OVERVIEW ARTICLE

Approaches to head movement: A critical assessment

Éva Dékány

Research Institute for Linguistics, Hungarian Academy of Sciences, Benczúr utca 33, 1068 Budapest, HU
dekany.eva@nytud.mta.hu

Heads can be spelled out higher than their merge-in position. The operation that the transformationalist generative literature uses to model this is called head movement. Government and Binding posited that the operation in question involves adjunction of a lower head to a higher head in narrow syntax. It has been noted early on, however, that this approach is highly problematic because head-adjunction violates several well-motivated constraints on syntactic structures.

This overview article surveys the problems raised by the adjunction approach as well as the recent alternative analyses that were suggested in its stead in the Minimalist program. It evaluates the strengths and weaknesses of the alternatives, discusses to what extent they are able to eliminate the problems raised by the adjunction analysis, and also points out the new problems that they give rise to.

1 Introduction

1.1 *Head movement and the controversy surrounding it*

Heads can be spelled out higher than their merge-in position. For instance, the exponent of a verb is expected to appear in the verb phrase, but in certain languages (and certain types of clauses) it appears higher, in the inflectional or the complementizer domain. Naturally, this has an effect on the word order of the clause.

Let us consider the specific examples in (1).

(1)   a.   *English*
               John doesn't always **like** Mary.

        b.   *French* (Pollock 1989: 367)
               Jean (n') **aime** pas Marie.
               John not   love   not Mary
               'John doesn't love Mary.'

        c.   *German* (András Bárány, p.c.)
               Gestern   **sah** Hans Maria   nicht.
               yesterday   saw   Hans Maria.ACC   not
               'Yesterday Hans did not see Maria.'

In English, the verb is in the vP because it follows negation and adverbs, the sign-post elements that mark the left edge of the verb phrase. In French and German, however, the exponents of the verbs appear outside of the vP. In French the verb follows the subject but precedes negation (1b), so it is in a position above NegP but still in the IP-domain. In German non-subject-initial root clauses, on the other hand, the verb is obligatorily in the

second position; it precedes the canonical positions of the subject, object, and negation (1c), and so it is in the CP-domain.[1]

Pre-theoretically, we can use the term head movement (HM) as the name of the operation that the transformationalist generative literature uses to model the word order difference between (1a), (1b) and (1c). There is currently great controversy over what grammatical mechanism this operation exactly corresponds to: whether it is movement or not, and if it is movement, what exactly moves (only a head or a whole phrase), where it moves, and in which component of the grammar. While both ingredients of the label "head movement" are under debate, it will serve as a useful descriptive term for the operation involved in the word order difference between (1a) to (1c) thoughout this paper.

The proposed alternatives of what HM exactly involves fall into four main groups.

The operation of HM is:
  • syntactic movement
  • a combination of syntactic movement and a post-syntactic operation
  • post-syntactic movement
  • post-syntactic and involves no movement; it falls out from the way the syntactic hierarchy is translated into linear oder

The aim of this paper is to offer a balanced discussion of these alternatives and to evalue their strengths and weaknesses.

1.2 Constraints on HM

While the exponent of a head can occur higher than the merge-in position of that head, not all logically possible types of patterns are attested. Researchers have come to the conclusion early on that the operation that is used to model the data must be subject to three constraints, or else overgeneration cannot be avoided.

The first constraint applies to morphologically complex heads. The HM operation always displaces the exponent of the head from the merge-position of the head. In some cases the displaced exponent remains morphologically unaltered. We can observe this in English matrix yes-no questions, where the preposing of the auxiliary from T to C does not change the form of the auxiliary.

(2)   a.   The cat will eat the mouse.
       b.   Will the cat eat the mouse?

In other cases, the change of the base-generated word order is also accompanied by word formation: affixation of a stem with derivational and/or inflectional suffixes (3) or incorporation of a noun into a(n inflected) verb (4b).

---

[1] We can observe similar data in the domain of noun phrases as well. In English the noun appears to the right of all NP-modifiers (except complements), therefore it is standard to assume that it is in situ. In the Hebrew Construct state, on the other hand, the noun sits in the left periphery of the DP. It is the first element in the DP, and it precedes the adjective that modifies it as well as its possessor (ii), a sign that it (or a projection containing it) has undergone movement.

(i)   *English*
      the Italian **invasion** of Albania

(ii) *Hebrew* (Ritter 1988: 916)
      beyt   ha-mora   ha-yafe
      house-M the-teacher-F the-pretty-M
      'the teacher's pretty house'

(3)   *Dutch* (Marcel den Dikken, p.c.)
      Ze   **speel-de**-n   op straat.
      they play-PST-3PL on street
      'They were playing in the street.'

(4)   *Mohawk* (Baker 1988: 81–82)
      a.   Yao-wir-aʔa   ye-nuhweʔ-s   ne ka-**nuhs**-aʔ.
           PRE-baby-SUF 3FS/3N-like-ASP the PRE-house-SUF
           'The baby likes the house.'

      b.   Yao-wir-aʔa   ye-**nuhs-nuhweʔ**-s.
           PRE-baby-SUF 3FS/3N-house-like-ASP
           'The baby likes the house.'

As first discussed in Baker (1985), the internal structure of complex words created by HM reflects the underlying syntactic structure of the given expression.

(5)   The Mirror Principle
      Morphological derivations must directly reflect syntactic derivations (and vice versa). (Baker 1985: 375)

What the Mirror Principle says is that an affix that spells out a lower head will end up closer to the root than an affix that spells out a higher head.[2] This way the morphological make-up of words allows insights into the syntactic hierarchy of functional projections. For instance, on the basis of the morphology of the Hungarian verb in (6a) we can conclude that the syntactic hierarchy of the causative, modal, and tense heads is as in (6b).

(6)   *Hungarian*
      a.   Ír-at-tat-hat-t-ak   gyógyszer-t.
           write-CAUS-CAUS-POT-PST-3PL medicine-ACC
           'They may have made somebody have medication prescribed.'

      b.   tense > modality > outer causative > inner causative > V

The first constraint on the HM operation is that it cannot break up the morphologically complex heads that it creates at a later point in the derivation: there is no excorporation from complex heads. This restriction is a subcase of the Lexical Integrity Hypothesis.

(7)   Lexical Integrity Hypothesis
      No syntactic rule can refer to elements of morphological structure. (Lapointe 1980: 8)

The no excorporation condition amounts to saying that head movement is always "roll-up" movement and there is no successic cyclic head movement; chains of head movement are maximally two-member chains.

   The second constraint on the operation is that it is strictly local: it always establishes a relation between two structurally adjacent heads. This has been formulated as the Head Movement Constraint (Travis 1984).[3]

---

[2] As noted in Brody (2000a), the name in (5) involves the word "Principle". It is, however, not a syntax-internal genuine principle like the Projection Principle or the Empty Category Principle. Instead, it is an empirical generalization, and so the name "Mirror Generalization" would be a better fit for it.

[3] Chomsky (1986); Baker (1988) and Rizzi (1990) proposed that the HMC can be derived from the Empty Category Principle (ECP).

   (i)   The Empty Category Principle
         A nonpronominal empty category must be properly governed. (Lasnik & Saito 1984: 240)

(8)   Head Movement Constraint
         An X^0 may only move into the Y^0 which properly governs it. (Travis 1984: 131)

In cases in which a head apparently moves to a structurally non-adjacent higher head position, e.g. when V ends up in C, we have the successive creation of separate local chains: V-to-T and then T-to-C. V ends up in C because it is pied-piped by T when T moves to C.

   Thirdly, the HM operation is clause-bound, or more generally, it applies only within but not across extended projections.[4]

   Importantly, these constraints do not apply to phrasal movement. Phrasal movement can skip intermediate phrasal positions; it is not the case that it has to target the next higher specifier. Phrasal movement can also be successive cyclic: phrases can touch down in intermediate positions and move further on without pied-piping any other material with them. Finally, phrasal movement can cross clausal boundaries giving rise to so-called long movement (long *wh*-movement, long topicalization, long focalization, etc.).

1.3 The structure of the paper

This paper is organized as follows. Section 2 discusses the GB-style syntactic adjunction analysis proposed for (1b) through (4b), as well as the theory-internal arguments that were leveled against this analysis. Approaches that maintain that data like (1b)–(4b) should be captured with a syntactic operation will be surveyed in Section 3. A model in which such data arise as a result of a syntactic movement followed by a post-syntactic operation will be discussed in Section 4. Theories that account for these data with a post-syntactic displacement operation will be the topic of Section 5. The analysis of (1b) to (4b) as positioning via syntax-phonology mapping is taken up in Section 6. I will discuss to what extent these theories can eliminate the problems posed by the adjunction analysis as well as the new problems they give rise to.[5] In the current Minimalist framework, the most important question is whether (1b) through (4b) should be modeled by a narrow syntactic operation or not. Section 7 addresses arguments related to this issue. Section 8 concludes the paper.

2 The head-to-head adjunction analysis

In Section 1.2 three types of data were discussed: i) upward displacement of a head's exponent without morphological growth of that head, ii) upward displacement of a head's exponent accompanied by morphological growth of the head (i.e. displacement + affixation), and iii) incorporation. In the GB period all three types of data were modeled with the same syntactic operation, whereby a lower head moves up to and adjoins to a
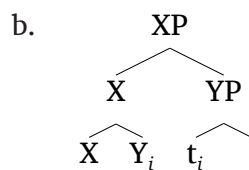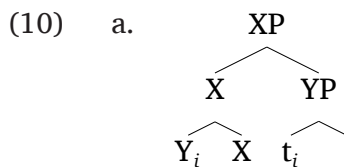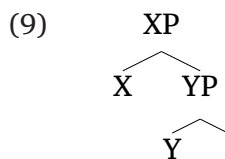

   Roberts (2001) suggests that the Head Movement Constraint is instead a special case of Relativized Minimality (RM). As Vicente (2007) points out, however, this would require some feature that is shared by all heads (as RM makes reference to features, not positions).

[4] From the early days of GB, however, the validity of all three constraints has been called into question: there has been discussion about cases of "long head movement" and excorporation (see Sections 3.2.2 and 3.4.2) as well as non-clause bounded HM (see Landau 2006; Vicente 2007; Harizanov 2016 and Harizanov & Gribanova accepted).

[5] The focus of this paper is the head adjunction operation posited in GB, the problems with this operation and the the various alternative operations that were meant to replace head-adjunction. There are cases in which the exponent of a head appears lower than its merge-in position. Such data are customarily modeled with Lowering or Affix Hopping. Head-adjunction and its alternatives on the one hand and Lowering/Affix Hopping on the other hand model a complementary set of data. As Lowering is not meant to replace head adjunction, it will not be discussed here. I refer the interested reader to Halle & Marantz (1994); Bobaljik (1995); Embick & Noyer (2001); Embick & Marantz (2008) and Skinner (2009) for relevant detailed discussion.

higher head (Koopman 1984; Travis 1984; Baker 1985).[6] The output of this adjunction is a complex head, as in (10).

(9)   XP
        /\
       X   YP
           /\
          Y

(10)   a.   XP
                 /\
                X   YP
               /\   /\
             Y_i X t_i

       b.   XP
                 /\
                X   YP
               /\   /\
              X Y_i t_i

Data of type i) arise when the host of adjunction (X in (10)) has a zero exponent, while types ii) and iii) result from head-adjunction to a host that has an overt exponent.

That after adjunction neither the moved head nor the target can move out (no excorporation) does not follow from the structure itself; this must be taken care of by a separate constraint (see Baker 1988 for early discussion, who suggests that words cannot contain traces).

There are several well-known problems with the head-adjunction approach, however. Firstly, it does not obey the Extension Condition of Chomsky (1993: 22–23).[7]

(11)   The Extension Condition
        GT and Move α extend K to K*, which includes K as a proper part. (Chomsky 1993: 22)[8]
        Substitution operations always extend their target. (Chomsky 1993: 23)

In other words, while Merge and phrasal movement are cyclic operations, head-adjunction is not (it does not extend the tree at the root).

Secondly, (10) complicates the definition of c-command. The Proper Binding Condition requires traces to be bound, that is, a moved constituent must c-command its extraction site from the landing site.

---

[6] There are also cases in which morphological growth (affixation) of a head occus without displacement (see Brody 2000a; Adger et al. 2010; Harley 2013; Harizanov & Gribanova accepted). It has always been acknowledged that the head-adjunction analysis cannot account for this; such data were assumed to involve a different operation such as Affix Hopping or Lowering.

[7] Note, however, that in Chomsky (1993) and Chomsky (1995: Chapter 3) the Extension Condition does not apply to adjunction, which is the mechanism of HM in (10). In order to allow for HM, Chomsky (2000) replaces the Extension Condition with the Least Tampering Condition.

(i)   The Least Tampering Condition
      Given a choice of operations applying to A and projecting its label L, select one that preserves R(L, g). (Chomsky 2000: 137)

This, however, raises more problems than it solves. See Surányi (2005) for discussion.

[8] GT stands for Generalized Transformation.

(12)   Proper Binding Condition
         In surface structure S_α, if $[e]_{NPn}$ is not properly bound by $[\ldots]_{NPn}$, then S_α is not
         grammatical. (Fiengo 1977: 45)

Applied to (10), this means that Y must c-command its trace, which, in turn, means that
c-command must be defined in such a way that the moved head c-commands out of the
complex head it is part of. Baker's (1988) definition of c-command, for instance, is given
in (13).

(13)   Baker's revised definition of c-command
         A C-COMMANDS B iff A does not dominate B and for every maximal projection
         C, if C dominates A then C dominates B. (Baker 1988: 36, original emphasis)

(13) effectively replaces c-command with m-command as the crucial relationship holding
between moved elements and traces.[9]
  Thirdly, (10) also violates the Chain Uniformity Condition.

(14)   Chain Uniformity Condition
         A chain is uniform with regard to phrase structure status. (Chomsky 1995: 253)

In Bare Phrase Structure (BPS) heads and phrases are defined relationally: heads are cat-
egories that are not projected, while phrases are categories that do not project. On this
definition, in (10) the lower copy of the moving Y is a head, while the higher copy is both
a head and a phrase. Therefore the movement in (10) produces a non-uniform chain, in
violation of (14). Furthermore, in BPS the higher X in (10) (dominating X and the moved
Y) is neither a head nor a phrase. As intermediate categories are generally thought to be
inert, it is predicted that the complex head X will not be able to undergo movement to the
next higher head. This is undesirable, as "roll-up" HM does occur.
  Fourthly, (10) violates the A-over-A Principle on movement. The A-over-A Principle is a
sort of minimality condition: it states that if a category A contains another category A (i.e.
[A … [A …]]), then it is not possible to extract the lower category A across the higher
category A containing it. If movement of category A is required, then it is the higher one
that needs to move. For instance, if a DP embeds another DP, then it is not possible to
move the lower DP out of and across the higher DP.

(15)   A-over-A Principle
         If a transformation applies to a structure of the form $[\alpha \ldots [_A \ldots]_A \ldots]\alpha$, where
         α is a cyclic node, then it must be so interpreted as to apply to the maximal
         phrase of the type A. (Chomsky 1973: 235)

In BPS there are no category labels like X-bar and XP. Instead, intermediate and maximal
categories inherit the category label of their head. This means that a head and its maximal
projection bear the same label, which gives rise to an [A … [A …]] configuration. In
(10) the head raises out of its own maximal category. With the lower Y moving across the
higher, containing Y(P), an A-over-A violation is incurred. (See Section 7.2 for a proposal

---

[9] Baker (1988: 449, fn. 10) also discusses an alternative, however, which he attributes to personal
  communication with Chomsky. According to this version, in (10) X does not contain Y because it has a
  segment that does not contain Y. The smallest category that properly dominates Y is XP. Since XP contains
  the trace of Y, there is c-command between the head and the tail of the chain. While this is more restrictive
  than m-command, it still requires a complex definition of c-command that makes reference to segments.
  Note, however, that a complicated definition of c-command may be needed independently of head-to-head
  adjunction, too, cf. Kayne's (1994: Chapter 3) discussion of adjunction (and specifiers) in general.

by Preminger why this violation happens, and how it can be used to argue that HM takes place in narrow sytnax).

Fifthly, (10) violates anti-locality. Anti-locality bans movements that are too local/too short (cf. Grohman 2001; 2002; 2003a; b; 2011 and Abels 2003, among others).

(16)   Anti-Locality Hypothesis
         Movement must not be too local. (Grohman 2003b: 269)

Abels (2003: Chapter 2.4) proposes that all movements must lead to feature satisfaction that was impossible before the movement. Certain local movements are such that they cannot lead to new feature satisfaction by definition. Movement from the complement of a head to the specifier of the same head is a case in point: any feature that can be satisfied between a head and its specifier can also be satisfied between the head and its complement, so anti-locality rules out this type of movement. Head-adjunction also runs afoul of anti-locality. In BPS all the featues of the head are assumed to be present on the phrase projected by the head. Feature satisfaction between a head X and the next lower head Y thus can take place immediately upon merger of X with YP, and adjunction of Y to X does not allow feature satisfaction that was impossible before the movement.[10,11] It should be pointed out, however, that the anti-locality constraint on movements has not been unanimously adopted in the literature, and so this is not necessarily a strong argument against head-adjunction.

Sixthly, at least in some (but not necessarily all) cases (10) needs a special triggering feature that is different from the feature triggering phrasal movement to specifiers. If this were not the case, there would be no cases in which movement to both the head and the specifier of a projection are simultaneously necessary. But such cases do exist: for instance in English matrix *wh-* questions the specifier of C is filled by a *wh-* element, while C is filled by T-to-C head movement.

Seventhly, in the checking theory of movement (10) also complicates the definition of checking domain (Surányi 2005). "Checking domain" must have a disjunctive defintion because we must allow features of heads to be checked either by a specifier (spec-head agreement) or by an adjoining head. Disjunctive definitions are always suspect, however, of missing an important generalization.[12]

Finally, the HM operation does not seem to affect semantic interpretation in a consistent/systematic way. Here the term **consistent/systematic** is of key importance. For instance, while A-movement may affect interpretation, e.g. by altering scope relations between constituents, not every instance of A-movement does so (the movement of the subject to Spec, TP, for instance, does not appear to have an effect on the interpretation of

---

[10] A reviewer points out that anti-locality does not rule out all cases of HM, though. If a lower head bears an uninterpretable feature that should be valued by the next higher head, and upward Agree is excluded, then the maximally local configuration of the heads in their base-generated position is not enough for feature valuation to happen.

[11] Grohman (2003a) has a different view on what constitutes anti-local movement. He suggests that any movement that takes place within a so-called Prolific Domain is too short, where the Profilic Domains are the thematic, agreement and discourse domains. If coupled with the HMC, this view excludes all cases of head movement except when the highest head of a domain moves to the lowest head of the next domain. In the following sections we will apply Abels' definition of anti-locality.

[12] In more recent Minimalism the checking theory (understood as a relation between two valued features) has been replaced by valuation (of an unvalued feature by a valued feature) under Agree. While checking required a local relationship, Agree can also take place long-distance (provided there is no intervener or phase head between the valuing and the to-be-valued feature). Given this shift in the theory, I will not discuss how the different alternatives to head-adjunction can solve the problem with the disjunctive definition of checking.

the sentence). We would expect any core syntactic operation to have the ability to affect interpretation, and so it is unexpected that HM systematically fails to do so.[13]

In spite of these problems with (10), however, the head-adjunction analysis is not universally rejected. Pesetsky (2013: Chapter 4) defends (10) on theoretical grounds. He notes that the root of all problems with head movement is that it is a **complement-creating** rule: the moving element lands in a complement position (in (10), the moved Y becomes the complement of X). He labels complement-creating movement Undermerge, and points out that this type of movement has also been proposed in the realm of phrasal movement. Sportiche (2005), for instance, argues that D is merged not within the extended noun phrase, but among clausal functional projections, and NPs combine with D by moving up to the D head and becoming a complement of D. Raising to Object is another instance of complement-forming phrasal movement: here the subject of an embedded clause raises to the complement (object) position of the matrix verb (Rosenbaum 1967; Postal 1974 and later work).[14] Finally, McCloskey (1984) argues that modern Irish features complement-creating phrasal movement to the P head (see also Postal 2004: Chapter 2 on English *rely on*). In view of these proposals for "head movement-like phrasal movement", Pesetsky fully embraces head-adjunction. Baker (2009) also argues that (10) is needed: he suggests that this is the best model of noun incorporation in Mohawk and Mapudungun.[15,16]

A large body of literature, however, considers HM qua adjunction as a highly problematic operation, and seeks other alternatives to model the data in (1b) through (4b). Many researchers are exploring narrow syntactic alternatives, in which some version of HM is part of the core syntactic module of grammar. We will turn to these theories in the next section.

3 HM as a syntactic operation

In this section we look at approaches that consider HM to be part of narrow syntax. We will start with theories in which the final output of HM is an adjunction structure very much like in HM qua adjunction: sideward movement (Section 3.1) and Agree with a defective goal (Section 3.2). Then we turn to the reprojective movement analysis (Section 3.3). Some proponents of this theory hold that complex words are composed via head adjunction, but the core of the analysis can be maintained without this assumption, too. Finally, we look at the phrasal movement analysis, which has a different output from head-adjunction (Section 3.4).

---

[13] But see Mahajan (2003) and Matushansky (2006) for proposals why verb movement lacks interpretive effects, and Section (7) for proposals that HM can have an effect on interpretation.

[14] See, however, Lasnik & Saito (1991) and Runner (1998) for early analyses in which the embedded subject lands in a specifier position in the matrix clause (Spec, AgrOP; recast as the outer specifier of vP in later minimalist work).

[15] In Mapudungun the incorporated noun follows the verb rather than precedes it, so deriving these data by the mechanism of (10) requires right-adjunction. That is, one cannot maintain both the head-adjunction analysis of (ib) and Kayne's (1994) LCA (as the latter exludes right-adjunction).

  (i)   *Mapudungun* (Baker 2009: 149)
        a. Ñi chao kintu-le-y   ta-chi   pu   waka.
            my father seek-PROG-IND.3SS the-ADJ COLL COW
            'My father is looking for the cows.'

        b. Ñi chao **kintu-waka**-le-y.
            my father seek-cow-PROG-IND.3SS
            'My father is looking for the cows.'

[16] Chomsky, too, holds that while HM generally takes place post-syntactically, incorporation involves syntactic head movement: "There are some reasons to suspect that a substantial core of head-raising processes, excluding incorporation in the sense of Baker (1988), may fall within the phonological component" (Chomsky 2001: 37).
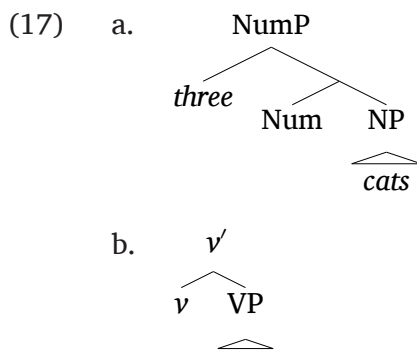
3.1 Sideward movement

The first syntactic alternative to head-adjunction to be discussed here is sideward movement, i.e. movement of heads between different derivational spaces. This approach is pursued in Nunes (1995; 2001; 2004); Bobaljik & Brown (1997) and Uriagereka (1998).

3.1.1 The mechanics

Mainstream generative grammar holds that structure building proceeds in a bottom-up fashion.[17] A consequence of this approach is that whenever a head merges with an internally complex specifier (or a phrase merges with an internally complex adjunct), syntax must make use of two different workspaces (aka. derivational spaces) in parallel. Consider the case in which v takes a subject NP/DP in which the noun has a modifier, e.g. *three cats*. Before merger of the subject and v', we have the two syntactic objects in (17).

(17)   a.   NumP
                 /   \
            three   / \
                    Num   NP
                          /\
                        cats

       b. v'
                    /\
                  v VP
                       /\

Crucially, (17a) is internally complex and must have been built independently of (17b). So in order for the derivation to reach the stage with the two objects in (17), syntax must have used two parallel workspaces: one to construct (17a), and another, independent one to create (17b).[18]

   Nunes (1995; 2001; 2004); Bobaljik & Brown (1997) and Uriagereka (1998) suggest that movement can take place between two parallel workspaces. They call this type of movement sideward movement, interarboreal movement or paracyclic movement, and suggest that HM also proceeds in this fashion.

   The standard approach holds that a head Y can move to the next higher head, X, only after X has merged with its phrasal complement YP (see Section 2). The sideward movement analysis abandons this assumption and suggests that the order of operations is exactly the other way around. Consider the case of v-to-T movement, for instance. (The internal structure of v resulting from V-to-v is ignored for expository purposes). Once the vP is constructed, the head T is placed into a workspace separate from vP (18).

(18)   a.   workspace 1
                 vP
                 /\
              v VP

       b.   workspace 2
                 T

---
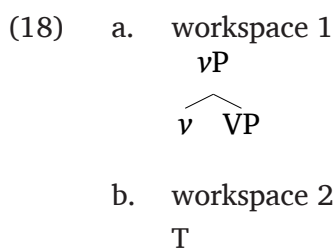
[17] See Phillips (1996; 2003); Chesi (2004; 2015) and Den Dikken (2018) for some exceptions that advocate top-down structure building.

[18] Of course, at the point when (17a) is merged with (17b) and vP is created, the two objects become part of a single workspace.

In the sideward movement approach, the next step is that v moves out of workspace 1 into workspace 2 and adjoins to T. This creates a complex head in workspace 2. At this stage, the two instances of v are not in a c-command relationship.

(19)　　a.　workspace 1

vP

v　VP

　　　b.　workspace 2

T

v　T

Next, the trees in (19a) and (19b) are merged with each other. At this stage, we have a structure that contains two non-distinct instances of v, and (on Kayne's definition of c-command) these instances are in a c-command configuration. Consequently, they are interpreted as forming a chain and Chain Reduction silences the lower copy at PF.

(20)　　　　TP

　　　T　　vP

　v　T　v　VP

The final output of the sideward movement approach is the same as the output of the head-adjunction analysis discussed in Section 2.

3.1.2 The pros and cons of this approach

The sideward movement approach is compliant with the Extension Condition. In (19), the movement of v to T extends the root in workspace 2, and the merger depicted in (20) also extends the root of vP. In this analysis the HM operation is fully cyclic. No problem arises with anti-locality either: the movement brings two heads into the same workspace, allowing feature satisfaction between them that was not possible when they were in different workspaces.[19]

　At the same time, the approach retains many of the problems of the head-adjunction analysis. The sideward movement in (19) still violates the formulation of the A-over-A Principle in (15).[20] The structure still requires a complication in the definition of c-command (the v adjoined to T must be assumed to c-command out of T), the assumed movement operation still needs a trigger that is different from phrasal movement, and the problem with the Chain Uniformity Condition is not solved either. A new problem that arises in this approach is how to keep the theory constrained enough to admit only attested cases of movement.

3.2 Agree with a defective goal

In an analysis developed in Roberts (2010) and taken up in Livitz (2011); Aelbrecht & Den Dikken (2013); Walkden (2014) and Iorio (2015), among others, head-adjunction is replaced by Agree between a probing head and a head that serves as its defective goal.

---

[19] The theory has to ensure that HM obeys the head movement constraint, however. Bobaljik & Brown (1997) argue that while locality does not apply to the movement itself, the notions "shortest" or "closest" apply to the links of the chain in (20).

[20] If the A-over-A Principle is viewed as a kind of minimality condition, and holds because AP is always closer to higher probes than A, then this approach has no problem with the A-over-A Principle, though: in (19) vP is not closer to T than v.

3.2.1 The mechanics

In this analysis HM happens when a probe and a goal enter into a syntactic Agree relationship, and the goal's formal features are a proper subset of the probe's. A goal in such a relationship is called a defective goal.[21]

(21)   Defective goal
         A goal G is defective iff G's formal features are a proper subset of those of G's
         Probe P. (Roberts 2010: 62)

After Agree, all of the features of the defective goal are also present on the probe, and the goal incorporates into the probe. As a result of this mechanism, the goal's features are pronounced at the probe.

Let us consider the case of v-to-T as a specific example.[22] This movement happens when T has an interpretable Tense-feature and an uninterpretable V-feature (as well as ϕ-features to be valued by the subject), while v has an uninterpretable T-feature and an interpretable V-feature. In the trees below, the ϕ-features and the internal structure of v (after V-to-v movement) are ignored for simplicity of exposition.

(22)   Environment for Agree

$$T^{min} \quad v^{min}$$
$$[iT, uV] \quad [iV, uT]$$
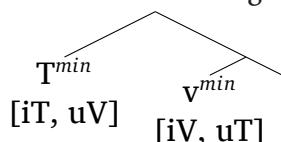
The Agree relationship beween T and v exhausts the goal's features: after Agree the label of T contains valued versions of v's features. Crucially, originally unvalued features are assumed not to undergo deletion at the transfer of the phase. So as a result of Agree, the same set of features will be present both in v's label and within T's label at the final stage of the derivation. As T c-commands v, the two sets of identical V and T features will be interpreted to form a chain. This means that the output of Agree with a defective goal is formally indistinguishable from the output of Move. The resulting configuration allows v to adjoin to T in an incorporation operation. The result is a derived minimal head, a T^{min} rather than a T^0.

(23) v-to-T: valuation, incorporation
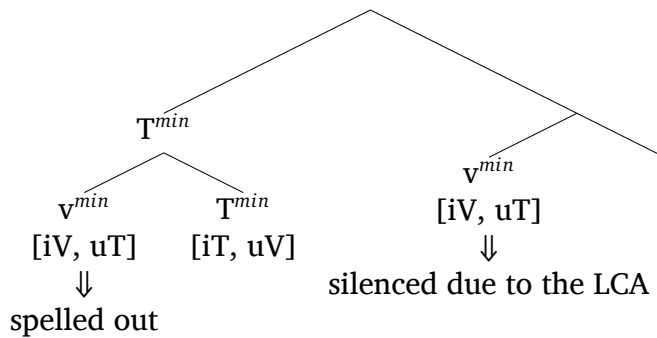
$$T^{min} \quad v^{min}$$
$$v^{min} \quad T^{min} \quad [iV, uT]$$
$$[iV, uT] \quad [iT, uV]$$
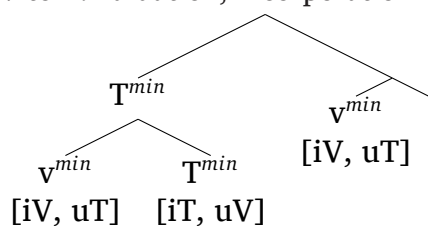
At this point the iV and uT features are present at two places, and the higher instances (in T^{min}) c-command the lower instances (in v). They are therefore subject to regular chain reduction when the structure is linearized. As usual, it is the head of the chain that receives phonetic form and the tail remains silent. In other words, we have "the PF effect of movement" (Roberts 2010: 61).

---

[21] By (21), defectivity is a relative rather than an inherent notion. It does not mean that the goal is somehow defective in its feature content; it simply means that the goal has no formal features that its probe does not also have. This also means that a goal can be defective with respect to one probe but non-defective with respect to another. For this reason, Aelbrecht & Den Dikken (2013) use the term "subset goal" to refer to such goals.

[22] It should be emphasized that in the end, Roberts adopts a partial reprojection analysis for Romance and Celtic V-to-T, and a remnant movement approach for Norwegian verb movement (see Chapter 4 for details). The sample derivations here feature v-to-T for better comparability with the other proposals.

(24)   Chain reduction

$$T^{min}$$
$$v^{min} \quad T^{min} \qquad v^{min}$$
$$[iV, uT] \quad [iT, uV] \qquad [iV, uT]$$
$$\Downarrow \qquad\qquad\qquad \Downarrow$$
$$\text{spelled out} \qquad \text{silenced due to the LCA}$$

3.2.2 The pros and cons of this approach

This approach solves several problems raised by the head-adjunction analysis. There is no need for a specific movement-triggering feature; the trigger is the unvalued features that trigger any Agree relationship. The definition of c-command, Roberts argues, does not need to be complicated, because the goal incorporates into the probe, and the probe c-commands the base-position of the goal.

This approach involves incorporation, and incorporation is restricted to heads. This means that the phrase projected by the defective goal is not a possible target for movement to begin with, and so (23) does not violate the A-over-A Principle.

The incorporation of the probe into the goal violates the Chain Uniformity Condition: the lower copy of the incorporee is a head but its higher copy is both a head and a phrase. Roberts suggests, however, that this condition may have to be abandoned independently of head movement, and that it is also possible that the notion of chain is unnecessary in general.

In this analysis the output of Agree is an X^{min} rather than an X^0; the lower head is not added extraneously to the target, but becomes part of the higher head. As a result, Roberts argues, the higher head is not extended, and so the Extension Condition is not violated. We will see below, however, that in addition to Agree, this approach also involves ordinary movement of the goal to the probe, and this movement does not extend the root of the tree, so the problem with the Extension Condition is not resolved. (To be fair, however, Roberts argues that the Extension Condition is not even relevant here: this condition is only forced by edge features, which are not involved in Agree with a defective goal).

The Agree with a defective goal approach does not derive the HMC. Agree can take place between structurally non-adjacent heads, and if the goal is defective, its features will end up being pronounced on the goal. In other words, this approach predicts that cases of long HM will occur. Roberts argues that this is desirable because such cases indeed exist, e.g. in the case of English Quotative Inversion, Breton long verb movement or Mainland Scandinavian V2 (see Chapter 5). In the latter case, for instance, V ends up in C apparently without stopping in T (in embedded clauses V stays in the vP and in main clauses we do not have direct evidence for an intermediate position in T). That HM often has a local character is derived from the Phase Impenetrability Condition and the locality conditions on Agree rather than a condition specific to heads. The proposal even entertains the possibility that if the functional hierarchy is fine-grained enough, HM never targets the next head up, and so it does not ever violate the anti-locality constraint on movement. At the same time, in the purported cases of long HM it is difficult to construct empirical arguments regarding the presence or absence of an intermediate copy, and cases of long HM can always be recast as cases of remnant vP/VP movement.

This approach does not derive the no excorporation condition either. Suppose that a defective goal incorporates into its probe, and then that same goal is probed by a higher head, such that the goal is defective with respect to that higher probe as well. In this case the goal's features will end up pronounced on the higher probe, which yields the effect of successive cyclic movement. The system therefore predicts that excorporation is possible and relevant cases should be attested. Roberts argues that excorporation of the incorporee (the "moved" head) indeed exists (e.g. in the case of clitic climbing).[23] However, the cases in which this is suggested to apply, namely cliticisation, are extremely contentious cases for excorporation, as it is not clear that incorporation is involved in the first place. Indisputable cases of incorporation (e.g. those discussed in Baker 1988) do not allow excorporation, and so the fact that excorporation is allowed is not advantageous.
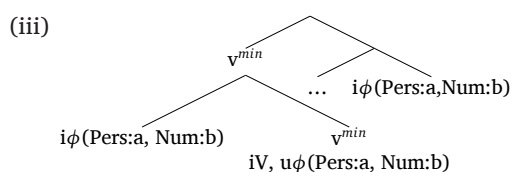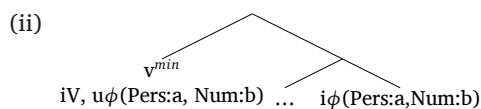
The analysis leaves doubt about whether it involves syntactic movement or not. On the one hand, it is suggested that it does not. In the relevant cases "Agree and Move/Internal Merge are formally indistinguishable" (Roberts 2010: 60) and "given that copying the features of the defective goal exhausts the feature content of the goal, Agree/Match is in effect indistinguishable from movement. For this reason we see the PF effect of movement" (Roberts 2010: 160). On the other hand, the fact that Agree is followed by incorporation suggests that some form of movement is involved, after all (see also Matushansky 2011). In (23), for instance, we see that v has adjoined to T.[24] Incorporation restricts the operation involved in the analysis to heads. In principle, it should be possible for a phrase to be a featural subset of a higher probe, in which case Agree with a defective goal followed by chain formation and chain reduction (as detailed above) should yield the PF effect of XP-movement. But if Agree with a defective goal is obligatorily followed by incorporation, then this would involve adjunction of a phrase to a head; an illicit configuration. The movement step following Agree is thus needed, but this has to be stipulated because it does not follow from the mechanism of Agree.

---

[23] This approach does not allow excorporation of the original incorporation host. See Roberts (2010: Chapter 5.2) for a detailed explanation.

[24] In this specific case one might argue that the movement is motivated by the unvalued T feature on v, which must end up c-commanding T in order to get valued. The analysis of object clitic movement to v shows that incorporation also happens when the goal has no uninterpretable features. In this case v has unvalued ϕ-features and the clitic only has (valued) ϕ-features (i). (The other features of v and its internal structure after V-to-v are again ignored for expository purposes).

(i)

v^{min}
iV, uϕ(Pers:_,Num:_) ... iϕ(Pers:a,Num:b)
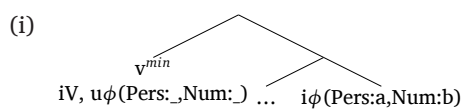
In the proposed analysis after Agree we do not simply get (ii) but (iii), with the interpretable ϕ-features as the highest adjunct to v.

(ii)

v^{min}
iV, uϕ(Pers:a, Num:b) ... iϕ(Pers:a,Num:b)

(iii)

v^{min}
... iϕ(Pers:a,Num:b)
iϕ(Pers:a, Num:b) v^{min}
iV, uϕ(Pers:a, Num:b)

Here this extra step is required to get the word order: on the surface, the clitic object appears to the left of the verb, and it would be difficult to achieve this if the clitic were the spellout of the uninterpretable features inherent in the label of v.

It should also be mentioned that Agree with a defective goal is not a global alternative to the head-adjunction analysis. It is suggested to co-exist with three other mechanisms that deliver upward movement of heads: reprojective movement of a compound head formed in the Numeration, A′-movement and *wh*-movement (see Chapter 5.3).[25]

3.3 Reprojective movement

In another recent analysis of data like (1b) through (4b), a head projects a phrase, moves up and adjoins to that phrase, and then projects another phrase with a different label. This approach is advocated, among others, in Koeneman (2000); Bury (2003); Fanselow (2004); Surányi (2005; 2008) and partly also in Biberauer & Roberts (2010) and Roberts (2010) for the verbal domain, in Donati (2006) for *wh*-movement in free relatives, and in Georgi & Müller (2010) for the nominal domain.

3.3.1 The mechanics

In the head-adjunction analysis complex heads arise as a result of syntactic movements. For instance, the verb in inserted in V, the past tense suffix is inserted in T, and they form a complex head only after movement. In the reprojection approach this is not the case: complex heads are merged into the structure already in their complex form. For a verb form like *kisses*, for instance, this means that what merges in the V position is not the verbal stem *kiss*, but the whole inflected verbal form *kisses*.[26] In the complex head, the affixes on the root have features that must be checked, but this will only be possible in a higher structural position. Therefore the complex head moves out of the phrase that contains it and merges into the structure again as the sister of that phrase. The moved head then projects the label of the newly formed syntactic object (hence the name reprojection).

Let us consider a specific example. In the hypothetical language *English′* with V-to-T movement, the complex verb form *kisses* that merges in the V position has three features: V, *v*, and T (25). In the VP, this complex head merges with the object. This satisfies V's requirement for an object complement, the V feature discharges its object theta-role, and the complex head projects the label V (because this feature's requirements have now been satisfied and it will be inactive in the rest of the derivation).
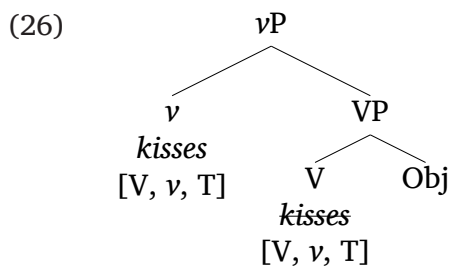
(25)

```
            VP
           /  \
         V     Obj
       kisses
      [V, v, T]
```
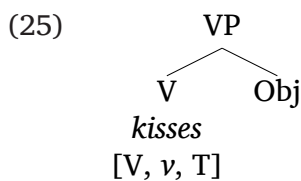
The *v* and T features on *kisses* remain active, however, as they have their own selectional requirements (*v* wants a V complement, T wants a *v* complement) that have not yet been satisfied. In the next step of the derivation *kisses* moves out of VP and is merged as a sister to it, as in (26). After the movement the selectional feature of *v* is satisfied, and so the moved head projects this label.[27]

---

[25] Interestingly, the first of these is still argued to be a form of incorporation, but it is dubious that it could be viewed as a form of Agree with a defective goal.
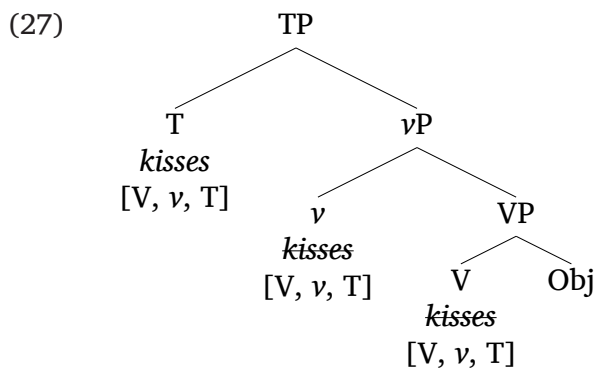
[26] Opinions differ as to how this morphological complex comes about. Koeneman (2000) and Fanselow (2004) take the lexicalist position of Chomsky (1995) that this complex comes from the lexicon, Biberauer & Roberts (2010) and Roberts (2010) propose that complex heads are assembled in the Numeration, while Surányi (2005; 2008) argue that complex heads are put together in syntax. We will abstract away from these differences here and represent complex heads without any internal syntactic structure.

[27] Fanselow (2004) and Surányi (2005; 2008) argue that the features that drive the movement are c-selectional features (a position also shared by Matushansky 2006), while Koeneman (2000) proposes that the trigger is that ultimately all of the categorial features of the complex head need to project. The two approaches are not incompatible with each other.

(26)

```
                    vP
              ┌──────┴──────┐
              v             VP
           kisses      ┌────┴────┐
          [V, v, T]    V        Obj
                     kisses
                    [V, v, T]
```

In the final step of the derivation the complex head moves out of its phrase again and merges with the root node, now projecting its T feature. With this *kisses* has no active features left, and the derivation continues with the external merge of a new head.[28]

(27)

```
                        TP
              ┌──────────┴──────────┐
              T                     vP
           kisses            ┌──────┴──────┐
          [V, v, T]          v             VP
                          kisses      ┌────┴────┐
                         [V, v, T]    V        Obj
                                    kisses
                                   [V, v, T]
```

Importantly, in the reprojection analysis the complex head does not move into an already pre-existing head position; that head position is created by the movement.

3.3.2 The pros and cons of this approach

In this approach no problem with the Extension Condition arises (the movement extends the tree at the root), the Uniformity Condition on chains (the moved element is a head at both the tail and the head of the chain), and the definition of c-command. Reprojective movement does not violate anti-locality either. While the movement of the complex head is short, it leads to new feature satisfaction: after the movement a selectional feature (in (26) the *v* feature's requirement for a V complement) can be satisfied.

It is also possible to argue that the A-over-A Principle is not violated by this movement. It is true that in (26) the node V is extracted from VP, which, in BPS, also has the V label. However, the trigger of the movement is the *v* and T features of the complex head; V is only pied-piped along with these features. Therefore technically, we are dealing with the movemet of *v* and T over V, rather than the movement of V over V.

The locality constraint on HM naturally falls out from the assumption that all features of the complex head must be discharged before another external merge can take place. This analysis thus predicts that genuine long HM cannot occur (cases that look like long HM could be captured by remnant phrasal movement, though). Cases of excorporation are also excluded: all the active features of a complex head must be discharged before a new head is merged.

---

[28] A valid question that arises here is how a Tense suffix and a lexical verb can form a complex verb in languages without V-to-T movement, for instance in the case of English *kisses*. At least two analyses are possible. One is that both English and *English'* have a verb-moving syntax, as in (26), but they differ in which copy of V is pronounced (the lowest one in English and the highest one in *English'*). The other possible analysis is that in contrast to *English'*, English has no verb movement. This is possible if *-es* is not a real tense suffix (if it were, it would trigger movement of the complex verb to T, as in (26)). In this approach *-es* can be viewed as e.g. agreement with a morphologically free zero tense morpheme merged in T (see Fanselow 2004 and Biberauer & Roberts 2010).

Reprojection potentially makes an interesting and correct prediction when applied to the nominal domain. DP/NP-internal movements are subject to a well-known restriction: only projections that contain N can be displaced. For instance, it is possible to move N(P), or the constituent comprising N and Adj, but it is not possible to move Adj on its own (see Cinque 2005 and Abels & Neeleman 2009 for an exhaustive list of the movements that are allowed and disallowed by this restriction). Georgi & Müller (2010) show that if DP/NP internal movements are modeled with reprojection, then this pattern is predicted. The hedge "potentially" is used here because this prediction is made only if reprojection is coupled with some non-standard assumptions about nominal structures (e.g. the maximal extension of nominal phrases is NP, not DP, and AP, NumP, and DP are NP-specifiers).

While this approach offers a comprehensive solution to the problems raised by head-adjunction, it also raises some new problems. For instance, if the complex head *kisses* has V, *v* and T features, then it is not entirely clear why *v*'s selectional requirement for V (and T's selectional requirement for *v*) cannot be checked already within the complex head, by the V (and *v*) feature. Incorporation might also pose problems for this approach. Surányi (2008: 313) argues that "When a functional head F is morphologically free, it will not be generated as part of the inflected head H, which will then never raise to F by HM". If this is to be maintained, then in incorporating languages nouns have to be listed both as free and as bound elements in order to capture both non-incorporated and incorporated cases. One possible track to take here is to assume that incorporation is always pseudo-incorporation (i.e. incorporation of phrases). Whether this is a plausible approach or not will have to be settled on the basis of the data (but noun incorporation in Mohawk and Mapudungun always leave behind NP-modifiers, which supports Baker's original head incorporation analysis, cf. Baker 2009: 153).

3.4 *Phrasal movement*

The final syntactic alternative to head-adjunction is (complete or remnant) phrasal movement. This view is advocated by Koopman & Szabolcsi (2000); Massam (2000); Rackowski & Travis (2000); Kayne & Pollock (2001); Mahajan (2003); Nilsen (2003); Müller (2004); Pollock (2006) and Bentzen (2007) for verb movement, and by Shlonsky (2004); Cinque (2005) and Cinque (2010) for noun movement, to mention just a few.

Some proponents of the phrasal movement analysis hold that syntax cannot move heads at all; data like (1b) to (4b) always involve phrasal movement (Mahajan 2003). Others allow for syntactic movement of heads under restricted circumstances, while maintaining that the majority of the relevant data are derived by phrasal movement (e.g. Koopman & Szabolcsi 2000: 41–42).

3.4.1 The mechanics

In this approach data like (1b) to (4b) arise when a phrase whose last (and possibly only) overt element is a head moves to the specifier of the next higher head (29).

(28)

XP
├ X   YP
      Y

(29)

XP
├ YP$_i$ X′
  Y   X   t$_i$
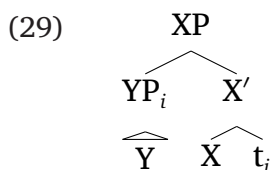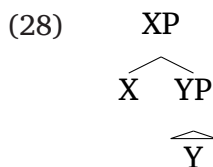
If in (29) Y is a free morpheme and X is a bound morpheme, then after linearization their order is "Y precedes X", and X can simply lean onto Y for phonological support in the linear string. (The fact that in the syntactic hierarchy there is a phrase boundary between Y and X does not affect this.) That Y-X is a morphological word is not reflected in the syntactic structure.

In several cases the final output should be a Y-X morphological word (where X and Y are exponents of syntactic heads), but the (originally lower) Y head already has a phrasal complement. In this case suffixation is achieved by remnant movement. First the complement of Y must move out of the way, creating a phrase that contains the head Y as the last overt element. The remnant YP then moves to the next higher specifier position. As the output of this operation two heads become string-adjacent and affixation can take place. Depending on the final word-order, evacuating movements may be required for the specifier and the adjuncts of YP, too.[29]
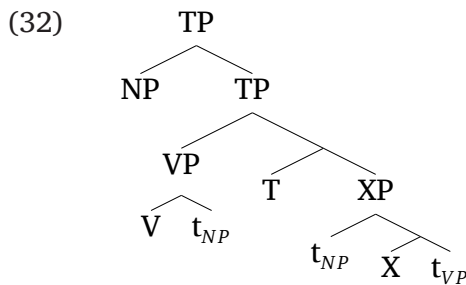
Let us consider V-to-T as a specific example. If we are dealing with an unaccusative verb, and all phrases are generated with a head-first order, then in order for V to pick up a tense suffix, the first step of the derivation involves evacuation of the sole argument out of the V-complement position to the specifier of a higher projection (30). The remnant VP can then move to Spec, TP, whereby V and T end up adjacent on the surface (31). In the final step the deep structure object moves above the remnant VP, either to a second, outer specifier of TP or the specifier of a higher projection (32).[30]

(30)

(31)

---
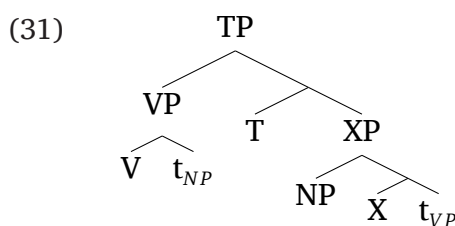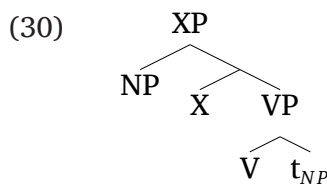
[29] While HM qua adjunction can create complex words with both suffixation and prefixation, on the assumption that movement is always up and to the left, phrasal movement analyses can only account for suffixation. This is not to say that such analyses cannot capture prefixation: it is possible to view them as base-generated structures, with the higher head leaning onto the lower head for phonological support. (It is important that in these cases the specifier position must not be occupied by an overt constituent.) The point is that the complex words of prefixing and suffixing languages will have different internal structures.

(i)
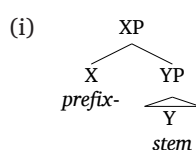
[30] Alternatively, the NP could move to Spec, VP. Then XP would not be required in the structure: VP movement to Spec, TP would directly produce the surface word order "subject precedes verb". NP movement to Spec, VP would violate anti-locality, however.

(32)

```
                    TP
              ┌──────┴──────┐
             NP            TP
                       ┌────┴─────┐
                      VP          │
                    ┌──┴──┐    ┌──┴──┐
                   V    t_NP  T     XP
                              ┌─────┼─────┐
                            t_NP    X    t_VP
```

With a transitive verb, there is also a *v*P in the structure. In this case the remnant VP could move to Spec, TP over *v*P followed by subject movement to a higher position. This would be a case of long HM (see Mahajan 2003 for such an analysis). Alternatively, there could be object-evacuation from VP and subject-evacuation from *v*P, followed by remnant *v*P movement to Spec, TP.

3.4.2 The pros and cons of this approach

This approach straightforwardly solves the problem of the Extension Condition (the movement extends the root of the tree) as well as the problem with the Chain Uniformity Condition (both the head and the foot of the chain are unambiguously phrasal) and the c-command condition (the moved phrase c-commands its trace). It also solves the problem with the A-over-A Principle: the head is not extracted from its maximal projection bearing the same label and same features. Instead, the whole phrase moves. The phrasal movement analysis can also potentially avoid violation of anti-locality; whether this is the case depends on how far the phrase moves from its base position. Cases that involve movement from the complement to the specifier of the same head do violate anti-locality.

Remnant XP-movement, however, does not straightforwardly predict the constraints that have been observed on the HM operation and that make it different from garden variety XP-movement (Section 1.2). Firstly, HM is more local than phrasal movement: the former cannot skip intervening heads, while the latter can move across phrasal positions on the way. The phrasal movement approach thus predicts massive anti-mirror effects for morphologically complex heads. While relevant cases have been argued to exist (see e.g. Muriungi 2008), they are not as frequent as one might expect in this analysis. One way to tackle this issue is to assume that the relevant features are arranged in the tree such that their checking/valuation will always give rise to a short movement. Another possibility is to assume that the Head Movement Constraint is wrong. Long head movement has been defended in Rivero (1991; 1993); Rivero & Terzi (2005); Roberts (2010); Harizanov (2016) and Preminger (2017), among others.

As discussed in Section (1), it has long been assumed that the HM operation must be roll-up, as this can capture the observation that there is no excorporation from complex heads. Phrasal movement, on the other hand, can be either roll-up or successive cyclic. One could assume that excorporation would leave a suffix dangling without the proper host, and would violate a morphological constraint. This would not extend to all cases of incorporation, however, as in some cases neither the incorporee (the noun or noun phrase) nor the lexical item incorporated into (the verb) is a morphologically bound element. Koopman & Szabolcsi (2000: 40) argue for another solution, namely that if YP is the moving phrase that contains only an overt head, then "Either there is no higher head that attracts YP, or if there is one, YP is already buried in specifiers by the time that head is merged and thus cannot extract on its own anyway". On this view, however, it remains a coincidence that phrases that feature only an overt head are always involved in feature checking/valuation relations that yield such a configuration. One way out of

this problem is to admit excorporation into the grammar. This view is held by Roberts (1991; 2010) (but see Roberts 2010 for acknowledgement that the data in Roberts 1991 can be analyzed in other ways). This, however, remains controversial: Julien (2002) argues extensively that excorporation does not exist. For a possible solution to these problems, see Funakoshi (2014). Finally, there is some evidence from neurolinguistic experiments that Broca's and Wernicke's aphasics treat phrasal and head chains differently (Grodzinsky & Finkel 1998). This is again not predicted under the phrasal movement analysis.

This approach also raises new problems regarding the triggers and the landing sites of the evacuating movements. In most cases, there are no plausible triggers; the movement only takes place to create the right word order, and the phrase whose specifier serves as the landing site cannot be motivated on independent grounds. An important exception that explicitly addresses the issue of (object evacuating) triggers and landing sites is Mahajan (2003). His proposal adopts Sportiche's (1997) idea that verbs combine with bare nouns only, and determiner heads occupy positions in the clausal spine. In order for the noun to be associated with its determiner, it has to move out of the VP to the DET head. Mahajan proposes that this is the trigger for object evacuation.[31] In many cases, however, constituents other than objects also have to undergo evacuating movements, and these movements also require triggers and landing sites. The evacuating movements also often do not show reconstruction effects, which raises the question if the surface orders in question should really be modeled via movements.

4 HM as a combination of a syntactic and a post-syntactic operation

In this section we turn to an analysis in which data like (1b) to (4b) arise via syntactic movement of a head to a specifier position, followed by a rebracketing operation that creates the morphologically complex head. This approach is pursued in Matushansky (2006); Vicente (2007) and Gallego (2010), among others.
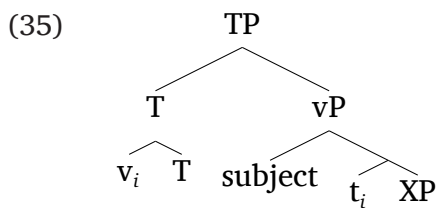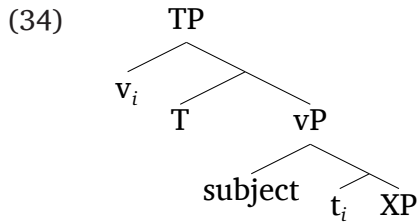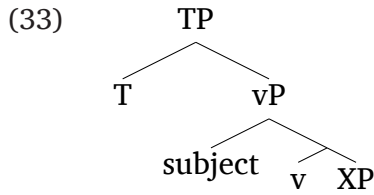
4.1 The mechanics

In Section 2 we have seen that in BPS head-adjunction violates the Chain Uniformity Condition. Some researchers suggest that this condition is too strong (or wrong), however. Matushansky (2006); Vicente (2007); Gallego (2010) and others propose that heads move to phrasal (specifier) positions rather than adjoin to higher heads. v-to-T, for instance, involves v moving to Spec, TP, as in (34) (the results of the earlier V-to-v step are not shown here).[32] This step is followed by a rebracketing operation called morphological merger (m-merger). M-merger forms a complex head out of the moving head and the head whose specifier serves as the landing site (35). The output of HM to specifier followed by rebracketing is a head adjunction structure, just like in the case of the GB-style head-adjunction approach.[33]

---

[31] In Mahajan's analysis object movement out of VP happens only in SVO languages; in SOV languages the object moves only as far as Spec, VP and is carried along with VP-fronting to Spec, TP. He suggests that SOV languages have no object evacuation because they have no DET heads. This proposal would be falsified by SOV languages with a definite article.

[32] See Preminger (2017) for an analysis in which this movement is allowed only if the moved head is in a non-branching phrase, and Harizanov (2014) for a proposal in which this movement can also affect branching maximal projections.

[33] Toyoshima (2001) proposes a strong lexicalist analysis in which there is no m-merger: the complete inflected word is inserted into the low head position and then moves to the specifier of an empty head.

(33)

```
          TP
         /  \
        T    vP
            /  \
      subject   \
              v   XP
```

(34)

```
          TP
         /  \
      v_i    \
           T    vP
               /  \
         subject   \
                 t_i   XP
```

(35)

```
            TP
           /  \
          T    vP
         / \   / \
      v_i  T subject \
                   t_i   XP
```

M-merger is subject to the constraints in (36). When these conditions are met, m-merger is obligatory. This means that movement of a head is always followed by m-merger.

(36)   Morphological merger (Vicente 2007: 49)
        Two constituents y and x may undergo m-merger if
        a.   y and x form a complex word, or a subpart of one
        b.   y and x are linearly adjacent
        c.   y and x stand in a spec-head configuration

There is some disagreement as to which grammatical component m-merger is part of. Matushansky (2006) argues that m-merger takes place post-syntactically. She suggests that syntax interfaces with the post-syntactic component after every Merge operation. This means that while movement and rebracketing take place in different components of the grammar, rebracketing can still immediately follow the movement. She argues that rebracketing returns a feature bundle, and it also involves partial spellout of the rebracketed structure, i.e. the complex head. The rebracketed and spelled out structure is then handed back to narrow syntax, where the derivation can continue.

Importantly, in this approach both the syntactic and the post-syntactic component are implicated in every individual step of HM. There are other approaches, too, in which both syntax and post-syntax have a role to play (see esp. Harizanov 2016; Gribanova 2017b and Harizanov & Gribanova accepted and the references in fn. 58). However, in these models syntax and post-syntax do not work together in any individual step of head movement: each HM step is either purely syntactic or purely post-syntactic. For instance, in these models V-to-v could be syntactic and v-to-T could be post-syntactic, but the syntactic and post-syntactic components are never both involved in V-to-v or v-to-T.

Vicente (2007) and Gallego (2010), on the other hand, argue that m-merger takes place in narrow syntax; post-syntax is not involved in the HM operation at all. Their analysis thus forms a natural class with the approaches surveyed in Section 3 in a way that Matushansky's does not. Vicente argues that complex heads are not opaque to syntactic operations in general. For instance, parts of complex heads are accessible for binding and

coreference relations (Section 2.2.2) and possibly also variable-binding (Vicente 2007: 17, fn. 11). Therefore spellout cannot be involved in m-merger. It is only movement that complex heads are opaque to. For Vicente (2007: 48), m-merger has a morpho-phonological trigger: word formation ("it happens so that two morphemes can be spelled out as a word"); and the no excorporation condition is a phonological restriction.[34]

4.2 *The pros and cons of this approach*

The movement to specifier plus rebracketing analysis solves many problems raised by head-adjunction. It eliminates the problem with the Extension Condition because it involves a cyclic movement: it extends the root of the tree. This analysis does not require a cumbersome definition of c-command either, as the moved head straightforwardly c-commands its lower copy from the specifier position. Matushansky further proposes that if m-merger is taken to return a feature bundle, then we can derive the no excorporation condition, as syntax can move whole feature bundles but it cannot subextract from them.[35] She suggests that the Head Movement Constraint can also be derived if we assume that the movement takes place to check a c-selectional feature on the higher head by the lower head. Since c-selection is a relation between a head and (the head of) its complement, the moving head will always target the next higher, selecting head. Gallego (2010) proposes that the anti-locality problem can also be solved if after head movement both the moved head and the target head project, creating a hybrid label, e.g. after *v*-to-T movement, the label is the composite *v*-T. As the creation of this new label becomes possible only as a result of the movement, anti-locality is not violated. Movement to specifier plus rebracketing still violates the A-over-A Principle, however.

There is also an important new problem that arises with this particular approach: rebracketing does not look like a licit syntactic operation, and it also violates the Extension Condition. One answer to this problem is to relegate it to the post-syntactic component, as Matushansky does. This requires syntax to interface with the post-syntactic component at every step. Richards (2011), however, shows that this is difficult to reconcile with phase theory.[36]

5 HM as post-syntactic movement

Given the problems with the head-adjunction analysis, and the conviction that the HM operation does not have semantic effects,[37] a body of literature suggests that the operation producing data like (1b) through (3) takes place post-syntactically. These approaches are often referred to as "PF movement" analyses. Before we begin the discussion of these approaches, it will be useful to clarify what is exactly meant by "PF" in "PF movement". Once this has been done in Section 5.1, I will turn to the motivations and arguments for placing the HM operation into the post-syntactic part of grammar in Section 5.2. The possible mechanics and a sample derivation will be the topic of Section 5.3, with the pros and cons discussed in Section 5.4. In Section 5.5 I will briefly discuss some analyses of (1b)–(3) that call themselves "phonological" but do not, in fact, involve post-syntactic displacement, and so do not belong to the same group as the analyses in Section 5.2.

---

[34] The trigger of phrasal movement, on the other hand, is the valuation of formal features.

[35] Matushansky also proposes that m-merger involves (partial) Spell Out. Therefore excorporation would be excluded even if the output of m-merger had analyzable syntactic structure.

[36] As pointed out by a reviewer, this approach also raises the problem of how to explain the fact that movement paths are punctuated rather than uniform (Abels 2012: Chapter 2; van Urk 2015).

[37] But see Section 7.

5.1 Post-syntactic movement, PF movement

According to current Minimalist ideas, syntax manipulates abstract units without phonological content. When the syntactic derivation reaches the point of Spell Out, the derivation splits into two branches. One branch ships the syntactic structure to the LF interface (LF branch), while the other branch ships the syntactic information to the PF interface (PF branch).[38]

An explicit theory of the mapping from Spell Out to the PF interface is provided by Distributed Morphology (DM). According to DM, in the early phase of the PF branch nodes are still in the hierarchical arrangement that syntax has created. This hierarchy can be slightly modified by a few types of morphological processes (Fusion, Fission, Lowering, etc.) that happen due to language-specific morphological or morpho-phonological requirements. This is followed by Vocabulary Insertion (the pairing of abstract nodes with Vocabulary Items) and Linearization. After Linearization, phonological information is present in the representations, but hierarchy is not; morphemes are related to each other by precedence and subsequence rather than c-command or dominance. Linearization is followed by various phonological processes, e.g. the building of prosodic domains. The final product of the mapping from Spell Out to PF is Phonological Form (see Embick & Noyer 2001: Fig. 1.).

To the best of my understanding, all analyses that claim that HM is PF movement mean that HM takes place after Spell Out, in the PF branch of grammar, but before Vocabulary Insertion/Linearization. What this means is that HM applies in a part of grammar that follows narrow syntax, but which still retains the hierarchy produced by syntax. There is no analysis, as far as I understand, that claims that HM takes place after Vocabulary Insertion/Linearization, when only precedence/subsequence information is accessible to the representations. In the rest of this paper, post-syntactic movement and PF movement will be used as synonymous terms and will refer to movement that takes place after Spell Out but before Linearization.

5.2 Arguments that HM takes place after syntax

Chomsky suggests that since HM has no semantic effect, it plausibly takes place at PF and is "conditioned by the phonetically affixal character of the inflectional categories" (Chomsky 2001: 38). He observes that none of the movements that he assumes to take place at PF can iterate, and raises the possibility that lack of iteration is a property that characterizes PF operations in general.[39]

Beyond this, he says very little about the nature of PF movements. He suggests that one of the PF movements, namely Thematization/Extraposition, adjoins the internal argument to vP in the case of rightward movement, while it substitutes the internal argument in Spec, vP in the case of leftward movement (Chomsky 2001: 23). From this, it is clear that this type of PF movement takes place in the part of the PF branch where hierarchy is still retained (i.e. before Linearization), but this is never made explicit. We can assume that heads also move in this part of the PF branch (but this is again not claimed explicity). This is consistent with his proposal that HM does not to create a chain (Chomsky 2001: 39): post-syntactic movements in DM (such as Lowering) are assumed not to leave a trace.

While the theoretical arguments for HM being a post-syntactic movement come from the lack of semantic effects, the empirical arguments mainly involve data from ellipsis constructions. Boeckx & Stjepanović (2001) argue that pseudogapping constructions support

---

[38] This is called the Y-model of grammar.

[39] HM is not interable in the sense that once a head has moved to the next higher head, it cannot move further on its own. In the case of v-to-C, for instance, v can move to T, but it cannot simply move on to C from this position; after v-to-T, it is the complex T head that moves on to C rather than v on its own.

the idea that heads move at PF. Lasnik (1999) observes that in English non-elliptical sentences the verb has to raise, but in pseudogapping constructions the verb may either raise or stay put and be part of the elided constituent. In Boeckx & Stejepanović' interpretation, this means that ellipsis can apply before HM, and since ellipsis takes place at PF, it follows that so must HM. They do not make any suggestions as to how PF movement works, however.[40]

Schoorlemmer & Temmerman (2012) study verb-stranding VP-ellipsis, i.e. ellipsis that affects all arguments and adjuncts in the verb phrase but not the main verb itself. (This is attested e.g. in Portuguese, Irish and Semitic.) The literature agrees that this kind of ellipsis is regular VP-ellipsis except for the fact that the verb moved out of the ellipsis site. This kind of ellipsis, however, presents an apparent paradox. It is well known that there is a general identity requirement on elided elements: in order to be recoverable, the element in the ellipsis site and the corresponding element in the antecedent have to be identical. For verbs, for instance, this means that the antecedent and the ellipsis site must contain the same verb(al root). The apparent paradox is that in verb-stranding VP-ellipsis the verb is **not** elided, but (modulo inflectional affixes) it still has to be identical to the verb in the antecedent. Goldberg (2005) proposes that this is because at LF, the verb is within the ellipsis site, and this forces it to be interpreted identically to the verb in the antecedent. Schoorlemmer & Temmerman (2012) pick up on this suggestion and argue that the verb is in the ellipsis site at LF because it does not move out of the VP either in narrow syntax or at LF; the movement takes place at PF only (see also McCloskey 2017). This movement thus affects the linear order but not the interpretation (and so being in the ellipsis site at LF, the verb is subject to the identity requirement on ellipsis).[41]

It is interesting to note that both Boeckx & Stjepanović (2001) and Schoorlemmer & Temmerman (2012) use VP-ellipsis and verb movement to argue that (at least in certain cases) heads move at PF, however, in order for the Boeckx & Stjepanović (2001) analysis to work, ellipsis has to be able to precede HM at PF (or ellipsis and HM apply at the same time, and one has to choose between them), while for the Schoorlemmer & Temmerman (2012) analysis, HM must be able to precede ellipsis at PF.[42]

That data from ellipsis support placing the HM operation into the PF component has been seriously challenged in the recent literature. In direct opposition to Schoorlemmer & Temmerman (2012), Gribanova (2017a) argues that the identity condition in verb-stranding ellipsis is actually an argument for HM being a narrow syntactic operation. Stripped to its essentials, her argument proceeds like this: HM influences the way parallel domains are calculated between the antecedent and the ellipsis site, therefore it feeds LF, therefore it must be syntactic. (The paper offers detailed argumentation which I cannot reproduce here for reasons of space.) Relatedly, Lipták (2017) shows that the identity condition on verbs studied by Schoorlemmer & Temmerman cannot be due to the special nature of HM. On the one hand, the same condition also holds in answers to questions that involve no ellipsis, therefore it is not an effect of ellipsis or HM out of an ellipsis site.

---

[40] Both Chomsky (2001) and Boeckx & Stjepanović (2001) cite Grodzinsky & Finkel's (1998) neurolinguistic study as support for the view that HM takes place at PF. This work shows that aphasic patients treat head-chains differently from XP-chains. However, the finding that HM is processed differently from XP-movement does not automatically mean that the two types of movements take place in different modules of the grammar. The approaches surveyed in Section 6 are also compatible with Grodzinsky & Finkel's findings, but unlike Chomsky (2001), these approaches do not assume displacement at PF. It is worth noting, however, that the results of this neurolinguistic study are problematic for phrasal movement approaches.

[41] Schoorlemmer & Temmerman (2012) do not suggest, however, that every instance of HM takes place at PF, and like Boeckx & Stjepanović (2001), make no suggestions regarding the mechanics of HM at PF.

[42] Note that the two orderings are not necessarily in contradiction: both analyses could be accommodated if HM and VP-ellipsis are freely ordered.

On the other hand, an identity condition also holds of certain XPs moving out of ellipsis sites (see also Gribanova 2017a). This means that the identity requirement is independent of heads, and so it cannot be used as an argument for the special nature of the movement of heads.

Gribanova (2017b) points out, however, that authors who study ellipsis and come to contradictory conclusions about the place of the HM operation in grammar (syntax or post-syntax) are looking at different languages. She suggests that it is possible that all of them are right for the particular cases they are studying. Specifically, two different operations may have data like (1b) to (4b) as their output, and while one of these operations takes place in syntax, the other happens in the PF branch.[43] If this is so, then data from different languages must be looked at on a case-by-case basis, and a compelling argument from one case does not warrant definitive conclusions about the HM operation across the board.

5.3 The mechanics

In general, proponents of the PF movement analysis do not address the question of what the exact mechanism of post-syntactic HM is. As Boeckx & Stjepanović (2001: 353) acknowledge, "a full-fledged theory of PF operations remains to be worked out before the view that head movement falls outside the core computational system can be fully endorsed". To my knowledge, the only work that addresses this issue is Harizanov & Gribanova (accepted).[44] This paper suggests that data which GB analyzed with head-adjunction can arise either as a result of a syntactic operation or as a result of a post-syntactic operation, and it offers an explicit discussion of what the post-syntactic operation consists of.

Harizanov & Gribanova suggest that heads that (for language-specific reasons) have to form a morphological word with another head are endowed with the binary morphological selection feature M. The [M:–] specification triggers adjunction to the structurally adjacent lower head. This is, in effect, a formalization of DM's well-known Lowering operation. The [M:+] specification, on the other hand, triggers adjunction to the structurally adjacent higher head. This operation, called Raising, is basically the upward counterpart of Lowering, and its effect is that a head is spelled out higher than its syntactic merge-in position.[45]
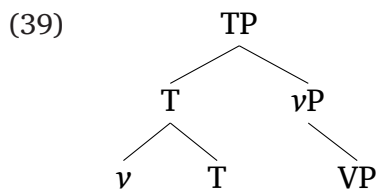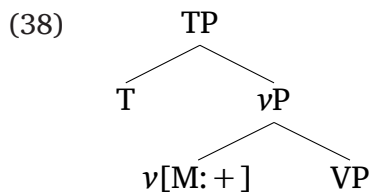
(37)   Post-syntactic head Raising (Harizanov & Gribanova accepted)
         $[_{XP} \ldots X \ldots [_{YP} \ldots Y [_{ZP} \ldots]]] \rightarrow [_{XP} \ldots [_X Y X] [_{YP} \ldots [_{ZP} \ldots]]]$
         (where Y and X are heads, X c-commands Y, and there is no head Z that
         c-commands Y and is c-commanded by X)

In this analysis the case of post-syntactic *v*-to-T, for instance, is triggered by a [M:+] feature on *v* (38). After Raising, *v* and T form a complex head as in (39). It is always the head that is amalgamated into that projects the label of the complex head, so in this case the label will be T.

---

[43] Note that Chomsky also advocates two types of HM: he suggests that head incorporation is a narrow syntactic operation. Chomsky and Gribanova, however, cut the pie of syntactic vs. post-syntactic HM operations differently.

[44] See, however, Parrott (2001) for related ideas that remained in a manuscript form. The two papers share the core idea that PF movement happens after Spell Out but before the hierarchy is converted into linear order. The details of execution are significantly different in the two papers, however.

[45] As Harizanov & Gribanova note, Lowering and Raising could be unified into a single operation, Amalgamate, but they leave working out the details for further research.

(38)

```
        TP
       /  \
      T    vP
          /  \
      v[M:+]  VP
```

(39)

```
          TP
         /  \
        T    vP
       / \     \
      v   T    VP
```

After Raising (or Lowering) has taken place, the [M] feature on the moved head is erased or becomes inactive. In (39) this is reflected by the lack of [M:+] on *v*. Note also that in (39) there is only one instance of *v*. This is because Raising/Lowering does not leave a trace, and so it does not involve chain formation.

The output of Raising is very similar to the output of the head-adjunction analysis in Section 2, but the operation that produces the complex head resides in the post-syntactic component rather than in narrow syntax.

5.4 *The pros and cons of this approach*

In Section 2 we have seen that the GB-style head-adjunction analysis violates several syntactic principles. If the HM operation does not take place in narrow syntax, then by definition, it cannot violate any syntactic principles and can pose no problems for syntax. There is, therefore, little point in checking post-syntactic HM against our original list of problems: all of them will be eliminated.[46] In and of itself, however, this does not mean that placing the HM operation into the post-syntactic component is superior. The posited operation must fit with what we independently know about post-syntactic operations, and the theory working with it should be internally consistent and constrained, making the right empirical predictions.

In DM, all post-syntactic operations are highly local in nature. Raising fits with this view because it operates only on structurally adjacent heads. The Head Movement Constraint and the no excorporation condition are baked into the definition of Raising, so these constraints will always be obeyed. Raising also delivers the Mirror Generalization: Harizanov & Gribanova suggest that like other post-syntactic operations, Raising also proceeds bottom up, which means that lower heads will be closer to the stem than higher heads.

There are, however, some general architectural issues with DM that naturally, hold of this approach, too. The first issue is that post-syntactic operations that work on the hierarchical representation (i.e. all operations before Vocabulary Insertion) are triggered by morpho-phonological properties of Vocabulary Items, but at the point that these operations are in effect, Vocabulary Items have not yet been paired with terminals. In other words, these operations, including Raising, involve look-ahead. One might argue that Raising (and Lowering) is exempt from this problem. Harizanov & Gribanova (accepted: 22) argue that the M feature triggering post-syntactic head-adjunction is "one of the features in the feature bundle that constitutes the lexical item and … the lexical item comes

---

[46] It may seem that the Extension Condition, the c-command condition, the Chain Uniformity Condition and the A-over-A Principle are violated in the PF branch. If these are constraints on hierarchical representations in general, then post-syntactic HM violates them. If, however, these constraints hold strictly of narrow syntax, then post-syntactic movements do not fall under their purview.

specified with a value for M from the lexicon". Therefore the derivation can simply make reference to a feature that is inherently part of the terminal; it does not have to know about properties of actual Vocabulary Items, and so no look-adhead is involved.[47] This is problematic because M is not a syntactic feature in the sense that syntax does not make reference to or manipulate it, and recent Minimalist work aims to eliminate all non-syntactic features from narrow syntax.[48] The second issue is whether post-syntactic operations are indeed indispensable, or we could make do without them and so simplify the model or grammar. Theories of syntax without post-syntactic operations are pursued both within DM (Julien 2002) and outside of DM (e.g. in Nanosyntax, see Caha 2009 and Starke 2009; 2014). Whether post-syntactic operations are well motivated is, in the end, an empirical question: they are if their properties can be shown to be systematically different from those in syntax.

5.5 Analyses that involve PF movement in name only

Some use the term "PF movement" or "phonological movement" to characterize their alternative to head-adjunction, but the operation they posit does not, in fact, involve post-syntactic movement. Here I will briefly discuss two of them, because it is important to see how they differ from the analyses surveyed earlier in this section.

Building on ideas in Hale & Keyser (2002), Harley (2004) proposes an analysis that she considers to be "a natural candidate for a Minimalist, phonological head-movement mechanism" (Harley 2004: 240). In short, the analysis works as follows. Each head in syntax is endowed with a position-of-exponence; a kind of place-holder for phonological features that will be filled in during post-syntactic Vocabulary Insertion. In line with BPS, an XP is assumed to have all the features that its head X does, including the position-of-exponence of the head. When XP is merged with a head α whose position-of-exponence is defective, a Conflation mechanism takes effect: XP's position-of-exponence is merged into α's position-of-exponence. As XP has the same position-of-exponence as its head, Conflation means that α acquires the position-of-exponence of X, the next head down. After conflation X's position-of-exponence will be present in the tree both at X and at α, and as usual for elements with multiple copies, only the highest one is pronounced. This means that X will be spelled out at the next higher head, α.

It is clear that Harley's analysis does not involve movement in the PF branch: the operation that is responsible for a head being pronounced higher than its merge position takes place during the narrow syntactic derivation. In fact, Harley claims that her analysis does not involve any movement at all (Harley 2004: fn. 5 and Harley 2013: 73).[49] Conflation can be said to be a phonological operation in the sense that only the phonology-related subpart of the head is affected by it.

Zwart (2001) argues that head movement has two subtypes: syntactic and phonological head movement. The term "phonological head movement", however, is potentially misleading, as this type of movement, too, takes place in syntax rather than in the PF branch

---

[47] In order to eliminate the look-ahead problem, this approach could be extended to the features that trigger Fusion and Fission as well.

[48] It is subject to ongoing debate to what extent this is necessary or possible, however. Marantz (1995) and Haugen & Siddiqi (2013) argue for the complete elimination of morpho-phonological features from all syntactic terminals, while Embick (2000) and Embick & Noyer (2007) suggest that this should hold for only a subset of the terminals.

[49] It is not clear that this is the case, however: Conflation, in Harley's words, involves copying during the derviaton plus pronunciation of the higher instance of an element – exactly the operations that are implicated in syntactic movement, too. It is true, however, that Conflation affects only part of a head (namely its position-of-exponence) rather than the whole head.

after Spell Out. Zwart suggests that syntactic terminals may have two types of features: all of them have F[ormal] features, and in addition, many of them also have LEX[ical] features. Agree chains always involve movement of formal features. If this is all that happens, then the movement has no phonological reflex in the phonological component, thus we get covert head movement. This is what he terms "syntactic movement". If the highest head in an Agree chain is defective in the sense that it has no LEX-features of its own, then the LEX-features of the bottom of the chain also move along with the formal features. LEX-feature movement has a reflex in the phonological component, i.e. it yields overt head movement. Zwart calls this type of movement "phonological movement".

It is phonological, however, only to the extent that it is "triggered by requirements of the spell-out procedure only" (Zwart 2001: 38). The movement's target is set by syntactic mechanisms (feature-valuation), and the movement takes place in narrow syntax. LEX-movement always accompanies movement of formal features, and in this sense, ""phonological verb movements" are a subset of "syntactic verb movements"" (Zwart 2001: 60).

Harley's and Zwart's analyses differ in significant details. The following points, however, are shared by their proposals: 1) syntactic terminals are endowed with phonology-related features (Harley's position-of-exponence, Zwart's LEX-features), 2) if a higher head's phonology-related features are defective, a lower head's phonology-related features move up to the higher head, 3) the movement takes place in syntax, and 4) morphemes and linear order are not manipulated in the PF branch itself.

Another potentially misleading use of the term "PF" appears in Platzack (2013). The title of this paper is "Head movement as a phonological operation", but his analysis does not involve any movement either during or after syntax. Instead, it belongs to the group of analyses discussed in Section 6.

6 Post-syntax, no movement: Direct Linearization Theories

In this section we turn to Direct Linearization Theories. These theories posit that no actual movement is involved when a head is pronounced higher than its merge-position; the illusion of movement arises as a result of the way syntactic structures are linearized. This approach is pursued in Brody (2000a); Adger (2013); Ramchand (2014) and Hall (2015), among others.

It has been an accepted thesis for a long time that syntactic representations contain only hierarchical information, and the way the hierarchy maps onto a linear order must be stated separately from syntactic rules. The most influential mapping rule in Minimalism is Kayne's (1994) Linear Correspondence Axiom (LCA), which translates asymmetric c-command relations into linear precedence relations.
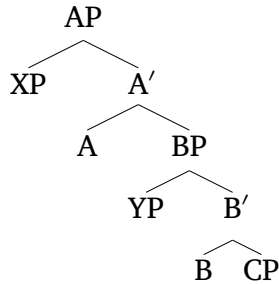
A well-known feature of Kayne's system is that it requires many semantically empty movements in order to create the structures that will translate into the correct word order. These movements have no plausible syntactic trigger and do not show reconstruction effects (see Section 3.4 on phrasal movement). So-called Direct Linearization Theories (DLTs, a term coined by Ramchand 2014) address this problem by i) using syntactic representations different from the familiar GB or BPS trees (so-called Telescopic representations), ii) using mapping rules different from the LCA, and iii) base-generating Kayne's roll-up structures (thus eliminating the need for a movement trigger and predicting the lack of reconstruction effects). These theories are of interest to us here because they model the upward displacement of a head's exponent by a specific mapping rule (or in Ramchand's terms, Direct Linearization Statement) from syntax to linear order without involving any movement (syntactic or post-syntactic) in the process.
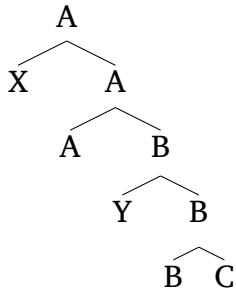
6.1 The mechanics

In this section we first look at syntactic representations in DLTs, then discuss how these structures are mapped onto linear order. This discussion will lead to the analysis of data like (1b) to (4b) in DLTs.

In both GB-style and BPS representations, if a head has both a complement and a specifier, then the head, the intermediate projection, and the phrase are represented by separate levels of projection in the structure.

(40)   Government and Binding

```
              AP
            /    \
          XP      A′
                /    \
               A      BP
                    /    \
                  YP      B′
                          /\
                         B  CP
```

(41)   Bare Phrase Structure

```
              A
            /   \
           X     A
               /   \
              A     B
                  /   \
                 Y     B
                     /   \
                    B     C
```

The Telescope principle, given in (42), states that this is not necessary: one node can represent both the head and the maximal projection even if the head has both a complement and a specifier.

(42)   Telescope
        A single copy of a lexical item can serve both as a head and as a phrase. (Brody 2000a: 41)
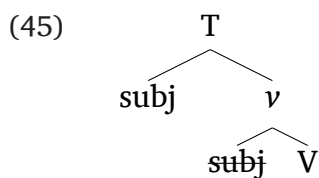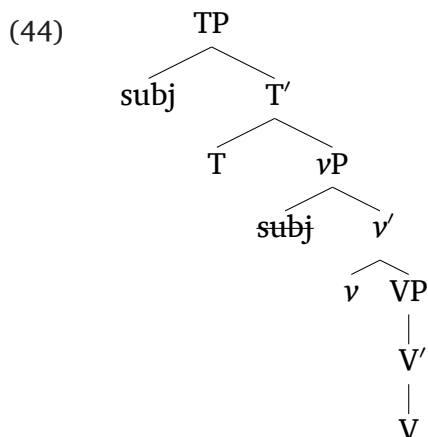
Telescope allows representations like (40) and (41) to be replaced by (43). By convention, in DLT trees specifiers are represented with leftward sloping lines, while complements are represented with rightward sloping lines. In (43) the node A in and of itself represents a head, while taken together with its dependents, the specifier X and the complement B, it represents the phrasal level.

(43)   Telescopic representation

```
              A
            /   \
           X     B
                /  \
               Y    C
```
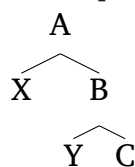
As shown by (43), the structural relationship between a selecting head and a selected head is that of immediate domination rather than c-command. The heads in a structure form one uninterrupted line; specifiers dangle from this line to the left rather than intervene between heads, as in traditional representations.

Applied to a specific example, the standard representation of (44), featuring an unergative verb, is replaced in DLTs by (45).

(44)

```
              TP
         ┌────┴────┐
       subj        T′
              ┌─────┴─────┐
              T          vP
                    ┌─────┴─────┐
                  s̶u̶b̶j̶        v′
                         ┌─────┴─────┐
                         v          VP
                                    │
                                    V′
                                    │
                                    V
```

(45)

```
              T
         ┌────┴────┐
       subj        v
              ┌─────┴─────┐
            s̶u̶b̶j̶        V
```

In both cases, the subject moves from the specifier of *v*P to the specifier of TP, which gives rise to a phrasal chain.

Let us now turn to the mapping rules from hierarchy to linear order. The first Direct Linearization Statement says that when mapped to a linear order, a specifier precedes its head.

(46)   Direct Linearization Statement for specifiers
        The specifier and its constituents precede the head. (Brody 2000a: 40)

The second Direct Linearization Statement regulates the linearization of the head and its complement.

(47)   Direct Linearization Statement for complements
        The complements and its constituents follow [the head]. (Brody 2000a: 40)

The two Direct Linearization Statements in (46) and (47) yield a specifier-head-complement order, like the LCA, but the hierarchical relation that they take to be the basis for linearization is immediate dominance rather than c-command.[50]

The third Linearization Statement regulates morpheme order within morphologically complex words.

(48)   Mirror Axiom
        The syntactic relation "X complement of Y" is identical to an inverse-order
        morphological relation "X specifier of Y." (Brody 2000a: 42)

A morpheme is the morphological specifier of the morpheme that it immediately precedes within a morphologically complex word. Thus in a morphological word of the form V + *v* + T, V is the morphological specifier of *v*, and *v*, in turn, is the morphological specifier of T. It follows from Mirror that if the exponents of a series of heads form a morphologically

---

[50] See Bury (2003) for an exception: he allows the complement-head-specifier linearization as well.

complex word (i.e. are involved in an affixation or incorporation relationship), then the order of the morphemes within the morphological word will be the inverse of the syntactic hierarchy. In other words, (48) ensures that the relationship between morphology and syntax obeys Baker's Mirror Principle. As Brody points out, Mirror in Baker's work is a generalization over the observed data. In DLTs, on the other hand, Mirror is a genuine principle; morphological structures that do not conform to it cannot be generated.[51]

A sample representation of V-to-T is given in (49). Here the subject *John* has moved from Spec, *v*P to Spec, TP. In order to make the exposition more transparent, I follow Bowers (1993); Hale & Keyser (1993); Arad (1996) and Den Dikken (2015), among others, and represent the object *Mary* as a specifier of V (hence the left-sloping line), but nothing crucial hinges on this.[52] The heads V, *v*, and T have separate exponents; those of T and *v* have an affixal requirement.

(49)
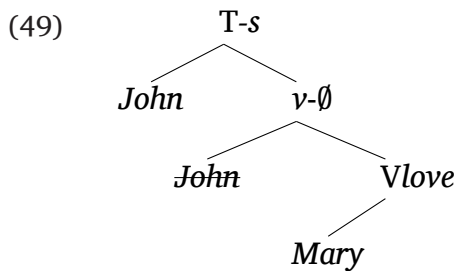
Due to the language-specific morphological rules of English, when mapped to a linear order, the exponents of T, *v* and V will have to form a morphological word. By (48), the order of the morphemes within the morphological word will be the inverse of the syntactic hierarchy, that is, V + *v* + T = *love* + ∅ + s. So in the sentence *John loves Mary*, the morphologically complex word *loves* spells out all of T, *v*, and V.

The question that arises now is in which of the three positions *loves* will be pronounced. In DLT any head that a complex word spells out is a potential spell-out position for that word: in (49) all of T, *v*, and V are possible in principle as a spell-out position for *loves*. What the actual spell-out position will be is regulated by the Positioning algorithm.

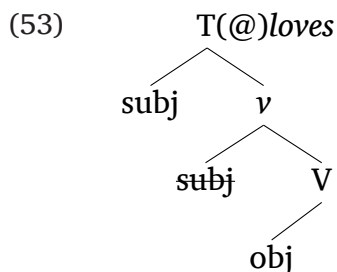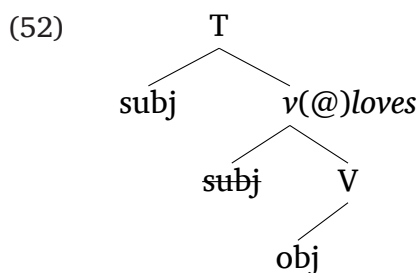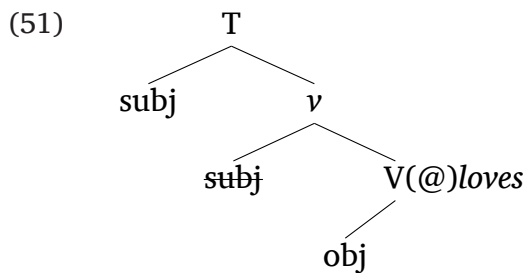(50)   Positioning Algorithm
        Pronounce an element E (a word or a chain) in the lowest position P such that
        all higher positions P' of E are weak. (Abels 2003: 270)

Positioning says that the actual spell-out position depends on whether there is a strong head in the complement line. If so, then the spell-out position will be at the highest strong head. For graphic convenience, strong heads are marked with a diacritic in syntactic trees (@, →, or *, varying across works in the DLT family). If there is no strong head in the complement line, then the morphological word in question spells out in the lowest head.

That is, in absence of a strong head, or if the highest strong head is V, then *loves* spells out down in V, and (if the object stays low and the subject moves to Spec, TP, as in English), SOV order arises (51). If the highest strong head is *v*, then *loves* spells out in *v*, yielding the SVO word order of English (52). (By (46) the object still precedes V, but now *loves* spells out in a higher head, and so it precedes V and everything that V dominates). Finally, if the highest strong head is T, then *loves* spells out at T, delivering the SVO word order of French (53).

---

[51] Exceptional cases in which the morpheme order does not plausibly correspond to the mirror order of the syntactic projections must involve phrasal movement, as discussed in Brody (2000a: 34).
[52] See also Brody (2000a).

(51)

```
              T
         ┌────┴────┐
       subj        v
              ┌─────┴─────┐
            s̶u̶b̶j̶      V(@)loves
                        └──┐
                          obj
```

(52)

```
              T
         ┌────┴────┐
       subj    v(@)loves
              ┌─────┴─────┐
            s̶u̶b̶j̶         V
                        └──┐
                          obj
```

(53)

```
         T(@)loves
      ┌──────┴──────┐
    subj            v
              ┌─────┴─────┐
            s̶u̶b̶j̶         V
                        └──┐
                          obj
```

As shown by the examples above, when the complement line contains a strong head that is not the lowest head, as in (52) and (53), then the exponents of the heads below @ appear higher than the heads themselves. However, no real movement takes place; no chain formation is involved. The syntax of languages with "no V movement", with "V-to-*v* movement" and with "V-to-T movement" is exactly the same.

It is important that (52) and (53) do not involve PF movement either. The morphemes that make up a morphologically complex word do not come together by movement under one terminal at any point; they are simply placed next to each other at Phonetic Form by the mapping rule that translates syntactic hierarchy into linear order.[53]

Readers are also encouraged to check Platzack (2013) for a related theory. Like the DLTs in this section, Platzack's theory uses direct linearization statements (his Spell-out Principles 1 and 2) that i) make it possible for a head to be spelled out in a higher head within its own extended projection, with the spell-out position marked with a diacritic (which he calls an EPP feature), and ii) ensure that within the morphological word, suffixes mirror the syntactic hierarchy. This approach differs from the DLTs reviewed here in two respects: it does not use Telescopic structures and it suggests that the spell-out marking diacritic is always on a head that enters into an Agree relation with a lower head.

---

[53] DLTs come in two types: lexicalist (Brody 1997; 2000a; b; 2004; Brody & Szabolcsi 2003) and non-lexicalist (Abels 2003; Bury 2003; Adger et al. 2010; Bye & Svenonius 2012; Adger 2013; Ramchand 2014; Hall 2015). The two approaches lead to different syntactic representations when the selecting and the selected head do not form a morphological word. This is the case with modal auxiliaries and the *v*P in English, for instance. I have glossed over this difference here and used the structures of non-lexicalist approaches. This does not affect the main points.

6.2 The pros and cons of this approach

As in the case of post-syntactic movement, many considerations raised in Section 2 are not applicable to DLTs: in this approach the operation that displaces (the exponents of) heads upwards does not take place in syntax, therefore no syntactic principles are violated or syntax-internal problems arise. What needs to be considered instead is if this theory is internally consistent, if it makes the right predictions, and if it gives rise to new problems of its own.

DLTs capture Baker's Mirror Generalization via Mirror in (48). The locality effect of the HMC and the no excorporation condition also fall out from this axiom automatically. Structures that do not conform to the HMC or which involve genuine excorporation simply cannot be generated.

The DLT approach captures the long-standing observation that the upward displacement of (the exponent of) a head is local in the sense that phrasal movement is not, and in contrast ot phrasal movement, it can only operate in a roll-up fashion (no excorporation). In DLTs, these differences arise because phrasal movement and the displacement of (the exponents of) heads are completely different mechanisms: the former takes place in syntax and gives rise to chains, while the latter does not. This view is also compatible with Grodzinsky & Finkel's (1998) findings that aphasics treat phrasal movement and the displacement of (the exponents of) heads differently.

In DLTs heads are spelled out in positions other than their merge position as a by-product of a spell-out instruction to phonology. This makes the strong prediction that for heads, the dissociation between the position of merge and the position of exponence will never have any semantic effects and will never alter syntactic locality domains. We will see in Section 7 that there are arguments both for and against interpretive and locality-changing effects of the HM operation, and arriving at definite conclusions is not easy because the arguments are often quite involved and rely on rather subtle judgments. It is clear, however, which side of the debate DLTs come down on: similarly to the post-syntactic movement approach, they predict that such effects will not arise.

There are also some problems that arise internal to DLTs. The spell-out instruction @ is probably part of the featural content of the relevant heads, and is therefore present in the syntactic representation. But since the position of pronunciation becomes relevant only after narrow syntax, a pronunciation-related feature should have no place in narrow syntax.[54] Furthermore, whether a particular head is strong or weak still cannot be reduced to an independent property.

7 How can we tell if HM is in syntax or not?

As already mentioned in Section 1, currently the biggest question is whether data like (1b) through (4b) should be modeled by a narrow syntactic operation or not. Two types of evidence have been brought to bear on this question: possible semantic effects and interaction with syntactic locality. In this section we will briefly look at these in turn.

7.1 Semantic effects

The idea that the HM operation is not part of syntax can be entertained at all because it appears to have no semantic effects. As semantic interpretation is computed at LF, and LF has a syntactic (hierarchical) representation, if HM turns out to have semantic effects, it

---

[54] A reviewer mentions that "a spelled-out sequence of heads implies that there are grammatical operations applying to non-constitutents" as a further problem. While the sequence of heads that will form a morphological word is indeed a non-constituent, there is no grammatical (or other) operation that applies to that non-constituent. Each head is mapped onto linear structure on its own. After linearization they end up next to each other in the string, and their phonological exponents amalgamate into a morphological word.

must be a syntactic operation and cannot be relegated to the PF branch or the translation rules between the syntactic hierarchy and linear order.

Matushansky (2006) suggests that in spite of taking place in narrow syntax, head movement often lacks a semantic effect because several cases (including verb movement) involve displacement of a non-scopal element, and so the logical possibility that movement leads to a new interpretation does not arise in the first place. In this respect, HM is similar to A-movement, which also characteristically lacks semantic effects, but it is nevertheless part of narrow syntax.

There are also several cases in which HM has been argued to have interpretive effects. The relevant empirical domains include i) the generic vs. existential interpretation of determinerless plural subjects in English and Spanish (Benedicto 1998), ii) the scope of modals relative to negation (Lechner 2006; 2007; Iatridou & Zeijlstra 2013), iii) the licensing of NPI subjects by English subject-auxiliary inversion (Roberts 2010; Matyiku 2014), iv) quantifier scope interaction between aspectual raising verbs and quantified subjects (Szabolcsi 2010: Chapter 3 and Szabolcsi 2011), v) the way parallelism domains are calculated in ellipsis (Hartman 2011; Gribanova 2017a), and vi) verb cluster formation in German long passives (Bhatt & Keine 2015; Keine & Bhatt 2016). Space prevents me from summarizing their arguments here; I refer the interested reader to the cited papers for the details.

Whether these works offer conclusive evidence for HM being a syntactic operation depends on the correctness of their basic (as well as auxiliary) assumptions and the correctness of the details of their analyses (cf. also Platzack 2013: 34, fn. 12). Hall (2015), for instance, argues in detail that the interpretive effects studied in Benedicto (1998); Lechner (2007); Roberts (2010) and Hartman (2011) can be captured my means other than movement. McCloskey (2016) reviews the arguments for semantic effects in Lechner (2007) and Iatridou & Zeijlstra (2013) and suggests that they are not conclusive, in part because the two studies need to make contradictory assumptions about the obligatoriness of reconstruction. Another case when two different proposals for semantic effects actually weaken each other is Lechner (2007) versus Roberts (2010). Both papers use the scopal interaction of negation with some other element in English to make a case for HM taking place in syntax, neither explicitly assumes more than one NegP in the language, and the position of NegP is crucial for both authors. However, while Lechner (2006; 2007) places NegP above TP, Roberts (2010) uses the opposite hierarchy.[55] As pointed out by a reviewer, the NPI-licensing effects discussed in Roberts (2010) and Matyiku (2014) are not necessarily strong either, as (at least in some dialects) English NPIs do not have to be c-commanded by their licensor (Henry 1995; Hickey 2007).

7.2 Interaction with locality

There are two different types of arguments in the literature that use locality to make a case that HM is a syntactic operation. The first type of argument is that the HM operation is licensed by and complies with syntactic locality principles. Preminger (2017) argues that the locality principle called Principle of Minimal Compliance (Richards 1998; 2001) is directly relevant for the licensing of HM. This principle states that if a certain position is involved in more than one syntactic dependency (movement or Agree relation), then only the first one has to meet locality criteria; the next dependencies targeting that position need not. The reader will recall from Section 2 that head-adjunction poses a problem for the A-over-A Principle. Preminger argues that HM happens only when a head H is either

---

[55] Lechner posits two NegPs when there are two morphologically negative expressions in the clause, e.g. in *No guest didn't show up*, but in these cases both NegPs are above T.

c-selected or agreed with by a higher head, and there is a second dependency between a higher head and H, too. In such a case, the first dependency, c-selection or Agree, will target HP, in compliance with the A-over-A Principle. But the Principle of Minimal Compliance allows the second dependency between the higher head and H to violate locality, therefore this second dependency can target (and move) the head in violation of the A-over-A Principle. Crucially, the Principle of Minimal Compliance is a general locality condition that also holds of phrasal movement. Therefore the fact that the HM operation also complies with it means that it must take place where general locality principles apply: in narrow syntax.[56]

The second type of argument is that the operation of head movement interacts with locality in a dynamic way, specifically, it can change locality domains in syntax. While executing the details differently, Den Dikken (2007); Gallego (2010); Stepanov (2012) and Mathew (2015) share the core idea that movement of a phase head "Ph" extends the phase boundary up to the landing site (this has been called phase extension or phase sliding).[57] This movement affects locality because the specifier of PhP is on the phase edge before movement of Ph but ends up in the phase domain after the movement. Due to the Phase Impenetrability Condition, movement of Ph makes Spec, PhP inaccessible for further operations that would have been able to affect it had Ph stayed in situ. If HM can indeed shift the phase boundary and affect locality, then it must be a syntactic operation.

Similarly to the semantic effects discussed above, locality-based arguments make a case for syntactic HM only to the extent that the proposed analyses are on the right track.

7.3 Interim summary

Whether the HM operation can have semantic effects or it can interact with locality domains are crucial questions for two reasons. Firstly, debates about HM are often too focused on the theory-internal issues, and leave room for the individual's perception of what is a more elegant or parsimonius solution to a problem. The issues brought up in this section bring empirical data into the discussion.

Secondly, if the answer to either of the above questions is a clear, unambiguous "yes", then the PF movement approach and DLTs become non-starters. These approaches are built on the premise that HM is semantically and syntactically inert, and so they predict that it cannot have an effect on interpretation or locality. Any strong evidence to the contrary will serve as a knock-down argument against these approaches, leaving only syntactic alternatives in the competing arena.

8 Conclusion: Where are we now and how to proceed?

This overview article surveyed the traditional head-adjunction model of HM, the problems with this model, and the various alternative analyses that the problems have lead to in the literature. The different approaches that have been discussed are summarized in (54).

(54) • syntax, movement
        – head-adjunction
        – same output as head-adjunction but via a different mechanism
           * sideward movement
           * defective goal
        – different output than head-adjunction
           * reprojection
           * phrasal movement

---

[56] See also Harizanov & Gribanova (accepted) on syntactic head movement and the A-over-A Principle.
[57] This idea goes back to Chomsky (1986) where it is suggested that verb movement can alter a barrier.

- interplay of syntactic movement and a post-syntactic operation (movement to specifier plus rebracketing)
- post-syntax, movement (Raising)
- post-syntax, no movement (DLTs)

Right now, the most crucial discussion in the literature is whether the HM operation is part of narrow syntax or not. This question can be probed by examining if the operation has effects on interpretation or locality, but there is disagreement in the literature about how compelling the arguments for such effects are. We can proceed forward by finding the best analysis for the data in the literature cited in Section 7, a task of future research. If HM can be conclusively shown to obey different constraints and have different properties than syntactic operations, then there is motivation to place it outside syntax. If not, then it should be treated as part of narrow syntax.

Recent research suggests that in the end, we might not be able to give a simple answer to the question of whether HM is a syntactic operation or not because the data that GB-style head-adjunction was meant to capture have heterogenous properties; they should not (and cannot) be accounted for with a single operation. Harizanov (2016); Gribanova (2017b) and Harizanov & Gribanova (accepted) propose that on the one hand, there is genuine syntactic movement of heads, which is characterized by the following constellation of properties: i) it is not subject to the HMC, ii) it targets specifiers, iii) it has semantic effects, iv) it is driven by non-morphological properties of heads (e.g. by discourse properties) and relatedly v) it does not result in morphological word formation. On the other hand, there is also a post-syntactic operation on heads (the Raising discussed in Section 5.3), which has the opposite properties: i) it obeys the HMC, ii) it yields head-adjunction structures, iii) it has no semantic effects, iv) it is driven by morphological properties of heads, and v) it results in "morphological growth" of the head involved (word formation).[58,59] If these properties indeed cluster together this way, then we have an empirically grounded, new and exciting perspective on HM. Checking the strong predictions (e.g. the lack of data for which morphological word formation goes together with semantic effects) of this approach on a large sample of empirical material will be the task of future research.

Abbreviations

3FS = third person feminine subject agreement, 3N = third person neutral agreement, 3S = third singular agreement, 3SS = third singular subject agreement, ACC = accusative, ADJ = adjectival suffix, ASP = aspect, CAUS = casuative, COLL = collective, F = feminine, IND = indicative mood, M = masculine, PL = plural subject agreement, POT = potential suffix, PRE = prefix, PROG = progressive, PST = past, SUF = nominal inflectional suffix

---

[58] There are also other researchers who suggest that more than one type of operation might be needed to cover all the data. As we have seen in Section 5.2, Chomsky advocates PF movement of heads, but he suggests that incorporation should be kept in narrow syntax. Bury (2003) also makes use of both a syntactic and a post-syntactic HM operation: he suggests that reprojection and the DLT-style approach are both necessary. Roberts (2010) is an example that uses different types of syntactic operations (the Agree-based defective goal appraoch as well as reprojection and remnant movement), but he does not exclude the possibility that there are also some cases that involve post-syntactic movement of heads. However, it is Harizanov & Gribanova (accepted) that is the most constrained of these proposals.

[59] If there is indeed such a bifurcation in HM, then showing that some cases of HM have an effect on semantics or locality will not be a knock-down argument against post-syntactic movement and DLT approaches, but it will give them a limited area of application.

Acknowledgements

Competing Interests

The author has no competing interests to declare.

References

Abels, Klaus. 2003. *Successive cyclicity, anti-locality, and adposition stranding*. Storrs, CT: University of Connecticut dissertation.

Abels, Klaus. 2012. *Phases. An essay of cyclicity in syntax* (Linguistische Arbeiten 543). Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110284225

Abels, Klaus & Ad. Neeleman. 2009. Universal 20 without the LCA. In José M. Brucart, Anna Gavarró & Jaume Solà (eds.), *Merging features: Computation, interpretation, and acquisition*, 60–79. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199553266.003.0004

Adger, David. 2013. *A syntax of substance* (Linguistic Inquiry Monograph 64). Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262018616.001.0001

Adger, David, Daniel Harbour & Laurel J. Watkins. 2010. *Mirrors and microparameters: Phrase structure beyond free word order* (Cambridge Studies in Linguistics 122). Cambridge: Cambridge University Press.

Aelbrecht, Loebke & Marcel den Dikken. 2013. Preposition doubling in Flemish and its implications for the syntax of Dutch PPs. *Journal of Comparative Germanic Linguistics* 16(1). 33–68. DOI: https://doi.org/10.1007/s10828-013-9054-2

Arad, Maya. 1996. A minimalist view of the syntax-lexical semantics interface. *UCL Working Papers in Linguistics* 8. 1–30.

Baker, Mark C. 1985. The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry* 16. 373–416.

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing*. Chicago, IL: University of Chicago Press.

Baker, Mark C. 2009. Is head movement still needed for noun incorporation? *Lingua* 119(2). 148–165. DOI: https://doi.org/10.1016/j.lingua.2007.10.010

Benedicto, Elena E. 1998. Verb movement and its effects on determinerless plural subjects. In Armin Schwegler, Bernard Tranel & Myriam Uribe-Etxebarria (eds.), *Romance linguistics: Theoretical perspectives*, 25–40. Amsterdam and Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/cilt.160.04ben

Bentzen, Kristine. 2007. *Order and structure in embedded clauses in Northern Norwegian*. Tromsø: CASTL, University of Tromsø dissertation.

Bhatt, Rajesh & Stefan Keine. 2015. Verb cluster formation and the semantics of head movement. In Ulrike Steindl, Thomas Borer, Huilin Fang, Alfredo Garcia Pardo, Peter Guekguezian, Brian Hsu, Charlie O'Hara & Iris Chuoying Ouyang (eds.), *Proceedings of the 32nd West Coast Conference on Formal Linguistics*, 82–91. Somerville, MA: Cascadilla Proceedings Project.

Biberauer, Theresa & Ian Roberts. 2010. Subjects, tense and verbmovement. In Theresa Biberauer, Anders Holmberg & Michelle Sheenan (eds.), *Parametric variation: Null subjects in minimalist theory*, 263–302. Cambridge: Cambridge University Press.

Bobaljik, Jonathan. 1995. *Morphosyntax: The syntax of verbal inflection*. Boston, MA: MIT dissertation.

Bobaljik, Jonathan David & Samuel Brown. 1997. Interarboreal operations: Head movement and the extension requirement. *Linguistic Inquiry* 28(2). 345–356.

Boeckx, Cedric & Sandra Stjepanović. 2001. Head–ing toward PF. *Linguistic Inquiry* 32(2). 345–355. DOI: https://doi.org/10.1162/00243890152001799

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24(4). 591–656.

Brody, Michael. 1997. Mirror theory. *UCL Working Papers in Linguistics* 9. 1–48.

Brody, Michael. 2000a. Mirror Theory: Syntactic representation in Perfect Syntax. *Linguistic Inquiry* 31(1). 29–56. DOI: https://doi.org/10.1162/002438900554280

Brody, Michael. 2000b. Word order, restructuring and Mirror Theory. In Peter Svenonius (ed.), *The derivation of VO and OV* (Linguistik Aktuell/Linguistics Today 31), 27–44. Amsterdam and Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/la.31.02bro

Brody, Michael. 2004. "Roll-up" structures and morphological words. In Katalin É. Kiss & Henk C. Riemsdijk (eds.), *Verb clusters: A study of Hungarian, German and Dutch*, 147–171. Amsterdam and Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/la.69.09bro

Brody, Michael & Anna Szabolcsi. 2003. Overt scope in Hungarian. *Syntax* 6(1). 19–51. DOI: https://doi.org/10.1111/1467-9612.00055

Bury, Dirk. 2003. *Phrase structure and derived heads*. London: University College London dissertation. http://ling.auf.net/lingbuzz?_s=ca3SQsYc01fgKfn_&_k=HRjxGJCOGuZloPSv.

Bye, Patrik & Peter Svenonius. 2012. Non-concatenative morphology as an epiphenomenon. In Jochen Trommer (ed.), *The morphology and phonology of exponence*, 427–495. Oxford: Oxford University Press.

Caha, Pavel. 2009. *The Nanosyntax of Case*. Tromsø: CASTL, University of Tromsø dissertation.

Chesi, Christiano. 2004. *Phases and cartography in linguistic computation: Toward a cognitively motivated computational model of linguistic competence*. Siena: University of Siena dissertation.

Chesi, Christiano. 2015. On directionality of phrase structure building. *Journal of Psycholinguistic Research* 44(1). 65–89. DOI: https://doi.org/10.1007/s10936-014-9330-6

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

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

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36(3). 315–332. DOI: https://doi.org/10.1162/0024389054396917

Cinque, Guglielmo. 2010. *The syntax of adjectives: A comparative study* (Linguistic Inquiry Monograph 57). Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262014168.001.0001

den Dikken, Marcel. 2007. Phase extension: Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33(1). 1–41. DOI: https://doi.org/10.1515/TL.2007.001

den Dikken, Marcel. 2015. Raising the subject of the "object-of" relation. In Ángel J. Gallego & Dennis Ott (eds.), *MIT Working Papers in Linguistics 77. 50 years later: Reflections on Chomsky's Aspects*, 85–98. Cambridge, MA: MIT Press.

den Dikken, Marcel. 2018. *Dependency and directionality* (Cambridge Studies in Linguistics 154). Cambridge: Cambridge University Press.

Donati, Catherina. 2006. On *wh*-head movement. In Lisa Cheng & Norbert Corver (eds.), *Wh-movement: Moving on*, 21–46. Cambridge, MA: MIT Press.

Embick, David. 2000. Features, syntax, and categories in the Latin perfect. *Linguistic Inquiry* 31(2). 185–230. DOI: https://doi.org/10.1162/002438900554343

Embick, David & Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39(1). 1–53. DOI: https://doi.org/10.1162/ling.2008.39.1.1

Embick, David & Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4). 555–595. DOI: https://doi.org/10.1162/002438901753373005

Embick, David & Rolf Noyer. 2007. Distributed morphology and the syntaxmorphology interface. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 289–324. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199247455.013.0010

Fanselow, Gisbert. 2004. Münchhaussen-style head movement and the analysis of verb second. In Ralf Vogel (ed.), *Linguistics in Potsdam 22: Three papers on German verb movement*, 9–49. Potsdam: University of Potsdam.

Fiengo, Robert. 1977. On trace theory. *Linguistic Inquiry* 8(1). 35–62.

Funakoshi, Kenshi. 2014. *Syntactic head movement and its consequences*. College Park, MD: University of Maryland dissertation.

Gallego, Ángel J. 2010. *Phase theory* (Linguistic Aktuell/Linguistics Today 152). Amsterdam and Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/la.152

Georgi, Doreen & Gereon Müller. 2010. Noun-phrase structure by reprojection. *Syntax* 13(1). 1–36. DOI: https://doi.org/10.1111/j.1467-9612.2009.00132.x

Goldberg, Lotus M. 2005. *Verb-stranding VP ellipsis: A cross-linguistic study*. Montreal: McGill University dissertation.

Gribanova, Vera. 2017a. Head movement and ellipsis in the expression of Russian polarity focus. *Natural Language and Linguistic Theory* 35(4). 1079–1121. DOI: https://doi.org/10.1007/s11049-017-9361-4

Gribanova, Vera. 2017b. Head movement, ellipsis, and identity. Ms., Stanford University.

Grodzinsky, Yosef & Lisa Finkel. 1998. The neurology of empty categories: Aphasics' failure to detect ungrammaticality. *Journal of Cognitive Neuroscience* 10(2). 281–292. DOI: https://doi.org/10.1162/089892998562708

Grohman, Kleanthes K. 2001. On predication, derivation and anti-locality. *ZAS Papers in Linguistics* 26. 87–112.

Grohman, Kleanthes K. 2002. Anti-locality and clause types. *Theoretical Linguistics* 28(1). 43–72. DOI: https://doi.org/10.1515/thli.2002.28.1.43

Grohman, Kleanthes K. 2003a. *Prolific domains: On the anti-locality of movement dependencies*. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/la.66

Grohman, Kleanthes K. 2003b. Successive cyclicity under (anti-)local considerations. *Syntax* 6(3). 260–312. DOI: https://doi.org/10.1111/j.1467-9612.2003.00063.x

Grohman, Kleanthes K. 2011. Anti-locality: Too-close relations in grammar. In Cedric Boeckx (ed.), *The Oxford handbook of linguistic minimalism*, 260–290. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199549368.013.0012

Hale, Ken & Samuel Jay Keyser. 1993. On the argument structure and the lexical expression of syntactic relations. In Ken Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 53–109. Cambridge, MA: MIT Press.

Hale, Ken & Samuel Jay Keyser. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, MA: MIT Press.

Hall, David. 2015. *Spelling out the Noun Phrase: Interpretation, word order, and the problem of "meaningless movement"*. London: Queen Mary, University of London dissertation.

Halle, Moris & Alec Marantz. 1994. Some key features of Distributed Morphology. *MIT Working Papers in Linguistics* 21. 275–288.

Harizanov, Boris. 2014. Clitic doubling at the syntax-morphophonology interface: A-movement and morphological merger in Bulgarian. *Natural Language and Linguistic Theory* 32(4). 1033–1088. DOI: https://doi.org/10.1007/s11049-014-9249-5

Harizanov, Boris. 2016. Head movement to specifier positions in Bulgarian participle fronting. *Handout of a talk presented at the 90th Annual Meeting of the LSA*. January 2016, Washington DC. https://stanford.edu/~bharizan/pdfs/Harizanov_2016_LSA_handout.pdf.

Harizanov, Boris & Vera Gribanova. accepted. Whither head movement? *Natural Language and Linguistic Theory*.

Harley, Heidi. 2004. Merge, conflation and head movement: The First Sister Principle revisited. In Keir Moulton & Matthew Wolf (eds.), *Proceedings of NELS 34* 1. 239–254. Amherst, MA: GLSA. http://dingo.sbs.arizona.edu/~hharley/PDFs/HarleyNELS2003.pdf.

Harley, Heidi. 2013. Getting morphemes in order: Merger, affixation, and head movement. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *Diagnosing syntax* (Oxford Studies in Theoretical Linguistics), 44–74. Oxford: Oxford University Press.

Hartman, Jeremy. 2011. The semantic uniformity of traces: Evidence from ellipsis parallelism. *Linguistic Inquiry* 42(3). 367–388. DOI: https://doi.org/10.1162/LING_a_00050

Haugen, Jason D. & Daniel Siddiqi. 2013. Roots and the derivation. *Linguistic Inquiry* 44(3). 493–517. DOI: https://doi.org/10.1162/LING_a_00136

Henry, Alison. 1995. *Belfast English and standard English. Dialect variation and parameter setting*. Oxford and New York: Oxford University Press.

Hickey, Raymond. 2007. *Irish English*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511551048

Iatridou, Sabine & Hedde Zeijlstra. 2013. Negation, polarity, and deontic modals. *Linguistic Inquiry* 44(4). 529–568. DOI: https://doi.org/10.1162/LING_a_00138

Iorio, David Edy. 2015. *Subject and object marking in Bembe*. Newcastle: Newcastle University dissertation.

Julien, Marit. 2002. *Syntactic heads and word formation* (Oxford Studies in Comparative Syntax). New York: Oxford University Press.

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

Kayne, Richard & Jan-Yves Pollock. 2001. New thoughts on stylistic inversion. In Aafke Hulk & Jan-Yves Pollock (eds.), *Subject inversion in Romance and the theory of Universal Grammar*, 107–162. Oxford: Oxford University Press.

Keine, Stefan & Rajesh Bhatt. 2016. Interpreting verb clusters. *Natural Language and Linguistic Theory* 34(4). 1445–1492. DOI: https://doi.org/10.1007/s11049-015-9326-4

Koeneman, Olaf Nicolaas Cornelis Johannes. 2000. *The flexible nature of verb movement*. Utrecht: University of Utrecht dissertation.

Koopman, Hilda. 1984. *The syntax of verbs: From verb-movement rules in the Kru languages to Universal Grammar*. Dordrecht: Foris.

Koopman, Hilda Judith & Anna Szabolcsi. 2000. *Verbal complexes* (Current Studies in Linguistics 34). Cambridge, MA: MIT Press.

Landau, Idan. 2006. Chain resoltuion in Hebrew V(P)-fronting. *Syntax* 9(1). 32–66. DOI: https://doi.org/10.1111/j.1467-9612.2006.00084.x

Lapointe, Steven. 1980. *A theory of grammatical agreement*. Boston, MA: MIT dissertation.

Lasnik, Howard. 1999. On feature strength: Three minimalist approaches to overt movement. *Linguistic Inquiry* 30(2). 197–217. DOI: https://doi.org/10.1162/002438999554039

Lasnik, Howard & Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15(2). 235–289.

Lasnik, Howard & Mamoru Saito. 1991. On the subject of infinitives. In Lise Dobrin, Lynn Nichols & Rosa Rodriguez (eds.), *Papers from the twentyseventh regional meeting of the Chicago Linguistic Society*, 324–343. Chicago: Chicago Linguistic Society.

Lechner, Winfried. 2006. An interpretive effect of head movement. In Mara Frascarelli (ed.), *Phases of interpretation*, 45–71. Berlin and New York: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110197723.2.45

Lechner, Winfried. 2007. Interpretive effects of head movement. Ms., University of Tübingen, version 2. http://ling.auf.net/lingBuzz/000178.

Lipták, Anikó. 2017. Identifying identity conditions in and outside ellipsis: A case study of Hungarian. *Talk delivered at BLINC 2*. Budapest, June 2017.

Livitz, Inna. 2011. Incorporating PRO: A defective-goal analysis. In Neil Myler & Jim Wood (eds.), *NYU Working Papers in Linguistics 3*, 95–119. New York, NY: New York University.

Mahajan, Anoop. 2003. Word order and (remnant) VP movement. In Simin Karimi (ed.), *Word order and scrambling* (Explaining Linguistics 4), 217–237. Malden, MA: Blackwell.

Marantz, Alec. 1995. A late note on Late Insertion. In Young-Sun Kim, Byung-Choon Lee, Kyoung-Jae Lee, Kyun-Kwon Yang & Jong-Kuri Yoon (eds.), *Explorations in generative grammar*, 396–413. Seoul: Hankuk.

Massam, Diane. 2000. VSO and VOS: Aspects of Niuean word order. In Andrew Carnie & Eithne Guilfoyle (eds.), *The syntax of verb-initial languages*, 97–116. Oxford: Oxford University Press.

Mathew, Rosmin. 2015. *Head movement in sytax*. John Benjamins. DOI: https://doi.org/10.1075/la.224

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37(1). 69–109. DOI: https://doi.org/10.1162/002438906775321184

Matushansky, Ora. 2011. Ian Roberts, Agreement and head movement: Clitics, incorporation, and defective goals (Linguistic Inquiry Monograph 59). Cambridge, MA: MIT Press, 2010. Pp. x + 290. *Journal of Linguistics* 47(2). 538–545. DOI: https://doi.org/10.1017/S0022226711000120

Matyiku, Sabina. 2014. Semantic effects of head movement in negative auxiliary inversion constructions. *Handout of a talk delivered at WCCFL* 32. March 4, 2014. http://campus-press.yale.edu/sabina/files/2015/05/Matyiku-2014-WCCFL32-2fvz1xi.pdf.

McCloskey, James. 1984. Raising, subcategorization and selection in modern Irish. *Natural Language and Linguistic Theory* 1(4). 441–485.

McCloskey, Jim. 2016. Interpretation and the typology of head movement: A re-assessment. *Handout of a talk presented at the Workshop on the Status of Head Movement in Linguistic Theory*. http://ohlone.ucsc.edu/~jim/papers.html.

McCloskey, Jim. 2017. Ellipsis, polarity, and the cartography of verb-initial orders in Irish. In Enoch Aboh, Eric Haeberli, Manuela Schönenberger & Genoveva Puskás (eds.), *Elements of comparative syntax: Theory and description* (Studies in Generative Grammar 127), 99–151. Berlin: De Gruyter. DOI: https://doi.org/10.1515/9781501504037-005

Müller, Gereon. 2004. Verb-second as vP-first. *Journal of Comparative Germanic Linguistics* 7(3). 179–234. DOI: https://doi.org/10.1023/B:JCOM.0000016453.71478.3a

Muriungi, Peter Kinyua. 2008. *Phrasal movement inside Bantu verbs. Deriving affix scope and order in Kîîtharaka*. Tromsø: CASTL, University of Tromsø dissertation.

Nilsen, Øystein. 2003. *Eliminating positions*. Utrecht: Utrecht University dissertation.

Nunes, Jairo. 1995. *The copy theory of movement and linearization of chains in the Minimalist Program*. College Park, MD: University of Maryland dissertation.

Nunes, Jairo. 2001. Sideward movement. *Linguistic Inquiry* 32(2). 303–344. DOI: https://doi.org/10.1162/00243890152001780

Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Cambridge, MA: MIT Press.

Parrott, Jeffrey. 2001. A morphological theory of head movement. Ms., Georgetown University, Washington D.C.

Pesetsky, David. 2013. *Russian case morphology and the syntactic categories* (Linguistic Inquiry Monograph 66). Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262019729.001.0001

Phillips, Colin. 1996. *Linear order and constituency*. Boston, MA: MIT dissertation.

Phillips, Colin. 2003. Linear order and constituency. *Linguistic Inquiry* 34(1). 37–90. DOI: https://doi.org/10.1162/002438903763255922

Platzack, Christer. 2013. Head movement as a phonological operation. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *Diagnosing syntax* (Oxford Studies in Theoretical Linguistics), 21–43. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199602490.003.0002

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.

Pollock, Jean-Yves. 2006. Subject clitics and complex inversion. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax* IV. 601–659. Oxford: Blackwell. DOI: https://doi.org/10.1002/9780470996591.ch67

Postal, Paul. 1974. *On raising: An inquiry into one rule of English grammar and its theoretical implications* (Current Studies in Linguistics 5). Cambridge, MA: MIT Press.

Postal, Paul. 2004. *Skeptical linguistic essays*. New York: Oxford University Press.

Preminger, Omer. 2017. What the PCC tells us about "abstract" agreement, head movement, and locality. Ms., University of Maryland. http://ling.auf.net/lingbuzz/003221.

Rackowski, Andrea & Lisa Travis. 2000. V-initial languages: X or XP movement and adverbial placement. In Andrew Carnie & Eithne Guilfoyle (eds.), *The syntax of verb-initial languages*, 117–142. Oxford: Oxford University Press.

Ramchand, Gillian. 2014. Deriving variable linearization: A commentary on Simpson and Syed (2013). *Natural Language and Linguistic Theory* 32(1). 263–282. DOI: https://doi.org/10.1007/s11049-013-9225-5

Richards, Marc. 2011. Deriving the edge: What's in a phase? *Syntax* 14(1). 74–95. DOI: https://doi.org/10.1111/j.1467-9612.2010.00146.x

Richards, Norvin. 1998. The Principle of Minimal Compliance. *Linguistic Inquiry* 29(4). 599–629. DOI: https://doi.org/10.1162/002438998553897

Richards, Norvin. 2001. *Movement in language: Interactions and architectures*. New York: Oxford University Press.

Ritter, Elizabeth. 1988. A head-movement approach to construct-state noun phrases. *Linguistics* 26(6). 909–929. DOI: https://doi.org/10.1515/ling.1988.26.6.909

Rivero, Maria-Luisa. 1991. Long head movement and negation: Serbo-Croatian vs. Slovak. *The Linguistic Review* 8(2–4). 319–351. DOI: https://doi.org/10.1515/tlir.1991.8.2-4.319

Rivero, Maria-Luisa. 1993. Long head movement vs. V2 and null subjects in Old Romance. *Lingua* 89(2–3). 113–141. DOI: https://doi.org/10.1016/0024-3841(93)90053-Y

Rivero, Maria-Luisa & Arhonto Terzi. 2005. Imperatives, V-movement and logical mood. *Journal of Linguistics* 31(2). 301–332. DOI: https://doi.org/10.1017/S0022226700015620

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.

Roberts, Ian. 1991. Excorporation and minimality. *Linguistic Inquiry* 22(1). 209–218.

Roberts, Ian. 2001. Head movement. In Mark Baltin & Chris Collins (eds.), *Handbook of syntactic theory*, 112–147. Oxford: Blackwell. DOI: https://doi.org/10.1002/9780470756416.ch5

Roberts, Ian. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals* (Linguistic Inquiry Monograph 59). Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262014304.001.0001

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

Runner, Jeffrey. 1998. *Noun Phrase licensing and interpretation*. New York: Garland.

Schoorlemmer, Erik & Tanja Temmerman. 2012. Head movement as a PFphenomenon: Evidence from identity under ellipsis. In Jaehoon Choi, E. Alan Hogue, Jeffrey Punske, Deniz Tat, Jessamyn Schertz & Alex Trueman (eds.), *Proceedings of the 29th West Coast Conference on Formal Linguistics*, 232–240. Somerville, MA: Cascadilla Proceedings Project.

Shlonsky, Ur. 2004. The form of Semitic noun phrases. *Lingua* 114(12). 1465–1526. DOI: https://doi.org/10.1016/j.lingua.2003.09.019

Skinner, Tobin R. 2009. *Investigations of downward movement*. Montreal: McGill University dissertation.

Sportiche, Dominique. 1997. Reconstruction and constituent structure. *Handout of a talk presented at MIT*. October 1997. http://www.linguistics.ucla.edu/people/sportich/papers/mittalk97.pdf.

Sportiche, Dominique. 2005. Division of labor between merge and move: Strict locality of selection and apparent reconstruction paradoxes. In Nathan Klinedinst & Greg Kobele (eds.), *Proceedings of the Workshop Divisions of Linguistic Labor: The la Bretesche workshop*, 159–262. lingbuzz/000163.

Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd: Tromsø University Working Papers in Linguistics* 36(1). 1–6. DOI: https://doi.org/10.7557/12.213

Starke, Michal. 2014. Towards elegant parameters: Language variation reduces to the size of lexicaly-stored trees. In Carme M. Picallo (ed.), *Linguistic variation in the Minimalist Framework* (Oxford Linguistics), 140–152. New York: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780198702894.003.0007

Stepanov, Arthur. 2012. Voiding island effects via Head Movement. *Linguistic Inquiry* 43(4). 680–693. DOI: https://doi.org/10.1162/ling_a_00111

Surányi, Balázs. 2005. Head movement and reprojection. In *Annales Universitatis Scientiarum Budapestenesis de Rolando Eötvös Nominatae. Sectio linguistica. Tomus XXIV*, 313–342. Budapest: ELTE.

Surányi, Balázs. 2008. The theory of head movement and cyclic spell out. In Jutta Hartmann, Veronika Hegedűs & Henk van Riemsdijk (eds.), *Sounds of silence: Empty elements in syntax and phonology*, 289–332. Amsterdam: Elsevier.

Szabolcsi, Anna. 2010. *Quantification* (Research Surveys in Linguistics). New York: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511781681

Szabolcsi, Anna. 2011. Certain verbs are syntactically explicit quantifiers. In Barbara H. Partee, Michael Glanzberg & Jurģis Šķilters (eds.), *Formal semantics and pragmatics. Discourse, context and models. The Baltic international yearbook of cognition, logic and communication* 6. 1–26. Manhattan, KS: New Prairie Press. DOI: https://doi.org/10.4148/biyclc.v6i0.1565

Toyoshima, Takashi. 2001. Head-to-Spec movement. In Galina M. Alexandrova & Olga Arnaudova (eds.), *The minimalist parameter. Selected papers from the Open Linguistics Forum, Ottawa, 12–23 March 1997*, 115–136. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/cilt.192.10toy

Travis, Lisa. 1984. *Parameters and effects of word order variation*. Boston, MA: MIT dissertation.

Uriagereka, Juan. 1998. *Rhyme and reason: An introduction to Minimalist syntax*. Cambridge, MA and London: MIT Press.

van Urk, Coppe. 2015. *A uniform syntax for phrasal movement: A case study in Dinka Bor*. Boston, MA: MIT dissertation.

Vicente, Luis. 2007. *The syntax of heads and phrases: A study of verb (phrase) fronting*. Leiden: Leiden University dissertation. http://www.lotpublications.nl/publish/articles/002244/bookpart.pdf.

Walkden, George. 2014. Language contact and the loss of strict V2. *Paper presented at the 16th International Diachronic Generative Syntax Conference*. Budapest, July 2014.

Zwart, Jan-Wouter. 2001. Syntactic and phonological verb movement. *Syntax* 4(1). 34–62. DOI: https://doi.org/10.1111/1467-9612.00036