$ ein Buch $[_{PP}$ über wen $]_1$ $]_2$	CED	FC	EOC	LR
O_1: $[_{CP}$ wh_1 $C_{[+wh]}$ $[_{TP}$ $[_{NP}$ ... t_1 $]_2$ $[_{vP}$ PP t_2 ...]]]	*!			
O_2: $[_{CP}$ wh_1 $C_{[+wh]}$ $[_{TP}$ – $T_{[EPP]}$ $[_{vP}$ PP $[_{NP}$ ... t_1 $]_2$ ...]]]		*!		
O_3: $[_{CP}$ $C_{[+wh]}$ $[_{TP}$ $[_{NP}$ ... wh_1 $]_2$ $[_{vP}$ PP t_2 ...]]]		*!		
★O_4: $[_{CP}$ wh_1 $[_{TP}$ $[_{NP}$ ... t_1 $]_2$ t_1' $[_{vP}$ PP t_2 ...]]]				*
O_5: Ø			*!	

approach in terms of "neutralization" of features. On this view, the role of the EOC would be played by a Feature Faithfulness (FF) constraint that prohibits, i.a., the deletion of *wh*-features on C and *wh*-phrases. The optimal candidate in T_6 might then reinterpret $C_{[+wh]}$ as $C_{[-wh]}$, and the *wh*-phrase as a *wh*-indefinite. This way, an initial feature distinction [±wh] is neutralized in the output. A neutralization approach to ineffability in syntax is pursued by Legendre, Smolensky & Wilson (1998) and Baković & Keer (1999), among others. The two approaches do not seem to make different predictions for the cases discussed in this article. Note in particular that FF introduces a dividing line into constraint rankings in exactly the way that the EOC does. The neutralization approach gives rise to systematic derivational ambiguity (in the sense that sentences can be optimal outputs of more than one derivation) that can only be remedied by additional assumptions (see Prince & Smolensky's concept of "input optimization," a second-order optimization procedure). For that reason, we adopt an EOC-based approach here and in what follows.

Outputs O_1, O_2, and O_3 incur fatal violations of high-ranked constraints (the CED or the FC – note that O_2 fails to apply NP movement and leaves the strong EPP feature of T feature unchecked, whereas O_3 fails to apply *wh*-movement, so that the strong *wh*-feature of C remains unchecked). The empty output O_5 is the one that should be optimal (compare T_6) but is not; its EOC violation is fatal because the competing output O_4 that involves chain interleaving avoids an EOC violation by violating the lower-ranked LR instead. Thus, O_4 qualifies as optimal. Hence, (14) is wrongly predicted to be well formed if t_1' is present. Crucially, whereas nothing is to be gained by repair-driven *wh*-scrambling to vP during vP optimization in the local approach, and it is impossible to apply this movement during subsequent CP optimization, when it would actually pay off, there is every reason to apply repair-driven *wh*-scrambling to vP (thereby inducing chain interleaving) in the global approach: The "early" LR violation can "later" be justified by the avoidance of a more severe EOC violation because "early" and "late" are non-existing concepts in an approach that simultaneously subjects all parts of a sentence to optimization.

3. Repair-Driven Successive-Cyclic Wh-Movement

3.1. The Problem

A well-known problem in theories that incorporate LR is to find a trigger for the intermediate steps of successive-cyclic *wh*-movement in sentences like (17) in German.[24]

(17) Wen$_1$ hat sie gesagt [$_{CP}$ t_1' dass Maria t_1 liebt] ?
     whom has she said that Maria loves

[24] 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)), *wh*-agreement in Chamorro (see Chung (1994)), 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)), obligatory V-to-C raising with (certain types of) *wh*-phrases in Spanish (see Torrego (1984) and Baković (1998)) and Basque (see Ortiz de Urbina (1989)), the selection of subject pronouns by C in Ewe (see Collins (1994)), tonal downstep in Kikuyu (see Clements, McCloskey, Maling & Zaenen (1983)), *meN* deletion in colloquial Singapore Malay (see Cole & Hermon (2000)), *wh*-copying in German (see Fanselow & Mahajan (2000)), and obligatory CP extraposition in German (see Müller (1998)).

A standard solution is to postulate an optional, strong operator feature on $C_{[-wh]}$ that triggers the movement of a *wh*-phrase to $\mathrm{SpecC}_{[-wh]}$ (see Collins (1997), Chomsky (1998), Sabel (1998), and Fanselow & Mahajan (2000), among others). This assumption faces both a conceptual and an empirical problem. A conceptual problem is that there does not seem to be a way to correlate the presence of the 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 $C_{[-wh]}$ that triggers *wh*-movement, it is unclear what precludes partial *wh*-movement of an embedded *wh*-phrase to the embedded $\mathrm{SpecC}_{[-wh]}$ position in a multiple question. This problem is especially pressing in a language like German, which exhibits such partial *wh*-movement in the presence of a *wh*-scope marker *was* instead of a real *wh*-phrase in the matrix clause (following Cheng (2000) and Sabel (2000), we assume that the scope marker is the realization of a moved bare *wh*-feature, i.e., it signals successive-cyclic *wh*-movement just like (17) does; see the last note). Thus, compare (18-a) ("partial" *wh*-movement with scope marking, an instance of successive-cyclic *wh*-movement) and (18-b) (*wh*-in situ) with the ill-formed (18-c) (true partial *wh*-movement without scope marking).

(18) a. Was$_1$ hat sie gesagt [$_{\mathrm{CP}}$ wen$_1$ Maria t$_1$ liebt] ?
what has she said whom Maria loves

b. Wer$_1$ hat t$_1$ gesagt [$_{\mathrm{CP}}$ – dass Maria wen$_2$ liebt] ?
who has said that Maria whom loves

c. *Wer$_1$ hat t$_1$ gesagt [$_{\mathrm{CP}}$ wen$_1$ Maria t$_1$ liebt] ?
who has said whom Maria loves

The contrast between legitimate *wh*-movement to $\mathrm{SpecC}_{[-wh]}$ in (17) and (18-a) and illegitimate *wh*-movement to $\mathrm{SpecC}_{[-wh]}$ in (18-c) is striking. Note that at the stage of the derivation where the embedded TP and $C_{[-wh]}$ are merged, these two categories are nearly identical in all four cases under consideration.[25] This strongly suggests that

[25] To be sure, $C_{[-wh]}$ 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

it cannot be an internal property of either the embedded TP or the $C_{[-wh]}$ head that is merged with it that forces or disallows *wh*-movement. 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 $SpecC_{[-wh]}$ is not feature-driven movement, but repair-driven movement that incurs a non-fatal violation of LR. This approach does not involve genuine look-ahead capacity; however, it makes crucial use of the concept of numeration (see Chomsky (1995)). More specifically, assuming that the numeration is accessible at each step of 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.[26]

3.2. Phase Balance

We propose that the constraint in question is (19).

(19) PHASE BALANCE (PB):
Every phase P has to be balanced: If P is a phase candidate, then for every strong feature F in the numeration there must be a distinct potentially available item to check F.

The notions *phase* and *potentially available* remain to be clarified. Following Chomsky (1998), we assume that every CP constitutes a phase, i.e., a special derivational unit; but deviating from Chomsky, we postulate that vP does not constitute a phase.[27] Syntactic material counts as potentially available within the current phase P if it is

sentences, without any consequence for wellformedness.

[26]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 – 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.

[27]The reason for this deviation is the following. As will be shown momentarily, PB forces movement of a *wh*-phrase to the specifier of a (declarative) CP in certain contexts. If vP also were to qualify as a phase, PB would systematically move a *wh*-phrase to a specifier of vP in simple questions. But then, the prohibition against chain interleaving that was discussed in section 2.4 cannot follow from LR anymore. On the contrary, we would expect that local vP optimization as in table T_5 produces O_2 (rather than O_1) as the optimal output, which would eventually make chain interleaving possible in the derived island construction (14), clearly an unwanted result.

either part of the numeration, or at the left edge (i.e., in SpecC) of P.[28] Thus, PB forces material from the current phase P that is supposed to check a feature within a higher (though as yet non-existent) phase P′ to move to the edge of P, in violation of LR. Given a ranking FC, PB ≫ EOC ≫ LR, successive-cyclic movement now emerges as a repair strategy.

As an illustration, let us consider the derivation that underlies (17), focussing on optimization of the embedded CP first. The competing CP outputs are generated by merging the optimal TP with a declarative complementizer $C_{[-wh]}$. One possible output, O_1, leaves the *wh*-phrase *wen* in situ. This, however, violates PB: For the strong *wh*-feature of the $C_{[+wh]}$ head that is still part of the numeration (and that will eventually become the root head), there is no potentially available item left that might check it; this is so because (a) there is no other *wh*-phrase left in the numeration, and (b) the *wh*-phrase *wen* is not at the left edge of CP. In contrast, in output O_2, *wh*-movement of *wen* has applied, in violation of LR. Consequently, *wen* is now potentially available for checking the *wh*-feature of $C_{[+wh]}$ in the numeration, CP is balanced, and PB is respected. Finally, O_3 is the empty output, which is suboptimal because of the ranking EOC ≫ LR. The competition from which repair-driven *wh*-movement emerges as optimal in the course of long-distance *wh*-dependencies is shown in table T_8.

T_8: Successive cyclic movement: local optimization of embedded CP

Input: $[_{\text{TP}}$ ... wh_1 ...], $C_{5[-wh]}$ Numeration = $\{C_{7[+wh]}, ...\}$	FC	PB	EOC	LR
O_1: $[_{\text{CP}_5}$ – $C_{[-wh]}$ ... wh_1 ...]		*!		
☞O_2: $[_{\text{CP}_5}$ wh_1 $C_{[-wh]}$ ... t_1 ...]				*
O_3: Ø			*!	

Only O_2 in T_8 can serve as the input for the next steps in the derivation. In the following optimization cycles, nothing spectacular happens, and table T_9 shows

[28] Material that is part of a tree that has been created earlier and has not yet been used in the derivation is not included in either the numeration or the current phase. Items in these external trees belong to the work space of the derivation, just like items in the numeration; we assume that they also count as potentially available. In what follows, we understand the notion of numeration in this extended sense, as comprising all derivational material outside the current tree. Also see Frampton & Gutman (1999) for related discussion.

that the root CP employs regular, feature-driven *wh*-movement from the embedded SpecC position (see O_{22}) rather than leaving the *wh*-phrase in the embedded SpecC 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 vacuously respected throughout.[29]

T_9: Successive cyclic movement: local optimization of matrix CP

Input: $[_{\text{TP}}$ ... $[_{\text{CP}_5}$ wh_1 $C_{[-wh]}$ ... t_1 ...]], $C_{7[+wh]}$ Numeration = { ... }	FC	PB	EOC	LR
O_{21}: $[_{\text{CP}_7}$ – $C_{7[+wh]}$ ... $[_{\text{CP}_5}$ wh_1 $C_{5[-wh]}$ ... t_1 ...]]	*!			
☞O_{22}: $[_{\text{CP}_7}$ wh_1 $C_{7[+wh]}$ ... $[_{\text{CP}_5}$ t'_1 $C_{5[-wh]}$ ... t_1 ...]]				
O_{23}: Ø			*!	

To sum up, an intermediate step in successive-cyclic *wh*-movement violates LR in order to provide a potentially available item for a strong *wh*-feature in the numeration and ensure that the embedded CP respects PB. The prediction is that no such repair-driven *wh*-movement to $\text{SpecC}_{[-wh]}$ should be possible if there is another *wh*-phrase left at this point that qualifies as potentially available for the *wh*-feature of C in the numeration. We now show that this prediction is borne out.

3.3. Long-Distance Superiority

German does not exhibit superiority effects with two *wh*-phrases that are clause-mates (see Bayer (1990) and Haider (1993), among many others):

(20) a. Wer$_1$ hat t_1 wen$_2$ getroffen ?
who has whom met

[29] As noted before, "partial" *wh*-movement as in (18-a) can be assumed to differ minimally from "full" *wh*-movement as in (17) in that the final *wh*-movement step affects a *wh*-feature in the former case, and a complete *wh*-phrase in the latter; the first option is not available in English. A similar analysis can be given for languages like Bahasa Malay and Kikuyu, where sentences are possible that superficially take the form of O_{21} in T_9, if we assume that the bare *wh*-feature does not have to be phonologically realized after movement here, so that an analysis as O_{22} is available after all. See Cole & Hermon (2000) and Sabel (1998), respectively.

b. Wen$_2$ hat wer$_1$ t$_2$ getroffen ?
whom has who met

However, as noted by Büring & Hartmann (1994), German does exhibit superiority effects with long-distance *wh*-movement. Compare (21-a) (= (18-b)), in which the matrix *wh*-phrase is moved to the matrix SpecC$_{[+wh]}$ position, with (21-b), in which the embedded *wh*-phrase undergoes such movement.

(21) a. Wer$_1$ hat t$_1$ gesagt [$_{CP}$ – dass Maria wen$_2$ liebt] ?
who has said that Maria whom loves

b. *Wen$_2$ hat wer$_1$ gesagt [$_{CP}$ (t$'_2$) dass Maria t$_2$ liebt] ?
whom has who said that Maria loves

For present purposes, we can neglect the issue of how standard (clause-bound) superiority effects are derived, and why such effects are absent in German. What is important here is the existence of the contrast between (20) and (21). This contrast strongly suggests that long-distance and clause-bound superiority phenomena should not be treated in a uniform way.[30] In what follows, we show that long-distance superiority effects can be reduced to the unavailability of successive-cyclic *wh*-movement. Let us begin by considering the local optimization of the embedded CP in (21-ab). The central question is whether the embedded *wh*-phrase *wen*$_2$ must move to SpecC$_{[-wh]}$ to fulfill PB, or is forced to stay in situ by LR. Given that there is another *wh*-phrase left in the numeration at this point that is potentially available for checking the *wh*-feature of C$_{[+wh]}$ (viz., *wer*$_1$, which is later merged in the matrix vP), CP can fulfill PB without the cost of a LR violation incurred by *wh*-movement. Hence, an output that involves *wh*-movement is suboptimal at this stage; the optimal output has *wh*-in situ. This is shown by table T$_{10}$.

T$_{10}$: Long-distance superiority: local optimization of embedded CP

Input: [$_{TP}$ ... wh_2 ...], $C_{5[-wh]}$ Numeration = {$C_{7[+wh]}$, wh_1, ...}	FC	PB	EOC	LR
☞O$_1$: [$_{CP_5}$ – $C_{5[-wh]}$ ... wh_2 ...]				
O$_2$: [$_{CP_5}$ wh_2 $C_{5[-wh]}$ ... t$_2$ ...]				*!
O$_3$: Ø			*!	

[30] Also see Takahashi (1993) and Pesetsky (1999) on a similar phenomenon in Japanese.

Turning next to the root CP, it is clear that only descendants of O_1 in T_{10} need to be considered. If neither of the two *wh*-phrases moves, this incurs a fatal violation of FC; the empty output fatally violates the EOC. The two remaining possibilities are *wh*-movement of the matrix *wh*-phrase and *wh*-movement of the lower *wh*-phrase. Consider the second option. Since the SCC precludes a use of the embedded SpecC position (independently of LR, which would also preclude it at this later stage, for the same reasons as before), movement must proceed in one step, thereby giving rise to an instance of non-local movement that we can expect to violate a high-ranked locality constraint. The two locality constraints adopted thus far (the NIC and the CED) do not constrain such one-step movement across a finite clause, though. However, the following constraint from Chomsky (1998) does.

(22) PHASE IMPENETRABILITY CONDITION (PIC):
Move operates only on locally available items.

Suppose that an item counts as *locally available* within the current phase P if P is the minimal phase that contains it, or if it is at the left edge of the previous phase P′.[31] As illustrated in table T_{11}, the contrast between (21-a) (= O_{12}) and (21-b) (= O_{13}) follows: *Wh*-movement of the higher *wh*-phrase respects the PIC, whereas *wh*-movement of the lower *wh*-phrase does not (the latter is not locally available in the matrix CP cycle). O_{11} and O_{14} fatally violate the FC and the EOC, respectively. Thus, the long-distance superiority effect is derived.[32]

[31] This analysis builds on Chomsky's (1998) proposal. The interaction of the PIC and PB in the present system is roughly comparable to that of the PIC and condition (24) in Chomsky's feature-based system, with PB a strengthened form of his (24).

[32] *Wh*-movement from infinitives embedded by restructuring verbs do not give rise to superiority effects in German (see Fanselow (1991)), modulo a non-identity requirement on *wh*-phrases (see Haider (2000)). However, infinitives embedded by non-restructuring verbs yield deviance in this case:

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

A standard way to capture the restructuring/non-restructuring distinction is to attribute CP (hence, phase) status to the complement of non-restructuring verbs, and vP or TP status to complements of restructuring verbs. If β is CP with *zögern* ('hesitate'), and vP/TP with *versuchen* ('try'), the PIC blocks *wh*-movement in (i-b) in the former case, but not in the latter.

T_{11}: *Long-distance superiority: local optimization of matrix CP*

Input: $[_{TP}$ ... wh_1 ... $[_{CP_5}$ – $C_{5[-wh]}$ ... wh_2 ...]], $C_{7[+wh]}$ Numeration = { ... }	FC	PB	PIC	EOC	LR
O_{11}: $[_{CP_7}$ – $C_{7[+wh]}$ ... wh_1 ... $[_{CP_5}$ – $C_{5[-wh]}$ ... wh_2 ...]]	*!				
☞O_{12}: $[_{CP_7}$ wh_1 $C_{7[+wh]}$ ... t_1 ... $[_{CP_5}$ – $C_{5[-wh]}$ ... wh_2 ...]]					
O_{13}: $[_{CP_7}$ wh_2 $C_{7[+wh]}$ ... wh_1 ... $[_{CP_5}$ – $C_{5[-wh]}$ ... t_2 ...]]			*!		
O_{14}: Ø				*!	

3.4. Further Long-Distance Intervention Effects

A conspicuous property of the present approach to long-distance superiority is that the actual hierarchical position of the second *wh*-phrase is irrelevant; what is relevant is the fact that it enters the derivation at a later stage. Hence, we expect superiority-like effects to arise with *wh*-phrases that are merged late but do not end up in a position that c-commands the lower *wh*-phrase. More specifically, long-distance *wh*-movement should also be impossible if there is another *wh*-element in the numeration (more generally, outside the current tree, see note 28) that eventually ends up in a more deeply embedded position, e.g., in an island. This prediction is borne out.[33] The evidence that shows this seems to have gone unnoticed so far, but it is clear enough. (23) illustrates a long-distance intervention effect with *wh*-in situ inside an adjunct.

(23) a. Wen$_1$ hat Fritz $[_{CP}$ nachdem er was$_2$ gemacht hat] t_1 getroffen ?
whom has Fritz after he what done has met

b. *Wen$_1$ hat Fritz $[_{CP}$ nachdem er was$_2$ gemacht hat] gesagt $[_{CP}$ (t'_1) dass Maria t_1 liebt] ?
whom has Fritz after he what done has said that Maria loves

(23-a) involves clause-bound, feature-driven *wh*-movement of *wen*$_1$. (23-b) is the interesting case. This sentence is as ill formed as the long-distance superiority example in

[33] The situation is different with clause-bound superiority effects, where non-c-commanding intervening *wh*-phrases do not block *wh*-movement; see Fiengo (1980). Again, this suggests that the two phenomena should not be treated on a par.

(21-b), even though the "intervening" *wh*-phrase was_2 does not c-command the base position of wen_1.[34] This is unexpected under standard 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 movement of the *wh*-phrase wen_1 in the lower CP is blocked by LR (PB is fulfilled by the potential availability of was_2), and feature-driven movement of wen_1 in the higher CP fatally violates the PIC. Still, there is a difference: In the superiority case, the higher *wh*-phrase moves to the matrix $\text{SpecC}_{[+wh]}$ position. In the present case, was_2 cannot move either; compare (23-b) with (24).

(24) *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

This case of absolute ungrammaticality (ineffability) follows without further ado from the assumptions made thus far, in particular, from the fact that the locality constraints CED and PIC both outrank the EOC. Table T_{12} shows that the derivation crashes in the very last step, with Ø emerging as the optimal output (PB is ignored here since it can never play a role in the last cycle of a derivation).

T_{12}: Long-distance intervention: local optimization of matrix CP

Input: [$_{\text{TP}}$ .. [$_{\text{CP}}$ .. wh_2 ..] [$_{\text{CP}_5}$ – $\text{C}_{5[-wh]}$ .. wh_1 ..]], $\text{C}_{7[+wh]}$ Numeration = { ... }	FC	CED	PIC	EOC	LR
O_1: [$_{\text{CP}_7}$ – $\text{C}_{7[+wh]}$ [$_{\text{CP}}$ ... wh_2 ...] ... [$_{\text{CP}_5}$ – $\text{C}_{5[-wh]}$ ... wh_1 ...]]	*!				
O_2: [$_{\text{CP}_7}$ wh_2 $\text{C}_{7[+wh]}$ [$_{\text{CP}}$ ... t_2 ...] ... [$_{\text{CP}_5}$ – $\text{C}_{5[-wh]}$ ... wh_1 ...]]		*!			
O_3: [$_{\text{CP}_7}$ wh_1 $\text{C}_{7[+wh]}$ [$_{\text{CP}}$ ... wh_2 ...] ... [$_{\text{CP}_5}$ – $\text{C}_{5[-wh]}$ ... t_1 ...]]			*!		
☞O_4: Ø				*	

The same type of long-distance intervention effect as in (23) shows up in (25) and (26). Here, the *wh*-in situ item shows up within a complex NP, in a relative clause (see (25)) or in a PP (see (26)). Given that extraction from the complex NP violates the

[34] The example improves drastically if was_2 is replaced by a non-*wh*-item like *das* ('this').

CED, this observation can be accounted for in exactly the way sketched in table T_{12}.

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

b. *Wen$_1$ hat Fritz [$_{NP}$ einem Mann [$_{CP}$ der was$_2$ kennt]] gesagt [$_{CP}$ (t$'_1$)
whom$_{acc}$ has Fritz a man$_{dat}$ that what knows said
dass er t$_1$ einladen soll] ?
that he invite should

(26) 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
whom$_{acc}$ has Fritz a friend$_{dat}$ of whom said that
Maria t$_1$ liebt] ?
Maria loves

3.5. Long-Distance Intervention and Local vs. Global Optimization

Thus far, we have argued that successive cyclicity supports an optimality-theoretic approach since it presupposes a violability of LR in favour of PB. Now we show that the data argue for the local system of optimization developed here, and against the global approach. In the local optimization procedure sketched in T_{12}, there is no candidate that involves successive-cyclic *wh*-movement of the lower *wh*-phrase. Such a candidate 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_{10}), 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 long-distance intervention effects. Again, 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 some higher-ranked constraint must be violated. This is shown in table T_{13}.

3.6. A Problem

To end this section, we discuss an empirical problem for the present approach that involves a case of undergeneration and thus shows that the system is slightly too

T_{13}: Long-distance intervention: global optimization

Input: $C_{5[-wh]}$, $C_{7[+wh]}$, wh_1, wh_2, ... Numeration = { ... }	FC	CED	PIC	EOC	LR
O_1: $[_{CP_7} - C_{7[+wh]}\ [_{CP} \ldots\ wh_2 \ldots] \ldots$ $[_{CP_5} - C_{5[-wh]} \ldots\ wh_1 \ldots]]$	*!				
O_2: $[_{CP_7}\ wh_2\ C_{7[+wh]}\ [_{CP} \ldots\ t_2 \ldots] \ldots$ $[_{CP_5} - C_{5[-wh]} \ldots\ wh_1 \ldots]]$		*!			
O_3: $[_{CP_7}\ wh_1\ C_{7[+wh]}\ [_{CP} \ldots\ wh_2 \ldots] \ldots$ $[_{CP_5} - C_{5[-wh]} \ldots\ t_1 \ldots]]$			*!		
★O_4: $[_{CP_7}\ wh_1\ C_{7[+wh]}\ [_{CP} \ldots\ wh_2 \ldots]] \ldots$ $[_{CP_5}\ t_1'\ C_{5[-wh]} \ldots\ t_1 \ldots]]$					*
O_5: Ø				*!	

restrictive as it stands. (27) may sound clumsy, but is undeniably well formed.

(27) $[_{NP}$ Die Frage $[_{CP_9}$ wer$_1$ $C_{[+wh]}$ t_1 was$_2$ mitbringt]]] ist relevant für die
the quetion who what brings is relevant to the
Frage $[_{CP_7}$ wie$_3$ Fritz denkt $[_{CP_5}$ t_3' dass die Party t_3 wird]
question how Fritz thinks that the party will be

The problem raised by (27) is that it is unclear how movement of *wie*$_3$ to SpecC$_5$ can be forced during CP$_5$ optimization. At the point where it has to be decided whether *wie*$_3$ moves or not, there are two *wh*-features left outside the current derivation that need potentially available items for checking, viz., the *wh*-features of C$_7$ and of C$_9$. For these two *wh*-features, there are two items left in the numeration (more generally, in the work space) that may check them, viz., *wer*$_1$ and *was*$_2$. Hence, it seems that CP$_5$ can fulfill PB without repair-driven *wh*-movement of *wie*$_3$. Consequently, *wh*-movement of *wie*$_3$ to SpecC$_7$ later in the derivation should be blocked, the optimal CP$_7$ should be Ø, and (27) should be ill formed, contrary to fact.

The problem here is that either *was*$_2$ or *wer*$_1$ "fools" *wie*$_3$: The *wh*-phrases do not compete for the same target position. Accordingly, we would like to suggest that *was*$_2$ and *wer*$_1$ do not qualify as potentially available for the *wh*-feature of C$_7$ that is to be checked by *wie*$_3$. To execute this idea, let us assume that *wh*-features are accompanied by scope indices in the numeration. Then, $XP_{[+wh],i}$ does not count as potentially available for a contra-indexed feature on $C_{[+wh],j}$, due to a feature mismatch. Hence, repair-driven *wh*-movement of wie_3 to SpecC_5 is the only way to fulfill PB during CP_5 optimization, and subsequent feature-driven movement of wie_3 to its target position SpecC_7 does not violate the PIC anymore, thereby rendering (27) well formed.

4. Repair-Driven Multiple Wh-Movement

4.1. The Ban on Multiple Wh-Movement in German

Only one *wh*-phrase moves overtly in German multiple questions; compare (28-a) (= (20-a)) with (28-b):[35]

(28) 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 strong features that are responsible for multiple *wh*-movement in languages like Bulgarian (see note 7) 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 multiple sluicing constructions. Before we address this issue, let us introduce an analysis of simple sluicing.

4.2. Simple Sluicing

(29) 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

[35] Grewendorf (1999) 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.

matrix clause if it is an argument).[36]

(29) Der Fritz hat irgend wen gesehen aber ich weiss nicht [$_{\text{CP}}$ wen$_1$ [$_{\text{TP}}$ ~~der~~
ART Fritz has someone seen but I know not whom ART
~~Fritz~~ t$_1$ ~~gesehen~~ ~~hat~~]]
Fritz seen has

We will make the following assumptions. First, sluicing involves TP deletion at PF (see, e.g., Lasnik (1999) and Merchant (1999)), not TP insertion at LF (as argued in Chung, Ladusaw, & McCloskey (1995)). Second, only complete constituents can be deleted (this contrasts with Ross's (1969) original approach that permits discontinuous deletion).[37] Finally, a T head can optionally bear a deletion feature [Δ]; PF deletion then affects the maximal projection of $T_{[\Delta]}$. Simple sluicing as in (29) does not yet involve repair-driven movement – the *wh*-phrase moves to the embedded SpecC position to check the *wh*-feature of $C_{[+wh]}$. The case is different with sluicing in multiple questions in German.

4.3. 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 (30-a), which is well formed, with (30-b), which is not a possible realization of an embedded multiple question:[38]

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

[37]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. 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 can 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.

[38]The PF-string in (30-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 has a [–wh] feature on the embedded object, not a [+wh] feature. Hence, such a derivation does not

(30) a. Irgend jemand hat irgend etwas geerbt aber der Fritz weiss nicht
someone has something inherited but ART Fritz knows not
mehr [$_{\text{CP}}$ wer$_1$ was$_2$ [$_{\text{TP}_{[\Delta]}}$ ~~t_1 t_2 geerbt~~ ~~hat~~]]
more who what inherited has

b. *Irgend jemand hat irgend etwas geerbt aber der Fritz weiss nicht
someone has something inherited but ART Fritz knows not
mehr [$_{\text{CP}}$ wer$_1$ [$_{\text{TP}_{[\Delta]}}$ ~~t_1 was$_2$ geerbt~~ ~~hat~~]]
more who what inherited has

Obligatory multiple *wh*-movement in (30-a) strongly suggests an analysis in terms of repair-driven movement in violation of LR. 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 (31), an instance of the Recoverability condition proposed in Chomsky (1965) and Pesetsky (1998):

(31) WH-RECOVERABILITY (WH-R):
Within a phase, a *wh*-phrase must not be dominated by a [Δ]-marked category.

Thus, the second *wh*-phrase in (30-a) moves to escape the [Δ]-marked TP. If it did not, WH-R would be fatally violated during CP optimization.[39] The local optimization procedure that underlies the data in (30) is shown in table T_{14}.

T_{14}: Multiple sluicing and clause-bound wh-movement: local optimization of CP

Input: ~~[$_{\text{TP}_{[\Delta]}}$ ... wh_1 ... wh_2 ...]~~, C$_{[+wh]}$	WH-R	EOC	LR
O_1: [$_{\text{CP}}$ wh_1 ~~[$_{\text{TP}_{[\Delta]}}$ ... t_1 ... wh_2 ...]~~]	*!		
☞O_2: [$_{\text{CP}}$ wh_1 wh_2 ~~[$_{\text{TP}_{[\Delta]}}$ ... t_1 ... t_2 ...]~~]			*
O_3: Ø		*!	

compete with the derivations that generate the sentences in (30).

[39]Two remarks. First, note that WH-R must be formulated in terms of [Δ]-marking rather than in terms of deletion in the present approach. The reason is that, by assumption, deletion applies at PF, whereas the movement that occurs to escape deletion applies before that, in overt syntax. Second, given that WH-R is restricted to phases, the question does not arise of whether repair-driven multiple *wh*-movement is an issue during TP optimization – at this stage of the derivation, WH-R is satisfied vacuously, and a violation of LR incurred by multiple *wh*-movement hence impossible.

Repair-driven successive-cyclic *wh*-movement differs substantially from repair-driven multiple *wh*-movement.[40] Still, the roles played by PB in the first case, and by WH-R in 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.

[40]In fact, an alternative would be to assume that multiple sluicing involves *wh*-scrambling to TP rather than *wh*-movement to SpecC. If so, the evidence for repair-driven movement would remain unaffected as such; but it would be evidence for repair-driven *wh*-scrambling, not for repair-driven *wh*-movement. There are some problems with such an approach, though. First, if multiple sluicing involves *wh*-scrambling, PF deletion must apply to a lower TP segment, leaving the higher TP segment in place. This complicates the theory of deletion; in the present approach, it requires additional, non-trivial restrictions to stop [Δ] feature percolation from T to the highest TP segment, and to force it to apply up to the second-highest segment. Second, assuming that *wh*-phrases must show up to the left of $C_{[+wh]}$ to be interpretable at LF (see, e.g., Beck (1996)), a scrambling analysis of multiple sluicing must posit an additional movement operation from an outer SpecT position to SpecC. Third, the non-scrambling language English exhibits multiple sluicing to some degree (see Bolinger (1978)). If the possibility of repair-driven α-movement can be tied to the availability of feature-driven α-movement in a given language, this may argue for a *wh*-movement analysis of multiple sluicing. (Japanese exhibits multiple sluicing; but Takahashi (1993) argues that overt *wh*-movement is optionally available in this language.)

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).

(i) Irgend jemand hat versucht/?*gezögert irgend etwas zu klauen aber ich weiss nicht [$_{CP}$ wer$_1$
    someone has tried/hesitated something to steal but I know not who
    was$_2$ [$_{TP_{[\Delta]}}$ t$_1$ ~~versucht/gezögert~~ ~~hat~~ [$_\beta$ t$_2$ ~~zu~~ ~~klauen~~]]]
    what tried/hesitated has to steal

However, this also follows under the *wh*-movement approach to multiple sluicing adopted here. As remarked in note 32, evidence from superiority effects points to a CP (phase) status of infinitival complements of non-restructuring verbs, and a vP/TP status of infinitives embedded by restructuring verbs. In the following section, we will see that PB and the PIC systematically predict *wh*-movement across a CP in multiple sluicing constructions to be impossible.

4.4. 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) 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.

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

(32) Die Maria hat behauptet dass sie irgend etwas geerbt hat aber der Fritz
ART Maria has claimed that she something inherited has but ART Fritz
weiss nicht mehr [$_{CP}$ was$_1$ [$_{TP_{[\Delta]}}$ ~~die Maria behauptet hat~~ [$_{CP}$ ~~t_1' dass sie t_1~~
knows not more what ART Maria claimed has that she
~~geerbt hat~~]]]
inherited has

This construction involves repair-driven *wh*-movement. However, the LR violation is not forced by WH-R (in the second step); rather, it is forced by PB (in the first step), 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_{15}.

T_{15}: Simple sluicing, local optimization of embedded CP

Input: [$_{TP}$ Maria was$_1$ geerbt hat], $C_{5[-wh]}$ Numeration = {$C_{7[+wh]}$, ... }	FC	WH-R	PB	LR
O$_1$: [$_{CP_5}$ – $C_{5[-wh]}$ [$_{TP}$ ... *wh*$_1$ ...]]			*!	
☞O$_2$: [$_{CP_5}$ *wh*$_1$ $C_{5[-wh]}$ [$_{TP}$ ... t$_1$ ...]]				*

O$_2$ is then the sole input for subsequent optimization. The next CP phase involves merging a [Δ]-marked TP and $C_{[+wh]}$. As shown in table T_{16}, 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.

Consider now the case of multiple long-distance sluicing in (33):

T_{16}: *Simple sluicing, local optimization of matrix CP*

Input: $\cancel{[_{\text{TP}_{[\Delta]}}\ \text{M. behauptet hat}\ [_{\text{CP}_5}\ \text{was}_1\ [_{\text{TP}}\ \text{M. t}_1\ \text{geerbt hat}\ ]}]$, $\text{C}_{7[+wh]}$ Numeration = { ... }	FC	WH-R	PB	LR
O_{21}: $[_{\text{CP}_7} - \text{C}_{7[+wh]}\ \cancel{[_{\text{TP}_{[\Delta]}} \ldots [_{\text{CP}_5}\ wh_1\ \text{C}_{5[-wh]}\ [_{\text{TP}} \ldots \text{t}_1 \ldots]}]$	*!	*		
☞O_{22}: $[_{\text{CP}_7}\ wh_1\ \text{C}_{7[+wh]}\ \cancel{[_{\text{TP}_{[\Delta]}} \ldots [_{\text{CP}_5}\ \text{t}_1'\ \text{C}_{5[-wh]}\ [_{\text{TP}} \ldots \text{t}_1 \ldots]}]$				

(33) *Irgend jemand hat behauptet dass die Maria irgend etwas geerbt hat aber
someone has claimed that ART Maria something inherited has but
der Fritz weiss nicht mehr $[_{\text{CP}}$ wer$_1$ was$_2$ $[_{\text{TP}_{[\Delta]}}$ ~~t$_1$~~ ~~behauptet~~ ~~hat~~ $[_{\text{CP}}$ ~~dass~~
ART Fritz knows not more who what claimed has that
$[_{\text{TP}}$ ~~die~~ ~~Maria~~ ~~t$_2$~~ ~~geerbt~~ ~~hat~~]]]]
ART Maria inherited has

Here, 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 – 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_{17}.

T_{17}: *Multiple long-distance sluicing, local optimization of embedded CP*

Input: $[_{\text{TP}}$ M. was$_2$ geerbt hat], $\text{C}_{5[-wh]}$ Numeration = $\{\text{C}_{7[+wh]}, wh_1, \ldots \}$	FC	WH-R	PB	EOC	LR
☞O_1: $[_{\text{CP}_5} - \text{C}_{5[-wh]}\ [_{\text{TP}} \ldots wh_2 \ldots]]$					
O_2: $[_{\text{CP}_5}\ wh_2\ \text{C}_{5[-wh]}\ [_{\text{TP}} \ldots \text{t}_2 \ldots]]$					*!
O_3: Ø				*!	

The optimization of the next CP phase involves outputs that are generated by merging $\text{TP}_{[\Delta]}$ (a descendant of O_1) with $\text{C}_{[+wh]}$. The situation is now similar to that in the well-formed multiple question (21-a). In that case, the *wh*-phrase in the lower clause stays in situ, and the *wh*-phrase in the higher clause moves to $\text{SpecC}_{[+wh]}$; see table T_{11}. 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 ≫ EOC, the empty output will win. Since (33) is in fact ungrammatical, this is the right result. Thus, the effect here is comparable to that shown in table T_{12}: Just as moving the higher *wh*-phrase induces a fatal violation of the CED in T_{12}, 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_{18}.[41]

T_{18}: Multiple long-distance sluicing, local optimization of matrix CP

Input: ~~[$_{\text{TP}[\Delta]}$ wer$_1$ behauptet hat [$_{\text{CP}_5}$ C$_{5[-wh]}$ [$_{\text{TP}}$ M. was$_2$ geerbt hat]]]~~, C$_{7[+wh]}$ Numeration = { ... }	FC	WH-R	PIC	EOC	LR
O$_{11}$: [$_{\text{CP}_7}$ – C$_{7[+wh]}$ ~~[$_{\text{TP}[\Delta]}$ *wh*$_1$ ... [$_{\text{CP}_5}$ C$_{5[-wh]}$ [$_{\text{TP}}$ ... *wh*$_2$ ...]]]~~]	*!	**			
O$_{12}$: [$_{\text{CP}_7}$ *wh*$_1$ C$_{7[+wh]}$ ~~[$_{\text{TP}[\Delta]}$ t$_1$ ... [$_{\text{CP}_5}$ C$_{5[-wh]}$ [$_{\text{TP}}$ ... *wh*$_2$ ...]]]~~]		*!			
O$_{13}$: [$_{\text{CP}_7}$ *wh*$_1$ *wh*$_2$ C$_{7[+wh]}$ ~~[$_{\text{TP}[\Delta]}$ t$_1$ ... [$_{\text{CP}_5}$ C$_{5[-wh]}$ [$_{\text{TP}}$ ... t$_2$ ...]]]~~]			*!		*
O$_{14}$: [$_{\text{CP}_7}$ *wh*$_2$ C$_{7[+wh]}$ ~~[$_{\text{TP}[\Delta]}$ *wh*$_1$ ... [$_{\text{CP}_5}$ C$_{5[-wh]}$ [$_{\text{TP}}$ ... 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$_{[+wh]}$, as we have seen to be the best strategy in T_{11}, 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 ensured.[42]

Suppose now that we were to adopt a global optimization approach to repair-driven *wh*-movement in sluicing constructions. This would accomodate the well-formed (i.e., clause-bound) cases of multiple *wh*-movement without problems. However, as with

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

[42] One might think that a [Δ]-marking on the embedded and on the matrix TP 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_{18}. However, this strategy is not available since it would imply successive-cyclic deletion, and hence, deletion of a non-constituent after the first step.

long-distance intervention effects (see T_{13}), the global approach invariably overgenerates: A variant of O_{13} in T_{18} 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 (33) well formed, is sketched in table T_{19}.

T_{19}: Multiple long-distance sluicing, global optimization

Input: $C_{5[-wh]}$, $C_{7[+wh]}$, wh_1, wh_2, T, $T_{[\Delta]}$, ... Numeration = { ... }	FC	WH-R	PIC	EOC	LR
O_{11}: $[_{CP_7}$ – $C_{7[+wh]}$ $\cancel{[_{TP_{[\Delta]}}\ wh_1 \ldots [_{CP_5}\ C_{5[-wh]}\ [_{TP} \ldots wh_2 \ldots]]]]}$	*!	**			
O_{12}: $[_{CP_7}$ wh_1 $C_{7[+wh]}$ $\cancel{[_{TP_{[\Delta]}}\ t_1 \ldots [_{CP_5}\ C_{5[-wh]}\ [_{TP} \ldots wh_2 \ldots]]]]}$		*!			
O_{13}: $[_{CP_7}$ wh_1 wh_2 $C_{7[+wh]}$ $\cancel{[_{TP_{[\Delta]}}\ t_1 \ldots [_{CP_5}\ C_{5[-wh]}\ [_{TP} \ldots t_2 \ldots]]]]}$			*!		*
O_{14}: $[_{CP_7}$ wh_2 $C_{7[+wh]}$ $\cancel{[_{TP_{[\Delta]}}\ wh_1 \ldots [_{CP_5}\ C_{5[-wh]}\ [_{TP} \ldots t_2 \ldots]]]]}$		*!	*		
O_{15}: Ø				*	
★O_{16}: $[_{CP_7}$ wh_1 wh_2 $C_{7[+wh]}$ $\cancel{[_{TP_{[\Delta]}}\ t_1 \ldots [_{CP_5}\ t_2'\ C_{5[-wh]}\ [_{TP} \ldots t_2 \ldots]]]]}$					**

5. Repair-Driven Quantifier Raising

Fox (1995) argues that whereas semantically vacuous quantifier raising at LF is normally blocked by economy, such movement must apply in certain VP ellipsis constructions in English in order to respect a Parallelism constraint. This can naturally be viewed as an instance of constraint violability and constraint ranking. Our goal in this section is to recast Fox's analysis in the system outlined in the previous sections, and to show that it supports an approach to optimization that is local, not global. Before we turn to that, we introduce Fox's (1995) analysis.

5.1. Scope and VP Ellipsis in Fox (1995)

(34) is a VP ellipsis construction in English. Assuming that VP ellipsis involves PF deletion, the material that is crossed out remains unpronounced at PF, but is accessible for LF operations.

(34) [$_{\text{CP}_1}$ Some boy admires every teacher], [and [$_{\text{CP}_2}$ some girl does [$_{\text{VP}}$ ~~admire every teacher~~] too]]

Assuming further that the operation of quantifier raising (QR) is available in the LF component, moving a quantified NP to some propositional category (vP or TP), we would now in principle expect (34) to have four different readings: The subject may have scope over the object in both conjuncts (∃∀ & ∃∀); the object may have scope over the subject in both conjuncts (∀∃ & ∀∃); the subject may have wide scope in the first conjunct, and narrow scope in the second; and the object may have wide scope in the first conjunct, and narrow scope in the second. However, the last two readings are systematically absent (i.e., *(∃∀ & ∀∃), *(∀∃ & ∃∀)). These readings can be excluded by a Parallelism constraint that demands that the scopal relationships among the elements of the two conjuncts are identical in ellipsis constructions.

In addition, Fox (1995) argues that QR is subject to an Economy constraint: QR may cross another item only to yield a different interpretation. Given Parallelism and Economy, it follows that unlike (34), (35) is unambiguous:

(35) [$_{\text{CP}_1}$ Some boy admires every teacher], [and [$_{\text{CP}_2}$ Mary does [$_{\text{VP}}$ ~~admire every teacher~~] too]] (∃∀, *∀∃)

If QR applies only in the first conjunct, Parallelism is violated; if it also applies in the second conjunct, Economy is violated (raising of *every teacher* across the proper name *Mary* does not give rise to a new interpretation). Hence, QR does not apply at all, and a wide scope of the subject is the only interpretation that is available. Interestingly, this cannot yet be the whole story. Scope ambiguity does arise if semantically vacuous QR is required not in the second conjunct (as in (35)), but in the first one; cf. (36).

(36) [$_{\text{CP}_1}$ Mary admires every teacher], [and [$_{\text{CP}_2}$ some boy does [$_{\text{VP}}$ ~~admire every teacher~~] too]] (∃∀, ∀∃)

Thus, from (34) we can conclude that QR is restricted by Parallelism; (35) shows that QR is restricted by Economy; and (36) suggests that semantically vacuous QR can nevertheless apply to respect Parallelism in certain contexts. Assuming a phrase structure of coordinate structures in which linear precedence translates into asymmetric c-command (such that CP_1 asymmetrically c-commands CP_2 in the examples discussed so far), Fox (1995) argues that the decision of whether vacuous QR is permitted or not has to be made on a local basis, regulated by the SCC (also see Sternefeld (1997)). On this view, vacuous QR in the lower CP_2 cycle is never permitted because Parallelism is irrelevant at this point; but vacuous QR in the higher CP_1 cycle is permitted if non-vacuous QR has taken place in the lower CP_2 cycle already, and Parallelism then requires the same LF movement operation in the higher CP_1 cycle.

Since this analysis implicitly relies on constraint ranking and constraint violability, it lends itself to an optimality-theoretic implementation according to which semantically vacuous QR in (36) emerges as a repair operation. Furthermore, the asymmetry between (35) and (36) suggests that optimization is local, not global. The following section implements Fox's (1995) analysis within the present approach.

5.2. A Local Optimization Approach

5.2.1. Premisses

For present purposes, we can adopt a simplified definition of PARALLELISM:[43]

(37) PARALLELISM (PAR):
In an ellipsis construction [CP_1 & CP_2], LF operations apply in the first conjunct iff they apply in second conjunct.

Furthermore, to obtain a local version of the Economy constraint (rather than a transderivational one, as in Fox's (1995) analysis), we will follow Chomsky (1995, 377) in assuming that QR is triggered by a [Q] feature. QR that does not check a [Q] feature violates LR. Given that [Q] is weak, it does not trigger overt movement. However, an

[43]Note that the status of [CP_1 & CP_2] as an ellipsis (i.e., PF-deletion) construction can be determined at LF by [Δ]-marking (see the previous section), without direct PF/LF interaction.

analogue of the FC demands that [Q] is checked at LF:[44]

(38) WEAK FEATURE CONDITION (WFC):
Features must be checked by movement.

Next, restrictions on the insertion of [Q] features must ensure that semantically vacuous QR violates LR. Suppose that a feature $[Q]_i$ can be merged with a propositional category α, attracting a co-indexed quantified NP_i, only if it minimally c-commands a contra-indexed quantified NP_j (this is adapted and simplified from Heck (2000)). As a consequence, QR of an NP respects LR only if it is triggered by a co-indexed [Q] feature; given the restriction on [Q] feature insertion, this implies that QR can only apply if it is locally non-vacuous. Suppose finally that the SCC applies to the LF part of the derivation as well, again with each XP as a cyclic node, subject to optimization. Given the asymmetrical nature of coordinate structures, the SCC ensures that optimization proceeds from bottom to top, beginning with the lower conjunct.[45] Assuming the ranking PAR, WFC $\gg$ EOC $\gg$ LR, we can now turn to the case of repair-driven QR in (36).

5.2.2. Repair-Driven Vacuous QR in the Higher Conjunct

(36) is repeated here as (39).

(39) $[_{CP_1}$ Mary admires every teacher], [and $[_{CP_2}$ some boy does $[_{VP}$ ~~admire every teacher~~] too]] ($\exists\forall$, $\forall\exists$)

[44]Here and in what follows, we maintain the classical assumption that covert movement induced by weak features applies after Spell-Out, in a separate LF component. However, most of what follows could be made compatible with an approach according to which covert movement is overt movement, accompanied by a pronunciation of traces. See Bobaljik (1995), Brody (1995), Groat & O'Neill (1996), Pesetsky (1998; 1999), Grewendorf (1999), and Fanselow & Ćavar (2000), among others.

[45]Or does it? If nothing else is said, the SCC in (1) would in fact not rule out an LF derivation in which first the XPs of the higher conjunct are being optimized, and then the XPs of the lower conjunct (see, e.g., Fox (2000)). The reason is that the first CP does not dominate the second CP, even though the former is higher in the structure. This problem disappears if we assume that specifiers are deficient in the sense that XP is a cyclic node as soon as SpecX is; this view might be supported by the observation that SpecX does not block c-command of its specifier into XP (see Reinhart (1983)). Alternatively, one might take this as evidence for a cascade structure (see Pesetsky (1995)) in which the higher CP actually dominates the lower CP.

It remains to be shown that there is a well-formed derivation for both readings. Let us first assume that a [Q] feature is not present in the numeration (and hence, absent in overt and covert syntax). In this case, the LF part of the derivation is straightforward. The two relevant optimization procedures of the LF derivation based on (39) involve the TP of the second conjunct and the TP of the first conjunct; for it is in the TP cycle that QR would have to apply to create wide scope for the object. It turns out that QR is not legitimate at any point in this derivation, which accounts for the reading ∃∀: Since there is no [Q] feature that might force QR by the WFC, the constraints Par, WFC, EOC, and LR are vacuously respected without QR.[46]

However, suppose now that a [Q] feature is present in the numeration. Given the above assumptions about [Q] feature insertion, the [Q] feature can only be merged in the second conjunct, and it must be co-indexed with the object quantifier. Suppose it is merged with T, after subject raising to SpecT.[47] Then, given the WFC, QR must apply in the lower conjunct, violating none of the constraints currently under consideration; in particular, Par is vacuously satisfied because the derivation has not yet reached the higher conjunct. This is shown in table T_{20}.[48]

T_{20}: Feature-driven QR, local LF optimization of TP in CP_2

Input: [$_{TP}$ Q_4 [$_{TP}$ some boy_3 [$_{vP}$ t_3 admires every $teacher_4$]]]	Par	WFC	EOC	LR
☞O_1: [$_{TP}$ every $teacher_4$ [$_{TP}$ some boy_3 [$_{vP}$ t_3 admires t_4]]]				
O_2: [$_{TP}$ Q_4 [$_{TP}$ some boy_3 [$_{vP}$ t_3 admires every $teacher_4$]]]		*!		
O_3: [$_{TP}$ some b.$_3$ [$_{TP}$ every t.$_4$ [$_{TP}$ t'_3 [$_{vP}$ t_3 admires t_4]]]]				*!
O_4: Ø			*!	

However, Par becomes active when the derivation has reached the higher conjunct, and given that Par outranks LR, semantically vacuous (i.e., non-feature-driven) QR must apply in the higher conjunct, where a [Q] feature cannot possibly show up; see

[46]An underlying assumption here is that object quantifiers can be interpreted in situ, i.e., QR is not type-driven. A simple way to ensure this is to assume that verbs are interpreted as open formulae; see, e.g., Heim (1982) and Sternefeld (1998).

[47]In principle, it could also be merged with v, but given overt subject raising to SpecT, this would not yield a new interpretation, and we can disregard such a derivation.

[48]Only those parts are considered that play a role at a given step. Thus, CP_1 is not yet represented because the LF derivation has not yet reached this domain.

table T_{21}. Thus, the reading ∀∃ of (39) is accounted for.

T_{21}: Repair-driven QR, local LF optimization of TP in CP_1

Input: [$_{TP}$ Mary$_5$ [$_{vP}$ t$_5$ admires every teacher$_6$]] and [$_{CP_2}$ [$_{TP}$ every teacher$_4$ [$_{TP}$ some boy$_3$ [$_{vP}$ t$_3$ admires t$_4$]]]]	PAR WFC	EOC	LR
O_{11}: [$_{TP}$ Mary$_5$ [$_{vP}$ t$_5$ admires every teacher$_6$]] and [$_{CP_2}$ [$_{TP}$ every t.$_4$ [$_{TP}$ some b.$_3$ [$_{vP}$ t$_3$ admires t$_4$]]]]	*!		
☞O_{12}: [$_{TP}$ every teacher$_6$ [$_{TP}$ Mary$_5$ [$_{vP}$ t$_5$ admires t$_6$]]] and [$_{CP_2}$ [$_{TP}$ every t.$_4$ [$_{TP}$ some b.$_3$ [$_{vP}$ t$_3$ admires t$_4$]]]]			*
O_{12}: Ø		*!	

5.2.3. *The Ban on Vacuous QR in the Lower Conjunct*

The unambiguous example (35) is repeated here as (40).

(40) [$_{CP_1}$ Some boy admires every teacher], [and [$_{CP_2}$ Mary does [$_{VP}$ ~~admire every teacher~~] too]] (∃∀, *∀∃)

Again, two LF derivations must be considered. In one derivation, [Q] is not present in the numeration. This derivation does not involve QR and yields the reading ∃∀. The more interesting derivation is one in which a [Q] feature shows up on the higher TP (in accordance with the restrictions on [Q] feature insertion mentioned above). As shown in table T_{22}, the quantified object NP must stay in situ in the lower TP.

T_{22}: No repair-driven QR, local LF optimization of TP in CP_2

Input: [$_{TP}$ Mary$_3$ [$_{vP}$ t$_3$ admires every teacher$_4$]]	PAR WFC	EOC	LR
☞O_1: [$_{TP}$ Mary$_3$ [$_{vP}$ t$_3$ admires every teacher$_4$]]			
O_2: [$_{TP}$ every teacher$_4$ [$_{TP}$ Mary$_3$ [$_{vP}$ t$_3$ admires t$_4$]]]			*!
O_3: Ø		*!	

Subsequent QR in the higher TP (as forced by the WFC) will therefore have to violate PAR. Hence, given the ranking PAR, WFC ≫ EOC, the empty output is optimal, i.e., the derivation crashes at LF. Consequently, the reading ∀∃ is not available. This is shown in table T_{23}.

Consider finally (34), repeated here as (41).

T_{23}*: No feature-driven QR, local LF optimization of TP in* CP_1

Input: $[_{TP}$ Q_6 $[_{TP}$ some b.$_5$ $[_{vP}$ t_5 admires every t.$_6$]]] and $[_{CP_2}$ $[_{TP}$ Mary$_3$ $[_{vP}$ t_3 admires every teacher$_4$]]]	PAR	WFC	EOC	LR
O_{11}: $[_{TP}$ Q_6 $[_{TP}$ some b.$_5$ $[_{vP}$ t_5 admires every t.$_6$]]] and $[_{CP_2}$ $[_{TP}$ Mary$_3$ $[_{vP}$ t_3 admires every t.$_4$]]]		*!		
O_{12}: $[_{TP}$ every t.$_6$ $[_{TP}$ some b.$_5$ $[_{vP}$ t_5 admires t_6]]] and $[_{CP_2}$ $[_{TP}$ Mary$_3$ $[_{vP}$ t_3 admires every t.$_4$]]]	*!			
☞O_{13}: Ø			*	

(41) $[_{CP_1}$ Some boy admires every teacher], [and $[_{CP_2}$ some girl does $[_{VP}$ ~~admire every teacher~~] too]]

First, if both conjuncts lack a [Q] feature, QR is unavailable in either conjunct, and the uniform reading ∃∀ is derived. Second, if both conjuncts exhibit a [Q] feature (which is possible because the insertion can be non-vacuous in both cases), feature-driven QR applies in both conjuncts, yielding the uniform reading ∀∃. Third, if there is a [Q] feature in the lower conjunct, but not in the higher one, PAR forces repair-driven QR in the higher conjunct, in violation of LR. The derivation proceeds essentially as shown in T_{21}, with a uniform ∀∃ reading emerging. Finally, if there is a [Q] feature in the higher conjunct, but not in the lower one, matrix optimization selects Ø as the optimal candidate, as illustrated in T_{23}. Thus, the unavailability of the non-parallel readings in (41) is derived.

5.3. Quantifier Raising and Local vs. Global Optimization

Note that it is important that PAR is not yet an issue at the stage of the derivation where it must be decided whether QR is possible in the lower conjunct in T_{22}. Later in the derivation, in the higher conjunct, PAR does become relevant (cf. T_{23}), but then it is too late for QR to apply in the lower conjunct. Thus, the analysis relies on the assumption that optimization proceeds locally. Otherwise QR could apply early (with only a LR violation) to avoid a problem (i.e., a PAR or WFC violation) that shows up later in the derivation. Hence, it is not surprising that a global optimization approach would incorrectly predict a wide scope reading for the object to be possible in (40): If the whole sentence is evaluated in one step (i.e., T_{22} and T_{23} are combined), the optimal output can respect Par, WFC, and EOC, and violate LR by applying repair-driven QR in CP_2 and feature-driven QR in CP_1. We would then wrongly expect (40) to be ambiguous. This is shown in T_{24}.

T_{24}: Repair-driven QR, global LF optimization

Input: $[_{CP_1}$ $[_{TP}$ Q_6 $[_{TP}$ some b.$_5$ $[_{vP}$ t_5 admires every t.$_6$]]]] and $[_{CP_2}$ $[_{TP}$ Mary$_3$ $[_{vP}$ t_3 admires every t.$_4$]]]	Par	WFC	EOC	LR
O_1: $[_{CP_1}$ $[_{TP}$ Q_6 $[_{TP}$ some b.$_5$ $[_{vP}$ t_5 admires every t.$_6$]]]] and $[_{CP_2}$ $[_{TP}$ Mary$_3$ $[_{vP}$ t_3 admires every t.$_4$]]]		*!		
O_2: $[_{CP_1}$ $[_{TP}$ Q_6 $[_{TP}$ some b.$_5$ $[_{vP}$ t_5 admires every t.$_6$]]]] and $[_{CP_2}$ every t.$_4$ $[_{TP}$ Mary$_3$ $[_{vP}$ t_3 admires t_4]]]	*!	*		*
O_3: $[_{CP_1}$ $[_{TP}$ every t.$_6$ $[_{TP}$ some b.$_5$ $[_{vP}$ t_5 admires t_6]]]] and $[_{CP_2}$ $[_{TP}$ Mary$_3$ $[_{vP}$ t_3 admires every t.$_4$]]]	*!			
★O_4: $[_{CP_1}$ $[_{TP}$ every t.$_6$ $[_{TP}$ some b.$_5$ $[_{vP}$ t_5 admires t_6]]]] and $[_{CP_2}$ every t.$_4$ $[_{TP}$ Mary$_3$ $[_{vP}$ t_3 admires t_4]]]				*
O_5: Ø			*!	

This concludes the presentation of the four instances of repair-driven movement that are the main topic of this article.[49] Needless to say, the local approach to opti-

[49]It should be noted that Fox (2000, ch. 3) abandons his earlier approach in Fox (1995). We will briefly address two counter-arguments that he presents. First, Jacobson (1997) points out that the same asymmetry as that between (35) and (36) occurs in intersentential contexts; see (i-ab). This suggests a uniform analysis, which implies that the SCC applies intersententially. Furthermore, it is not clear why (i-c) is unambiguous even though the last (hence, by assumption, lowest) sentence should permit semantically non-vacuous QR.

(i) a. At least one critic from the *Times* admires every movie. Pauline Kael does, too. (∃∀, *∀∃)
 b. Pauline Kael admires every movie. At least one critic from the *Times* does, too. (∃∀, ∀∃)
 c. At least one critic from the *Times* admires every movie. Pauline Kael does, too; and at least one critic from the *Post* does, too. (∃∀, *∀∃)

It is not clear how severe the problem posed by (i) is for Fox's (1995) account. An intersentential application of (something like) the SCC does not seem to be inherently implausible. Furthermore, it is reasonable to assume that the SCC does not indiscriminately apply to a complete discourse, but only to parts of it. Then, (i-c) might instantiate a situation where the last, afterthought-like sentence represents a separate subdiscourse; this would suffice to block QR in the first sentence by the SCC.

Second, Fox (2000) notes that the effect in (35)/(36) does not show up in other deletion constructions like (ii-ab); the absence of the inverted reading in (ii-b) is a priori unexpected.

mization in syntax that we argued for on this basis raises a number of further questions. We will discuss some of these in the following section.

6. Further Issues

6.1. Other Repair Phenomena

In this article, the focus has been on repair phenomena that involve movement. However, several other syntactic repair phenomena (most of which involve repair-driven lexicalization in a broader sense) have successfully been addressed in the literature on the basis of standard (global) optimization. The question arises to what extent these approaches are compatible with the more restricted system of local optimization developed here. It turns out that some of the existing analyses can be integrated without problems; others demand modifications; and yet others may be fundamentally incompatible. Schmid's (1998) analysis of German Ersatzinfinitiv (Infinitivus Pro Participio, IPP) constructions belongs to the first class. Here, all repair operations (repair-driven replacement of a selected participle by a non-selected infinitive, repair-driven VP extraposition) apply locally, and there is no problem for the present approach. The repair phenomena in the following two sections are not quite as straightforwardly compatible, and hence, more interesting in the present context.

(ii) a. A boy talked to every teacher, and a girl did ~~talk~~ to Jane (∃∀, *∀∃)

b. A girl talked to Jane, and a boy did ~~talk~~ to every teacher (∃∀, *∀∃)

But again, it seems that this evidence does not necessarily undermine the analyis in Fox (1995). There are at least two fundamental differences between (35)/(36) and (ii-ab) that can be exploited in the present approach. For one thing, the deletion affects only a bare verb in (ii). Suppose that only XPs can undergo deletion. Then, (ii-ab) must be analyzed in terms of object movement plus remnant VP deletion (see Kayne (1998) and Johnson (1998)). As a consequence, the object in the lower conjunct must occupy a derived position, and derived positions are known to freeze the overt scope. The other important difference concerns the object in the higher conjunct; it is a quantified NP in (35)/(36), and a proper name in (ii-b). Hence, to ensure that (ii-b) is unambiguous it suffices to assume that proper name NPs cannot undergo semantically vacuous QR as easily as quantified NPs. We would like to conclude from this that, whatever the merits of the revised approach in Fox (2000) are, the original approach is still a viable alternative.

6.1.1. Do-Support

Do-support is obligatory in (non-subject-initial) root *wh*-questions in English (see (42-ab)); it is blocked in embedded *wh*-questions (see (42-cd)).

(42) a. [$_{\text{CP}}$ Which book$_1$ do$_2$ [$_{\text{TP}}$ you t$_2$ like t$_1$]] ?
b. *[$_{\text{CP}}$ Which book$_1$ – [$_{\text{TP}}$ you T like t$_1$]] ?
c. *I wonder [$_{\text{CP}}$ which book$_1$ do$_2$ [$_{\text{TP}}$ you t$_2$ like t$_1$]]
d. I wonder [$_{\text{CP}}$ which book$_1$ – [$_{\text{TP}}$ you T like t$_1$]]

Based on Chomsky (1957; 1991), Grimshaw (1997) analyses *do*-support in root *wh*-clauses as a repair phenomenon. First, there is a low-ranked constraint that prohibits the insertion of expletives, including (semantically empty) *do* (FULL-INT). Second, two higher-ranked constraints (OP-SPEC and OB-HD) conspire to force *wh*-movement to SpecC and a filling of C by verb movement. Third, there is a high-ranked constraint blocking movement of a full lexical verb (NO-LEX-MVT). Since, of these four constraints, FULL-INT is ranked lowest in English, *do*-support is optimal in (42-ab).

Note that *do* is not directly inserted in C in (42-a) in this analysis; rather, it is inserted in T and then moved to C. This assumption is necessary in the system of Grimshaw (1997): *Do*-support is blocked in embedded clauses in favour of a null C position violating OB-HD; see (42-cd). This is so because of a high-ranked constraint PURE-EP that prohibits movement to an embedded C. Hence, in order for (42-c) to fatally violate PURE-EP, *do* must be moved from T to C, and cannot simply be inserted in T. This feature of Grimshaw's analysis is interesting in the present context because it turns out to be problematic for a local optimization account. The trigger for repair occurs in the CP domain (C must be filled if a *wh*-phrase is in SpecC), but the repair itself must affect the lower TP domain (*do* is inserted in T). In other words: The repair operation does not seem to be local.

There are various ways to avoid this problem while maintaining the essentials of Grimshaw's (1997) approach. For one thing, the problem disappears under Grimshaw's (1999) factorization of PURE-EP into a system of more elementary constraints; for another, one might minimally enlarge optimization domains (as argued on independent grounds in Fanselow & Ćavar (2000)). Here we would like to sketch a third possibility, a modification of Grimshaw's (1997) analysis based on Roberts (1998). Suppose that *do*-support in *wh*-constructions instantiates an obligatory PF realization of feature movement from T to C. Under this view, the cost of *do*-support is not a violation of FULL-INT, but of a constraint against bare feature movement. OB-HD can be replaced by a constraint that requires a filled SpecC position to agree with verbal features; and the work of OP-SPEC is done by the FC. Assuming finally that the nature of C as root or embedded is encoded by appropriate features on C in the numeration, Grimshaw's repair analyis of *do*-support can be transferred into the present approach: *Do*-support does not apply in the optimal TP in (42-ab), but the optimal CP is one with bare feature movement, which triggers PF-realization as *do*. Bare feature movement (hence, *do*-support) is blocked in (42-cd), though, because of a higher-ranked constraint that prohibits movement to an embedded C.

6.1.2. Resumptive Pronouns

The examples in (43-ab) show that resumptive pronouns are only available as a repair strategy, to avoid a violation of a locality constraint like the COMPLEX NOUN PHRASE CONDITION (CNPC) that a trace would incur; see Chomsky (1982).[50]

(43) a. the man [$_{\text{CP}_1}$ who(m)$_1$ I saw *him$_1$/t$_1$]
b. the man [$_{\text{CP}_2}$ who(m)$_1$ I don't believe [$_{\text{NP}}$ the claim [$_{\text{CP}_1}$ (t$'_1$) that anyone saw ?him$_1$/*t$_1$]]]

Optimality-theoretic analyses are given in Legendre, Smolensky & Wilson (1998) and Pesetsky (1998); a similar repair approach is developed in Hornstein (2000). Although the details of these analyses differ a great deal, the gist of the explanation is identical: The CNPC (or whatever derives its effects) is ranked higher than a constraint against (resumptive or, perhaps, all) pronouns (AVOID PRONOUN, in Chomsky (1982)). Thus, where a trace can fulfill both constraints, resumptive pronouns are impossible (see (43-a)); and where a trace must violate the CNPC, the violation of AVOID PRONOUN that is incurred by the resumptive pronoun becomes non-fatal, and the respective output turns out to be optimal (see (43-b)).

[50]To simplify matters, we give examples from English here. The resumptive pronoun effect is even clearer in other languages (like Hebrew, Chinese, and Polish), where the resumptive pronoun strategy results in complete wellformedness (rather than slight deviance, as is the case in English).

If nothing else is said, this type of analysis is incompatible with the present local optimization approach. To see this, consider the optimization procedure affecting CP_1 in (43-b). At this stage of the derivation, the NP that erects the island is not yet present in the structure, and the CNPC is irrelevant. Consequently, a CP_1 output that moves *who(m)* to $SpecC_1$ (because of PB, see section 3) and leaves a trace t_1 blocks a competing CP_1 output with a resumptive pronoun (the latter fatally violates AVOID PRONOUN). When CP_2 finally becomes subject to optimization, the CNPC is relevant; but now it is too late to replace the initial trace with a pronoun – if anything, we should expect a resumptive pronoun in $SpecC_1$ in (43-b).

The following steps can be taken in view of this problem. First, one might exempt chain-internal operations from the SCC, such that the foot of a chain is accessible to syntactic operations if the head of the chain is. Depending on the precise characterization of the operation Move, this might be unavoidable anyway. Second, one might take this problem to indicate that although resumptive pronouns are repair strategies to be handled in an optimality-theoretic manner, the optimization procedure does not affect syntax proper, but a post-syntactic (PF) component, as proposed in Pesetsky (1998). We will leave the choice between these (and, possibly, other) options open. On a more general note, we can conclude that, for better or worse, local optimization imposes severe restrictions on analyses in a way that global optimization does not.

6.2. Parametrization

Thus far, we have not yet addressed parametrization. As it stands, there are two options that the present approach has inherited from its two main sources (the minimalist program and optimality theory) – the strong/weak feature distinction on the one hand, and constraint re-ranking on the other. At this point, we would like to avoid taking a firm stand on the issue. What follows is a brief sketch of one possibility.

Suppose that parametrization is solely effected by constraint re-ranking, and that the class of strong and weak features is cross-linguistically uniform (e.g., [+wh] on C is universally strong, [Q] is universally weak).[51] Suppose further that the FC is

[51]It has been claimed that QR applies overtly in some languages, but it seems just as likely that these cases reflect an isomorphism constraint on QR at LF. See Lakoff (1971), Huang (1995); references cited there; and Müller (1999) and Heck (2000) for optimality-theoretic implementations. Note also

relativized with respect to feature classes (e.g., $FC_{[+wh]}$, $FC_{[\Sigma]}$). Suppose finally that there is a general constraint CHECKING ECONOMY:

(44) CHECKING ECONOMY (CE):
Avoid overt feature checking.

CE is analogous to Procrastinate in Chomsky (1995). The relative ranking of $FC_{[\alpha]}$ and CE in a language then determines whether overt α-movement is possible or not.

As it stands, languages are not systematically excluded in which feature-driven α-movement is impossible (because of a ranking CE $\gg$ $FC_{[\alpha]}$), and repair-driven α-movement is possible (because of a ranking $X_{[\alpha]} \gg$ LR, where $X_{[\alpha]}$ is a constraint that may trigger α-movement). Thus, there might be a language which does not have feature-driven scrambling (CE $\gg$ $FC_{[\Sigma]}$), but employs repair-driven scrambling in weak crossover contexts (WCC $\gg$ LR). It is not clear whether such patterns exists. English, for instance, lacks feature-driven scrambling, and it lacks repair-driven scrambling as well. Should it turn out that the availability of repair-driven α-movement is tied to the availability of feature-driven α-movement, CE could be generalized so as to prohibit movement to the domain of a head if that head is not a potential host in a language (where a head is a potential host if it can be equipped with a feature triggering movement). We will leave this question undecided.

6.3. An Alternative Approach to Optimization?

In this article, we have pursued two main goals: First, we tried to show that the phenomenon of repair-driven movement shows that LR must be assumed to be violable and ranked below other constraints: NIC, WCC, PGC $\gg$ LR triggers repair-driven scrambling; PB $\gg$ LR induces repair-driven successive-cyclic *wh*-movement; WH-R $\gg$ LR gives rise to repair-driven multiple *wh*-movement in sluicing constructions; and PAR $\gg$ LR is responsible for repair-driven QR. Second, we argued that the optimization procedures underlying the repair operations must apply locally, not globally. In view of this, we suggested to implement the analysis in a local version of optimality theory. To the extent that this enterprise was successful, it is an argument for an optimality-theoretic approach to syntax in general, and to establish this point can be viewed as a

that this implies that the EPP feature of T is strong throughout, but can be absent in German.

third goal of the article. However, at first glance it may seem that the two main goals can also be achieved in a slightly simpler way, without invoking the full optimality-theoretic machinery. The underlying reason is that there are no crucial rankings among the constraints dominating the EOC in sections 2–5.

Thus, suppose that we were to adopt the following approach. All constraints universally belong to one of three classes: those that are inviolable but do not create non-convergence if violated; those that are inviolable and lead to non-convergence if violated; and those that are violable to ensure convergence. In the first class are constraints that we have so far (more or less tacitly) assumed to be ranked below LR; these contraints do not trigger repair-driven movement. In the second class are the constraints that we have shown to be ranked above LR; these constraints can trigger repair-driven movement. In the last class is LR. It may seem as though most of the empirical evidence in sections 2–5 can be accounted for in this way. Note also that there would be no need for a constraint like the EOC on this view. Nevertheless, closer inspection reveals that this simpler approach to optimization faces severe difficulties. In what follows, we give four arguments against it.

First, the impression that an approach to repair phenomena in syntax can make do with only one type of violable constraint (viz., LR) is mainly due to the fact that we have focussed on exactly those cases where LR is violated in a well-formed sentence, because these are the cases where movement that is normally impossible is exceptionally possible after all. As soon as other repair phenomena are taken into consideration, it becomes clear that the simpler system must fail because other constraints than LR emerge as violable, and these constraints are in turn sometimes dominated by constraints that must themselves be assumed to be violable (see, e.g., the reconstruction of Grimshaw's (1997) analysis of *do*-support in section 6.1).

Second, it is unclear how to handle parametrization with respect to syntactic repair operations in the simpler system. It does not seem to be a straightforward possibility to maintain that one and the same constraint leads to non-convergence under violation in one language (thereby triggering repair), and to convergence under violation in another (thereby prohibiting repair).

Third, if it looks as though the constraints that outrank the EOC in sections 2–5 are inviolable, this is at least in part so because they often stand for more articulate systems of constraints. For instance, under the approach to parametrization sketched above, the FC is to be decomposed into a set of more specific $\text{FC}_{[\alpha]}$ constraints, some of which will typically be dominated by CE and violable in well-formed outputs in a given language. Similarly, Legendre, Smolensky, & Wilson (1998) have argued that the CED should be split up into a subhierarchy of constraints Bar^1, Bar^2, ... Bar^n with a fixed internal ranking; Bar^i is violated by a movement operation that crosses at least i barriers. Furthermore, each Bar^i subconstraint is specified along the referential/non-referential distinction ($\text{Bar}^{i[-ref]}$, $\text{Bar}^{i[+ref]}$). In such a theory of locality, more XPs can be assumed to be barriers, including vP and TP (which qualify as barriers under the most straightforward interpretation of Chomsky (1986), in which the notion of "non-L-marked category" is adopted instead of the notion "non-complement" in (15-b)); in fact, the resulting system comes close to what is arguably the null theory of locality – any XP is a barrier (see, e.g., Koster (1987)). But crucially, there must now be $\text{Bar}^{i[\pm ref]}$ constraints that are ranked below the EOC, in addition to those that are ranked above the EOC if extraction is to be possible at all. As Legendre, Smolensky, & Wilson (1998) show, even $\text{Bar}^{i[\pm ref]}$ constraints that are ranked lower than (their analogue of) the EOC can have important effects.[52]

Finally, note that we have not discussed the relative ranking of the EOC and the NIC, the WCC, and the PGC. Indeed, in all these cases there is evidence that the EOC is ranked higher. This implies that LR can be violated to fulfill NIC, WCC, and PGC; and that these constraints are themselves violable in well-formed derivations. This is incompatible with the simpler approach envisaged above.

Let us illustrate this with the NIC. As observed in note 12 of section 2.2, the NIC is not surface-true in well-formed sentences where a negative element and a *wh*-phrase are not clause-mates to begin with. This is particularly obvious if the LF position of the *wh*-phrase is below negation, as in (45-a); but it also holds if the position from which the *wh*-phrase takes scope is above negation, as in (45-c) (compare (18-b)).[53] As shown in (45-bd), repair-driven long-distance scrambling is impossible in this context.

---

[52] The constraints replacing the CED and the PIC in the analyses given in this article might be $\text{Bar}^{2[-ref]}$ (section 2) and $\text{Bar}^{3[+ref]}$ (sections 3, 4).

[53] See Stechow (1996). Note that Beck's (1996) original MNSC would predict (45-c) to be ill formed.

(45) a. dass er [$_{\text{NegP}}$ niemals weiss [$_{\text{CP}}$ was$_1$ er t$_1$ will]]
that he never knows what he wants

b. *dass er [$_{\text{NegP}}$ was$_1$ niemals weiss [$_{\text{CP}}$ t$'_1$ (dass) er t$_1$ will]]
that he what never knows that he wants

c. Wer$_1$ hat t$_1$ [$_{\text{NegP}}$ niemals gesagt [$_{\text{CP}}$ dass Maria wen$_2$ liebt]] ?
who has never said that Maria whom loves

d. *Wer$_1$ hat t$_1$ [$_{\text{NegP}}$ wen$_2$ niemals gesagt [$_{\text{CP}}$ dass Maria t$_2$ liebt]] ?
who has whom never said that Maria loves

Consider (45-cd) first. If we assume that the NIC is ranked below the EOC (hence, below the PIC), the facts fall into place. Local optimization of the embedded CP yields an optimal candidate with *wen*$_2$ in situ (CP fulfills PB without *wh*-movement, so LR is decisive). The next relevant step concerns matrix NegP optimization. If *wen*$_2$ remains in situ, the NIC is violated; but if *wen*$_2$ raises to the matrix NegP domain, the higher-ranked PIC is fatally violated (and Ø violates the higher-ranked EOC). Hence, (45-c) is optimal even though it violates the NIC. Note that we thus have yet another argument against a global approach to optimization: If the complete structures of (45-cd) were evaluated in one step, (45-c)'s NIC violation would wrongly be predicted to be fatal as opposed to (45-d)'s early LR violation incurred by repair-driven *wh*-movement to the embedded SpecC (so as to avoid a later PIC violation).

Turning to (45-ab), the main difference is that the embedded *wh*-phrase *was*$_1$ undergoes feature-driven *wh*-movement to the embedded SpecC position; hence, it should in principle be available for further repair-driven movement without a PIC violation. However, it seems reasonable to assume that if such NIC-driven scrambling to NegP applies, a high-ranked constraint is fatally violated that demands items that have reached their LF target position in overt syntax to remain there during the rest of the derivation (see Chomsky (1995) and Collins (1997), among others).[54] Note finally that we expect one case of repair-driven long-distance *wh*-scrambling triggered by the NIC to be legitimate: In (46), *wh*-scrambling to NegP is forced during NegP optimization under present assumptions. But since such scrambling can only be an intermediate step, followed by subsequent *wh*-movement (otherwise, PB and the PIC would not have permitted this operation in the first place), the strict ban on *wh*-phrases showing

[54]Interestingly, this latter constraint may also be violable in selected contexts; see Reis & Rosengren (1992) on the (somewhat marginal) phenomenon of *wh*-imperatives in German.

up in long-distance scrambling positions in German can be maintained.

(46) Wen$_1$ hat sie [$_{\text{NegP}}$ t$_1''$ niemals gesagt [$_{\text{CP}}$ t$_1'$ dass Maria t$_1$ liebt]] ?
whom has she never said that Maria loves

The case is similar with the WCC. First, recall from section 2.3 that the WCC must be violable non-fatally in the derivation if a potential binder is not yet present. Second, bound-variable pronouns can show up in positions that are accessible to parasitic gaps. In English, which does not have the "mixed" movement type scrambling (L-related but A-bar), this suggests that there is a way to avoid a fatal WCC violation here (see Stowell (1991)). Interestingly, even though German has scrambling, a bound variable pronoun does not trigger repair-driven movement of a *wh*-phrase in this context; see (47-a). Furthermore, the bound variable pronoun in (47-b) does not give rise to ungrammaticality despite lacking an overt binder (see Sternefeld (1993)).

(47) a. Wann hat die Maria [$_{\text{CP}}$ ohne es$_1$ zu lesen] dem Fritz welches Buch$_1$ zurückgegeben ?
when has ART Maria without it to read ART Fritz$_{dat}$ which book$_{acc}$ returned

b. [$_{\text{CP}}$ Dass sein$_1$ Fahrrad da gestohlen werden könnte]$_2$ hat t$_2$ keinem$_1$ eine schlaflose Nacht bereitet
that his bicycle there stolen be could has no-one$_{dat}$ a sleepless night$_{acc}$ given

The correct descriptive generalization appears to be that if there is no potential binder for a bound variable pronoun within the latter's minimal CP, the WCC can be violated in German, even at the end of an overt derivation. We will not account for this here, but the data seem to lend themselves to an optimality-theoretic analysis.

Finally, the discussion of (11-a) in section 2.4 has shown that the PGC can be violated during the derivation; furthermore, English sentences like (48) (see Nissenbaum (1999)) suggest that non-fatal PGC violations can show up in surface forms.

(48) ?Which paper$_1$ did you assign t$_1$ to which student$_2$ [$_{\text{CP}}$ after talking about t$_1$ with an advisor of t$_2$] ?

More generally, then, we can conclude that the approach that implements the findings in sections 2–5 in a non-optimality-theoretic way is hardly tenable: LR is violable in favour of the NIC, the WCC, and the PGC, which are themselves violable in favour of other constraints.

References

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

Alexiadou, Artemis. 1997. *Adverb Placement.* Amsterdam: Benjamins.

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

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. 1999. Optionality and Ineffability. Ms., Harvard University & UMass., Amherst. To appear in *Optimality-Theoretic Syntax*, ed. Géraldine Legendre, Jane Grimshaw & Sten Vikner. Cambridge, Mass.: MIT Press.

Bayer, Josef. 1990. Notes on the ECP in English and German. *Groninger Arbeiten zur Germanistischen Linguistik* 30:1-51.

Beck, Sigrid. 1996. *Wh-Constructions and Transparent Logical Form.* Doctoral dissertation, Universität Tübingen.

Bennis, Hans, & Teun Hoekstra. 1984. Gaps and Parasitic Gaps. *The Linguistic Review* 4:29-87.

Bobaljik, Jonathan. 1995. *Morphosyntax: The Syntax of Verbal Inflection.* Doctoral dissertation, MIT, Cambridge, Mass.

Bolinger, Dwight. 1978. Asking More Than One Thing At A Time. In *Questions*, ed. Henry Hiz, 107-150. Dordrecht: Reidel.

Bošković, Željko. 1999. What is Special about Multiple Wh-Fronting? Ms., University of Connecticut. To appear in *Proceedings of NELS 30.*

Bošković, Željko, & Howard Lasnik. 1999. How Strict is the Cycle? *Linguistic Inquiry* 20:691-703.

Brody, Michael. 1995. *Lexico-Logical Form.* Cambridge, Mass.: MIT Press.

Browning, M.A. 1991. Bounding Conditions on Representations. *Linguistic Inquiry* 22:541-562.

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. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.

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. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1986. *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1991. Some Notes on Economy of Derivation and Representation. In *Principles and Parameters in Comparative Grammar*, ed. Robert Freidin, 417-454. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1995. Categories and Transformations. Chapter 4 of *The Minimalist Program*, 219-394. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1998. Minimalist Inquiries. Ms., MIT, Cambridge, Mass.

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.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads*. Oxford: Oxford University 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.

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. 1995. A Minimalist Approach to Free Constituent Order. Ms., Universität Potsdam.

Fanselow, Gisbert. 1998. Features, Theta-Roles, and Free Constituent Order. Ms., Universität Potsdam. To appear in *Linguistic Inquiry.*

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.

Felix, Sascha. 1983. Parasitic Gaps in German. *Groninger Arbeiten zur Germanistischen Linguistik* 22:1-46.

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

Fox, Danny. 1995. Economy and Scope. *Natural Language Semantics* 3:283-341.

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

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

Frank, Robert, Young-Suk Lee, & Owen Rambow. 1995. Scrambling, Reconstruction and Subject Binding. *Rivista di Grammatica Generativa* 21:67-106.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, & Ivan Sag. 1985. *Generalized Phrase Structure Grammar.* Oxford: Blackwell.

Grewendorf, Günther. 1988. *Aspekte der deutschen Syntax.* Tübingen: Narr.

Grewendorf, Günther. 1999. Multiple Wh-Fronting. Ms., Universität Frankfurt/Main.

Grewendorf, Günther, & Joachim Sabel. 1999. Scrambling in German and Japanese. *Natural Language and Linguistic Theory* 17:1-65.

Grimshaw, Jane. 1997. Projection, Heads, and Optimality. *Linguistic Inquiry* 28:373-422.

Grimshaw, Jane. 1999. Clause Structure and Competition in Syntax. Ms., Rutgers University.

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.

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. 2000. Towards a Superior Account of Superiority. In *Wh-Scope Marking*, ed. Uli Lutz, Gereon Müller, & Arnim von Stechow, 231-248. Amsterdam: Benjamins.

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.

Heim, Irene. *The Semantics of Definite and Indefinite Noun Phrases.* Doctoral dissertation, University of Massachusetts, Amherst.

Heim, Irene, & Angelika Kratzer. 1998. *Semantics in Generative Grammar.* Oxford: Blackwell.

Hermans, Ben, & Mark van Oostendorp (ed.). 2000. *The Derivational Residue in Phonological Optimality Theory.* Amsterdam: Benjamins.

Hornstein, Norbert. 2000. Is the Binding Theory Necessary? Ms., University of Maryland. (Chapter 5 of forthcoming book.)

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.

Jacobson, Pauline. 1997. Where (If Anywhere) is Transderivationality Located? In *The Limits of Syntax*, ed. Peter Culicover & Louise McNally, 303-336. San Diego: Academic Press.

Johnson, Kyle. 1998. Ways to Hide Quantifier Raising. Ms., UMass., Amherst.

Kayne, Richard. 1998. Overt vs. Covert Movements. *Syntax* 1:128-191.

Kitahara, Hisatsugu. 1997. *Elementary Operations and Optimal Derivations.* Cambridge, Mass.: MIT Press.

Koster, Jan. 1987. *Domains and Dynasties.* Dordrecht: Foris.

Lakoff, George. 1971. On Generative Semantics. In *Semantics*, ed. Danny Steinberg & Leon Jakobovits, 232-296. Cambridge: Cambridge University 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. 1999. Harmonic Serialsm and Parallelism. Ms., University of Massachusetts, Amherst.

McCloskey, James. 1979. *Transformational Syntax and Model Theoretic Semantics. A Case Study in Modern Irish.* Dordrecht: Reidel.

Mahajan, Anoop. 1990. *The A/A-bar Distinction and Movement Theory.* Doctoral dissertation, MIT, Cambridge, Mass.

Merchant, Jason. 1999. *The Syntax of Silence.* Doctoral dissertation, UCSC.

Müller, Gereon. 1998. *Incomplete Category Fronting.* Dordrecht: Kluwer.

Müller, Gereon. 1999. Order Preservation, Parallel Movement, and the Emergence of the Unmarked. To appear in *Optimality Theoretic Syntax*, ed. Géraldine Legendre, Jane Grimshaw & Sten Vikner. Cambridge, Mass.: MIT Press.

Müller, Gereon. 2000. Shape Conservation and Remnant Movement. Ms., IDS Mannheim. To appear in *Proceedings of NELS 30.* Amherst, Mass.: GLSA.

Müller, Gereon, & Wolfgang Sternefeld. 1993. Improper Movement and Unambiguous Binding. *Linguistic Inquiry* 24:461-507.

Nissenbaum, Jon. 1999. Covert Movement and Parasitic Gaps. Ms., MIT, Cambridge, Mass. To appear in *Proceedings of NELS 30.* Amherst, Mass.: GLSA.

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. 1999. *Phrasal Movement and Its Kin.* Ms., MIT, Cambridge, Mass.

Prince, Alan & Paul Smolensky. 1993. *Optimality Theory. Constraint Interaction in Generative Grammar*. Ms., Rutgers University.

Reinhart, Tanya. 1983. *Anaphora and Semantic Interpretation.* London: Croom Helm.

Reis, Marga, & Inger Rosengren. 1992. What do Wh-Imperatives Tell us About Wh-Movement? *Natural Language and Linguistic Theory* 10:79-118.

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.

Roberts, Ian. 1998. Have/Be Raising, Move F, and Procrastinate. *Linguistic Inquiry* 29:113-125.

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.

Sabel, Joachim. 1998. *Principles and Parameters of Wh-Movement.* Habilitation thesis, Universität Frankfurt/Main.

Sabel, Joachim. 2000. A Unified Analysis of Wh- and Non-Wh-Expletives. Ms., Universität Frankfurt/Main. To appear in *Proceedings of WCCFL 19.* 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. 1998. West Germanic "Infinitivus Pro Participio" (IPP) Constructions in Optimality Theory. In *Proceedings of Console VII*, ed. Tina Cambier-Langeveld, Anikó Lipták, Michael Redford & Erik Jan van der Torre, 229-244. Leiden: SOLE.

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. 1996. Partial Wh-Movement and Logical Form. In *Papers on Wh-Scope Marking*, ed. Uli Lutz & Gereon Müller, 1-36. Stuttgart & Tübingen: Bericht 76 des SFB 340.

Sternefeld, Wolfgang. 1993. Anaphoric Reference. In *Syntax. An International Handbook of Contemporary Research*, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, & Theo Vennemann, 940-966. Berlin: de Gruyter.

Sternefeld, Wolfgang. 1997. Comparing Reference Sets. In *Economy in Linguistic Theory*, ed. Chris Wilder, Hans-Martin Gärtner & Manfred Bierwisch, 81-114. Berlin: Akademieverlag.

Sternefeld, Wolfgang. 1998. Reciprocity and Cumulative Predication. *Natural Language Semantics* 6:303-337.

Stowell, Tim. 1991. Adjuncts, Arguments, and Crossover. Ms., UCLA.

Takahashi, Daiko. 1993. Movement of Wh-Phrases in Japanese. *Natural Language and Linguistic Theory* 11:655-678.

Takahashi, Daiko. 1994. Sluicing in Japanese. *Journal of East Asian Linguistics* 3:265-300.

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

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.

Webelhuth, Gert. 1992. *Principles and Parameters of Syntactic Saturation.* Oxford: Oxford University Press.

Wexler, Ken, & Peter Culicover. 1980. *Formal Principles of Language Acquisition.* Cambridge, Mass.: MIT Press.

Wilson, Colin. 1998. Bidirectional Optimization and the Theory of Anaphora. Ms., Johns Hopkins University. To appear in *Optimality Theoretic Syntax*, ed. Géraldine Legendre, Jane Grimshaw & Sten Vikner. Cambridge, Mass.: MIT Press.

Wunderlich, Dieter. 2000. The Force of Lexical Case: German and Icelandic Compared. Ms., Universität Düsseldorf.