

Some implications of improper movement for Cartography*

Klaus Abels
Universitetet i Tromsø – CASTL
University College London
klaus.abels@gmx.net

Draft 1.0 – Please do not distribute
November 22, 2007

Abstract

This paper briefly reviews the theory of movement from Abels (to appear). I start with the observation that my Universal Constraint on the Ordering of Operations in Language is massively redundant if not isomorphic to existing cartographic hierarchies for phrase structure. Thus, operational ordering can be viewed as a new source of evidence for or against particular cartographic proposals. Given the redundancy between the two hierarchies, I then investigate a number of possible ways of reducing my theory of movement to cartography. The failure of these attempts is instructive in that they bring out an apparently irreducible difference between external and internal merge. They also point to a rather abstract conception of cartographic hierarchies.

*This paper owes its existence to Edwin Williams' work (1974; 2002). I would also like to thank audiences in Tromsø, Leipzig, London, Geneva, Konstanz, and Bruxelles for the chance to present and discuss related material in various forums between 2005 and 2007.

The goal of this paper is to discuss the relation between structural, in particular cartographic, hierarchies and the ordering of movement operations. The ban on improper movement is the simplest example of a relevant ordering relation between movement operations. The paper makes two claims. First, the cartographic hierarchy of projections and the hierarchy of operations are massively redundant with each other. Despite this redundancy, the hierarchy of operations cannot be reduced to the structural hierarchy. The second claim is that therefore cartographic hierarchies cannot simply be structural primitives, they cannot be understood as structural building blocks, ‘lego blocks’ whose construction and internal make-up need not concern the derivation of sentences.

Another way of phrasing the second claim is that the hierarchies discovered by cartographic research must express very abstract ordering principles. These principles are accessed by the grammar for at least two purposes: structure building by external merge and structure building by internal merge/move. Both ways of merging material are constrained by the ordering principles, but in somewhat different ways.

The paper is organized as follows. In section 1 I briefly review some facts concerning the ordering of operations as discussed in Abels (to appear). In section 2 I discuss various ways one might try to capture the relevant generalizations without invoking an abstract version of the cartographic hierarchy as a constraint on the ordering of operations and show that and why they fail. A brief conclusion follows.

1 On the notion of Improper Movement

The ban on improper movement as usually understood (classic references are Chomsky (1973, p. 243 fn. 24, p. 253 # 110c); May (1979); Chomsky (1981, p. 195-204)) has the effect of ruling out any derivation in which A-movement applies to constituent C after \bar{A} -movement has applied to C . This formulation makes a number of explicit and implicit claims. The formulation assumes, for example, that there is a binary distinction between types of movement operations, A- vs. \bar{A} -movement (Postal, 1971). It assumes identity between the constituents that undergo consecutive movements; the same constituent undergoes both movements. The ban crucially applies across cycles; movement of C to

COMP on the lower cycle must not be followed by movement to [NP; S] on the higher cycle. Finally, operations are – somewhat trivially – ordered linearly.

We can ask whether these assumptions are all equally well motivated empirically and whether all of them are equally necessary to derive the desired result. It turns out that not all of them stand up to scrutiny equally well. The next subsections therefore evaluate the empirical assumptions behind the standard understanding of the ban on improper movement in the order in which these assumptions were just introduced.

1.1 Are there really only two types of movement?

The assumption that an adequate description of movement phenomena or even the feeding/bleeding relations between movements can be given on the basis of just a binary distinction between A- and \bar{A} -movement is dubious. Finer grained distinctions are frequently invoked in the literature (see for example Williams (1974, 2002); Cinque (1990); Sternefeld (1992); Müller and Sternefeld (1993); Bošković (1994); Hornstein (1999a); Starke (2001); Friederici et al. (2003); Abels (to appear)). Even the staunchest defenders of a unified view of movement regularly face the problem of having to assume derivational and/or representational machinery that introduces distinctions between types of movement through the back door. A prominent case in point is Lasnik and Saito's (1984) distinction between γ -marking of moved arguments and moved adjuncts, which happens at S-structure and at LF respectively.

Postal (1971) introduces the A-/ \bar{A} -movement dichotomy on the basis of cross-over effects, that is, the possibility or impossibility of establishing a binding relation from a derived position. Standard cases of A-movement, (1a), allow binding of a derived co-argument and binding into a derived clausemate argument, while standard cases of \bar{A} -movement, (1c), allow neither. However, (1b) shows that this does not exhaust the space of possibilities. There are movement processes that, like A-movement, allow binding into a derived clausemate argument but that, like \bar{A} -movement, disallow binding of such derived clausemate arguments. Example (1b) illustrates this for English on Postal's (1971) assumption that the three examples in (1b) are related by a movement operation. So-called A-scrambling in Japanese famously ex-

hibits the binding pattern in (1b), *mittelfeld* scrambling in German does, too, and so does clause-internal scrambling in many of the Slavic languages.

- (1) 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.

The paradigm in (1) can be reproduced with quantifier bound pronouns instead of reciprocals. To see this, simply replace the reciprocal in the examples by a pronoun (or a reflexive, where this is required by the binding theory) and the moving element by a universal quantifier. Examples (1c.ii) and (1c.iii) must then be replaced by a restrictive relative clause headed by a universally quantified noun. The pattern found with reciprocals and quantifier-bound pronouns point to the same conclusion. A-movement allows binding of both types from the derived position, \bar{A} -movement disallows both types. But, there is a third type of movement which allows binding *into* clausemate arguments from the derived position but disallows binding *of* clausemate arguments from the derived position. This shows clearly that even to deal with the binding facts, more than two types of movement need to be distinguished. I will show that the same conclusion can be based on feeding/bleeding relations.¹

¹To the best of my knowledge, a movement operation with the opposite binding profile of (1b) has not been attested.

1.2 Is identity necessary for the ordering to apply?

The second assumption inherent in the paraphrase of the ban against improper movement that opened this section was that one and the same constituent undergoes both movements. Like the dichotomous partitioning of movement into A- and \bar{A} -movement, this assumption has been called into question before (Sakai (1994); Williams (2002); Grewendorf (2003); Abels (to appear)). In addition to the standard case of identity, figure 1, I will discuss two configurations here: extraction from a moved constituent, figure 2, and remnant movement, figure 3.

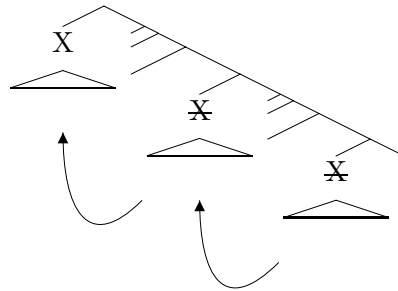


Figure 1: The standard case – identity

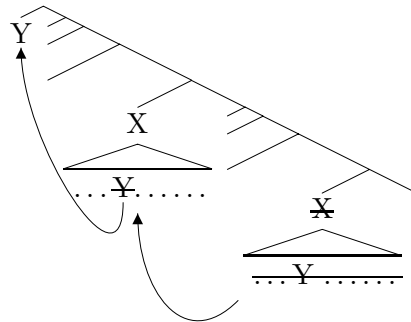


Figure 2: Subextraction – surfing path, descendant moves second

The derivations in figures 2 and 3 used to be thought to be ruled out by the freezing principle and the proper binding condition, respectively. However, there is increasing consensus that both of those principles in their original formulations were too strong. Both of the derivations in figures 2 and 3 involve two movements. In figure 2, a large constituent

- c. (i) ✓It is known [how likely t_{Oscar}^A to win] Oscar is $t_{\text{how likely } t_{\text{Oscar}}^A \text{ to win}}^{wh}$
(ii) *[A picture of $t_{\text{which king}}^{wh}$] is known which king to have
been sold $t_{\text{a picture of which king}}^A$

In Abels (to appear) I argue that this triple of contrasts is not accidental. It systematically reflects the ordering of movement operations relative to each other. The same triplicated asymmetry can be found for example when the application of mittelfeldscrambling to a pre-subject position in German is compared to *wh*-movement. I sketch the argument here.

Making the work a bit difficult is the fact that *wh*-phrases do not usually scramble in front of the subject in German, (3a) (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)). This could be taken as an argument that mittelfeldscrambling does not in fact feed *wh*-movement. However, under certain conditions, discussed in Beck (1996, 6-7), Sauerland (1996), Fanselow (2001, 2004), it becomes possible, (3b). I will assume that case like (3b) show that *wh*-phrases can in principle undergo mittelfeld scrambling. To address the question whether mittelfeld scrambling may feed *wh*-moved or the other way around, we need to introduce a few additional facts concerning these operations.

- (3) 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?

First, mittelfeld scrambling, unlike *wh*-movement is clause-bounded, as (4) shows. Second, mittelfeld scrambling does not give rise to weak-crossover effects, (4a). In fact, it has the binding profile of (1b) above. Clear cases of *wh*-movement on the other hand are subject to weak-crossover, (4b).

- (4) a. *Gestern hat {jeden | diesen} Studenten Hans
 Yesterday has every.ACC this.ACC student.ACC Hans
 gesagt, dass seine Eltern unterstützen
 said that his parents $t_{\{every|this\} student}$ support
- b. Welchen Studenten glaubt Hans, dass Maria
 which.ACC student.ACC believe Hans that Maria
 unterstützt?
 supports
 Which student does Hans believe that Maria supports?
- (5) a. (i) *Früher haben seine_k Eltern jeden
 earlier have his parents every.ACC
 Studenten_k unterstützt.
 student.ACC supported
 In the past his_k parents supported every student_k.
- (ii) ✓Früher haben jeden Studenten_k seine_k
 earlier have every.ACC student.ACC his
 Eltern unterstützt.
 parents supported
 In the past his_k parents supported every student_k.
- b. (i) *Welchen Studenten_k glauben seine_k Eltern,
 which.ACC student.ACC believe his parents
 dass Maria unterstützt?
 that Maria supports
 Which student_k believe his_k parents that Maria sup-
 ports?

Curiously, unlike the long instance of *wh*-movement in (5b), short *wh*-movement does not give rise to weak crossover effects. Likewise, long *wh*-movement is not subject to the weak-crossover constraint within the clause where the *wh*-phrase originates. This is illustrated in (6a) and (6b). These facts suggests strongly that scrambling feeds *wh*-movement (Wiltschko (1997, 1998), also Fanselow (2004); Pesetsky (2000); Grewendorf (2001) for discussion). Thus, examples like (6) involve an initial step of mittelfeld scrambling; binding is possible from the scrambled position; then, *wh*-movement takes place.

- (6) a. ✓Welchen Studenten_k unterstützen seine_k Eltern?
 which.ACC student.ACC support his_k parents?

- Which student_k do his_k parents support?
- b. ✓ Welchen Studenten_k glaubt Maria, dass seine_k
 which.ACC student.ACC believes Maria that his_k
 Eltern unterstützen?
 parents support?
 Which student_k does Maria believe that his_k parents sup-
 port?

While the binding facts above lead to the conclusion that mittelfeld-scrambling may feed *wh*-movement in the identity case, *wh*-movement never feeds scrambling. The assumption that *wh*-movement does not feed scrambling is necessary to account for the status of (5bi) above. We need to rule out the derivation where *welchen Studenten* undergoes one step of *wh*-movement to the embedded [Spec, CP], then scrambling to enable binding, and then another step of *wh*-movement to [Spec, CP] in the matrix.² A less indirect case is given in (7). Likewise with remnant movement.

- (7) *Gestern hat welchen Studenten nur ein Meisterdetektiv
 yesterday has which student only a master detective
 wissen können, Hans beim Schummeln erwischt hat.
 know could Hans at.the cheating caught has
 Yesterday only a world-class detective could have known which
 student Hans caught cheating.

The same asymmetry is found for surfing paths. While a mittelfeld scrambled constituent can be *wh*-moved out of, (8a-b), a *wh*-moved constituent cannot be scrambled out of, (8c). A remnant created by mittelfeld scrambling may *wh*-move, (9a), but a remnant created by *wh*-movement may not mittelfeld scramble, (9b).³

- (8) a. Worüber kann einen Südkurier-Artikel
 what-about can [a Südkurier-article t_{worüber}]
 selbst Peter nicht am Strand verfassen?
 even Peter not at.the beach write
 For which topic is it the case that even Peter cannot write

²Thanks to Tarald Taraldsen for pointing this out.

³Example (8c) is independently ruled out by the locality of scrambling.

an article about it for the Südkurier when he is at the beach. (from (Kuthy and Meurers, 1999, p. 27) attributed to Fanselow (1991))

- b. Über welchen deutschen Kaiser hatte ein fertiges
 about which German emperor had [a done
 Manuskript leider keiner der anwesenden
 manuscript $t_{PP}]_{NP}$ unfortunately none the present
 Historiker anzubieten.
 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,
 yesterday has [about Charlemagne] $_{PP}$ nobody asked
 was für eine Arbeit er schreiben soll
 what for a work t_{PP} he write should
 Nobody asked yesterday what kind of paper about Charle-
 magne he should write.
- (9) a. Was für ein Buch hat über die Liebe
 [what for a book $t_{PP}]_{DP}$ has [about the love] $_{PP}$
 niemand gelesen.
 nobody t_{DP} read
 What type of book about love has nobody read? (based on
 Müller (1998, p. 221-222))
- b. *Auf dem Kongress hat über die Liebe niemand gewusst
 on the congress has about the love nobody known
 was für ein Buch er empfehlen soll.
 what for a book he recommend should
 At the congress nobody knew what kind of book about love
 he should recommend.

We have seen that in all three conditions, scrambling feeds *wh*-movement but not the other way around. For English we saw that the same was true for the pair of A-movement and *wh*-movement. These observations lead to the conjecture, that the same ordering of operations must be imposed under all three conditions. Identity is not an essential part of the ban on improper movement.

1.3 Linear order across cycles!

The facts reviewed so far suggest that the dichotomy of movement operations might be too coarse and that ordering of operations need to be imposed not only in the identity case, but the same ordering needs to be imposed on the case of extraction from a moved constituent and on remnant movement. Both of these claims are empirically underscored by the data in table 1, taken from Abels (to appear), which shows the feeding and bleeding relations between four different operations.

		A-mvt	Scrambling	<i>wh</i> -movement	Topicalization
A-mvt	identity		✓	✓	✓
	surfing path		✓	✓	✓
	diving path		?	✓	✓
Scrambling	identity	*		✓	✓
	surfing path	*		%	%
	diving path	*		✓	✓
<i>wh</i> -mvt	identity	*	*		*
	surfing path	*	*		??
	diving path	*	*		??
Topicalization	identity	*	*	*	
	surfing path	*	*	*	
	diving path	*	*	*	

Table 1: Interactions of four types of movement

Importantly, feeding is transitive and asymmetric, and total. We are thus dealing with a linear order of four different movement types across all three conditions. It remains crucial that the order is applied globally across cycles (phases, CPs, etc.), otherwise *wh*-movement on one cycle might illicitly end up feeding A-movement on the next cycle up.

In Abels (to appear) I propose to treat these facts by introducing the assumptions in (10).

- (10) a. Universal Constraint on the Ordering of Operations in Language (UCOOL)
 $\theta^4 \prec \dots \prec \text{Case} \prec \dots \prec \text{Scrambling} \prec \dots \prec wh \prec \dots \prec \text{topicalization} \prec \dots$
 (If $\alpha \prec \beta$, we say that α precedes β .)
- b. Generalized Prohibition against Improper Movement (Gen-PIM)
 No constituent may undergo movement of type τ if it has been affected by movement of type σ , where $\tau \prec \sigma$ under (10a).
- c. A constituent α is *affected* by a movement operation iff
- i. α is reflexively contained in the constituent created by movement, and
 - ii. α is in a (reflexive) domination relation with the moved constituent.

Condition (10a) is a statement of the ordering of movement operations. I hope that future work will validate the claim that this ordering is indeed universal. Given that the classic ban on improper movement appears to hold universally, there is hope that this more fine-grained hierarchy will turn out to be universal, too. Condition (10b) is a generalization of the ban on improper movement to the three conditions discussed above. The standard formulation, when coupled with the fine-grained hierarchy of movement types in (10a), would posit that

no constituent may undergo movement of type τ if it *has undergone* movement of type σ , where $\tau \prec \sigma$ under (10a).

The difference between the standard formulation and (10b) lies in the fact that the effects of movement are extended from the moving constituent to all those constituents *affected* by that movement. The notion of affectedness is defined in (10c).

Figure 4 illustrates how the definition of affectedness works. In all cases the circled node X is affected by α -movement of the moving

⁴I include θ relations here. If they are only ever established in the position of first merge, then no interesting feeding/bleeding relations will be observed. However, if θ -role assignment to derived positions is possible, as claimed in Bošković (1994); Hornstein (1999b); Boeckx and Hornstein (2003) a.o., then we can ask whether movement to θ -positions precedes or follows other movement types. The answer seems to be that movement to θ -positions precedes all other types of movement.

constituent. Figure 4a) is the case of identity. The moving constituent X is affected by its own movement, since it is contained in the structure created by movement and since X is in a reflexive domination relation with itself. Figure 4b) shows that all subparts of a moving constituent are affected by the movement. Such subparts are clearly contained in the target of movement and are dominated by the moving constituent. This case is relevant for extraction out of moved constituents, i.e., diving paths. Finally, figure 4c) shows that ancestor nodes of a moving constituent in the pre-movement tree are affected by the movement. They are contained in the constituent created by movement and they dominate the moving constituent in its original position. This is the case relevant for remnant movement.

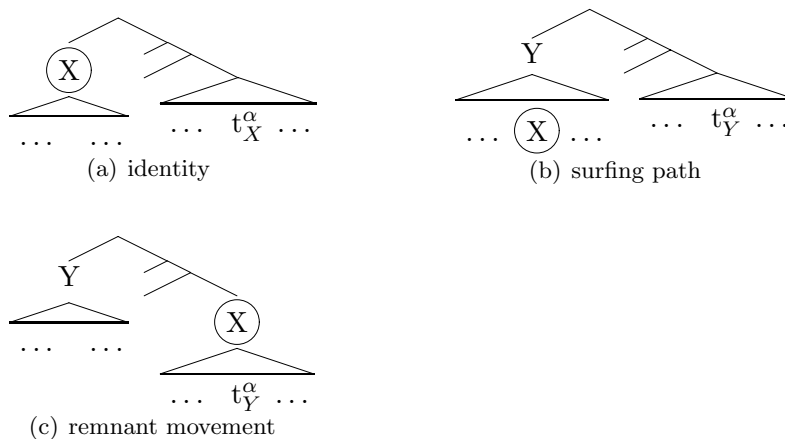
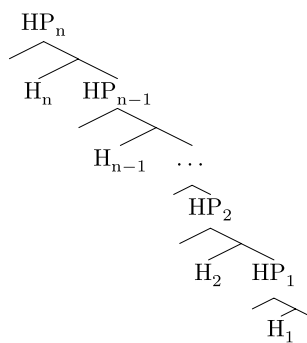


Figure 4: Affectedness

How the system incorporating UCOOL and GenPIM works is discussed in some detail in Abels (to appear). What is important for the present discussion is that the system captures the twofold generalization of the ban on improper movement. I will now turn to a discussion of the relation between UCOOL and cartographic hierarchies.

2 UCOOL and cartographic hierarchies

Even the most superficial look at UCOOL, (10a), reveals that is massively redundant with existing cartographic hierarchies. Within the clausal structure, the θ -domain is at the bottom, then there is the case layer in the IP-domain, finally UCOOL faithfully reproduces the fine-structure of the CP in many languages, with topics above the landing site of *wh*-movement. Although further reserach might reveal discrepancies between UCOOL and the universal cartographic template, I will assume here that they are, in fact, the same hierarchy. We should then ask whether GenPIM and the notion of *affectedness* can be dispensed with. That is, assume that UCOOL is identical to the cartographic hierarchy, can the effects of the rest of (10) be derived by relating movement/feature checking with structure in a simple and direct way? An even more radical approach might be to try to derive the effects of (10) without invoking the cartographic template=UCOOL as an independent constraint, independent, that is, from the phrase structures projected according to it. The answer will be that both attempts fail. There is an apparently irreducible difference between external merge and internal merge/move. While external merge is constrained by the cartographic template=UCOOL per cyclic domain, internal merge is constrained across cyclic domains: the θ domain is ordered below the case domain within a clause, but across clauses the θ -domain of a higher clause may come to be structurally higher than the case domain of a lower clause. Movement on the other hand always proceeds from θ -position to case position both within and across clauses.



To facilitate the following discussion, let me introduce some notational and terminological conventions. Cartographic hierarchies will be represented by sequences of heads, $H_1, H_2, \dots, H_{n-1}, H_n$, where the $>$ relation of subscripts corresponds to c-command within the hierarchy, as shown in the tree.

Cartographic hierarchies define templates of bounded size. To generate unbounded structures, (pieces of) these templates are embedded in each other recursively. I call every token of (a maximal fragment of) a cartographic hierarchy a *cyclic domain*.

There are two possible types of solution to the problem posed by the ordering of operations. We can either use the hierarchy of functional heads outside of the moving elements to derive the order of operations, or we may employ the hierarchy of functional heads within the moving element. I will discuss the two options in the order given. Wherever possible, I frame the discussion in terms of analyses actually found in the literature.

2.1 Solutions based on external syntax

2.1.1 Landing Site > Launching Site (Sternefeld, 1992)

If we assume that the landing site of movement c-commands the launching site, then it follows for movement within a cyclic domain that (the head projecting) the landing site of movement > (the head projecting) the launching site of movement. This means that for movement within a given cyclic domain, the ordering of operations falls out from the fact that movement is upward in the tree. Sternefeld (1992) suggested to generalize this to movement across cyclic domains and to demand that movement in general, that is, even across cycles, obey the constraint that (the head projecting) the landing site of movement > (the head projecting) the launching site of movement. This constraint now, of course, no longer follows from the direction of movement.

Sternefeld's condition derives the ban on improper movement for the identity case, but there are problems both with the case of surfing paths, figure 5, and with remnant movement, figure 6. The correct ordering of operations is not derived in the case of surfing paths, because Sternefeld's condition only guarantees that $HP_a > HP_b$ and $HP_u > HP_v$. From this it does not follow that $HP_a > HP_u$. Similarly, in (11b) Sternefeld's condition makes sure that $HP_a > HP_b$ and $HP_u > HP_v$. This again does not guarantee $HP_a > HP_u$. Sternefeld's is a classical theory of improper movement, dealing only with the identity case.

While Sternefeld's condition overgenerates with respect to surfing and diving paths, it also undergenerates elsewhere under certain assumptions. If we assume a version of the raising to object analysis of ECM in English (Postal (1974); Lasnik and Saito (1991); Davies and Dubinsky (2004)), the exceptionally case-marked argument almost certainly lands in a position in the matrix domain which is projected by

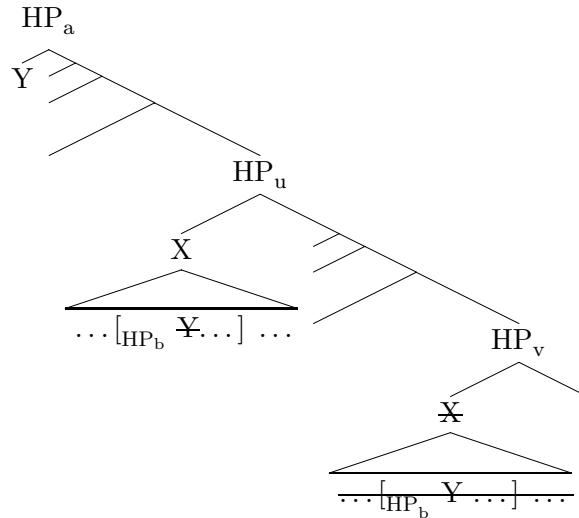


Figure 5: Indexed structure for surfing paths

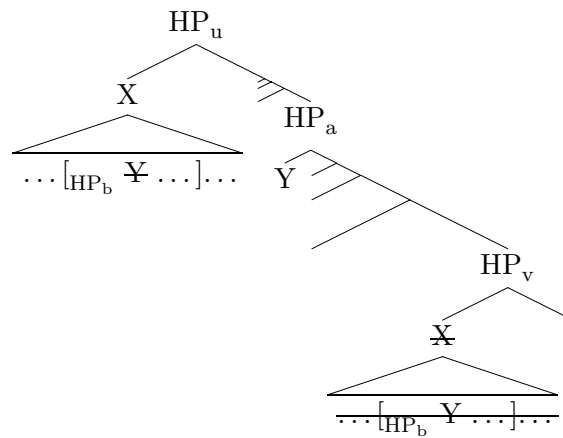


Figure 6: Indexed structure for diving paths

a head (AgrO) < the head projecting the highest position occupied by that argument in the embedded domain (T). This is shown in (11).

- (11) I want [_{AgrOP} John ... [_{TP} t_{John} to leave]]

I now turn to Williams' (2002) theory, which is considerably stronger than Sternefeld's. Anticipating the conclusion of the next subsection, we will see that Williams' proposal solves the overgeneration issues facing Sternefeld's proposal, but the undergeneration issues multiply.

2.1.2 For all H_n along the path of movement, the head projecting the landing site $\geq H_n$ (Williams, 2002, LEC)

Williams' (2002) Level Embedding Conjecture (LEC) together with the extension condition has the following as a corollary: For all H_n along the path of movement, the head projecting the landing site $\geq H_n$. A head H counts as *along the path of movement* iff some projection of H dominates the launching site but does not dominate the landing site of movement. In other words, every head that m-commands the launching site but does not c-command the landing site of a given movement counts as part of the path. This definition of a movement path ensures that the head projecting the launching site and the head including the landing site are both included in the path. This formulation thus entails Sternefeld's condition from the previous subsection (modulo any difference that arises from the difference between $>$ and \geq). Hence, Williams' formulation rules out a proper superset of the structures ruled out by Sternefeld's condition.

Thus, the LEC derives the ban on improper movement for the identity case for the same reason that Sternefeld's condition does. The LEC also derives the ban on improper movement for the case of surfing paths. To see this consider again the structure in figure 5. The LEC guarantees not only that $HP_a > HP_b$ and $HP_u > HP_v$, it crucially makes sure that $HP_a > HP_u$. Under the paraphrase of the LEC in the preceding paragraph this is true technically, because the head H_u^0 m-commands the launching site of Y and thus counts as lying along the path of Y . Finally, the LEC derives the ordering of operations for the case of remnant movement. A look at figure 6 reveals that, as before, $HP_a > HP_b$ and $HP_u > HP_v$, but crucially the LEC also demands that $HP_a > HP_u$. This is so because under the LEC the head projecting the landing site must be higher in the functional sequence than all of the heads along the path including, in particular H_a^0 . Thus, the overgeneration problems facing Sternefeld's approach are solved by the LEC.

Exceptional case marking in English again spells trouble for this analysis, at least if raising to object analysis is correct (Postal (1974); Lasnik and Saito (1991); Davies and Dubinsky (2004)). The next example illustrates the problem on the assumption that negation is in the same hierarchical position in the embedded infinitive and in the matrix. *Mary* crosses negation in the infinitive but lands below it in the matrix. For concreteness assume that *Mary* ends up in [Spec, Agr_OP] in the matrix. On the way from the θ -position in the embedded VP to matrix [Spec, Agr_OP] *Mary* crosses negation in the embedded clause. Therefore, Agr_O⁰ > Neg⁰. However, the linear order of matrix negation and *Mary* indicates that negation is higher than Agr_O, i.e., Neg⁰ > Agr_O⁰. A similar contradiction arises with respect to the ordering of T⁰, hosting *to*, and Agr_O⁰. There does not seem to be a way of avoiding this contradiction.⁵

(12) I expect John not to want Mary not to be told _{t_{Mary}} the truth.

The same issue arises for hyperraising (Nevins (2004); Williams (to appear)) if the movement analysis of hyperraising is correct, for quantifier movement across a complementizer in French (noted as a problem in Williams (2002)) and negative quantifiers in Scandinavian, for scrambling in Russian (Glushan, 2006), and for long *wh*-movement in the Bantu language Kĩĩtharaka (see Muriungi (2005); Abels and Muriungi (to appear), also Bergvall (1987); Schwarz (2003) for a similar facts from Kikuyu),... The problem always has the same shape: an element high in the hierarchy, typically T⁰ or C⁰, is crossed on the path of movement but the landing site of the movement is below T⁰ or C⁰. This is quite obvious in the case of quantifier movement in French, (13a-b), where the quantifiers *rien* and *tout* might originate as low as the object position of the embedded verb, cross T⁰ and C⁰ along the way and end up in a position transparently below T⁰. Similarly for the Icelandic examples in (14). These cases are only a problem for Williams' approach, Sternefeld's theory sketched above allows them.

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

⁵Williams (2002) is aware of the issue and weakens the impact of the LEC by allowing what he calls *mismappings* between levels.

- 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.
- (14) Icelandic ((Kayne, 1998, p. 142 #55-6) attributed to Rognvaldsson (1987, p. 44) and Jonsson (1996, p.86))
- a. Hann mun ekkert hafa getadh gert.
 they will nothing have could done
 They haven't promised to do anything.
- b. Their hafa ekkert lofadh adh gera.
 they have nothing promised to do
 They haven't promised to do anything.

The same point is made again by the Kĩĩtharaka examples in (15). Kĩĩtharaka is a Bantu language spoken in Kenya. This language allows *wh*-questions to be formed by leaving the *wh*-word in situ, (15a), by moving it partially, (15b-d), and by moving it all the way to its scope position, (15e). The *wh*-word *ũũ* - 'what' is boxed in the examples for easier orientation. There is apparently no semantic difference between these five examples (see Muriungi (2003, 2005); Abels and Muriungi (2006) for discussion), the translation in (15e) is appropriate for all of them. The crucial example in the current context is (15d). The *wh*-word moves across the underlined complementizer *atĩ* in the lower clause and ends up right after the circled complementizer *atĩ* in the next higher clause. This is then another example where an element crosses over a position which is higher in the template than the position where it ends up.⁶

- (15) Kĩĩtharaka (Peter Muriungi, p.c.)
- a. U- ri- thugania atĩ John a- ug- ir- e
 2nd.SG- pres- think that John sm- say- perf- FV

⁶Non-obvious glosses are as follows:

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

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

In this section we have seen that Williams' theory based on the LEC and the extension condition avoids the overgeneration problems that remained unresolved under Sternefeld's approach, but in doing so Williams' theory gives rise to quite severe undergeneration problems.

2.1.3 Conclusion: External syntax

Both of these potential solutions attempt to state the generalization derived above by (10) using conditions over launching and landing sites and/or paths of movement, that is, material external to the moving item. In both cases the statement makes reference to the hierarchy as an independent object, independent of the actual tokens that form a given syntactic object; reference to the resulting syntactic object with its structural hierarchy (c-command, dominance) alone is insufficient. The independent hierarchy of labels is necessarily invoked in stating the restriction. Both approaches have serious empirical disadvantages not shared by the system adopting GenPIM.

Independently of whether Sternefeld's, Williams', or the present approach are to be preferred eventually, all of them require the hierarchy to be accessible for the syntactic computation beyond the narrow realm of building up phrase structure trees whose local structure is isomorphic to the hierarchy. All cartographic work that I am familiar with, uses the hierarchy in this very narrow and as we have seen now too insufficient sense.

2.2 Solutions based on internal syntax

The next two subsections discuss the option of deriving the ban on improper movement from the internal structure of the moving element. Unlike the approaches based on the external syntax, possible approaches based on the internal syntax have not been proposed in the literature, though a version of the top-down condition on feature accessibility under the name of 'peeling' has been suggested to me as a possible account of the ban on improper movement by Michal Starke in personal communication. Given the lack of a more or less worked out proposal, which would make use of the internal syntax to deal with improper movement, the discussion is necessarily more schematic.

2.2.1 Bottom-up condition on feature accessibility

Since at least Abney's work (1987) it has often been assumed that there is a strong parallelism between the noun phrase/DP and the clause/CP. Such parallelism would suggest that θ -related functional structure occupies the lower stretches of the noun phrase, case is located

higher up, and *wh*- and information structure related features are even more peripheral to the NP. The functional hierarchy within the noun phrase might then be topic > *wh* > scrambling > Case > θ . Since θ -features have to be satisfied first, followed by case, and then and only then by operator features, one might conjecture that the features of a head H_x become accessible for checking/attraction only after the features of all heads H_y that are c-commanded by H_x (alternatively: that are dominated by a projection of H_x , i.e., that are m-commanded by H_x) have been checked/attracted.

This condition will directly rule out improper movement for fairly simple instances of the identity condition.⁷ The bottom-up condition on feature accessibility does not derive the ban on improper movement for the case of diving paths. This can be shown by considering example (2cii) above, repeated here for convenience.

(2cii) * $[A \text{ picture of } t_{\text{which king}}^{wh}]$ is known which king to have been sold t_a^A picture of which king

Structures like this are not ruled out, because first the more deeply embedded feature is checked and the higher feature is checked later in conformity to the bottom-up condition. The problem is that the features reside in different cyclic domains, thus cartography imposes no ordering on them. The original problem that motivated the ban on improper movement, the problem of enforcing an ordering across cycles, comes back here. When the moving constituent is bigger than a single cyclic domain, the bottom-up condition on feature accessibility does not guarantee an ordering any more.

In addition, the bottom-up condition on feature accessibility categorically rules out surfing paths, (2bi) above. The Case feature in the larger DP is checked first. This (wrongly) makes the more deeply embedded *wh*-feature of the complement inaccessible.

The fact that the bottom-up condition on feature accessibility rules out surfing paths (i.e., derives the freezing principle) might seem relatively harmless even in the face of the apparent counterexamples given above, since every derivation with surfing paths can, in a sufficiently unrestrictive theory of movement (e.g. Koopman and Szabolcsi (2000)),

⁷The hedge ‘fairly simple’ will be discussed in the conclusion to this section, 2.2.3

be emulated by a derivation with diving paths. This is illustrated in figure 7. The symbol \rightsquigarrow is to be read as *can be emulated by*. The

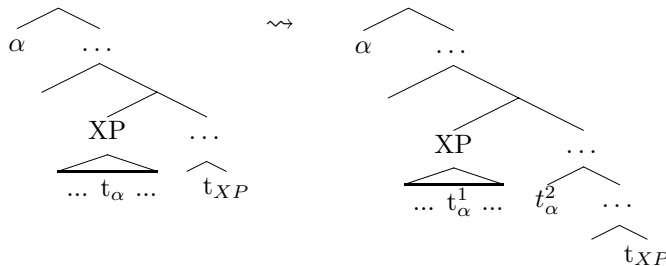


Figure 7: Emulating surfing through diving

idea in figure 7 is to move α out of XP before XP moves for the first time (‘stacking it’ in Koopman and Szabolcsi’s terms), then letting XP undergo remnant movement, and finally moving α past XP. The structures are identical modulo the additional intermediate trace of α , t_α^2 .

However, this solution inherits the problem that no ordering is imposed on the various steps of remnant movement. As long as the features reside in different cyclic domains A-movement before *wh*-movement as well as *wh*-movement before A-movement ought to be allowed. And even if this problem could be overcome somehow, the putative unification of subextraction and remnant movement in figure 7 would miss the generalization from table (1, since the cases of real, surface true, remnant movement would require the opposite ordering of operations from the cases of apparent remnant movement, i.e., subextractions on the surface, suggesting that the unification of subextraction with remnant movement would be a step in the wrong direction.

2.2.2 Top-down condition on feature accessibility

We can ask ourselves whether the serious problems of the bottom-up condition on feature accessibility might have stemmed from having assumed the wrong arrangement of features within the DP. We might then put the hypothesis that there is parallelism between CPs and DPs in terms of feature organization on its head and assume that the featural organization of CPs and DPs are mirror images of each other:

DPs have their θ -features outermost, followed by a Case-layer lower down, followed by a layer where *wh*- and discourse related features reside. Such an organization of the DP could then be coupled by a top-down condition on feature accessibility: Features of a head H_x become accessible for checking/attraction only after the features of all heads H_y that c-command H_x (alternatively: whose projections dominate H_x , i.e., that m-command H_x) have been checked/attracted.⁸

On the positive side, the top-down condition on feature accessibility derives the ban on improper movement for simple cases of the identity condition.⁹ However, the top-down condition on feature accessibility does not derive the ban on improper movement for surfing paths. Recall example (2bii), repeated here for convenience.

(2bii) *Oscar is known [how likely t_{Oscar}^A to win] it was $t_{\text{how likely } t_{\text{Oscar}} \text{ to win}}^{wh}$.

Checking is top-down. Since the *wh*- and the Case-feature reside in different cyclic domains, the cartographic hierarchy is also respected. The top-down condition on feature accessibility thus does not derive the ordering of operations observed with subextraction.

Worse still, the top-down condition on feature accessibility disallows remnant movement (diving paths) categorically. To see this consider again example (2ci), repeated below.

(2ci) ✓It is known [how likely t_{Oscar}^A to win] Oscar is $t_{\text{how likely } t_{\text{Oscar}} \text{ to win}}^{wh}$.

Such structures cannot be derived since first the more deeply embedded feature is checked/attracted creating the remnant. Later the higher feature is checked/attracted leading to remnant movement. This sequence violates the top-down condition on feature accessibility. More generally, no remnant movement derivation can be allowed under the

⁸Such a condition might be derivable for the simplest of cases by assuming that once a feature has been satisfied the relevant layer of the DP is stranded and not taken along under movement. Thus, the topmost θ -layer of the DP would be stranded in the θ position when movement to the Case-position takes place, etc. This is what Starke calls ‘peeling.’ The ‘peeling’ solution apparently makes the prediction that moved *wh*-phrases should be caseless, seeing as they peel off their case layer. This expectation is not borne out. The ‘peeling’-solution also inherits the problems to be discussed in the main text. These are shared by all variants of the top-down condition on feature accessibility.

⁹ The hedge ‘simple’ will be discussed in the conclusions to this section, 2.2.3

top-down condition on feature accessibility, unless the features are not in a c-command relation to begin with. But in that case no ordering of operations can be derived from the top-down condition, counter to fact.

The prohibition against remnant movement is very damaging. We saw above the subextraction could be emulated through remnant movement plus ‘stacking.’ The opposite is not true, leaving potential defenders of the top-down condition without any account of remnant movement.

2.2.3 Conclusion: Internal syntax

We have seen that both solutions based on internal syntax face immediate serious problems. Stated generally, the problem with both attempts is that they do not generalize across cyclic domains within the moving item.

As hinted at in the text, the solutions apparently only work in the most trivial of cases. Thus compare (16a), which is just like (2a_{ii}) except with a relative clause, with (16b) and (16c). While in the trivial case of (16a) the correct templatic arrangement of features together with the top-down or bottom-up condition on feature accessibility provides a solution to the improper movement puzzle, the hierarchical relations between the features are arguably altered in (16b) and (16c). But if this is so, then the top-down or bottom-up condition on feature accessibility fail even for the identity condition as soon as more than the most trivial examples are taken into consideration.

- (16) a. *the man who is known t_{who}^A (that) it was told t_{who}^{wh} that he would be fired
 b. *the man whose brother is known $t_{\text{whose brother}}^A$ (that) it was told $t_{\text{whose brother}}^{wh}$ that he would be fired
 c. *the man the brother of whom is known $t_{\text{the brother of whom}}^A$ (that) it was told $t_{\text{the brother of whom}}^{wh}$ that he would be fired

Of course these examples with DPs inside of DPs, that is, cyclic domains embedded within each other, underscore the problem just noted. The putative solutions based on internal syntax fail because they do not generalize across cyclic domains within the moving item. At least for

the top-down condition this generalization can be achieved by replacing the c-command or m-command relation in the formulation of the condition on feature accessibility by the $>$ relation. Given the apparent isomorphism between $>$ and \succ , the top-down condition turns into a variant of (10) and the cartographic hierarchy is stated and accessible independently of the structural tokens implementing it.

This alternative theory would be just like (10), but replace the definition of affectedness in (10c) by the following definition.

- (17) A constituent α is *affected* by a movement operation iff α is contained in the constituent created by movement.

This alternative theory has the advantage of offering a more symmetrical definition of affectedness. It derives all three cases of improper movement and shares with the LEC solution that all specifiers and adjuncts along the movement path are affected by the movement of α . The additional predictions of this formulation are largely correct (and follow from the LEC, too); it runs into difficulties with quantifier movement in French and Scandinavian.

3 Conclusion

It should have become clear throughout this paper that the ordering of operations provides a novel source of evidence for – and testing ground of – cartographic hierarchies. To achieve this generalization, cartographic hierarchies must be stated independently of templates; they are more abstract hierarchies. These hierarchies seem to regulate the order of application of operations (merge and move), but in slightly different ways: For external merge recursion is possible, for internal merge (move) no recursion is allowed in the *affected* domain.

On the whole this view leads to a more integrated theory of structure and movement: they are regulated by the same global constraint, namely the hierarchy, which applies to them in slightly different ways. In addition, we end up with an improved theory of movement. Under the current standard view, movement operations affect only the moving constituent itself and the attractor. This opens the door to very stipulative accounts of movement since movement of one constituent is completely independent of movement of any other constituent. The

idea that movement affects parts of the structure distinct from the moving constituent and the attractor itself adds constraints on this system. The system sketched in (10) cuts considerably into the expressive power of the theory (see Abels (to appear) for discussion) in a way which is, apparently, empirically desirable.

If the hierarchy is more abstract than often thought, there must be an interface between the hierarchy and the structure projected from it. In the case discussed here, that of internal merge, this interface is provided by GenPIM and affectedness. Similarly, some interface will be necessary between the hierarchy and the structures projected from it via external merge. This interface is usually construed either as an isomorphism or a homomorphism, but other interfaces are imaginable that might solve some of the open problems such as Bobaljik paradoxes and transitivity failures that arise for cartographic hierarchies at present.

References

- Abels, Klaus. to appear. Towards a restrictive theory of (remnant) movement: Improper movement, remnant movement, and a linear asymmetry. *Linguistic Variation Yearbook* 7.
- Abels, Klaus, and Peter Muriungi. 2006. The focus particle in Kĩĩtharaka. 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. to appear. The focus particle in Kĩĩtharaka: Syntax and semantics. *Lingua* .
- Abney, Steve. 1987. The english noun phrase in its sentential aspect. Doctoral dissertation, MIT.
- Beck, Sigrid. 1996. Quantified structures as barriers for LF movement. *Natural Language Semantics* 4:1–56.
- Bergvall, Victoria Lee. 1987. Focus in Kikuyu and universal grammar. Doctoral Dissertation, Harvard University, Cambridge, Mass.
- Boeckx, Cedric, and Norbert Hornstein. 2003. Reply to "Control is not movement". *Linguistic Inquiry* 34:269–280.
- Bošković, Željko. 1994. D-structure, theta-criterion, and movement into theta-positions. *Linguistic Analysis* 24:247–286.
- 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.
- Cinque, Guglielmo. 1990. *Types of A'-dependencies*. Cambridge, MA: MIT Press.
- Davies, William D., and Stanley Dubinsky. 2004. *The grammar of raising and control*. Blackwell.
- 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. 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.
- Friederici, Angela D., Matthias Schlesewsky, and Christian J. Fiebach. 2003. WH-movement versus scrambling: the brain makes a difference. In *Word order and scrambling*, ed. Simin Karimi, 325–344. Malden, Mass.: Blackwell Publishing.
- Glushan, Zhanna. 2006. Japanese style scrambling in Russian: myth and reality. Master's thesis, Universitetet i Tromsø.
- Grewendorf, Günther. 2001. Multiple *wh*-fronting. *Linguistic Inquiry* 32:87–122.
- Grewendorf, Günther. 2003. Improper remnant movement. *Gengo Kenkyu* 123:47–94.
- Hornstein, Norbert. 1999a. Minimalism and quantifier raising. In *Working minimalism*, ed. Samuel David Epstein and Norbert Hornstein, 45–75. Cambridge, MA: MIT Press.
- Hornstein, Norbert. 1999b. Movement and control. *Linguistic Inquiry* 30:69–96.
- Jonsson, J. G. 1996. Clausal architecture and case in Icelandic. Ph.d. dissertation, University of Massachusetts, Amherst.
- Kayne, Richard. 1998. Overt vs. covert movement. *Syntax* 1:128–191.
- Koopman, Hilda, and Anna Szabolcsi. 2000. *Verbal complexes*. Cambridge, MA: MIT Press.
- Kutny, K. de, and Walt Detmar Meurers. 1999. On partial constituent fronting in German. In *Tübingen studies in head-driven phrase structure grammar*, ed. V. Kordoni, 22–73.
- Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15:235–289.
- 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.
- May, Robert. 1979. Must comp-to-comp movement be stipulated? *Linguistic Inquiry* 10:719–725.
- Müller, Gereon. 1998. *Incomplete category fronting: A derivational approach to remnant movement in German*. Dordrecht, Boston: Kluwer Academic Publishers.
- 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. 2003. Wh-questions in Kiitharaka. Master's thesis, University of the Witwatersrand, Johannesburg.
- 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.
- Pesetsky, David. 2000. *Phrasal movement and its kin*. Linguistic Inquiry Monograph 37. Cambridge, MA: MIT Press.
- 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.
- Rognvaldsson, E. 1987. OV word order in Icelandic. In *Proceedings of the Seventh Biennial Conference of Teachers of Scandinavian in Great Britain and Northern Ireland*, 33–49. University College London.
- 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.
- Schwarz, Florian. 2003. Focus marking in Kikuyu. In *Questions and focus*, ed. Regine Eckardt, volume 30 of *ZAS Papers in Linguistics*, 41–118. Berlin: Zentrum für allgemeine Sprachwissenschaft.
- Starke, Michal. 2001. Move dissolves into merge: a theory of locality. Doctoral dissertation, University of Geneva.
- von Stechow, Arnim, and Wolfgang Sternefeld. 1988. *Bausteine syntaktis-*

- chen Wissens: ein Lehrbuch der generativen Grammatik*. Opladen: Westdeutscher Verlag.
- Sternefeld, Wolfgang. 1992. Transformationstypologie und strukturelle Hierarchie.
- Williams, Edwin S. 1974. Rule ordering in syntax. Doctoral Dissertation, MIT.
- Williams, Edwin S. 2002. *Representation theory*. Cambridge, Mass.: MIT Press.
- 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.