Remnant Movement in Local Derivational Grammar

Gereon Müller (Universität Leipzig)*

February 22, 2014

Abstract

This paper sets out to substantiate three claims. First, there is (conceptual and emiprical) evidence for adopting a local derivational approach to syntax (as opposed to non-local derivational, local declarative, and non-local declarative approaches). Second, the general counter-bleeding effect with remnant movement provides an argument for a derivational approach as such, but raises problems for a *local* derivational approach as it is presupposed in phase-based minimalist syntax. Third, this challenge can be met if the movement-related features of moved items (like [wh]) have values that act as *buffers*, storing minimal aspects of the derivational history on the moved item.

1. Background

1.1. Two Dichotomies

Current theories of syntax can roughly be grouped into four classes according to two dichotomies: An approach can be *derivational* (such that syntactic structures are generated by structure-building and structure-manipulating operations) or *declarative* (such that syntactic structures are viewed as potential objects that have to comply with the grammatical constraints of a language); and an approach can be *local* (such that all syntactic dependencies are confined to local domains, e.g., subtrees, and syntactic constraints have access only to these subtrees) or *non-local* (such that syntactic dependencies can span larger portions of structure, potentially the whole sentence, and syntactic constraints also have access to these large domains). A cross-classification yields four models of grammar: Current *local derivational* approaches include most prominently the Minimalist Program (Chomsky (2001; 2008)), where syntactic structures are generated bottom-up, by alternating operations like Merge, Move, and Agree, and the accessible window of a derivation is quite small throughout (it is standardly assumed to be confined to the minimal phase). GPSG (Gazdar (1981), Gazdar et al. (1985)) and HPSG (Pollard & Sag (1994), Sag & Wasow (1999), Müller, St. (2007)) belong to the class of *local declarative* approaches: Dependencies and constraints are evaluated in small subtrees (essentially, mother-daughter relations), and syntactic structures are assumed to be licensed by constraint satisfaction, rather than being generated successively by syntactic operations. Next, there are *non-local derivational* models, like the classical Principles-and-Parameters theory as developed in Chomsky (1981; 1986)

and Chomsky & Lasnik (1993). Here, syntactic structures are generated by generalized rules of the phrase-structure component (X-bar theory) and operations like Move-α or (even more generally) Affect-α (Lasnik & Saito (1984; 1992)); by assumption, syntactic dependencies can in principle span large domains (even though, e.g., movement is, in practice, required to stop in all intermediate clause-initial positions by locality constraints), and constraints can in principle access the entire structure of a given, arbitrarily large syntactic object. Finally, LFG (Bresnan (2001)) is an approach that is *non-local declarative*: Syntactic structures are licensed (rather than generated), and dependencies and constraints that restrict them have access to the full structure (or, more precisely, structures, given the central assumption that sentences are simultaneously represented at several parallel levels of representation).

Against this background, the question arises what potential empirical and conceptual evidence for or against a particular model looks like.

1.2. *Local vs. Non-Local Approaches*

As for the difference between local and non-local approaches to syntax, empirical arguments for the former are provided by the existence of (morphological, syntactic or semantic) reflexes of seemingly non-local dependencies like movement (see, e.g., Sag & Wasow (1999), Fox (2000), Bouma et al. (2001), McCloskey (2002), Asudeh (2004), Huybregts (2009), Assmann et al. (2010), Lahne (2009), and references cited in all these works). These reflexes of movement initially favour a local modelling of non-local dependencies because they suggest a partitioning of the structure affected by movement into subparts (either by assuming series of small, successive-cyclic movement steps, or by assuming SLASH features that are percolated locally), and the availability of the relevant information (viz., that some domain has been affected by movement) for other operations (e.g., insertion of special morphological exponents, application of certain syntactic processes, semantic reconstruction, etc.) Reflexes of movement can be captured in non-local approaches to movement dependencies like the standard LFG accounts (Bresnan (2001), Dalrymple (2001, ch. 14)), earlier accounts that envisage unbounded movement transformations (Chomsky (1965; 1975), Ross (1967), Bresnan (1976a;b)), or *path*-based approaches as they have been developed in Principles-and-Parameters theory as an alternative to successive-cyclic movement (Kayne (1982), Pesetsky (1985), Koster (1987), and Longobardi (1985)), but only at a cost.[1] Further empirical arguments for a local approach to syntax are given in Heck & Müller (2000; 2003) on the basis of syntactic repair operations.

In addition, conceptual arguments for a local modelling of non-local dependencies have been advanced. These arguments typically center around notions like complexity (see Gazdar (1981) on preserving context-freeness by avoiding transformations as tools to model non-

---

[1] Dalrymple (2001, ch. 14) develops a non-local LFG approach to movement that addresses this issue. Here, non-local movement dependencies can be stated as identity relations between two grammatical functions; what qualifies as a legitimate identity relation in a non-local dependency is then encoded as a *regular expression* in the phrase-structure component. This analysis does not break up a movement dependency into smaller parts (as with successive-cyclic movement steps of SLASH feature percolation), and, as such, it does not imply any *record*, or *track-keeping device*, of a non-local movement dependency in the syntactic structure that shows up between the displaced item and its base position. However, in view of the existence of (morphological) reflexes of movement, Dalrymple (2001) proposes that a track-keeping device (a feature [LDD], for 'long-distance dependency') can be *added* to phrase structures after all, so as to provide a point of reference for a (morphological) reflex of movement. This solution may be viewed as satisfactory from a purely technical point of view; but it seems clear that it is inferior to local modellings of the phenomenon: The sole purpose of the [LDD] feature is to permit accounts of (morphological) reflexes of movement; the feature does not play any role in bringing about, or restricting, movement dependencies per se.
The same conclusions can be drawn for path-based approaches; as McCloskey (1988, 30) notes, here the movement path is "not formally marked in any way".

local dependencies, and Chomsky (2001; 2005b; 2007) on introducing phases as a means of bringing about "efficient computation"). Heck & Müller (2010) present an argument from learning theory, based on the fact that the set of possible grammars that the language learner needs to consider is greatly reduced in a local approach.

1.3. Derivational vs. Declarative Approaches

Turning next to the second fundamental distinction, viz., that between derivational and declarative approaches to syntax, it can be noted that central arguments for the former are typically based on *opacity* phenomena, i.e., *counter-bleeding* and *counter-feeding* (Chomsky (1951; 1975), Kiparsky (1973)). In derivational terms, *counter-feeding* describes rule interactions where a rule A would feed another rule B (i.e., create the context for B to apply) but does not actually do so because it applies too late (B's chance to apply has already passed): Thus, A counter-feeds B. *Counter-bleeding* captures rule interactions where a rule A would bleed another rule B (i.e., destroy the context in which B can apply) but, again, does not actually do so because it applies too late (B has already applied): A counter-bleeds B. From a derivational perspective, opacity is entirely unproblematic: Operations that induce counter-feeding or counter-bleeding apply late in the derivation (put another way, operations that are counter-fed and counter-bled apply early). However, from a declarative perspective that does without postulating sequences of operations, counter-feeding and counter-bleeding raise potential problems: Rule interaction in counter-feeding and counter-bleeding environments is opaque because it cannot be determined by simply looking at the respective output representations why some operation (or process) B has not applied even though its context of application would seem to be present (*counter-feeding*), and why some operation B has applied even though its context of application would not seem to be present (*counter-bleeding*). Thus, a declarative approach to syntax faces initial problems of underapplication and overapplication in the area of opaque interactions.

A classical instance of counter-feeding in syntax involves *wanna*-contraction in English. This operation is possible in control contexts but impossible in exceptional case marking (ECM) contexts; see (1-a) vs. (1-b).

(1)   a.   Who do you want to/wanna meet ?
      b.   Who do you want to/*wanna meet Mary ?

A derivational analysis is straightforward (Bresnan (1972), Pullum (1979), Arregi & Nevins (2012)): There are two relevant operations, viz. (i) *wanna*-contraction (which contracts *want* and *to*, yielding *wanna*, under strict phonological adjacency) and (ii) wh-movement to the specifier of an interrogative C node. Suppose next that *wanna*-contraction necessarily precedes wh-movement.[2] On this view, in (1-a), *want* and *to* are adjacent to begin with; so *wanna*-contraction can apply without problems. In (1-b), however, the ECM-subject *who* initially shows up between *want* and *to*. Wh-movement of *who* would result in a feeding of *wanna*-contraction, but it doesn't because (by assumption) this operation applies too late. Thus, (1-b) instantiates *counter-feeding*.[3]

---

[2] Such an order would follow immediately if derivational order is governed by the Strict Cycle Condition, and cyclic domains are sufficiently small (as suggested by McCawley (1984; 1998); also see footnote 13 below): *Wanna*-contraction affects the matrix VP, but wh-movement of the subject of the infinitive affects the matrix CP.

[3] Note that this reasoning presupposes that at least one of the following statements holds: (i) There is no structural subject in control infinitives. (ii) There is an empty subject in control infinitives (e.g., PRO, as in Chomsky (1981)), which does not block phonological adjacency. (iii) If there is an overt subject in control infinitives, and this is deleted by Equi NP deletion (or some other rule), then this latter rule must apply early, so

A standard argument for counter-bleeding in syntax involves reconstruction for Principle A of the binding theory. The core observation is that reflexive pronouns can undergo movement, and they can occur in phrases that undergo movement. After the movement operation, they should therefore not be able to satisfy Principle A of the binding theory; but they clearly can satisfy this constraint in the relevant examples. Relevant examples that show this are given in (2-a) (with scrambling of the reflexive to a pre-subject position in German), (2-b) (with topicalization of a reflexive in English), and (2-c) (with topicalization of an XP containing a reflexive in German).

(2)   a.   dass [$_{DP_1}$ sich] der Fritz$_1$ gestern   im   Spiegel gesehen hat
             that   REFL the Fritz yesterday in the mirror   seen   has
       b.   [$_{DP_1}$ Himself] John$_1$ does not really like
       c.   [$_{DP}$ Bücher über   sich$_1$] hat er$_1$ keine gelesen
                 books   about REFL   has he none read

Principle A requires a reflexive pronoun to be bound within its minimal clause, the constraints triggering movement require the reflexive to show up in a displaced position outside the c-command domain of its binder, and the wellformedness of the examples in (2) can then be taken to indicate that Principle A can be checked, and satisfied, before subsequent movement would bleed this operation.[4]

As a third case exhibiting opaque interaction in syntax, consider the analysis of the prohibition against movement of ergative DPs given in Assmann et al. (2012). The generalization to be derived is that in many ergative systems, an ergative DP cannot undergo A-bar movement (wh-movement, topicaliztion, relativization, focus movement), in contrast to an absolutive DP. Crucially, no comparable restriction holds in accusative systems, where accusative and nominative DPs are both mobile in principle. This can be derived as follows. First, for locality reasons, ergative and accusative DPs must undergo an intermediate movement step to SpecT in the course of A-bar movement. Second, T assigns (absolutive/nominative) case to an internal argument DP in ergative systems, and to an external argument DP in accusative systems; v assigns structural case to an external argument in ergative systems (ergative), and to an internal argument in accusative systems (accusative) (see Murasugi (1992)), as an instance of of Agree (Chomsky (2001)). Third, movement of some DP α to SpecT makes assignment of absolutive/nominative to the remaining DP β impossible because SpecT counts as *closer* to T than the vP-internal domain. Thus, if an intermediate movement step of α occurs *before* absolutive/nominative assignment by T, such assignment is blocked, and ungrammaticality results because β remains without structural case, as an instance of *bleeding*. In contrast, if an intermediate step of α occurs *after* absolutive/nominative assignment by T, it comes too late to create problems, and there is *counter-bleeding*. In Assmann et al. (2012), it is argued that there is independent evidence that ergative DPs move early (before Agree operations have taken place), whereas accusative DPs move late (after Agree operations have taken place; see Müller (2009)), yielding a bleeding effect in the first case, and a counter-bleeding effect in the second. This latter consequence poses a problem from a declarative perspective: The resulting two TP structures *are identical* in the relevant respects, with an

that it can feed *wanna*-contraction. (iv) If the movement theory of control (MTC) is adopted (Hornstein (2001), Boeckx, Hornstein & Nunes (2010)), control movement to the matrix θ-position must take place early, so that it can feed *wanna*-contraction.

[4] Again, the required order would follow from the Strict Cycle Condition (with satisfaction of Principle A taking place in a smaller domain than movement), but eventually a bit more would have to be said, given that there is evidence that Principle A can be satisfied at any point in the derivation; see Barss (1986), Belletti & Rizzi (1988).

ergative/accusative DP in SpecT that should uniformly block case assignment by T to an item that still needs case but does so only in one case.

Some opaque interactions in syntax can be captured in declarative approaches by enriching representations of linguistic expressions with abstract material that encodes (what would otherwise qualify as) earlier derivational steps (e.g., by assuming traces/copies, and various type of empty pronouns, like pro and PRO), and by then postulating (more complex) constraints that refer to the enriched structure. Thus, counter-feeding of *wanna*-contraction by wh-movement can be captured representationally if movement leaves a trace (or copy) in the original position, and the *wanna*-contraction rule is formulated in a more abstract way, such that traces (copies) can block phonological adjacency (whereas an empty subject pronoun PRO in control infinitives for some reason cannot do so). However, in other cases, deriving opacity effects in syntax by enriching representations is not as straightforward because it does not suffice to postulate abstract elements; there also have to be *special constraints* or *mechanisms* that refer to them, and these typically do not qualify as simple and elegant, and thus lead to an inferior analysis. More specifically, the problem is that the *effects* of constraint interaction are integrated into the definition of a single constraint, which makes this constraint extremely implausible (see Grimshaw (1998), Chomsky (2001; 2008)). The counter-bleeding effect with reflexivization and movement is a case in point. As shown by Barss (1984), this effect can be derived in a declarative approach that does rely on an order of syntactic operations by assuming a version of Principle A as in (3-a) (with binding replaced by chain-binding), and postulating the (doubly) disjunctive notion of chain-binding as in (3-b), which bascially undoes movement for the purposes of binding theory evaluation; thus, the effects of an interaction of operations (movement counter-bleeds Principle A) are not derived from the application of the operations themselves, but are stipulated as part of a more complex constraint.[5]

(3) a. *Principle A*:
         At S-structure, an anaphor is chain-bound in its binding domain.
     b. *Chain-Binding*:
         α chain-binds β iff (a), (b), and (c) hold:
         (i) α and β are co-indexed.
         (ii) α occupies an A-position.
         (iii) α c-commands β, or α c-commands a trace of γ, where $\gamma = \beta$ or γ dominates β.

Concerning the third case discussed above – viz., the interaction of case assignment by T and movement of some case-bearing DP to SpecT –, the situation is even worse. It is completely unclear how the counter-bleeding effect with nominative assignment by T in contexts with accusative movement to SpecT could be accounted for in a declarative approach by enriching representations and postulating more complex constraints.

These three examples may suffice as illustrations of what form opacity arguments for a derivational approach to syntax typically take. It should be emphasized that the arguments considered here are by no means conclusive: It is of course possible to come up with a completely different account of *wanna*-contraction that is compatible with a declarative analysis without requiring postulation of a different behaviour of two types of empty categories (see Pullum (1997)); or to devise a non-structural theory of reflexivization in which movement is not expected to potentially give rise to bleeding effects (see, e.g., Pollard & Sag (1992),

---

[5] Barss (1986) ultimately adopts a concept of *chain accessibility sequences*, which extends the chain binding proposal sketched here; but this analysis suffers from the same shortcomings and is in fact even more complex.

Reinhart & Reuland (1993), Büring (2005)); or to account for the ban on ergative movement (and the concurrent lack of a comparable ban on accusaive movement) in a different way that does not involve an interaction of movement and case assignent (see Campana (1992), Stiebels (2006), Coon et al. (2011)). These considerations notwithstanding, I take the derivational approach to syntax to be well supported by opacity arguments, and will presuppose it in what follows. Against this background, let me now turn to remnant movement constructions in German.

2. A Non-Local Derivational Approach to α-over-β Configurations: Bleeding and Counter-Bleeding

2.1. Remnant Movement

Remnant movement constructions like the German example in (4) have also been taken as an argument for a derivational approach to syntax since they exhibit opacity – more specifically, a counter-bleeding effect (see Müller (1998)).

(4)　　[$_{VP_2}$ t_1 Gelesen] hat das Buch$_1$ keiner t_2
　　　　　　　read　　　 has the book　no-one

(4) involves a combination of two movement operations. First, there is scrambling from VP (to a pre-subject position which is arguably a specifier of vP – it follows the position occupied by weak pronouns in German but can precede a subject DP that has not undergone optional movement to SpecT). Second, there is remnant VP topicalization to SpecC (or to the specifier of some other functional head in the left periphery). More generally, remnant movement constructions are characterized by a pre-movement α-over-β configuration, as in (5), in which both α and β undergo separate movement operations (and β targets an α-external position).

(5)　　... [$_\alpha$ ... β ...] ...

Movement in α-over-β configurations of this type has a number of conspicuous properties. Assuming that all movement operations leave traces (or copies), remnant movement (i.e., movement of α in the presence of β-movement in (5)) creates an unbound trace (or copy). Accordingly, the properties of remnant movement constructions can be accounted for by postulating specific *restrictions* on unbound traces (see Thiersch (1985), den Besten & Webelhuth (1987; 1990), Müller (1993), Grewendorf & Sabel (1994), Saito (2003), Collins & Sabel (2007) for attempts along these lines). However, it is of course a priori preferable to account for the properties of multiple movement in α-over-β configurations without invoking designated constraints referring to unbound traces; such constraints will eventually qualify as construction-specific. In fact, from a current, third-factor-based minimalist perspective (see Chomsky (2005a; 2007; 2008; 2013), it would seem to be impossible to postulate constraints referring to traces (like the Empty Category Principle (ECP) of Chomsky (1981)) – let alone constraints referring to unbound traces; and this conclusion holds independently of whether displacement in syntactic derivations is or is not assumed to leave a trace (copy) in the first place (Epstein & Seely (2002)).

As argued in Müller (1998), it is possible to come up with a straightforward, reasonably simple analysis of three conspicuous properties of remnant movement constructions (which I will call *freezing*, *anti-freezing*, and *Müller-Takano generalization*) that does not have to resort to concepts like "unbound trace", provided that a *derivational* (rather than *declarative*) approach is adopted. I address the relevant generalizations and the derivational analysis in

the next subsection.[6]

2.2. Three Properties of α-over-β Configurations

2.2.1. Freezing

The first relevant generalization is that movement of α and β in α-over-β configurations as in (5) induces a *freezing* effect if movement of β ends up in a higher position than movement of α; i.e., if the trace of β (assuming there is one) is bound in the final representation: A trace in a moved item leads to illformedness when its antecedent is outside of the moved item and c-commands the trace. Relevant German data illustrating this interaction of movement operations are given in (6-a) (with VP topicalization accompanied by wh-movement from VP) and (6-b) (with VP scrambling accompanied by wh-movement from VP).

(6) a. *Was$_1$ denkst du [$_{VP_2}$ t$_1$ gelesen] hat keiner t$_2$?
         what think   you   read   has no-one
     b. *Was$_1$ hat [$_{VP_2}$ t$_1$ gelesen] keiner t$_2$?
         what has   read   no-one

Freezing effects as in (6) instantiate a case of transparent (rather than opaque) interaction of operations. More precisely, we are dealing with a *bleeding* effect here: XP$_2$ movement bleeds XP$_1$ movement. Consequently, deriving freezing effects in α-over-β configurations is unproblematic under both a declarative and a derivational approach (provided the approach is not strictly local; see below). As shown by Browning (1991), in a declarative approach it suffices to look at the output representations in (6-ab) to correctly determine illformedness. Given a constraint like the Condition on Extraction Domain (CED) in (7) (see Huang (1982), Chomsky (1986), Browning (1987), Cinque (1990)), VP$_2$ in (6) qualifies as a barrier between t$_1$ and its antecedent *was*$_1$ ('what') (since it shows up in a specifier position), and movement therefore illegitimately crosses a barrier.

(7) *Condition on Extraction Domain* (CED):
     a. Movement must not cross a barrier.
     b. An XP is a barrier iff it is not a complement.

Similarly, in a derivational approach, the freezing effect can be straightforwardly derived: Given the Strict Cycle Condition (SCC; see Chomsky (1973), Perlmutter & Soames (1979), among many others), a moderately updated version of which is given in (8), movement of XP$_2$ (which targets a lower position) must precede movement of XP$_1$ (which targets a higher position) in (6). Consequently, movement of XP$_1$ takes place from XP$_2$ when the latter has already become a barrier, and a CED violation will be unavoidable (see (7)): Movement of

---

[6] Throughout this paper, I presuppose that examples like (4) do indeed involve remnant movement from an α-over-β configuration, with scrambling preceding topicalization, and that the three properties discussed in the next subsection do indeed properly characterize α-over-β configurations. Both assumptions have been called into question. Thus, De Kuthy & Meurers (2001), Fanselow (2002), Hale & Legendre (2004) and Thoms & Walkden (2013), among others, argue against a remnant movement analysis of examples like (4). In contrast, Grewendorf (2003; 2004) and Abels (2008), while basically assuming a remnant movement analysis of these kinds of data, adopt some generalizations about α-over-β configurations that are not co-extensive with those I adopt in what follows (involving, e.g., systematic exceptions from freezing and anti-freezing). For reasons of space and coherence, and since the main focus of the present paper is on a reconstruction of an existing theoretical analysis in a slightly different (i.e., more local) framework rather than on justifying this analysis, I will not discuss these alternative proposals in what follows. Suffice it to say that I think there is neither strong empirical evidence against remnant movement in German, nor strong evidence for relativizing the generalizations that I will now turn to in the main text.

α bleeds extraction of β from α.[7]

(8) *Strict Cycle Condition* (SCC):
Within the current XP α, a syntactic operation may not target a position that is included within another XP β that is dominated by α.

2.2.2. Anti-Freezing

The second generalization about movement in α-over-β configurations as in (5) is that an *anti-freezing* effect arises if movement of β ends up in a lower position than movement of α; i.e., if the trace of β is not bound in the final representation because remnant movement has applied: A trace in a moved item does not have to lead to illformedness when its antecedent is outside of the moved item and does not c-command the trace. Relevant data from German are (9-a) (= (4), with scrambling from VP accompanied by VP topicalization in the same clause), (9-b) (with scrambling from VP accompanied by long-distance VP topicalization), and (9-c) (with wh-movement from VP accompanied by long-distance VP topicalization).[8]

(9) a. [$_{VP_2}$ t_1 Gelesen] hat das Buch$_1$ keiner t_2
             read   has the book   no-one
     b. [$_{VP_2}$ t_1 Zu lesen] glaubte sie [$_{CP}$ t'_2 habe   [$_{DP_1}$ das Buch] keiner t_2
             to read   believed she   has$_{subj}$ the book   no-one
         versucht]
         tried
     c. ?[$_{VP_2}$ t_1 Zu lesen] weiß ich nicht [$_{CP}$ was$_1$ sie t_2 versucht hat]
             to read   know I   not   what she   tried   has

In contrast to freezing effects, anti-freezing effects in α-over-β configurations distinguish between declarative and derivational approaches because here, the interaction of operations is opaque. More specifically, there is a *counter-bleeding* effect that can receive a straightforward account in a derivational approach: Given the SCC, movement of XP$_2$ (which targets a higher position) must follow movement of XP$_1$ (which targets a lower position). Thus, extraction of XP$_1$ from XP$_2$ takes place when XP$_2$ is still in its base (complement) position, in accordance with the CED; and subsequent movement of XP$_2$, which turns XP$_2$ into a barrier, comes too late to prevent extraction. In other words: Movement of α counter-bleeds extraction of β from α. However, from a declarative perspective, anti-freezing poses a problem: By only looking at the output representations, it is not clear why the sentences in (9) can be grammatical; we should expect a straightforward violation of the CED because t_1 is separated from its antecedent (*das Buch*$_1$ ('the book') or *was*$_1$ ('what')) by a barrier. The only possible way out in a declarative approach, it seems, would be to artificially enrich the CED in the same way that Barss (1984; 1986) modified Principle A of the binding theory in view of the counter-bleeding effect with moved items that are, or contain, anaphors, and this would imply, again, stipulating the effects of constraint interaction (here: the SCC and the

---

[7] There is a qualification, though, that foreshadows the problems with a *local* derivational approach I will address below. As observed by Collins (1994), to derive illformedness in (6), it must be ensured that a derivation in terms of 'chain interleaving' is blocked where β is moved from α to an intermediate position first, α is moved to its target position next, and β is finally moved from the intermediate position to its ultimate landing site. This derivation respects both the CED and the SCC. Collins (1994) argues that it can be excluded by (transderivational) economy considerations; in Müller (1998, ch. 4), I suggest that such a derivation is ruled out as an instance of a constraint on improper movement, which is itself derivable from a combination of derivational and transderivational economy constraints.

[8] The last case gives rise to a wh-island effect, but these effects are often not very strong in German, particularly if the item that moves from the wh-island does so by topicalization; see Fanselow (1987).

CED) by integrating the interaction into the formulation of a single, much more complex, constraint.

2.2.3. Müller-Takano Generalization

The third generalization about movement in α-over-β configurations that I will focus on here is about cases where the two items undergo the same kind of movement. The generalization is due to Müller (1993) and Takano (1994), and has sometimes been referred to as the "Müller-Takano generalization" (see, e.g., Sauerland (1999), Pesetsky (2012)). It states that remnant XPs cannot undergo Y-movement if the antecedent of the unbound trace has also undergone Y-movement, where Y stands for a movement-related feature (like [wh] for wh-movement, [top] for topicalization, [Σ] for scrambling, etc.). The sentences in (10) fall under this generalization. In the German example (10-a), DP scrambling from VP to a position following the subject is accompanied by remnant VP scrambling; in the German example (10-b), DP scrambling from VP to a position preceding the subject is accompanied by remnant VP scrambling (cf. Fanselow (1991), Grewendorf & Sabel (1994), Frank et al. (1992), Stechow (1992), and Haider (1993), among others, for these kinds of data); and in the Japanese example (10-c), long-distance scrambling of DP (which is an option as such in Japanese) is accompanied by remnant CP scrambling (cf. Saito (1992)). In all these cases, ungrammaticality results; i.e., anti-freezing is somehow suppressed.

(10)   a. *dass [$_{VP_2}$ t_1 zu lesen] keiner [$_{DP_1}$ das Buch   ] t_2 versucht hat
            that   to read   no-one   the book$_{acc}$ tried   has

        b. *dass [$_{VP_2}$ t_1 zu lesen] [$_{DP_1}$ das Buch   ] keiner t_2 versucht hat
            that   to read   the book$_{acc}$ no-one   tried   has

        c. *[$_{CP_2}$ Mary-ga t_1 yonda-to   ] [$_{DP_1}$ sono hon-o   ] John-ga t_2 itta (koto)
            Mary$_{nom}$ read-COMP   that book$_{acc}$ John$_{nom}$ said   fact

Again, there is a very simple analysis in a derivational approach. As has been observed by Kitahara (1994; 1997), Fox (1995), Koizumi (1995), Müller (1998), and others, Müller-Takano generalization effects follow from the Minimal Link Condition. Here is why: Suppose that movement of of XP$_2$ and XP$_1$ is triggered by the same movement-related feature. In this case, XP$_2$ is invariably closer to the attracting head, and must therefore move first; early movement of the lower XP$_1$ would give rise to a violation of the (generalized) Minimal Link Condition (MLC; the generalization consists of an extension of minimality from c-command to dominance contexts, thereby incorporating a *relativized* A-over-A principle as it has been proposed in Chomsky (1973), Bresnan (1976b), Fitzpatrick (2002)). Therefore, a CED effect is unavoidable in these contexts; in addition, subsequent movement of XP$_1$ will also have to violate the SCC if it ends up in a lower position that is included in another phrase. Again, a declarative reconstruction of this analysis (that maintains its gist) does not seem to be forthcoming.

2.2.4. Dilemma

To sum up so far, in a *non-local derivational* approach like standard Principles-and-Parameters theory, the freezing, anti-freezing and Müller-Takano generalizations for α-over-β configurations can straightforwardly be derived by the unstipulated interaction of three well-established constraints (CED, SCC, MLC) – assuming, that is, that there are no intermediate traces (except perhaps for those in SpecC postulated for non-clause bound movement). Unfortunately, this is not the case anymore if a *local derivational* approach is adopted, as it is required under more recent minimalist assumptions.

Given the Phase Impenetrability Condition (PIC; Chomsky (2001)), XP movement from

a phase can only take place via its specifier, and given that vP and CP are phases, at least some of the relevant movement types will have their eventual landing sites in an area that is beyond the minimal phase in which they originate. If more XPs qualify as phases, more movement types will have their ultimate landing sites in a higher phase. In what follows, I will in fact assume that *all* phrases are phases, which will then amplify (but not substantially change) the problem that arises with α-over-β configurations in a strictly local approach. And the problem is this:

In the legitimate cases (anti-freezing), extraction of XP_1 from XP_2 will have to take place immediately to an intermediate phase edge position, before XP_2 undergoes an intermediate movement step itself. Thus, suppose that Y is the first phase head above the base position of XP_2, such that XP_1 (β) and XP_2 (α) must both move to SpecY because of the PIC (given a non-recursive definition of edge, pace Richards (2010; 2011)). Then, an account of the anti-freezing effect in remnant movement constructions would seem to necessitate a derivation as in (11), with XP_1 moving first to SpecY, and XP_2 moving to another SpecY position after that, in line with the CED.[9]

(11) *Anti-freezing, first intermediate steps*:
   a. $[_{Y'} \; Y \; [_{XP_2{}^a} \; XP_1{}^b \; [_{X_2{}'} \; X_2 \; ... \;]] \;]]$
   b. $[_{Y'} \; XP_1{}^b \; [_{Y'} \; Y \; [_{XP_2{}^a} \; t_1 \; [_{X_2{}'} \; X_2 \; ... \;]] \;]]$
   c. $[_{YP} \; [_{XP_2{}^a} \; t_1 \; [_{X_2{}'} \; X_2 \; ... \;]] \; [_{Y'} \; XP_1{}^b \; [_{Y'} \; Y \; t_2 \;]]]$

In contrast, in the illegitimate cases (freezing and Müller-Takano generalization), it looks as though extraction of XP_1 from XP_2 will have to follow the first intermediate movement step of XP_2, so as to produce a CED violation. This is shown in (12) and (13), respectively.

(12) *Freezing, first intermediate steps*:
   a. $[_{Y'} \; Y \; [_{XP_2{}^b} \; XP_1{}^a \; [_{X_2{}'} \; X_2 \; ... \;]] \;]]$
   b. $[_{Y'} \; [_{XP_2{}^b} \; XP_1{}^a \; [_{X_2{}'} \; X_2 \; ... \;]] \; [_{Y'} \; Y \; t_2 \;]]$
   c. *$[_{YP} \; XP_1{}^a \; [_{Y'} \; [_{XP_2{}^b} \; t_1 \; [_{X_2{}'} \; X_2 \; ... \;]] \; [_{Y'} \; Y \; t_2 \;]]$

(13) *Müller-Takano generalization, first intermediate steps*:
   a. $[_{Y'} \; Y \; [_{XP_2{}^a} \; XP_1{}^a \; [_{X_2{}'} \; X_2 \; ... \;]] \;]]$
   b. $[_{Y'} \; [_{XP_2{}^a} \; XP_1{}^a \; [_{X_2{}'} \; X_2 \; ... \;]] \; [_{Y'} \; Y \; t_2 \;]]$
   c. *$[_{YP} \; XP_1{}^a \; [_{Y'} \; [_{XP_2{}^a} \; t_1 \; [_{X_2{}'} \; X_2 \; ... \;]] \; [_{Y'} \; Y \; t_2 \;]]$

However, the problem is that it looks as though the decision must be made at a point in the derivation (viz., at the YP level) when the relevant information (concerning where XP_1 and XP_2 will eventually end up, and concerning the question of whether XP_1 and XP_2 will eventually check the same kind of movement-related feature or not) does not yet seem to be present: In all three cases, there is just PIC-driven intermediate movement of both XP_1 and XP_2. This problem could only be solved if look-ahead were permitted in syntax (in violation of locality). On the other hand, at a later point, the relevant distinction is lost; both XPs are in specifiers of phase edges from the first cycle onwards, and invoking the CED will be impossible from then on. This problem could only be solved if back-tracking were permitted in syntax (in violation of both locality and strict cyclicity).[10]

---

[9] Here and in the following two subderivations, the letters "a" and "b" stand for different movement-related features that will eventually be checked in a higher position but are not checked in the intermediate positions. The sole purpose of these symbols is to enhance perspicuity; they play no role in the analysis.

[10] One might think that reference to the movement-related feature on the intermediately moved item ("a" and "b" in the above subderivations) could help to avoid these problems. This is not the case. Even if one were to assume that edge features triggering intermediate movement steps to phase edges are sensitive to the nature of

For these reasons, we end up with a severe dilemma: Core properties of movement in α-over-β configurations do in fact not follow anymore if a local derivational (e.g., phase-based) approach is adopted. In view of this, I will develop a new analysis, one that does not account for the three generalizations in terms of the CED anymore (because this constraint's activity window in α-over-β configurations comes too early). I take it to be unavoidable (in a local derivational approach) that the relevant information needed to ultimately decide whether movement in α-over-β configurations is legitimate or not does indeed (potentially) arise only late in the derivation, when a moved item has reached its ultimate, criterial landing site. Against this background, I would like to suggest that crucial aspects of (recent) movement steps in α-over-β configurations are recorded on a *buffer*, in the form of a list that acts as the value of the movement-related feature.

3. A Local Derivational Approach to α-over-β Configurations: Counter-Feeding and Feeding

3.1. *Defective Valuation of Movement-Related Features*

The starting point of the new analysis is the hypothesis that remnant movement is not completely unproblematic from a theory-internal point of view (pace Stabler (1999), Koopman & Szabolcsi (2000)). On this view, languages ideally want to do without situations in syntax where some item β moves out of a category α that itself needs to undergo movement, because of the intermingling of dependencies and the potential ambiguities that ensue. However, this does not mean that remnant movement is excluded, and that α-over-β configurations are ruled out by stipulation as possible sources of well-formed derivations; it just means that there is a price to pay. For concreteness, I would like to suggest that if β moves out of α in an α-over-β configuration like (5), α is *contaminated*: β provides a defective value for α's movement-related feature (e.g., [wh], [top], [Σ]), which invariably brings about a crash of the derivation if it is not removed in time, before a criterial position is reached; thus, the movement-related feature acts as a buffer that stores a crucial aspect of an earlier part of the syntactic derivation. This does not keep β from undergoing movement itself; a *temporary* contamination of a movement-related feature is unproblematic as long as a criterial position has not yet been reached. A moved item β can in principle *decontaminate* a category α again by removing the defective symbol; but this only happens when β itself reaches a criterial position, under c-command. Thus, the *timing* of movement steps of α and β will be crucial. Criterial remnant movement of α is legitimate if β has been able to remove the fatal value from α's feature list before the criterial movement step; otherwise criterial remnant movement of α is illegitimate.

At this point, two questions arise. First, what is this feature value that turns an XP α from which extraction of β has taken place into an illegitimate item? And second, why does such a feature value of a movement-related feature on α lead to illformedness unless it is deleted before α reaches a criterial position? The answer to the first question that I would like to give here is that the incriminating feature value that α gets from β when β moves out of α is β's *index*.[11] Consequently, contamination and decontamination of movement-related features in

the criterial position that the movement will ultimately end up in (see Abels (2012) for such a proposal), this could at best help with the Müller-Takano generalization, not with freezing and anti-freezing: In these latter cases, the crucial factor is the *relative height* of the two criterial positions (which do not even need to be in the same clause), and this information cannot be encoded by a single feature.

[11] This implies that I do not follow Chomsky (1995) in assuming that indices on syntactic categories do not exist. However, indices seem to be needed anyway for semantic interpretation (see Heim & Kratzer (1998)); and it is also worth pointing out that indices do not violate the Inclusiveness Condition or the No Tampering Condition, given that they are present before the syntactic derivation starts. That said, other options would be

α-over-β configurations can be defined as in (14).

(14)   a.   *Contamination*:
             Movement of β from a position within α to a position outside of α values a
             movement-related feature γ on α with β's index.
        b.   *Decontamination*:
             Movement of β to a criterial position deletes β's index on all movement-related
             features of items that c-command it.

Removal of a defective value under c-command in (14-b) can be viewed as an instance of
Agree, with the feature bearing the defective value on α acting as a probe. Crucially, this
only becomes possible when β has reached a criterial position; before that, β's index does
not qualify as a proper goal.

    Now that it has been clarified how α registers extraction of β in α-over-β configurations
by acquiring β's index as a value of its movement-related feature, the second question can
be addressed: How can such a feature value lead to illformedness? For present purposes, I
will simply assume that this is due to a specific constraint stating that such values are not
tolerated in criterial positions; see the Index Filter in (15).[12]

(15)   *Index Filter*:
        A movement-related feature (like [wh], [top], [Σ]) must not have an index as (part
        of) its value in a criterial position.

It remains to be shown how the Index Filter (ultimately, the Williams Cycle, cf. footnote 12)
is violated in freezing and Müller-Takano generalization environments (because the incrim-
inating index on the remnant XP α's movement-related feature has not yet been removed
when α reaches a criterial position), and respected in anti-freezing contexts (because the in-
criminating index has successfully been removed when α enters a criterial position). Thus,
everything now boils down to the *timing* of syntactic operations. As I will argue in the next
subsection, independently motivated constraints on the timing of syntactic operations cor-
rectly predict *feeding* in the good (anti-freezing) contexts, and *counter-feeding* in the bad
ones (freezing, Müller-Takano).

---

readily available, as long as the value on α's movement-related feature that α gets from β uniquely identifies
β. (Note incidentally that this role cannot be played by *category labels*; otherwise any category with the same
label as β could decontaminate α, not just β.)

[12] This is a gross simplification. Ultimately, the Index Filter should be derivable from more basic assumptions.
In line with this, note that I have been silent so far as to what a regular, well-formed value for movement-related
features is. In Müller (2014), I argue that the value of movement-related features is not a simple symbol like "+"
or "–", but rather a list – more precisely, a queue of category symbols (a first-in/first-out list) that is successively
generated by movement steps, recording all category labels passed by the moved item on its way to the ultimate
landing site and deleting symbols at the bottom of the list under identity with new incoming symbols. This way,
proper and improper movement can be locally determined (in improper movement configurations, the symbol
list that acts as the value of a movement-related feature on some moved item does not conform to the functional
sequence (f-seq) of heads in a clause); essentially, this approach makes it possible to formulate a local version
of the Williams Cycle approach to improper movement (Williams (1974; 2003)): This constraint then states
that categorial information on a list of a movement-related feature must conform to f-seq when a moved item
reaches a criterial position.
Against the background of this analysis, it is clear why the index of β as (part of) a value of a movement-related
feature on α will lead to illformedness in criterial positions: Such an index invariably brings about an f-seq
violation because it is not a category label and can therefore never show up in any well-formed f-seq (which
consists only of category labels by definition). In sum, then, the Index Filter can be derived as a theorem of a
local derivational approach to improper movement based on the Williams Cycle.

3.2. On the Order of Movement Operations

Minimalist syntax envisages certain basic operations, most notably Merge ('external Merge'), Move ('internal Merge') and Agree. These operations may interact with one another in syntactic derivations, thereby potentially giving rise to feeding, bleeding, counterfeeding and counter-bleeding relations. The question then arises of whether there are restrictions on the order of elementary operations. This issue is highly relevant in derivational approaches to syntax since resolving the order in one way or the other may make radically different empirical predictions. A well-known example involves English expletive constructions; consider the sentence pair in (16).

(16)   a.   There$_1$ seems [$_{TP}$ t$_1$ to be [$_{PP}$ someone$_2$ in the room]]
        b.   *There$_1$ seems [$_{TP}$ someone$_2$ to be [$_{PP}$ t$_2$ in the room]]

As observed by Chomsky (2000), if Merge (of the expletive *there*$_1$) precedes Move (of *someone*$_2$) on the TP cycle, (16-a) is predicted to be the correct output; in contrast, if Move precedes Merge on the TP cycle, we expect (16-b) to be well formed. The distribution of grammaticality in (16) can therefore be taken to support a general principle *Merge over Move*. Closer scrutiny reveals a surprising number of general principles determining the order of grammatical operations that have been suggested over the last decades.[13]

   In what follows, I will make the following assumptions about the order of operations in the case of multiple movement to phase edges.

(17)   *Movement to phase edge*:
        a.   If α c-commands β in the pre-movement structure, then α moves first, and β moves after that, to a *lower* specifier.
        b.   If α does not c-command β in the pre-movement structure, the order is not fixed; the second item that moves ends up in a *higher* specifier.

(17-a) states that in cases of multiple movement of the same type, of items that are in a c-command relation, the derivation proceeds by tucking in; see Richards (2001) and Branigan (2013), among many others (and Fox & Pesetsky (2005), Stroik (2009), Unger (2010) and Assmann & Heck (2013) for related concepts). (17-b) will permit movement of β to apply first in α-over-β environments (where α, β both initially undergo intermediate movement), which is a precondition for CED satisfaction of any derivation in which this configuration occurs.[14] The assumptions in (17) can arguably be shown to follow from more elementary concepts, but for present purposes I will simply take them as given.[15] On this basis, let

---

[13] Here are some proposals: (i) *extrinsic ordering* (Chomsky (1965), Perlmutter & Soames (1979)), possibly with free variation to capture cross-linguistic variation (Georgi (2014)); (ii) *obligatoriness vs. optionality* (Pullum (1979)); (iii) *specificity* (Sanders (1974), Pullum (1979), van Koppen (2005), Lahne (2012), Georgi (2013)); (iv) *anti-specificity* (Chomsky (2000; 2001) – note that ordering Merge before Move has originally been justified by assuming that Merge is the more general, i.e., less specific, operation); (v) *strict cyclicity* (McCawley (1984; 1998)); (vi) *strata/levels* (Chomsky (1981): D-structure, S-structure, LF/PF; Riemsdijk & Williams (1981): NP-structure); (vii) *rule vocabulary* (Arregi & Nevins (2012, ch. 6)); (viii) *minimal search* (Chomsky (2013)).

[14] Thus, the MLC must not force movement of α in [$_\alpha$ ...β...] configurations, at least not if both items undergo intermediate movement steps driven by edge features. More generally, the MLC will play no role anymore in the account of the Müller-Takano generalization to be developed below. See Müller (2011) for arguments against the MLC, and for a proposal of how to derive most of the intervention effects that it is supposed to cover.

[15] See Müller (2013) for discussion. Basically, (17) can be understood in such a way that it brings about a *minimization of changes to existing structures*, as required under a (non-categorical) version of the No Tampering Condition (NTC, Chomsky (2007; 2008; 2013)) that incorporates Pullum's (1992) assumptions about the ori-

me now turn to the three generalizations about movement in α-over-β configurations. It will turn out that the interaction of the assumptions about the order of criterial and intermediate movement in (17) and the assumptions about contamination and decontamination of movement-related features in (14) correctly predicts the distribution of anti-freezing, freezing, and Müller-Takano effects in α-over-β configurations.

3.3. Feeding vs. Counter-Feeding: Generalizations Derived

3.3.1. Initial Steps

To begin with, recall that I assume that all phrases qualify as phases. Thus, the PIC forces intermediate movement through every intervening phrase edge domain on the way to the criterial landing site of a moved item. Furthermore, the CED and the SCC continue to hold (in contrast to the MLC).

In all cases, the decisive stage of the derivation starts when the higher XP_2 (α) has merged with a head Y, as a complement, where XP_1 (β) has earlier undergone movement to XP_2's specifier; cf. (18-a). In the first step, XP_1 must now move out of XP_2, to SpecY (if XP_2 moves first, the CED will be violated); cf. (18-b). Such movement is typically unproblematic because a fixed order is not required here; cf. (17-b). However, movement of XP_1 contaminates XP_2 by valuing X_2's movement-related feature (γ) with XP_1's index, thereby creating a situation that must be remedied before XP_2 reaches a criterial position. In the next step, XP_2 moves to an outer SpecY (because of PIC and (17-b)); cf. (18-c), with the movement-related feature γ of XP_2 now having a contaminated value: [γ:1].

(18)   *Initial steps in α-over-β configurations*
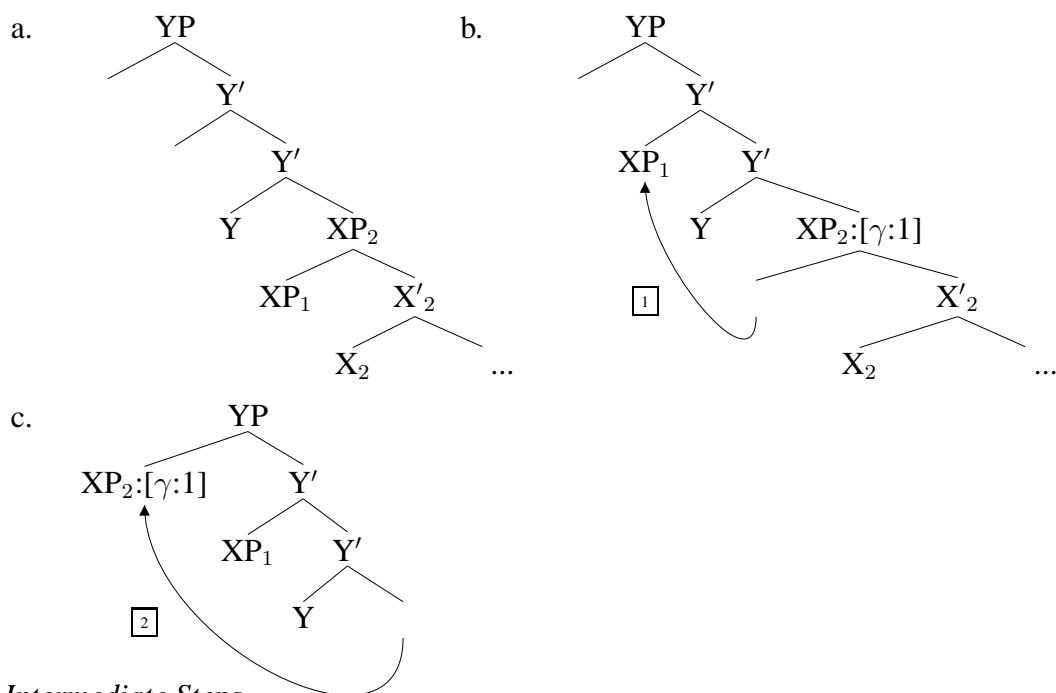
3.3.2. Intermediate Steps

Consider next what happens on intermediate cycles, where both XP_2 (α) and XP_1 (β) undergo intermediate (edge feature-driven) movement to specifier domains, as required by the PIC. Given (17-a), the two intermediate movement steps must be order-preserving, with the first, higher item (the remnant category XP_2) moving first, and the second, lower item (XP_1)

---

gins of cyclicity: Once established, linear order is preserved throughout the derivation as much as possible (i.e., as long as triggers for movement can be satisfied, etc.), even at the cost of counter-cyclic tucking in operations (cf. (17-a)); but such violations are not permitted when maintaining linear order is not an issue (cf. (17-b)).

moving afterwards to a lower specifier, by tucking in. Of course, this pattern can be applied recursively, leading to order-preservation with multiple movement spanning arbitrarily long distances.

(19) *Intermediate steps in α-over-β configurations*
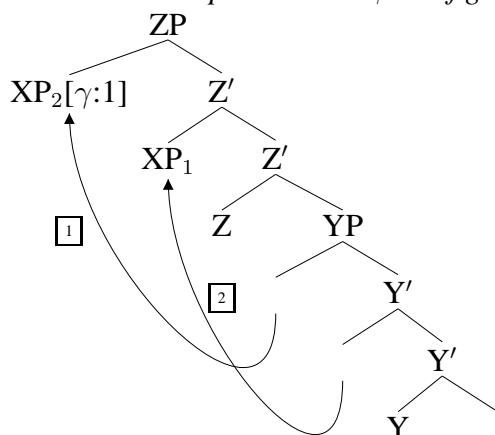
Finally, building on either (18-c) or (19), criterial movement steps of either XP_1 (β), XP_2 (α), or both XP_2 and XP_1 can take place, giving rise to anti-freezing, freezing, and Müller-Takano effects.

3.3.3. Anti-Freezing

In the case of anti-freezing as in typical remnant movement constructions like those in (9) (with a relevant example repeated in (21)), again XP_2 undergoes intermediate movement first. Subsequently, criterial movement of XP_1 takes place to an inner specifier, which then removes XP_1's index from XP_2. XP_2 is free to undergo criterial movement in accordance with the Index Filter (ultimately, the Williams Cycle) from now on. This *feeding* effect in anti-freezing configurations (decontamination feeds criterial remnant movement) is shown in (20) (1 signals intermediate movement of XP_2, 2 subsequent criterial movement of XP_1; and 3 index removal (while XP_1 is c-commanded by XP_2); a box around a category indicates that the category has reached a criterial position).

(20) *Criterial steps in α-over-β configurations: XP_1*
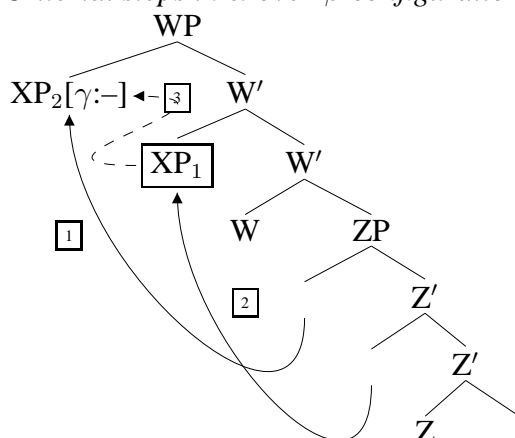
(21) [$_{VP_2}$ t_1 Gelesen] hat das Buch$_1$ keiner t_2
          read   has the book   no-one

3.3.4. Freezing

In contrast, freezing configurations as in (6) (also see (23)) involve *counter-feeding*: Criterial movement of XP_1 comes *far* too late to remove the fatal index from XP_2: XP_2 undergoes criterial movement when it still has a contaminated value on its movement-related feature and thereby violates the Index Filter; see (22). Here, XP_2 has a defective value 1 in a criterial position, and XP_1 is far from being able to remove the incriminating symbol 1 because it has undergone intermediate rather than criterial movement itself. In addition, note that movement of XP_1 would have come too late anyway (this will become relevant below).

(22)   *Criterial steps in α-over-β configurations: XP_2*
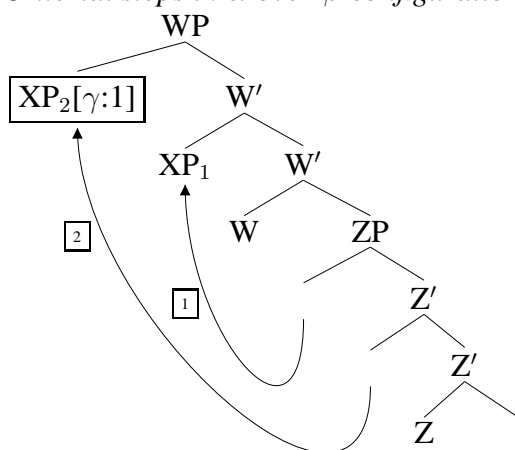
(23)   *Was$_1$ denkst du [$_{VP_2}$ t$_1$ gelesen] hat keiner t$_2$?
           what think you   read   has no-one

3.3.5. Müller-Takano Generalization

The third possible continuation of a derivation involving criterial movement is that both XP_2 and XP_1 undergo criterial movement to a given specifier domain. This is the situation underlying Müller-Takano effects as in (10) (a relevant example is repeated here as (25)), and it also involves *counter-feeding* under present assumptions: This time, criterial movement of XP_1 comes *a bit* too late to be able to remove the fatal index from XP_2: Given (17-a), XP_2 undergoes criterial movement first and thereby violates the Index Filter; cf (24). Subsequent criterial movement of XP_1 (via tucking in) creates a configuration in which the defective index on XP_2 could be removed (signalled by ③), but at this point of the derivation, the damage has already, and irrevocably, been done.

(24)   *Criterial steps in α-over-β configurations: XP_1 & XP_2*
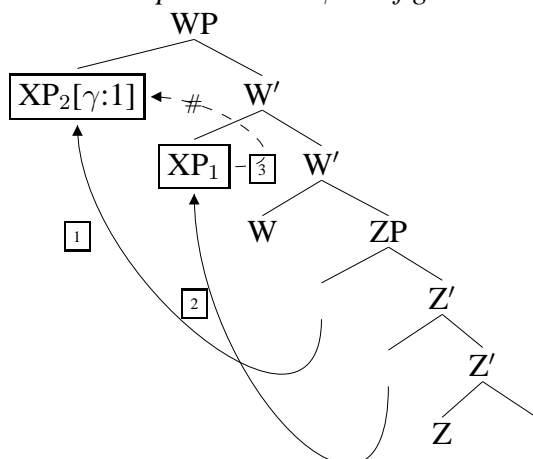
(25)   *dass [$_{VP_2}$ t_1 zu lesen] [$_{DP_1}$ das Buch   ] keiner t_2 versucht hat
          that   to read   the book$_{acc}$ no-one   tried   has

3.3.6.   Criterial Initial Steps

The discussion so far has presupposed that the initial operations in (18-a) are two interme-
diate movement steps, and this is arguably the prototypical situation. Still, it is worth in-
vestigating whether the conclusions concerning anti-freezing, freezing, and Müller-Takano
effects can be maintained if XP_1, XP_2, or both XP_1 and XP_2, undergo extremely local crite-
rial movement to SpecY positions in (18).[16]

There are three cases to be addressed. First, XP_2 undergoes criterial movement to SpecY
in (18-a), and XP_1 undergoes intermediate movement. If XP_2 moves first, and XP_1 moves
after that, this violates the CED (it does not produce a counter-feeding effect of the same type
as seen in (22) because the value of XP_2's movement-related feature is not yet contaminated
by extraction of XP_1). Alternatively, XP_1 is extracted first, thereby contaminating XP_2's
movement-related feature with its index. Subsequent criterial movement of XP_2 will then
violate the Index Filter.

The second possibility is that XP_1 undergoes criterial movement to SpecY in (18-a),
and XP_2 undergoes intermediate movement. XP_1 extraction can (given (17-b)) and must
(given the CED) apply first, which instantiates XP_1's index on the value of XP_2's movement-
related feature. Next, XP_2 undergoes intermediate movement to a higher position, and the
incriminating symbol is deleted again on XP_2, under c-command. So, this derivation is
legitimate, as suggested by the empirical evidence. For instance, a sentence like (21) is also
possible in German if the scrambled XP_1 follows rather than precedes the subject DP, which
may in turn be assumed to be in situ, in Specv (as evidenced by the option of a preceding
unstressed pronoun in minimally different double object constructions; see section 4 below);
cf. (26).

(26)   [$_{VP_2}$ t_1 Gelesen] hat keiner das Buch$_1$ t_2
            read   has no-one the book

The third and final possibility to be discussed here involves extremely local criterial move-
ment of both XP_1 and XP_2 to SpecY in (18-a). Again, XP_1 needs to extract first (because of
the CED), which contaminates XP_2. Since decontamination can only take place when XP_2
c-commands XP_1, XP_2 must move next, thereby immediately giving rise to a violation of the
Index Filter. Thus, the order of operations is (i) criterial movement of XP_1, (ii) valuation of
XP_2 with XP_1's index, (iii) criterial movement of XP_2 (giving rise to a violation of the Index
Filter), and (iv) deletion of XP_1's index on XP_2 (which comes too late).

To sum up, assuming that remnant-creating movement of β in α-over-β configurations
does not come entirely for free but involves the creation of a temporary contaminated value
on the movement-related feature of the remnant XP α which must eventually be removed by

---

[16] Such a situation cannot occur if extremely local criterial movement is blocked; see Bošković (1997), Abels
(2003; 2012), and Grohmann (2003)) for proposals to this effect. Throughout the present paper, the analyses
presuppose that extremely local intermediate movement is available (given that every phrase is a phase); but
this does not necessarily imply that extremely local criterial movement is also an option. As a matter of fact,
given that the structure-building features triggering the operation are provided in the course of the derivation in
the former case, whereas they are intrinsically present on heads in the latter case, one might plausibly argue that
whatever constraint blocks extremely local criterial movement (e.g., economy, as in Abels (2003)) does not also
have to block extremely local intermediate movement. Still, these considerations notwithstanding, in view of
the apparently highly local nature of some of the criterial scrambling operations involved in relevant examples,
I will assume in what follows that extremely local movement is an option, in both criterial and intermediate
contexts; i.e., there is no constraint whatsoever demanding anti-locality.

criterial movement of β, the freezing, anti-freezing, and Müller-Takano generalizations can all be shown to follow from independently established restrictions on multiple movement: On this view, examples subsumed under the freezing and Müller-Takano generalizations emerge as counter-feeding effects, with criterial movement of β applying too late in the derivation.

4. Consequences

The analysis makes a number of further predictions. I will discuss three of them, concerning (i) multiple remnant movement, (ii) the scope of Müller-Takano effects, and (iii) temporary defectivity.

4.1. Multiple Extraction, Multiple Remnant Movement

The analysis is compatible with multiple extraction from a remnant XP, as in (27). Here, there is both pronominal object shift of *es*$_1$ ('it') and scrambling of *dem Fritz* ('the Fritz'), and both items are extracted from VP$_2$ while the latter is still in situ, thereby contaminating it with their separate indices (recall that the value of a movement-related feature is a list). Both indices are subsequently removed from VP$_2$ when DP$_1$ and DP$_2$ reach their respective criterial landing sites, and when VP$_2$ reaches the topic position, it has the value of its movement-related feature restored to normalcy.

(27)   [$_{VP_3}$ t$_2$ t$_1$ Gegeben] hat es$_1$ die Maria   am Ende [$_{DP_2}$ dem Fritz] t$_3$
                         given   has it$_{acc}$ the Maria$_{nom}$ in the end   the Fritz$_{dat}$

The analysis is also compatible with multiple remnant movement, as in examples like (28-a) vs. (28-b) in German (see den Besten & Webelhuth (1990)).

(28)   a.   [$_{VP_3}$ t$_2$ Gerechnet] hat da$_1$ wie immer keiner t$_3$ [$_{PP_2}$ t$_1$ mit]
                       counted   has there as   always no-one   with
        b. *[$_{VP_3}$ t$_2$ Gerechnet] hat da$_1$ wie immer [$_{PP_2}$ t$_1$ mit] keiner t$_3$
                       counted   has there as   always   with no-one

Under the analysis in Müller (1998), (28-a) involves a combination of scrambling of *da*$_1$ from PP$_2$, (globally string-vacuous) extraposition of the remnant PP$_2$ from VP$_3$, and topicalization of the remnant VP$_3$. Given the present assumptions, this derivation is unproblematic (an instance of anti-freezing): The R-pronoun *da*$_1$ is moved from PP$_2$ while the latter is in situ, thereby creating a defective value on PP$_2$'s extraposition feature that can be removed after *da*$_1$ has reached its criterial scrambling position, before PP$_2$ has reached the (higher) extraposition position. Furthermore *da*$_1$ and *PP*$_2$ both undergo movement from VP$_3$ while the latter is still in situ, thereby contaminating it with their indices; but decontamination can take place for both defective values of VP$_3$'s [top] feature before VP$_3$ actually reaches the criterial position (which is higher than that of PP$_2$). In contrast, in (28-b), PP$_2$ undergoes scrambling, and reaches this position before *da*$_1$ shows up in *its* criterial position and has a chance to remove the illicit feature value on PP$_2$. Thus, there is a freezing effect.

4.2. The Scope of Müller-Takano Effects

The present approach makes a prediction that is different from the accounts of Müller-Takano generalization effects highlighted in section 2 above. The prediction is that remnant movement should be possible, in violation of the Müller-Takano Generalization as it is formulated above, if the remnant XP$_2$ has the same movement-related feature as XP$_1$, but checks this with some higher head in the clause. The reason is that ungrammaticality can only occur

if the feature is checked in the same domain; if XP_2 does not check the feature in the same phase edge as XP_1, it can be decontaminated before it reaches its criterial position.[17] At present, I take it to be an open question whether this loophole for identical movement operations in α-over-β configurations is desirable or not. A straightforward way to close it would be to assume that there is a one-to-one correspondence between movement-attracting features and the functional heads on which they occur. Still, an approach recognizing this loophole might possibly receive empirical confirmation. For instance, Fanselow (2002) and Hale & Legendre (2004) argue that sentences like (25) (repeated here as (29)) improve with a rise/fall (I-topicalization) intonational pattern that might indicate that a higher functional head is targetted by the VP fronting operation than by prior DP scrambling (see Frey (2000; 2004)).

(29) *dass [$_{VP_2}$ t_1 zu lesen] [$_{DP_1}$ das Buch   ] keiner t_2 versucht hat
          that   to read   the book$_{acc}$ no-one   tried   has

If different movement-triggering features are involved in the two positions targetted by DP_1 and VP_2, both analyses of Müller-Takano effects discussed in the present paper would be able to accomodate a possible amelioration of (29) under I-topicalization. If, however, the same movement-inducing feature is involved (viz., a feature that uniformly triggers all kinds of scrambling in German, like [Σ]), then an improvement under I-topicalization would follow under the local derivational approach in terms of contamination and decontamination, but not under the non-local derivational approach based on the MLC.[18]

4.3. Temporary Defectivity

As a consequence of the buffer-based analysis of α-over-β constructions, a remnant XP_2 (α) is not uniformly the same kind of syntactic object throughout a derivation. At the very beginning, before extraction from it takes place, α is a legitimate object; and it also qualifies as a legitimate object as soon as the item that has undergone movement out of XP_2 (viz., β) has reached a criterial position, and index removal is effected. However, α is defective in the intermediate parts of the derivation, after extraction of β and before the criterial movement step of β. Therefore, it seems plausible to venture the hypothesis that the temporary defectivity of a remnant XP in the middle of its active life cycle should be reflected in a special behaviour with respect to other syntactic processes like binding and scope.

Consider binding first. For concreteness, suppose that a temporarily defective remnant category α cannot participate in the computation of binding principles. With this in mind, consider the sentence in (30).

(30) ?[$_{VP_2}$ Dem Peter t_3 zu geben] weiß ich nicht [$_{CP_4}$ was$_3$ Maria   denkt [$_{CP_5}$
              the Peter$_{dat}$ to give   know I   not   what$_{acc}$ Maria$_{nom}$ thinks
          dass Karl t_2 versuchen sollte]]
          that Karl   try   should

In (30), there is long-distance wh-movement of DP_3 to the least deeply embedded SpecC position, plus long-distance topicalization of the remnant VP_2. The resulting sentence is only mildly degraded, like other cases of (non-adjunct) topicalization across a wh-island in German. The present analysis predicts that, as a result of contamination by DP_1 extraction, VP_2 has a defective value on its movement-related feature ([top]) in the subpart of the derivation

---

[17] Note that this consequence is in fact one that also arises in Takano's (1994) original approach.

[18] Similar conclusions hold for allegedly acceptable cases of dislocation of remnant VPs; cf. Fanselow (2002) vs. Müller (1998).

that starts with the most deeply embedded vP and ends with CP_4. Given the above assumption about the consequences of temporary defectivity for binding theory, the prediction is that the R-expression *dem Peter*, as part of the remnant VP_2, cannot participate in the computation of binding principles at these stages of the derivation. More specifically, it is predicted to be immune to Principle C effects at these stages (but not before or after these stages). Interestingly, this prediction seem to be borne out. This is shown by the three examples in (31).

(31)   a. *[$_{VP_2}$ Dem Peter$_1$ t_3 zu geben] weiß ich nicht [$_{CP_4}$ was$_3$ Maria   denkt
               the   Peter$_{dat}$ to give   know I   not   what$_{acc}$ Maria$_{nom}$ thinks
          [$_{CP_5}$ dass er$_1$ (selbst) t_2 versuchen sollte]]
              that he (self)   try   should

    b. ?[$_{VP_2}$ Dem Peter$_1$ t_3 zu geben] weiß ich nicht [$_{CP_4}$ was$_3$ er$_1$ (selbst) denkt
               the   Peter$_{dat}$ to give   know I   not   what$_{acc}$ he (self)   thinks
          [$_{CP_5}$ dass Maria t_2 versuchen sollte]]
              that Maria$_{nom}$ try   should

    c. *[$_{VP_2}$ Dem Peter$_1$ t_3 zu geben] weiß   er$_1$ (selbst) nicht [$_{CP_4}$ was$_3$
               the   Peter$_{dat}$ to give   knows he (self)   not   what$_{acc}$
         Maria   denkt [$_{CP_5}$ dass ich t_2 versuchen sollte]]
         Maria$_{nom}$ thinks   that I   try   should

In (31-a), there is a clear Principle C effect, in addition to the mild wh-island effect. Here, the incriminating co-indexed subject pronoun *er* ('he') shows up in the most deeply embedded CP_5; it is merged in Specv *before* DP_3 and VP_2 move to (outer) specifiers of vP (given Merge over Move). This means that *dem Peter*$_1$ violates Principle C before DP_1 is extracted from VP_2, contaminating it with its index and rendering it defective.[19]

In (31-c), there is also a clear Principle C effect. Here, the incriminating co-indexed pronoun *er* ('he') shows up in the matrix vP; it successfully binds *dem Peter* in VP_2 *after* DP_3 (*was*, 'what') has reached its criterial position in CP_4.

However, there is a subtle but clearly discernible improvement in (31-b). Here, the subject pronoun *er* ('he') shows up in the intermediate clause, CP_4; and throughout the derivational stages where the pronoun might effect c-command of the co-indexed DP *dem Peter*, the latter is sheltered by the defective remnant VP_2. Thus, we get a reconstruction effect in (31-ac) and an anti-reconstruction effect in (31-b) that follows straightforwardly under present assumptions but must, it seems, remain a complete mystery under virtually all other theories of movement.

Next consider scope. According to *Barss's generalization* (see Barss (1986), Lechner (1998), Sauerland & Elbourne (2002), Bhatt & Dayal (2007), Neeleman & van de Koot (2010), Heck & Assmann (2012)), a quantified item in a moved remnant XP α cannot take scope, via reconstruction, over an item β that has undergone movement from α. Thus, (32-a) is ambiguous, but (32-b), with remnant VP topicalization, is not: The dative DP *jedem Studenten* ('every student') cannot take scope over the accusative DP *ein Buch* ('a book').

(32)   a. Jedes Buch   hat sie   einem Studenten gegeben   ($\forall > \exists$)
          every book$_{acc}$ has she$_{nom}$ a   student$_{dat}$ given

    b. Jedem Studenten gegeben hat sie   ein Buch   (*$\forall > \exists$, $\exists > \forall$)
          every   student$_{dat}$ given   has she$_{nom}$ a   book$_{acc}$

---

[19] The example is pragmatically complex but by no means implausible: I don't know for which thing (e.g. a price, an assignment, a job) it is the case that Maria believes that Peter should try to give it to himself.

Suppose now that a temporarily defective XP_2 in intermediate stages of a derivation (i.e., after contamination, and before decontamination) is not visible for the purposes of Barss's generalization. Then we expect that it can in fact take scope over β in this syntactic domain.[20] Needless to say, relevant examples are not completely trivial to construct, and quite hard to classify. Here is an attempt. First, (33) shows that a remnant VP with a quantified dative object can undergo long-distance topicalization in German (again, with a weak wh-island effect arising).

(33) ?[$_{VP_2}$ Jedem Studenten t_1 gegeben] weiß ich nicht [$_{CP}$ was$_1$ sie t_2 hat]
       every student$_{dat}$ given know I not what$_{acc}$ she has

Second, as observed by Beck (1997), sentences like (34) are ambiguous.

(34) Wieviele Bücher$_1$ hat sie jedem Studenten$_2$ t_1 gegeben ?   ($\exists > \forall, \forall > \exists$)
      how many books$_{acc}$ has she every student$_{dat}$ given

(34) can mean: For which number n: There are n books that she gave to every student ($\exists > \forall$). Alternatively, it can mean: For which number n: For every student there are n books that she have to him ($\forall > \exists$).

Third, by combining the two contexts, we expect a Barss's generalization effect. As shown in (35), this does indeed seem to be the case; (35) can hardly have a reading where *jedem Studenten* ('every student') takes scope over the existential quantifier.[21]

(35) ??[$_{VP_2}$ Jedem Studenten$_3$ t_1 gegeben] weiß ich nicht [$_{CP}$ wieviele Bücher$_1$ sie t_2
      every student$_{dat}$ given know I not how many books$_{acc}$ she
    hat]   (*$\forall > \exists$)
    has

Against this background, consider the example in (36). Given the present assumptions, the prediction is that a Barss's generalizations effect could be absent here, i.e., that the sentence can have a reading with the universal quantifier outscoping the existential quantifier since the remnant VP_2 is defective at the relevant stages of the derivation (i.e., in the intermediate CP).

(36) ??[$_{VP_2}$ Jedem Studenten$_3$ t_1 gegeben] weiß ich nicht [$_{CP}$ wieviele Bücher$_1$ Fritz
      every student$_{dat}$ given know I not how many books$_{acc}$ Fritz
    denkt [$_{CP}$ dass sie t_2 hat]
    thinks that she has

A similar, parallel example pair is given in (37). Again, the prediction would be that (37-b) can have a wide-scope reading for the universal quantifier more easily than (37-a).

(37) a. ??[$_{VP_2}$ Jedem Besucher$_3$ t_1 zu zeigen] weiß ich nicht [$_{CP}$ wieviele Bilder$_1$
        every visitor$_{dat}$ to show know I not how many pictures$_{acc}$
       sie t_2 versuchen will]
       she try wants

---

[20] In the approach of Heck & Assmann (2012), this could be so because a temporarily defective XP_2 might exceptionally make acyclic reconstruction possible, where reconstruction normally has to be cyclic.

[21] For reasons unclear to me, the wh-island effect seems to be stronger with *wieviele Bücher* ('how many books') than with *was* ('what'), which is an orthogonal factor that nevertheless seems to complicate judgements somewhat.

b. ??[$_{VP_2}$ Jedem Besucher$_3$ t$_1$ zu zeigen] weiß ich nicht [$_{CP}$ wieviele   Bilder$_1$
       every visitor$_{dat}$ to show   know I   not   how many pictures$_{acc}$
    Maria sagt [$_{CP}$ dass sie t$_2$ versuchen will]
    Maria says   that she   try   wants

However, since judgements are extremely subtle in this case, and the exact theoretical implementation of Barss's generalization is far from clear (and certainly not innocuous), I will leave it at that, and draw a conclusion.

The gist of the analysis I have developed in this paper is that movement of some category β from another category α that needs to undergo movement itself triggers a contamination of the movement-related feature of α; thus, there is a price to be paid for remnant movement in syntax. A contamination of α is temporarily acceptable, but decontamination must take place before α moves to a criterial position; and the required index removal can only apply when β has moved to a criterial position. On this basis, independently motivated assumptions about the order of movement operations ensure that freezing and Müller-Takano generalization configurations violate the Index Filter (ultimately, a local constraint against improper movement demanding adherence to f-seq; the Williams Cycle); and that anti-freezing configurations do not. Interestingly, there is opacity (and hence a potential argument for a derivational approach to syntax) in both the standard approach to movement in α-over-β configurations sketched in section 2, and the new local approach developed in section 3. However, whereas it is the anti-freezing effect that creates opacity as an instance of counter-bleeding in the former case, it is the freezing and Müller-Takano effects that create opacity as an instance of counter-feeding in the latter case.[22]

---

[22] To be sure, there is still counter-bleeding with remnant movement in the new approach since the CED would be violated with extraction of β after α moves; but this does not distinguish legitimate remnant movement from illegitimate kinds of movement in α-over-β constructions anymore.

Bibliography

Abels, Klaus (2003): Successive Cyclicity, Anti-Locality, and Adposition Stranding. PhD thesis, University of Connecticut, Storrs, Connecticut.

Abels, Klaus (2008): Towards a Restrictive Theory of (Remnant) Movement, *Linguistic Variation Yearbook* 7, 53–120.

Abels, Klaus (2012): *Phases. An Essay on Cyclicity in Syntax*. Vol. 543 of *Linguistische Arbeiten*, De Gruyter, Berlin.

Arregi, Karlos & Andrew Nevins (2012): *Morphotactics: Basque Auxiliaries and the Structure of Spellout*. Springer, Heidelberg.

Assmann, Anke & Fabian Heck (2013): Opaque Intervention. Ms., Universität Leipzig.

Assmann, Anke, Doreen Georgi, Fabian Heck, Gereon Müller & Philipp Weisser (2012): Ergatives Move Too Early. Ms., Universität Leipzig (*Linguistische Arbeits Berichte* 90, 2013). To appear in *Syntax*.

Assmann, Anke, Fabian Heck, Johannes Hein, Stefan Keine & Gereon Müller (2010): Does Chain Hybridization in Irish Support Movement-Based Approaches to Long-Distance Dependencies?. In: S. Müller, ed., *The Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar*. CSLI Publications, Stanford, pp. 27–46.

Asudeh, Ash (2004): Resumption as Resource Managment. PhD thesis, Stanford University.

Barss, Andrew (1984): Chain Binding. Ms., MIT, Cambridge, Mass. (Talk at GLOW).

Barss, Andrew (1986): Chains and Anaphoric Dependence. Ph.d. thesis, MIT, Cambridge, Mass.

Beck, Sigrid (1997): Wh-Constructions and Transparent Logical Form. PhD thesis, Universität Tübingen.

Belletti, Adriana & Luigi Rizzi (1988): Psych-Verbs and Θ-Theory, *Natural Language and Linguistic Theory* 6, 291–352.

Bhatt, Rajesh & Veneeta Dayal (2007): Rightward Scrambling as Rightward Remnant Movement, *Linguistic Inquiry* 38, 287–301.

Boeckx, Cedric, Norbert Hornstein & Jairo Nunes (2010): *Control as Movement*. Cambridge University Press, Cambridge.

Bouma, Gosse, Robert Malouf & Ivan Sag (2001): Satisfying Constraints on Extraction and Adjunction, *Natural Language and Linguistic Theory* 19, 1–65.

Bošković, Željko (1997): *The Syntax of Nonfinite Complementation. An Economy Approach*. MIT Press, Cambridge, Mass.

Branigan, Phil (2013): Cyclicity and the Approach the Probe Principle. Ms., Memorial University of Newfoundland.

Bresnan, Joan (1972): Theory of Complementation in English Syntax. PhD thesis, MIT, Cambridge, Mass.

Bresnan, Joan (1976a): Evidence for a Theory of Unbounded Transformations, *Linguistic Analysis* 2, 353–399.

Bresnan, Joan (1976b): On the Form and Functioning of Transformations, *Linguistic Inquiry* 7, 3–40.

Bresnan, Joan (2001): *Lexical-Functional Syntax*. Blackwell, Oxford.

Browning, Marguerite (1987): Null Operator Constructions. PhD thesis, MIT, Cambridge, Mass.

Browning, Marguerite (1991): Bounding Conditions on Representations, *Linguistic Inquiry* 22, 541–562.

Büring, Daniel (2005): *Binding Theory*. Cambridge University Press, Cambridge.

Campana, Mark (1992): A Movement Theory of Ergativity. PhD thesis, McGill University.

Chomsky, Noam (1951): Morphophonemics of Modern Hebrew. Master's thesis, University of Pennsylvania.

Chomsky, Noam (1965): *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.

Chomsky, Noam (1973): Conditions on Transformations. In: S. Anderson & P. Kiparsky, eds., *A Festschrift for Morris Halle*. Academic Press, New York, pp. 232–286.

Chomsky, Noam (1975): *The Logical Structure of Linguistic Theory*. Plenum Press, New York.

Chomsky, Noam (1981): *Lectures on Government and Binding*. Foris, Dordrecht.

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

Chomsky, Noam (1995): *The Minimalist Program*. MIT Press, Cambridge, Mass.

Chomsky, Noam (2000): Minimalist Inquiries: The Framework. In: R. Martin, D. Michaels & J. Uriagereka, eds., *Step by Step*. MIT Press, Cambridge, Mass., pp. 89–155.

Chomsky, Noam (2001): Derivation by Phase. In: M. Kenstowicz, ed., *Ken Hale. A Life in Language*. MIT Press, Cambridge, Mass., pp. 1–52.

Chomsky, Noam (2005a): On Phases. Ms., MIT, Cambridge, Mass.

Chomsky, Noam (2005b): Three Factors in Language Design, *Linguistic Inquiry* 36, 1–22.

Chomsky, Noam (2007): Approaching UG from Below. In: U. Sauerland & H.-M. Gärtner, eds., *Interfaces + Recursion = Language?*. Mouton de Gruyter, Berlin, pp. 1–31.

Chomsky, Noam (2008): On Phases. In: R. Freidin, C. Otero & M. L. Zubizarreta, eds., *Foundational Issues in Linguistic Theory*. MIT Press, Cambridge, Mass., pp. 133–166.

Chomsky, Noam (2013): Problems of Projection, *Lingua* 130, 33–49.

Chomsky, Noam & Howard Lasnik (1993): Principles and Parameters Theory. In: J. Jacobs, A. von Stechow, W. Sternefeld & T. Vennemann, eds., *Syntax. Ein internationales Handbuch zeitgenössischer Forschung*. Vol. 1, de Gruyter, Berlin, pp. 506–569.

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

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

Collins, Chris & Joachim Sabel (2007): An LF Interface Constraint on Remnant Movement. Ms., NYU and Université Catholique de Louvain.

Coon, Jessica, Pedro Mateo Pedro & Omer Preminger (2011): The Role of Case in A-Bar Extraction Asymmetries: Evidence from Mayan. Ms., McGill University.

Dalrymple, Mary (2001): *Lexical Functional Grammar*. Vol. 34 of *Syntax and Semantics*, Academic Press, San Diego.

De Kuthy, Kordula & Walt Detmar Meurers (2001): On Partial Constituent Fronting in German, *Journal of Comparative Germanic Linguistics* 3(3), 143–205.

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

den Besten, Hans & Gert Webelhuth (1990): Stranding. In: G. Grewendorf & W. Sternefeld, eds., *Scrambling and Barriers*. Benjamins, Amsterdam, pp. 77–92.

Epstein, Samuel David & T. Daniel Seely (2002): Rule Applications as Cycles in a Level-Free Syntax. In: S. D. Epstein & T. D. Seely, eds., *Derivation and Explanation in the Minimalist Program*. Blackwell, Oxford, pp. 65–89.

Fanselow, Gisbert (1987): *Konfigurationalität*. Narr, Tübingen.

Fanselow, Gisbert (1991): Minimale Syntax. Habilitation thesis, Universität Passau.

Fanselow, Gisbert (2002): Against Remnant VP Movement. In: *Dimensions of Movement*. Benjamins, Amsterdam, pp. 91–125.

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

Fox, Danny (1995): Untitled Ms.. Unpublished, MIT.

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

Fox, Danny & David Pesetsky (2005): Cyclic Linearization of Syntactic Structure, *Theoretical Linguistics* 31, 1–45.

Frank, Robert, Young-Suk Lee & Owen Rambow (1992): Scrambling as Non-Operator Movement and the Special Status of Subjects. In: S. Barbiers, M. den Dikken & C. Levelt, eds., *Proceedings of the Third Leiden Conference for Junior Linguists*. University of Leiden, Leiden, pp. 135–154.

Frey, Werner (2000): Über die syntaktische Position der Satztopiks im Deutschen. In: *ZAS Papers in Linguistics*. Vol. 20, pp. 137–172.

Frey, Werner (2004): A Medial Topic Position for German, *Linguistische Berichte* 198, 153–190.

Gazdar, Gerald (1981): Unbounded Dependencies and Coordinate Structure, *Linguistic Inquiry* 12, 155–184.

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

Georgi, Doreen (2013): Opaque Interaction of Internal Merge and Agree. In: *Rule Interaction in Grammar*. Vol. 90 of *Linguistische Arbeits Berichte*, Institut für Linguistik, Universität Leipzig, pp. 413–461.

Georgi, Doreen (2014): Opaque Interactions of Merge and Agree. PhD thesis, Universität Leipzig.

Grewendorf, Günther (2003): Improper Remnant Movement, *Gengo Kenkyo: The Journal of the Linguistic Society of Japan* 123, 47–94.

Grewendorf, Günther (2004): Die linke Bewegung von Restkategorien. Ms., Universität Frankfurt/Main.

Grewendorf, Günther & Joachim Sabel (1994): Long Scrambling and Incorporation, *Linguistic Inquiry* 25, 263–308.

Grimshaw, Jane (1998): Constraints on Constraints in Optimality Theoretic Syntax. Ms., Rutgers University, New Brunswick, New Jersey.

Grohmann, Kleanthes K. (2003): *Prolific Domains: On the Anti-Locality of Movement Dependencies*. John Benjamins, Amsterdam & Philadelphia.

Haider, Hubert (1993): *Deutsche Syntax – generativ*. Narr, Tübingen.

Hale, John & Géraldine Legendre (2004): Phrase Impenetrability and Wh-Intervention. In: A. Stepanov, G. Fanselow & R. Vogel, eds., *Minimal Links, Remnant Movement, and (Non-) Derivational Grammar*. Mouton de Gruyter, Berlin, pp. 177–203.

Heck, Fabian & Anke Assmann (2012): Barss' Generalization and the Strict Cycle at LF. Ms., Universität Leipzig.

Heck, Fabian & Gereon Müller (2000): Repair-Driven Movement and the Local Optimization of Derivations. Ms., Universität Stuttgart and IDS Mannheim.

Heck, Fabian & Gereon Müller (2003): Derivational Optimization of Wh-Movement, *Linguistic Analysis* 33, 97–148. (Volume appeared 2007).

Heck, Fabian & Gereon Müller (2010): Lokale Modellierung nicht-lokaler Abhängigkeiten in der Syntax. Ms., Universität Leipzig.

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

Hornstein, Norbert (2001): *Move. A Minimalist Theory of Construal*. Blackwell, Oxford.

Huang, Cheng-Teh James (1982): Logical Relations in Chinese and the Theory of Grammar. PhD thesis, MIT, Cambridge, Mass.

Huybregts, Riny (2009): The Minimalist Program: Not a Bad Idea. Ms., Universiteit Utrecht.

Kayne, Richard (1982): *Connectedness and Binary Branching*. Foris, Dordrecht.

Kiparsky, Paul (1973): Abstractness, Opacity and Global Rules. In: O. Fujimura, ed., *Three Dimensions in Linguistic Theory*. TEC, Tokyo, pp. 57–86.

Kitahara, Hisatsugu (1994): Restricting Ambiguous Rule-Application. In: M. Koizumi & H. Ura, eds., *Formal Approaches to Japanese Linguistics*. Vol. 1, MITWPL, Cambridge, Mass., pp. 179–209.

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

Koizumi, Masatoshi (1995): Phrase Structure in Minimalist Syntax. PhD thesis, MIT, Cambridge, Mass.

Koopman, Hilda & Anna Szabolcsi (2000): *Verbal Complexes*. MIT Press, Cambridge, Mass.

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

Lahne, Antje (2009): Where There is Fire There is Smoke. Local Modelling of Successive-Cyclic Movement. PhD thesis, Universität Leipzig.

Lahne, Antje (2012): Specificity-driven Syntactic Derivation. In: M. Uribe-Etxebarria & V. Valmala, eds., *Ways of Structure Building*. Oxford University Press, Oxford, pp. 271–296.

Lasnik, Howard & Mamoru Saito (1984): On the Nature of Proper Government, *Linguistic Inquiry* 15, 235–289.

Lasnik, Howard & Mamoru Saito (1992): *Move α*. MIT Press, Cambridge, Mass.

Lechner, Winfried (1998): Two Kinds of Reconstruction, *Studia Linguistica* 52, 276–310.

Longobardi, Giuseppe (1985): Connectedness and Island Constraints. In: J. Guéron, H.-G. Obenauer & J.-Y. Pollock, eds., *Grammatical Representation*. Foris, Dordrecht, pp. 169–185.

McCawley, James (1984): Exploitation of the Cyclic Principle as a Research Strategy in Syntax. In: W. de Geest & Y. Putseys, eds., *Sentential Complementation*. Foris, Dordrecht, pp. 165–183.

McCawley, James (1998): *The Syntactic Phenomena of English*. University of Chicago Press, Chicago.

McCloskey, James (1988): Syntactic Theory. In: F. Newmeyer, ed., *Linguistics: The Cambridge Survey. Vol. I. Linguistic Theory: Foundations*. Cambridge University Press, Cambridge, pp. 18–59.

McCloskey, James (2002): Resumptives, Successive Cyclicity, and the Locality of Operations. In: S. D. Epstein & T. D. Seely, eds., *Derivation and Explanation in the Minimalist Program*. Blackwell, Oxford, pp. 184–226.

Müller, Gereon (1993): On Deriving Movement Type Asymmetries. PhD thesis, Universität Tübingen.

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

Müller, Gereon (2009): Ergativity, Accusativity, and the Order of Merge and Agree. In: K. K. Grohmann, ed., *Explorations of Phase Theory. Features and Arguments*. Mouton de Gruyter, Berlin, pp. 269–308.

Müller, Gereon (2011): *Constraints on Displacement. A Phase-Based Approach*. Vol. 7 of *Language Faculty and Beyond*, Benjamins, Amsterdam.

Müller, Gereon (2013): On the Order of Syntactic Operations. Ms., Universität Leipzig. (Lecture Notes for a Compact Course, University of Cambridge).

Müller, Gereon (2014): A Local Approach to the Williams Cycle, *Lingua* 140, 117–136.

Müller, Stefan (2007): *Head-Driven Phrase Structure Grammar: Eine Einführung*. Stauffenburg, Tübingen.

Murasugi, Kumiko (1992): Crossing and Nested Paths. PhD thesis, MIT, Cambridge, Mass.

Neeleman, Ad & Hans van de Koot (2010): A Local Encoding of Syntactic Dependencies and its Consequences for the Theory of Movement, *Syntax* 13, 331–372.

Perlmutter, David & Scott Soames (1979): *Syntactic Argumentation and the Structure of English*. The University of California Press, Berkeley.

Pesetsky, David (1985): Morphology and Logical Form, *Linguistic Inquiry* 16, 193–246.

Pesetsky, David (2012): Phrasal Movement and its Discontents: Diseases and Diagnoses. Ms., MIT, Cambridge, Mass. (To appear in Lisa Cheng and Norbert Corver (eds.), *Diagnostics in Syntax*. Oxford University Press.).

Pollard, Carl & Ivan Sag (1992): Anaphors in English and the Scope of Binding Theory, *Linguistic Inquiry* 23, 261–303.

Pollard, Carl J. & Ivan A. Sag (1994): *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.

Pullum, Geoffrey (1979): *Rule Interaction and the Organization of a Grammar*. Garland, New York.

Pullum, Geoffrey (1992): The Origins of the Cyclic Principle. In: *Proceedings of the Regional Meeting of the Chicago Linguistic Society*. Vol. 28.2, pp. 209–236.

Pullum, Geoffrey (1997): The Morpholexical Nature of English to-Contraction, *Language* 73, 79–102.

Reinhart, Tanya & Eric Reuland (1993): Reflexivity, *Linguistic Inquiry* 24, 657–720.

Richards, Marc (2010): Deletability of Edge Features. Ms., Universität Leipzig.

Richards, Marc (2011): Deriving the Edge: What's in a Phase?, *Syntax* 14, 74–96.

Richards, Norvin (2001): *Movement in Language*. Oxford University Press, Oxford.

Riemsdijk, Henk van & Edwin Williams (1981): NP-Structure, *The Linguistic Review* 1, 171–217.

Ross, John (1967): Constraints on Variables in Syntax. PhD thesis, MIT, Cambridge, Mass.

Sag, Ivan & Thomas Wasow (1999): *Syntactic Theory. A Formal Introduction*. CSLI Publications, Stanford University.

Saito, Mamoru (1992): Long Distance Scrambling in Japanese, *Journal of East Asian Linguistics* 1, 69–118.

Saito, Mamoru (2003): Some Asymmetries in Japanese and Their Theoretical Implications, *Lingua* 113, 481–518.

Sanders, Gerald (1974): Precedence Relations in Language, *Foundations of Language* 11, 361–400.

Sauerland, Uli (1999): Erasability and Interpretation, *Syntax* 3, 161–188.

Sauerland, Uli & Paul Elbourne (2002): Total Reconstruction, PF Movement, and Derivational Order, *Linguistic Inquiry* 33, 283–319.

Stabler, Edward (1999): Remnant Movement and Complexity. In: G. Bouma, G.-J. Kruijff, E. Hinrichs & R. Oehrle, eds., *Constraints and Resources in Natural Language Syntax and Semantics*. CSLI Publications, Stanford University, pp. 299–326.

Stechow, Arnim von (1992): Kompositionsprinzipien und grammatische Struktur. In: P. Suchsland, ed., *Biologische und soziale Grundlagen der Sprache*. Niemeyer, Tübingen, pp. 175–248.

Stiebels, Barbara (2006): Agent Focus in Mayan Languages, *Natural Language and Linguistic Theory* 24, 501–570.

Stroik, Thomas (2009): *Locality in Minimalist Syntax*. MIT Press, Cambridge, Mass.

Takano, Yuji (1994): Unbound Traces and Indeterminacy of Derivation. In: M. Nakamura, ed., *Current Topics in English and Japanese*. Hituzi Syobo, Tokyo, pp. 229–253.

Thiersch, Craig (1985): VP and Scrambling in the German Mittelfeld. Ms., University of Tilburg.

Thoms, Gary & George Walkden (2013): VP preposing and the PBC. Ms., University of Manchester.

Unger, Christina (2010): A Computational Approach to the Syntax of Displacement and the Semantics of Scope. PhD thesis, Universiteit Utrecht, LOT.

van Koppen, Marjo (2005): One Probe – Two Goals. Aspects of Agreement in Dutch Dialects. PhD thesis, Leiden University.

Williams, Edwin (1974): Rule Ordering in Syntax. PhD thesis, MIT, Cambridge, Mass.

Williams, Edwin (2003): *Representation Theory*. MIT Press, Cambridge, Mass.