Shape Conservation and Remnant Movement

Gereon Müller

Universität Tübingen

1. Introduction

Remnant movement is movement of an XP β from which extraction of α has taken place; cf. (1). This phenomenon has been argued to support a derivational approach to syntax (cf. Chomsky (1998)): Since remnant movement creates an unbound α trace that is separated from its antecedent by an XP in non-selected position (i.e., a barrier), the wellformedness of the resulting structure is unexpected under representational approaches that require proper binding of traces and check locality constraints at S-structure; but nothing is wrong with (1) under a strictly derivational approach in which proper binding is replaced by strict cyclicity and locality is checked directly after each movement operation.

(1) $[_{\beta_2}$ ... t_1 ...] ... [... α_1 ... [... t_2 ...]]

Remnant movement has been suggested for two different kinds of constructions. On the one hand, Thiersch (1985) and den Besten & Webelhuth (1987; 1990)) have argued that cases of incomplete category fronting like (2-a) in German should be analyzed as involving scrambling of NP_1 and remnant VP_2 topicalization.[1] On the other hand, it has recently been proposed that remnant movement is a much more general phenomenon that also underlies certain other constructions where this may not be immediately obvious. Most notably, Kayne (1998) analyzes constructions like (2-b) in English as involving obligatory overt negative NP_1 preposing followed by TP-internal remnant VP_2 fronting.[2] Henceforth, I will refer to the two constructions as "primary" and

For comments and discussion, I would like to thank Artemis Alexiadou, Darcy Berry, Jane Grimshaw, Fabian Heck, Kyle Johnson, and the audiences of the workshop on Remnant Movement & Feature Movement (Universität Potsdam, July 1999) and NELS 30 (Rutgers University, October 1999). The research reported here was supported by DFG grants MU 1444/1-1,2-1.

[1]Also see Stechow & Sternefeld (1988), Bayer (1996), Müller (1998), Grewendorf & Sabel (1999), and references cited in these works.

[2]Also see den Dikken (1996), Hinterhölzl (1997), Ordóñez (1997), Johnson (1998), Koopman & Szabolcsi (1999), Noonan (1999) on related analyses for other constructions. In what follows, I will focus on Kayne's analysis of negative NP preposing. What I will have to say can straightforwardly be extended to Johnson's and Noonan's analyses. The other cases may require additional assumptions.

"secondary" remnant movement, respectively.

(2) a. [$_{VP_2}$ t_1 Gelesen] hat das $Buch_1$ keiner t_2
read has the book no-one
"No-one read the book."
b. John [$_{VP_2}$ reads t_1] no $novels_1$ t_2

The goal of this paper is twofold. In section 2, I will show that the two constructions exhibit radically different properties. In section 3, I will argue that a unified analysis is possible despite these differences if we assume that shape conservation (Williams (1999)) can be a trigger for movement, in addition to feature checking (Chomsky (1995)). In particular, we will see that whereas primary remnant movement is feature-driven, secondary remnant movement is a repair strategy that is triggered by shape conservation. This latter idea will be implemented in a restrictive model of optimality theory ("local optimization"), for which I will present empirical support.

2. The Properties of Primary and Secondary Remnant Movement

2.1. Independent Availability

In primary remnant movement constructions, movement of both β_2 and α_1 in (1) must be independently available. Thus, German remnant VP topicalization as in (2-a) presupposes that VP topicalization and NP scrambling are independent options in the language, which indeed they are:

(3) a. [$_{VP_2}$ Das $Buch_1$ gelesen] hat keiner t_2
the book read has no-one
"No-one read the book."
b. daß das $Buch_2$ keiner [$_{VP}$ t_2 gelesen] hat
that the book no-one read has
"that no-one read the book."

Similarly, the English primary remnant movement construction in (4-a) relies on the independent existence of VP topicalization and NP raising of the subject, as in (4-bc).

(4) a. [$_{VP_2}$ Criticized t_1 by his boss] $John_1$ has never been t_2
b. [$_{VP_2}$ Criticize John] he wouldn't t_2
c. $John_1$ has never been [$_{VP_2}$ criticized t_1 by his boss]

In line with this, English lacks the counterpart to the German remnant movement construction in (2-a) for the simple reason that it does not have scrambling:

(5) a. *[$_{VP_2}$ Kicked t_1] John never has the dog_1 t_2
b. [$_{VP_2}$ Kicked the dog_1] John never has t_2
c. *John never has the dog_1 [$_{VP_2}$ kicked t_1]

In contrast, in secondary remnant movement constructions like (6-a) (= (2-b)), movement of neither α_1 nor β_2 is independently available. This is clear for negative NP preposing; cf. (6-c). Given that independent VP_2 fronting in (6-b) would be string-vacuous, the question arises of whether this is an option. Since Kayne assumes that the "more emphatic, less neutral character" of sentences like (6-a) "must be correlated with VP-movement," and since it is unclear which feature could trigger TP-internal VP fronting in (6-b), we may conclude that it is not.[3] Consequently, none of the two movement operations in (6-a) is independently available in secondary remnant movement constructions.

(6) a. John [$_{\mathrm{VP}_2}$ reads t_1] no novels$_1$ t_2
b. *John [$_{\mathrm{VP}_2}$ likes that novel$_1$] t_2
c. *John no novels$_1$ [$_{\mathrm{VP}_2}$ reads t_1]

2.2. Secondary Object Fronting

Double object constructions reveal a second difference. Primary remnant VP topicalization in German may carry along or strand (by scrambling) any of the two objects:

(7) a. [$_{\mathrm{VP}_2}$ t_1 Ein Buch zum Geburtstag geschenkt] hat sie dem Jason$_1$ t_2
a book$_{acc}$ for the birthday given has she ART Jason$_{dat}$
"She gave Jason a book as a birthday present."
b. [$_{\mathrm{VP}_2}$ Dem Jason$_1$ t_3 zum Geburtstag geschenkt] hat sie ein Buch$_3$ t_2
ART Jason$_{dat}$ for the birthday given has she a book$_{acc}$
c. [$_{\mathrm{VP}_2}$ t_1 t_3 Zum Geburtstag geschenkt] hat sie dem Jason$_1$ ein Buch$_3$ t_2
for the birthday given has she ART Jason$_{dat}$ a book$_{acc}$

In contrast to this, whether secondary remnant VP fronting carries along an NP in a double object construction or strands it prior to VP fronting depends on whether the pre-movement order is maintained. If the negative NP is the first object, the second object cannot be fronted together with the verb, but must leave the VP by an earlier operation that I will call "secondary object fronting" (indicated here by underlining); this operation targets a position below that of the negative NP, thereby restoring the pre-movement order:[4]

(8) a. *John [$_{\mathrm{VP}_2}$ gave t_1 to Mary$_3$] no books$_1$ t_2
b. John [$_{\mathrm{VP}_2}$ gave t_1 t_3] no books$_1$ $\underline{\text{to Mary}_3}$ t_2
c. *John [$_{\mathrm{VP}_2}$ gave t_1 a book$_3$] no-one$_1$ t_2
d. John [$_{\mathrm{VP}_2}$ gave t_1 t_3] no-one$_1$ $\underline{\text{a book}_3}$ t_2

If, on the other hand, the negative NP is the second object, the first object must be fronted together with the verb, and cannot undergo secondary object fronting:

(9) a. John [$_{\mathrm{VP}_2}$ gave Mary$_1$ t_3] no books$_3$ t_2
b. *John [$_{\mathrm{VP}_2}$ gave t_1 t_3] no books$_3$ $\underline{\text{Mary}_1}$ t_2

2.3. Extraction

Both the remnant XP β_2 and the antecedent of the unbound trace α_1 in (1) are barriers for further extraction in primary remnant movement constructions. This is

[3] Kayne states that negative NP preposing will "in turn ... require the ... VP to prepose," which suggests that TP-internal VP fronting is not independently available in English.

[4] Derivations of the type in (8-a) have sometimes been argued to underlie heavy NP shift; but this issue is clearly not at play in the case at hand.

a standard freezing effect that is expected if (a) moved items end up in non-selected positions, where they are barriers (cf. Cinque (1990) vs. Lasnik & Saito (1992)), and (b) strict cyclicity ensures that extraction from these items cannot take place before they undergo movement to a lower position (cf. Chomsky (1995) and references cited there). This is shown for β in (10-ab), and for α in (10-c) (barriers are underlined).

(10) a. *Wem$_3$ denkst du [$_{CP}$ [$_{VP_2}$ t$_3$ t$_1$ gegeben] hat das Buch$_1$ keiner t$_2$] ?
whom think you given has the book no-one
"To whom do you think that no-one gave a book?"
b. *Children$_3$ I think that [$_{CP}$ [$_{VP_2}$ written t$_1$ for t$_3$] those books$_1$ could not possibly be t$_2$]
c. *[$_{VP_2}$ t$_1$ Gerechnet] hat da$_3$ gestern [$_{PP_1}$ t$_3$ mit] wieder keiner t$_2$
counted has there yesterday with again no-one
"Again, no-one reckoned with it yesterday."

In contrast, neither β_2 nor α_1 is a barrier for further extraction in secondary remnant movement constructions; cf. (11-a) and (11-b), respectively.[5] Given the interaction of barriers theory and strict cyclicity, this anti-freezing effect is a priori unexpected.

(11) a. Which book$_3$ did John [$_{VP_2}$ give t$_3$ t$_1$] [$_{PP_1}$ to no-one] t$_2$?
b. About Nixon$_3$ John [$_{VP_2}$ read t$_1$] [$_{NP_1}$ only one book t$_3$] t$_2$

2.4. Movement Types

It has often been noted that not all movement types seem to be able to affect (primary) remnant XPs equally well, the crucial distinction being that between middle field-external and middle field-internal movement operations. E.g., whereas topicalization of a remnant infinitival VP is possible in German (cf. (12-a)), scrambling of the same remnant VP leads to ungrammaticality (cf. (12-b)).[6]

(12) a. [$_{VP_2}$ t$_1$ Zu lesen] hat das Buch$_1$ keiner t$_2$ versucht
to read has the book no-one tried
"No-one tried to read the book."
b. *daß [$_{VP_2}$ t$_1$ zu lesen] das Buch$_1$ keiner t$_2$ versucht hat
that to read the book no-one tried has
"that no-one tried to read the book."

Again, things are different with secondary remnant movement. Indeed, secondary remnant VP$_2$ fronting is not just permitted to target a middle field-internal (post-subject) landing site (cf. (13-a) = (2-b)); it is required to do so (cf. the failed attempt at topicalization in this context in (13-b)).

(13) a. John [$_{VP_2}$ reads t$_1$] no novels$_1$ t$_2$
b. *[$_{VP_2}$ Reads t$_1$] (I think that) John t$'_2$ no novels$_1$ t$_2$

To sum up, we have seen that primary and secondary remnant movement constructions differ radically. One might want to take this to indicate that one of the two approaches

[5] Note that Kayne (1998) treats *only*-phrases on a par with negative NPs.
[6] See Fanselow (1991), Frank, Lee & Rambow (1992), Haider (1993), Grewendorf & Sabel (1999).

should be abandoned. Given that both approaches have their virtues, I will not draw this conclusion here. Rather, I will develop a unified approach that explains the diverging properties of primary and secondary remnant movement constructions by distinguishing between feature-driven movement and repair-driven movement.

3. A Unified Approach

3.1. Shape Conservation and Local Optimization

All movement operations can plausibly be viewed as being feature-driven in primary remnant movement constructions. Thus, (14-a) involves a combination of NP raising (triggered by the EPP feature) and VP topicalization (triggered by a topic feature); and (14-b) has NP scrambling (which I will here assume to be triggered by a specific scrambling feature[7]) followed by VP topicalization (again triggered by a topic feature). In contrast, in secondary remnant movement constructions, it looks as though only one movement operation is feature-driven; in the construction at hand, this is negative NP preposing. All other movement operations are parasitic – they depend on the first operation having taken place. The absence of a feature that triggers secondary remnant movement and secondary object fronting is illustrated in (14-c).

(14) a. [$_{VP_2}$ Criticized t_1 by his boss]-[top] $John_1$[-D] has never been t_2
b. [$_{VP_2}$ t_1 t_3 Zum Geburtstag geschenkt]-[top] hat sie dem $Jason_1$-[scr]
for the birthday given has she ART Jason
ein $Buch_3$-[scr] t_2
a book
c. John [$_{VP_2}$ gave t_1 t_3]-Ø no $books_1$-[neg] to $Mary_3$-Ø t_2

Then, given constraints like the FEATURE CONDITION (FC) in (15) and LAST RESORT (LR) in (16) (cf. Chomsky (1995)), a problem arises: Some instances of movement in secondary remnant movement constructions are not triggered by FC, and they thus violate LR. Consequently, a different trigger must be involved, and respecting this trigger must permit a violation of LR, which is otherwise impossible. Thus, secondary remnant movement emerges as a repair strategy: Exceptionally, LR can be violated so as to prevent even greater damage.

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

(16) LAST RESORT (LR):
Overt movement must result in checking of a strong feature.

I would like to suggest that the trigger in question is the SHAPE CONSERVATION (SC) constraint that is proposed on independent grounds in Williams (1999). For the sake of concreteness, I will assume that SC basically demands that the shape of predicate

[7]See Sauerland (1997) and Grewendorf & Sabel (1999). Arguably, there is more than one possible trigger for scrambling in German, and this fact might be formally encoded by assigning a complex internal structure to the scrambling feature. This would not affect the issue at hand, though.

phrases, or vPs, must be preserved in derivations:[8]

(17) SHAPE CONSERVATION (SC):
Feature checking in the domain of a head Y must not change the linear order of lexical items established in vP within YP.

The English vP shape that will be relevant is completely standard, and given in (18).[9]

(18) $[_{vP}$ NP_1 $[_{v'}$ v+V $[_{VP}$ NP_2 $[_{V'}$ t_V $\{NP_3/PP_3\}$]]]]

The analysis then relies on three assumptions. First, feature-driven movement of the negative NP_1 in (14-c) ends up in the specifier of a functional head Neg that bears a strong [neg] feature. Given SC, it follows that vP_2-[Ø] (and not VP, as assumed thus far) must be fronted to an outer specifier of Neg (i.e., to a position that precedes NP_1-[neg] within the same projection), as an instance of repair-driven movement.[10] It also follows that repair-driven movement of PP_3-[Ø] in (14-c) must end up in an inner specifier of the very same domain, NegP.

The second assumption concerns a qualification. Evidently, whereas negative NP preposing requires vP shape conservation, other movement operations do not. This is obvious in the case of *wh*-movement in English: Checking of [wh] with an object NP in the C domain does not trigger repair-driven movement of TP_4 to an outer specifier of C; cf. (19-a) vs. (19-b) (the latter would correspond to a *wh*-in situ language in which there is evidence that *wh*-movement is in fact overt).

(19) a. $What_1$-[wh] did $[_{TP_4}$ you_3 $[_{vP_2}$ t_3 see t_1]] ?
b. *$[_{TP_4}$ You $[_{vP_2}$ t see t_1]]-Ø $what_1$-[wh] did t_4 ?

This means that SC either does not hold for *wh*-movement in English (and many other movement operations), or that it holds, but in a much weaker form. I will draw the second conclusion here and suggest that SC is to be split up, and made sensitive to feature classes: Features like [neg] obey a strong SC constraint that permits a violation of LR (cf. the references in footnote 2 for other possible features with this property), whereas features like [wh] obey only a weaker SC constraint that does not permit a violation of LR (other features in this class include [top] and [scr]). It is tempting to conclude that the relevant distinction is between features that trigger A-movement and features that trigger A-bar movement. Indeed, most cases of NP raising to SpecT will automatically satisfy SC. Successive-cyclic NP raising may initially look problematic; but assuming that the absence of intermediate vP projections is exactly the property that makes such raising possible, SC is respected in this case as well.

[8]For predecessors of this constraint, see Lakoff (1971), Kroch (1974), Huang (1982), Reinhart (1983), Lasnik & Saito (1992), Watanabe (1992), Haegeman (1995), Meinunger (1995), and Müller (1997). In general, these constraints are defined in structural rather than linear terms. This would not be sufficient for the present analysis – SC-driven movement restores linear order, not c-command.

[9]Whether NP_2 occupies SpecV as a result of movement or base-generation in dative shift constructions is immaterial for present purposes – as long as there is no vP yet, all movement (including V-to-v raising) satisfies SC vacuously. As for German, I will postulate essentially the same structure, the only difference being that v+V is right-peripheral in vP.

[10]This position follows typical adverb positions; cf. the evidence against V-to-T raising in English.

Similarly, Scandinavian object shift is well known for its rigid order preservation.[11] Thus, at least for present purposes, we can assume that there are only two general SC constraints – SC_A (including [D], [Neg]) and $SC_{\overline{A}}$ (including [wh], [top], [scr]).

Third, since the analysis involves the notion of repair and depends on the violability and ranking of constraints, it lends itself to an optimality-theoretic implementation. The implicit ranking just sketched can be made explicit as follows (the ranking of FC and SC_A is not determined by the evidence discussed here):

(20) $\{FC, SC_A\} \gg LR \gg SC_{\overline{A}}$

Repair phenomena are certainly among those constructions where optimality theory has proven most successful, and the notion of repair itself can be given a precise characterization in this approach: A repair is a competition in which the optimal candidate incurs an (otherwise fatal) violation of a high-ranked constraint C_i in order to respect an even higher-ranked constraint C_j. However, it is clear that standard global optimization procedures as laid out in Prince & Smolensky (1993) induce complexity of a type that more recent versions of the minimalist program manage to avoid. In view of this, and deviating from the vast majority of work in optimality-theoretic syntax, I would like to suggest that syntactic optimization is local, not global, and takes place repeatedly throughout the derivation.[12]

For the sake of concreteness, suppose that syntactic derivations proceed as in Chomsky (1995): Merge and Move alternate, with each XP a cyclic node. Crucially, the subderivation from one cyclic node α to the next cyclic node β ($\alpha \rightarrow \beta$) is subject to input/output optimization. An XP is optimal if the subderivation that creates it best satisfies an ordered set of violable constraints and respects inviolable constraints (like strict cyclicity), which can be conceived of as parts of the definitions of Merge and Move. Thus, an XP that is the optimal output of a subderivation forms the input for the next subderivation, together with a new lexical item Y (and possibly another optimal ZP if SpecY is to be filled by Merge). Optimization determines the new optimal output YP, which in turn shows up in the input of the next subderivation, and so on, until the optimal root is reached.

Based on these assumptions, the differences between primary and secondary remnant movement can now be accounted for.

3.2. Independent Availability and Secondary Object Fronting Revisited

Consider again a typical secondary remnant movement example like (21-c):

(21) a. $[_{vP_2}$ John$_3$ reads $[_{VP_3}$ t_V no novels$_1$]] + Neg →
b. $[_{NegP}$ $[_{vP_2}$ John$_3$ reads t_1] $[_{Neg'}$ no novels$_1$ $[_{Neg'}$ Neg t_2]]] + T →

[11] Multiple object shift strictly preserves vP shape, and it seems possible to reanalyze double object NP_1-Pronoun$_2$ orders as the result of feature-driven pronominal object shift accompanied by by SC-driven NP_1 fronting. See Müller (1997), Williams (1999), and references cited there.

[12] Versions of multiple optimization in phonology are discussed in Prince & Smolensky (1993, ch.2) and McCarthy (1999). Heck (1999) and Wilson (1999) assume multiple (but non-local) optimization in syntax – three times per sentence in the former case (to determine D-structure, S-structure, and LF), and twice in the latter case (to determine interpretation and syntactic expression). Also note that there is a trade-off: Whereas there is more complexity with global optimization than there is with local optimization, local optimization in turn requires a large number of optimization procedures.

c. [$_{\text{TP}}$ John$_3$ T [$_{\text{NegP}}$ [$_{\text{vP}_2}$ t$_3$ reads t$_1$] [$_{\text{Neg}'}$ no novels$_1$ [$_{\text{Neg}'}$ Neg t$_2$]]]]

What we want to derive is that NP$_1$ moves to SpecNeg to check a strong [neg] feature and thereby respect FC, and that vP$_2$ then raises to an outer SpecNeg position without feature checking in order to respect SC$_A$, even if this violates LR. The optimization procedure that ensures this outcome is the one that takes the optimal vP$_2$ in (21-a) and Neg as inputs and creates a set of NegPs as output candidates. The optimal NegP is the one in (21-b), which violates LR but respects FC and SC$_A$, and thus has a better constraint profile than its competitors, which fatally violate either FC (by not applying negative NP$_1$ preposing) or SC$_A$ (by not applying secondary remnant vP$_2$ movement). The local competition is shown in tableau T$_1$.

T$_1$: vP → NegP Optimization: Secondary remnant movement

Input: Neg, [$_{vP_2}$ John$_3$ reads [$_{VP}$ t$_V$ no novels$_1$]]	FC	SC$_A$	LR	SC$_{\overline{A}}$
⇒O$_1$: [$_{NegP}$ [$_{vP_2}$ John$_3$ reads [$_{VP}$ t$_V$ t$_1$]] no novels$_1$ Neg t$_2$]			*	
O$_2$: [$_{NegP}$ no novels$_1$ Neg [$_{vP_2}$ John$_3$ reads [$_{VP}$ t$_V$ t$_1$]]]		*!		
O$_3$: [$_{NegP}$ – Neg [$_{vP_2}$ John$_3$ reads [$_{VP}$ t$_V$ no novels$_1$]]]	*!			
O$_4$: [$_{NegP}$ [$_{vP_2}$ John$_3$ reads [$_{VP}$ t$_V$ no novels$_1$]] Neg t$_2$]	*!		*	
O$_5$: [$_{NegP}$ reads$_4$ no novels$_1$ Neg [$_{vP_2}$ John$_3$ t$_4$ [$_{VP}$ t$_V$ t$_1$]]]		*!	*	

The optimal NegP O$_1$ is then merged with T, and subsequent NegP → TP optimization produces the expected result: The best subderivation fronts the subject NP$_3$ to SpecT and has v+V in situ (this output violates none of the constraints at hand). Note that *only* O$_1$ can be in the input for the next optimization procedure, not O$_2$–O$_5$ or other suboptimal outputs. It is this property that minimizes complexity: Under standard, global optimization, all these suboptimal outputs would have to be continued to the end (in representational terms: considered as substructures of the whole sentence) and would thereby give rise to exponential growth of the candidate set.

In addition to this conceptual difference, local optimization turns out to also yield a desirable empirical difference. In the present system, it is clear that V raising is not an alternative to remnant vP movement: Local V raising to SpecNeg as in O$_5$ does not satisfy SC$_A$, leading to VOS instead of SVO order; and non-local V-to-T raising can never satisfy SC$_A$ within NegP. In contrast, under global optimization there would be no SC$_A$ violation, due to subsequent NP$_3$ raising to SpecT (which ultimately restores SVO order), and repair-driven V raising might incorrectly (given adverb placement facts) be permitted along with (or instead of) remnant vP movement.[13]

Next consider the case where secondary remnant movement is accompanied by secondary object fronting, as in the double object construction (22-c).

(22) a. [$_{\text{vP}_2}$ John$_4$ gave [$_{\text{VP}}$ no books$_1$ t$_V$ to Mary$_3$]] + Neg →
b. [$_{\text{NegP}}$ [$_{\text{vP}_2}$ John$_4$ gave [$_{\text{VP}}$ t$_1$ t$_V$ t$_3$]] no books$_1$ to Mary$_3$ Neg t$_2$] + T →
c. [$_{\text{TP}}$ John$_4$ T [$_{\text{NegP}}$ [$_{\text{vP}_2}$ t$_4$ gave [$_{\text{VP}}$ t$_1$ t$_V$ t$_3$]] no books$_1$ to Mary$_3$ Neg t$_2$]]

[13]Of course, V raising could still independently be filtered out by stipulating a higher-ranked constraint that, e.g., bans movement of a lexical category (cf. Grimshaw (1997), Vikner (1999), and Kayne (1998, fn. 11), who notes: "The lexical verb in English cannot raise by head movement, yet it must move, consequently the whole VP moves"). Still, the point remains that local and global optimization differ empirically, and the former approach offers a simpler account in the case at hand.

Again, the important subderivation is the step from vP in (22-a) to NegP in (22-b), and essentially the same reasoning applies as before. The optimal NegP is one in which NP_1 moves to SpecNeg to check the [neg] feature and thereby respect FC, and PP_3 and vP_2 undergo repair-driven movement to inner and outer specifiers of NegP, respectively, to respect SC_A. This incurs two violations of LR, but as shown in tableau T_2, all competing subderivations fatally violate higher-ranked constraints. Note in particular that O_1 blocks O_5 as suboptimal; O_5 has secondary remnant vP movement but fails to apply secondary object fronting.

T_2: vP → NegP Optimization: Secondary remnant movement and object fronting

Input: Neg, $[_{vP_2}$ $John_4$ gave $[_{VP}$ no $books_1$ t_V to $Mary_3$]]	FC	SC_A	LR	$SC_{\overline{A}}$
⇒O_1: $[_{NegP}$ $[_{vP_2}$ J_4 gave $[_{VP}$ t_1 t_V t_3]] no $books_1$ to M_3 Neg t_2]			**	
O_2: $[_{NegP}$ no $books_1$ Neg $[_{vP_2}$ J_4 gave $[_{VP}$ t_1 t_V to M_3]]]		*!*		
O_3: $[_{NegP}$ Neg $[_{vP_2}$ J_4 gave $[_{VP}$ no $books_1$ t_V to M_3]]]	*!			
O_4: $[_{NegP}$ $[_{vP_2}$ J_4 gave $[_{VP}$ no $books_1$ t_V to M_3]] Neg t_2]	*!		*	
O_5: $[_{NegP}$ $[_{vP_2}$ J_4 gave $[_{VP}$ t_1 t_V to M_3]] no $books_1$ Neg t_2]		*!	*	
O_6: $[_{NegP}$ no $books_1$ to M_3 Neg $[_{vP_2}$ J_4 gave $[_{VP}$ t_1 t_V t_3]]]		*!***	*	

As before, the step from (22-b) to (22-c) is straightforward because a constraint conflict does not arise and FC, $SC_{A,\overline{A}}$, and LR can all be satisfied.

Furthermore, a second argument for local optimization can be gained. Suppose that PP_3 in (22) bears a [top] feature. Then, local optimization proceeds exactly as shown here, creating (22-b) from (22-a) as in T_2, and then (22-c) from (22-b). The only difference is that later in the derivation, PP_3 is moved to the topic position, yielding (23).[14]

(23) $[_{CP}$ To $Mary_3$ $[_{TP}$ $John_4$ T $[_{NegP}$ $[_{vP_2}$ t_4 gave $[_{VP}$ t_1 t_V t_3]] no $books_1$ t'_3 Neg t_2]]]

Viewed globally, SC_A cannot be fulfilled by this sentence. This would threaten to undermine the motivation for remnant vP movement in this context.[15] In contrast, no problem arises if optimization is local: The subderivation vP → NegP respects FC and SC_A by violating the lower-ranked LR, and the subderivation TP → CP respects FC and LR by violating the lower-ranked $SC_{\overline{A}}$. Instead of giving a tableau that shows this latter optimization procedure, let me proceed to the case of primary remnant movement, where exactly the same reasoning applies. A simple example is (24-d) from German, with its derivation in (24-abc).[16]

(24) a. $[_{vP_2}$ der $Fritz_3$ ein $Buch_1$ gelesen] + $[_V$ hat] →
b. $[_{VP}$ ein $Buch_1$ $[_{vP_2}$ der $Fritz_3$ t_1 gelesen] $[_V$ hat]] + T →
c. $[_{TP}$ der $Fritz_3$ $[_{VP}$ ein $Buch_1$ $[_{vP_2}$ t_3 t_1 gelesen] $[_V$ t_V]] $[_T$ hat]] + C →

[14] I assume here that English topicalization is movement to SpecC, but the same argument can be made if topicalization is adjunction to TP, movement to SpecTop, etc.

[15] It would not help to assume that SC can be fulfilled by traces like t'_3 because, if nothing else is said, this would mean that SC is trivially respected by all sentences, vP order always being recoverable with the help of vP-internal traces.

[16] The derivation given here rests on some decisions that are controversial and, to some extent, arbitrary (concerning the projection of auxiliaries, subject raising to SpecT, V raising to a right-peripheral T, etc.). The only important assumption is that both NP_1 scrambling and vP_2 topicalization are triggered by features that obey $SC_{\overline{A}}$.

d. [$_{CP}$ [$_{vP_2}$ t_3 t_1 Gelesen] hat [$_{TP}$ der Fritz$_3$ ein Buch$_1$ t_2] t_V]
read has Fritz a book

Consider first the subderivation $vP_2 \rightarrow$ VP in (24-ab); cf. tableau T_3. Assuming that the object NP_1 has an optional [scr] feature that is matched by [$_V$ *hat*] (and the subject NP_3 does not), the optimal VP is O_1. Here, NP_1 moves to SpecV (respecting FC), and NP_3 stays in situ (respecting LR and violating SC): Whereas [neg] obeys SC_A, [scr] obeys $SC_{\overline{A}}$. This precludes repair-driven movement as in O_3.

T_3: vP → VP Optimization: Scrambling

Input: [$_V$ hat], [$_{vP_2}$ der Fritz$_3$ ein Buch$_1$ gelesen]	FC	SC_A	LR	$SC_{\overline{A}}$
⇒O_1: [$_{VP}$ ein Buch$_1$ [$_{vP_2}$ der Fritz$_3$ t_1 gelesen] [$_V$ hat]]				*
O_2: [$_{VP}$ [$_{vP_2}$ der Fritz$_3$ ein Buch$_1$ gelesen] [$_V$ hat]]	*!			
O_3: [$_{VP}$ der Fritz$_3$ ein Buch$_1$ [$_{vP_2}$ t_3 t_1 gelesen] [$_V$ hat]]			*!	
O_4: [$_{VP}$ der Fritz$_3$ [$_{vP_2}$ t_3 ein Buch$_1$ gelesen] [$_V$ hat]]	*!		*	

The optimal VP O_1 (= (24-b)) is subsequently merged with T. Assuming that the EPP feature can optionally be strong in German, and is strong in the case at hand, the optimal output of the subderivation VP → TP is (24-c), in which the subject NP_3 moves to SpecT (in addition, V moves to T). Since this subderivation respects FC, LR, and $SC_{A,\overline{A}}$, it is not necessary to illustrate the competition by a tableau. Finally, the optimal TP in (24-c) is merged with C. In V/2 languages, an empty finite declarative C bears a [top] feature (and a feature attracting V). Assuming that this feature is also instantiated on vP_2, the optimal output of the subderivation TP → CP is (24-d), which involves remnant vP_2 movement to SpecC and respects FC and LR at the cost of violating the lower-ranked $SC_{\overline{A}}$ (cf. O_1 vs. O_3 in T_4).

T_4: TP → CP Optimization: Primary remnant VP movement

Input: C, [$_{TP}$ der Fritz$_3$ ein Buch$_1$ [$_{vP_2}$ t_3 t_1 gelesen] hat]	FC	SC_A	LR	$SC_{\overline{A}}$
⇒O_1: [$_{CP}$ [$_{vP_2}$ t_3 t_1 Gelesen] hat [$_{TP}$ der Fritz$_3$ ein Buch$_1$]]				*
O_2: [$_{CP}$ Hat [$_{TP}$ der Fritz$_3$ ein Buch$_1$ [$_{vP_2}$ t_3 t_1 gelesen]]]	*!			
O_3: [$_{CP}$ der Fritz$_3$ ein Buch$_1$ [$_{vP_2}$ t_3 t_1 gelesen] hat [$_{TP}$ t'_3 t'_1]]			*!*	
O_4: [$_{CP}$ der Fritz$_3$ [$_{vP_2}$ t_3 t_1 gelesen] hat [$_{TP}$ t'_3 ein Buch$_1$]]			*!	*

$SC_{\overline{A}}$ has not yet been fatally violated by a candidate; i.e., it has played no role in the analysis so far. However, there is evidence for a low-ranked $SC_{\overline{A}}$: As soon as two or more subderivations behave identically with respect to higher-ranked constraints, the decision is passed on to the low-ranked $SC_{\overline{A}}$. A particularly obvious case is the superiority effect in English:[17]

(25) a. (I wonder) [$_{CP}$ who$_1$ C [$_{TP}$ t_1 bought what$_2$]]
b. *(I wonder) [$_{CP}$ what$_2$ C [$_{TP}$ who$_1$ bought t_2]]

Suppose that C bears a strong [wh] feature here which is matched by weak [wh] features on both *wh*-phrases. T_5 then shows that the subderivation TP → CP must involve movement of one *wh*-phrase to SpecC, so as to fulfill FC (cf. O_3), and must

[17] Other phenomena that lend themselves to the same kind of analysis are German weak pronoun fronting and multiple *wh*-movement in Bulgarian. These phenomena are covered by PAR-MOVE in Müller (1997); it seems that $SC_{\overline{A}}$ can do all the work that was attributed to that constraint.

leave one *wh*-phrase in situ, so as to fulfill LR (cf. O_4). O_1 and O_2 (= (25-ab)) meet both requirements, and they vacuously fulfill SC_A. However, only O_1 respects $SC_{\overline{A}}$ by maintaining vP order with [wh] feature checking; therefore, it blocks O_2. Thus, the superiority effect is derived without recourse to constraints like the ECP or the MLC.

T_5: TP → CP Optimization: The superiority effect

I: C-[wh], $[_{TP}$ who$_1$ $[_{vP}$ t$_1$ bought what$_2$]]	FC	SC_A	LR	$SC_{\overline{A}}$
⇒O_1: $[_{CP}$ who$_1$ C $[_{TP}$ t$'_1$ $[_{vP}$ t$_1$ bought what$_2$]]]				
O_2: $[_{CP}$ what$_2$ C $[_{TP}$ who$_1$ $[_{vP}$ t$_1$ bought t$_2$]]]				*!
O_3: $[_{CP}$ – C $[_{TP}$ who$_1$ $[_{vP}$ t$_1$ bought what$_2$]]]	*!			
O_4: $[_{CP}$ who$_1$ what$_2$ C $[_{TP}$ t$'_1$ $[_{vP}$ t$_1$ bought t$_2$]]]			*!	

3.3. Extraction Revisited

Recall that both the remnant XP β_2 in (26-a) and the antecedent of the unbound trace α_1 in (26-b) are barriers for extraction of some element δ_3 in primary remnant movement constructions (freezing; cf. (10)), and that neither β_2 nor α_1 is a barrier for extraction in secondary remnant movement constructions (anti-freezing; cf. (11)).

(26) a. δ_3 ... $[_{\underline{\beta_2}}$ ... t$_3$ ... t$_1$ ...] ... α_1 ... [... t$_2$ ...]
b. δ_3 ... $[_{\underline{\alpha_1}}$ ... t$_3$ ...]$_{...}$ [... t$_2$ ...]

Assuming that XPs in derived positions are barriers, the freezing effect with primary remnant movement can be accounted for. But how can secondary remnant movement escape this effect? The key to a solution is that secondary remnant movement is triggered by SC_A rather than by FC. Hence, it always restores local relations that existed earlier in the derivation. Thus, if α_1, β_2 are not barriers in situ, they will not be turned into barriers in secondary remnant movement constructions because each selected XP will still be in the same minimal domain as the head that selects it. To execute this idea, let us assume the Barriers Condition (BC) in (27-a), and define barriers as in (27-b); this definition differs from standard approaches (cf. Cinque (1990) and references cited there) only in replacing the notion of *sisterhood* in (27-b.(ii)) by the slightly more liberal notion of *same minimal domain*.

(27) Barriers Condition (BC):
a. Movement must not cross a barrier.
b. An XP γ is a barrier unless there is a non-derived head σ such that:
(i) σ selects γ.
(ii) σ and γ are in the same minimal domain.

Thus, extraction from α_1, β_2 does not violate BC in secondary remnant movement constructions. However, given that feature-driven movement in primary remnant movement constructions typically has the effect that an XP γ and its selecting head σ are not in the same minimal domain anymore, extraction from α_1, β_2 violates BC in this case.[18] To derive ungrammaticality from this violation, one could postulate that BC is an inviolable constraint (part of the definition of Move), or that it is ranked

[18]The confinement to non-derived heads in (27-b) ensures that γ may not become transparent by accidentally ending up in the same domain as σ after non-local movement; cf. (10-a).

high. Let us assume the latter. The optimal subderivation YP → ZP (where SpecZ is the landing site of δ_3 in (26)) can then be one that yields an empty output (which vacuously respects BC/FC and violates a lower-ranked ban on empty outputs) – the derivation cannot continue; it crashes.[19]

We expect that movement in primary remnant movement constructions does not create barriers if it is extremely local. As noted by den Besten & Webelhuth (1990), this prediction is borne out. Whereas PP_1 is a barrier for extraction in (28-a) (= (10-c)), it is transparent in (28-b), where it has undergone string-vacuous scrambling.

(28) a. *[$_{vP_2}$ t_1 Gerechnet] hat da_3 gestern [$_{\underline{PP_1}}$ t_3 mit] wieder keiner t_2
counted has there yesterday with again no-one

b. [$_{vP_2}$ t_1 Gerechnet] hat da_3 gestern wieder keiner [$_{\underline{PP_1}}$ t_3 mit] t_2
counted has there yesterday again no-one with

3.4. Movement Types Revisited

Based on examples like those in (12), I have so far assumed that middle field-internal movement (e.g., scrambling) cannot affect remnant XPs, whereas middle field-external movement (e.g., topicalization) can. This generalization has proven problematic in the light of secondary remnant movement, which is obligatorily middle field-internal; cf. (13-a) vs. (13-b). The illformedness of (13-b) now follows from the fact that SC_A-driven movement is strictly local (accompanied by the standard assumption that finite vPs cannot bear a [top] feature); but the difference between illegitimate primary remnant scrambling in (12-b) and legitimate local secondary remnant movement in (13-a) still calls for an explanation. This turns out to be straightforward. Note that the above generalization is not quite correct: Remnant scrambling is in fact possible if the antecedent of the unbound trace has not also undergone scrambling, but another type of movement, e.g., weak pronoun fronting; cf. (29-a) (= (12-b)) vs. (29-b).

(29) a. *daß [$_{vP_2}$ t_1 zu lesen] das $Buch_1$ keiner t_2 versucht hat
that to read the book no-one tried has
"that no-one tried to read the book."

b. daß [$_{vP_2}$ t_1 zu lesen] es_1 keiner t_2 versucht hat
that to read it no-one tried has
"that no-one tried to read it."

Similarly, middle field-external remnant *wh*-movement is impossible if the antecedent of the unbound trace has also undergone *wh*-movement, and possible if it has undergone another type of movement, e.g., scrambling; cf. (30-a) vs. (30-b).

(30) a. *[$_{NP_2}$ Was für ein Buch t_1] fragst du dich [$_{CP}$ [$_{PP_1}$ über wen] du t_2
what for a book ask you REFL about whom you
lesen sollst] ?
read should

[19]Alternatively, the optimal subderivation could remove the feature that triggers δ_3-movement and, e.g., change a [+wh] *wh*-element into a [–wh] indefinite. Then, δ_3 can remain in situ without violating BC or FC, at the cost of a violation of a lower-ranked faithfulness constraint; this amounts to neutralization of a [±wh] distinction in the input.

"*What kind of book do you wonder about whom to read?"

b. [$_{NP_2}$ Was für ein Buch t_1] hast du [$_{PP_1}$ über die Liebe] t_2 gelesen ?
what for a book have you about the love read
"What kind of book did you read about love?"

Thus, the data suggest a constraint like UNAMBIGUOUS DOMINATION (UD) in (31), rather than a stipulation as to which movement type may affect remnant XPs.[20]

(31) UNAMBIGUOUS DOMINATION (UD):
In ... [$_\alpha$ ... β ...] ..., α and β cannot check the same kind of feature (outside α).

It can easily be verified that UD is violated in cases like (29-a) and (30-a), but respected in (29-b), (30-b), and typical primary remnant movement constructions that involve a combination of scrambling (or NP raising) and topicalization. Furthermore, it is now clear why secondary remnant movement as in (13-a) can never violate UD: α and β cannot check the same feature if α does not check a feature at all.

4. Conclusion and Outlook

I have tried to show that the different properties of primary and secondary remnant movement follow from the fact that the former operation is feature-driven, whereas the latter is not: It is a repair strategy forced by SHAPE CONSERVATION and the FEATURE CONDITION, in violation of LAST RESORT. As a consequence of this, secondary object fronting may also be required; BARRIERS CONDITION violations can be avoided; and UNAMBIGUOUS DOMINATION violations do not show up.

On a more general note, I have argued that since repair-driven secondary remnant movement presupposes constraint violability and ranking, it lends itself to an optimality-theoretic implementation. What is more, it provides evidence that syntactic optimization is local, not global (as is standardly assumed): On the one hand, there are ill-formed derivations that are indeed locally suboptimal, but globally optimal (cf. T_1). And on the other hand, there are well-formed derivations that are locally optimal, but globally suboptimal (cf. T_2). In general, it seems that syntactic repair is typically a local phenomenon: An "offending" property is removed instantaneously, not at some earlier or later stage in the derivation. This holds for other cases of repair-driven movement that have been proposed in the literature; cf. Heck & Müller (1999), where arguments are given for local analyses of, e.g., semantically vacuous QR that is forced by a higher-ranked parallelism constraint (Fox (1995)), and *wh*-scrambling that is forced by a higher-ranked Neg-intervention constraint (Beck (1996)). Moreover, many other cases of syntactic repair that have been approached in terms of global optimization (cf., e.g., Grimshaw (1997) on *do*-support, Pesetsky (1998) and Legendre, Smolensky, & Wilson (1998) on resumptive pronouns, Schmid (1998) on the Westgermanic "Ersatzinfinitiv") can be treated by local optimization. It remains to be seen, though, whether local optimization can (or should) do all the work that global optimization has been held responsible for in syntax.

[20]UD is from Müller (1998). To ensure ungrammaticality in cases where UD would have to be violated by a subderivation, the same reasoning applies as in the case of BC. For more empirical evidence and attempts to derive (something like) this constraint from even more general assumptions, see also Takano (1993), Koizumi (1995), and Kitahara (1997).

References

Bayer, Josef. 1996. *Directionality and Logical Form. On the Scope of Focussing Particles and Wh-in-situ.* Dordrecht: Kluwer.

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

Besten, Hans den, & Gert Webelhuth. 1987. Adjunction and Remnant Topicalization in the Germanic SOV-Languages. Paper presented at GLOW (Venice).

Besten, Hans den, & Gert Webelhuth. 1990. Stranding. In *Scrambling and Barriers*, ed. Günther Grewendorf and Wolfgang Sternefeld, 77-92. Amsterdam: Benjamins.

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

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

Cinque, Guglielmo. 1990. *Types of A'-Dependencies.* Cambridge, Mass.: MIT Press.

Dikken, Marcel den. 1996. Extraposition as Intraposition, and the Syntax of English tag Questions. Ms., Vrije Universiteit Amsterdam.

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

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

Frank, Robert, Young-Suk Lee, & Owen Rambow. 1992. Scrambling as Non-Operator Movement and the Special Status of Subjects. In *Proceedings of the Leiden Conference for Junior Linguists* 3:135-154. Leiden: University of Leiden.

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.

Haegeman, Liliane. 1995. *The Syntax of Negation.* Cambridge: CUP.

Heck, Fabian. 1999. Quantifier Scope in German and Cyclic Optimization. Ms., Universität Stuttgart. To appear in *Competition in Syntax*, eds. Gereon Müller & Wolfgang Sternefeld. Berlin: Mouton/de Gruyter.

Heck, Fabian, & Müller, Gereon. 1999. Repair-Driven Movement and the Local Optimization of Derivations. Ms., Universität Stuttgart.

Hinterhölzl, Roland. 1997. An XP-Movement Account of Restructuring. Ms., USC.

Huang, Cheng-Teh James. 1982. *Logical Relations in Chinese and the Theory of Grammar.* Doctoral dissertation. MIT, Cambridge, Mass.

Johnson, Kyle. 1998. Ways to Hide Quantifier Raising. Ms., University of Mass., Amherst.

Kayne, Richard. 1998. Overt vs. Covert Movement. Ms., New York University. Also in *Syntax* 1.

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

Koizumi, Masatoshi. 1995, *Phrase Structure in Minimalist Syntax.* Doctoral dissertation, MIT, Cambridge, Mass.

Koopman, Hilda, & Anna Szabolcsi. 1999. Verbal Complexes. To appear: Cambridge, Mass.: MIT Press.

Kroch, Anthony. 1974. *The Semantics of Scope in English.* Doctoral dissertation, MIT, Cambridge, Mass.

Lakoff, George. 1971. On Generative Semantics. In *Semantics*, ed. Danny Steinberg & Leon Jakobovits, 232-296. Cambridge: Cambridge University Press.

Lasnik, Howard, & Mamoru Saito. 1992. *Move α*. Cambridge, Mass.: MIT Press.
Legendre, Géraldine, Paul Smolensky, & Colin Wilson. 1998. When is Less More? Faithfulness and Minimal Links in Wh-Chains. In *Is the Best Good Enough?*, eds. Pilar Barbosa et al., 249-289. Cambridge, Mass.: MIT Press & MITWPL.
McCarthy, John. 1999. Harmonic Serialsm and Parallelism. Ms., University of Massachusetts, Amherst.
Meinunger, André. 1995. Prominence Hierarchy and Phrase Ordering. *FAS Working Papers*, vol. 2. ZAS: Berlin.
Müller, Gereon. 1997. Parallel Movement. In *Zur Satzstruktur im Deutschen*, ed. Franz d'Avis & Uli Lutz, 171-214. Tübingen: SFB 340 Working Papers no. 90. Revised version to appear in *Optimality-Theoretic Syntax*, eds. Géraldine Legendre, Jane Grimshaw, & Sten Vikner. Cambridge, Mass.: MIT Press.
Müller, Gereon. 1998. *Incomplete Category Fronting*. Dordrecht: Kluwer.
Noonan, Maire. 1999. CP-Pied Piping and Remnant VP Movement in Long-Distance Wh-Movement. Ms., York University.
Ordóñez, Francisco. 1997. *Word Order and Clausal Structure in Spanish and Other Romance Languages*. Doctoral dissertation, CUNY.
Pesetsky, David. 1998. Some Optimality Principles of Sentence Pronunciation. In *Is the Best Good Enough?*, eds. Pilar Barbosa et al., 337-383. Cambridge, Mass.: MIT Press & MITWPL.
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.
Schmid, Tanja. 1998. Optional, Obligatory, and Impossible IPP Constructions in West Germanic. Ms., Universität Stuttgart. To appear in *Proceedings of Console*.
Sauerland, Uli. 1997. Scrambling and Interpretability. Ms., MIT, Cambridge, Mass.
Stechow, Arnim von, & Wolfgang Sternefeld. 1988. *Bausteine syntaktischen Wissens*. Opladen: Westdeutscher Verlag.
Takano, Yuji. 1993. Minimalism and Proper Binding. Ms., UC Irvine.
Thiersch, Craig. 1985. VP and Scrambling in the German Mittelfeld. Ms., University of Tilburg.
Vikner, Sten. 1999. V-to-I Movement and 'Do'-Insertion in Optimality Theory. Ms., Universität Stuttgart. To appear in *Optimality-Theoretic Syntax*, eds. Géraldine Legendre, Jane Grimshaw, & Sten Vikner. Cambridge, Mass.: MIT Press.
Watanabe, Akira. 1992. Wh-in-Situ, Subjacency, and Chain Formation. Ms., MIT, Cambridge, Mass. (*MIT Occasional Papers in Linguistics* 2.)
Williams, Edwin. 1999. Economy as Shape Conservation. Ms., Princeton University.
Wilson, Colin. 1998. Bidirectional Optimization and the Theory of Anaphora. Ms., Johns Hopkins University. To appear in *Optimality Theoretic Syntax*, eds. Jane Grimshaw, Géraldine Legendre, & Sten Vikner. Cambridge, Mass.: MIT Press.

Seminar für Sprachwissenschaft
Universität Tübingen
D-72074 Tübingen, Germany

gereon.mueller@uni-tuebingen.de