Head Movement and the Minimalist Program

Ian Roberts

Downing College, University of Cambridge

(igr20@cam.ac.uk)

0. Introduction

In this paper I will begin by recapitulating the essentials of the analysis of head-movement as it was largely agreed on in mainstream syntactic theory by the late 1980s. This approach was in essence unaltered in the earlier versions of minimalism (Chomsky (1993, 1995, aside from section 4.10)). In Section 2, I consider the reasons which lead Chomsky (2001: 37-8) to suggest excluding head-movement from the core operations of the narrow syntax. Section 3 reviews the various alternatives to the earlier conception of narrow-syntactic head-movement which have been put forward: PF movement, remnant phrasal-category movement and "reprojective movement", focussing on a case study of each alternative. Finally, in Section 4, I will consider the conceptual status of head-movement in relation to the general goals of the minimalist program.

1. The GB approach[1]

Earlier versions of generative grammar often featured head-movement operations; see for example Affix Hopping in Chomsky (1957), McCawley's (1971) Tense-attraction rule, Emonds' (1971, 1976) *have/be*-raising and his (1978) verb-movement rule for French, den Besten's (1983) analysis of Germanic verb second, French subject-clitic inversion and English subject-auxiliary inversion. But it was only in the GB period that these ideas were systematised and a series of theoretical postulates were put forward that together provided a clear characterisation of head-movement, arising primarily from the work of Koopman (1984), Travis (1984) and Baker (1985, 1988).

The central idea in these approaches is (1):

(1) Head movement is the case of Move-α where α is X°.

[1] This section summarises some of the main points in Roberts (2001).

In most versions of X'-theory assumed in GB theory, X° was defined as the head of XP. This was the position into which terminals could be substituted; unlike the bare-phrase-structure notion of minimal X, however, it could have internal structure (beyond simply being a bundle of features of some kind), in part thanks to the possibility of head-movement.

As an instance of Move-α, head-movement was argued to be subject to the standard well-formedness conditions applying to movement operations and their outputs generally. These conditions were of three main types, not necessarily exclusive in their empirical effects: structure preservation, locality and the requirement that the trace created by the movement operation meet the relevant well-formedness conditions on traces. Let us look at each of these in turn.

Concerning structure preservation, Chomsky (1986:4) posits two general conditions on movement: "only X° can move to a head position" and "only a maximal projection can move to a specifier position". He remarks that these "would follow from an appropriate form of Emonds' Structure Preservation Hypothesis" (the second given "the X-bar theoretic assumption that heads cannot be base-generated without a maximal projection so that a bare head cannot appear in the specifier position to receive a moved X^o category"). Later, Chomsky (1986:73) suggests that only maximal projections may adjoin to maximal projections, ruling out adjunction of heads to maximal projections. This proposal ("a kind of generalization of Emonds' Structure Preserving Hypothesis" (*ibid*)) follows if "we were to regard movement of a lexical category as analogous to NP-movement, barring [it] either on the grounds that *t* [the trace of this movement – IGR] is an unlicensed free variable or that there is 'improper movement' with *t* ultimately bound in the domain of the head of its chain" (*ibid*). Given the assumptions in Chomsky (1986), this would violate Principle C of the binding theory. If the head moves on to an "A-position" (i.e. a position adjoined to another head), then, we have improper movement. If it does not, the trace of head-movement counts as an unlicensed free variable. This proposal does not, however, rule out head-to-head adjunction, and was not intended to. In fact, the upshot of Chomsky's reasoning is that head-movement can only move a head to another head position. It was generally assumed that head-movement adjoined the moved head to the host head, forming a structure like (2):

(2) $[_Y\ X\ Y\ ]$

(But see Rizzi & Roberts (1989) for a more elaborate proposal). Kayne (1991, 1994) proposed that head-adjunction is always left-adjunction, as depicted in (2).

Concerning locality, the central condition on head movement was the Head Movement Constraint (HMC), first explicitly formulated in Travis (1984). I give it in the following form:

(3) Head movement of X to Y cannot skip an "intervening" head Z.

(Roberts (2000:113))

"Intervention" is understood in terms of asymmetric c-command in the usual way (Z intervenes between Y and X iff Y asymmetrically c-commands both X and Z, while Z asymmetrically c-commands X). The HMC has the effect of forcing head-movement to be cyclic, in an obvious sense. Moreover, it has typically been assumed that formation of the complex head in (2) could not be undone by a later step of movement. Hence further movement of Y to a higher head W would form the complex head $[_W [_Y \ X \ Y \ ] \ W \ ]$. In other words, iterated head-movement always involves cyclic "roll-up," forming a successively more complex head. On the other hand, "excorporation" of X from Y (or W), or of Y from W (or X), is not allowed (but see Roberts (1991) for the observation that this assumption did not follow from any aspect of the theory of movement assumed at the time, and a discussion of two empirical candidates for excorporation).

Finally, and perhaps most importantly in the GB context, the trace of head-movement was subjected to the standard conditions on traces. Indirectly, many of the conditions on head movement were derived in this way, since a structure containing an ill-formed trace would be ruled out. The movement itself was allowed to overgenerate illicit representations, which were filtered out by specific conditions. The most important of these conditions was the Empty Category Principle (ECP), which required all traces to be properly governed. Definitions of proper government varied somewhat in detail, and for present purposes it is simpler to break down the requirements imposed by the ECP into a number of separate cases, bearing in mind that the ECP provided a unified characterisation of this range of cases.

First, one effect of the ECP was that head-movement of X to Y out of an XP not contained in the structural complement of Y is impossible. Thus head-movement

from subjects and adjuncts was impossible. Baker (1988), in particular, showed in detail that the various forms of incorporation he proposed satisfy this condition. Second, "downward" head-movement is not allowed, since a fundamental requirement imposed by the ECP is that an antecedent c-command its trace. This implies that the Affix Hopping, as conceived in Chomsky (1957) and elsewhere, could not be an instance of head-movement if this is seen as a core syntactic operation. (Pollock (1989) and Chomsky (1991), among others, sought to avoid this consequence by treating the ECP as holding of LF representations and allowing downward movement in the overt syntax as long as the effects of this were obliterated by the time the ECP applied). Third, to the extent that, following Rizzi (1990), the ECP featured some form of relativised minimality constraint, the HMC itself can be derived from the ECP. Hence the local nature of head-movement follows.

The GB conception of head-movement, then, was that this was a core syntactic operation raising a head X to an immediately superjacent (governing) head Y where X is contained in Y's immediate structural complement. The effects of this highly articulated and restricted conception were observed in a very wide range of empirical phenomena: noun-incorporation, many kinds of morphologically complex causative constructions, applicatives, passives, verb-movement within the clause of the French/Romance kind, to C of the Germanic kind and to clause-initial position of the kind found in VSO languages, English subject-auxiliary inversion, French subject-clitic and complex inversion, Italian Aux-to-Comp, inversion of inflected infinitives in European Portuguese, a whole range of phenomena involving movement of the Noun within DP (including Semitic construct states, Balkan and Scandinavian postposed articles and the relative ordering of Nouns in relation to possessors and modifiers of various kinds; see Cinque (1994), Longobardi (1994)), clitic-movement, and many other phenomena (see Roberts (2001) for overview, illustration and further references).

2. *The minimalist program*

In the early versions of the minimalist program, the GB conception of head-movement was by and large retained. The discussion of V-movement to T and Agr and related issues in Chomsky (1993:27-32/1995:195-199) introduces checking theory, and makes it clear that V-movement, like other forms of movement, obeys the

core constraints that this theory imposes. The same is true for the notions of checking domain, internal domain and complement domain; moreover, head-movement plays a role in giving rise to equidistant positions, which is central to capturing the generalisation that objects move only when the verb moves (an early version of Holmberg's (1986) generalisation); see the discussion in Chomsky (1993:10-19/1995:176-186). In Chomsky (1995: 4.10), the picture changes somewhat, partly as a consequence of the abandonment of Agr as a syntactic category. Here, Chomsky proposes analysing "multiple subject constructions" (e.g. Germanic transitive expletive constructions such as *There painted a student the house* or passives like *There have some cakes been baked for the party*) in terms of multiple specifiers of T, since SpecAgrP is not available if Agr is not a functional head. Chomsky argues for an analysis which features the substring *Expletive Subject T* ... If V is in T, this is clearly the wrong order, the attested order being *Expletive V Subject*. Taking this order to be a direct reflex of the verb-second property of the languages in question, Chomsky suggests (1995:368) that "the V-second property .. may belong to the phonological component. If that is the case, the observed [i.e. V2 – IGR] order is formed by phonological operations … and may observe the usual constraints (V --> C), but need not, as far as we know". Although V-to-T movement is assumed (Chomsky (1995:367)), the possibility that V2 orders are derived by something other than syntactic head-movement of T to C is at least questioned here.

But it was in Chomsky (2001:37-8) that a series of arguments of a range of types are presented that, together, lead Chomsky to conclude that "a substantial core of head-raising processes, excluding incorporation in the sense of Baker (1988), may fall within the phonological component" (37).[2]

First, Chomsky claims that head-movement never affects interpretation: "the semantic effects of head-raising in the core inflectional system are slight or non-existent, as contrasted with XP-movement" (2001:37). The core point here is that, while French or Icelandic verbs occupy a different structural position in finite clauses from their English or Mainland Scandinavian counterparts, analysing this in terms of different extents of head-movement as was standard in GB or early Minimalism (see in particular Vikner (1995)), leads to the expectation that there may be some LF-

[2] Chomsky excludes incorporation here because it has rather different properties from the other cases of head-movement (in particular, according to Baker, it is implicated in the core cases of grammatical-function changing phenomena).

related differences between verbs – perhaps involving scope or reconstruction effects -- in the two classes of languages. Such effects are not found, leading to the suggestion that head-movement is confined to the PF part of the grammar.

Second, Chomsky raises the question of the nature of the trigger for head-movement. The issue arises when we consider, for example, T in a language such as French, which has consistent DP-movement into SpecTP and consistent V-movement to T (following Pollock (1989)). Hence T must contain the relevant triggers for these movements: (uninterpretable/unvalued) φ-features and an EPP feature to trigger DP-movement and, presumably, some form of V-features combined with a movement-triggering feature triggering V-movement. All other things being equal, the system has to have sufficiently rich featural information to be able to correctly distinguish the two sets of triggers: an XP-movement trigger for D and head-movement trigger for V. Similarly in V2 languages: T must move to C and XP to SpecCP, but not vice versa. Note that the prediction is not that the inverse properties may not exist as parametric options: perhaps they do. For example, D-movement to T can be seen as a form of subject-cliticisation, while VP-movement to SpecTP, satisfying T's EPP feature, was argued for for Niuean by Massam (2000) and others (see Section 3.2 below). The point is that the movement-triggering mechanism needs to be enriched in such a way that head-movement has a special kind of triggering feature. Chomsky suggests that such a complication is not needed if head-movement is treated as something outside of the core computational system of narrow syntax.

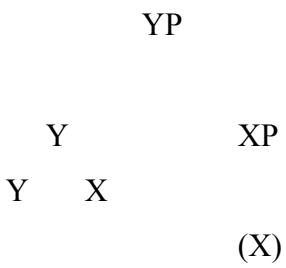
Third, Chomsky points out that the derived structure of head-movement, as construed in Section 1, is countercyclic; in fact, it violates the Extension Condition.[3] The Extension Condition requires that all movement operations extend the root of the structure that they apply to. For example, a standard case of A-movement raising the subject to SpecTP (triggered by T's EPP feature in the system in Chomsky (2001)) applies at the point in the derivation after T is combined with its complement vP. T Agrees with the nearest DP that it asymmetrically c-commands, which, in a simple transitive clause, is the DP merged in SpecvP, the external argument. In virtue of this Agree relation and T's EPP feature, this DP is raised, forming SpecTP. The formation of SpecTP extends the root at this point in the derivation. It is fairly clear that wh-movement to SpecCP, as well as various kinds of adjunct operations, can be seen in

[3] This was noticed in Chomsky (2000:137), where it is concluded that the Extension Condition should be weakened in this case.

the same light, whatever the precise details concerning the triggers for these operations. However, as we have see, head-movement was thought to derive structures such as that in (2), by adjoining one head to another. Such an operation does not involve extension of the root, at least in any obvious way without appeal to a special notion of "root" (which is imaginable but has not been proposed; the assumed notion of root is that node X such that there is no node Y that irreflexively dominates X). maybe HM is simult w/ merge?

Fourth, Chomsky makes the related point that, owing the fact that head-movement adjoins one head to another, in the derived structure the moved head is unable to c-command its trace/copy. This is true if we maintain the most natural definition of c-command: that it is the transitive closure of sisterhood and containment (this is the natural definition since both sisterhood and containment can be directly defined in terms of Merge; see Chomsky (2000: 116). If we adopt a definition of the kind assumed in Kayne (1994: 18), which allows an adjoined category to c-command both the category to which adjoins, and out of that category,[4] then the moved head would be able to c-command its trace in a typical head-movement configuration such as that shown in (4):

(4)



But Chomsky suggests that such complications of the definition of c-command are unnecessary and undesirable (they do not "fall under the notion of c-command derived from Merge", i.e. transitive closure of sisterhood and containment (Chomsky (200): 116)). If so, then head-movement features a major anomaly in relation to other types of movement in that the moved category does not c-command its trace.

Fifth, Chomsky pointed out that head-movement was suspect as a core-syntactic operation since onward cyclic movement is never successive-cyclic, instead

[4] The definition is as follows:
(i) X c-commnads Y iff *X and Y are categories* and X excludes Y and every category that dominates X also dominates Y (emphasis in original).

it always involves "roll-up" (i.e. movement of the entire derived constituent formed by movement, of the type in (2)). We commented on this point in Section 1: after adjunction of X to Y, forming (2), further movement of Y to a higher head W forms the complex head $[_W [_Y \ X \ Y \ ] \ W]$. In other words, iterated head-movement always forms a successively more complex head. Successive-cyclic head-movement, on the other hand, would involve excorporation of X from [$_Y$ X Y], moving X on to form $[_W \ X \ W]$. As already pointed out, Roberts (1991) observed that nothing prevented this in the GB conception of head-movement, and that it was in fact empirically desirable. The general view, however, has remained that this possibility is not found (the empirical cases Roberts adduced can be analysed in other ways). If so, then an explanation is required. Chomsky says that if head-movement were seen as a morphological operation, then this might be why we do not observe excorporation ("iterability is a general property of operations of narrow syntax, and these alone" (p. 38)). But if we treat head-movement as syntactic movement, then we have to explain why successive-cyclic movement, so clearly available for phrasal movement (both A and A'-movement) is not available to head-movement.

Chomsky's arguments have given rise to various reactions, as we shall see. In general they have been influential, in that many researchers have been led to look for alternatives to the earlier approach to head-movement, either by eliminating it altogether, eliminating it from the core computational system of narrow syntax or radically redefining it. In many cases, new phenomena have been brought to bear on the issues, or at least older data has been reconsidered in a new light. Two points should, however, be made here. First, although Chomsky's arguments naturally lead to a re-evaluation, at least, of the account of head-movement sketched in Section 1, he does not articulate a theoretical principle which would force, either directly or as a deductive consequence, the elimination of head-movement from narrow syntax. The question that remains open if we accept Chomsky's conclusions is then: why is head-movement not part of narrow syntax? The second point is related: to what extent do these questions bear on the conceptual goals of the minimalist programme? To put the question, in a sense, the other way around (albeit tendentiously): could this discussion regarding the nature of head-movement have been GB-internal? I will return to these points in Section 4.

3. *Alternatives to core-syntactic head-movement*

Three main alternatives to the earlier form of syntactic head-movement have been proposed since Chomsky (2001), one of which developed to some extent independently of Chomsky's remarks. These are the PF-movement approach, which Chomsky himself advocated, the remnant-movement approach, which partly stems from Kayne (1994), and the "reprojective" approach. I will look at each of these turn, focussing on one case study of how an earlier analysis or family of analyses involving core-syntactic head-movement is replaced by the alternative mechanism.

3.1 PF-movement

To judge from Chomsky's (1995:4.10; 2001:37-8) comments, the alternative he has in mind to syntactic head-movement is a PF–operation. This becomes clear when we consider that the PF-movement alternative is unproblematic in relation to all the arguments Chomsky makes: clearly we do not expect PF-movement to have to obey the Extension Condition or the c-command condition (English Affix-Hopping could be a case of PF head-movement but cannot be syntactic movement (see Section 1)); we expect it to be triggered quite separately from syntactic XP-movement, perhaps to involve morphological "roll-up", as already mentioned, to be subject to special, non-syntactic, locality constraints, and, of course, to lack LF effects. The existence of head-movement(-like) operations in PF is frequently assumed: alongside Affix Hopping one can point to Halpern's (1992) operation of Prosodic Inversion, which switches the positions of a clause-initial enclitic and a potential host, so that the enclitic can "lean left", as required; this operation may underlie many 2nd-position clitic phenomena, and the general approach is characteristic of Distributed Morphology outlined in Embick & Noyer (2001).[5]

[5] On the other hand, Affix Hopping can be handled as a purely morphological reflex of Agree among local heads in the verb-auxiliary system of English, and Prosodic Inversion may well be similar to verb second in being a case where an inflection-bearing head is attracted to C with concomitant XP-attraction to SpecCP (see Starke (1993) on the similarities between clitic-second and V2). It may be that PF movement is not found. If, as suggested in Chomsky (2004), movement is Internal Merge this would make sense to the extent that Merge is not a PF operation. In Distributed Morphology, Merge of feature-bundles in narrow syntax is distinguished from Vocabulary Insertion, which takes place post-syntactically. Vocabulary Insertion should not be seen as a case of Merge, because it is not combinatorial, it does not build structure and it is not recursive. It is thus formally quite distinct from Merge. There is also evidence that LF is sensitive to Affix Hopping; see Siegel (1984).

The question becomes, then, one of providing evidence that head-movement processes which appear to be syntactic are really PF processes. Here, decisive evidence is somewhat lacking. One interesting argument is made by Boeckx and Stjepanović (2001), who propose that pseudogapping, as in (5), provides evidence for PF verb-movement in English:

(5) Although John doesn't eat pizza, he does – pasta.

Starting from Lasnik (1995), examples of this kind have been taken as evidence for syntactic object-shift in English, combined with remnant VP-deletion after object shift (i.e. deletion of *[$_{VP}$ eat (pasta)]* in (5)).[6] Boeckx & Stjepanović (2001) observed that Lasnik's original account of why the verb moves when there is no pseudogapping (to derive *John eats pasta*) in the second conjunct in (5) by the combination of object shift and V-movement), which relies on PF "feature strength" requiring either V-movement or V-deletion by PF, cannot be maintained in the Agree-based theory of movement of Chomsky (2000, 2001). Assuming all three operations (object shift, verb-movement and ellipsis) to be intrinsically unordered, the question then becomes why V-movement followed by ellipsis is not possible, giving the ungrammatical (6):

(6) * … he eats ~~[$_{VP}$ (eats) pasta].~~ .

Boeckx & Stjepanović argue that the question concerns ordering, and point out that object shift must precede both head-movement and ellipsis, while the latter two can appear in either order:

(7) a. Object Shift > ellipsis (head-movement bled) → pseudogapping:
,, he does pasta ~~[$_{VP}$ eats (pasta)]~~.

b. Object shift > head-movement > ellipsis:
.. he eats pasta ~~[$_{VP}$ (eats) (pasta)]~~.

c. *Head-movement > ellipsis (object shift bled):
* .. he eats ~~[$_{VP}$ (eats) pasta]~~.

[6] Traces/copies of moved elements are in round brackets.

They conclude that the right result can be guaranteed if object shift is a syntactic operation, with both ellipsis and V-movement taken to be PF-processes. Hence ellipsis can either precede or follow PF V-movement; in the former case, as in (7a), pseudogapping results, in the latter, VO order results, as in (7b). (7c) is impossible since object shift, as a syntactic operation must precede verb-movement.

However, Baltin (2002: 655) observes that the same movement/deletion options apply to non-verbal predicates such as *fond* in (8) and to phrasal categories as in (9):

(8) Although he isn't fond of pizza, he is (fond) of pasta.

(9) a. Although he isn't very fond of pizza, he is (very fond) of pasta.

b. Although he didn't try to persuade Mary, he did (/tried to persuade) Martha.

In (9a) the gapped string is *very fond*, presumably an AP, and (9b) it is *try to persuade*. Baltin further observes that it seems that the *of*-PP has undergone "object shift" in (8) and (9a), raising questions about Lasnik's initial conclusion. The following examples underscore both points:

(10) a. Although John isn't easier to please than Mary, he is – than Bill.

b. Although John isn't easier to convince the students to talk to than Mary, he is – than Bill.

Here, *than Bill* must have undergone putative "object shift", which is surprising since this category is usually taken to be either a PP an elliptical CP and the pseudogapped constituent is the complex AP, containing a possibly unbounded A'-dependency.

In fact, it appears that the "object shift" operation should really be seen as an optional focussing operation, moving an XP to the left edge of vP (see Belletti (2004) on the idea that the vP, like CP, may have an extended left periphery). This operation seems to be like scrambling in other West Germanic languages, in that it can apply to many XPs, but not readily to small clauses, particles or small-clause predicates (see Johnson (2001: 463) for the same suggestion, and his Note 41 for one or two provisoes):

(11) a. Even though John didn't put Mary down, he did put her up.
b. * … he did up –.
c. Even though John didn't get Mary drunk, he did get her angry.
d. * … he did her angry --.
e. * ... he did angry --.

Let us suppose, then, that English has an XP-movement operation, a highly restricted residue of scrambling, that moves an XP out of VP to the left edge of the vP phase, subject to that element receiving a special interpretation. This operation is associated with VP-deletion, which then applies to the remnant VP, giving pseudogapping. Nothing further needs to be said. In particular, V-movement plays no role in accounting for the salient facts of this construction.

In fact, head-movement may be relevant in one respect, and this points to exactly the opposite conclusion from that drawn by Boeckx & Stjepanović. In examples where VP is headed by a main verb, V-to-T movement is impossible and *do* is inserted in the standard way, in order to bear T's φ and Tense features. Examples like (8, 9a, 10, 11) can also be seen as involving VP-ellipsis combined with obligatory *be*-raising to T. The ungrammaticality of the corresponding examples without *be* can then be taken to argue that V-to-T movement must apply before VP-ellipsis, and hence is a syntactic operation (the ungrammaticality of "*do*-support" here further implies that that operation, too, cannot be a purely PF matter).

The one open question concerns the relation between leftward XP-movement and VP-ellipsis. The latter can clearly apply without leftward XP-movement, but leftward XP-movement appears to be conditioned by VP-ellipsis, in that *he pasta eats/he does pasta eat* are ungrammatical.[7] This fact seems to be connected to the intrinsic link between VP-ellipsis and focus, also manifest in the very well-known fact that the auxiliary cannot be contracted here:

(12) a. *John is fond of pizza, and Bill's – too.
b. *Although he isn't fond of pizza, he's -- of pasta.

A focus feature on v seems required for both VP-ellipsis and optional XP-movement.

[7] Similarly, Boeckx & Stjepanović have no obvious way of ruling out (i):
(i) *Debbie ate chocolate, and Kazuko milk drank.

It appears, then, that Boeckx & Stjepanović's argument does not support the postulation of PF head-movement. Many other cases of head-movement could be treated as PF phenomena, in part for the reasons given Chomsly as summarised in Section 2. However, PF head-movement must be entirely without LF effects, and a number of arguments showing that some cases of head-movement have LF effects have been given, notably by Lechner (2005) (see also Cinque (1999: 184, n. 8), Roberts (forthcoming, Chapter One), Zwart (2001)) have adduced cases where apparent head-movement has LF effects. Roberts (forthcoming, Chapter One) points to the following paradigm (see also McCloskey (1996:89), Kayne's (2000:44)):

(13) a. *Which one of them does anybody like?
b. Which one of them doesn't anybody like?
c. *They succeeded in finding out which one of them anybody liked.
d. *They succeeded in finding out which one of them anybody didn't like.
e. They succeeded in finding out which one of them wasn't liked by anybody.

Here it appears that the NPI *anybody* in subject position in (13b) is licensed by the auxiliary raised to C. This argument depends on the standard assumption that NPIs must be c-commanded by their licensers at LF. Movement of the auxiliary in examples like (13b) above affects LF by altering c-command relations involving the moved item, and as such is head-movement analogue of raising in (14):

(14) a. After the meeting, nobody seemed to anybody to be satisfied with the outcome.
b. *After the meeting, it seemed to anybody that nobody was satisfied with the outcome.

Furthermore, Matushansky (2006:102-4) provides a plausible reason for why it should be the case that verb-movement, in particular, often lacks semantic effects: essentially this is because verbs are predicates. To quote Matushansky "whether we assume that predicates must reconstruct .. or allow them to be interpreted in their final position, the outcome is the same: predicate movement is not reflected at LF" (103).

There may well be reasons, then, to think that not all head-movement takes place at PF. This does not imply that no head-movement takes place at PF, of course, although unambiguous evidence to this effect is lacking (and if the suggestion in Note 5 that Internal Merge cannot take place at PF is correct then it may be that PF head-movement is impossible, after all).

3.2 Remnant phrasal movement

To some degree as a direct response to Chomsky's (2001) arguments, summarised in Section 2, and to some extent as a consequence of the re-evaluation of the status of clitic pronouns following on from Kayne (1994), a number of authors have proposed remnant-movement accounts for some of the phenomena previously handled as head-movement, including verb-movement of various kinds (see Koopman & Szabolcsi (2000), Nilsen (2003), Müller (2004), Wiklund & Bentzen (2007), Wiklund, Hrafnbjargarson, Bentzen & Hróarsdóttir (2007), Bentzen (2007, to appear, and several of the contributions in Mahajan (2003)); see also the recent treatments of various forms of inversion in French in Kayne & Pollock (2001), Poletto & Pollock (2004), Pollock, Poletto & Munaro (2003), Pollock (2006), and several of the papers on verb-initial languages in Carnie, Harley & Dooley (2005), and, on the syntax of nominals, Shlonsky (2004), Cinque (2005, forthcoming) and the references given there).

These approaches share the central idea that analyses positing head-movement relations of the type schematised in (15) should be replaced by analyses of the general type in (16):

(15) … H … [$_{XP}$ Z (H) Y] …

(16) … XP … Z .. Y .. [$_{XP}$ (Z) H (Y)] …

Other things being equal, both scenarios convert underlying *–ZHY-* order to surface *–HZY-*. In (15), this is achieved by H-movement out of the category XP containing H, Z and Y prior to movement. In (16), on the other hand, H does not move: instead XP moves, but thanks to presumably independent operations moving Z and Y, the moved XP contains only H; all the other material has been moved out of XP before XP-

movement takes place. XP is thus a "remnant category", in that it contains only a subset of the elements it contained at an earlier stage of the derivation (this point should really be stated in terms of the categories realised at PF, since copies/traces are presumably present in core syntax but deleted in PF). Movement of XP in scenarios like that schematised in (16), is thus referred to as "remnant movement".

Strictly speaking, the term "remnant-movement" does not denote a form of movement, but rather a (subpart of) a derivation where, given a complex constituent $[_{XP}\ Y\ Z\ ]$, both movement of Y or Z from XP and movement of XP itself take place. Derivations of this type are allowed and attested quite independently of the issues surrounding head-movement. Typically, this movement is subject to certain constraints, though. In particular, various notions of Freezing and (strict) cyclicity are relevant. Freezing (originally put forward by Ross (1967), Wexler & Culicover (1980)) bans movement out of moved constituents; this forces movement of Y or Z to take place before XP-movement in the derivation (if XP is a cyclic domain, then the Strict Cycle has the same effect). Moreover, the Strict Cycle, on many formulations, requires XP to move to a higher position than Y or Z. The Extension Condition, combined with Freezing, will also have this effect. The schema in (16) reflects this order of operations.

Perhaps the best-known independent motivation for remnant-movement comes from so-called "remnant topicalisation" in German, as in examples such as the following, discussed by den Besten & Webelhuth (1990):

(17) a. Gelesen hat er das Buch nicht.
Read has he the book not.
"He hasn't read the book."

b. [$_{VP}$ (das Buch) gelesen] hat er das Buch nicht (VP).

Here, *das Buch* has undergone scrambling, an operation which productively applies to definite DP objects, raising them outside VP to some TP-internal position in German (the exact nature and trigger for scrambling in German and elsewhere is much debated; see Grewendorf & Sternefeld (1990), Corver & van Riemsdijk (1994), Thráinsson (2000)). This is followed by VP-fronting to the first position in the clause, usually thought to be SpecCP, satisfying the V2 constraint here. This combination of operations is entirely licit, and explains what would otherwise be an anomalous V2

construction, involving just a participle preceding the inflected verb. (18), from Müller (1998), illustrates the interaction of remnant-movement with Freezing and the Strict Cycle:

(18) *Worüber hat [$_{DP}$ ein Buch (worüber)] keiner (DP) gelesen?
What-about has a book noone read
"What did noone read a book about?"

(18) can be derived by moving the DP *ein Buch worüber* first, and then by subextracting *worüber*; this violates Freezing. Alternatively, if *worüber* is first moved, and then DP, movement of the DP violates the Extension Condition. We see, then, that remnant-movement and the constraints relevant to it are motivated.

Müller (2004) argues that the analysis of verb second (V2) constructions which postulates two separate movements, one head-movement of the verb and the other XP-fronting, originally proposed by den Besten (1983), should be replaced by a single operation of remnant-fronting. Specifically, Müller proposes that a vP evacuated of all overt material other than the verb and a single constituent on the left edge undergoes this fronting operation; note that here the remnant category must contain more than just an unmoved head and so the schema in (16) does not exactly apply. The creation of the appropriate initial domain is achieved by what Müller calls the Edge Domain Pied-piping Condition, which states exactly this: only one maximal constituent, occupying the left edge of vP, can be present in a vP which moves (the definition of Edge Domain is given in (21) below). This analysis is claimed to have certain interesting empirical advantages, and, notably to have the theoretical advantage of allowing us to dispense with a well-known case of head-movement.

The central innovation in Müller's analysis is the idea that V2 is derived by a single movement operation, remnant vP-fronting, rather than by the interaction of movement of the finite verb and movement of an XP. Thus, instead of the standard derived structure for an object-initial V2 clause as in (19) we have (20):

(19) [$_{CP}$ Das Buch [$_{C'}$ hat-C [$_{TP}$ Fritz [$_{vP}$ (Fritz) [$_{VP}$ (das Buch) gelesen] (hat)] (hat)]]]
the book has Fritz read

(20) [$_{CP}$ [$_{vP}$ Das Buch (Fritz) (VP) hat] [$_{C'}$ C [$_{TP}$ Fritz [$_{T'}$ [$_{VP}$ (das Buch) gelesen] [$_{T'}$ (vP) T]]]]]

the book has Fritz read

"Fritz has read the book."

As Müller points out:

In this approach, the pre-V/2 position is occupied by whatever category happens to be at the left edge of vP earlier in the derivation – this will typically be the subject NP or an adverb, but, after scrambling, it may also be an object NP, a PP, a CP, or a VP (complete or remnant ..). (pp. 182-3)

In addition, there is no reason to postulate head-movement; in examples of the kind in (20), *hat* is assumed to have merged directly in v. Where a main verb appears in second position, it has not moved to v, but rather counts as being on the edge of vP owing to the first clause of the definition of Edge Domain, which runs as follows (Müller's (6), p. 184):

(21) *Edge Domain:*

A category α is in the edge domain of a head X iff (a) or (b) holds:

a. α is the highest overt head reflexively c-commanded by X.

b. α is a specifier that is not c-commanded by any other specifier in XP, and that precedes the edge domain of X.

Müller claims a number of empirical advantages for his approach, and it has inspired other analyses of verb-movement, including V2, in Germanic, notably in Wiklund & Bentzen (2007), Wiklund, Hrafnbjargarson, Bentzen & Hróarsdóttir (2007), Bentzen (2007, to appear).

In an independent development, a number of papers on the syntax of inversion and verb-movement in French and other Romance varieties have suggested replacing earlier head-movement analyses with alternatives based on remnant-movement. Poletto & Pollock (2004), Pollock, Poletto & Munaro (2003) and Pollock (2006) argue for a remnant-movement analysis of verb-movement into the C-system in French (and other Romance varieties). Their arguments are based on Kayne's

(1994:42-46) discussion of the landing site of clitics in Romance. Consider first a basic example with a direct-object clitic (here from Italian):

(22) Voi lo vedete.
You(pl) him/it see
"You see him/it."

Kayne adopts three postulates. First, that morphologically derived forms such as *vedete* are syntactically formed, possibly by syntactically combining the root *ved-* with the theme vowel *–e-* and the ending *–te*. Second, that the LCA applies to sub-word-level operations, and, third, that the LCA bans multiple head-adjunction. Given these three postulates, the clitic would have to adjoin to the verb root *ved-*, followed by adjunction of *[lo ved-]* to (the functional head occupied by) *–e-* and then adjunction of *[[lo ved-] –e-]* to *–te*. Where the verb bears a prefix, as in *lo prevedete* ("you foresee it"), the clitic would have to attach to the prefix.

Kayne goes on to suggest that a more plausible option is to assume that clitics adjoin to empty functional heads. Kayne further observes enclisis to infinitives and imperatives of the type in (23):

(23) a. Fais-le. (French)
Do it
b. Parlargli sarebbe un errore. (Italian)
To-speak.to-him would-be a mistake

Since it is very likely that the verb moves to C in imperatives like (23a) (see among others Rivero (1994a,b)), and that the infinitive is in a "high" position in (23b) (Belletti (1990), Kayne (1991)), Kayne concludes that in general verb-movement to C does not "carry along" clitics. It then follows that, in a French example like (24), involving "subject-clitic inversion" with an object proclitic on the inverted auxiliary, the clitic+auxiliary combination has not moved to C:

(24) L'as-tu fait?
It.have-you done?
"Have you done it?"

Kayne follows Sportiche's (1999) proposal that there may be V-movement to C at LF, hence accounting for the root nature of the construction (he suggests that the clitic may delete at LF; see his Note 16). Finally, Kayne observes high-register examples, which show the order Clitic-Adverb-Infinitive (e.g. .. *le bien faire* .. "it well to-do") support the idea that the clitic and the verb do not have to combine.

Poletto & Pollock (2004), Pollock, Poletto & Munaro (2003) and Pollock (2006) endorse Kayne's general conclusion that clitics and verbs cannot and do not combine in syntax, but propose that, instead of covert verb-movement into the C-system in examples like (24), there is overt remnant movement. The derivation of (24) would proceed as follows:

(25) a. Tu $[_{XP}$ le $[_{YP}$ as $[_{ZP}$ fait]]] → (movement of ZP)
b. Tu $[_{ZP}$ fait] $[_{XP}$ le $[_{YP}$ as (ZP)]]] → (remnant movement of XP)
c. $[_{XP}$ le $[_{YP}$ as (ZP)]] tu $[_{ZP}$ fait] (XP)

Remnant XP-movement is triggered by the interrogative feature of the attracting head, which is part of an articulated C-system. It is unclear what the trigger (or the landing site) of ZP-movement (probably vP-movement) is, as well as the cliticisation operation itself.[8] Kayne & Pollock (2001) propose a similar analysis of French Stylistic Fronting, and Pollock, Poletto & Munaro (2003) extend the approach to interrogatives in various Romance varieties.

A third case where remnant-movement has influentially replaced an earlier head-movement analysis comes from verb-initial languages, in particular Macronesian languages showing an alternation between VSO and VOS orders (Massam & Smallwood (1997), Massam (2000), Rackowski & Travis (2000), and many of the papers in Carnie, Dooley & Harley (2005)). In her study of VOS and VSO in Niuean, for example, Massam (2000) argues that there is an operation fronting a verbal constituent, and that this constituent is fronted to a position within TP. She then shows that there is a general operation which fronts non-verbal

[8] It is also worth pointing out that although Poletto & Pollock (2004) and Pollock (2006) share with Müller (2004) the basic idea that V2-type verb movement (full in the latter case, residual in the former), should be reanalysed as remnant movement, they do not suggest that remnant movement should be a global replacement for head-movement, as they continue to assume both V-to-T movement of the type argued for in Pollock (1989) and a head-movement analysis of (subject) cliticisation, hence *tu* in (25c) moves to the head whose specifier the fronted remnant XP occupies. In that case, XP-movement can be seen as remnant TP-movement.

predicates which are clearly larger than heads, e.g. relative clauses. Third, Massam shows that what has been called noun-incorporation in Niuean (e.g. by Baker (1988)) cannot be movement of N into V (*pace* Baker) since there are clear cases where a constituent larger than N undergoes this operation. She proposes instead that putative noun-incorporation is really the absence of object-shift to a VP-external position. In that case, the fact that the apparently incorporated noun moves with the verb shows that what is moved is VP rather than V. VOS order is thus derived by VP-fronting, and VSO by object-shift to a VP-external position combined with remnant VP-fronting, as shown in (26):[9]

(26) a. $[_{TP}\,[_{VP}\,V\;\;O\,]\;T\;\;[_{vP}\;\;S\;\;v\;..\;(VP)\,]]$ -- VOS

b. $[_{TP}\,[_{VP}\,V\;\;(O)\,]\;T\;\;[_{vP}\;\;S\;\;v\,[_{AbsP}\;\;O\;\;(VP)\,]]]$ -- VSO

As (26) shows, the landing-site of VP-fronting is taken to be SpecTP; Massam argues that this is motivated by essentially the same property as that which causes the subject to raise to SpecTP in languages like English, French and Mainland Scandinavian: the operations "can be seen as two reflections of a single EPP predication feature" (Massam (2000:111)). This type of analysis, first put forward by Massam & Smallwood (1997), and developed by Rackowski & Travis (2000) as well as several of the papers in Carnie, Dooley & Harley (2005), has been applied to a number of languages which display both VOS and VSO orders (mainly but not exclusively Macronesian and Mayan languages; unlike rigidly VSO languages such as the Celtic languages, where it is at the very least much harder to motivate a remnant VP-fronting analysis). Here too, though, the question of the trigger for some of the movements arises (see Chung (2005)).

Let us now briefly evaluate remnant-movement approaches against Chomsky's (2001) arguments, given in Section 2. First, it is clear that remnant-movement avoids the problems head-movement causes for the Extension Condition and the definition of c-command. Since it is XP-movement, it presumably extends the root. Similarly, as XP-movement, the issues concerning the Head Movement Constraint and successive-cyclic vs. roll-up movement are directly solved by the

[9] AbsP here stands for Absolutive Phrase, which Massam suggests may correspond to AgrOP in more familiar languages. My summary here glosses over the complication that Niuean is an ergative language and Massam's treatment of the assignment of ergative and absolutive case.

postulation of XP-movement (although it is not clear why remnant-movement should, even apparently, obey a strong locality condition such as the Head Movement Constraint). On the other hand, remnant-movement is expected to have LF-effects; it is entirely unclear why this particular type of XP-movement should be exempt from these.

Where the remnant-movement approaches are problematic, as has frequently been pointed out, is in relation to the question of movement triggers. This may not always be a serious problem in the case of the movement of the remnant itself, but it is frequently difficult to see what drives the other movements (indicated as movement of Z and Y in (16)) exactly where the larger XP moves to a higher position. In the case of the derivation in (25), for example, this would apply to ZP-movement, as mentioned above.

A further issue arises in connection with the apparent LF effects of head-movement mentioned at the end of the previous section. Recall that we observed that the NPI *anybody* in subject position in (13b) is licensed by the auxiliary raised to C. This argument depends on the assumption that NPIs must be c-commanded by their licensers at LF. Movement of the auxiliary in examples like (13b) above affects LF by altering c-command relations involving the moved item. However, if we consider the remnant-movement alternative, this conclusion would not follow. To see this, suppose that English subject auxiliary inversion does not have the form in (27a), the "traditional" T-to-C movement analysis, but rather that in (27b), involving remnant TP-movement:

(27) a. T+C .. [$_{TP}$ Subj (T) vP]

b. TP … C … Subj … vP … ([$_{TP}$ (Subj) T (vP)])

Obvious questions arise here concerning the nature and landing sites of subject-movement and vP-movement, but let us leave these aside here. Now, if remnant-movement of the general type in (27b) were involved in subject-aux inversion, and if c-command is retained as the relation determining polarity-licensing, then the definition of this relation would have to be complicated so as to allow the auxiliary to c-command out of TP into the complement domain of TP here. So negation inside TP, possibly attached to an auxiliary in T, does not c-command the subject, and cannot do so on any plausible, simple definition of c-command. That T rather than TP counts as

the relevant licensing element can be seen from cases where *not* does not raise with the auxiliary and subject-auxiliary inversion does not license the NPI, such as **Which of them does anybody not like?* To the extent that there are LF effects associated with head-movement (*pace* Chomsky), then, remnant-movement approaches may have difficulties with c-command effects.

In conclusion, reanalysing head-movement as remnant-movement avoids a number of the problems Chomsky pointed out for head-movement, although if anything the "trigger problem" may be exacerbated. Reconsidering head-movement in this light has been productive in the cases of N-to-D movement, creation of verb-clusters of the West Germanic/Hungarian type (see Koopman & Szabolsci (2000) and the papers in Kiss and van Riemsdijk (2004)), and has led to new ideas in the case of V2 and certain cases of inversion in Romance. This approach is almost certainly, then, an alternative in some cases, but it remains unclear to what extent it represents a global alternative to head-movement, and, given the "trigger problem", it is not clear that it is conceptually simpler.

3.3 "Reprojective" movement

Still another strand which has been pursued as part of the general reconsideration of the nature of head-movement is represented by a class of analyses which we can collectively label "reprojective". This approach has been developed primarily by Bury (2003, 2007), Donati (2006), Koeneman (2000) and Surányi (2005, 2007, 2008). The basic idea is to take head-movement to be syntactic movement, but to treat it as arising from a different set of conditions from XP-movement. Chomsky (1995:256-60) argues that where a new category γ is formed by movement of α to β, γ must always project the target of movement: hence DP-movement attracted by T will create a new projection of T, wh-movement attracted by C will create a new projection of C, etc. Bury, Donati and Surányi suggest that this may not always be the case, and that "reprojective" movement may arise, where the moving category gives its label to the new category formed by movement. Bury develops this proposal in a very interesting way in connection with phenomena connected to both V-initial and V2 languages, treating V-movement as "reprojective" in this sense. He also applies the approach to free relatives. Here I will briefly summarise the main proposals made by Donati (2006) in her analysis of free relatives and related constructions.

Donati's main concern is the basic conceptual question of why we find phrasal movement at all, since head–movement involves moving less material. She suggests that Chomsky's (2001) proposal to eliminate head-movement is inadequate, since it cannot in principle rule out head-movement to a specifier position (this point is also made by Matushansky (2006), Roberts (2005), Toyoshima (2000), Vicente (2006)). Again similar to what we have seen here, she suggests that the Head Movement Constraint is irrelevant to the question of the existence of head and phrasal movement, in that locality constraints act on the search operation, not on movement itself, hence a single set of locality constraints should govern both types of movement. She further observes that there are empirical doubts about the HMC, citing "long verb-movement" in Breton (see Borsley, Rivero & Stephens (1996)) and the similar cases in South Slavic and archaic Romance discussed by Ćavar & Wilder (1994), Lema & Rivero (1990, 1991), Rivero (1991, 1993a,b, 1994a,b, 1997), Rivero & Terzi (1995)). Instead, she adopts the chain uniformity condition of Chomsky (1995:253)) in (28) and the minimality condition on Merge in (29):

(28) "A chain is uniform with regard to phrase structure status" (where "phrase structure status" means the "(relational) property of maximal, minimal or neither").

(29) Merge just enough material for convergence.

(29) applies to both internal and external merge. Finally, she assumes that a head, when merged either externally or internally, projects; XPs, on the other hand, do not. Thus, for Donati, head-movement is always and only "reprojective".

Donati then goes on to show a minimal contrast involving movement of a [+wh]D. This gives the two possibilities in (30):

(30) a. CP

DP C

b. DP

D C

In (30a), we have the derived structure of a wh-interrogative. Donati argues that the wh-feature cannot move as a head in this case as it would turn the interrogative clause into a DP. Thus, requirements of LF convergence (the structure must be interpretable as an interrogative clause) cause the non-minimal DP-movement option to be taken. However, "in a context compatible with DP-selection and showing no phrasal pied-piping" (Donati (2006:32)), the option in (30b) should be available. This, Donati argues, is what we find in free relatives and comparatives.

For free relatives, Donati's evidence comes from paradigms like the following:

(31) a. *I will visit [what town] you will visit.
b. I wonder [what town] you will visit.
c. I will visit [what] you will visit.

(31b) clearly contains an indirect question, i.e. a CP-complement to *wonder* with *what town* in its Specifier. On the other hand, *visit* does not take an indirect-question complement, or indeed any kind of CP, but only a DP. The complement in (31c) is thus a DP, a free relative. Pied-piping to the edge of a free relative is impossible, as (31a) shows (following Kayne (1994), Donati assumes that *whatever*-type relatives – as in *I will visit whatever town you will visit* – are not in fact free relatives). (31c) thus involves "reprojective" movement of a [+wh]D, giving rise to a derived structure like (31b).

Donati goes on to argue the same for comparatives, known to involve wh-movement since Chomsky (1977). The idea that comparatives are complex nominals is supported by the fact that they express a description of a degree, and by the fact that this expression can enter into scope ambiguities of the type first discussed in Russell (1905):

(32) I thought your yacht was bigger than it is.

If the comparative expression (*than it is*) is outside the scope of *think*, we have a non-contradictory reading for (32); if it is inside the scope of *think* we have contradictory reading (cf. "I thought your yacht was bigger than itself"). Third, comparatives are strong islands for extraction, suggesting they are complex DPs (here *eat* is elided while *what* and *x-quickly* are wh-moved):

(33) *What do you eat the soup more quickly than Paul does (eat) (what) (x-quickly)?

Finally, Donati gives evidence that in Romanian and Bulgarian the same wh-element moves overtly as a head in comparatives but as a phrase in interrogatives. In fact, the same can be shown with non-standard varieties of English which allow *what* to appear in comparative subdeletion and to act as an adnominal wh-determiner:[10]

(34) a. Mary ate more cookies that what she ate [(what) candies].
b. * Mary ate more cookies that what candies she ate (what candies).
c. What candies did she eat (what candies) ?
d. *What did she eat [(what) candies]?

Donati (2006:39) concludes "there is no principled reason for wh-movement to be restricted to phrases".

Once again, let us consider these proposals in relation to Chomsky's (2001) arguments. "Reprojective" movement does not target heads, and so the Extension Condition and c-command problems Chomsky raises do not apply in this case. The triggering problem appears to be dealt with by Donati by LF: if the movement is not reprojective, one kind of structure and interpretation must result; if it is, then a different one results. The syntax itself allows either option. In a sense, then, Donati has LF act as a filter on the syntactic derivation. Regarding onward movement, presumably the possibilities here are determined by the reprojection option. Sticking

[10] The grammaticality of (34a) implies that even some varieties of English allow left-branch extraction, at least in comparative subdeletion cases like this. The comments in Donati (2006:37-8) could provide a basis for understanding why the left-branch extraction is possible in (34a) but not (34d). For discussion of whether comparative subdeletion involves movement or unbounded deletion, see Bresnan (1976), Chomsky (1977).

to Donati's example with [+wh]D, we can note that free relatives are unbounded and subject to island constraints, and therefore must involve standard, non-reprojective wh-movement on earlier cycles prior to a last step of reprojective movement:

(35) a. I will visit what Tom says Bill thinks Mary believes you will visit.
b. ?*I will visit what Tom believes the claim you will visit.

Conversely, the entire DP formed by reprojection can move, and indeed undergo A-movement:

(36) a. What you will visit, (Tom says) I will visit.
b. What you will visit seems to have been visited by many tourists.

So it seems that each step of movement can in principle be either reprojective or not, but once reprojection takes place, it cannot be "undone". The latter constraint can arguably be seen as an instance of the general "no-tampering" condition, in that once D has projected the label cannot be "unprojected" but, conversely, until reprojection takes place, it is always possible in principle. The locality properties are directly tied to the nature of projected category: DPs, as in (36b) can undergo A-movement and the wh-phrases can undergo A'-movement, with the moved category obeying standard locality conditions in each cases. The LF properties of the structure resulting from movement category are crucial, as we have seen; on the other hand, PF appears to play no role in this approach.

Again, reprojection appears to be a valid alternative approach which avoids the general difficulties discussed by Chomsky. It leads to an interesting account of free relatives, and, in Bury (1993), of some cases of verb-movement. How far it can be extended as a global alternative to head-movement remains to be seen, however (see Koenemann (2000), Biberauer & Roberts (2008) for a reprojective account of V-to-T movement).

3.4 Conclusion

Here we have looked at the three main alternatives to standard head-movement that have been discussed in the literature, in many cases directly responding to Chomsky's

(2001) comments. No single version is entirely free of problems, and none appears to be a global alternative to "traditional" head-movement in the sense that it is clear that all former cases of head-movement can and should be reanalysed in the relevant terms. This may in fact be a good state of affairs: it is quite possible that the mechanisms of head-movement were overextended in the earlier approach. At least in the DP and in the areas of verb-clustering it really seems that XP-movement analyses represent a valid alternative, while reprojection looks promising for some cases of verb-movement, with perhaps PF-movement valid for others. Here I leave these questions open: now I want to turn to more conceptual issues raised by some of the thinking behind the minimalist programme.

4. Head movement and the Minimalist Programme

What the above sections have shown, I hope, is that there are at present a range of views on the question of head-movement. One could claim that for certain core cases, say French "V-to-T" movement or Germanic V2, up to four competing analyses are available: the traditional, GB-style head-movement one, a remnant vP-movement one, a PF one and a reprojection one. The question is then, obviously, which of the available analyses is the most successful, both empirically and theoretically? To be unable to answer this question in any immediately straightforward way seems to me to be a healthy state of affairs: the phenomena are complex and the implications of and relations among the various types of analysis not easy to tease out. So there is no reason to expect an immediate or simple answer.

But the question I want to address here is a slightly different one: which, if any, of these approaches is likely to be the most successful one given the overall goals of the Minimalist Programme, as Chomsky has articulated this in his recent work (Chomsky (1993, 1995, 2000, 2001, 2002, 2004, 2005a,b, 2007, 2008)? To put it another way, do the particular goals of minimalist theory contribute anything to deciding which, if any, of the alternatives we have seen might be the best overall approach to the phenomena of head-movement?

Obviously we cannot answer that question without reminding ourselves of the goals of the minimalist programme. The simplest way to put this is to say that, having got an inkling of the nature of UG through the GB version of principles-and-parameters theory, the minimalist programme has as its goal to refine and axiomatise

that conception by asking, why, among all conceivable UGs, we have the one we have? To an extent, as Chomsky (2005b:1) has pointed out, "the issues can be recast as metaphysical rather than epistemological: Is that how the world works?" We move from asking questions about knowledge of language, questions whose answers have led to the postulation of UG, to asking questions about the kind of world which produces a mental object like UG with the properties we observe it to have.

This, in turn, has led to an emphasis on the "third factor" determining the nature of the adult language faculty. To see this what this means, observe that adult competence is the result of the interaction of three factors: (i) experience of the primary linguistic data (PLD), we need this to learn the vocabulary and set the parameters of our native language; (ii) Universal Grammar, the innate endowment which makes it all possible, construed as a set of principles with parameters initially open; (iii) Principles not specific to the faculty of language.[11] These principles constitute the third factor in language design, and include "(a) principles of data analysis that might be used in language acquisition and other domains; (b) principles of structural architecture and developmental constraints that enter into canalization, organic form, … , including principles of efficient computation .. It is the second of these subcategories that should be of particular significance in determining the nature of attainable languages" (Chomsky (2005a: 6)). The first and second questions may answer the epistemological question, but third-factor postulates seem to be implied in answering the metaphysical question. In a sense, we have to move beyond UG and, so, beyond explanatory adequacy (in the Chomsky (1964) sense).

In pursuing the axiomatisation of GB principles, we subject every postulate to a "minimalist critique": do we really need it? Can it be reduced to something else? We want to get back to the first principles of syntax. We want to reduce the theoretical postulates to those which are (virtually) conceptually necessary. At the same time, we want our explanatory postulates to relate to the higher level of explanation constituted by the attempt to answer the metaphysical question by invoking third-factor considerations. The Strong Minimalist Thesis (SMT) expresses one hypothesis which can do this:

---

[11] This isn't an entirely new idea: cf. "there is surely no reason today for taking seriously a position that attributes a complex human achievement entirely to months (or at most years) of experience, rather than to millions of years of evolution or *to principles of neural organization that may be even more deeply grounded in physical law*" (Chomsky (1965:59), emphasis mine).

(37) "Language is an optimal solution to legibility conditions" (Chomsky 2000: 97)

The notion of "legibility conditions" here relates to interface properties. So the idea is that the core computational system of syntax provides the optimal way of relating an arbitrary set of lexical items to the interfaces (PF and LF, for simplicity) in such a way as to satisfy whatever conditions the intrinsic properties of the lexical items and the interfaces may impose.

Evaluating head-movement, in any of its potential technical guises, against the SMT is difficult, since movement in general appears to be an unnecessary complication. Surely a system which lacked movement operations (of any kind, A', A, head- or anything we might imagine) is simpler and more optimal than a system with such operations. Chomsky (2004:110) provided a compelling negative response to this conjecture: "SMT entails that Merge of α, β is unconstrained, therefore either *external* or *internal*. Under external Merge, α and β are separate objects; under internal Merge, one is part of the other Merge yields the property of 'displacement'". To the extent that movement reduces to Internal Merge (IM), then, we expect to find it in natural language.

Concerning the status of narrow-syntactic head-movement, we might then reason that, all other things being equal, IM and EM are supposed to be exactly the same operation except that IM takes place "within" a structure in the process of being while EM introduces the element to be merged from outside. Since EM quite uncontroversially applies to heads, i.e. single lexical items or feature bundles, we need a very good reason to treat IM in a different way (this point is made by both Donati (2006) and Roberts (forthcoming)). If head-movement is absent altogether, or restricted to the PF interface, there must be an explanation for this in terms of what differentiates IM and EM.

Of course, Merge is restricted to a search space. EM can only look to the Numeration; IM is subject to syntactic locality constraints. So if we can find a reason in the theory of locality for the absence of syntactic head-movement, we would have a principled reason to exclude it from narrow syntax. The A-over-A principle is a good candidate.

Consider the formulation of this condition given in (3) (from *Language and Mind* (Chomsky (2006:45)):

(38) If a transformation applies to a structure of the form

$[_S \ldots [_A \ldots] \ldots] \ldots$

for any category A, then it must be interpreted so as to apply to the *maximal* phrase of the type A.

(Here "maximal" is not intended in the X'-theoretic sense, but simply as the largest phrase of type A, in the sense that A should not be dominated by further occurrences of A). A non-maximal occurrence of A in (38) could be construed as the head of A. Then (38) would in general block head-movement. Hence the principled exclusion of head-movement from narrow syntax might depend on the extent to which a version of (38) can be integrated into the theory of locality. Rackowski & Richards (2005) make an interesting proposal in this direction, arguing that a version of the A-Over-A Principle is a condition on Probe-Goal Agree, and may derive some of the effects of the Phase Impenetrability Condition (PIC). If this can be fully achieved, then we can eliminate many cases of head-movement from core syntax on principled grounds, while maintaining that movement is IM. However, at least some of the cases discussed by Donati would remain. There is also evidence that some kinds of predicate-cleft constructions in some languages involve unbounded, island-sensitive, and hence A'-like, verb-movement to a specifier position (Vicente (2006, 2007), Landau (2006)). Assuming the verb is extracted from VP in these cases, the A-Over-A Condition would be violated; hence it is not in fact clear that the A-Over-A Condition applies in such a way as to ban head-movement in general.

Where does this leave the different approaches we have looked at? If Move is reduced to IM, it arguably cannot apply in PF. Therefore "PF-head-movement" must really be something else, such as concatenation of heads and affixes, or rebracketing along the lines of Marantz's (1984, 1988) conception of morphological merger (see also Embick & Noyer (2001), Roberts (2005), Matushansky (2006)), and cannot have LF effects (see Note 5). As already suggested, the remnant-movement approach is compatible with the total elimination of head-movement from PF and narrow syntax: however, this approach too may have problems accounting for some kinds of LF

effects, as we saw in Section 3.2. On the other hand, "reprojection" approaches may involve a complication of the theory of movement.

Very tentatively, then, we can perhaps conclude that the "pure PF" and remnant-movement alternatives are equally attractive in terms of the SMT, since they add nothing to what we appear to have to assume anyway regarding movement/IM, and this seems to be most compatible with the SMT. However, both approaches appear to have problems with some LF effects of head-movement. This suggests a combined approach: we could extend the operation of head-to-specifier movement (independently needed for some types of predicate cleft) and combine it with a PF rebracketing operation along the lines of Marantz's (1984, 1988) notion of merger. This would allow us retain the idea that IM can apply to heads, and allow an account of the observed LF effects, while at the same time acknowledging that these cases of head-movement are partially morphological. This kind of approach is advocated in Roberts (2005) and Matushansky (2006); the difficulty with it is that the merger operation really has to be part of the head-movement operation in order to avoid the difficulties with c-command, the Extension Condition and successive cyclicity pointed out by Chomsky. This entails something of a departure from "pure" IM, and so again creates a conceptual difficulty. It seems, then, that some alternative notion of incorporation may after all be needed.[12]

References

Baker, M. (1985) *Incorporation: A Theory of Grammatical-Function Changing*. PhD dissertation, MIT.

Baker, M. (1988) *Incorporation: A Theory of Grammatical-Function Changing*. Chicago: University of Chicago Press.

Baltin, M. (2002) "Movement to the higher V is remnant movement," *Linguistic Inquiry* 33: 653-9.

Belletti, A. (1990) *Generalized Verb Movement*. Turin: Rosenberg & Sellier.

Belletti, A. (2004) Aspects of the Low IP Area. In L. Rizzi (ed) *The Structure of IP and CP*. Oxford: Oxford University Press, pp. 16-51.

Bentzen, K. (2007) *Order and Structure in Embedded Clauses in Northern Norwegian*. PhD dissertation, University of Tromsø.

Bentzen, K. (to appear) 'Subject positions and their interaction with verb movement.' To appear in *Studia Linguistica*.

[12] A further issue concerns Agree: why is this operation not subject to the A-Over-A Condition and what, if any, is its connection to head-movement? One might expect there to be some connection since Agree is a head-head relation. Perhaps we could envisage an approach which derives all the properties of head-movement from Agree, adding essentially nothing to our conception of that operation. Such an approach would be conceptually very appealing, and is developed in Roberts (forthcoming). But this is not the place to evaluate that alternative.

den Besten, H. (1983) “On the interaction of root transformations and lexical deletive rules,” in W. Abraham (ed) *On the Formal Syntax of the Westgermania*. Amsterdam: John Benjamins, pp. 47-131.
den Besten, H. & G. Webelhuth (1990) “Stranding.” In G. Grewendorf & W. Sternefeld (eds) *Scrambling and Barriers.* Amsterdam: John Benjamins, 77-92.
Biberauer, T. & I. Roberts (2008). Subjects, Tense and Verb movement in Germanic and Romance. *Cambridge Occasional Papers in Linguistics* 3: 24-43.
(Available on-line at: http://www.ling.cam.ac.uk/COPIL/Vol3(Online)/vol3.html)
Boeckx, C. & S. Stjepanović (2001) “Head-ing towards PF.” *Linguistic Inquiry* 32: 345-355.
Borsley, Robert D., Maria-Luisa Rivero and Janig Stephens (1996) “Long head movement in Breton,” in R. Borsley & I. Roberts (eds) *The Syntax of the Celtic Languages*, Cambridge: Cambridge University Press, 53-74.
Bresnan, J. (1976) “Evidence for a theory of unbounded transformations,” *Linguistic Analysis* 2: 353-393.
Bury, D. (2003) *Phrase Structure and Derived Heads.* PhD dissertation, University College, London.
Bury, D. (2007) Verb movement and VSO-VOS alternations. *UCL Working Papers in Linguistics* 19.
Carnie, A., H. Harley & S. Dooley (2005) *Verb First.* Amsterdam: John Benjamins.
Ćavar, D. & C. Wilder (1994) “Long head movement? Verb movement and cliticization in Croatian.” *Lingua* 93: 1-58.
Chomsky, N. (1957) *Syntactic Structures*. The Hague: Mouton.
Chomsky, N. (1964) *Current Issues in Linguistic Theory*. The Hague: Mouton.
Chomsky, N. (1965) *Aspects of the Theory of Syntax.* Cambridge, Mass.: MIT Press.
Chomsky, N. (1977) “On Wh-Movement.” In A. Akmajian, P. Culicover & T. Wasow (eds) *Formal Syntax*. New York: Academic Press, pp. 71-132.
Chomsky, N. (1986) *Barriers.* Cambridge, Mass.: MIT Press.
Chomsky, N. (1991) “Some Notes on Economy of Derivation and Representation,” in R. Freidin (ed) *Principles and Parameters in Comparative Grammar*. Cambridge, Mass.: MIT Press, pp. 417-454.
Chomsky, N. (1993) “A Minimalist Program for Linguistic Theory,” in K. Hale & S.J. Keyser (eds) *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press, pp. 1-52.
Chomsky, N. (1995) *The Minimalist Program*. Cambridge, Mass.: MIT Press.
Chomsky, N. (2000) “Minimalist Inquiries: The Framework,” in R. Martin, D. Michaels & J. Uriagereka (eds) *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, Mass.: MIT Press, pp. 89-156.
Chomsky, N. (2001) “Derivation by Phase,” in M. Kenstowicz (ed) *Ken Hale: A Life in Language*. Cambridge, Mass.: MIT Press, pp. 1-52.
Chomsky, N. (2002) *On Nature and Language*. Cambridge: Cambridge University Press.
Chomsky, N. (2004) “Beyond Explanatory Adequacy,” in A. Belletti (ed) *Structures and Beyond: The Cartography of Syntactic Structures, Volume 3,* Oxford: Oxford University Press, pp. 104-131.
Chomsky, N. (2005a) “Three Factors in Language Design,” *Linguistic Inquiry* 36: 1-22.
Chomsky, N. (2005b) “Biolinguistics and the Human Capacity,” lecture given at MTA, Budapest, May 17, 2004; ms. MIT

Chomsky, N. (2007) “Approaching UG from below,” in *Interfaces + Recursion = Language? Chomsky’s Minimalism and the View from Syntax-Semantics*, U. Sauerland and H.-M. Gärtner. Berlin: Mouton de Gruyter, pp. 1-29.
Chomsky, N. (2008) “On Phases,” in R. Freidin, C.P. Otero and M.-L. Zubizarreta (eds) *Foundational Issues in Linguistic Theory, Essays in Honor of Jean-Roger Vergnaud*. Cambridge, Mass.: MIT Press.
Chomsky, N. (2006) *Language and Mind*. (3rd edition). Cambridge; Cambridge University Press.
Chung, S. (2005) What Fronts? On the VP Raising Account of Verb-Initial Order' In *Verb-First: Papers from the Tucson Workshop*, ed. Andrew Carnie, Heidi Harley, and Sheila Ann Dooley, 9-29. Benjamins.
Cinque, G. (1994) “On the evidence for partial N-movement in the Romance DP.” In G. Cinque *et al* (eds) *Paths Towards Universal Grammar: Studies in Honor of Richard S. Kayne*, Georgetown: Georgetown University Press, 85-110.
Cinque, G. (1999) *Adverbs and Functional Projections*. Oxford/New York, Oxford University Press.
Cinque, G. (2005) “Deriving Greenberg’s Universal 20 and its exceptions,” *Linguistic Inquiry* 36: 315-332.
Cinque, G. (forthcoming) *The Syntax of Adjectives: A Comparative Study*. Cambridge, Mass.: MIT Press.
Corver, N. & H. van Riemsdijk (1994) *Studies on Scrambling : Movement and Non-Movement Approaches to Free Word-Order Phenomena.* Berlin: Mouton de Gruyter.
Donati, C. (2006) “On *Wh*-Head Movement,” in L. Cheng & N. Corver (eds) *Wh-Movement: Moving On*. Cambridge, Mass.: MIT Press, pp. 21-46.
Embick, D. & R. Noyer (2001) “Movement operations after syntax,” *Linguistic Inquiry* 32:555-596.
Emonds, J. (1971) *Root, Structure-Preserving and Local Transformations.* Phd dissertation, MIT.
Emonds, J. (1976) *A Transformational Approach to English Syntax: Root, Structure-Preserving and Local Transformations*. New York: Academic Press.
Emonds, J. (1978) "The Verbal Complex V-V' in French," *Linguistic Inquiry* 9:151-175.
Grewendorf, G. & W. Sternefeld (1990) *Scrambling and Barriers.* Amsterdam: John Benjamins. Grewendorf & Sternefeld.
Halpern, A. (1992) *Approaching Second*. Stanford: CSLI Publications.
Holmberg, A. (1986) *Word Order and Syntactic Features in Scandinavian Languages and English*. PhD Dissertation, University of Stockholm.
Johnson, K. (2001) “What VP-Ellipsis Can Do, What it Can’t, But Not Why” in M. Baltin & C. Collins (eds) *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, pp. 439-480.
Kayne, R. (1991) “Romance Clitics, Verb Movement and PRO,” *Linguistic Inquiry* 22: 647-686.
Kayne, R. (1994) *The Antisymmetry of Syntax*, Cambridge, Mass.: MIT Press.
Kayne, R. (2000) *Parameters and Universals*, Oxford: Oxford University Press.
Kayne, R. & J.-Y. Pollock (2001) “New Thoughts on Stylistic Inversion,” in A. Hulk & J.-Y. Pollock (eds) *Subject Inversion in Romance and the Theory of Universal Grammar*, Oxford: Oxford University Press, pp. 107-162.
Kiss, E. & H. van Riemsdijk (2004) *Verb Clusters*. Amsterdam: John Benjamins.

Koeneman, O. (2000) *The Fkexible Nature of Verb Movement*. Utrecht, LOT Publications.
Koopman, H. (1984) *The syntax of verb-movement: from verb movement rules in the Kru languages to Universal Grammar*. Dordrecht: Foris.
Koopman, H. & A. Szabolsci (2000) *Verbal complexes*, Cambridge, Mass.: MIT Press.
Landau, I. (2006) "Chain Resolution in Hebrew V(P)-Fronting," *Syntax* 9: 32-66.
Lasnik, H. (1995) "A Note on Pseudogapping," In *MIT Working Papers in Linguistics 27: Papers in Minimalist Syntax*, pp. 143-63. Reprinted with minor corrections in H. Lasnik (1999) *Minimalist Analysis*. Oxford: Blackwell, pp. 151-174.
Lechner, W. (2005) "Interpretive effects of head-movement" ms University of Tübingen (lingBuzz/000178).
Lema, J. & M.-L. Rivero (1990) Long head-movement: ECP vs. HMC. *Proceedings of NELS* 20, 333-47.
Lema, J. & M.-L. Rivero. 1991. "Types of verbal movement in Old Spanish: modals, futures and perfects." *Probus* 3: 237-78.
Longobardi, G. (1994) "Proper names and the theory of N-movement in syntax and logical form," *Linguistic Inquiry* 25: 609-665.
Mahajan, A. (2003) *Syntax at Sunset 3: Head Movement and Syntactic Theory*. UCLA Working Papers in Linguistics, 10.
Marantz, A. (1984) *On the N ature of Grammatical Relations*. Cambridge, Mass.: MIT Press.
Marantz, A. (1988) "Clitics, Morphological Merger and the Mapping to Phonological Structure" in M. Hammond & M. Noonan (eds) *Theoretical Morphology*. New York: Academic Press, pp. 253-270.
Massam, D. (2000) "VSO and VOS: Aspects of Niuean Word Order," Carnie, A. & E. Guilfoyle (2000) *The Syntax of Verb-Initial Languages*. Oxford: Oxford University Press, pp. 97-116.
Massam, D. & C. Smallwood (1997) "Essential Features of Predication in Niuean and English" in K. Kusumoto (ed) *Proceedings of NELS 27*, GLSA, University of Massachusetts, Amherst. pp. 236-272.
Matushansky, O. (2006) "Head movement in linguistic theory," *Linguistic Inquiry* 37:69-110.
McCawley, J. (1971) "Tense and time reference in English," in C. Fillmore & D. Langendoen (eds) *Studies in Linguistic Semantics*. New York: Holt, Rinehart and Winston, pp. 95-113. Reprinted in J. McCawley (1973) *Grammar and Meaning*, pp. 257-272.
McCloskey, J. (1996) "The scope of verb-movement in Irish," *Natural Language and Linguistic Theory* 14: 47-104.
Müller,G. (1998) *Incomplete Category Fronting*. Dordrecht: Kluwer.
Müller, G. (2004) "Verb-Second as vP-First" *Journal of Comparative Germanic Linguistics* 7: 139-274.
Nilsen, Ø. (2003) *Eliminating Positions: Syntax and Semantics of Sentence Modification*. PhD Dissertation, University of Utrecht.
Poletto, C. & J.-Y. Pollock (2004) "On the left periphery of some Romance Wh-Questions," in L. Rizzi (eds) *The Structure of IP and CP; The Cartography of Syntactic Structures, Volume 2*. Oxford: Oxford University Press, pp. 251-296.
Pollock, J.Y. (1989) Verb Movement, Universal Grammar, and the Structure of IP, *Linguistic Inquiry* 20, 365-424.

Pollock, J.-Y. (2006) “Subject Clitics and Complex Inversion,” in M. Everaert & H. van Riemsdijk (eds) *The Blackwell Companion to Syntax, Volume IV*. Oxford: Blackwell, pp. 601-659.
Pollock, J.-Y., Poletto, C. & N. Munaro (2003) “Eppur si muove! On comparing French, Portuguese and Bellunese *Wh*-Movement,” in P. Pica (ed) *Linguistic Variation Yearbook*. Amsterdam: John Benjamins, pp. 147-180.
Rackowski, A. & N. Richards (2005) “Phase Edge and Extraction: A Tagalog Case Study”, *Linguistic Inquiry* 36: 565-599.
Rackowski, A. & L. Travis (2000) “V-initial Languages: X or XP Movement and Adverbial Placement,” in Carnie, A. & E. Guilfoyle (2000) *The Syntax of Verb-Initial Languages*. Oxford: Oxford University Press, pp. 117-142.
Rivero, M.-L. (1991) "Patterns of V-raising in Long Head Movement, and Negation: Serbo-Croatian vs. Slovak." *The Linguistic Review* 8.
Rivero, M.-L. (1993a) “Long Head Movement vs, V2 and Null Subjects in Old Romance,” *Lingua* 89: 113-141.
Rivero, M.-L. (1993b) “Finiteness and Second Position in Long Head Movement Languages,” ms. University of Ottawa.
Rivero, M.-L. (1994a) “Negation, Imperatives and Wackernagel Effects,” *Rivista di Linguistica* 1: 91-118.
Rivero, M.-L. (1994b) “Clause structure and V-movement in the languages of the Balkans”. *Natural Language and Linguistic Theory* 12, 63-120.
Rivero, M.-L. (1997) “On Two Locations for Complement Clitic Pronouns: Serbo-Croatian, Bulgarian and Old Spanish,” in A. van Kemenade & N. Vincent (eds) *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, pp. 170-208.
Rivero, M.-L. & A. Terzi (1995) “Imperatives, V-Movement and Logical Mood,” *Journal of Linguistics* 31: 301-332.
Rizzi, L. (1990) *Relativized Minimality*. Cambridge, Mass.: MIT Press.
Rizzi, L. & I. Roberts (1989) “Complex Inversion in French,” *Probus* 1, 1-30.
Roberts, I. (1991) "Excorporation and Minimality," *Linguistic Inquiry*, 22, 209-218.
Roberts, I. (2001) “Head Movement,” in Mark Baltin and Chris Collins (eds) *Handbook of Syntactic Theory*. Oxford: Blackwells, pp. 113-147.
Roberts, I. (forthcoming) “Head Movement: Clitics, Incorporation and Verb-Movement”, ms. University of Cambridge.
Ross, J. (1967) *Constraints on Variables in Syntax*, MIT PhD Dissertation.
Russell, B. (1905) “On Denoting,” *Mind* 14: 479-493.
Siegel, M. (1984) “Gapping and Interpretation,” *Linguistic Inquiry* 523-530.
Shlonsky, U. (2004) “The form of Semitic Noun Phrases,” *Lingua* 114: 1465-1526.
Sportiche, D. 1999. Subject clitics in French and Romance, complex inversion and clitic doubling. In K. Johnson and I. Roberts (eds.). *Beyond Principles and Parameters: essays in memory of Osvaldo Jaeggli*. Dordrecht: Kluwer, 189-222.
Starke, M. (1993) *En deuxième position en Europe Centrale.* BA Dissertation, University of Geneva.
Surányi, B. (2005) Head movement and reprojection. In: *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio Linguistica.* Tomus XXVI. Budapest: ELTE. 313-342.
Surányi, B. (2007) On Phase Extension and head movement. *Theoretical Linguistics* 33(1): 121-132.

Surányi, B. (2008) Cyclic Spell Out and reprojection in head movement. In: J. Hartmann, V. Hegedűs and H. van Riemsdijk (eds.) *Sounds of Silence: Empty Elements in Syntax and Phonology*. Amsterdam: Elsevier, pp. 293-337.
Thráinsson, H.(2000) "Object Shift and Scrambling," in Mark Baltin and Chris Collins (eds) *Handbook of Syntactic Theory*. Oxford: Blackwell, pp. 148-202.
Toyoshima, T. (2000) *Head-to-Spec Movement and Dynamic Economy*. PhD dissertation, Cornell University.
Travis, L. (1984) *Parameters and Effects of Word Order Variation*. PhD Dissertation, MIT.
Vicente, L. 2006. *The syntax of heads and phrases: a study of verb (phrase) fronting*. Ph.D. dissertation: Leiden.
Vikner, S. (1995) *Verb Movement and Expletive Subjects in the Germanic Languages*, Oxford University Press, Oxford.
Wexler, K. & P. Culicover (1980) *Formal Principles of Language Acquisition*. Cambridge, Mass.: MIT Press.
Wiklund, A. & K. Bentzen (2007) "Rethinking Scandinavian Verb Movement," *Journal of Comparative Germanic Syntax* 10: 203-233.
Wiklund, A., G. Hrafnbjargarson, K. Bentzen & T. Hróarsdóttir (2007) "The Afterglow of Verb Movement," ms. University of Tromsø.
Zwart, J.-W. (2001) "Syntactic and Phonological Verb Movement," *Syntax* 4: 34-62.