Towards a Restrictive Theory of (Remnant) Movement: Improper Movement, Remnant Movement, and a Linear Asymmetry*

Klaus Abels
Universitetet i Tromsø

October 22, 2006

Contents

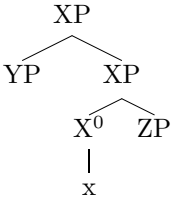
*I would like to thank the participants of the Left-Right Seminar (Tromsø, Spring 2006), the Ph.D. course on remnant movement and left-right asymmetries (Leipzig, Summer 2006), the participants of the remnant movement reading group (Tromsø, Fall 2006), and Andrew Nevins for invaluable comments, discussion, and challenges to the proposals made here and their predecessors. Among the aforementioned I am particularly indebted to Øystein Nilsen and Gereon Müller.

1 Introduction: A plea for a restrictive theory of operations

Syntactic theory should explain linear asymmetries in language. Why is leftward movement so much less restricted than rightward movement (Perlmutter and Postal (1983/1972), Bresnan (1970))? Why aren't OV-languages simply the mirror image of VO-langues (see Kayne (1994) for some illustrations)? Why are there verb-second languages but no verb-second-last languages? Why are noun phrases asymmetric in a typological perspective (Greenberg (1963), Cinque (1996, 2005); Abels and Neeleman (2006))? Questions like these are in the background throughout this paper.

Over 10 years ago now, Kayne (1994) suggested that part of the answer to such questions could be the Linear Correspondence Axiom (LCA), that has as one of its consequences that the unique head of any phrase must precede the unique complement of that head and that the unique specifier/adjunct in a phrase must precede both the head and the complement. According to the LCA, all phrases in natural language follow the template in (1). The LCA also entails that all movement operations are leftward.[1]

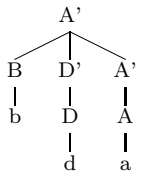
(1)



Obviously, the LCA, if true, can be no more than a partial answer to the questions raised in the first paragraph. No matter how restricted the theory of phrase structure is and no matter how serious the restriction to leftward movement may seem, an unduly permissive theory of movement can still undermine the explanatory effort expressed in the LCA. All of the structures that cannot be directly generated under the LCA can be brought back through movement operations, (2). The symbol $\rightsquigarrow$ can be read as follows: Any LCA-incompatible analysis containing the structure on the left can be made LCA-compatible by replacing all occurrences of the structures on the left hand side of the arrow by the structures on the right hand side. Head-final structures can

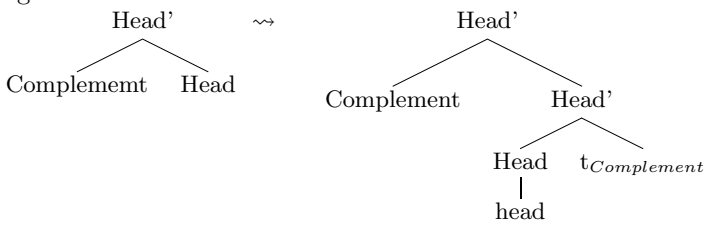
[1] This is not the place to discuss the mechanics of the LCA and the various interpretations it has been given, for which see Kayne (1994); Chomsky (1995a); Uriagereka (1999); Koopman (1996); Wilder (1998) among others. It might be curious to note, though, that Kayne's original formulation of the LCA does allow ternary branching structures despite claims in Kayne (1994) to the contrary. Ternary branching structures like (i) are compatible with Kayne's formulations, as the reader may easily verify.

(i)

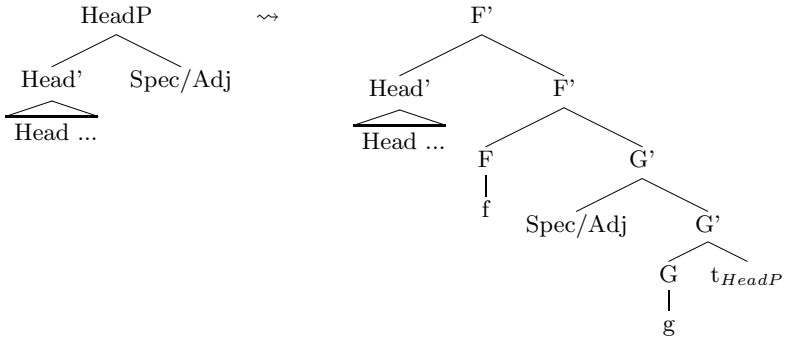


be emulated through short movement operations, (2a), right specifiers and adjuncts can likewise be brought into the fold through the introduction of silent functional elements, f and g in (2b). Finally, any rightward movement operation can be emulated by a two-movement sequence, where the second step involves remnant movement, as in (2c). The structural and movement types indicated in (2) are not just hypothetical figments of my imagination; Kayne (1994, p. 50-54) chooses (2a) to analyze right headedness–although the analysis was later revised in Kayne (1999, 2004); Cinque (2005) uses structures like (2b) to treat right adjuncts, and Kayne (1998) employs structures abstractly like (2c) to treat certain cases of quantifier scope.

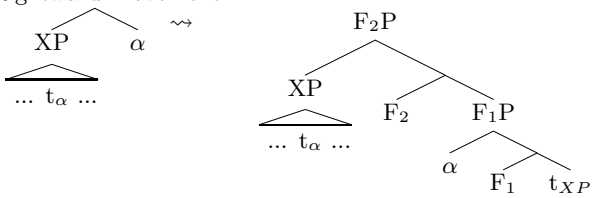
(2) a. Right Heads:



b. Right Sepcifiers and Adjuncts:



c. Rightward Movement:



Given this, the LCA is more restrictive and predictive than the older, symmetric theory of structure and movement only in conjunction with a restrictive theory of movement.[2]

Precisely such a restrictive theory of movement was in the background when Kayne's book

[2]In addition to a restrictive theory of the architecture of phrases we also need to assume tight constraints on possible base-generated hierarchies of multiple phrases. I assume that some version of this hypothesis, variously called the *Hierarchy of Projections* Adger (2003), the Cinque hierarchy, the functional sequence (fseq in Starke (2004)), or Pollock-Cinque functional hierarchy (PCFH in Williams (2002)), is correct although current formulations face a number of problems (see Bobaljik (1999) and Nilsen (2004) for clear statements of the two most pressing issues). I will usually refer to this assumed fixed hierarchy of phrasal projections as the Cinque hierarchy.

was first published, hence, the LCA was viewed as a great explanatory success. In particular, the proper binding condition was widely assumed. But this has changed over the years.

I concentrate on remnant movement in this paper. Some of the other issues, those raised by the movements in (2a) and (2b), are discussed in some detail in Abels and Neeleman (2006). The conclusion drawn there is that the massive use of roll-up movements in the current literature removes the empirical teeth of the claim that complements follow heads and that specifiers precede heads (Richards (2004, chapter 2) comes to a similar assessment). Moreover, Abels and Neeleman argue that such roll-up movements violate otherwise valid generalizations concerning pied-piping, stranding, and anti-locality. They conclude that generating complements to the left or to the right of a head and allowing left and right specifiers and adjuncts allows a more restrictive theory of movement that can dispense with entire classes of movement. This helps in the ultimate explanation of linear asymmetries. I assume this reasoning here and make use of trees with left and right complements and left and right specifiers. Nothing in this paper, except for the discussion of head movement in section 5.1, hinges on this choice and adherents of the LCA may mechanically transform all non-LCA-compatible structures in this paper into fully LCA-compatible ones by assuming the required functional and agreement heads as indicated in (2a) and (2b). In other words, this paper does not bear on the question of base structures. It is about the shape of a restrictive theory of movement.

What remains of the LCA under the view sketched above is the ban against rightward movement—for an attempt at deriving this ban independently of the LCA see Abels and Neeleman (2006). The ban against rightward movement might still have empirical teeth and explanatory force. It all depends on the theory of movement it is embedded in.

In the early and mid nineties the Proper Binding Condition (PBC) was still widely assumed, remnant movement was considered exotic, and Müller (1998) felt compelled to spend an entire chapter on reconciling PBC-effects with the existence of remnant movement. The PBC – in the formulation argued for in (Lasnik and Saito, 1992, p. 90) – says: "Traces must be bound throughout a derivation" (see Fiengo (1977, p. 45) for the original formulation). Under Lasnik and Saito's formulation, the PBC rules out any derivation along the lines in (2c), because the trace of alpha within the fronted XP is not bound at the stage of the derivation depicted. With the PBC as background the ban against rightward movement predicts the complete absence of what is observationally rightward movement. In that sense the ban against rightward movement (and thus the LCA) together with the PBC predict a linear asymmetry.

Two things have become clear since the publication of Kayne (1994) though. First, the PBC cannot be entirely correct; the arguments for the existence of remnant movement are just too strong (for an overview of the arguments from German VP-, PP-, and NP-topicalization see (Müller, 1998, chapter 1), see also Hiraiwa (2002) Abels (2002)). Second, the existence of what appears to be rightward movement cannot be completely denied; there is, for example, a consensus now that the structure for Heavy NP Shift under Kayne's 1994 analysis is incorrect. Under that analysis the "heavy-shifted NP" simply remained in its low VP internal position. Rather, it seems that the "heavy shifted NP" does move out of VP. There are two implementations for this: the traditional rightward movement analysis and a remnant movement analysis. Both of them give rise to gross constituent structures along the lines of those generated by the traditional rightward movement analysis (den Dikken (1995); Rochemont and Culicover (1997); Kayne (1998)

among others). Of course, the long-standing observation still stands that what observationally looks like rightward movement has very different properties of leftward movement.

These remarks show that, if the explanatory project behind the LCA, to account for linear asymmetries in languages, is to succeed, a restrictive theory of operations, movement in particular, is required.

If remnant movement is not constrained by the proper binding condition, what is it constrained by? To investigate this question, I look at improper movement and related phenomena that bear on the sequencing of movement operations. My suggestion is to make explicit a tacit assumption in much current theorizing, namely, that operations must apply in a strictly ordered sequence. I hypothesize that there is a Universal Constraint on Operational Ordering in Language, UCOOL for short, operative in all languages. As a first and very coarse approximation this constraint takes the form in in (3).

(3) $\theta << \text{Case} << Op$

I go on to suggest a way that grammar makes use of this hierarchy below in (9). (9) is a Generalized Prohibition against Improper Movement (GenPIM) and it constrains regular as well as remnant movement. This way of deploying UCOOL derives a number of feeding and bleeding relationships between movement operations, not last in the arena of remnant movement. The insights gained this way are then brought to bear on the analysis of cross-serial dependencies. It emerges that remnant movement derivations of unbounded cross-serial dependencies (see Nilsen (2003)) necessarily violate the Generalized Prohibition against Improper Movement. The same is true of a number of analyses in Koopman and Szabolcsi (2000), of which their analysis of order preservation is discussed below as an example.

The paper is structured as follows. First I discuss improper movement in various guises. Then I turn to cross-serial dependencies in Swiss German and Dutch and claim that there is a hitherto unnoticed linear asymmetry the domain of cross-serial dependencies, section 3. I then discuss two types of derivation for cross-serial dependencies: one with and one without remnant movement and show that the remnant movement derivation (a) fails to capture the asymmetry and (b) necessarily violates the Generalized Prohibition against Improper Movement. In sections 5, I show that the asymmetry falls out from the theory developed up to that point. The conclusion discusses open issues, methodological implications, and prospects for further inquiry.

2 UCOOL and GenPIM

In this section I argue that there is a, presumably universal, constraint that dictates the order in which operations have to apply to a given constituent (UCOOL), (4). It should come as no surprise to anybody that UCOOL demands that θ related operations have to apply first, followed by Case related operations, followed by operator movement. This is, of course, the received wisdom. UCOOL simply makes this consensual point explicit. I will now investigate some properties of the deployment of UCOOL.[3]

[3]I hasten to note that the hierarchy showcased in (4) is presumably far too coarse. Just like phrase-structure hierarchies have expanded from the simple $[_{CP}[_{IP}[_{VP}...]]]$ to encompass ever more fine-grained structures (Larson

(4) The Universal Constraint on Operational Ordering in Language (UCOOL) – *to be refined*
θ << Case << *Op*
I will say that θ-operations are a lower type of operations than Case, and that Case-operations are a higher type of operations than θ-operations.

We want to deploy UCOOL to make sure that elements that need θ-roles get them first, before they receive case or undergo operator movement. And to make sure that once an element has undergone operator movement, it cannot receive Case or pick up a θ-role. Many questions arise immediately. A first set of questions concerns the hierarchy itself: Is it a linear order? Or is it a partial order? What counts as a type of operation? How many such types are there? I will assume that UCOOL is a linear order. For the purposes of this paper, I will remain somewhat vague concerning the exact boundaries of my notion of operation. However, any movement will count as an operation. Finally, I assume that as a first approximation the feature-types that enter into the computation of relativized minimality/attract closest define the operational types.

A second set of questions is of a far reaching but somewhat technical nature: How exactly is this hierarchy deployed? Should it regulate the application of operations per derivation? Per cycle? Per phase? Or should it regulate the application of operations per feature? Per syntactic terminal? Per lexical item? Per word? Per constituent? I will attempt to provide answers to these questions in the remainder of this section.

A third set of question is less directly technical but even more far reaching: What is the relation between UCOOL and the Cinque hierarchy? Is one parasitic upon the other? If so, which one derives from the other? Can a (partial) shift of the explanation away from the Cinque hierarchy and onto UCOOL solve some of paradoxes and anomalies that arise under the current understanding of the Cinque hierarchy? While it is important to keep the third set of questions in mind as the exploration of UCOOL unfolds, trying to answer them strikes me as premature at the moment. We first need to establish a much firmer understanding of UCOOL itself.

2.1 Standard evidence for UCOOL: Improper Movement

Simple considerations suggest that movement from a non-θ- into a θ-position is barred, (5). Whether movement from θ-position to θ-position is possible is not a question I wish to address here (see Bošković (1994); Hornstein (1999); Landau (2003); Culicover and Jackendoff (2001); Boeckx and Hornstein (2003); Polinsky and Potsdam (2002) for discussion), but it is certainly true that for the authors who entertain movement into θ-positions this movement does not pass through Case or operator positions.

(5) a. *I asked who [$_{CP}$ t$_{who}$ [John should meet Mary]]
b. *I expected t$_I$ to be discovered a proof.

It appears as though (obligatory) control relations in general are immutable after Case has been assigned, (6) ((6a-b) are from (Williams, 1974, p. 4)). In (6b) it is controlled by *John*. In examples (6a) the subject of *tell* is controlled by *the police*. The same is true in (6c), despite the

(1988); Pollock (1989); Alexiadou (1997); Rizzi (1997); Cinque (1999, 2002); Belletti (2004); Rizzi (2004b); Cinque (2006) among many others), so our inventory of operations presumably needs to be refined (Sternefeld (1992); Müller and Sternefeld (1993); Starke (2001); Williams (2002); Grewendorf (2003)).

fact that *who* has moved. The control relation into the adjunct apparently requires c-command and cannot be established by an argument that has already received Case. Hence (6a) and (6c) behave the same way as far as control into the adjunct is concerned.

(6) a. The police arrested John six times without ever telling anyone about it.
b. John was arrested six times by the police without ever telling anyone about it.
c. Who did the police arrest six times without ever telling anyone about it.

The examples in (7) control for distance, a potential interfering factor in (6) (thanks to M. Starke, p.c., for pointing this out). It might be assumed, following the idea of the minimal distance principle (Rosenbaum (1967, 1970)) that *who* cannot act as a controller in (7c) because of intervention caused by the closer potential controller *the police*. In (7) *without asking for a permit* can attach either in the higher or in the lower clause. (7a) allows only low-clause attachment, because there is no potential controller in the matrix. However, (7b) ought to be ambiguous between (7b-i), which represents the reading the sentence has, and (7b-ii). (7b-ii) has a perfectly coherent interpretation: Who is the person such that it is illegal to visit that person without him/her asking for a permit? It should be available if *wh*-movement could feed control. Apparently it cannot. This suggests grouping (obligatory) control together with θ-role assignment very high on the Universal Constraint on Operational Ordering in Language.

(7) a. It is illegal to visit your grandmother without asking for a permit.
b. Who is it illegal to visit without asking for a permit?
(i) [Who is it illegal [PRO_i to visit t_{who} [without PRO_i asking for a permit]]]
(ii) *[Who_k is it illegal [PRO_i to visit t_{who}] [without PRO_k asking for a permit]]

A standard example of improper movement might suffice to illustrate that movement from $\overline{A}$-position to A-position (and back to $\overline{A}$-position in (8b)) is illict, (8).

(8) a. ✓[$_{\text{CP}}$ What [$_{\text{IP}}$ t^3_{what} seems [$_{\text{IP}}$ t^2_{what} to have been said t^1_{what} ?]]]]
b. *[$_{\text{CP}}$ What [$_{\text{IP}}$ t^3_{what} seems [$_{\text{CP}}$ t^2_{what} that [$_{\text{IP}}$ it was said t^1_{what} ?]]]]
c. John asked whose book to read.
d. (i) *Whose book was asked [$\text{t}_{whosebook}$ [{$\text{t}_{whosebook}$ | PRO | there} to be read $\text{t}_{whosebook}$]].
(ii) *Whose book was asked [$\text{t}_{whosebook}$ {whether | that | $\emptyset$} [{ it | there} was read $\text{t}_{whosebook}$]].

Let this act as a reminder of what type of facts the Universal Constraint on Operational Ordering in Language, (4), is accountable for. But, again, how exactly do we deploy UCOOL?

2.2 How UCOOL is deployed: GenPIM

Under a theory that generates syntactic structures bottom up, we can safely discard the option of deploying UCOOL at the level of the entire derivation, that is, it doesn't make much sense empirically to demand that all θ-role assignment happens before all A-movement before all $\overline{A}$-movement; θ-role assignment, A-movement, and $\overline{A}$-movement will happen in embedded clauses

before they can happen in higher clauses.[4] This suggests that UCOOL must apply to smaller domains than entire derivations. An obvious candidate would be the derivational cycle (the phase). However, construing of UCOOL as a constraint on operations per cycle gives rise exactly to the problems we are trying to solve. Consider example (8b). On the lower cycle, in the domain of the embedded clause, UCOOL is not violated. Though *what* fails to get Case, it receives a θ-role first and later undergoes $\overline{A}$-movement in total conformity with UCOOL. On the upper cycle, *what* does not receive a θ-role but it is affected by movement to an A-position followed by movement to an $\overline{A}$-position in conformity with UCOOL. Cyclicity in this sense, then, cannot be the answer.[5]

Looking at the problem from the other end of the size scale, relativizing the Universal Constraint on Operational Ordering in Language to features will certainly not do. A single feature under standard assumptions is responsible for a single type of operation. There is no point in imposing an ordering on a single operation, it will vacuously comply with the order, i.e., (5) and (8d) cannot be ruled out like this.[6] For ordering statements like the Universal Constraint on Operational Ordering in Language to have any bite, they have to apply at least to bundles of features. A moment's reflection reveals that head-sized, morpheme-sized, or word-sized bundles of features are not large enough to deal with improper movement. Consider again (8d). Everything goes fine for *whose*. It may or may not receive a θ role from *book*, it receives genitive Case internally to the DP *whose book* and then triggers $\overline{A}$-movement. UCOOL is satisfied. Likewise, *book* receives a θ-role and later Case in the matrix [Spec, TP]. Again, UCOOL is satisfied. So what went wrong?

What went wrong is that the entire phrase (qua pied-piping) was involved in high type $\overline{A}$-operation, which by UCOOL should ban the entire phrase with all its parts from entering into a lower type of operation later. I thus suggest to deploy UCOOL as stated in the Generalized Prohibition against Improper Movement, GenPIM, (9). GenPIM, together with the definition of affectedness in (10) generalizes the ban against improper movement from the trigger of movement to the moving constituent itself and beyond it to all constituents affected by the movement.

(9) Generalized Prohibition against Improper Movement (GenPIM)
No constituent may undergo movement of type τ if it has been affected by movement of type σ, where $\tau << \sigma$ under UCOOL.

(10) A constituent α is affected by a movement operation iff
   i. α is reflexively contained in the constituent created by movement, and
   ii. α is in a (reflexive) domination relation with the moved constituent.

What (10) says is that when a constituent X is moved, this movement affects not only the moving

[4] Williams' (2002) Representation Theory – and its partial translation into a derivational theory sketched in (Nevins, 2004, p. 298-301), also a few remarks in Nevins and Hornstein (2005) – can treat UCOOL as a constraint on entire derivations. Representation Theory is therefore incompatible with root-extending merge, which I would like to hold on to.

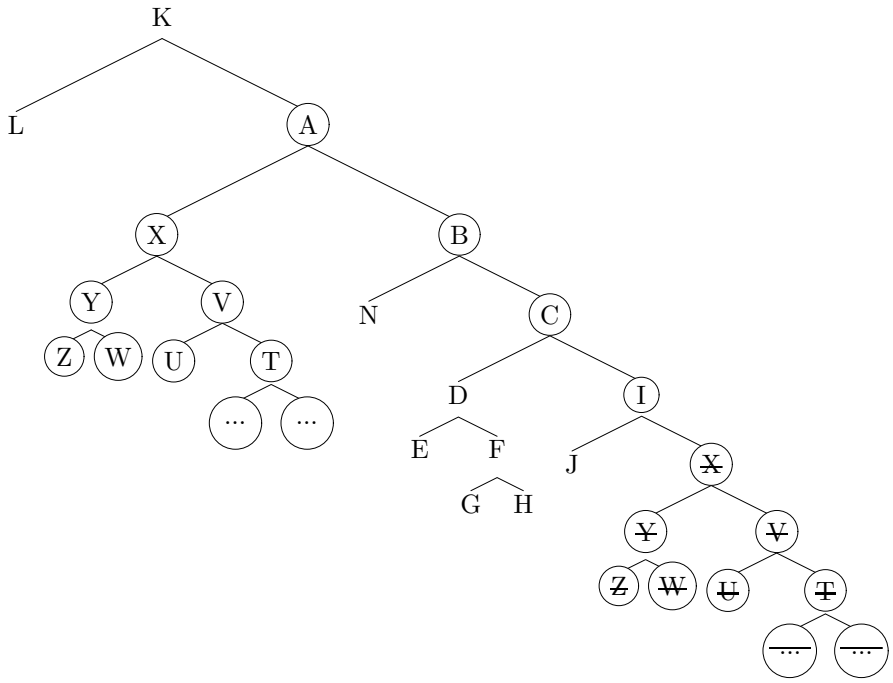
[5] Chomsky's (2001b; 2001a) notion of the cycle, where information about what happened in the lower phase remains accessible during the higher phase until the next higher strong phase head is introduced, overcomes these problems only partly. The loophole created by cyclic phase-edge-to-phase-edge movement is just too large.

[6] I do not mention (8b) here, because it might be argued to violate some independent condition, like Collins' (1994) prohibition against chain-interleaving.

constituent itself (it is in a reflexive domination relation with itself) but also the constituents that make up X (the elements that X dominates), and the nodes along the path of movement (understood in terms of domination) since those dominate X in the pre-movement configuration. Every node which is not affected is unaffected. Therefore, specifiers and heads along the path of movement are unaffected and so are all constituents "higher up" in the tree.

This is illustrated in (11), where the nodes affected by movement of X are circled and the unaffected nodes are not.

(11)



GenPIM generalizes the ban against improper movement in the following sense. The standard understanding of the ban against improper movement imposes the condition in (12), which differs from GenPIM only in the boldfaced part. This standard understanding makes sure that movements undergone by a particular constituent do not violate UCOOL. GenPIM generalizes this. Not only the moving constituent is subject to the operational ordering but also its component parts and the constituents along the path of movement.

(12) No constituent may undergo movement of type σ if it has undergone movement of type τ, where $\sigma << \tau$ under UCOOL.

Although GenPIM is not a derivational constraint as given, a derivational metaphor might make it more tangible. A constituent, X in (12), has a number of things it needs to do in its derivational lifespan. These are ordered in terms of urgency by UCOOL. X may not undergo a less urgent (higher) operation unless all the more urgent (lower) ones have been completed. Similarly, and this is the contribution of GenPIM over (13), X may not undergo a less urgent (higher) operation

if any of its parts still have more urgent business to do. Conversely, if one of the constituents that contain X in the pre-movement structure still has urgent business to do, then X will have to wait with its less urgent business.

The generalization from (12) to GenPIM will be justified in the remainder of this section. Before we look at the empirical justification for it, a few comments and clarifications are in order.

First, the way it is presently formulated, GenPIM has to be understood as being about final landing sites of movement. Despite the derivational metaphor above, it cannot be understood directly as a constraint on derivations (unless the standard assumption is given up that successive cyclic movement is launched before the target of movement is merged into the tree). I make no attempt to reformulate GenPIM in derivational terms here. The primary goal of the present paper is to investigate the correctness of the proposal and to demonstrate a specific prediction of UCOOL together with GenPIM.

Second, there is an asymmetry in GenPIM which is absent in (12). (12) makes reference to moving constituents both in the protasis and in the apodosis. GenPIM on the other hand makes reference to moving constituents in the apodosis but to affected constituents in the protasis. I leave this asymmetry unmitigated here. Whether the stronger symmetrical formulation where both protasis and apodosis make reference to affected constituents is tenable will have to be investigated independently.

Third, I will give only a very brief and cursory discussion of the relation between the present theory and other theories of improper movement. Of particular interest is Williams' (2002) Representation Theory, which derives a condition that is in many ways even stronger than the symmetrical version of GenPIM in terms of affected constituents. Williams' proposals were the inspiration for many of the ideas in this paper.

Clearly, GenPIM makes predictions far beyond the realm usually considered to fall under improper movement. Consider the following simple table, which represent a presumed subpart of the Universal Constraint on Operational Ordering in Language. I have left out θ-operations here. By A-movement I want to understand movement to a structural Case position for the purposes of discussion. I also added the operation of scrambling (by which I specifically refer to movement into a pre-subject position in the German mitteldfeld) and split the entry for $\overline{A}$-movement in (4) into *wh*-movement and topicalization. The left-to-right and top-to-bottom ordering of operations corresponds to a (part) of the hierarchy of movement operations one would postulate for German (see Grewendorf (2003) for a similar proposal regarding the hierarchy of operational types).[7]

(13)

	A-mvt	Scrambling	*wh*-mvt	Topicalization
A-mvt		✓	✓	✓
Scrambling	*		✓	✓
wh-mvt	*	*		✓
Topicalization	*	*	*	

Suppose we interpret the table as a table of the combinability of operations applying to one and the same constituent. The rows represent the first operation, the columns – the second. Thus the first row says that a constituent that has undergone A-movement may subsequently

[7] A structural formulation of this hierarchy is presented in Sternefeld (1992).

undergo scrambling, *wh*-movement, or Topicalization. The second row says that a constituent that has undergone scrambling may not subsequently undergo A-movement, but may undergo subsequent *wh*-movement or topicalization. The third row says that *wh*-moved constituents may – according to the hierarchy – undergo topicalization bot not A-movement or scrambling. Finally topicalization does not feed any other operation.

The diagonal of the table is left open, mostly because it is unclear to me whether the individual steps of successive cyclic movement should count, that is, whether a step in successive cyclic A-movement should count as A-movement, whether a step in successive cyclic *wh*-movement should itself count as *wh*-movement, and so on. If so, then the cells along the diagonal of the table should presumably get a checkmark. If, on the other hand, we interpret the fact that DPs do not move from structural case to structural case position and that *wh*-phrases do not leave their scope positions for further *wh*-movement once they have reached them as evidence that movement of the same type may not iterate (see Rizzi's (2004a) Criterial Freezing), then we predict the diagonal of the table to be populated by asterisks. Since the issue is not directly relevant to the purposes of this paper, I leave it unresolved.

The predictions encoded in this table are by and large borne out. In particular there is no reason to believe that the operational sequences marked with an asterisks ever occur. It is also clear that A-movement, that is Case marking, may precede scrambling, *wh*-movement, and topicalization. In other words, the first and the fourth row of the table are not disputed. It is less clear whether the predictions of the second and third row a correct. *Wh*-movement does not seem to feed topicalization (the topic marker in Japanese apparently does not attach to *wh*-phrases – see Kuroda (1965), Heycock (to appear) for a useful overview), this might be attributed to semantic incompatibility. The facts then would not bear on the ordering of operations. Whether scrambling feeds *wh*-movement is a disputed issue. The observation (von Stechow and Sternefeld (1988, p. 466), Fanselow (1990), Müller and Sternefeld (1993, 1996), (Müller, 1998, p. 39-40), (Pesetsky, 2000, p. 70-83), (Grewendorf, 2001, p. 110 fn. 37)) that *wh*-phrases in German are not usually allowed to scramble in multiple *wh*-questions suggests that scrambling does not feed *wh*-movement. On the other hand, the ban against scrambling *wh*-phrases is not rigidly observed ((Beck, 1996, 6-7), Sauerland (1996), Fanselow (2001, 2004)). In particular it is possible to scramble *wh*-words in front of quantified and other operator-like subjects, (15) (from Fanselow (2001, p. 414)). The fact that scrambling is clause bound and does not give rise to weak cross-over effects can then be used to explain two asymmetries in German *wh*-movement: (i) that short *wh*-movement is not subject to superiority but long *wh*-movement is and (ii) that short *wh*-movement does not give rise to weak cross-over effects but long *wh*-movement does (see Fanselow (2004); Wiltschko (1997, 1998) for ideas along these lines and Pesetsky (2000); Grewendorf (2001) for criticism).

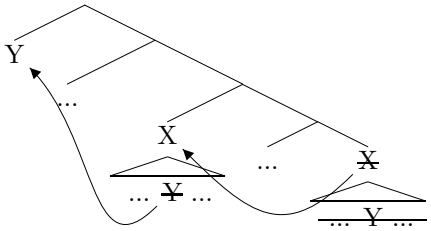
(14) a. ?*Wann hat wem der Mann geholfen?
when has who$_{DAT}$ the$_{NOM}$ man helped
When did the man help whom?

b. Wann würde wem nur ein Held helfen?
when would who$_{DAT}$ only a hero help
When would only a hero help whom?

Although the discussion here is necessarily somewhat inconclusive, since the field has not reached a consensus on the issue, it is not unreasonable to suppose that the predictions of the table are borne out. Even if scrambling fed neither *wh*-movement nor topicalization, this would not disconfirm the hypotheses formulated in UCOOL and GenPIM, since the relevant constructions might be ruled out by factors that are orthogonal to operational ordering. The hypothesis would be disconfirmed if feeding relations between operations were discovered that correspond to the part of the table below the diagonal. No such cases exist.

In the discussion above I assumed that the rows and columns of table (13) are to be interpreted as subsequent movements of the *same* constituent. Suppose that instead we were to interpret the rows as movement of some constituent and the columns as movement of a subconstituent out of the moved one (the *surfing*-paths in Sauerland's 1996 terminology). In other words, the rows label the prior movement of X while the columns label the subsequent movement of Y in (15).

(15)



To be sure, such structures have traditionally been ruled out categorically by the freezing prinicple (Ross (1967/1986); Wexler and Culicover (1980)), which postulates that moved constituents are islands for extraction. I will now review some evidence showing that the freezing principle is too strong and that the predictions that emerge from UCOOL together with GenPIM are closer to being correct.

The first row of (13) now claims that A-movement should feed scrambling, *wh*-movement, and topicalization. An example of the first predicted is furnished by (16b) on the assumption that objects in German move to Case positions. We can demonstrate that the prepositional adverbial *darüber* has moved out of the NP by observing the specificity effect in (16b) (for discussion of this point see e.g. Müller (1998, p. 10-15)).

(16) a. ...weil niemand {✓ ein |✓ das} Buch darüber lesen wollte
...because nobody [a the book there-about] read wanted
...because nobody wanted to read {✓a|*the} book about that

b. ...weil darüber niemand {✓ ein | das} Buch $t_{darüber}$ lesen wollte
...because there-about nobody [a the book $t_{darüber}$] read wanted
...because nobody wanted to read {✓a|*the} book about that

Examples bolstering the second claim, that A-movement may feed *wh*-movement of a constituent contained in the moving item, are easy to come by. (17) furnishes a simple example of extraction from an A-moved argument. The point made by this example stands independently of whether or not the ECM subject remains in the embedded infinitival or moves into the matrix. The fact that A-movement can feed *wh*-movement can also be made on the basis of the existence of exceptions to the subject island condition, (18b) (from (Starke, 2001, p. 36)), which,

according to Starke, is at worst slightly worse than the corresponding object extraction case, (18a). The same appears to be true in the English - and, to my ear, German - translations. (Notice that although the NPs from which extraction happens here are formally definite, they are semantically unspecific because of the inherently relational nature of the nouns.)

(17) Which football team do you want [the manager of $t_{\text{which football team}}$] to pay a large fine?

(18) a. De quel film as tu raté la première partie
of which film have you missed [the first part t]
Which film did you miss the first part of?

b. De quel film est-ce que tu crois que la premiére partie va créer une scandale?
of which film is-it that you think that [the first part t] goes create a scandale
Which movie do you think that the first part of would create a big scandal?

Similar examples with topicalization from a subject are easy to construct, (19). The discourse particle *doch* and the temporal adverbial are introduced to make sure that the subject cannot remain VP-internally (for discussion of VP-internal subjects and discourse particles see Haider (1993); Diesing (1992) among many others).

(19) Von diesem Film hat der erste Teil doch letztes Jahr einen grossen Skandal ausgelöst.
of this film has the first part PRT last year a big scandal caused
The first part of this film caused a big scandal last year.

This shows that the prediction generated by UCOOL and GenPIM, that A-movement should feed scrambling, *wh*-movement, and topicalization is borne out.

The predictions of the second row are similarly borne out. Scrambling feeds both *wh*-movement and topicalization. This is shown in (20) and (21a).[8] There is a sharp contrast between (21a) with topicalization from a scrambled constituent and (21b) with scrambling from a topicalized one.

(20) a. Worüber kann einen Südkurier-Artikel selbst Peter nicht am
what-about can [a Südkurier-article $T_{\text{worüber}}$] even Peter not at.the
Strand verfassen?
baech write
For which topic is it the case that even Peter cannot write an article about it for the Südkurier when he is at the beach. (from (Hale and Legendre, 2004, p. 187) attributed to (Kuthy and Meurers, 1999, p. 5 [sic!]))

b. Über welchen deutschen Kaiser hatte ein fertiges Manuskript
about which German emperor had [a done manuscript $t_{PP}]_{NP}$

[8] These examples are not acceptable to all speakers.
The fact that many examples in the literature where scrambling is followed by subextraction are unacceptable is probably due to a specificity island (see (Fanselow, 2001, p. 413) for this suggestion).

leider keiner der anwesenden Historiker anzubieten.
unfortunately none the present historians t_{NP} to.offer
About which German emperor could none of the historians present offer a completed manuscript?

(21) a. Über Karl den Grossen hatte ein fertiges Manuskript leider
[About Charlemagne]$_{PP}$ had [a done manuscript t_{PP}]$_{NP}$ unfortunately
keiner der anwesenden Historiker anzubieten.
none the present historians t_{NP} to.offer
About Charlemagne, none of the historians present could offer a done manuscript.

b. *Auf dem Kongresse hat über Karl den Grossen keiner gesagt ein fertiges
at the congress has about Charlemagne nobody said a done
Manuskript habe er anzubieten.
manuscript have he to.offer
At the congress nobody said that he had a completed manuscript about Charlemagne to offer.

The final case, topicalization out of a *wh*-moved constituent is degraded, (22a). However, there is still a clear contrast between (22a) and *wh*-movement out of a topicalized constituent, (22b).

(22) a. ??Über Karl den Grossen weiß ich nicht, was für ein Buch er
[about Charlemagne]$_{PP}$ know I not [what for a book t_{PP}]$_{NP}$ he t_{NP}
schreiben will.
write wants.
About Charlemagne I don't know what kind of book he wants to write.

b. *Über welchen deutschen Kaiser sagt er ein fertiges Manuskript hat
[about which German emperor]$_{PP}$ says he [a done manuscript t_{PP}]$_{NP}$ has
keiner anzubieten.
nobody t_{NP} to.offer
About which German emperor does he say that, a completed manuscript, nobody can offer?

Interestingly, note that there is a predicted asymmetry between *wh*-movement from an A-moved item and A-movement from a *wh*-moved item. The former is possible as the extraction from ECM subjects and the exceptions to the subject condition above illustrate. The opposite seems to be impossible. Thus, consider example (23) from Sakai (1994, p. 300), which is discussed as a problem in Collins (2005). On the intended but ungrammatical construal where *Oscar* A-moves out of the *wh*-moved clause, (23b) ought to mean something like *It was asked how likely it was that Oscar would win.* What has goes wrong from the present perspective is that A-movement of *Oscar* and $\overline{A}$-movement *how likely to win* occur in an order violating GenPIM. A-movement out of an $\overline{A}$-moved constituent is correctly predicted to be impossible.

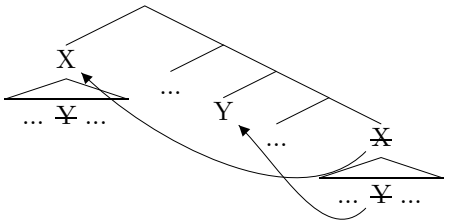
(23) a. Max asked how likely to win Oscar was
[$_{\text{IP1}}$ Max asked [$_{\text{CP2}}$ [$_{\text{AP}}$ how likely t_j to win]$_i$ [$_{\text{IP2}}$ Oscar$_j$ was t_i]]]

b. *Oscar was asked how likely to win it was.[9]
[$_{IP1}$ Oscar$_j$ was asked [$_{CP2}$ [$_{AP}$ how likely t$_j$ to win]$_i$ [$_{IP2}$ it was t$_i$]]]

The very simple theory laid out above predicts the possible and impossible patterns of extraction out of moved constituents with a surprisingly high degree of accuracy. It is clearly superior to a theory that bans all extraction from moved constituents by appeal to the freezing principle.

Finally, let us turn to the third interpretation we can give to the table in (13). Consider (24). We can interpret the rows as indicating a first movement step of Y and the columns as the following remnant movement step of X.[10]

(24)



The first row of the table in (13) now claims that A-movement may be followed by remnant scrambling, remnant *wh*-movement, and remnant topicalization. All three predictions are true. For *wh*-movement and remnant topicalization this can be demonstrated even for English, (25).

(25) a. [How likely t$_{John}$ to win the race]$_{AP}$ is John t$_{AP}$?
b. [Criticized t$_{John}$ by his boss]$_{VP}$ John has never been t$_{VP}$.

Scrambling of predicates in front of the subject in German is always somewhat degraded, (26a) (Müller, 1998, p. 226). It may be debatable whether the scrambled category α in (26a) contains the trace of the subject, since it may be that a smaller constituent than the full VP that is fronted. The same is not true in (26b), where the subject is almost certainly extracted from α. To my ear, these examples have a similar status. It is unclear what to conclude from these facts. A possibility consistent with the theory developed here is that the degradation stems from a mismatch between the categorial features of the scrambled constituent and the requirements imposed by the target.

(26) a. ??... dass das Buch gelesen keiner hat
... that [t$_{keiner}$ the book read]$_\alpha$ nobody t$_\alpha$ has
... that nobody has read the book
b. ??... dass von Maria angefasst kein einziges Buch werden sollte
... that [by Maria t$_{NP}$ touched]$_\alpha$ [no single book]$_{NP}$ t$_\alpha$ become should

[9] This sentence is of course possible if we construe *Oscar* as the object of *ask* and *it* referentially.

[10] In other words, I am discussing Sauerland's (1996) *diving paths* now. The movement of Y is what Grewendorf (2003) somewhat misleadingly calls *internal* movement. The system proposed here incorporates Grewendorf's (2003) condition on improper remnant movement as a special case, (i):

(i) *Improper Remnant Movement* (Grewendorf, 2003, p. 67)
Remnant movement is prohibited unless it is of a higher type than internal movement.

. . . that not a single book should be touched by Mary

The rest of the table under this interpretation is unproblematic. Müller (1998) has argued in great detail that scrambling may feed both remnant *wh*-movement, (27) (based on Müller (1998, p. 221-222)), and remnant topicalization, (28) (Müller, 1998, p. 10).[11]

(27) Was für ein Buch hat über die Liebe niemand gelesen.
[what for a book $t_{PP}]_{DP}$ has [about the love]$_{PP}$ nobody t_{DP} read
What type of book about love has nobody read?

(28) Ein Buch gelesen hat darüber keiner
[[a book t_{PP}] read]$_\alpha$ has there-about no-one
A book about that, nobody has read.

Finally, remnant topicalization can be fed by *wh*-movement. Such examples involve extraction from a *wh*-island, which always involves a certain amount of degradation. However, the example in (29a) is no more or less degraded than that in (29b) (based on (Müller, 1998, p. 301)).

(29) a. ??Hans das Buch zu geben weiß ich nicht warum Fritz abgelehnt hat.
[Hans the book to give]$_\alpha$ know I not why Fritz t_α refused has
I don't know why Fritz refused to give Hans the book.
b. ??Hans zu geben weiß ich nicht was Fritz abgelehnt hat.
[Hans t_{was} to give]$_\alpha$ know I not what Fritz t_α refused has
I don't know what Hans refused to give Fritz.

The novel approach to improper movement suggested here ties together, via UCOOL and GenPIM, improper movement with extraction out of moved phrases and remnant movement. It makes immediate and surprisingly accurate predictions about the various feeding and bleeding relationships between movement operations.

2.3 Discussion

In this section I discuss some alterantive approaches to improper movement and how they deal with the facts reviewed above.

Notice first that the classical government and binding theoretic approach to improper movement in terms of the binding theory is insufficient. This was already observed in Müller and Sternefeld (1993), who charged that the standard account of improper movement countenances

[11] It could be countered that the examples predicted to be bad under this interpretation of table (13) are ungrammatical simply because the putative second operation is clause bound. True. But question begging. On usual assumptions clause boundedness is described in terms of a prohibition against using [Spec, CP] as an intermediate landing site and that [Spec, CP] is not the right kind of landing site for the bounded movement. In other words, the explanation comes back to improper movement anyway.

Moreover, in the system of Chomsky (2001b) long movement does not need to pass through the phase edge as an escape hatch any more. The lower phase is only spelled out when the next higher phase-head is introduced. This allows movement to proceed from phase-medial to phase-medial position (see Abels (2003, ch 2) for discussion), which makes even the description of clause-boundedness a tricky and unresolved issue.

Some of the possible variations on GenPIM discussed in the next subsection handle such cases quite easily. Without a viable counterproposal then, this objection has little force.

too few types of movement. They did not think that it was necessary to order movement types along a hierarchy, though. Asymmetries like those between *wh*-movement out of A-moved items, (25a), and A-movement out of *wh*-moved items, (23b), between topicalization out of scrambled items, (21a), and scrambling out of topicalized items, (21b), etc., strongly suggest that there are not only more types of movement, but that there is a hierarchical order between them. Certainly the standard GB approach to improper movement is insufficient.

Second, certain cases of improper movement, (8d) for example, can be ruled out either by the current theory or by Rizzi's (2004a) principle of criterial freezing, which demands that an item which has reached an operator position where it is interpreted for scope purposes cannot be moved from that position.[12] However, cases like (8b) cannot be ruled out by criterial freezing, because the *wh*-phrase is not interpreted for scope in the intermediate [Spec, CP]. Moreover, criterial freezing cannot account for examples like (23b) because of the existence of examples like the following from (Rizzi, 2004a, p. 13 ex. 36b) (see also Chomsky (1986, p. 26); Starke (2001, p. 54)).[13] Thus criterial freezing though partly redundant with the current theory does not deliver a complete theory of improper movement.

(30) ?Di quale autore ti domandi quanti libri siano stati censurati
[by which author]$_i$ C$_Q$ you wonder [how-many books t$_i$]$_k$ C$_Q$ have been censored t$_k$]]
By which author do you wonder how many books have been censured?

Third, Sternefeld (1992) considers a theory of improper movement which relates UCOOL directly to the hierarchy of functional projections. Anachronistically, I will simply call this sequence the Cinque hierarchy. Sternefeld's idea is simple and appealing: Suppose that there is a fixed hierarchy of functional heads $H_1, H_2, \ldots H_n$ in that (bottom to top) order. Then, any movement that has as its launching site a projection with index i must land in a projection with index j where $i \geqslant j$. This rules out standard cases of improper movement on the unobjectionable assumption that complementizers are higher than tense. However, this formulation is both too strong and too weak. It is too weak, because it fails to rule out examples like (23b), where the launching site is an infinitival [Spec TP] and the landing site a finite [Spec, TP]. An unobjectionable configuration under the theory. It is too strong to leave room for the now standard account of English ECM as raising into the matrix clause pioneered by Postal (1974) and revived by Lasnik and Saito (1991); Lasnik (1998) (for references and clear discussion of the issues see Davies and Dubinsky (2004)). The problem is that the ECM subject ends up fairly low in the matrix clause but it must have passed through the embedded [Spec, TP] (see Bošković (2002) for evidence). The launching site, then, is higher than the landing site.

Fourth, we could strengthen Sternefeld's rule to say that any movement that has as its launching site a projection with index i must land in a projection with index j where $i \geqslant j$ and

[12] Under Rizzi's assumptions this might lead to the wrong expectation that the entire embedded CP becomes pied-pipable in examples like (8d), (i).

(i) *Whose book to be read was asked.

[13] Example (30) involves, in Rizzi's terms, subextraction from a phrase which is frozen in place by criterial freezing. As L. Rizzi (p.c.) clarifies, subextraction, from potentially abstract containing phrases, will also have to be invoked to account for cases where an item is interpreted in more than one position (split scope, binding reconstruction, etc.)

moreover there must be no projection along the path of this movement with index k such that $k > j$.[14] Path is to be understood in the obvious way here: the set of nodes that dominate the launching site and do not dominate the landing site. Although this formulation now does rule out (23b), the analysis of ECM becomes even more difficult. Similar problems arise with quantifier raising in French and Icelandic across complementizers, (31) (French is mentioned as a problem in Williams (2002)), long scrambling in Russian, which crosses complementizers but does not seem to end up in [Spec, CP], hyperraising (see Nevins (2004); Williams (to appear) for pertinent discussion).

(31) French ((Kayne, 1998, p. 141-142 #52-3))

a. Il n' a rien fallu que je fasse.
it neg has nothing been-necessary that i do
It hasn't been necessary for me to do anything.

b. Il a tout fallu que je leur enlève.
it has everything been-necessary that I them$_{dat}$ remove
I had to take everything off them.

Icelandic ((Kayne, 1998, p. 142 #55-6), attributed to Rögnvaldsson (1987, p. 44) and Jónsson (1996, p.86))

c. Hann mun ekkert hafa getadh gert.
they will have nothing have could done
They haven't promised to do anything.

d. Their hafa ekkert lofadh adh gera.
they have nothing promised to do
They haven't promised to do anytying.

Likewise cyclic movement across complementizers ending up in a position below the complementizer, (32), are ruled out by this condition.[15] The examples in (32) are from Kîîtharaka, a Bantu language spoken in Kenya. Kîîtharaka allows *wh*-in situ, partial *wh*-movement, and full *wh*-movement (see Muriungi (2005) for extensive discussion). The examples in next are therefore all synonymous. What is important here is that the *wh*-word moves across the complementizer *atî* and follows it in the partial *wh*-movement structures, (32d). Assuming *atî* is always in the same position and assuming that the position of the *wh*-word after the complementizer indicates that it is below the complementizer, we are left with the conclusion that *wh*-movement lands in a position just below the complementizer position that is crossed along the way. Examples of this kind are easy to multiply. It seems safe to conclude then, that this way of strengthening Sternefeld's account would be ill-advised.

[14] Incidentally, this is a corollary of Nevins' (2004) translation of Williams' Representation Theory. Williams' own theory puts this demand only on operations analyzed as movement, whereas other operations – including ECM – are analyzed as mismappings. Mismappings do not obey the constraint on movement formulated in the text. An appraisal of Williams' concept of mismapping would lead too far afield here.

[15] Non-obvious glosses are as follows:

pres	present tense	sm	subject agreement marker
perf	perfective	FV	final vowel
F	marker of focus and cyclicty		

(32) Kîîtharaka (Peter Muriungi, p.c.)

a. U- ri- thugania atî John a- ug- ir- e atî Pat a- ug- ir- e
2nd.SG- pres- think that John sm- say- perf- FV that Pat sm say pres FV
Lucy a- ring- ir- e ûû
Lucy sm- beat- perf- FV who
Who do you think that John said that Pat said that Lucy beat?

b. U- ri- thugania atî John a- ug- ir- e atî Pat a- ug- ir- e n-
2nd.SG- pres- think that John sm- say- perf- FV that Pat sm say pres FV F-
ûû Lucy a- ring- ir- e t_{wh}
who Lucy sm- beat- perf- FV
Who do you think that John said that Pat said that Lucy beat?

c. U- ri- thugania atî John a- ug- ir- e atî n- ûû Pat a- ug- ir-
2nd.SG- pres- think that John sm- say- perf- FV that F who Pat sm say pres
e Lucy n- a- ring- ir- e t_{wh}
FV Lucy F- sm- beat- perf- FV
Who do you think that John said that Pat said that Lucy beat?

d. U- ri- thugania atî n- ûû John a- ug- ir- e atî Pat n- a- ug-
2nd.SG- pres- think that F- who John sm- say- perf- FV that Pat F sm say
ir- e Lucy n- a- ring- ir- e t_{wh}
pres FV Lucy F- sm- beat- perf- FV
Who do you think that John said that Pat said that Lucy beat?

e. N- ûû u- ku- thugania atî John n- a- ug- ir- e atî Pat n- a- ug-
F who 2nd.SG- pres- think that John F sm- say- perf- FV that Pat F sm say
ir- e Lucy n- a- ring- ir- e t_{wh}
pres FV Lucy F- sm- beat- perf- FV
Who do you think that John said that Pat said that Lucy beat?

Although the discussion here is far from complete and although a number of variations on the general theme of how to deploy UCOOL come to mind, I will stick with the formulation offered above. That formulation was (9), repeated as (33).

(33) GenPIM
No constituent may undergo movement of type τ if it has been affected by movement of type σ, where $\tau << \sigma$ under UCOOL.

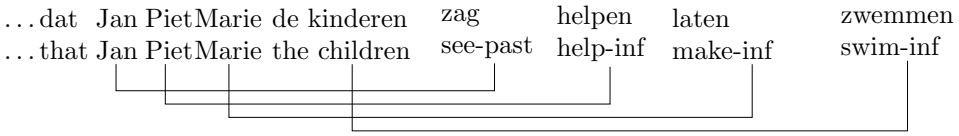
The asterisks in table (13) are straightforwardly testable consequences of the theory suggested here. The cases that receive an asterisk are predicted to be systematically absent.

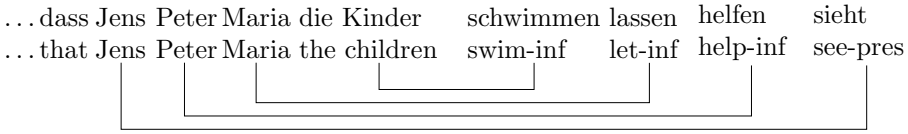
3 Cross-Serial Dependencies

In the debates concerning the weak and strong generative capacity of human languages (Pullum and Gazdar (1982); Bresnan et al. (1982); Huybregts (1984); Shieber (1985); Higginbotham

(1984); Culy (1985); Savitch et al. (1987)) the issue of unbounded cross-serial dependencies between non-identical elements played an important role. The other main issue were cross-serial dependencies between identical elements, that is, reduplicative processes in the broadest sense (see Stabler (2004) for important discussion and references). I ignore the case of reduplication here, since it does not bear directly on the issue of improper movement as far as I can see. Although unbounded cross-serial dependencies have remained important in the literature on mathematical linguistics (see Kracht (2003) and references given there), more descriptively oriented work (outside of West Germanic) has largely ignored them (though Tagalog is claimed to exhibit them in Maclachlan and Rambow (2002)).

In the so-called verb-raising construction Dutch famously exhibits crossing dependencies between verbs and their notional subjects, (34) from Bresnan et al. (1982). The standard German versions of such sentences show nesting dependencies, (35). The corresponding English construction is shown in (36). The same crossing pattern emerges for the dependencies between the verbs and their notional objects in Dutch. In addition there is, of course, a dependency between the verbs. This selectional dependency between the verbs goes from left to right in Dutch and English and from right to left in German.

(34) 

(35) 

(36) 

Huybregts (1976) argues that the verb-raising construction is unbounded.[16] Speakers usually have difficulties with sentences containing cross-serial dependencies of length four or more. To defend the claim that the grammar of Dutch should allow unbounded cross-serial dependencies Huybregts gives the following arguments: First, a number of verbs can come in either order under verb raising, (37). Second, for the triple *let–help–teach* Huybregts shows that the verbs can come in any of the six mathematically possible orders, (38). Finally, the construction is recursive since one and the same verb can occur more than once, (39). The apparent upper bound on the verb-raising construction is therefore usually attributed to processing complexity.

[16] Although these particular constructions in Dutch, unlike Swiss German (for which see Huybregts (1984); Shieber (1985)), are recognizable by a context-free device when viewed as a string language, there can be very little doubt indeed that structurally, context sensitivity is needed here (Huybregts (1976); Bresnan et al. (1982); Kracht (2005); Steedman (2000)). Huybregts (1984) discusses a slightly different construction in Dutch which is even weakly context sensitive.

(37) All examples from Huybregts (1976, p. 43-44 ex. 15-19)

a. let/see
   (i) . . . dat ik Cecilia een dokter zag laten komen
   . . . that I Cecilia a doctor saw let come
   . . . that I saw Cecilia let a doctor come
   (ii) . . . dat Cora ons twee nijlpaarden liet zien paren
   . . . that Cora us two hippos let see mate
   . . . that Cora let us see two hippos mate

b. let/try
   (i) . . . dat ik Cecilia wodka liet proberen te drinken
   . . . that I Cecilia vodka let try to drinken
   . . . that I let Cecilia try to drink vodka
   (ii) . . . dat Cora ons Arabisch probeerde te laten srpeken
   . . . that Cora us Arabic tried to let speak
   . . . that Cora tried to let us speak Arabic

c. try/help
   (i) . . . dat ik hem Cecilia probeerde te helpen verleiden
   . . . that I him Cecilia tried to help seduce
   . . . that I tried to help him seduce Cecilia
   (ii) . . . dat Cora ons de Petit Dru hielp proberen te beklimmen
   . . . that Cora us the Petit Dru helped try to climb
   . . . that Cora helped us try to climb the Petit Dru

d. help/teach
   (i) . . . dat ik Cecilia stijgijzers hielp leren aantrekken
   . . . that I Cecilia climbing-irons helped learn on-put
   . . . that I helped Cecilia learn to put on climbing-irons
   (ii) . . . dat Cora ons kinderen leerde helpen oversteken
   . . . that Cora us children taught help cross
   . . . that Cora taught us help children cross the street

e. teach/see
   (i) . . . dat ik Cecilia een microbe leerde zien bewegen
   . . . that I Cecilia a microbe taught see move
   . . . that I taught Cecilia to see a microbe move
   (ii) . . . dat hij ons Cora zag leren zwemmen
   . . . that he us Cora was teach swim
   . . . that he saw us teach Cora to swim

(38) Huybregts (1976, p. 44 ex. 20-25)

a. . . . dat we hem 'n druide nijlpaarden laten helpen leren paren
   . . . that we him a druide hippos let help teach mate
   . . . that we let him help a druid teach hippos mate

b. . . . dat we hem 'n druide nijlpaarden leren helpen laten paren
   . . . that we him a druid hippos teach help let mate
   . . . that we taught him help a druide let hippos mate

c. . . . dat we hem 'n druide nijlpaarden laten leren helpen paren
. . . that we him a druid hippos let teach help mate
. . . that we let him teach a druid help hippos mate

d. . . . dat we hem 'n druide nijlpaarden helpen leren laten paren
. . . that we him a druide hippos help teach let mate
. . . that we help him teach a druid let hippos mate

e. . . . dat we hem 'n druide nijlpaarden leren laten helpen paren
. . . that we him a druid hippos teach let help mate
. . . that we teach him let a druide help hippos mate

f. . . . dat we hem 'n druide nijlpaarden helpen laten leren paren
. . . that we him a druide hippos help let teach mate
. . . that we help him let a druid teach hippos mate

(39) Examples (a-bi) from Huybregts (1976, p. 45-46 ex. 26-28), (bii) from Bresnan et al. (1982, ex. 4), (c) provided by Hans van de Koot, p.c.

a. help

(i) Ik heb hem Cecilia helpen helpen verhuizen
I have him Cecilia help help move
I have helped him help Cecilia move.

(ii) Hij heeft me een leeuw leren leren blaten
he has me a lion teach teach bleat
He has taught me teach a lion how to bleat

(iii) Ik heb hem Cecilia det schaap een leeuw laten helpen leren helpen blaten
I have him Cecilia that sheep a lion let help teach help bleat
I have let him help Cecilia teaching that sheep how to help a lion bleat

b. teach

(i) Ik heb hem Cecilia iemand det shaap een leeuw laten helpen leren
I have him Cecilia someone that sheeep a lion let help teach
helpen leren blaten
help let bleat
I have let him help Cecilia teaching someone how to help that sheep teaching a lion how to bleat.

(ii) . . . dat de leraar Jan Marie de kinderen leerde laten leren zwemmen
. . . that the teacher Jan Marie the children taught let teache swim
. . . that the teacher taught Jan to make Marie teach the children to swim

c. see

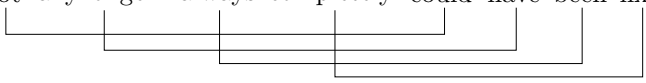
(i) . . . dat Jan de leraar Marie de kinderen ziet helpen zien zwemmen
. . . that Jan the teacher Marie the children sees help see swim
. . . that Jan sees the teacher help Marie see the children swim

I agree with the previous literature in assuming that the grammar of Dutch allows unbounded cross-serial dependencies. As is well-known, verb raising gives rise to clause-union effects. (See Evers (1975); Haegeman and Riemsdijk (1986) among many others and Wurmbrand (2006) for a recent overview.) Despite the clause-union effects, Huybregts' observations rule out an analysis

along the lines of the functional restructuring approach that Cinque (2006) advocates for other cases of clause union, because his approach is incompatible with free ordering and recursion.

Another example of cross-serial dependencies comes from certain dialects of Norwegian (Nilsen (2003, S. 72); Bentzen (2005)). The relevant construction is exemplified in (40) where all the adverbs must precede all the verbs despite the fact that the scope relations between verbs and adverbs cannot be read off this string directly. Scope relations rather follow the linear left-to-right order in the ungrammatical (40b).

(40) a. ... at det ikke lenger alltid helt kunne ha blitt ordnet
... that it not any.longer always completely could have been fixed

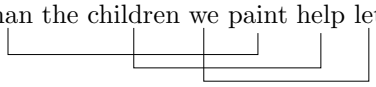


b. *... at det ikke kunne lenger ha alltid blitt helt ordnet
... that it not could any.longer have always been completely fixed

The cases of cross-serial dependencies exemplified here have one property in common: the left-to-right order of elements within each of the sub-sequences corresponds to the scope and/or selection within that sub-sequence. The verbs in the Dutch verb clusters come in their left-to-right order of selection. The same is true in the Norwegian example: both the auxiliaries and the adverbs take scope with respect to each other from left to right. I will call cross-serial dependencies where the left-to-right order corresponds to scope and selection within each sub-sequence *straight* cross-serial dependencies. Dutch and Norwegian exhibit straight cross-serial dependencies and so do Swiss German (Huybregts (1984); Shieber (1985)) and, according to the description in Maclachlan and Rambow (2002), Tagalog.

The opposite case would be an *inverse* cross-serial dependency, where the selectional and scope relations go from right to left. This hypothetical case is illustrated in (41), an inverted version of Shieber's famous Swiss German example, (42). To the best of my knowledge (see also Dehdari (2006) for a brief review of the literature) no case of an unbounded inverse cross-serial dependency has been discovered in human language.[17]

(41) Inverse cross-serial dependency:
... that the barn the man the children we paint help let



... that we let the children help the man paint the barn

(42) ... das mer d' chind em Hans es huus lönde hälfe
... that we.NOM the children.ACC the.DAT Hans the.ACC house let help
aastriiche
paint
... that we let the children help Hans paint the house

[17] The claim made here is quite different from that in Bobaljik (2004), who wonders about the apparent absence of predicate clusters in VO-languages. As noted in Wagner (2005) Norwegian (SVO) examples like (40) seem to be a counterexample to Bobaljik's generalization. Tagalog (VSO) is another apparent counterexample.

Despite the fact that little effort has been put into the description of unbounded cross-serial dependencies over the last 20 years, I claim that the absence of backwards cross-serial dependencies is not an accidental gap in our knowledge, but reflects a deep property of language. In other words, I claim that (43) is a true generalization concerning human languages. I suggest below that UCOOL together with GenPIM and a number of independently motivated assumptions derive the hypothesized systematic absence of unbounded inverse cross-serial dependencies. Universal (43) is another testable consequence of the theory proposed here.

(43) Universal:
UG does not allow unbounded inverse cross-serial depencies.

4 Deriving Straight Cross-Serial Dependencies

In this section I consider two types of derivation for straight cross-serial dependencies. Both of them rely strictly on leftward movement, but one relies heavily on remnant movement while the other one does not. I compare the two approaches with respect to their ability to predict the ban against inverse cross-serial dependencies, (43), and with respect to their compliance with GenPIM. The remnant movement account, it turns out, does not predict the asymmetry nor is it compatible with GenPIM. Before continuing, I need to justify the decision not to take into account any rightward-movement analyses of verb raising..

Traditional generative analyses of verb-raising in Dutch and Swiss German assume an underlying head-final structure and then transport the selected verb rightward around its selecting verb in some way (Evers (1975); Haegeman and Riemsdijk (1986)). The considerations of section 1 suggest that such analyses are wrong. An explanatory model of UG should countenance only leftward-movement, which is clearly independently justified, and remnant movement, which is also independently justified, but no rightward movement.[18] The two analyses of cross-serial dependencies considered below therefore do not invoke rightward movement.

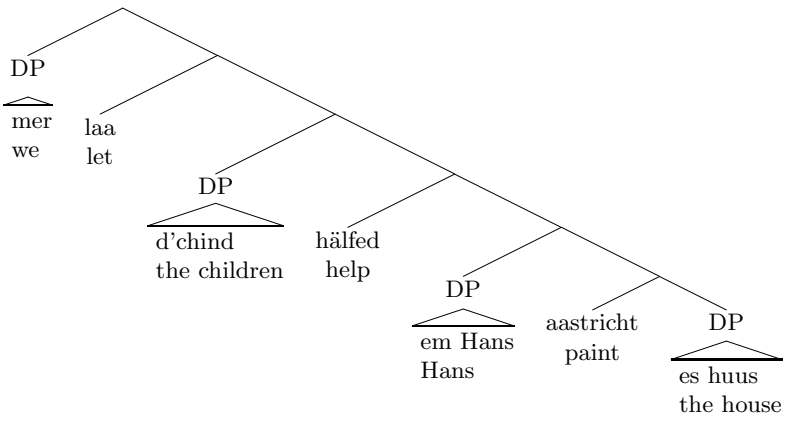
4.1 Straight Cross-Serial Dependencies without rightward movement and without remnant movement

In line with the argumentation so far, I will assume that there is a derivational point where the structure of (42) is roughly as in (44a). The details of the structure do not matter; there might be additional VP-shells present; there might be unpronounced controlled PROs; the c-command relations between the verbs and their respective arguments might be locally reversed... What is important is that the arguments are interspersed with – and locally related to – the verbs and that the c-command relations among the arguments and among the verbs are as given here. From an underlying structure with these characteristics, the cross-serial order can be derived by

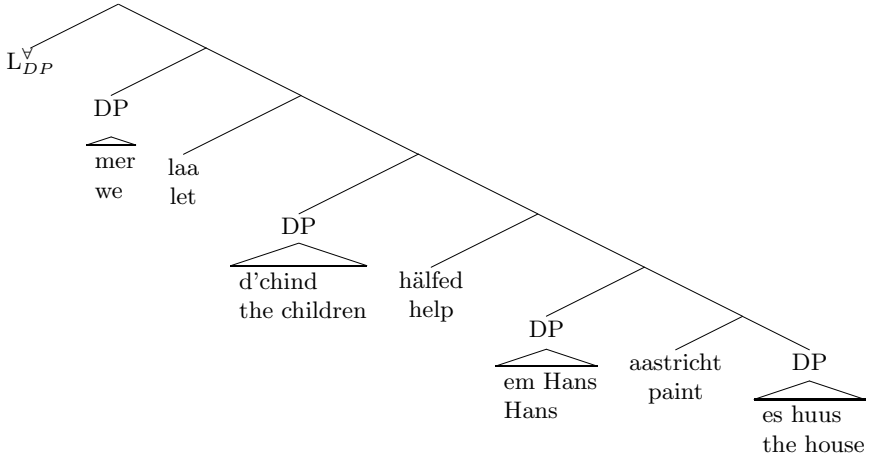
[18] A second reason for excluding rightward-movement as an analytic option is that such analyses almost invariable end up invoking a version of the right roof constraint, a constraint that, under a conception of the grammar with remnant movement and without rightward movement, ought to be derivable from constraints on remnant movement along the lines of UCOOL and GenPIM. The present formulation does not derive the right roof constraint. See the concluding remarks of this paper for further comment.

merging a licensing head with the property of attracting all DPs into its specifier positions, as shown in (44b-c).[19]

(44) a.

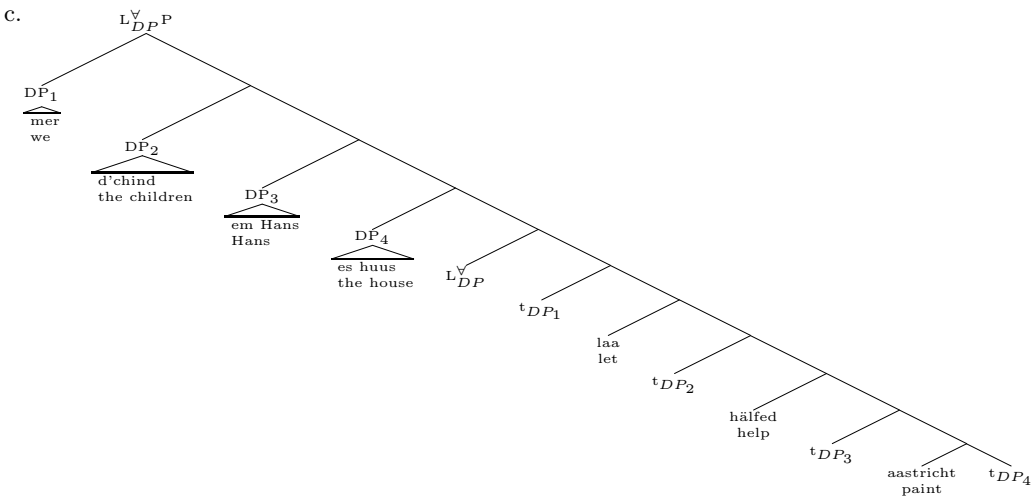


b.



[19]The constituent structure thus derived is not the one argued for in Bresnan et al. (1982) (see also Steedman (2000)). Bresnan et al.'s constituent structure can be derived if instead of lifting all arguments out of the verbal projections all verbs are lifted out of the initial structure in an order-preserving fashion with subsequent movement of the remnant of the tree in (44a). Such a derivation might indeed be called for. Nothing in the logic of this paper hinges on the choice. For reasons of expository simplicity, I assume the derivations given here.

c.



In order to derive the cross-serial order in (44c), I invoked two crucial assumptions. First, I assumed that there are heads that attract all (active) members of a given class, here DP: $L^{\forall}_{DP}$. The account draws on Bošković's (1999) theory of multiple *wh*-fronting in Slavic, where the existence of heads with the property "Attract all!" are postulated. Again the details are not particularly important here. An account that recursively stacks up as many DP-licensors as there are active DPs (constrained by convergence) would achieve the same result. Based on evidence like that reviewed in section 3 it was conjectured that cross-serial dependencies can be unbounded. If this is correct, then some assumption along these lines (recursion of the same head, recursion of a sequence of heads, or attract all) is necessary to account for these constructions. A grammar where each of the arguments is moved into a dedicated position, that is, where each argument's landing site has different properties and a different label, cannot describe the facts adequately. In other words, a linearly ordered, finite sequence of dedicated heads attracting a single DP is insufficient. On its usual construal, the Cinque hierarchy has exactly these properties: it is a linearly ordered, finite sequence of dedicated heads regulating the occurrence of single instances of a feature. Given their finite length, such sequences need recursion or the attract-all property to deal with unbounded cross-serial dependencies.

Second, I assumed that movement processes, though they disturb the order between elements of different classes, are generally order preserving when it comes to elements of the same class. This intuition lies behind the various formulations of Relativized Minimality (Rizzi (1990); Chomsky (1995b) for two prominent formulations). Relativized Minimality demands that pre- and post-movement structures – when relativized to a particular property, feature, or class of positions – be homomorphic: what was higher in the pre-movement structure is higher in the post-movement structure as well, (45a-b).[20] By contrast (45c) is disallowed by Relativized Minimality, because the pre- and post-movement orders are reversed, that is, there is no homomorphism here.

[20]Note that derivations of the type in (ib) might be impossible because of the extension condition. This does not affect the point that Relativized Minimality per se allows such derivations.

By the logic of order preservation derivations such as those in (45d) should also be allowed since the pre-movement c-command relations between α and β and the post-movement c-command relations are the same. This generalization of Relativized Minimality to chains is due to Starke (2001, chapter 8) (see also Richards' 1997 mechanism of *tucking in*), a concrete formulation is given in (46a).

(45) For $\alpha,\ \beta\ \in \Delta$, where Δ stands for some property, featurally defined class, or class of positions, $\prec$ the asymmetric c-command relation, and $\vDash_\Omega$ the relation of possible derivational subsequence through an operation targeting an element in class Ω:

a. $\alpha \prec \beta\ \vDash_\Delta \alpha \prec t_\alpha \prec \beta$
b. $\alpha \prec \beta\ \vDash_\Delta \alpha \prec \beta \prec t_\beta$
c. $\alpha \prec \beta\ \nvDash_\Delta \beta \prec \alpha \prec t_\beta$
d. $\alpha \prec \beta\ \vDash_\Delta \alpha \prec \beta \prec t_\alpha \prec t_\beta$

The empirical expectation this gives rise to is that order preservation should be common and that nesting dependencies should be rare. This expectation seems to be borne out (see Fanselow (1993); Richards (1997); Müller (2001); Starke (2001); Koopman and Szabolcsi (2000); Williams (2002); Fox and Pesetsky (2005) and the other contributions in that issue of Theoretical Linguistics among others). This view also makes concepts like equidistance Chomsky (1993) superfluous, which was designed specifically to allow an order preserving map between θ- and case-positions for subject and object. The cases (e.g., Kuno and Robinson (1972)) that were assumed to fall under the generalization that nesting paths are preferred to crossing paths (see Pesetsky's Path Containment Condition (1982)) will presumably have to be reanalyzed in terms of a more-fine grained typology of movement together with a syntactic feature geometry.

Again, any account of unbounded cross-serial dependencies must provide a general solution to the order-preservation problem. The unbounded nature of the phenomenon rules out a solution in terms of a linear sequence of dedicated projections. Relativized Minimality is the most plausible candidate.

The crucial assumptions are summarized again below.

(46) Assumptions:

a. If two or more items undergo the same type of movement, movement is order preserving
   (i) Relativized Minimality
   If X and Y are members of the same class, then X may not cross Y.
   (ii) *Crossing* – generalized version (see Starke (2001))
   X *crosses* Y iff all occurrences of Y asymmetrically c-command at least one occurrence of X; and at least one occurrence of X asymmetrically c-commands all occurrences of Y.
b. There are processes that correspond to the "attract-all" model in Bošković (1999).
c. Movement is to the left.

The Norwegian case, (40), can be dealt with along the same lines if we assume that there is a head which attracts all adverbs, $\mathrm{L}^{\forall}_{Adv}$.

Derivations like these are fully consistent with UCOOL and GenPIM. The obvious question to ask at this point is whether the ban against inverse cross-serial dependencies, (43), can be derived under the assumptions made in this subsection. I postpone discussion of this question until section 5 and turn to a second approach to straight cross-serial dependencies first.

4.2 Straight Cross-Serial Dependencies with remnant movement

Nilsen (2003) and Bentzen (2005) generate straight crossing dependencies through a series of remnant movements. The crucial assumptions driving the accounts are given in (47).

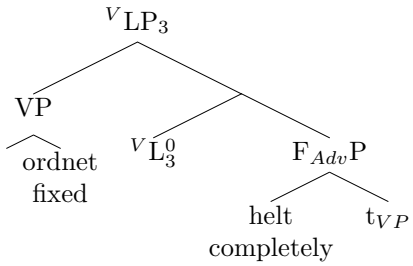
(47) a. Above every verbal head there is an adverb lifter, that is, a head which attracts the maximal projection of the closest head with an adverbial in its specifier, ${}^{Adv}\mathrm{L}^0$.

b. Above every functional head with an adverb in its specifier there is a verb lifter, ${}^{V}\mathrm{L}^0$, that is, a head which attracts the maximal projection of the closest verbal head (V, auxiliary, modal).

On these assumptions, the Norwegian cross-serial dependency, (40), is derived as follows. First, the VP with the main verb is formed, (48). Subsequently the adverb is merged in the specifier position of a dedicated functional head, F^0_{Adv}, the associated verb lifter, ${}^{V}\mathrm{L}^0_3$, is merged, and the original VP moves to its specifier position, (49).

(48)

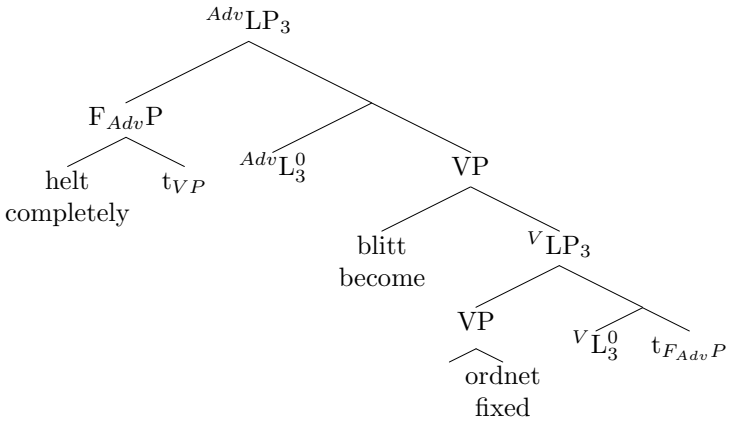
(49)



Notice that under the definition of affectedness given above, (10), the nodes along the path of VP-movement are affected by that movement. Any subsequent movement of any of those constituents cannot be of a type that is lower than the type of VP-movement without violating GenPIM; it will have to be of the same or of a higher type.

The next steps consist in merging the auxiliary *blitt* - 'become' and its associated adverb lifter, ${}^{Adv}\mathrm{L}^0_3$. Then $\mathrm{F}_{Adv}\mathrm{P}$, containing only the adverb *helt* - 'completely', moves to the specifier of ${}^{Adv}\mathrm{LP}_3$, (50).

(50)

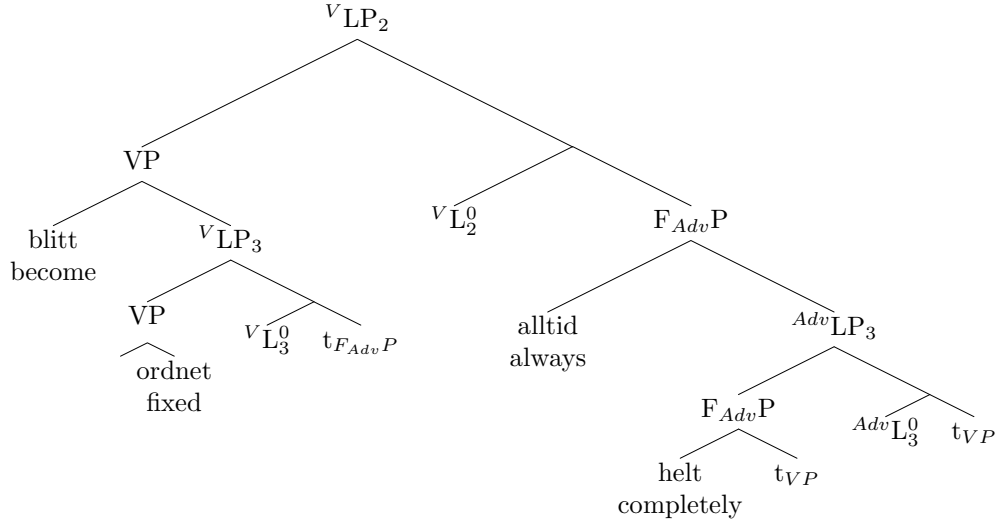


From the reasoning in the paragraph following (49), we now conclude that movement of F$_{Adv}$P is either of a higher type than movement of VP or of the same type. In fact, a stronger conclusion can be drawn: movement of F$_{Adv}$P is of a higher type than movement of VP. If both of the movements were of the same type, we would have a subextraction of type α followed by remnant movement of type α. As discussed in several places in the literature on remnant movement (Müller (1998, chapter 5); Kitahara (1997); Hiraiwa (2002); Grewendorf (2003)), remnant creating movement of type α cannot be followed by remnant movement of the same type α. This restriction is often referred to as Müller's generalization. As suggested in some of the literature mentioned, this generalization might follow from the minimal link condition, in other words, it is a case of order preservation in the broadest sense. I will assume that Müller's generalization holds. On the basis of Müller's generalization we can rule out the option that movement of VP in (49) and movement of F$_{Adv}$P in (50) are of the same type. The only option left is that UCOOL makes sure that Move$_{VP}$ $<<$ Move$_{F_{Adv}P}$.

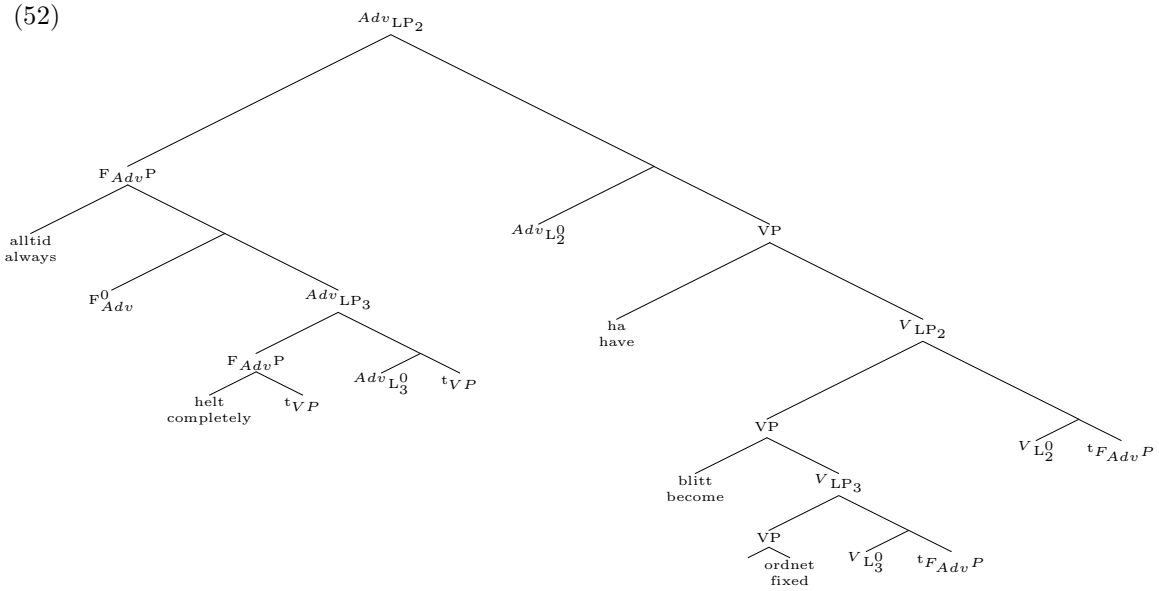
At this point the derivation appears to become an uneventful repetition of sequences of the same steps: A new adverb is merged in the specifier of a designated functional head, a verb lifter is merged on top of it, and the highest verb phrase, containing *blitt ordnet* – 'become fixed' moves, (51). Then another verb merges, an adverb lifter merges on top of it, and the closest F$_{Adv}$P, containing *alltid helt* – 'always completely' moves, (52).

However, given the conclusion just drawn about the order of movement of F$_{Adv}$P and VP, the movement step in (51) is illicit under GenPIM. We concluded that Move$_{VP}$ $<<$ Move$_{F_{Adv}P}$. In (50), VP among other nodes was affected by movement of F$_{Adv}$P. GenPIM now dictates that at this point VP is no longer accessible for operations that are ordered below Move$_{F_{Adv}P}$. Therefore (51) is an illicit derivational step and structures derived from (51) are ill-formed.

(51) V-lifter



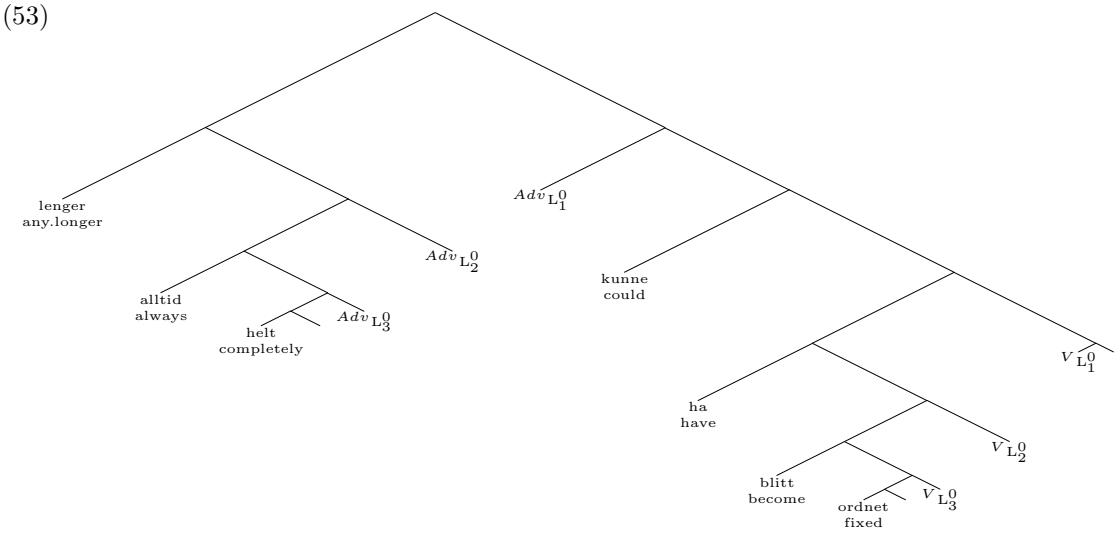
(52)



After another – equally illict under the assumptions entertained here – cycle through the "adverb, verb lifter, verb, adverb lifter" sequence, the structure is as in (53): The adverbs are collected in one constituent in an order-preserving fashion, the verbs are collected in another constituent, also in an order preserving fashion.[21]

[21] Note that, since Norwegian allows positioning the subject in between the various adverbs, there is little if any evidence that the adverb-cluster constituent exists. The remnant movement analysis is then not only theoretically ill-founded, it is also empirically poorly supported. Just what the theory outlined here would lead us to expect.

(53)

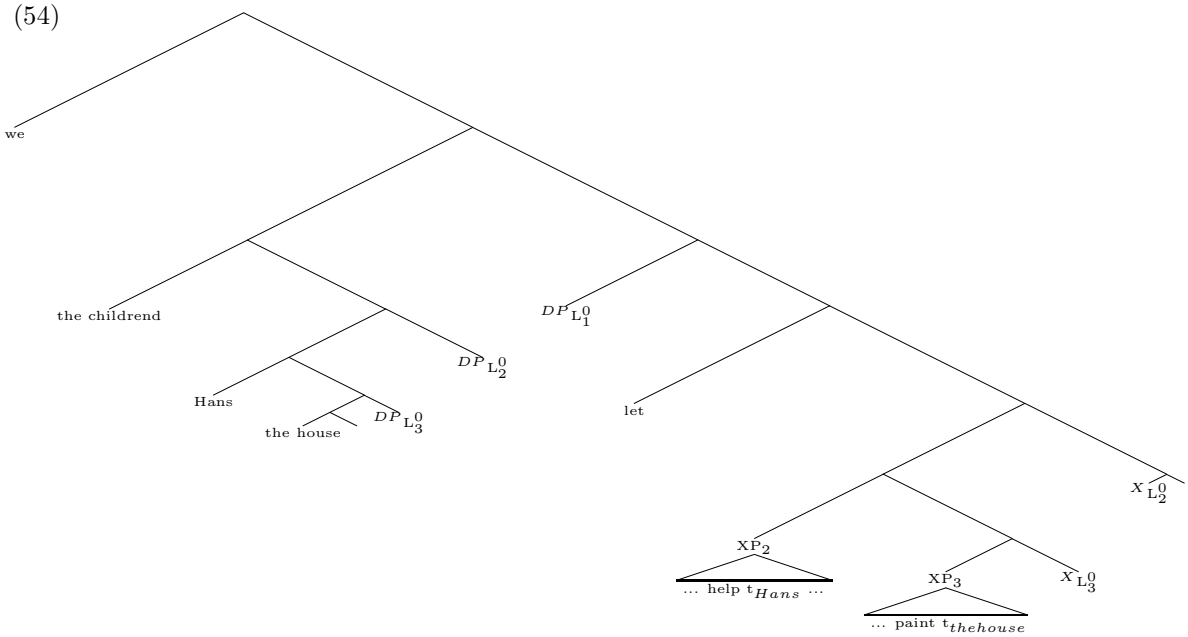


This is how the Nilsen-Bentzen system derives the cross-serial dependency between the adverbs and the verbs. The same type of analysis can be given for the Swiss German and Dutch cases, (54). All that is required for these languages are DP-lifters and XP-lifters. The latter of which bring about the clustering of the verbs. The unbounded nature of the phenomenon is captured here by recursion through fixed sequences of identical lifters.

The problems that arose under UCOOL and GenPIM for the analysis of Norwegian, arise again under the remnant movement analysis of Dutch and Swiss German. Movement of DP and of XP cannot be of the same type because of Müller's generalization. But they cannot be of different types either since then one of the remnant movements would sooner or later violate GenPIM. If anything, the problems for this analysis of the Dutch and Swiss German cross-serial dependencies is more severe than they were for the case of Norwegian. In Norwegian there is no reason to believe that the phenomenon is truly unbounded. This opens up a loophole. The analysis can be saved by assuming that the hierarchy of movements includes movement of VP_{ordnet}, F_{Adv_1}, VP_{blitt}, F_{Adv_2}, VP_{ha}, F_{Adv_3},...as separate types. They can then be hierarchically ordered as follows: $\mathrm{VP}_{ordnet} << \mathrm{F}_{Adv_1} << \mathrm{VP}_{blitt} << \mathrm{F}_{Adv_2} << \mathrm{VP}_{ha} << \mathrm{F}s_{Adv_3}\ldots$ This is an unappealing but possible fix. The same solution is not available in the case of Dutch and Swiss German cross-serial dependencies because of the recursive nature of the phenomenon. As we will see in a momement, the same problem plagues the analyses in Koopman and Szabolcsi (2000).

For reasons of space I cannot pursue the issue further.

(54)



Before turning to Koopmand and Szabocsi, though, let's consider the structures derived by the remnant movement account of cross-serial dependencies. A second look at them actually reveals two cross-serial dependencies each. In (53) the adverbs *lenger alltid helt* – 'any.longer always completely' and the verbs *kunne ha blitt* – 'could have become' form a straight cross-serial dependency; the abstract heads ${}^{Adv}\mathrm{L}^0_n$ and ${}^{V}\mathrm{L}^0_n$ – an inverse one. This inverse cross-serial dependency is formed as a by-product in the derivation. Similarly, in (54) the DPs *d'chind em Hans es huus* – 'the children Hans the house' are in a straight cross serial dependency with the verbs *laa hälfe aastriiche*, while the abstract lifters ${}^{DP}\mathrm{L}^0_n$ and ${}^{XP}\mathrm{L}^0_n$ enter into an inverse cross-serial dependency.

From the point of view of the Nilsen-Bentzen theory it is a coincidence that all heads involved in the inverse cross-serial dependency are abstract. In fact, I give a trivial adaptation below in section 5.2 that derives an overt inverse cross-serial dependency.

So far we had one reason to be suspicious of the remnant movement approach to cross-serial dependencies, namely, that it necessarily violated well-supported constraints on remnant movement that derive from UCOOL and GenPIM. Now we have a more directly empirical argument against this approach: it allows generating a type of the structures that is unattested. It is clearly not sufficiently restrictive since it fails to capture the left-right asymmetry regarding cross-serial dependencies.

I will digress for a moment to illustrate that the issue of inverse cross-serial dependencies as a derivational by-product is not limited to the Nilsen-Bentzen account of cross-serial orders in Norwegian alone. The same problem comes up under Koopman and Szabolcsi's (2000) theory of order preservation.

Koopman and Szabolcsi (2000) assume that while movement of specifiers can be order changing, the movement of complements is not because it is subject to the "move-one-category-at-a-time" constraint, (55a). Though the formulation is slightly cryptic, it is clear from context that Koopman and Szabolcsi intend (55a) to ensure that in a situation where XP contains YP as its complement, XP cannot move without previously evacuating YP. Since this is a general condition, if YP itself contains a complement ZP, then ZP will have to move prior to YP movement, etc. We need not concern ourselves with the exceptions in (55b) here.

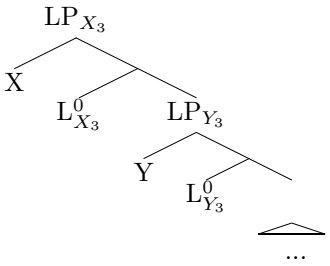
(55) Movement (Koopman and Szabolcsi, 2000, p. 43):

a. by default, move only one category at a time.

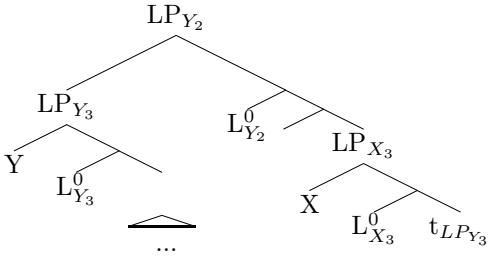
b. Exemption: Sequences of categories that may move together are (i) <YP+, YP>, where YP+ is a projection motivated by the modified LCA, (ii) sequences involving operator projections like RefP, DistP, F[oc]P, and the projection hosting the verb.

The move-one-category-at-a-time constraint entails, for example, that X and Y in (56) will not usually move by themselves but as part of their containing categories LP_X and LP_Y. These categories split under movement by (55a). In a first step, LP_{Y_3} would move, as in (57a); then LP_{X_3} moves, as in (57b).

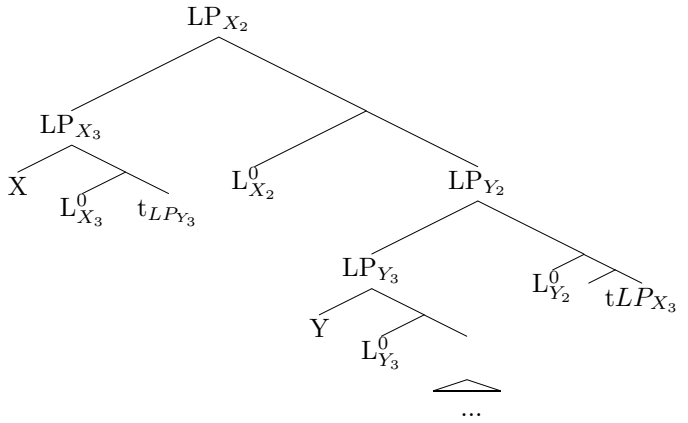
(56)



(57) a.

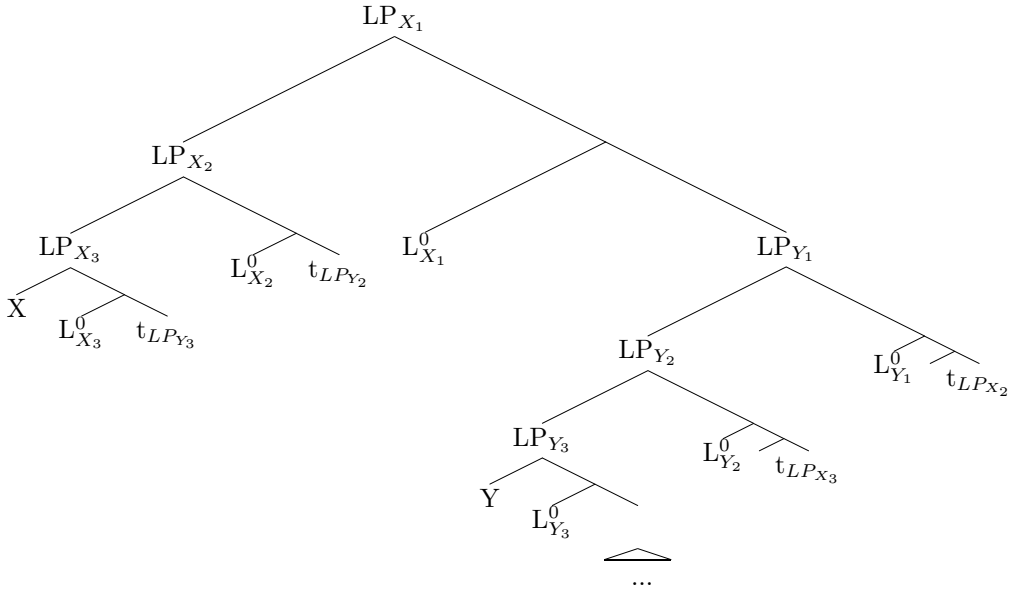


b.



After two further steps of movement, of which there are many in Koopman and Szabolcsi's model, the structure in (58) emerges.

(58)



What is important for the present discussion is that (58) contains an inverse cross-serial dependency between the $L^0_{X_n}$'s and the $L^0_{Y_n}$'s.[22] Again, these inverse cross-serial dependencies are simply a derivational by-product. There is no reason in the theory why they should be unpronounced. Theories that make unconstrained remnant movement available, thus fail to rule out inverse cross-serial dependencies even if they stick to the seemingly severe linear restrictions im-

[22]The careful reader will no doubt have noticed that the derivation currently under discussion also violates the conjunction of UCOOL and GenPIM for the same reason that the Nilsen-Bentzen derivations do.

posed by the LCA, as Koopman and Szabolcsi (2000); Nilsen (2003); Bentzen (2005) do.[23] This drives home the point made in the introduction: Without a restrictive theory of operations, the LCA loses its empirical bite.

I now turn to the question of whether there is a principled way of ruling out inverse cross-serial dependencies.

5 The Absence of Inverse Cross-Serial Dependencies

It should already be clear that in a system as the one entertained here, which disallows rightward movement, inverse cross-serial dependencies cannot be derived in a way totally symmetric to the derivation of straight cross-serial dependencies in section 4.1. The relevant movement operation that maps (44b) onto (44c) will necessarily line up the moving elements in an order where scope corresponds to left-to-right order and never the other way around. Two derivations are conceivable that would lead directly to an inverse cross-serial dependency: a nesting derivation and a rightward movement derivation. The former is disallowed under (46a), the latter – under (46c).

Let's see why the left-to-right order of *moved* elements always reflect scope even if *base-generated* specifiers and/or heads are allowed to be linearized to the right of their sisters. The reason lies in the structural rather than linear definition of order preservation, (46a) above, together with the assumption that movement is universally leftward. A simple linear reversal of the derivation for Swiss German sketched above in section 4.1 is therefore impossible. This is good news: the restriction to leftward movement induces an asymmetry.

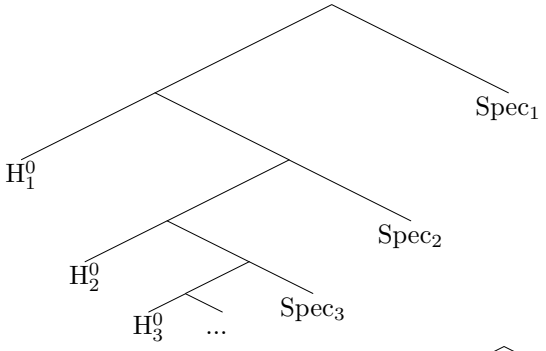
We need to see now how to block all remaining derivations for inverse unbounded cross-serial dependencies. Under the assumptions made so far, there are two candidates: a head-movement derivation and a remnant-movement derivation. I discuss them in turn.

5.1 Head Movement

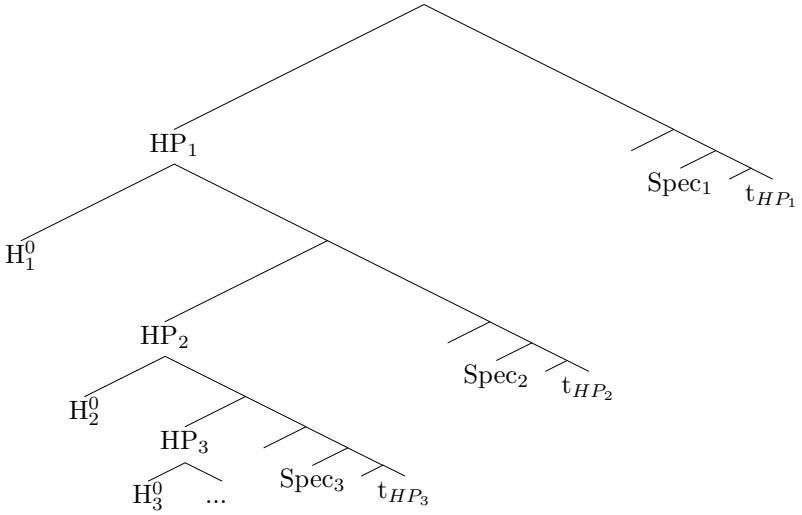
Head-to-head movement is problematic. Consider the structure in (59). This is a structure with a sequence of initial heads with their specifiers to the right. If we assume that the ban against rightward movement is the only linear asymmetry in syntax (see Abels and Neeleman (2006) for this view), then structures like (59a) can be base generated. If we adopt the LCA, then right specifiers are ruled out and must be derived through roll-up movement as in (59b) (see (2b) above). To make the structure in (59b) more easily readable, functional heads that are required by Kayne's LCA but irrelevant to the point under discussion are indicated as empty branches.

[23] As far as I can tell, the LFG account of the Dutch cross-serial dependencies in Bresnan et al. (1982) suffers from the same problem: there is no reason to expect a linear asymmetry in the realm of cross-serial dependencies. The same is true of Steedman's 2000 account in terms of type-lifting.

(59) a.

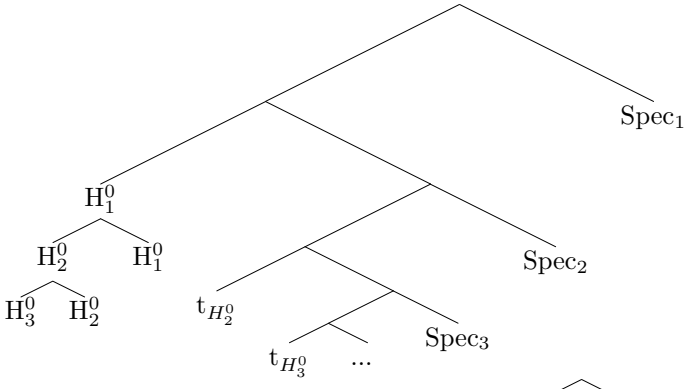


b.

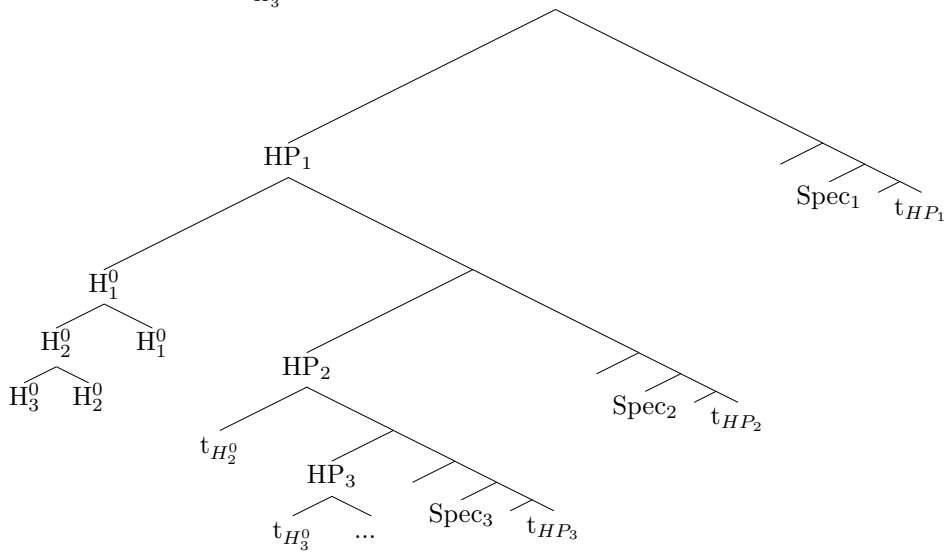


Standard (left adjoining) head movement derives (60a) with an inverse cross-serial dependency from (59a). Whether (60b) is available as a continuation of (59b) depends on whether head movement from specifiers is allowed or not, certainly the derivation is compatible with the LCA per se. Some restriction on movement in addition to the LCA must be invoked. This is, of course, just another instance of the more general lesson: the LCA has empirical teeth only in conjunction with a restrictive theory of movement.

(60) a.



b.



Since inverse cross-serial dependencies are universally unattested, the derivations in (60) must be ruled out. These facts might then be interpreted as a new empirical argument against head movement (see Hinterhölzl (1997); Brody (2000); Koopman and Szabolcsi (2000); Boeckx and Stjepanovič (2001); Chomsky (2000, 2001a); Williams (2002, 2004); Stabler (2003); Abels (2003); Nilsen (2003); Müller (2004); Lechner (2005); Matushansky (2006) for recent discussion of head movement from various angles). Certainly, in a system that bans rightward movement but allows base-linearization to the left and to the right, head movement must be banned. The easiest way of banning head movement in the usual sense is to demand that a moved item must strictly c-command its trace.

5.2 Remnant Movement

As discussed in the introduction to this section, the ban against rightward movement rules out derivations for inverse cross-serial dependencies which are fully symmetrical to (44). However,

we have also seen in section 1 that any rightward movement analysis can be mechanically re-done in terms of remnant movement, (2c). It is therefore no surprise that remnant movement may lead to inverse cross-serial dependencies, as was illustrated in section 4.2.[24]

In the discussion of Nilsen's and Koopman and Szabolcsi's theories I claimed that it is a pure coincidence under their assumptions that audible inverse cross-serial dependencies are unattested. I will illustrate this claim by giving a trivial reformulation of the Nilsen-Bentzen account of cross-serial orders in Norwegian which leads to inverse crossed orders.

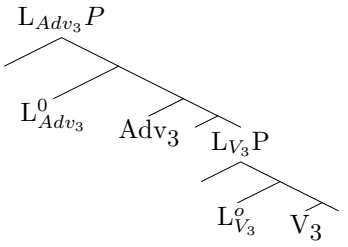
(61) Inverse Nilsen/Bentzen

a. Above every verbal element, there is a lifter head L^0_V, which lifts the closest $\mathrm{L}_V\mathrm{P}$ – if present – into its specifier.

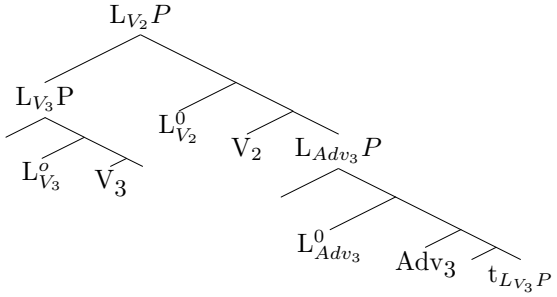
b. Above every adverbial element, there is a lifter head L^0_{Adv}, which lifts the closest $\mathrm{L}_{Adv}\mathrm{P}$ – if present – into its specifier.

These assumptions give rise to derivations like those depicted in (62a-c). Removing unnecessary clutter from (62c) yields (62d), where we see the beginning of an overt inverse cross-serial dependency emerging. The cycle of operations could be repeated; it easily generates unbounded inverse cross-serial dependencies.[25]

(62) a.



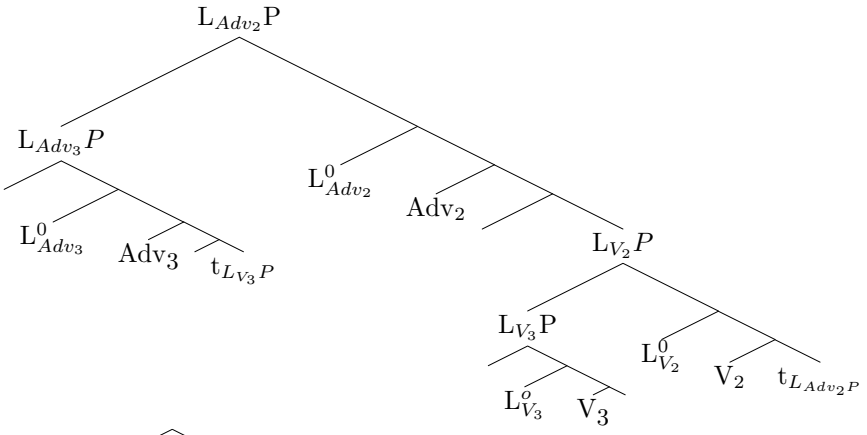
b.



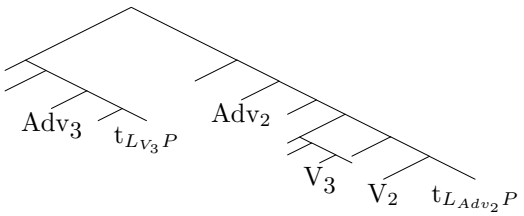
[24] The similarity between head movement and remnant movement with regard to cross-serial dependencies should come as no surprise to those who model head movement as remnant movement.

[25] Although I will not go through this in detail, Koopman and Szabolcsi's system can equally easily be modified to yield inverse cross-serial dependencies. A language that chooses some of the movement options assumed for Hungarian and some of the ones assumed for Dutch, that is, a cross between the two languages, will achieve exactly this. For lack of space, I cannot pursue the issue here.

c.



d.



A theory that predicts the ban against inverse cross-serial dependencies must rule out such derivations. Fortunately we are already in a position to do this. Movement of $L_V P$ and movement of $L_{Adv} P$ cannot be the same type of movement, otherwise the step creating (62c) would violate Müller's generalization. Thus, to enable the step from (62a) to (62b), $\text{Move}_{L_V P} << \text{Move}_{L_{Adv} P}$. This ordering is required to enable the derivation of (62b) from (62a) and of (62c) from (62b), but it rules out the continuation in (62d).

5.3 Interim Conclusion

If the empirical hypothesis that UG disallows inverse cross-serial dependencies is true, then syntactic theory must account for this. With assumptions (63-i-iii) in place, we can rule out all derivations for inverse cross-serial dependencies except for the remnant movement derivation.

(63) (i) Moved elements strictly c-command their traces.
(ii) Movement is uniformly leftward.
(iii) Movement operations are order preserving in the sense of (45).

If we add to these assumptions the assumption that movement operations are ordered, that is, if we assume UCOOL, and if we deploy UCOOL as in GenPIM, (9), then inverse cross-serial dependencies are correctly predicted to be impossible.

6 Conclusion

We have completed a journey. We started with the question how existing linear asymmetries in language can be explained and made the trivial observation that constraints on structures (like the Cinque hierarchy and the LCA) might be necessary to explain such asymmetries, but they are insufficient. We noticed in particular that giving up the Proper Binding Condition and allowing remnant movement into the theory runs the risk of robbing the LCA of its empirical bite. A restrictive theory of remnant movement must accompany any theory, LCA or otherwise, which aims at deriving linear asymmetries in language from a ban against rightward movement.

Section 2 represents a step in this direction. It is standardly assumed that operations in syntax are ordered and formulated this assumption as UCOOL. UCOOL is meaningless without an instruction of how to deploy it. We argued that an assumptions at least as strong as GenPIM, (9), repeated for convenience in (64), is necessary and yields a fine-grained and, as far as we could tell, correct framework for the analysis of phenomena in various domains: improper movement, effects of the freezing principle and exceptions from it, and remnant movement.

(64) Generalized Prohibition against Improper Movement (GenPIM)
No constituent may undergo movement of type τ if it has been affected by movement of type σ, where $\tau << \sigma$ under UCOOL.

We can synthesize the discussion of feeding and bleeding relations between operations in German from section 2 in the slightly more articulate hierarchy below.

(65) Universal Constraint on Operational Ordering in Language (UCOOL–revised)
$\theta << Case << scrambling << wh-movement << topicalization$

The near perfect match between the theory and the predictions is an encouraging direct result. It suggests that a whole range of important issues that were set aside in section 2 warrant further study: Is UCOOL a linear total order? How can it be integrated into a derivational theory of syntax? Is the asymmetry in the formulation of GenPIM an eliminable feature of the theory? What is the relation between UCOOL and the Cinque Hierarchy? What is the exact set and order of movement types referenced in UCOOL? ...

In section 3 we reviewed some well-known facts regarding cross-serial dependencies in Dutch and Swiss German and claimed that there exists a hitherto unnoticed linear asymmetry in this empirical domain: Inverse cross-serial dependencies are impossible, (43). We suggested a novel, restrictive analysis of forward cross-serial dependencies, 4.1, and then went on to show (section 4.2 and 5) that theories which derive cross-serial dependencies through remnant movement fail both theoretically and empirically. Theoretically, because they clash directly with the framework of UCOOL and GenPIM. Empirically, because they fail to predict the universal ban against backwards cross-serial dependencies. It now remains to be seen whether insights gained in analyses that make unrestricted use of remnant movement (such as Haegeman (2000); Hinterhölzl (1997)) can be captured under a restrictive theory of syntax, such as the one proposed here.

To finish, let me mention just two recalcitrant problems on which the framework of assumptions adopted here and including crucially UCOOL and GenPIM offers new perspectives.

First, den Besten and Webelhuth (1990) observe a curious restriction on remnant topicaliza-

tion in German, which they formulate as a ban against moving remnants that contain traces in non-argument positions. Subsequent research disagreed with the original formulation and converged on a formulation of the restriction in terms of intermediate traces: moved remnants may not contain unbound intermediate traces of successive cyclic movement (see Müller (1998, ch 6) for references, discussion, and some potential counterexamples). Assuming successive cyclicity and a remnant movement approach to apparent rightward movement, the right roof constraint and the constraint against movement of remnants containing unbound intermediate traces can be identified. To the best of my knowledge, the identity between the right-roof constraint and the ban against intermediate traces in remnants has not previously been observed. This type of problem clearly falls within the scope of the theory developed here. It is an open question though how GenPIM can be sharpened so as to yield the desired result. The alternative formulation of GenPIM based on Sternefeld (1992) contemplated above example (31), which we discarded there because of its over-restrictiveness, achieves almost that. It remains to be seen how the problems can be overcome. Alternatively, but less interestingly, the ban against unbound intermediate traces and the right roof constraint might simply follow from the details of the hierarchy under UCOOL. In this latter case the right-roof constraint and the ban against unbound intermediate traces would be facts about UCOOL, not about the way the grammar makes use of UCOOL, that is, GenPIM.

Second there are cases of ECM under *wh*-movement (see Rizzi (1982); Kayne (1984); Pesetsky (1992); Bošković (1997)). A fact that has proved difficult to capture in a principled way is that this kind of ECM must happen at the first step of *wh*-movement, that is, the items that undergo this restricted kind of ECM must be the subjects of the clause immediately embedded under the ECM verb. ECM does not happen at the intermediate landing site of successive cyclic *wh*-movement. This problem becomes tractable here if we assume that successive cyclic steps of an operation of type τ count as operations of type τ. A subject that has a θ-role but no Case may at the same time undergo movement for Case and the first step of successive cyclic movement without violating UCOOL and GenPIM. Although Case operations and *wh*-operations are ordered by UCOOL, if they happen simultaneously rather than one after the other, no violation of GenPIM ensues. However, after a single step of successive cyclic *wh*-movement, a Case-operation is no longer possible without violating GenPIM.

References

Abels, Klaus. 2002. On an alleged argument for the proper binding condition. In *Proceedings of humit 2001. mit working papers in linguistics 43*, ed. Tania R. Ionin, Heejeong Ko, and Andrew Nevins, 1–16. Cambridge, MA.

Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Doctoral dissertation, University of Connecticut.

Abels, Klaus, and Ad Neeleman. 2006. Universal 20 without the LCA. ms.

Adger, David. 2003. *Core syntax: A minimalist approach*. Oxford: Oxford University Press.

Alexiadou, Artemis. 1997. *Adverb placement: A case study in antisymmetric syntax*, volume 18 of *Linguistik Aktuell*. Amsterdam, Philadelphia: John Benjamins.

Beck, Sigrid. 1996. Quantified structures as barriers for lf movement. *Natural Language Semantics* 4:1–56.

Belletti, Adriana, ed. 2004. *Structures and beyond*, volume 3 of *Oxford studies in comparative syntax The cartography of syntactic structures*. Oxford: Oxford University Press.

Bentzen, Kristine. 2005. What's the better move? - on verb placement in Standard and Northern Norwegian. *Nordic Journal of Linguistics* 28:153–188.

den Besten, Hans, and Gert Webelhuth. 1990. Stranding. In *Scrambling and barriers*, ed. Günther Grewendorf and Wolfgang Sternefeld, 77–92. Amsterdam and Philadelphia: Academic Press.

Bobaljik, Jonathan D. 1999. Adverbs: the hierarchy paradox. *GLOT International* 4:27–28.

Bobaljik, Jonathan D. 2004. Clustering theories. In *Verb clusters: A study of Hungarian and Dutch*, ed. Katalin e Kiss and Henk van Riemsdijk, 121–145. Philadelphia: John Benjamins.

Boeckx, Cedric, and Norbert Hornstein. 2003. Reply to "Control is not movement". *Linguistic Inquiry* 34:269–280.

Boeckx, Cedric, and Sandra Stjepanovič. 2001. Heading toward PF. *Linguistic Inquiry* 31:345–355.

Bošković, Željko. 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge, Mass.: MIT Press.

Bošković, Željko. 1994. D-structure, theta-criterion, and movement into theta-positions. *Linguistic Analysis* 24:247–286.

Bošković, Željko. 1999. On multiple feature checking: Multiple *Wh*-fronting and multiple head movement. In *Working minimalism*, ed. Samuel David Epstein and Norbert Hornstein, 159–188. MIT Press.

Bošković, Željko. 2002. A-movement and the EPP. *Syntax* 5:167–218.

Bresnan, Joan W. 1970. On complementizers: Towards a syntactic theory of complement types. *Foundations of Language* 6:297–321.

Bresnan, Joan W., R. Kaplan, Stanley Peters, and Annie Zaenen. 1982. Cross-serial dependencies in Dutch. *Linguistic Inquiry* 13:613–635.

Brody, Michael. 2000. Mirror theory. syntactic representation in perfect syntax. *Linguistic Inquiry* 31:29–57.

Chomsky, Noam. 1986. *Barriers*, volume 13 of *Linguistic Inquiry Monograph*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from building 20*, ed. Kenneth Hale and Samuel J. Keyser, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995a. Bare phrase structure. In *Government and binding theory and the minimalist program*, ed. Gert Webelhuth, 383–439. Oxford: Basil Blackwell.

Chomsky, Noam. 1995b. Categories and transformations. In *The minimalist program*, 219–394. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalism in honor of howard lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001a. Beyond explanatory adequacy.

Chomsky, Noam. 2001b. Derivation by phase. In *Ken hale: a life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA.: MIT Press.

Cinque, Guglielmo. 1996. The antisymmetric programme: theoretical and typological implications. *Journal of Linguistics* 32:447–464.

Cinque, Guglielmo. 1999. *Adverbs and functional heads - a cross-linguistic perspective*. New York and Oxford: Oxford University Press.

Cinque, Guglielmo, ed. 2002. *Functional structure in DP and IP*. Oxford Studies in Comparative Syntax. Oxford University Press.

Cinque, Guglielmo. 2005. Deriving greenberg's universal 20 and its exceptions. *Linguistic Inquiry* 36:315–332.

Cinque, Guglielmo. 2006. *Restructuring and functional heads*. Oxford studies in comparative syntax: The cartography of syntactic structures. Oxford and New York: Oxford University Press.

Collins, Chris. 1994. Economy of derivation and the generalized proper binding condition. *Linguistic Inquiry* 25:45–61.

Collins, Chris. 2005. A smuggling approach to raising in English. *Linguistic Inquiry* 36:289–298.

Culicover, Peter W., and Ray Jackendoff. 2001. Control is not movement. *Linguistic Inquiry* 32:493–511.

Culy, Christopher. 1985. The complexity of the vocabulary of Bambara. *Linguistics and Philosophy* 8:345–351.

Davies, William D., and Stanley Dubinsky. 2004. *The grammar of raising and control*. Blackwell.

Dehdari, Jonathan. 2006. Crossing dependencies in Persian. MA-thesis, Brigham Young University.

Diesing, Molly. 1992. *Indefinites*. Cambridge, MA.: MIT Press.

den Dikken, Marcel. 1995. Extraposition as intraposition, and the syntax of english tag questions.

Evers, Arnold. 1975. The transformational cycle in dutch and german. Doctoral dissertation, University of Utrecht.

Fanselow, Gisbert. 1990. Scrambling as NP-movement. In *Scrambling and barriers*, ed. Günther Grewendorf and Wolfgang Sternefeld, 113–142. John Benjamins.

Fanselow, Gisbert. 1993. The return of the base generators. *Groninger Arbeiten zur Germanistischen Linguistik* 36:1–74.

Fanselow, Gisbert. 2001. Features, θ-roles, and free constituent order. *Linguistic Inquiry* 32:405–437.

Fanselow, Gisbert. 2004. The MLC and derivational economy. In *Minimality effects in syntax*, ed. Arthur Stepanov, Gisbert Fanselow, and Ralf Vogel, volume 70 of *Studies in Generative Grammar*, 73–124. deGruyter.

Fiengo, Robert. 1977. On trace theory. *Linguistic Inquiry* 8:35–62.

Fox, Danny, and David Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31:1–46.

Greenberg, Joseph. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of language*, ed. Joseph Greenberg, 73–113. Cambridge, Mass.: MIT Press.

Grewendorf, Günther. 2001. Multiple *wh*-fronting. *Linguistic Inquiry* 32:87–122.

Grewendorf, Günther. 2003. Improper remnant movement. *Gengo Kenkyu* 123:47–94.

Haegeman, Liliane. 2000. Remnant movement and ov order. In *The derivation of vo and ov*, ed. Peter Svenonius, volume 30 of *Linguistik Aktuell*, 69–96. Amsterdam and Philadelphia: John Benjamins.

Haegeman, Liliane, and Henk van Riemsdijk. 1986. Verb projection raising, scope and the typology of rules affecting verbs. *Linguistic Inquiry* 17:417–466.

Haider, Hubert. 1993. *Deutsche syntax- generativ: Vorstudien yur theorie einer projektiven grammatik*. Tübingen: Gunter Narr Verlag.

Hale, John, and Géraldine Legendre. 2004. Minimal links, remnant movement, and (non-)derivational grammar. In *Minimality effects in syntax*, ed. Arthur Stepanov, Gisbert Fanselow, and Ralf Vogel, volume 70 of *Studies in Generative Grammar*, 177–204. deGruyter.

Heycock, Caroline. to appear. Japanese *-wa*, *-ga*, and information structure.

Higginbotham, James. 1984. English is not a context-free language. *Linguistic Inquiry* 15:225–234.

Hinterhölzl, Roland. 1997. A VO-based approach to verb raising. In *Proceedings of NELS, McGill university*, ed. K Kusumoto, 187–202. GLSA.

Hiraiwa, Ken. 2002. Movement and derivation: Eliminating the PBC. In *Penn Linguistics Colloquium 26*. Philadelphia, PA.

Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30:69–96.

Huybregts, Riny. 1976. Overlapping dependencies in dutch. *Utrecht Working Papers in Linguistics* 1:24–65.

Huybregts, Riny. 1984. The weak inadequacy of Context-Free Phrase Structure Grammars. In *Van periferie naar kern*, ed. G de Haan, M Trommelen, and W Zonneveld, 81–99. Dordrecht: Foris.

Jónsson, J. G. 1996. Clausal architecture and case in Icelandic. Ph.d. dissertation, University of Massachusetts, Amherst.

Kayne, Richard. 1984. *Connectedness and binary branching*. Dordrecht: Foris.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA.: MIT Press.

Kayne, Richard. 1998. Overt vs. covert movement. *Syntax* 1:128–191.

Kayne, Richard. 1999. Prepositional complementizers as attractors. *Probus* 11:39–73.

Kayne, Richard. 2004. Prepositions as probes. In *Structures and beyond: The cartography of syntactic structures, volume 3*, ed. Adriana Belletti, 192–212. Oxford University Press.

Kitahara, Hisatsugu. 1997. *Elementary operations and optimal derivations*. Cambridge, MA: MIT Press.

Koopman, Hilda. 1996. The spec head configuration. In *Syntax at sunset*, ed. E Garrett and F Lee, 37–64. UCLA Working Papers in Linguistics.

Koopman, Hilda, and Anna Szabolcsi. 2000. *Verbal complexes*. Cambridge, MA: MIT Press.

Kracht, Marcus. 2003. *The mathematics of language*. Studies in Generative Grammar 63. Berlin, New York: Mouton de Gruyter.

Kracht, Marcus. 2005. The emergence of syntactic structure.

Kuno, Susumu, and Jane J. Robinson. 1972. Multiple wh-questions. *Linguistic Inquiry* 3:463–487.

Kuroda, S.-Y. 1965. Generative grammatical studies in the japanese language. Doctoral dissertation, MIT.

Kuthy, K. de, and Walt Detmar Meurers. 1999. On partial constituent fronting in German. In *Tübingen studies in head-driven phrase structure grammar*, ed. V. Kordoni, 22–73.

Landau, Idan. 2003. Movement out of control. *Linguistic Inquiry* 34:471–498.

Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19:335–391.

Lasnik, Howard. 1998. Exceptional case marking: Perspectives old and new. In *Formal Approaches to Slavic Linguistics (FASL) 6: The Connecticut Meeting*, ed. Zeljko Boskovic, Steven Franks, and William Snyder, 187–211. Michigan Slavic Publications.

Lasnik, Howard, and Mamoru Saito. 1991. On the subject of infinitives. In *Papers from the 27th regional Meeting of the Chicago Linguistic Society Part One: The General Session*, ed. Lise M. Dobrin, Lynn Nichols, and Rosa M. Rodriguez, 324–343. Chicago, Ill.: Chicago Linguistic Society, University of Chicago.

Lasnik, Howard, and Mamoru Saito. 1992. *Move α: Conditions on its application and output*. Cambridge, MA: MIT Press.

Lechner, Winfried. 2005. Interpretive effects of head movement, July 2005.

Maclachlan, Anna, and Owen Rambow. 2002. Cross-serial dependencies in Tagalog. In *Proceedings of the Sixth International Workshop on Tree Adjoining Grammar and Related Frameworks (TAG+6)*, 100–104.

Matushansky, Ora. 2006. Head-movement in linguistic theory. *Linguistic Inquiry* 37:69–109.

Müller, Gereon. 1998. *Incomplete category fronting: A derivational approach to remnant movement in German*. Dordrecht, Boston: Kluwer Academic Publishers.

Müller, Gereon. 2001. Order preservation, parallel movement, and the emergence of the unmarked. In *Optimality theoretic syntax*, ed. Geraldine Legendre, Jane Grimshaw, and Sten Vikner, 279–313. MIT Press.

Müller, Gereon. 2004. Verb-second as vp-first. *Journal of Comparative Germanic Syntax* 7:179–234.

Müller, Gereon, and Wolfgang Sternefeld. 1993. Improper movement and unambiguous binding. *Linguistic Inquiry* 24:461–507.

Müller, Gereon, and Wolfgang Sternefeld. 1996. A'-chain formation and economy of derivation. *Linguistic Inquiry* 27:480–511.

Muriungi, Peter. 2005. Wh-questions in Kitharaka. *Studies in African Linguistics* 34:43–104.

Nevins, Andrew Ira. 2004. Derivations without the activity condition. *MIT Working Papers in Linguistics* 49:287–310.

Nevins, Andrew Ira, and Norbert Hornstein. 2005. **Representation Theory** by Edwin Williams. *Language* 81:757–761.

Nilsen, Øystein. 2003. *Eliminating positions: Syntax and semantics of sentence modification*. Utrecht, NL: LOT.

Nilsen, Øystein. 2004. Domains for adverbs. *Lingua* 114:809–847.

Perlmutter, David M, and Paul M Postal. 1983/1972. The relational succession law. In *Studies in relational grammar 1*, ed. David M Perlmutter, 30–80. Chicago: University of Chicago Press.

Pesetsky, David. 1982. Paths and categories. Doctoral dissertation, MIT.

Pesetsky, David. 1992. Zero syntax vol. 2.

Pesetsky, David. 2000. *Phrasal movement and its kin*. Linguistic Inquiry Monograph 37. Cambridge, MA: MIT Press.

Polinsky, Maria, and Eric Potsdam. 2002. Backward control. *Linguistic Inquiry* 33:245–282.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of ip. *Linguistic Inquiry* 20:365–424.

Postal, Paul M. 1974. *On raising: One rule of english grammar and its theoretical implications*. Cambridge, MA: MIT Press.

Pullum, Geoffrey K., and Gerald Gazdar. 1982. Natural languages and context-free languages. *Linguistics and Philosophy* 4:471–504.

Richards, Marc David. 2004. Object shift and scrambling in North and West germanic. Ph.d. thesis, University of Cambrdige.

Richards, Norvin W. III. 1997. What moves where when in which language? Ph.d. dissertation, MIT.

Rizzi, Luigi. 1990. *Relativized minimality*, volume 16 of *Linguistic Inquiry Monograph*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar: Handbook in generative syntax*, ed. Liliane Haegeman, 281–337. Dordrecht, Netherlands: Kluwer.

Rizzi, Luigi. 2004a. On the form of chains: Criterial positions and ECP effects, January 2004.

Rizzi, Luigi, ed. 2004b. *The structure of CP and IP - the cartography of syntactic structures, volum 2*, volume 2 of *Oxford studies in comparative syntax The cartography of syntactic structures*. Oxford: Oxford University Press.

Rizzi, Luizi. 1982. *Issues in italian syntax*. Dordrecht: Foris.

Rochemont, Michael S., and Peter W. Culicover. 1997. Deriving dependent right adjuncts in english. In *Rightward movement*, ed. Dorothee Beerman, David Leblanc, and Henk van Riemsdijk, volume 17 of *Linguistik Aktuell*, 279–300. John Benjamins.

Rögnvaldsson, E. 1987. OV word order in Icelandic. In *Proceedings of the Seventh Biennial Conference of Teachers of Scandinavian in Great Britain and Northern Ireland*, 33–49. University College London.

Rosenbaum, Peter S. 1967. *The grammar of English predicate complement constructions*. Cambridge, Mass: MIT Press.

Rosenbaum, Peter S. 1970. A principle governing deletion in English sentential complementation. In *Readings in English transformational grammar*, ed. Roderick Jacobs and Peter Rosenbaum, 20–29. Ginn: Waltham, Mass.

Ross, John Robert. 1967/1986. *Infinite syntax*. Norwood: Ablex Publishing.

Sakai, Hiromu. 1994. Derivational economy in long distance scrambling. In *Formal approaches to Japanese linguistics 1*, ed. Masatoshi Kozumi and Hiroyuki Ura, volume 24 of *MIT Working Papers in Linguistics*, 295–314. MITWPL.

Sauerland, Uli. 1996. The interpretability of scrambling. In *Formal approaches to Japanese linguistics 2*, ed. Masatoshi Koizumi, Masayuki Oishi, and Uli Sauerland, volume 29 of *MIT Working Papers in Linguistics*, 213–234. MITWPL.

Savitch, Walter J., Emmon Bach, William Marsh, and Gila Sfran-Naveh, ed. 1987. *The formal complexity of natural language*. Dordrecht: Reidl.

Shieber, Stuart. 1985. Evidence against the context-freeness of natural languages. *Linguistics and Philosophy* 8:333–343.

Stabler, Edward P. 2003. Comparing 3 perspectives on head movement. In *Head movement and syntactic theory*, ed. Anoop Mahajan. UCLA/Potsdam Working Papers in Linguistics.

Stabler, Edward P. 2004. Varieties of crossing dependencies: structure dependence and mild context sensitivity. *Cognitive Science* 28:699–720.

Starke, Michal. 2001. Move dissolves into merge: a theory of locality. Doctoral dissertation, University of Geneva.

Starke, Michal. 2004. On the inexistence of specifiers and the nature of heads. In *Structures and beyond: The cartography of syntactic structures*, ed. Adriana Belletti, chapter 8. Oxford University Press.

von Stechow, Arnim, and Wolfgang Sternefeld. 1988. *Bausteine syntaktischen wissens: ein lehrbuch der generativen grammatik*. Opladen: Westdeutscher Verlag.

Steedman, Mark. 2000. *The syntactic process*. MIT Press.

Sternefeld, Wolfgang. 1992. Transformationstyplogie und strukturelle Hierarchie.

Uriagereka, Juan. 1999. Multiple spell-out. In *Working minimalism*, ed. Norbert Hornstein and Samuel D. Epstein, 251–282. Cambridge, MA: MIT Press.

Wagner, Michael. 2005. Review of verb clusters: A study of Hungarian, German and Dutch. *Linguistlist* 16.

Wexler, Kenneth, and Peter Culicover. 1980. *Formal principles of language acquisition*. Cambridge, MA: MIT Press.

Wilder, Chris. 1998. Transparent free relatives. In *Proceedings of wccfl xiii*.

Williams, Edwin. 2002. *Representation theory*. Cambridge, Mass.: MIT Press.

Williams, Edwin. 2004. The structure of clusters. In *Verb clusters*, ed. Katalin e Kiss and Henk van Riemsdijk, 173–201. John Benjamins.

Williams, Edwin S. 1974. Rule ordering in syntax. Doctoral Dissertation, MIT.

Williams, Edwin S. to appear. Subjects of different heights. In *Proceedings of FASL 14 – The Princeton meeting*, ed. Hana Filip, Steven Franks, James E. Lavine, and Mila Tasseva-Kurktchieve. Slavica.

Wiltschko, Martina. 1997. D-linking, scrambling and superiority in German. *Groninger Arbeiten zu germanistischen Linguistik* 41:107–142.

Wiltschko, Martina. 1998. Superiority in German. In *Proceedings of the sixteenth West coast conference on formal linguistics*, ed. E. Curtis, J. Lyle, and G. Webster, 431–445. Stanford, CA: CSLI Publication.

Wurmbrand, Susanne. 2006. Verb clusters, verb raising, and restructuring. In *The blackwell companion to syntax*, ed. Martin Everaert and Henk van Riemsdijk, volume 5, chapter 75, 229–343. Blackwell.