Towards a Restrictive Theory of (Remnant) Movement[*]

Klaus Abels

klaus.abels@gmx.net

Universitetet i Tromsø (CASTL) and University College London

Version 2.2
May 22, 2007

Contents

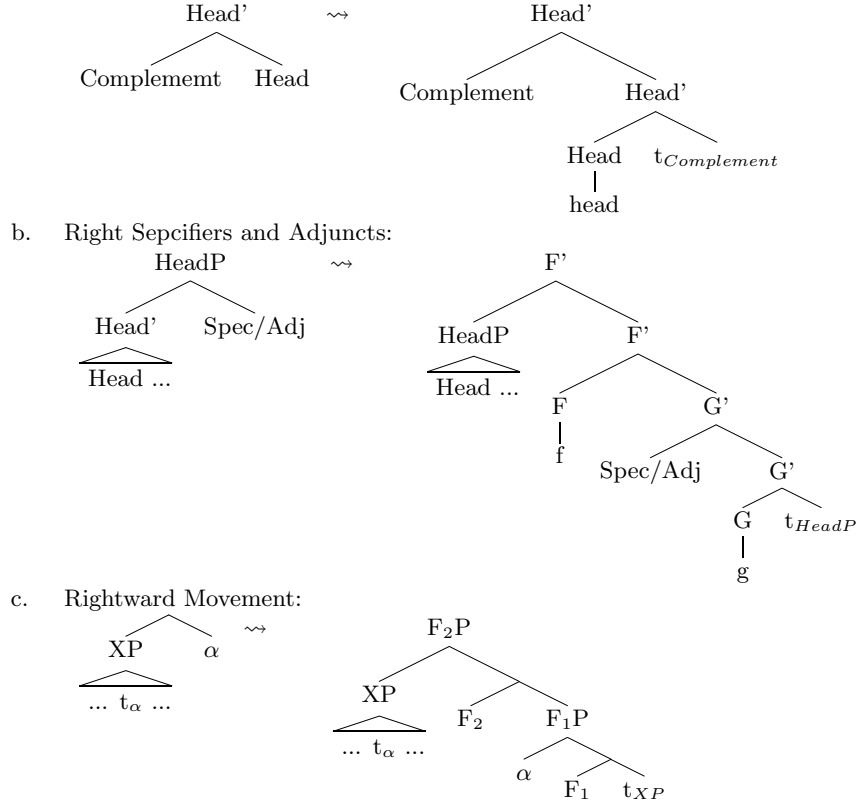# 1 Introduction: A plea for a restrictive theory of operations

Syntactic theory should explain linear asymmetries in language. Why is rightward movement so much more restricted than leftward movement (Ross (1967), Perlmutter and Postal (1983/1972), Bach (1971))? Why aren't OV-languages simply the mirror image of VO-langues (see Kayne (1994, 2005) 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). The LCA claims that all phrases in natural language follow the template in (1). I will refer to this as the rigid specifier-head-complement (S-H-C) hypothesis. The LCA also claims that all movement operations are leftward.

(1)
```
            XP
          /    \
       YP        XP
              /     \
           X⁰        ZP
           |
           x
```
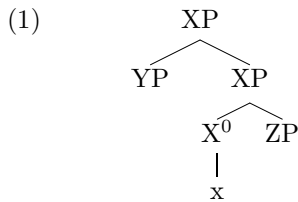
Obviously, the rigid S-H-C hypothesis, if true, can be no more than a partial answer to the questions raised in the first paragraph. No matter how restricted the theory of phrase structure is and no matter how serious the restriction to leftward movement may seem, an unduly permissive theory of movement can still undermine the explanatory effort expressed in the rigid S-H-C hypothesis. All of the structures that cannot be directly generated under the rigid S-H-C hypothesis can be brought back through movement operations, (2). The symbol ⤳ can be read as follows: Any analysis which is incompatible with the rigid S-H-C hypothesis because it contains the structure on the left can be made 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 be emulated through short movement operations, (2a), right specifiers and adjuncts can be simulated 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:

Head' ⤳ Head'

Complememt Head   Complement   Head'

                            Head   t$_{Complement}$

                            head

b.   Right Sepcifiers and Adjuncts:

HeadP ⤳   F'

Head'   Spec/Adj   HeadP   F'

Head ...   Head ...

                              F   G'

                              f   Spec/Adj   G'

                                          G   t$_{HeadP}$

                                          g

c.   Rightward Movement:

XP   α   ⤳   F$_2$P

... t$_\alpha$ ...   XP

                    ... t$_\alpha$ ...   F$_2$ F$_1$P

                                    α   F$_1$ t$_{XP}$

Excessive transformational power allows reformulating any analysis not conforming to the rigid S-H-C hypothesis in terms of it. Therefore, 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.[1]

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 movement in the current literature removes the empirical content 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 movement violates otherwise valid generalizations concerning pied-piping, stranding, and anti-locality. They conclude that generating complements to the

---

[1]In addition to a restrictive theory of the architecture of individual 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, chapter 1) 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.

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. Their claim is that a theory that dispenses with the rigid S-H-C hypothesis can reduce the amount of movement that needs to be assumed not only quantitatively but qualitatively. 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. However, nothing in this paper hinges on this choice and adherents of the rigid S-H-C hypothesis may mechanically transform all structures which are incompatible with it into compatible ones by assuming the required functional and agreement heads as indicated in (2a) and (2b). In other words, this paper does not bear directly on the question of base structures but is about the shape of a restrictive theory of movement.[2] I will assume throughout that rightward movement is banned (for an attempt at deriving this independently of the LCA see Abels and Neeleman (2006)). This ban might still have empirical and explanatory content. All depends on the theory of movement it is embedded in.

As we just saw, too permissive a theory of movement makes the strict S-H-C hypothesis empirically contentless. A theory of movement that was restrictive in certain ways was in the background when Kayne's book was first published, hence, the rigid S-H-C hypothesis was viewed as a great explanatory success.

In the early and mid nineties the proper binding condition was still widely assumed, remnant movement (as in (2c)) was considered exotic or impossible, and Müller (1998) felt compelled to spend an entire chapter on reconciling proper binding condition-effects with the existence of remnant movement. The proper binding condition – 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 proper binding condition 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 proper binding condition as background, the ban against rightward movement predicts that derivations like (2c) are unavailable, that is, there should be a complete absence of rightward movement even superficially.

A second restriction on movement that was widely assumed was the freezing principle. The freezing principle (Ross (1967); Wexler and Culicover (1980)) prohibits movement from a moved constituent. Given the VP internal subject hypothesis, the freezing principle predicts that subjects moved to [Spec,TP] should be islands for extraction – some counterexamples are discussed below in section (22). Together with the LCA the freezing principle also appears to rule out all extraction from preverbal obejcts in OV languages, (3), because all such extraction involves extraction from a moved constituent.

(3)   a.   X $\ldots [_{DP} \ldots t_X \ldots] \ldots$ V Aux $[\ t_V \ t_{DP}]$
         b.   X $\ldots [_{VP} \ldots [_{DP} \ldots t_X \ldots] \ldots$ V] Aux $\ldots t_{VP}$

This view is clearly too restrictive, as (4) shows – this example is discussed below in (23).

---

[2]It may appear that this treatment of right heads, specifiers, and adjuncts underestimates the actual differences between the structures on the left and on the right hand side in (2). For discussion of differences pertaining to derived c-command relations and the presence of additional functional structure, see Abels and Neeleman (2006).

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

The discussion above is slightly oversimplified. Every violation of the freezing principle can be circumvented if remnant movement is freely available, as (5) shows (see Collins (1994, 1997); Müller (1998) for pertinent discussion of chain interleaving).

(5)   Freezing Principle:
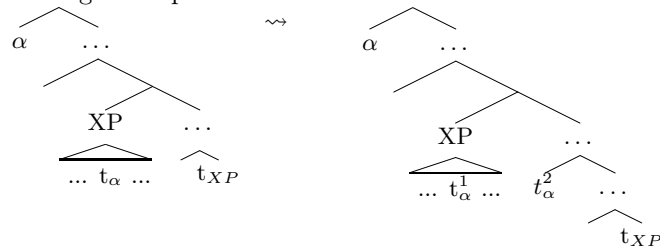
In a theory with the proper binding condition and without other restrictions on remnant movement, the freezing principle has no content. This underscores yet again that the power of remnant movement needs to be sharply restricted.

Thus while the proper binding condition and the freezing principle used to be accepted by many, it has become clear since the publication of Kayne (1994) that the proper binding condition 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)). Likewise, 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, p. 71-78) analysis is incorrect. Under that analysis the "heavy-shifted NP" simply remained in its low VP internal position. It seems now that the "heavy-shifted NP" moves out of VP. There are two implementations for this: the traditional rightward movement analysis and a remnant movement analysis (den Dikken (1995); Rochemont and Culicover (1997); Kayne (1998) among others). Both of them give rise to gross constituent structures along the lines of those generated by the traditional rightward movement analysis. Of course, the long-standing observation still stands that what observationally looks like rightward movement has properties very different from those of leftward movement. Finally, I review examples below that suggest that the freezing principle cannot generally be correct.

These remarks show that, if the explanatory project which motivated the LCA, to account for linear asymmetries in languages, is to succeed, a restrictive theory of operations, movement in particular, is required. The proper binding condition and the freezing principle need to be replaced by more empirically adequate constraints. This is the focal question of this paper. To investigate it, I look at improper movement and related phenomena that bear on the sequencing of movement operations and claim that there are non-trivial constraints on such sequencing. These constraints severely limit remnant movement (see Grewendorf (2003) for a related proposal) and extraction from moved constituents.

The suggestion I develop in this paper is to treat the ban against improper movement – at the very least for heuristic purposes – as a phenomenon rather than an epiphenomenon. Stated in the abstract, the phenomenon seems to be that movement operations must apply in a strictly ordered sequence to any given phrase. 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 (6). Here and throughout A-movement and movement for Case designate the same operation.

(6) $\theta << \text{A-mvt} << Op$

While (6) is just a restatement of the ban against improper movement, I suggest below that the phenomenon of improper movement is much more general than usually assumed. It is more general in two ways. First, more types of movement need to be taken into account. Second, while the ban against improper movement is usually construed as a ban against consecutive movements of the same phrase, I suggest below that it also regulates extraction out of moved phrases and remnant movement, that is, it will carry a lot of the empirical burden of the freezing principle and the proper binding condition. I give a unifying formulation for these three cases below, which is a Generalized Prohibition against Improper Movement (GenPIM).

The similarity to Williams' (2002) Generalized Ban On Improper Movement (GBOIM) is not accidental. This paper is the attempt of capturing important parts of William's Level Embedding Conjecture in a framework with bottom-up tree generation and to develop certain consequences not considered by Williams.

Drawing mostly on German with occasional glances at English, I give empirical justification for generalizing improper movement in just this way in section 2. The claim to universality for the fine-grained hierarchy of operations proposed below remains programmatic. Preliminary investigations of Czech, Serbo-Croatian, and Spanish reveal patterns of data in line with the expectations generated by UCOOL, but I cannot discuss them in this paper for reasons of space. I take the German and English facts discussed to be strong empirical evidence for the claims made here.

Some of the assumptions argued for in section 2 are then brought to bear on the issue of cross-serial dependencies. First I observe that of the two logically possible types of cross-serial dependencies only one is actually attested. This asymmetry is directly predicted by the conjunction of assumptions defended in the first part of the paper.

I also show that the remnant movement derivation of unbounded cross-serial dependencies proposed in Nilsen (2003) necessarily violates the conjunction of the Generalized Prohibition against Improper Movement with the assumption that operations are linearly ordered. The same is true of a number of analyses in Koopman and Szabolcsi (2000), of which their analysis of order preservation is discussed as an example. The results from the first part of the paper thus form an argument against these remnant-movement analyses. The argument is further buttressed by the fact that both of the remnant movement analyses discussed overgenerate in that they willy-nilly give rise to the unattested type of cross-serial dependencies.

2   UCOOL and GenPIM

In this section I show that an accurate description of a wide range of facts can be given if we assume that there is a constraint on the order in which operations, specifically movement operations apply to a given constituent. I call this constraint the Universal Constraint on Operational Ordering in Language or UCOOL, (7). 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.[3]

(7)   The Universal Constraint on Operational Ordering in Language (UCOOL) – *to be refined*
      $\theta << \text{A-mvt} << Op$
      I will say that θ-operations are a *lower* type of operations than A-movement (Case), and that Case-operations are a *higher* type of operations than θ-operations.

The discovery of the ban on improper movement was an automatic side-effect of the move from construction specific transformations to general operations and conditions on them (Chomsky (1973, p. 243 fn. 24, p. 253 # 110c); May (1979); Chomsky (1981, p. 195-204)). Move α run amok would make improper movement possible. The Government and Binding framework managed to keep the derivational part of the theory maximally general, maintaining Move α, by appealing to representational conditions and devices. Thus, in Lectures on Government and Binding Chomsky appeals to the projection principle and the theta criterion to rule out certain instances of improper movement and, following May (1979), to binding theoretic properties of traces to rule out others. However, the notions on which Chomsky's derivation of the ban on improper movement rested are no longer available. Deep structure has been abolished in minimalist theorizing and the projection principle with it. Trace theory has been given up in favor of the copy theory of movement; the binding theoretic assumptions concerning different types of traces are therefore no longer available. Under Chomsky's inclusiveness condition, the explicit manipulation of indeces required to make the binding theoretic account work is no longer available either. At present there is no minimalist account of the ban against improper movement.

UCOOL allows us to approach the relevant facts. It establishes three pairwise orders, none of which seem controversial: $\theta << \text{A-mvt}$, $\theta << \text{Op}$, and $\text{A-mvt} << \text{Op}$.

In the framework of Government and Binding theory θ-roles were assigned at deep structure, ensuring that θ-role assignment could not be fed by movement operations. This assumption is carried over into minimalism by Chomsky (1995, p. 312-313), who explicitly rules out movement into theta positions. He relies on a configurational approach to θ-theory (rather than a feature based one) and the curious additional claim (p. 313) that "[i]f α raises to a θ-position TH,

---

[3]Here and throughout I frame the discussion in terms of different types of movement operations and their sequencing: If $\sigma << \tau$ under UCOOL, then, once τ has applied to phrase P, σ can no longer apply to P. The anonymous reviewer cautions that this might be misread as a return to a construction specific view of movement. However, everything can easily be restated in terms of feature attraction. To do this, UCOOL can be construed as a feature hierarchy or as an extrinsically imposed ordering on feature checking. If we understand it as a feature hierarchy, my formulation in terms of operational sequencing can be restated as follows: If $\sigma << \tau$ under UCOOL, then once τ has been attracted and phrase P pied-piped, σ contained in P can no longer be attracted and/or induce pied-piping of P. A similar, even more clunky reformulation is available in a probe-goal system with the activity condition and the general (EP)P-property driving movement. Nothing contentive is at stake here and my own formulation in terms of operational sequencing seems to me to be maximally transparent.

forming the chain CH = (α, t), the argument that must bear a θ-role is CH, not α. But CH is not in any configuration, and α is not an argument that can receive a θ-role."

Others have argued that movement into θ-positions should be allowed (see Bošković (1994); Bošković and Takahashi (1998); Hornstein (1999, 2001); Landau (2003); Culicover and Jackendoff (2001); Boeckx and Hornstein (2003); Polinsky and Potsdam (2002); Ramchand (to appear) for discussion). A look at the derivations proposed by the proponents of movement into θ-positions shows that θ-roles are always assigned in the lowest positions an argument finds itself in, never in positions that are only reached via case or operator positions.

Simple factual considerations suggest that the consensual position is true, as (8) is intended to indicate.

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

Thus, although there is no consensus on whether movement from θ-position to θ-position exists, there is a consensus that movement from a case to a θ-position or from an operator to a θ-position does not exist. Thus we get $\theta << $ A-mvt and $\theta << $ Op.

The remaining case, the ordering between movement to a case position and movement to an operator position is usually simply stipulated. A typical example is Chomsky and Lasnik's (1995, p. 91) assumption that apart from uniform chains, "[t]he only other legitimate LF objects are operator-variable constructions (α, β), where α is in an $\bar{A}$ position and β heads a legitimate (uniform) chain." As a description this seems consensual. Simple examples suggesting the factual correctness of the consensus are given in (9). All of this motivates A-mvt $<< $ Op.

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

UCOOL makes the consensual point regarding operational sequencing explicit.

I hasten to add that it would be desirable to derive the hierarchy or its effects from deeper principles, but for the moment it is unclear how to do this. Anyway, the fact that an explanation at a deeper level is currently missing should not deter us from investigating and using the hierarchy. The situation is exactly parallel for structural hierarchies: although currently there is no explanation for the CP-IP-VP hierarchy and its cartographic avatars, such hierarchies are used to explain language internal and cross-linguistic facts as well as facts from language acquisition. This mode of reasoning is an explanation at one level, the level called explanatory adequacy in Chomsky (1965), but of course not at a deeper level. I assume that UCOOL can be an explanatory device just as much (or little) as the Cinque hierarchy can and relegate the search for a deeper explanation to further inquiry.

Setting aside the eternal troublemakers for the ban against improper movement (*tough*-movement and relative clauses) for some other occasion, the hierarchy may be taken to restate

what appears to be a fact across languages, namely, that an argument begins its derivational life in a θ-position. It may (or may not) then go on to receive case and move in connection to this. And it may (or may not) then move on to an operator position.

This classic trichotomy into θ-related, case-related, and operator-related operations was well suited to a trichotomous (CP-IP-VP) clause structure. The explosion of clause structure demands a more fine-grained approach to operational types as well. In other words, just like phrase-structure hierarchies have expanded from the simple $[_{CP}[_{IP}[_{VP}...]]]$ to encompass ever more fine-grained structures (see Larson (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)) and UCOOL be adjusted accordingly.

The dichotomy of A- vs. $\overline{A}$-movement, going back apparently to Postal (1971), is insufficient quite independently. It is often assumed that A-movement allows the moved element to act as the binder for anaphors and reciprocals from the landing site of movement, (10a), while $\overline{A}$-movement does not, (10c). From this perspective (10b) is incomprehensible; whatever operation is responsible for for inverting the order of the *to-* and the *about*-PP cannot be A-movement, because of the contrast between (10a-ii) and (10b-ii), but it cannot be $\overline{A}$-movement either, because of the contrast between (10b-iii) and (10c-iii). The binding pattern seen in (10b) is not particularly rare and it is found with mittelfeldscrambling in German and so-called A-scrambling in Japanese.[4]

(10)   a.   (i)   *It seems to each other ('s best friends) that the twins are smart.
             (ii) ✓The twins seem to each other to be smart.
             (iii) ✓the twins seem to each other's best friends to be smart.
       b.   (i)   *John talked to each other ('s best friends) about the twins.
             (ii) *John talked about the twins to each other. (see Postal (1971))
             (iii) ✓John talked about the twins to each other's best friends.
       c.   (i)   *John believes each other ('s best friends) to be fond of the twins.
             (ii) *Which twins does John believe each other to be fond of.
             (iii) *Which twins does John believe each other's best friends to be fond of.

A more comprehensive and fine-grained hierarchy of operations seems to be called for.

UCOOL allows us to approach and state the relevant facts. It is obvious that UCOOL actually compresses a number of logically independent hypotheses into a single statement.

(11)   a.   Movement operations are ordered asymmetrically or antisymmetrically;[5]on both construals (asymmetry and antisymmetry) the claim is that if a movement operation σ feeds a different movement operation τ, then τ does not feed σ; the difference is that an antisymmetric ordering allows movement of some type to feed movement of the same type, while this is ruled out under an asymmetric ordering;
       b.   the ordering of operations is total, i.e., for every pair, σ, τ, of movement operations σ feeds τ or τ feeds σ;

---

[4]I know of no operation whose binding-theoretic fingerprint is the opposite of (10b), that is, where (i) and (iii) are unacceptable and (ii) is acceptable.

    c.   the ordering of operations is universally fixed.

The three claims are logically independent. The conjunction of the three is the empirically strongest position, and it is the position I will adopt. As mentioned, for the three types of operations in (6), the claims in (11), including universality of the ordering, are uncontroversial.[6]

I will now investigate some properties of the deployment of UCOOL. Many questions arise immediately. A first set of questions concerns the hierarchy itself: What counts as a type of operation? How many such types are there? 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. The appeal to relativized minimality, to make sense, presupposes recent work on relativized minimality substantially refining the classes involved as compared to Rizzi's (1990) original three classes (see Starke (2001); Rizzi (2004a, 2006) and related work). To extend to θ-roles, it also presupposes the essential correctness of Hornstein's (1999) approach to the locality of control in terms of closest attract.

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 attempt to provide answers to these questions below.

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 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

---

[5]Asymmetry and antisymmetry are defined as follows (see Partee et al. (1993)):

(i)   a.   A relation R is asymmetric iff $\forall x, y[< x, y >\in R \rightarrow < y, x >\notin R]$
          b.   A relation R is antisymmetric iff $\forall x, y[[< x, y >\in R \wedge < y, x >\in R] \rightarrow x = y]$

[6]Every now and then one or another of the assumptions in (11) are revoked or questioned. Thus Nilsen's (2003) account of cross-serial dependencies effectively revokes (11a), as I will show below; Müller and Sternefeld's (1993) principle of unambiguous binding revokes (11b); and Grewendorf's (2001) account of the lack of superiority effects only in short distance questions in German revokes (11c).

before they can happen in higher clauses.[7] 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 (9b). In the lower phase, 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, which would conform to UCOOL if it were deployed at the phase level. In the upper phase, *what* does not receive a θ-role but it is moved to an A-position followed by movement to an $\overline{A}$-position, again, in conformity with UCOOL applied per phase. Cyclicity in this sense, then, cannot be the answer.[8]

Looking at the problem from the other end of the size scale, relativizing UCOOL 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. For an ordering statement to have consequences, it has 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 – despite the fact that at this level features are assumed to be ordered (Chomsky (1993); Brody (2000); Stabler (1997)). Consider again (9d). 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?
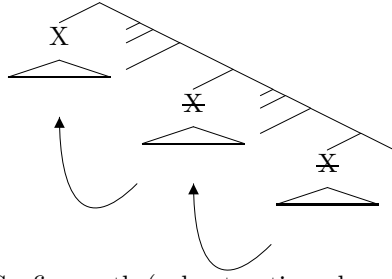
What went wrong is that the entire phrase (qua pied-piping) was involved in high type $\overline{A}$-operation first and later in a lower type of operation, an A-operation. UCOOL must be deployed so as to ban this. This entails that ordering features within a head or a word is at best irrelevant. The constraint has to apply to all parts of the moving constituent. I suggest below that it applies even beyond the moving constituent to all ancestor nodes of the moving constituent which are part of the phrase marker at the moment when movement takes place. In other words, I suggest to extend the ban against improper movement from the identity case, (12a), to cases that do not regularly come under its purview: subextraction out of a moved constituent, (12b), and remnant movement, (12c). Notice that under classical GB assumptions no ordering problems arose, because (12b) and (12c) were simply ruled out: (12b) – by the freezing principle, (12c) – by the proper binding condition.

Sauerland (1996) innovates aquatic metaphors for the configurations in (12b) and (12c). He calls the configuration in (12b) a *surfing* path, because constituent Y uses constituent X as its surfboard to ride up the tree. He calls (12c) a *diving* path, because constituent Y jumps ship and lets X pass without it. I will use these metaphors for their expressive value.

(12)   a.   The standard case (identity):

---

b. Surfing path (subextraction, descendant moves second):
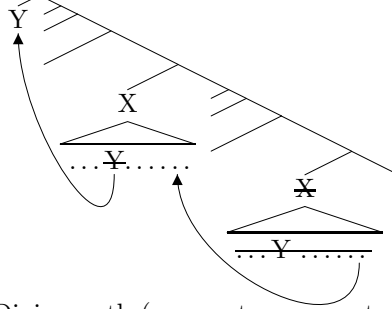
c. Diving path (remnant movement, ancestor moves second):
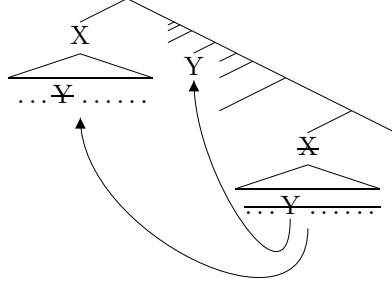
I suggest to deploy UCOOL as stated in the Generalized Prohibition against Improper Movement, GenPIM, (13). GenPIM, together with the definition of affectedness in (14) generalizes the ban against improper movement from the trigger of movement to the moving constituent itself and beyond it to all constituents affected by the movement. The notion of affectedness captures all and only the three cases mentioned above in (12): identity, subextraction, remnant movement.[9]

(13)   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.

(14)   A constituent α is affected by a movement operation iff

---

[9]Grewendorf's (2003) constraint on remnant movement therefore falls out as a subcase of the general ban on improper movement.

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 (14) says is that when a constituent X is moved, this movement affects not only the moving 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 (15), where the nodes affected by movement of X are circled and the unaffected nodes are not.

(15)
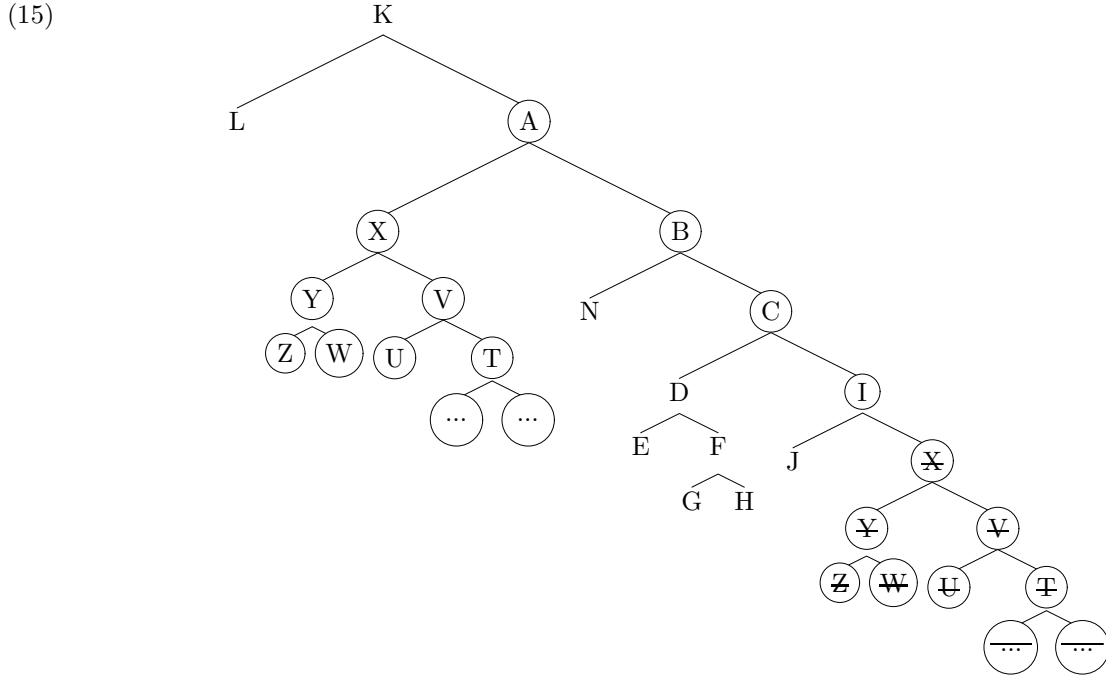
The notion of affectedness seems a bit complicated at first. However, it makes the conceptually desirable claim that when a constituent moves, this has an effect not only on the moving constituent alone but on the rest of the structure as well. Compare GenPIM with a standard formulation of the ban against improper movement in (16). They differ in the boldfaced part.

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

The standard understanding makes sure that movements undergone by a particular constituent do not violate UCOOL, but the effect of movement is sharply localized: the moved constituent is affected, the attracting head presumably is affected, but nothing else is. The standard un-

derstanding isolates the effect of movement and localizes it in the moved constituent and the attractor. This isolating view of movement opens the door to accounts where arbitrary formal features on abstract heads drive movements of arbitrarily picked constituents. Under the isolating view of movement movement never does any harm, since nothing else in the theory or the structure is affected by such movements. The notion of affectedness denies this isolating view of movement – as did the all too rigid old theory with the proper binding condition and the freezing principle. The denial of the isolating view of movement gives rise to a theory which is more strongly integrated, because every movement has an effect on a well-defined, potentially very big, part of the structure. Together with the idea that operations are ordered a(nti)symmetrically and drawn from a universal ordered set, GenPIM puts immediate constraints on what is possible. This seems clearly desirable to me, if only because the resulting theory will be much stronger in the sense of being more easily falsifiable than a theory without these assumptions. The sections of this paper that deal with cross-serial dependencies illustrate precisely this property: certain analyses become impossible in principle under the view advocated here.

but even if it is granted that a non-isolating approach to movement is conceptually desirable, I need to address the apparent complexity in the definition of affectedness. Assume in the background a theory of syntax which countenances only tree extending merge and move, that is, a theory which obeys Chomsky's (1993) extension condition. On a derivational view of syntactic relations (Epstein (1999); Epstein et al. (1998)) an operation can only ever have an effect on syntactic items which are already part of the representation when the operation is performed. In (15), the nodes K and L are not yet part of the representation when X moves, hence, they cannot be affected by that movement. Thus, bottom up tree generation together with a derivational view of syntactic relations derives the first part of (14), which need not be stipulated. With this said consider again (15) excluding K and L from consideration now. The predominant view of movement in the minimalist program is one of remerge: the constituent X is (externally) merged in the lower position and then remerged (internally merged) in the higher position. With this understanding of movement the nodes circled are simply all nodes that enter into a domination relation with X. This restatement alone is a considerable simplification of the notion of affectedness: Movement of X affects all nodes in a domination relation with X. Affectedness, in a derivational perspective, is simply the symmetric version of domination.

There still remains the flavor of a disjunctive definition in (14): "α is in a domination relation with β iff α dominates β **or** β dominates α." Disjunctive definitions are suspect because they merely list cases instead of bringing about a unification. However, the particular disjunction given here does not disjoin arbitrary terms without an obvious relation to each other. It rather creates symmetry from the asymmetric domination relation. Brody (2000, p. 29) reminds us that "[i]n the better developed sciences it is the departures from symmetry rather than the symmetries that are typically taken to be in need of explanation." If, as I argue below, affectedness is an important notion, the priority of symmetry over asymmetry would cast doubt on the assumption that the asymmetric domination relation is truly fundamental. What would need to be explained is not the symmetric notion of affectedness, but the asymmetric notion of dominance.

Developing this speculation just one step further, we might assume that there is an underlying symmetric relation "affect" and the asymmetry of the dominance comes from derivational subsequence: α dominates β iff α and β stand in the (symmetric) affect relation and α was derived

after β was derived. Affectedness might then – possibly together with the equally symmetrical relation of sisterhood – be *the* fundamental syntactic relation. To appreciate this, consider the c-command relation. Chomsky (2000) defines it as the composition of sisterhood with dominance. If affectedness is indeed a syntactic relation, then c-command need not be defined at all. It falls out: X c-commands Y iff at the time when X is (externally or internally) merged into the tree, X is not in the "affect" relation with Y (see Abels (2003, chapter 2 section 3) for further discussion).

I now turn to GenPIM. Although GenPIM is not a derivational constraint as given, a derivational metaphor might make it more tangible. A constituent, X in (15), has a number of things it needs to do in its derivational lifespan, that is, a given constituent might contain a number of unchecked or active features. 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 (16), X may not undergo a less urgent (higher) operation if any of its parts still have more urgent business to do, this will yield restrictions on (12b). 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, which will yield (12c).

The generalization from (16) to GenPIM will be justified in the next subsection. Before we look at the empirical justification, a few final 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 (16). (16) 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.

2.2 A case study

Clearly, the scope of GenPIM goes well beyond the realm usually considered to fall under improper movement. Consider the following simple table, which represents a presumed subpart of the Universal Constraint on Operational Ordering in Language. I have left out θ-operations here, but I 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 into *wh*-movement and topicalization – topicalization again of the German vorfeld variety[10] The left-to-

[10]See Frey (2000, 2005) for discussion and comparison to topicalization in English with German vorfeld topicalization. Frey discusses a number of interpretive differences between what is called "topicalization" in these

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 and Sternefeld (1992) for a structural formulation of the hierarchy).

(17)

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

The table is to be read as follows: Rows are taken as the first step of movement, columns – as the second. The first row claims that A-movement feeds the other three types of movement, the second – that scrambling feeds *wh*-movement and topicalization but not A-movement, the third – that *wh*-movement feeds topicalization but not the other two types of movement, and the forth – that topicalization feeds none of the other three types of movement.

The diagonal of the table is left blank. The reason is that I left open the question whether the ordering of operations is asymmetric or antisymmetric, (11a). If it is antisymmetric, the diagonal should be populated by checkmarks. If it is asymmetric, the diagonal should be filled in with asterisks. Some considerations that bear on this involve the question 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 we conclude that the ordering is antisymmetric. Similarly, if the movement analysis of control is correct, then the fact that subject control predicates can stack indicates that the ordering is antisymmetric and the cells along the diagonal of the table should 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 (2006) Criterial Freezing), then we are led to conclude that the ordering is asymmetric. In this case the diagonal of the table to be populated by asterisks.

Similar considerations apply to surfing paths. We might interpret the fact that *wh*-extraction out of a *wh*-moved constituents is possible (see Rizzi (2006, p. 13 ex. 36b), Chomsky (1986, p. 26); Starke (2001, p. 54)), as an indication that the ordering of operations is antisymmetric. It seems equally plausible to claim, however, that in these cases the two movements have different properties (see Starke (2001) for discussion), in which case the ordering would be asymmetric.

Finally, there is a well known restriction on remnant movement: remnant creating movement and remnant movement may not be of the same type (Müller (1998, chapter 5); Kitahara (1997); Hiraiwa (2002); Grewendorf (2003)): remnant creating *wh*-movement never feeds remnant *wh*-movement; remnant creating topicalization never feeds remnant topicaliztion; etc. This restriction is often referred to as Müller's generalization. Taken at face value, Müller's gener-

---

languages. Below I show that German topicalization does not feed *wh*-movement. The anonymous reviewer reminds me that English topicalization – at least marginally – does (Lasnik and Saito, 1992, p. 101-102), which suggests that English topicalization might be closer to what I call scrambling in German.

alization suggests that the ordering of operations is asymmetric, hence the diagonal should be filled with asterisks. However, some of the literature just mentioned argues that Müller's generalization follows from the minimal link or the A-over-A condition; in a structure like (18), the interrogative complementizer, $^2C^0$, cannot attract the lower wh-element pied-piping ZP and leave the higher wh-element to be attracted by $^1C^0$ with concomitant (remnant) pied-piping of YP, because the first step violates closest attract or the A-over-A condition. If this reasoning is correct, then it is unnecessary to derive Müller's generalization from an asymmetric ordering of operations and an antisymmetric ordering will do the job. I make crucial use of Müller's generalization later on in the paper, but either way of deriving it will serve the purposes of the argumentation at that point.

(18) $[\ ^1C^0_{[+\text{wh}]} \ldots [\ ^2C^0_{[+\text{wh}]} \ldots [_{\text{YP}} \ wh \ldots [_{\text{ZP}} \ldots wh \ldots] \ ] \ ] \ ]$

I therefore leave the decision whether the ordering of operations is asymmetric or antisymmetric open and the diagonal of the table blank.

Identity

Suppose we interpret the table as a table of the combinability of operations applying to one and the same constituent, the identity case in (12a). The first row of the table now says that a constituent that has undergone A-movement may subsequently 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 actual distribution of the data corresponds very closely to that given in the table. In particular there is no reason to believe that the operational sequences marked with an asterisk ever occur. It follows that no pair of operations σ, τ can be in a symmetric feeding relation. This observation strongly supports the assumption that all operations are a(nti)symmetrically ordered with respect to each other, (11a). The totality of the ordering is supported to the extent that all the cells with a checkmark in the table can be shown to correspond to grammatical sentences.

I take it to be clear that A-movement, that is case marking, may precede scrambling, wh-movement, and topicalization and that topicalization cannot feed any of the other operations (see Sternefeld (1992) for the relevant data). In other words, the first and the fourth row of the table are not disputed. It is less clear whether the second and third row are entirely correct. Wh-movement does not seem to feed topicalization in German.[11] Thinking along fairly standard lines, one may imagine that topicalization of an argument requires the existence of an identifiable discourse referent as topic, whereas a wh-question would be used to establish or identify such a referent. It would then follow that the same argument cannot be both questioned and topicalized

---

[11]Similarly for Japanese, where the topic marker apparently does not attach to wh-phrases – see Kuroda (1965), Heycock (to appear) for a useful overview. This might be attributed to a semantic or pragmatic incompatibility (see already Bach (1971) for this suggestion).

and the fact that *wh*-movement does not feed topicalization would not bear on the ordering of operations.[12]

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, (20) (from Fanselow (2001, p. 414)).

(19)   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?

A strong argument for the assumption that scrambling feeds *wh*-movment is given by Wiltschko (1997, 1998) (see Fanselow (2004); Pesetsky (2000); Grewendorf (2001) for discussion). As is well known scrambling does not give rise to weak cross-over effects, (20a) vs. (20b), and it is clause-bound, (21).

(20)   a. *Früher haben $seine_k$ Eltern jeden $Studenten_k$ unterstützt.
           earlier have his   parents every.ACC student.ACC supported
           In the past his_k parents supported every $student_k$.
    b. ✓Früher haben jeden $Studenten_k$ $seine_k$ Eltern unterstützt.
           earlier have every.ACC student.ACC his   parents supported
           In the past his_k parents supported every $student_k$.

(21)   *Gestern hat jeden   Studenten   Hans gesagt, dass seine Eltern
       Yesterday has every.ACC student.ACC Hans said   that his   parents $t_{everystudent}$
       unterstützen
       support

Short *wh*-movement does not give rise to weak crossover effects, (22a), but long *wh*-movement does, (22b), but only in the higher clause, not in the clause where the *wh*-word originates, (22c).

(22)   a. ✓Welchen $Studenten_k$ unterstützen $seine_k$ Eltern?
           which.ACC student.ACC support his_k parents?
           Which $student_k$ do his_k parents support?

---

[12]Rizzi (2006, p. 122) treats D-linked *wh*-phrases as topics. The reasoning in the text suggests that this is not correct. Indeed, in many languages the morphosyntax of D-linked *wh*-phrases is different from that of topics. Japanese, see fn. 11, is typical in this respect. Presumably D-linking has more to do with restricted quantification than with topicality.

b. *Welchen   Studenten$_k$ glauben seine$_k$ Eltern, dass Maria unterstützt?
   which.ACC student.ACC believe   his   parents that Maria supports
   Which student$_k$ believe his$_k$ parents that Maria supports?

c. ✓Welchen   Studenten$_k$ glaubt Maria, dass seine$_k$ Eltern unterstützen?
   which.ACC student.ACC believes Maria that his$_k$ parents support?
   Which student$_k$ does Maria believe that his$_k$ parents support?

This schizophrenic behavior of *wh*-movement in German, Wiltschko argues, can be straightforwardly explained if *wh*-movement gives rise to weak crossover effects but can be fed by scrambling. In (22a) and (22c) the *wh*-phrase would undergo cross-over obviating scrambling before undergoing *wh*-movement, but this possibility is barred in (22b): (i) scrambling is clause-bounded, (21), and (ii) an improper movement derivation where the initial step of *wh*-movement feeds a scrambling step into the higher clause which again feeds the final step of *wh*-movement is blocked by UCOOL.[13]

The paradigm in (22) can be replicated for topicalization by simply replacing the word *welchen* – 'which' by *jeden* – 'every'. The judgments remain as given above. Topicalization is as schizophrenic when it comes to weak crossover as *wh*-movement. By parody of reasoning I conclude that scrambling also feeds topicalization in German. These conclusions are not shared by all researchers; thus, the discussion here remains necessarily somewhat inconclusive.

The facts seem to pattern overall as presented in the table. But 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, as in the case of *wh*-movement and topicalization above. The hypothesis would be disconfirmed only if feeding relations between operations were discovered that correspond to the part of the table below the diagonal or if mutual feeding-relations, that is, symmetry, was discovered. No such cases exist.

Surfing Paths

In the discussion above I assumed that the rows and columns of table (17) are to be interpreted as subsequent movements of the *same* constituent, (12a). 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 (Sauerland's *surfing*-paths). In other words, the rows label the prior movement of X while the columns label the subsequent movement of Y in (12b).

Traditionally these structures were ruled out categorically by the freezing prinicple. I will now review some evidence showing that the freezing principle is too strong and that UCOOL together with GenPIM offer a much more accurate description of the situation.

The first row of (17) now claims that A-movement should feed scrambling, *wh*-movement, and topicalization. An example of the first is furnished by (23b) on the assumption that objects in German move to Case positions. We can demonstrate that the prepositional adverbial *darüber*

---

[13]Similar facts are found in Czech (Pavel Caha and Lucie Medova, p.c.). Clause-bound scrambling and *wh*-movement do not give rise to weak crossover effects, but long *wh*-movement does, but only for binding into arguments in the higher clauses.Within the clause where the *wh*-phrase originates there are no weak crossover effects. This pattern of data is as expected under the current setup. The explanation for Czech would run exactly parallel to Wiltschko's explanation of the German pattern in terms of scrambling feeding *wh*-movement.

has moved out of the NP by observing the specificity effect in (23b) (for detailed discussion of this point see e.g. Müller (1998, p. 10-15)).

(23) 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   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. (24) furnishes a simple example of extraction from an A-moved argument.[14] The point made by this example stands independently of whether 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 like (25b) (Starke, 2001, p. 36), which, according to Starke, is at worst slightly worse than the corresponding object extraction case, (25a). 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.)

(24) a. Which football team do you want/expect [the manager of t_which football team] to pay a large fine?
      b. Which football team do you believe [the manager of t_which football team] to have paid a large fine?
      c. Which politician do you believe the rumors about to be false?

(25) 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   première partie   va   créer un
         of which film is-it   that you think that [the first   part   t] goes create a
         scandale?
         scandale
         Which movie do you think that the first part of would create a scandal?

Similar examples with topicalization from a subject are easy to construct, (26). 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).

---

[14]Chomsky (1973, p. 249-250); Kayne (1984, p. 169); Pesetsky (1982, p. 319); and for German (Müller, 1995, p. 390 fn. 26) find similar examples unacceptable; my informants disagree and find them only slightly degraded. One important difference between the present examples and Chomsky's original one is that Chomsky's example contained a non-D-linked *wh*-phrase. Chomsky (2004, ex. # 19) has recently claimed that PPs (with a D-linked *wh*-phrase) can be extracted from ECM subjects. For some speakers pied-piping of *of* seems to be necessary to make these examples acceptable.

(26)   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 claims made in the table based on UCOOL and GenPIM, that A-movement should feed scrambling, *wh*-movement, and topicalization is correct. The many cases where this is impossible will have to be ruled out on independent grounds.

The opposite, that is, examples where scrambling, *wh*-movement, or topicalization feed A-movement are strikingly bad. I discuss just a single case here as an example, (27) (from Sakai (1994, p. 300), discussed as a problem in Collins (2005)). The example shows that A-movement from a *wh*-moved constituent is impossible. The other types of example are equally bad. On the intended but ungrammatical construal where *Oscar* A-moves out of the *wh*-moved clause, (27b) 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 the wrong order since A-movement out of an $\overline{A}$-moved constituent is ruled out under GenPIM. Example (27c) is a control suggested by the anonymous reviewer to check whether the infinitive can be pied-piped at all if there is no prior raising out of it.

(27)   a. Max asked how likely to win Oscar was
            $[_{IP1}$ Max asked $[_{CP2}$ $[_{AP}$ how likely t_j to win $]_i$ $[_{IP2}$ Oscar$_j$ was t_i $]]]$
        b. *Oscar was asked how likely to win it was.[15]
            $[_{IP1}$ Oscar$_j$ was asked $[_{CP2}$ $[_{AP}$ how likely t_j to win $]_i$ $[_{IP2}$ it was t_i $]]]$
        c. ✓Max asked how likely for Oscar to win it would be if he bought better equipment.
            (Gillian Ramchand, p.c.)

The second row of the table claims that scrambling should feed both *wh*-movement and topicalization. That this is correct is shown in (28a-b) and (29a) respectively.[16] There is a clear contrast between (28a-b) with *wh*-movement from a scrambled constituent and (28c) with scrambling from a *wh*-moved constituent. Likewise, there is a sharp contrast between (29a) with topicalization from a scrambled constituent and (29b) with scrambling from a topicalized one.

(28)   a. Worüber   kann einen Südkurier-Artikel   selbst Peter nicht am
            what-about can [a   Südkurier-article $t_{worüber}$] 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 (Kuthy and Meurers, 1999, p. 27)
            attributed to Fanselow (1991))

---

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

[16]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 and Pafel (1993) on the difficulty of delineating specificity effects precisely in German).

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

c. *Gestern hat über Karl den Großen keiner gefragt, was für eine Arbeit
yesterday has [about Charlemagne]$_{PP}$ nobody asked what for a work t_{PP}
er schreiben soll
he write should
Nobody asked yesterday what kind of paper about Charlemagne he should write.

(29) 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 Kongress hat über Karl den Großen keiner gesagt ein fertiges
at the congress has [about Charlemagne]$_{PP}$ nobody said a done
Manuskript habe er anzubieten.
manuscript t_{PP} have he to.offer
At the congress nobody said that he had a completed manuscript about Charle-
magne to offer.

The final case, topicalization out of a *wh*-moved constituent is degraded, (30a). However,
there is still a clear contrast between (30a) and *wh*-movement out of a topicalized constituent,
(30b).

(30) 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?

The generalized version of the ban against improper movement suggested above conforms
with the facts quite nicely. Empirically it is clearly superior to a theory that bans all extraction
from moved constituents by appeal to the freezing principle.

Diving Paths

Finally, let us turn to the third interpretation we can give to the table in (17). Consider (12c) again. We can now interpret the rows as indicating a first movement step of Y and the columns as the following remnant movement step of X.

The first row of the table in (17) now claims that A-movement may be followed by remnant scrambling, remnant *wh*-movement, and remnant topicalization. All three cases exist. For *wh*-movement and remnant topicalization this can be demonstrated even for English, (31).

(31)   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, (32a), based on Müller (1998, p. 226). It may be debatable whether the scrambled category α in (32a) contains the trace of the subject, since it may be that a smaller constituent than the full VP is fronted. The same is not true in (32b), where the subject is almost certainly extracted from α. To my ear, these examples have a similar, slightly degraded status.[17]

(32)   a.   ?... dass   ein Buch gelesen wohl   keiner   haben wird
            ... that [t$_{keiner}$ a   book read]$_\alpha$ probably nobody t$_\alpha$ have   will
            ... that in all likelyhood nobody has read a book
       b.   ?... dass von einem Studenten   angefasst kein einziges Reagenzglas   werden
            ... that [by a   student   t$_{NP}$ touched]$_\alpha$ [no single   test tube]$_{NP}$ t$_\alpha$ become
            durfte
            may.PAST
            ... that no student was allowed to tough a test tube

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, (33) (based on Müller (1998, p. 221-222)), and remnant topicalization, (34) (Müller, 1998, p. 10).

(33)   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?

(34)   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 (35a) is no more or less degraded than that in (35b) (based on (Müller, 1998, p. 301)).

(35)   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$_\alpha$ refused   has

---

[17]I would like to thank Eric Haeberli for pointing out some interfering factors with my original examples that made them worse than they needed to be.

23

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 opposite feeding relations are never possible.[18]

2.3   Conclusion

The following table summarizes the facts reviewed above. As can be seen, they fit the expectations generated by the conjunction of UCOOLand GenPIM very well. In particular, all of the feeding relations which are ruled out by the conjunction of the two principles fail to occur and all of the feeding relation which are predicted to exist by the conjunction of the two principles do indeed occur. The only data point that might remain a bit dubious is the feeding relation between *wh*-movement and topicalization, where the deviance of the examples seems to have different sources, though: a pragmatic incompatibility and the *wh*-island effect.

(36)

		A-mvt	Scrambling	*wh*-mvt	Topicalization
A-mvt	identity		✓	✓	✓
	surfing path		✓	✓	✓
	diving path		?	✓	✓
Scrambling	identity	*		✓	✓
	surfing path	*		%	%
	diving path	*		✓	✓
wh-mvt	identity	*	*		*
	surfing path	*	*		??
	diving path	*	*		??
Topicalization	identity	*	*	*	
	surfing path	*	*	*	
	diving path	*	*	*	

The facts strongly support the idea that operations are asymmetrically ordered, that the ordering is total, that it is as hypothesized for German, (37), and that the same ordering applies for the identity case, the case of surfing paths, and that of diving paths. As mentioned at the beginning, the claim to universality remains a promise at this stage.

---

[18]It could be countered that the examples predicted to be bad under this interpretation of table (17) are ungrammatical simply because the putative second operation is clause bound. This is true, but it is also 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.

Without a viable counterproposal, this objection has little force.

(37) $\theta << $ scrambling $<< $ A-mvt $<< wh << $ topicalization

Except for Williams' theory, to which I turn immediately below, none of the existing theories of improper movement extend from the identity case to the cases of surfing and diving paths. These approaches thereby miss the generalization illustrated in (36) and formulated in UCOOL and GenPIM. These are strong grounds for adopting a system which incorporates or derives UCOOL and GenPIM. The method used here makes immediate predictions. Once two distinct operations are identified, they will have to find their place on the hierarchy since the order is, by assumption, total. The prediction is then that the feeding and bleeding relations to all other operations in the hierarchy are automatically fixed and fixed in identical ways for the identity case, the case of diving paths, and the case of surfing paths.

I now turn to a brief discussion of Williams (2002). Williams' theory has a fairly different architecture from Chomskyan minimalism and a summary of it would lead too far afield here (see Nevins and Hornstein (2005) for a brief synopsis). I will therefore concentrate on a particular prediction that emerges from Williams' theory. In Williams' theory (as in an earlier proposal by Sternefeld (1992)) the ordering and scope of operations is directly tied to the structural hierarchy in the clause. Williams derives a number of particularly strong claims: (i) Movements that happen later in the overall derivation target position higher in the functional hierarchy than movements that happen earlier; (ii) movements that happen later may be longer than all movements that happened earlier; (iii) the launching site of movement is systematically lower (not only in terms of c-command but also on the functional hierarchy) than the landing site of movement; and (iv) no movement may cross positions on the functional hierarchy that are higher on that hierarchy than the landing site. These claims all follow under Williams' architecture.

I limit the discussion to (iii) and (iv) here since only they are directly relevant to the topic of this paper. We can formulate them slightly more formally by construing of the hierarchy of functional projections as a numbered series of heads where the lowest head is numbered 1 and numbers go up steadily to the highest head (presumably situated in the complementizer domain). Property (iii) now says that if a constituent occupies a specifier or complement position of a head indexed i and undergoes movement which targets the specifier (or head-adjoined) position of a head indexed j, then $i \leqslant j$. Property (iv) says that if a constituent X has landed in the specifier position of a head indexed j, then the movement path may not have crossed the projections of any heads indexed k, where $k > j$. In a minimalist syntax with a Cinque hierarchy, these results follow for clause internal movement but not for cross-clausal movement. Under Williams' theory the claim is completely general.

The ban against improper movement follows from (iii) and (iv) directly. Consider (38). (iii) rules out the derivation in (38a); the movement step from [Spec, CP] to [Spec, TP] is illicit under (iii) because C^0 is higher in the functional hierarchy than T^0. (iv) rules out the derivation in (38b); the movement step from the embedded object position directly to the matrix [Spec, TP] crosses along its way a head, C^0, which is higher in the hierarchy than the head which provides the landing site of that movement step, T^0.

(38)   *What seems that it was said?
    a.   [CP What [IP t^3_{what} seems [CP t^2_{what} that [IP it was said t^1_{what} ?]]]]
    b.   [CP What [IP t^2_{what} seems [CP that [IP it was said t^1_{what} ?]]]]

Claims (iii) and (iv) are sufficiently general so that their conjunction rules out all of the cases of improper movement discussed above including the diving and surfing paths.

Nevertheless, I will not adopt Williams' proposal because it seems overly restrictive. Consider exceptional case marking in English, (39). Under the now standard analysis of exceptional case marking as involving raising to a non-thematic object position in the higher clause (see Postal (1974); Lasnik and Saito (1991); Davies and Dubinsky (2004) for discussion and references), the subject of the embedded infinitival clause moves to a low position in the matrix clause. Movement from the original subject position in the lower clause violates (iii) or (iv) (or both). Suppose for concreteness that the embedded infinitival is a TP. The position of negation in the matrix clause, to the left of the ECM subject, indicates that the ECM subject has not moved up to TP in the matrix. But in the embedded clause the ECM subject must either move past TP without landing there, move through TP and land there, or originate in TP. In the first case, movement of the ECM subject would violate condition (iv). In the second two cases, movement of the ECM subject would violate condition (iii).[19]

(39)   John doesn't believe Mary not to have been telling the truth.

Similar problems arise with quantifier movement in French across complementizers, (40) (which is mentioned as a problem in Williams (2002); the Icelandic situation seems to be similar (Kayne, 1998, p. 142 #55-6)). The quantifier originates in object position in the embedded clause and ends up in a position preceding the non-finite main verb and following the finite auxiliary in the main clause, (40b). If the quantifier crosses the complementizer *que* in one step, then that movement violates condition (iv). If there is an intermediate movement step to [Spec, CP], then the movement step from [Spec, CP] to the final landing site violates (iii).

(40)   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.

Examples of this type are not too hard to come by. Thus, long scrambling in Russian crosses complementizers (the complementizer *kak* introducing direct perception reports) but does not seem to necessarily end up in [Spec, CP] in the higher clause (Glushan (2006)). Hyperraising (Ura, 1994) poses similar problems given that it apparently crosses complementizers but lands within TP (see Nevins (2004); Williams (to appear) for directly relevant discussion). Likewise cyclic movement across complementizers contained in lower clauses but ending up in a position to the right of the same type of complementizer in the higher clause are predicted to be impossible

---

[19]Williams (2002) suggests a solution to this particular problem in terms of what he calls mismapping, discussion of which would lead too far afield. If I understand Williams correctly, the examples discussed immediately below are not amenable to an analysis in terms of mismapping.

by the conjunction of (iii) and (iv). However, such cases do exist, (41).[20] The examples in (41) 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 (41) illustrate this pattern. They are synonymous, the only difference relevant to the discussion is the position of the *wh*-phrase *ûû* - 'what'.[21]

The declarative complementizer in Kîîtharaka is *atî* – 'that'. In (41b-d) the *wh*-word has undergone partial *wh*-movement. Notice that partial *wh*-movement is compatible with the occurrence of a declarative complementizer in the clause hosting the partially moved *wh*-phrase, (41c-d). In both (41c) and (41d) the *wh*-phrase appears to the right of the complementizer *atî* – presumably because the landing site of partial *wh*-movement is below the complementizer position. The crucial example is (41d). Here the *wh*-phrase has crossed one instance of the complementizer *atî*, underlined in the example, but ends up to the right, that is, below, the other instance of the complementizer, circled in the example. Assuming a fixed position in the Cinque hierarchy for the complementizer in Kîîtharaka, we end up with a violation of (iii) or (iv).

(41) 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 <u>atî</u> 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?

---

[20]Non-obvious glosses are as follows:

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

[21]The syntax, semantics, and morphology of the marker of focus and cyclicity glossed F is discussed in detail in Abels and Muriungi (2006a,b).

27

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?

I take such cases to be sufficient grounds for rejecting Williams' proposal as too restrictive and will stick with UCOOL and GenPIM, which do not make the problematic predictions.

In the next section I turn to cross-serial dependencies. The assumptions strongly supported by the facts so far (a(nti)symmetry, totality, extension to surfing and diving paths) predict a particular linear asymmetry in the area of cross-serial dependencies. The prediction appears to be borne out. These assumptions also rule out a family of remnant movement accounts of cross-serial dependencies found in the literature. These accounts, ruled out under present assumption, also necessarily fail to predict the observed asymmetry. We can therefore reject the accounts in question.

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, as in (42) from Bresnan et al. (1982). The standard German versions of such sentences show nesting dependencies, (43). The corresponding English construction is shown in (44). 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. The selectional dependency between the verbs goes from left to right in Swiss German, Dutch, and English and from right to left in German.

(42) . . . dat Jan Piet Marie de kinderen   zag   helpen   laten   zwemmen
      . . . that Jan Piet Marie the children   see-past help-inf   make-inf   swim-inf

(43)
   ...dass Jens Peter Maria die Kinder   schwimmen lassen helfen   sieht
   ...that Jens Peter Maria the children   swim-inf let-inf help-inf see-pres

(44)   ...that John saw Peter help Mary make the children swim

Huybregts (1976) argues that the verb-raising construction is unbounded.[22] Speakers usually have difficulties with sentences containing cross-serial dependencies of length four or more. To defend the claim that the grammar of Dutch allows unbounded cross-serial dependencies Huybregts gives the following arguments: First, a number of verbs can come in either order under verb raising, (45). Second, for the triple *let–help–teach* Huybregts shows that the verbs can come in any of the six mathematically possible orders, (46). Finally, the construction is recursive since one and the same verb can occur more than once, (47). The apparent upper bound on the verb-raising construction is therefore usually attributed to processing complexity.

(45)   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 drink
             ...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

---

[22]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 (2007); Steedman (2000)). Huybregts (1984) discusses a slightly different construction in Dutch which is even weakly context sensitive.

    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

(46)   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

(47)   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

With the previous literature I conclude that the grammar of Dutch allows unbounded cross-serial dependencies. The importance of the lack of a syntactic bound on the phenomenon can hardly be overestimated. As is well-known, verb raising gives rise to clause-union effects. (See Evers (1975); Haegeman and van 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. A Cinqueesque solution is incompatible with free ordering and recursion, that is, unboundedness.

Another example of cross-serial dependencies comes from standard Norwegian (Nilsen (2003, S. 72); Bentzen (2005)). The relevant construction is exemplified in (48) where all the adverbs must precede all the verbs. 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 (48b).

(48) 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 (49), an inverted version of Shieber's famous Swiss German example, (50). 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.[23] This lack cannot be attributed to the lack of inverse orderings of arguments per se, since inverse orders of arguments are attested (e.g., postverbally in Kiowa (Adger et al., 2007) and in Malagasy (Pearson, 2000)).

(49)   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

(50)   . . . 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

I claim that the absence of backwards cross-serial dependencies is not an accidental gap in the database, but reflects a deep property of language. I show below that the assumption that operations are a(nti)symmetrically, totally ordered and that that order applies in the identity, the surfing, and the diving cases alike (together with a number of independently motivated assumptions) predicts the absence of unbounded inverse cross-serial dependencies. Universal (51) is then another testable consequence of the assumptions argued for here.

(51)   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, (51), and with respect to their compliance with the a(nti)symmetric, total, ordering of operations and GenPIM. The remnant movement

---

[23]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 (48) seem to be a counterexample to Bobaljik's generalization. Tagalog (VSO) is another apparent counterexample.

account, it turns out, does not predict the asymmetry nor is it compatible with the assumptions regarding ordering of operations in conjunction with GenPIM. Before continuing, I need to justify the decision not to take into account any rightward-movement analyses of verb raising. This is crucial, because the prediction regarding inverse cross-serial dependencies follows only if rightward movement does not exist.

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 van 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.[24]

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 (50) is roughly as in (52a). 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 merging a licensing head with the property of attracting all DPs into its specifier positions, as shown in (52b-c).[25]

---

[24]A second reason for excluding rightward-movement as an analytic option is that such analyses invariably require 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. The present formulation of UCOOL and GenPIM does not derive the right roof constraint. Eventually this should be changed.

Notice that if all superficial rightward movement is analyzed as remnant movement, then the right roof constraint bans movement of remnants containing traces of remnant creating movement which originate more than one clause deep inside the remnant. Stated this way, the right roof constraint becomes part of a larger generalization known to constrain remnant movement in Germanic: Movement of remnants containing unbound intermediate traces is impossible (see den Besten and Webelhuth (1987, 1990); Grewendorf (1994); Fanselow (1993); Bayer (1996); Müller (1998, 1999) for relevant discussion). What this generalized version of the constraint follows from remains unclear.

[25]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 (52a). 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.

(52)　a.

DP
mer / we　　laa / let

DP
d'chind / the children　　hälfed / help

DP
em Hans / Hans　　aastricht / paint

DP
es huus / the house

b.

$L_{DP}^{\forall}$

DP
mer / we　　laa / let

DP
d'chind / the children　　hälfed / help

DP
em Hans / Hans　　aastricht / paint

DP
es huus / the house

c.

$L_{DP}^{\forall}P$

DP_1
mer / we

DP_2
d'chind / the children

DP_3
em Hans / Hans

DP_4
es huus / the house

$L_{DP}^{\forall}$

t_{DP_1}

laa / let

t_{DP_2}

hälfed / help

t_{DP_3}

aastricht / paint　　t_{DP_4}

34

In order to derive the cross-serial order in (52c), I invoked two crucial assumptions. First, I assumed that there are heads that attract all (active) members of a given class, here DP: $L_{DP}^{\forall}$. 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. Put yet another way, a linearly ordered, finite sequence of dedicated heads each attracting a single DP is insufficient.

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 (1995) 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, (53a-b). Example (53) is intended to capture in the abstract the spirit of all formulations of relativized minimality. (53a) claims that if α and β are in the same class targeted by operation Δ and α c-commands β before the application of Δ, then α c-commands β also after the operations applies. The class of items relevant to the operation have identical c-command relations before and after the operation applies – though of course, α might move past other elements that are not in the class targeted by Δ. (53b) makes the same claim for β: applying Δ may not change the c-command relations between α and β. Note that derivations of the type in (53b) might be impossible because of the extension condition. This does not affect the point that Relativized Minimality per se allows such derivations.

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

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. $\alpha \prec \beta \vDash_\Delta \alpha \prec \beta \prec t_\alpha \prec t_\beta$

By contrast (53c) is disallowed by Relativized Minimality, because the pre- and post-movement orders are reversed. There is no homomorphism between the structures before and after Δ applies.

By the logic of order preservation derivations such as those in (53d) 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), a concrete formulation is given in (54a).[26]

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 have to be reanalyzed.[27]

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 builds on the idea that movement is a homomorphic mapping (see above) and is therefore the most plausible candidate.

The crucial assumptions are summarized again below.

(54)   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 an operation targeting that class may not cross X and Y.

            (ii)   *Crossing* – generalized version to accommodate (54d) (see Starke (2001))

---

[26]Richards' (1997) mechanism of *tucking in* achieves the same result, but at the cost of violating the extension condition, a move that I would like to avoid.

[27]A possible way forward on the issue of nesting dependencies is to make use of feature geometries. Thus suppose there are two operations, Δ and Γ which target the features α and β respectively. Assume furthermore that β is a feature which depends on α in the sense that it is lower in the feature geometry than α. If Δ targeting α applies to the structure in (ia) followed by application of Γ targeting β a nesting configuration results, (ib). If the operations apply in the opposite order, Y will move first, (ic). Applying Δ to this structure will either move Y again or lead to a (defective) intervention effect. Either way, a nesting dependency is not derivable this way.

(i)   a. $[\,X_{[\alpha]}\,\ldots[\ldots Y_{[\alpha\,:\,\beta]}\,\ldots]\,]$

          b. $[\,Y_{[\alpha\,:\,\beta]}\,\ldots[\,X_{[\alpha]}\,[\,t_X\,\ldots[\ldots t_Y\,\ldots]\,]\,]\,]$

          c. $[\,Y_{[\alpha\,:\,\beta]}\,\ldots[\,X_{[\alpha]}\,\ldots[\ldots t_Y\,\ldots]\,]\,]$

Reversing the original arrangement of the features as in (iia), will give rise to movement of Y by Relativized Minimality whether Δ is applied first or Γ. If Δ is applied first, then Γ might move Y another step leaving X in place. If Γ is applied first, then Y will either move another step or give rise to a defective intervention effect and block movement of X. The only possible derivation where both X and Y move is the one in (ib), which yields the nesting path.

(ii)   a. $[\,Y_{[\alpha\,:\,\beta]}\,\ldots[\ldots X_{[\alpha]}\,\ldots]\,]$

         b. $[\,Y_{[\alpha\,:\,\beta]}\,\ldots[\,t_Y\,\ldots[\ldots X_{[\alpha]}\,\ldots]\,]\,]$

This is Starke's (2001) logic for weak island phenomena and it might be applicable in the relevant cases.

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, (48), can be dealt with along the same lines if we assume that there is a head which attracts all adverbs, $L_{Adv}^{\forall}$.
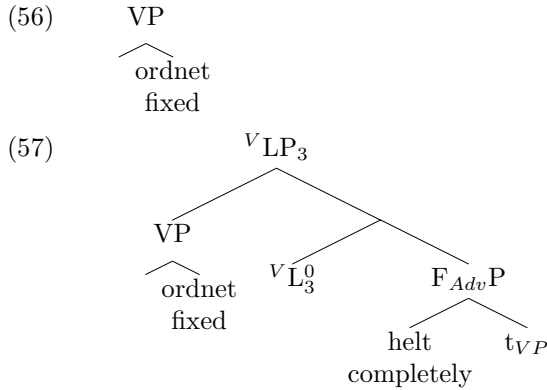
Derivations like these are fully consistent with the assumption that operations are ordered a(nti)symmetrically. The obvious question to ask at this point is whether the ban against inverse cross-serial dependencies, (51), can be derived. I postpone discussion of this question until section 5 and turn 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 (55).
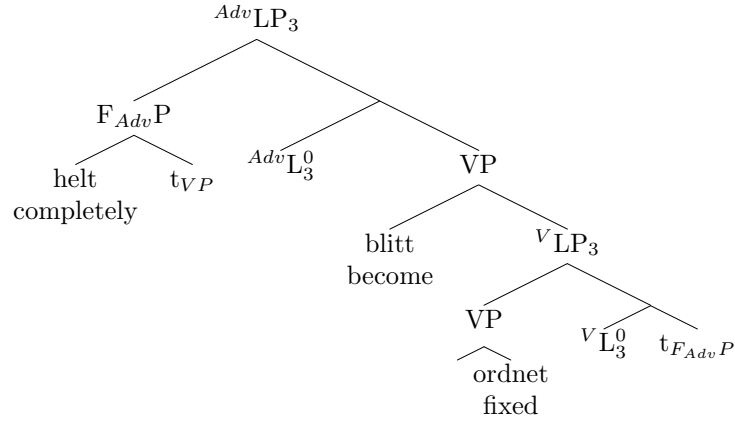
(55) a. Above every verbal head there is an adverb lifter, that is, a head which attracts the maximal projection of the closest head with an adverbial in its specifier, $^{Adv}L^0$.

b. Above every functional head with an adverb in its specifier there is a verb lifter, $^{V}L^0$, 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, (48), is derived as follows. First, the VP with the main verb is formed, (56). Subsequently the adverb is merged in the specifier position of a dedicated functional head, F_{Adv}^0, the associated verb lifter, $^{V}L_3^0$, is merged, and the original VP moves to its specifier position, (57).

(56)   VP
         ⌒
      ordnet
      fixed

(57) $^{V}LP_3$

         VP
          ⌒
       ordnet $^{V}L_3^0$
       fixed $F_{Adv}P$

                     helt t_{VP}
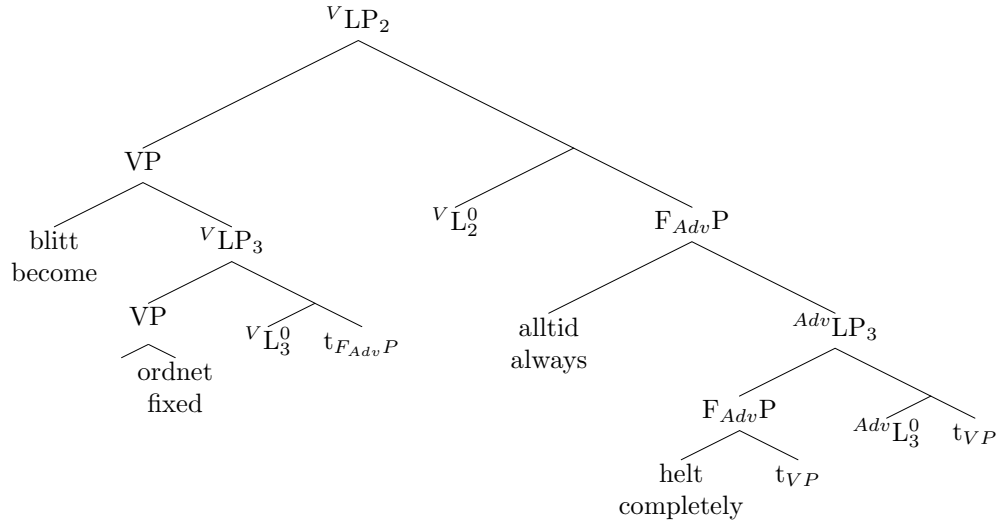   completely

The next steps consist in merging the auxiliary *blitt* - 'become' and its associated adverb lifter, $^{Adv}L_3^0$. Then $F_{Adv}P$, containing only the adverb *helt* - 'completely', moves to the specifier of $^{Adv}LP_3$, (58).

37

(58)

$^{Adv}LP_3$
- $F_{Adv}P$
  - helt / completely
  - t_{VP}
- $^{Adv}L_3^0$
- VP
  - blitt / become
  - $^{V}LP_3$
    - VP
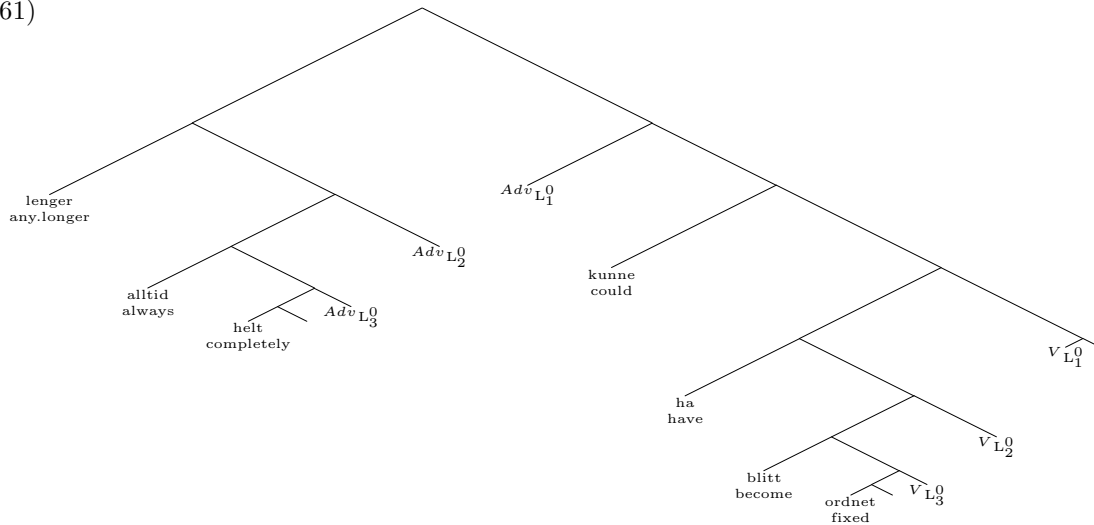   - ordnet / fixed
    - $^{V}L_3^0$
    - $t_{F_{Adv}P}$

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, (59). Then another verb merges, an adverb lifter merges on top of it, and the closest $F_{Adv}P$, containing *alltid helt* – 'always completely' moves, (60).

(59)   V-lifter

$^{V}LP_2$
- VP
  - blitt / become
  - $^{V}LP_3$
    - VP
      - ordnet / fixed
    - $^{V}L_3^0$
    - $t_{F_{Adv}P}$
- $^{V}L_2^0$
- $F_{Adv}P$
  - alltid / always
  - $^{Adv}LP_3$
    - $F_{Adv}P$
      - helt / completely
      - t_{VP}
    - $^{Adv}L_3^0$
    - t_{VP}

(60)

Adv_{LP_2}

$F_{Adv}P$

alltid
always

F^0_{Adv}

Adv_{LP_3}

$F_{Adv}P$

helt
completely

t_{VP}

$Adv_{L_3^0}$

t_{VP}

$Adv_{L_2^0}$

VP

ha
have

V_{LP_2}

VP

blitt
become

V_{LP_3}

VP

ordnet
fixed

$V_{L_3^0}$

$t_{F_{Adv}P}$

$V_{L_2^0}$

$t_{F_{Adv}P}$

After another cycle through the "adverb, verb lifter, verb, adverb lifter" sequence, the structure is as in (61): 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.[28]

(61)

lenger
any.longer

$Adv_{L_2^0}$

alltid
always

$Adv_{L_3^0}$

helt
completely

$Adv_{L_1^0}$

kunne
could

ha
have

blitt
become

ordnet
fixed

$V_{L_3^0}$

$V_{L_2^0}$

$V_{L_1^0}$

---

[28]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. There is then no empirical support for the adverb-cluster constituent hypothesized under the remnant movement analysis. For reasons of space I cannot pursue the issue further.
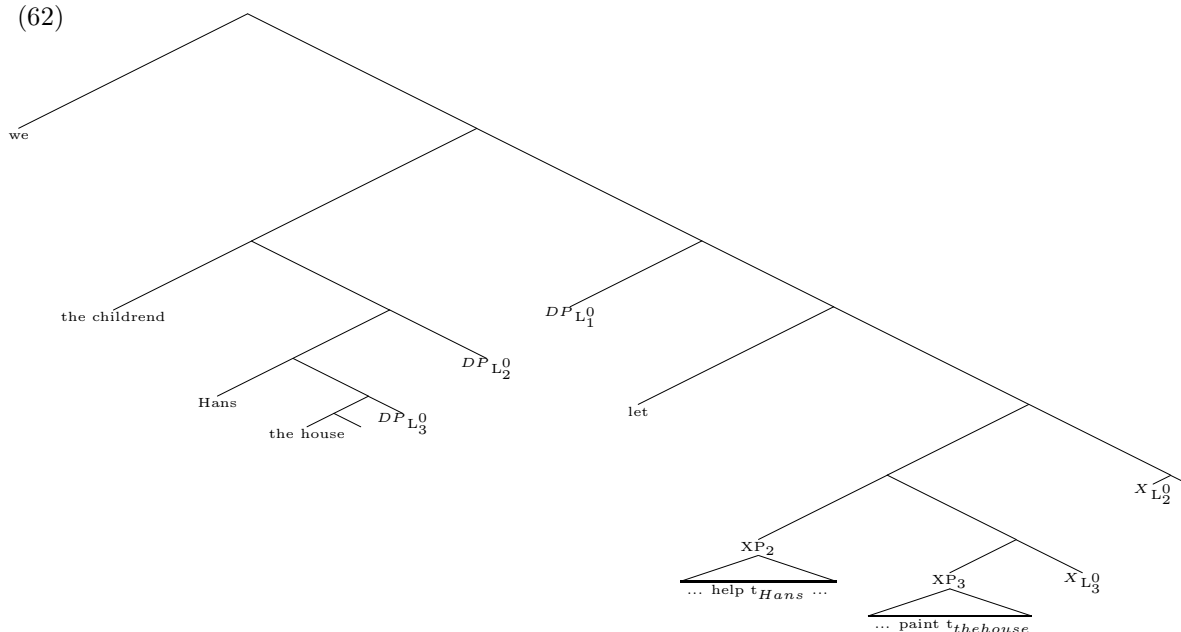
This is how the Nilsen-Bentzen system derives the cross-serial dependency between the adverbs and the verbs. The same type of analysis can be given for the Swiss German and Dutch cases, (62). All that is required for these languages are DP-lifters and XP-lifters, (62) below. The latter of which bring about the clustering of the verbs. The unbounded nature of the phenomenon is captured here by recursion through fixed sequences of identical lifters.

I will now show that the Nilsen-Bentzen approach to cross-serial dependencies is at odds with the assumptions argued for in section 2. We saw that all movement operations are ordered with respect to each other. If the Nilsen-Bentzen account of cross-serial dependencies were correct, then all postulated movement operations would have to find their place in the linear sequence of operations. This would have to be true, in particular, of the VP-movement operation in (57), (59), and (61), and the movement of the F_{Adv}Ps in (58), (60), and (61). In other words, assumption (11b) demands that VP-movement and F_{Adv}P-movement be ordered with respect to each other: VP-movement $<<$ F_{Adv}P-movement or F_{Adv}P-movement $<<$ VP-movement. Continuing the hypothetical argument, the step from (57) to (58) shows that VP-movement feeds F_{Adv}P-movement; hence, VP-movement $<<$ F_{Adv}P-movement. On the other hand, the step form (58) to (59) shows that F_{Adv}P-movement feeds VP-movement; hence, F_{Adv}P-movement $<<$ VP-movement. We end up with a violation of the asymmetric ordering of operations.[29] The contradiction is unavoidable. It is in principle impossible to harmonize the Nilsen-Bentzen account of cross-serial dependencies with the conjunction of (11a), (11b), and GenPIM. Given the evidence reviewed in section 2, this is a strong argument against the Nilsen-Bentzen approach to cross-serial dependencies. As we will see in a momement, the same problem plagues the analyses in Koopman and Szabolcsi (2000).

The problems that arose under (11a), (11b), and GenPIM for the analysis of Norwegian, arise again under the remnant movement analysis of Dutch and Swiss German (62). 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 of Norwegian 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: VP$_{ordnet}$ $<<$ F_{Adv_1} $<<$ VP$_{blitt}$ $<<$ F_{Adv_2} $<<$ VP$_{ha}$ $<<$ Fs_{Adv_3}... 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.

---

[29]Even if operations are only antisymmetrically ordered, F_{Adv}P-movement and VP-movement cannot be of the same type, because in that case Müller's generalization still rules out the derivation. As we saw, Müller's generalization follows from closest attract even if operations are only antisymmetrically ordered.

(62)

we

the childrend

$^{DP}\mathrm{L}_1^0$

Hans

$^{DP}\mathrm{L}_2^0$

the house

$^{DP}\mathrm{L}_3^0$

let

$^{X}\mathrm{L}_2^0$

XP$_2$

... help t$_{Hans}$ ...

XP$_3$

$^{X}\mathrm{L}_3^0$
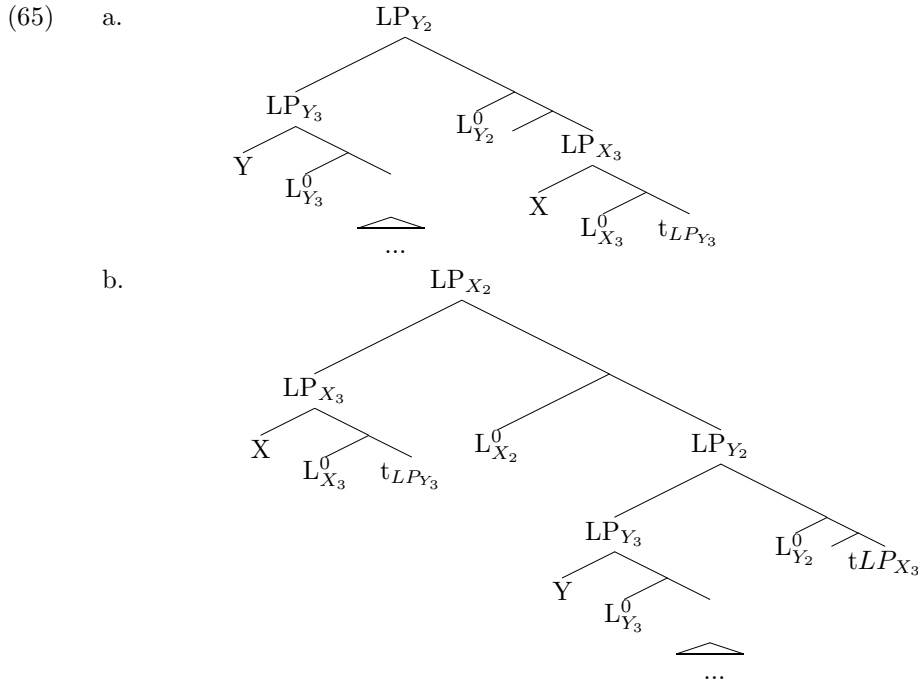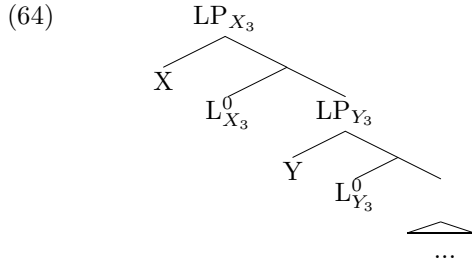
... paint t$_{thehouse}$

Let's consider the structures derived by the remnant movement account of cross-serial dependencies. A careful look at them reveals two cross-serial dependencies each. In (61) 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 $^{Adv}\mathrm{L}_n^0$ and $^{V}\mathrm{L}_n^0$ – an inverse one. This inverse cross-serial dependency is formed as a by-product in the derivation. Similarly, in (62) 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 $^{DP}\mathrm{L}_n^0$ and $^{XP}\mathrm{L}_n^0$ 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. The fact that the Nilsen-Bentzen approach cannot predict the linear asymmetry in (51) is a second argument against this approach.

I now turn to Koopman and Szabolcsi's (2000) theory of order preservation. It suffers from the same two problems as the Nilsen-Bentzen account of cross-serial dependencies: first it is incompatible with the conjunction of (11a), (11b), and GenPIM and second it generates inverse cross-serial dependencies as a derivational by-product.

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, (63a). Though the formulation is slightly cryptic, it is clear from context that Koopman and Szabolcsi intend (63a) 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 (63b) here.
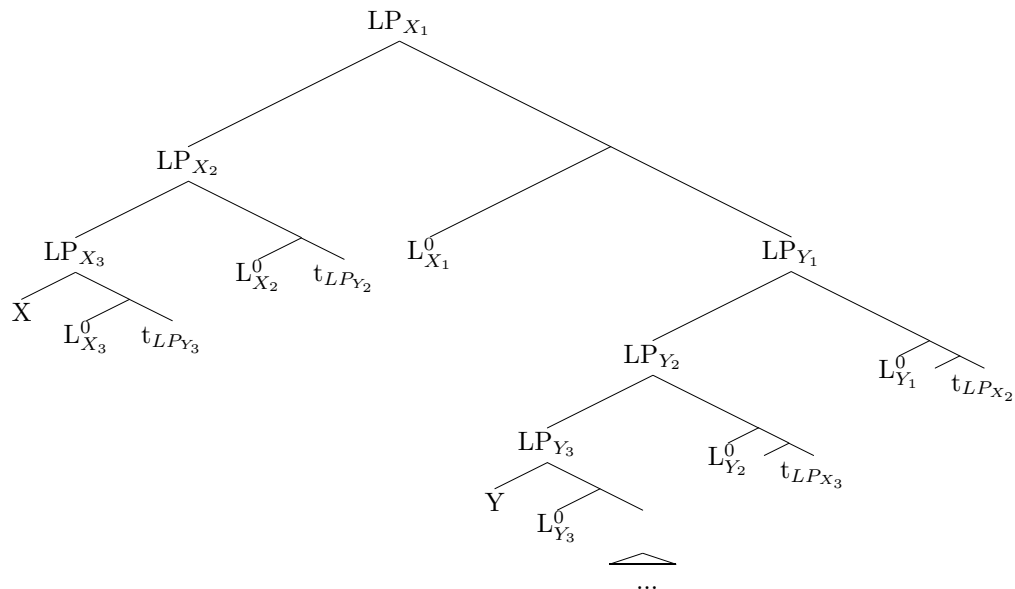
41

(63) 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 (64) will not usually move by themselves but as part of their containing categories LP_X and LP_Y. These categories split under movement by (63a). In a first step, LP_{Y_3} would move, as in (65a); then LP_{X_3} moves, as in (65b).

(64)

$$LP_{X_3}$$

X $L^0_{X_3}$ LP_{Y_3}

Y $L^0_{Y_3}$

...

(65) a.

$$LP_{Y_2}$$

LP_{Y_3} $L^0_{Y_2}$ LP_{X_3}

Y $L^0_{Y_3}$ X $L^0_{X_3}$ $t_{LP_{Y_3}}$

...

  b.

$$LP_{X_2}$$

LP_{X_3} $L^0_{X_2}$ LP_{Y_2}

X $L^0_{X_3}$ $t_{LP_{Y_3}}$ LP_{Y_3} $L^0_{Y_2}$ $t_{LP_{X_3}}$

Y $L^0_{Y_3}$

...

After two further steps of movement, of which there are many in Koopman and Szabolcsi's model, the structure in (66) emerges.

42

(66)

LP_{X_1}

LP_{X_2}

LP_{X_3} $L^0_{X_2}$ $t_{LP_{Y_2}}$ $L^0_{X_1}$ LP_{Y_1}

X $L^0_{X_3}$ $t_{LP_{Y_3}}$

LP_{Y_2} $L^0_{Y_1}$ $t_{LP_{X_2}}$

LP_{Y_3} $L^0_{Y_2}$ $t_{LP_{X_3}}$

Y $L^0_{Y_3}$

...

Two points are relevant here. First, as the careful reader will no doubt have noticed, the derivation currently under discussion violates the conjunction of (11a), (11b), and GenPIM – for the same reasons that the Nilsen-Bentzen derivations do. Second, (66) contains an inverse cross-serial dependency between the $L^0_{X_n}$'s and the $L^0_{Y_n}$'s. 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 are too unrestrictive even if they stick to the seemingly severe linear restrictions imposed by the LCA, as Koopman and Szabolcsi (2000); Nilsen (2003); Bentzen (2005) do.[30] This drives home the point made in the introduction: Without a restrictive theory of operations, the strict S-H-C hypothesis has no empirical content.

In the next section I show that the conjunction of (11a), (11b), GenPIM, and the assumption that movement is always leftward rules 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 (52b) onto (52c) will necessarily line up the moving elements in an order where scope corresponds to left-to-right order and never the other way around. The reason lies in the structural rather than linear definition of order preservation, (54a) above, together with the

---

[30]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.

assumption that movement is universally leftward. Another option might be a nesting derivation, but this is ruled out because Relativized Minimality imposes crossing rather than nesting. 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.
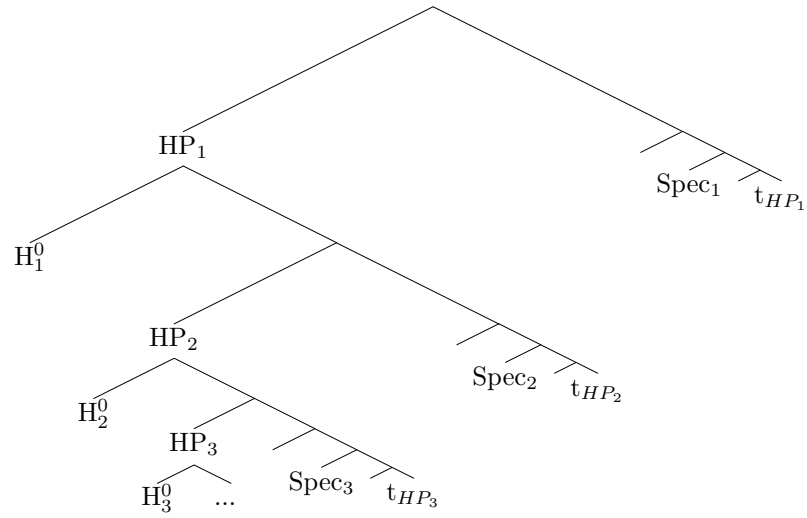
5.1   Head Movement

Head-to-head movement is problematic. Consider the structure in (67). 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 (67a) can be base generated. Alternatively, if we adopt the LCA, then right specifiers are ruled out and must be derived through roll-up movement as in (67b) (see (2b) above). To make the structure in (67b) more easily readable, functional heads that are required by Kayne's LCA but irrelevant to the point under discussion are indicated as empty branches.
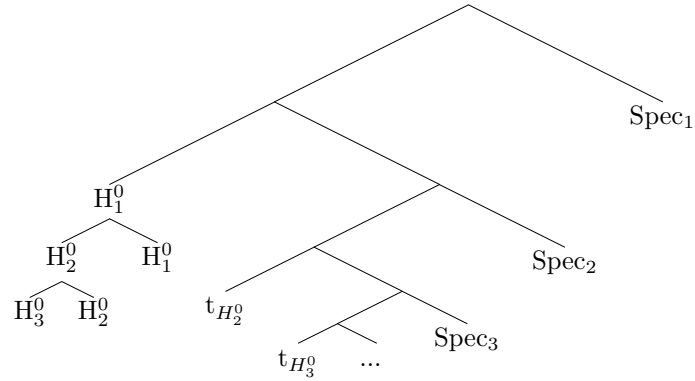
(67)   a.

b.

HP_1 H_1^0 HP_2 H_2^0 HP_3 H_3^0 ... $Spec_3$ t_{HP_3} $Spec_2$ t_{HP_2} $Spec_1$ t_{HP_1}
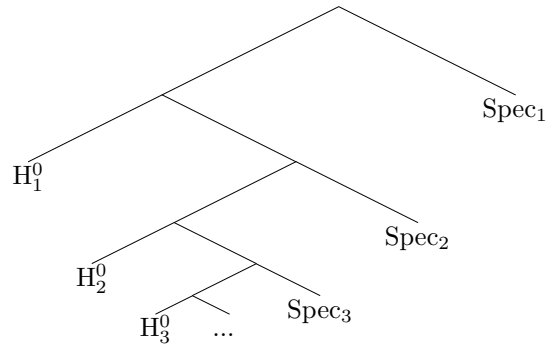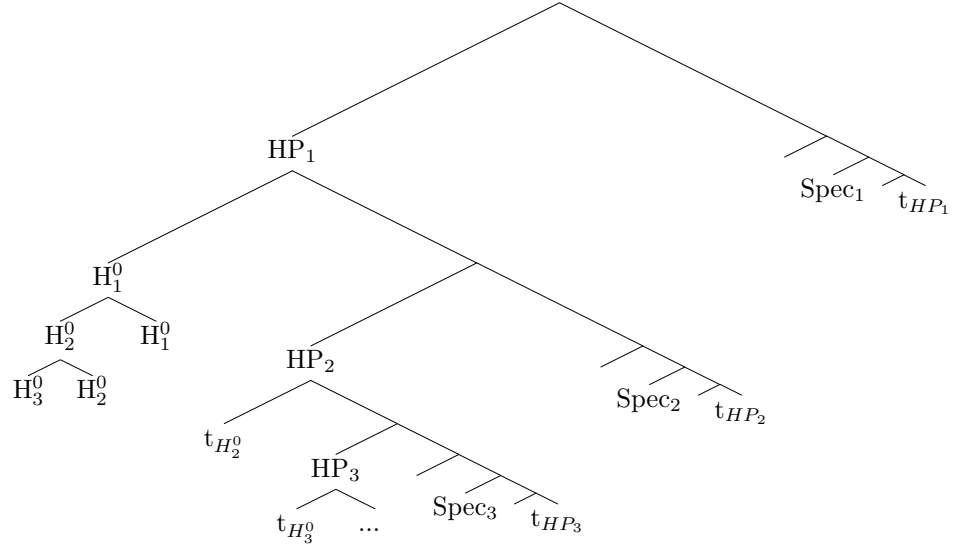
Standard (left adjoining) head movement derives (68a) with an inverse cross-serial dependency from (67a). Similarly, head movement derives (68b) from (67b). If (67b) is to be ruled out, independent assumptions about locality have to be invoked. This is not impossible, but it shifts the explanatory burden from the strict S-H-C hypothesis even further to the theory of movement.

(68)   a.

H_1^0 H_2^0 H_1^0 H_3^0 H_2^0 $t_{H_2^0}$ $t_{H_3^0}$ ... $Spec_3$ $Spec_2$ $Spec_1$

45

b.

HP_1

H_1^0

H_2^0 H_1^0

H_3^0 H_2^0

HP_2

$t_{H_2^0}$

HP_3

$t_{H_3^0}$ ...

$Spec_3$ t_{HP_3}

$Spec_2$ t_{HP_2}

$Spec_1$ t_{HP_1}

Since inverse cross-serial dependencies are universally unattested, the derivations in (68) must be ruled out. These facts might then be interpreted as a new empirical argument against head movement as head adjunction (see Hinterholzl (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). The easiest way of banning head movement in the sense of head adjunction is to demand that all movement obey the extension condition.
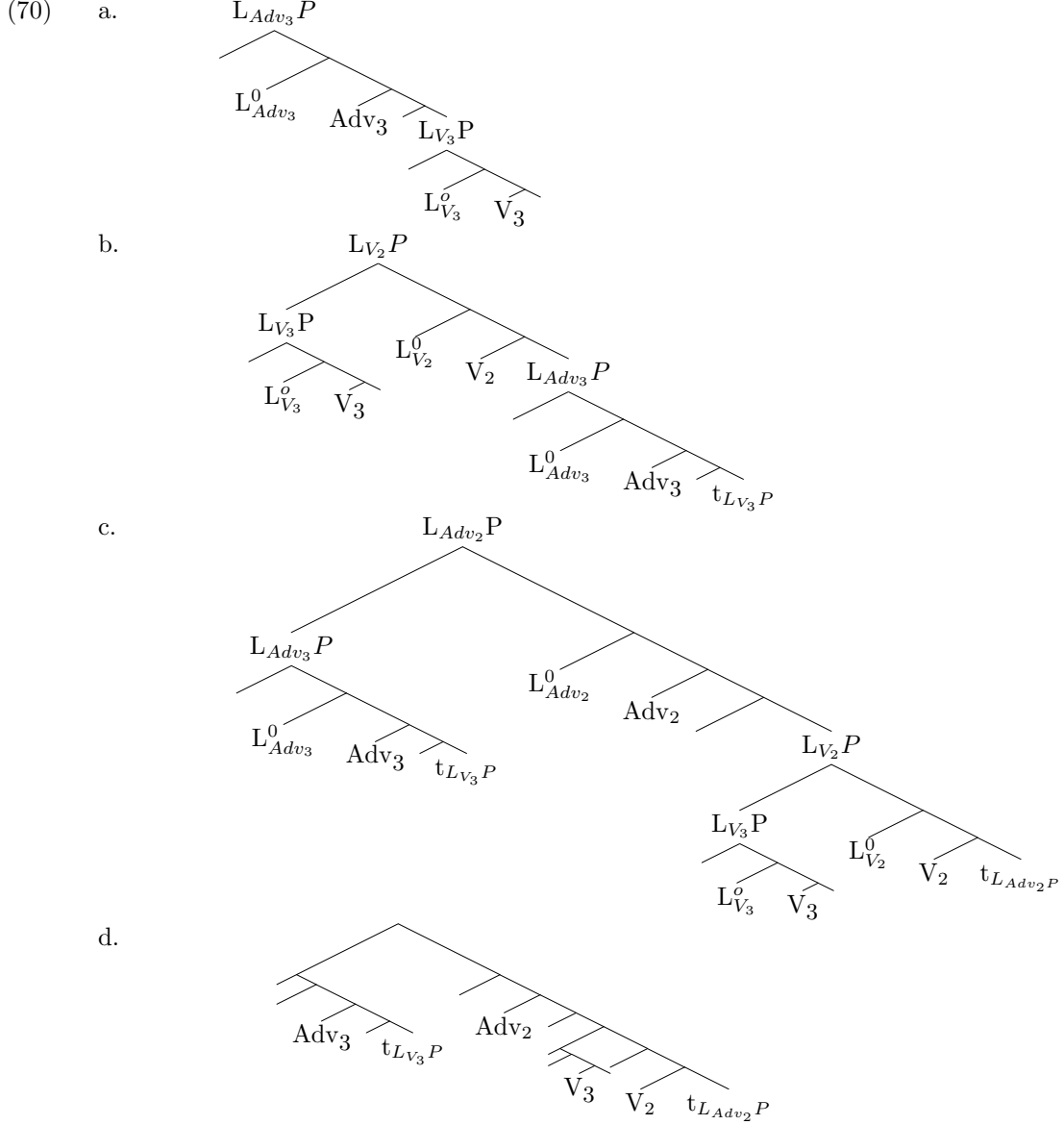
5.2   Remnant Movement

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 (52). 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.

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.

(69)   Inverse Nilsen/Bentzen

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

These assumptions give rise to derivations like those depicted in (70a-c). Removing unnecessary clutter from (70c) yields (70d), 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.

(70) a.

$L_{Adv_3}P$
- $L^0_{Adv_3}$
- Adv_3
- $L_{V_3}P$
  - $L^o_{V_3}$
  - V_3

b.

$L_{V_2}P$
- $L_{V_3}P$
  - $L^o_{V_3}$
  - V_3
- $L^0_{V_2}$
- V_2
- $L_{Adv_3}P$
  - $L^0_{Adv_3}$
  - Adv_3
  - $t_{L_{V_3}P}$

c.

$L_{Adv_2}P$
- $L_{Adv_3}P$
  - $L^0_{Adv_3}$
  - Adv_3
  - $t_{L_{V_3}P}$
- $L^0_{Adv_2}$
- Adv_2
- $L_{V_2}P$
  - $L_{V_3}P$
    - $L^o_{V_3}$
    - V_3
  - $L^0_{V_2}$
  - V_2
  - $t_{L_{Adv_2}P}$

d.

- Adv_3
- $t_{L_{V_3}P}$
- Adv_2
- V_3
- V_2
- $t_{L_{Adv_2}P}$

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. If the operations moving $L_V P$ and $L_{Adv}P$ exist at all, then they fall under the perview of (11a) and (11b), that is, the two operations must be ordered with respect to each other. But once the operations are ordered with respect to each other, the symmetric feeding relations between them are disallowed by GenPIM. Thus, even if such operations exist, they can never give rise to inverse cross-serial dependencies.

5.3 Conclusion

If the empirical hypothesis from section 3 is true and UG disallows inverse cross-serial dependencies, then syntactic theory must account for this. With assumptions (71-i-iii) in place, we can rule out all derivations for inverse cross-serial dependencies except for the remnant movement derivation.

(71) (i)   Movement obeys the extension condition.
      (ii) Movement is uniformly leftward.
      (iii) Movement operations are order preserving in the sense of (53).

If we add to these assumptions the assumption that movement operations are ordered, (11a) and (11b), and apply this ordering as in GenPIM, (13), then inverse cross-serial dependencies are correctly predicted to be impossible. I take this section of the paper to be an informal proof-by-case of this claim.

6   Conclusion

I opened the paper with the question of 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 that they are insufficient. I pointed out that giving up the Proper Binding Condition and the freezing principle runs the risk of robbing the LCA of its empirical content. A restrictive theory of 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. I took this as a phenomenon rather than an epiphenomenon and formulated the assumption as UCOOL. UCOOL, however, is meaningless without an instruction of how to deploy it. I argued on the basis of mostly German data that assumptions at least as strong as (11) and GenPIM are necessary. The universal claim made in UCOOL remains to be verified. The system yields a flexible, fine-grained, and apparently correct framework for the analysis of phenomena in various domains: improper movement, effects of the freezing principle and exceptions from it, and remnant movement with immediate predictions of what is possible and what is impossible.

It emerged, table (36), that there is a nearly perfect match between the three conditions unified under GenPIM (identity, surfing paths, diving paths). It also emerged that the ordering of operations does seem to be a(nti)symmetric and total. This suggests that a whole range of important issues that were set aside in section 2 warrant further study: Is UCOOL universally

fixed? 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? ...

Section 3 reviewed some well-known facts regarding cross-serial dependencies in Dutch and Swiss German. There exists a hitherto unnoticed linear asymmetry in this empirical domain: Inverse cross-serial dependencies are impossible, (51). I suggested a novel analysis of forward cross-serial dependencies, 4.1, and went on to show in sections 4.2 and 5 that theories which derive cross-serial dependencies through remnant movement are incompatible with the main assumptions argued for here and fail to predict the impossibility of inverse 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); Hinterholzl (1997)) can be captured under a more restrictive theory of operations, such as the one proposed here.

This paper has been somewhat programmatic. I had to raise many issues that I could not resolve in the space of this paper and mention constructions that I could not discuss. These issues and constructions include the right-roof constraint and the ban against unbound intermediate traces in remnants, *tough*-movement, clefts, and relative clauses. Other potentially relevant constructions have not even been mentioned. It might be interesting, for example, to consider ECM under *wh*-movement in Romance languages and the locality bounds on control into adjuncts from the current perspective. Both constructions might yield an answer to the question whether the ordering of operations is asymmetric or antisymmetric. I am planning to take up these questions in future work.

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 Peter Muriungi. 2006a. The focus particle in kîîtharaka. In *Papers on information structure in African languages*, ed. Ines Fiedler and Anne Schwarz, volume 42 of *ZAS Papers in Linguistics*, 1–20. ZAS.

Abels, Klaus, and Peter Muriungi. 2006b. The focus particle in Kîîtharaka: Syntax and semantics.

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

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

Adger, David, Daniel Harbour, and Laurel Watkins. 2007. Feature intervention, information structure and word order. In *GLOW Newsletter # 58*.

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

Bach, Emmon. 1971. Questions. *Linguistic Inquiry* 2:153–166.

Bayer, Josef. 1996. *Directionality and logical form. on the scope of focussing particles and wh-in-situ.*. Dordrecht: Kluwer.

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. 1987. Remnant topicalization and the constituent structure of vp in the germanic sov languages. *GLOW Newsletter* 18:15–16.

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. 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, and Daiko Takahashi. 1998. Scrambling and last resort. *Linguistic Inquiry* 29:347–366.

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. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. Stephen R Anderson and Paul Kiparsky, 232–286. New York: Holt, Rinehart and Winston.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

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. 1995. 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.

Chomsky, Noam. 2004. On phases, December 2004.

Chomsky, Noam, and Howard Lasnik. 1995. *The minimalist program*, chapter The Theory of Principles and Parameters, 13–127. Cambridge, Mass: 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. 1997. *Local economy*. Cambridge, Mass.: MIT Press.

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.

Epstein, Samuel David. 1999. Un-principled syntax and the derivation of syntactic relations. In *Working minimalism*, ed. Samuel David Epstein and Norbert Hornstein, 317–345. Cambridge, MA: MIT Press.

Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima, and Hisatsugu Kitahara. 1998. *A derivational approach to syntactic relations*. Oxford: Oxford University Press.

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. Gunther Grewendorf and Wolfgang Sternefeld, 113–142. John Benjamins.

Fanselow, Gisbert. 1991. Minimale syntax. Habilitationsschrift, Universität Passau.

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.

Frey, Werner. 2000. über die syntaktische position der satztopiks im deutschen. *ZAS Papers in Linguistics* 20:137–172.

Frey, Werner. 2005. Pragmatic properties of certain German and English left peripheral constructions. *Linguistics* 43:89–129.

Glushan, Zhanna. 2006. Japanese style scrambling in Russian: myth and reality. Master's thesis, Universitetet i Tromsø.

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. 1994. Variable binding and reconstruction.

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.

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.

Hinterholzl, 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.

Hornstein, Norbert. 2001. *Move! A minimalis theory of construal*. Cambridge, USA: Blackwell.

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.

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.

Kayne, Richard. 2005. chapter Antisymmetry and Japanese.

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

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. 2007. The emergence of syntactic structure. *Linguistics and Philosophy* 30:47–95.

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 *Tubingen 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, 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.

May, Robert. 1979. Must comp-to-comp movement be stipulated? *Linguistic Inquiry* 10:719–725.

Müller, Gereon. 1995. *A-bar syntax: A study of movement types*, volume 42 of *Studies in Generative Grammar*. New York: Mouton de Gruyter.

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

Müller, Gereon. 1999. Imperfect checking. *The Linguistic Review* 16:359–404.

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 Kiitharaka. *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.

Pafel, Jürgen. 1993. Ein überblick über die extraktion aus nominalphrasen im deutschen. In *Extraktion im deutschen i*, ed. Franz-Josef d' Avis, Sigrid Beck, Uli Lutz, Jürgen Pafel, and Susanne Trissler, Arbeiten des Sonderforschungsbereichs 340 34, 192–245. Universität Tübingen, Universität Stuttgart.

Partee, Barbara Hall, Alice ter Meulen, and Robert A. Wall. 1993. *Mathematical methods in linguistics*. Dordrecht, Boston, London: Kluwer Academic Publishers.

Pearson, Matthew. 2000. Two types of vo languages. In *The derivation of vo and ov*, ed. Peter Svenonius, Linguistik Aktuell, 327–364. Amsterdam and Philadelphia: John Benjamins.

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. 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. 1971. *Cross-over phenomena*. New York: Holt, Rinehart, and Winston.

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.

Ramchand, Gillian. to appear. *Verb meaning and the lexicon: A first phase syntax*. Cambridge University Press.

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. Locality and left periphery. In *Structures and beyond: Cartography of syntactic structures*, ed. Adriana Belletti, volume volume 3, 104–131. New York: Oxford University Press.

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, Luigi. 2006. On the form of chains: Criterial positions and ECP effects. In *Wh-movment: moving on*, ed. Lisa Lai Shen Cheng, 97–134. Cambridge, Mass.: MIT Press.

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.

Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT.

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. 1997. Derivational minimalism. In *Logical aspects of computational linguistics*, ed. Retore, 68–95. Springer.

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.

Ura, Hiroyuki. 1994. Varieties of raising and the feature-based bare phrase structure theory. *MIT Occasional Papers in Linguistics* 7:1–150.

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.

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. 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.

Address:
Dept. of Linguistics, CASTL
University of Tromsø
9037 Tromsø Norway

and
Department of Phonetics and Linguistics
University College London
Gower Street
London WC1E 6BT United Kingdom