

Head movement

Michael Barrie and Éric Mathieu

1 Introduction

This chapter discusses head movement (HM) as a distinct syntactic operation, as well as the empirical facts argued to be covered by such. We start with a brief history of the development of HM in the Government and Binding era and then go on to discuss how HM was enriched through to the beginning of Minimalism. It was at this point that HM began to be seriously questioned. We discuss the problems that HM raises for Bare Phrase Structure (BPS) and the solutions that have been proposed in the literature. Then we run through the current status of HM and some of its technical aspects. Finally, we discuss how HM is affected in patients with aphasia. We restrict ourselves principally to a discussion of HM within the Principles and Parameters framework (notably Minimalism and Government & Binding Theory, though we do not focus on the mechanics of the latter). HM, as a derivational process, does not play a role in representational theories such as Head-Driven Phrase Structure Grammar (HPSG) or Role and Reference Grammar (RRG), so we do not discuss these here. See Kiss and Wesche (1991), however, for a discussion of how to treat verb movement in an HSPG and Combinatory Categorical Grammar (CCG) framework.

The remainder of this chapter is structured as follows. Section 2 discusses the early inceptions of HM as it arose from Standard Theory and its successors. Section 3 discusses how HM changed in the wake of discussions on the Lexicalist Hypothesis and its role in these discussions. Section 4 presents HM in light of Minimalism and BPS. Specifically, it discusses problematic aspects of HM and how these were addressed. Section 5 presents the current status of HM, in particular, we highlight the lack of consensus of HM in current syntactic theory. Section 6 presents some current research on the comprehension of HM in aphasic individuals and briefly discusses how this is related to our current theoretical understanding of HM. Section 7 is a brief summary.

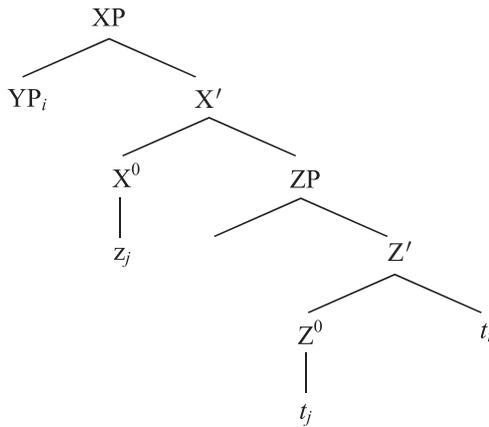
2 The birth of head movement

HM as a distinct operation was made explicit by Government and Binding Theory (Chomsky 1981). Previously movement operations in Revised Extended Standard Theory

and its predecessors were accomplished by language-specific operations targeting strings of specified lengths without regard as to whether the element moved was a single item or a whole phrase. Specifically, it was the implementation of X-Bar Theory that led to the distinction between HM and XP-movement (Chomsky 1970; Jackendoff 1977).

In X-Bar Theory, HM is accomplished by the terminal from one head detaching and raising to the terminal of the immediately c-commanding head. Unlike XP movement, which can target a c-commanding landing site at a distance, HM is constrained in this extremely local fashion under what came to be known as the Head Movement Constraint (HMC) (Travis 1984).

(1)



The HMC can be illustrated by the following pair of examples. In order to license a polarity question in English, an auxiliary or modal must raise to C. The following data show that only the higher head can raise (2a). It is not possible for the lower head to raise, skipping the intermediate head (2b).

- (2) a. Will John *t* have eaten?
 b. *Have John will *t* eaten?

Head movement as a distinct operation was given considerable cross-linguistic thrust by Koopman's (1984) landmark study on verb movement in the Kru languages of Africa, underscoring the universality of this operation.

The operation of HM served as a diagnostic for syntactic structure, especially in the derivation of V2 order in German and verb-initial order in Irish. V2 order in German is found only in matrix clauses. Embedded clauses typically always contain a complementizer and the verb appears at the right edge of the clause. Recall that HM operates mechanically on the terminals under the head nodes. The terminal raises to an immediately c-commanding empty head node. If this head node already contains a terminal, then head movement is blocked. Consider the following German data.

- (3) a. Er hat den Apfel gegessen.
 he has the apple eaten
 'He has eaten the apple.'

- b. ...dass er den Apfel gegessen hat.
 ...that he the apple eaten has
 ‘...that he ate the apple.’

Example (3a) illustrates the phenomenon of V2 in Germanic. The highest tensed head appears in “second” position – that is, immediately after the first constituent. This property of the highest tensed head appearing in this position is argued to arise by the contents of T^0 raising to C^0 (Thiersch 1978; den Besten 1977; 1983). The evidence for this analysis resides in the failure of V2 to hold in embedded clauses when an overt complementizer is present (3b). This kind of diagnostic was used repeatedly in the early to mid GB era to argue for particular clausal structures of a number of languages. The axiomatic foundation for this line of argumentation resides strongly in the notion that a head can undergo HM only to the next highest head, and only if that head is empty. We will revisit this second point later on in §3.

A major alteration to the theory of HM was made by a series of papers introducing Long Head Movement (LHM) (Lema and Rivero 1990a; 1990b; Rivero 1994; 1993). LHM appears to involve violations of the HMC presented above. Consider the following example.

- (4) Ver -te -ei [Portuguese]
 see -you -3.SG.FUT
 ‘I will see you.’

The future and conditional forms in Portuguese (as in Romance languages in general) is formed with the infinitive as a base to which is attached various agreement markers based on the verb ‘have’. In the future and conditional forms of modern literary Portuguese, however, the infinitival form of the verb is interrupted by clitics, as shown in (4). This was a pervasive feature of several older varieties of Romance languages but has since disappeared except in Portuguese, where it is found only in very formal speech today. Since the verbal base and the agreement marker are separated by a clitic, we have evidence that this structure is clearly formed in the syntax. The verbal base, however, originates low in the structure, below agreement. Thus, we have the following rough derivation.

- (5) [_{CP} ... [_{C^o} ver_i] [_{TP} [_{T^o} ei] [_{VP} [_V t_i]]]]

What’s noteworthy here, of course, is that the verb has raised to T^0 , across an intervening head in an apparent violation of the HMC. As mentioned, movement of this type was commonplace in older varieties of Romance, but is also found in some Slavic languages. Rivero’s original solution to this problem is as follows. She argues that the verb forms an agreement chain with T^0 as well as a movement chain with C^0 . Thus, there is a set of chains linking these three heads. This view requires us to understand the HMC as a condition on representations rather than as a condition on derivations.

To conclude this section, we have introduced the first discussions of HM in GB Theory, firmly couched within an X-Bar theoretic framework. HM proceeds by a head targeting an immediately c-commanding vacant head position. This extremely local movement is captured under the HMC. Finally, we saw that violations of the HMC are found in so-called LHM in Romance and Slavic languages. The next section discusses some further advances in the mechanism of HM at the dawn of early Minimalism as GB Theory was declining.

3 Expanding the role of head movement

A major shift in the role of HM in generative grammar came in the 1980s, when HM started to play a larger role in word formation. To appreciate the issues discussed in this section we need to understand the Lexicalist Hypothesis (Di Sciullo and Williams 1987). This hypothesis holds that the atoms of syntax are words, which possibly possess internal morphological complexity. This internal structure is not visible to the syntax. Crucially, this hypothesis posits a pre-syntactic word-formation module. Theories such as Distributed Morphology reject the Lexicalist Hypothesis by claiming that there is no such module (Marantz 1997). Rather, words are put together in the syntax (or additionally, in the case of DM, in a post-syntactic morphological module). The notion that words can be put together in the syntax is not new (Lees 1960; Chomsky 1970); however, it was the pioneering works of Baker (1985; 1988) and Pollock (1989) that set the stage for this development in HM.

The concept of building words by HM in the syntax was spurred primarily by the Mirror Principle of Baker (1985). Baker's observation was that the order of affixes in a word mirrors the order of functional projections. Thus, if the order of the affixes in a verbal complex is Verb-X-Y-Z, then the order of functional projections is ZP > YP > XP > VP. Consider the following Bembe (Niger Congo) example.

- (6) a. Naa- mon -an -ya [Mwape na Mutumba]
 1SGS- see -RECIP -CAUS Mwape and Mutumba
 'I made Mwape and Mutumba see each other.'
- b. [Mwape na Chilufya] baa- mon -eshy -ana Mutumba
 [Mwape and Chilufya] 3PS- see -CAUS -RECIP Mutumba.
 'Mwape and Chilufya made each other see Mutumba.'

In (6a) the causative suffix appears to the right of the reciprocal suffix. Thus, the causative scopes over the reciprocal marker, as indicated in the translation. In (6b), on the other hand, the reciprocal suffix appears to the right of the causative suffix. Here, the reciprocal takes scope outside the causative, giving rise to the reading indicated. This correlation between affix order and the functional hierarchy has been replicated in subsequent works (Julien 2002; Cinque 1999). The explanation for this phenomenon is simple. HM, restricted by the HMC, goes up the tree head by head, picking up affixes in order on the way. The proposal given above offers an attractive account of the strong correlation between the order of suffixes in the verbal complex and the order of functional projections; however, it is difficult to extend to prefixes. Harley (2011) proposes simply that affixes vary cross-linguistically as to whether they attach as prefixes or as suffixes (see also Wojdak 2008). See the Further Reading section for more comments on this point.

Pollock's (1989) discussion of verb movement in English and French offers a further illustration of word formation by HM. Pollock was chiefly concerned with verb movement and the structure of Infl, which he split into a tense phrase (TP) and an agreement phrase (AgrP). His proposal for verb movement accounts for the difference in word order in the following English and French sentences.

- (7) a. John often kisses Mary.
 b. *John kisses often Mary.

- (8) a. Jean embrasse souvent Marie.
 John kisses often Mary
 ‘John often kisses Mary.’
 b. *Jean souvent embrasse Marie.
 Jean often kisses Mary
 (‘John often kisses Mary.’)

Assuming that the adverb *often/souvent* consistently adjoins to VP, Pollock argued that the difference in word order can be captured by assuming that French has V-to-T movement, but English does not, as shown in the following bracketed illustrations.

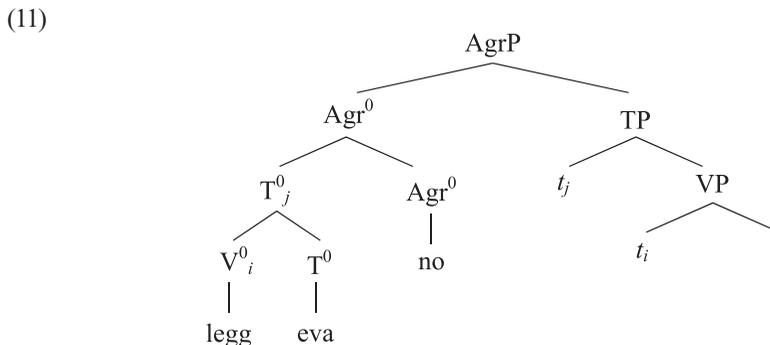
- (9) a. [CP [TP John T [VP often [VP kisses Mary]]]]. [English]
 b. [CP [TP Jean [T embrasse]_i [VP souvent [VP t_i Marie]]]]. [French]

Additional evidence for V-to-T raising in French and its absence in English comes from negation and yes/no question formation. We do not present these here for reasons of space. These diagnostics have become standard in syntactic theorizing. See, for instance, Adger’s (2003) textbook for a clear discussion and illustration of these diagnostics in various languages.

Based on the order of the morphemes in Romance verbs, Belletti (1990) proposed that the order of the functional heads is AgrP > TP. Consider the following Italian verb form.

- (10) legg-eva-no
 see-IMP-3.PL
 ‘They were reading.’

Note that the tense/aspect morphology is closer to the root than the agreement marking. The proposed structure is as follows.



Here, the verb raises to T⁰, picking up the tense/aspect affix. This complex head undergoes further HM to Agr⁰, picking up the agreement affix. Successive HM in this manner creates a larger and larger complex head.

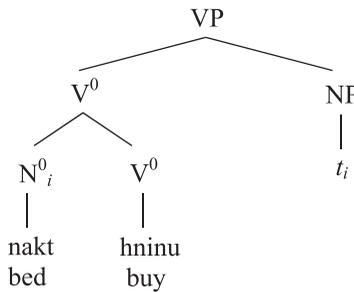
Let us now turn to Baker (1988), who develops a comprehensive theory of incorporation involving word formation via HM. Baker considered noun incorporation, causativization, applicativization, and restructuring. We will demonstrate the case of noun incorporation

here, as this phenomenon plays an important role in later discussions of HM (Barrie and Mathieu 2012; Roberts 2010; Baker 2009). Consider the following example.

- (12) Wa'- ke- nákt- a- hnínu -'
 FACT- 1sS- bed- EPEN- buy -PUNC
 'I bought the/a bed.'

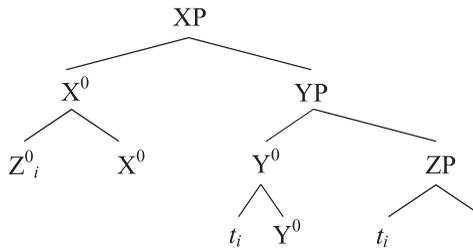
Here, the root *nakt* ('bed') appears inside the verbal complex. Note that the epenthetic vowel (EPEN) is not a morpheme, but rather is added phonologically to break up an illicit consonant cluster. Baker's proposal for a syntactic treatment of noun incorporation is quite simple. The head noun of the NP undergoes HM to V^0 .

- (13)



We end this section with a discussion on excorporation, illustrated in (14). Here, a head, Z^0 , has raised to Y^0 , and has then detached off, or excorporated, and raised to X^0 .

- (14)



First, it is important to note that XP movement allows a phrase to raise to a given position, and then to continue to raise, such as with long-distance *wh*-movement. An obvious question comes to mind: Is there an equivalent operation for HM? To answer this question, we need to understand two structure building operations from GB theory: adjunction and substitution. Adjunction is still a familiar operation today. Substitution referred to movement to a specifier position, but the term is no longer in general use. It had always been unclear whether HM proceeds by adjunction or by substitution. Roberts (1991) proposes that HM can proceed either by substitution or by adjunction. Substitution involves the satisfaction of some property of a head by the element that moves to its specifier. Likewise, Roberts proposes that HM that satisfies some property of the head targeted by HM is a kind of substitution. A typical case is subcategorization: for example, an I^0 head subcategorizes for a V^0 . Thus, V-to-I movement is an instance of substitution. However, if HM does not involve selection of any kind, then such HM proceeds by adjunction. Working under a Barriers

framework, Roberts showed that a head to which substitution HM has taken place acts as a barrier, while a head to which adjunction HM has taken place does not. Thus, excorporation can happen only with adjunction HM. The cases Roberts discusses involve clitic climbing and V2 effects in Dutch. We illustrate the first here with the following Italian example.

- (15) La_i volevo t_i chiamare t_i ieri
 her I-wanted to-call yesterday
 ‘Yesterday, I wanted to call her up.’

Roberts assumes that clitic movement is an HM operation. It is not selected, however, so it proceeds by adjunction. In non-finite clauses, the clitic can raise up to the higher, tensed verb. Roberts argues that this is excorporation, as illustrated above, made possible by virtue of the mechanism of adjunction HM.

To sum up, we have seen that HM firmly implanted itself as a fundamental operation vital to the role of syntax during the GB era. In particular, HM was shown to underpin word formation in a variety of situations, such as noun incorporation, causativization, and the formation of morphologically complex verb forms in Romance. Of course, many other empirical illustrations were discussed during this time and continue to be discussed in relation to the role of HM in word formation. Finally, we illustrated the possibility of excorporation with clitic climbing.

4 Head movement and the minimalist program

As GB Theory and X-Bar Theory were becoming fraught with more and more empirical and theory internal problems, Chomsky started developing what became known as the Minimalist Program (Chomsky 1993) and Bare Phrase Structure (BPS) (Chomsky 1994). BPS consists of a single operation, Merge, which takes two syntactic objects, α and β , and forms a single syntactic object as follows:

- (16) Merge (α, β) \rightarrow $\{\gamma, \{\alpha, \beta\}\}$

The resulting structure is an XP with a label, γ .

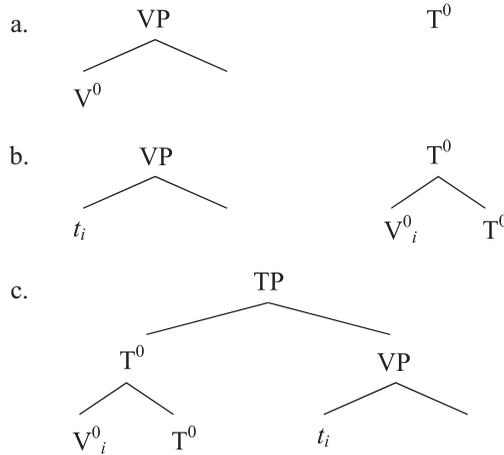
Several problems with HM came to light in the early 2000s (Mahajan 2003; Fanselow 2003). As noted by several authors, HM is unformulable in BPS (there were, in fact, also numerous problems within the X-bar theory/GB paradigm, so much so that Rizzi (1990: 117 n. 19), for example, decides that head movement must be substitution). The single operation, Merge, can only form an XP. Unless modifications or enrichments are made to BPS there is no set-theoretic way to distinguish a complex head from an XP. Another problem is that HM violates the Extension Condition. That is, HM does not target the root projection, but rather targets an embedded node. Finally, depending on how c-command is defined, a moved head does not c-command its trace. Several solutions have been advanced to some or all of these problems, and we take these up here. First, we examine those proposals in which HM is slightly altered to address these issues. Then we examine those proposals which radically alter HM by making it a PF operation or replacing it with XP movement.

4.1 Adjustments to head movement

An early solution to the problem of the Extension Condition violations (and that pre-dates the commonly cited papers that discuss the problems with HM) is that of Bobaljik and

Brown (1997). Bobaljik and Brown exploit the concept of parallel derivations taking place in the same workspace and sideward movement. Here is how they derive V-to-T movement (we ignore *v* here for simplicity). Once the VP has been formed, T^0 is selected and placed into the workspace (17a). V^0 is copied and merged to T^0 . At this point, T^0 is a root node because it is not yet attached to VP (17b). This complex T^0 head then merges with the VP and projects a TP (17c). Notice that at no point is the Extension Condition violated.

(17)



Although this proposal solves the problem of the violation of the Extension Condition, it does not solve the problem of c-command, nor does it address the fact that HM is incompatible with BPS.

With the shackles of X-Bar Theory gone, researchers were free to explore alternative conceptions of movement. Toyoshima (2000) develops a model of Head to Spec movement to account for a range of phenomena related to Economy. It is exactly this kind of movement that solves many of the problems related to HM discussed above (in BPS nothing bars head movement to a specifier position). Matushansky (2006) develops just such an analysis. She acknowledges the problems of HM discussed above, in particular the acute problem of the counter-cyclic aspect of HM. She invokes the mechanism of morphological merger in addition to Head-to-Spec movement. Morphological merger is an operation proposed under the framework of DM. It is a post-syntactic operation which locally dislocates a syntactic object with another string-adjacent syntactic object. Thus, as far as syntax is concerned, the Extension Condition and the requirement that a moved element c-command its trace are satisfied. Also, since the head moves to the specifier position, Matushansky’s conception of HM is compatible with BPS.

Finally, we mention “reprojection” analyses that also propose to keep head movement in the syntax (Bury 2003; 2007; Donati 2006; Koeneman 2000; Surányi 2005; 2007). The idea is to treat head movement as arising from a different set of conditions from XP-movement. Whereas Chomsky (1995: 256–60) argues that where a new category γ is formed by movement of α to β , γ must always project the target of movement, leading to a situation where, for example, DP-movement attracted by T will always create a new projection of T or *wh*-movement attracted by C will always create a new projection of C, the alternative is to suggest that this may not always be the case, and that “reprojective” movement may arise. In “reprojection” the moving category gives its label to the new category formed by

movement. This analysis has been applied to free relatives and related constructions as well as to V2 phenomena.

4.2 Removing Head Movement from Syntax

Boeckx and Stjepanovic (2001) argue from the analysis of pseudo-gapping that head movement is best viewed as a PF phenomenon. Their evidence is based on Lasnik's (1995) analysis of pseudo-gapping, exemplified in (18) where the object *emagazines* escapes the VP and undergoes object shift to the Spec of an Agr projection so that the rest of the VP can be deleted (namely, the verb *buy*).

(18) Although I wouldn't buy ebooks, I would __emagazines ~~buy~~.

The question that immediately arises is why the verb does not raise in pseudo-gapping constructions, given that in non-elliptical sentences it must, as shown in the following example.

(19) *John will emagazines_i buy *t_i*. (vs. John will buy_i emagazines_j *t_i t_j*)

Lasnik observes that if V fails to raise, and no relevant process takes place, the strong feature that is not checked overtly causes a crash at PF. But if the VP containing V is deleted in the PF component, then, patently, the strong feature cannot cause a PF crash, since the category containing the feature will be gone at that level. However, in Lasnik's system nothing stops the verb to raise and the object to stay put, as in (20). Since this sentence is clearly ungrammatical, the conclusion is that head movement is optional, but XP movement is not.

(20) *John bought ebooks, and Mary bought ~~ebooks~~ too.

Boeckx and Stjepanović believe that a solution to both problems – stipulating obligatory/optional movement, and the syntactic-feature-triggered head movement – is to assume that head movement (in this case, V-movement) takes place after Spell-Out, in the phonological component. They assume that XP-movement, like object shift, is syntactic (driven by the checking of some feature): it necessarily takes place in the syntax. They further argue that if head movement is a PF (post-Spell-Out) phenomenon, it necessarily follows all syntactic movement operations and could be “superseded” by an ellipsis rule: not being syntactically driven, head movement and ellipsis (both PF operations) compete.

In other words, if head movement is a PF phenomenon, we have an answer to Lasnik's puzzle as to why V need not raise in the pseudogapping cases. Suppose ellipsis is a PF operation. Then head movement and ellipsis become competing operations: V either moves or is deleted. In short, not only does the view of head movement as a PF operation remove any look-ahead and extrinsic ordering from the computational system, it also avoids the question of which head has the inadequacy that forces movement.

The advantage that this analysis and other similar ones have is that we can now understand why head movement has no effect on interpretation as argued by Chomsky (2000; 2001). Whether the verb moves to T, as in French, or whether it does not, as in English, makes no difference for the semantics, since the relevant sentences have the same truth conditions. The idea that head movement is a PF phenomenon explains this kind of semantic weight: PF movement, unlike syntactic movement (i.e., XP movement), has no bearing on interpretation.

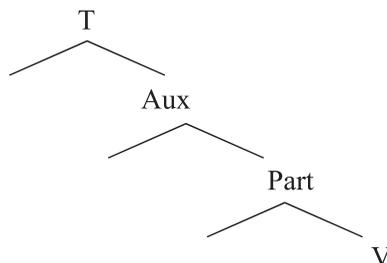
However, it must be noted that the idea that HM has no effect on interpretation is not uncontroversial. A few researchers have attempted to demonstrate that HM is necessarily a syntactic operation needed to account for changes in interpretations cross-linguistically (Benedicto 1998; Lechner 2006; 2007; Kishimoto 2007; 2008). According to Matushansky (2006), the reason why head movement has no LF effect is due to an independent factor (for her there is a syntactic side to head movement, in that the head moves to a specifier in the syntax): it lies in the very nature of the items moved by head movement. Most of them are predicates of some sort: verbs, nouns, most affixes, and so on.

While many claim head movement is a PF phenomenon, often the specifics of the PF operation are not spelled out. Harley (2004) is an exception. She proposes that the Conflation mechanism proposed in Hale and Keyser (2002) is a good candidate for the mechanism behind head movement *à la* PF. Conflation is a concomitant of Merge. It is an operation that occurs when a head X merges with a maximal constituent YP whose label is H(Y). The label of Y, by assumption, contains all the features of Y^o, including a copy of its phonological features that Hale and Keyser call a *p-signature* (p-sig for short). If X's p-sig is defective, Y's p-sig is conflated into X's when X and Y Merge, meaning that X is now pronounced with Y's phonological features. When X projects, the label of the whole constituent, H(X), will now contain Y's phonological features. This can be applied to noun incorporation, for example, but also, as argued by Harley, to V2 cases and other such head movement phenomena. As pointed out by Harley (2004), the conflation mechanism allows us implement the idea that head-movement is "phonological" while simultaneously ensuring that it is local.

A more radical proposal is that of Brody (1997; 2000) and his Mirror Theory. In allowing the projection of complementation structure directly from the lexical information encapsulated in the structure of words, Brody's Mirror Theory makes it possible to abandon head movement completely. While in traditional syntax the information pertaining to the structure of words is expressed both word-internally (i.e., X⁰-internally) and by the (inverse) structure of complementation (the Mirror Principle: Baker 1988), Brody suggests that the information carried by words can be directly projected in the syntax without the recourse of any additional matching/movement. This avoids redundancy in the system. Traditionally, redundancies are frowned upon, and Brody takes the example of the earlier coexistence of both Phrase Structure Rules (PSR) and the Projection Principle that led to a reappraisal and the abandonment of PSR.

In Brody's view, lexical items have projection lines (PLs) and words correspond to these projection lines. They are created pre-syntactically (either in the Lexicon or prior to the point of insertion). To illustrate, a word of the form V-Ptcpl-Aux-T will have the structure in (21). All the information that has to do with the spell out of the word as the mirror image of that sequence is contained in the lowest element: that is, V.

(21)



In this view, objects – that is, complements of verbs – are necessarily specifiers (like subjects). They never occur on the right side of the head: they are always projected on the left side.

This theory has a clear advantage over other theories in that it makes no distinction between words and phrases: both are created in the syntax. At the same time, it is a lexicalist theory, in that all the information about words (including morphophonological information) is contained in those words.

The theory has an appealing result: there is no need to explain why excorporation is generally not possible, why a moved head does not c-command its trace and why HM should be local. Excorporation is not possible, because there was no movement involved in the first place. The c-command problem does not arise, since nothing moved. Finally, locality is captured inside the word, since there is no possibility of a head raising and skipping another head.

Another radical way of dispensing with HM is to adopt a remnant movement approach to HM. In this view, syntax simply makes no reference to information related to heads. Remnant movement was originally proposed to handle topicalization of incomplete VPs, as in (22). Only the verb in this case moves to Spec-CP and the object remains below after having raised to a higher object position (AgroP or external specifier of v). For this structure to be possible in a given language, scrambling of objects must be an available option in the grammar.

- (22) [t_i Gelesen] $_j$ hat Hans [das Buch] $_i$ t_j nicht
 read has Hans the book not
 ‘Hans has not read the book.’

Since the mid-nineties it has been suggested (most notably by Kayne 1998) that remnant movement can be extended to many other domains. For example, Koopman and Szabolcsi (2000) propose that in Hungarian, Dutch, and German, the verb raises via XP movement generally. The need to appeal to HM becomes obsolete.

5 Head movement today

Roberts (2010) provides the most recent, thoroughly worked-out analysis of HM within the Minimalist Program, so we will cover his proposal in more detail here. Roberts recasts HM as a reflex of Agree between a Probe and a defective Goal, bringing it in line with mainstream Minimalist syntax. He takes a defective goal to be one whose features are a proper subset of the Probe. Following Roberts, we illustrate the operation with Romance clitics. Roberts assumes that an active, transitive v^* has an interpretable V-feature, [iV] (to categorize the lexical root) and [$u\varphi$]. Following Cardinaletti and Starke (1999) and Déchaine and Wiltschko (2002), he further assumes that a clitic is a bare φ P devoid of Case features – in other words, a bundle of interpretable φ -features, [$i\varphi$]. When the V- v^* complex enters into an Agree relation with the clitic, the [$u\varphi$] feature set on v^* is valued by the clitic (Roberts 2010: ex (30)).

- (23) a. *Trigger for Agree*
 v^* [Pers:____, Num:____] φ [Pers: a , Num : b]
 b. *Outcome of Agree*
 v^* [Pers: a , Num: b] (φ [Pers: a , Num : b])

Roberts argues that there is no formal difference between, on the one hand, Agree that exhausts all the features of the Goal and, on the other, Move/Internal Merge of the Goal to the Probe. In other words, the outcome of (23b), $\text{Agree}(v^*, \varphi)$ is indistinguishable from $\text{Merge}(v^*, \varphi)$ precisely because the set of features on φ is a subset of the features of v^* . Contrast the two structures below. In both cases exactly the same set of features appears on v^* after either Agree or Move.

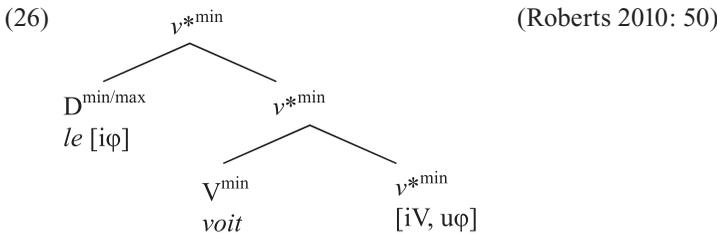
- (24) a. $[_{VP} v^* [_{\text{Pers}:a, \text{Num}:b}] \quad [_{\varphi P} \varphi [_{\text{Pers}:a, \text{Num}:b}]]]$ - outcome of Agree
 b. $[_{VP} [_{VP} \varphi [_{\text{Pers}:a, \text{Num}:b}]_i v^*] \quad [_{\varphi P} t_i]]$ - outcome of Move

Finally, HM fails with a full DP object because it contains features that are not part of v^* (such as D and Case).

Since the technical aspects of Roberts' proposal are quite different from that of standard HM, we will go over the derivation in more detail here. First, let us examine a structure containing a verb and a clitic.

- (25) Pierre le voit
 Pierre it.CL.3.SG sees
 'Pierre sees it.'

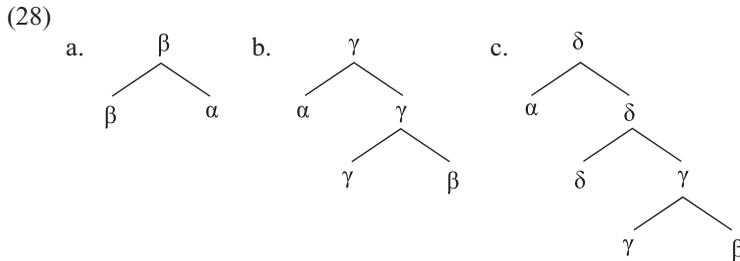
The final structure for the v^* head we will derive is shown in (26). As with Roberts, we assume that the underlying order of projections is $vP > VP > DP$.



There are two important points to consider here. First, both V and D have undergone HM to v^* : thus, they must be featurally non-distinct from v^* . Second, Roberts must ensure that V raises to v^* before D. It is easy to see in (26) that the features of D are a subset of the features of v^* . Roberts makes the reasonable assumption that a pronominal clitic is simply a bundle of phi-features, devoid of Case features. A full DP, of course, does have Case features, and so is not a subset of the features of the Probe, v^* . Roberts also assumes that lexical roots must raise to a category-defining head in the sense of Marantz (2001). In order to derive the correct surface order of heads, Roberts restates Strict Cyclicity in terms of “prominence”. Given the order projections, we expect the clitic to raise first (because it is more deeply embedded), followed by the verb, giving the incorrect order *verb-clitic-v. Roberts proposes the following definitions, where *merger* refers to the element that does not project upon Merge, and *mergee* refers to the element that does project (Roberts 2010: 52).

- (27) α is more prominent than β if either:
 (i) α is a merger for β (i.e., β projects the label of $\{\alpha, \beta\}$), or
 (ii) there is a category γ , γ a mergee for β , such that δ (reflexively) dominates γ , and α is a merger for δ .

This formulation gives rise to the following three scenarios. In all three cases, α is more prominent than β , and so must raise first if both α and β are targeted by the same Probe. (In fact, the definition above refers to any configuration akin to (28c) where α and β are separated by any number of projections.



Of course, for Roberts' proposal to work, he must show that a wide range of HM phenomena are amenable to this account. Specifically, he must show that the set of features of the 'moved' head is always a subset of those of the landing site. Indeed, he shows this with a wide range of phenomena, including second position clitics, clitic climbing, V-to-T movement, long head movement, verb second, and verb incorporation. He also suggests that noun incorporation may be handled under his approach (but see Barrie and Mathieu (2012) for an alternative view).

6 Head movement and aphasia

In this last section we make a few brief remarks on some current research on HM and Broca's aphasia. Broca's aphasia is usually characterized as slow, laboured speech which frequently omits function words. Van der Meulen (2004) summarizes previous research on comprehension deficits with movement in individuals with Broca's aphasia in addition to presenting her own research. We mention just two previous studies here before describing van der Meulen's research.

Pioneering research by Grodzinsky and Finkel (1998) indicates a significant difference in Broca's patients' understanding of HM versus XP movement. Their discussion centres on the following paradigm.

- (29) a. It seems likely that John will win.
 b. John_i seems likely t_i to win.
 c. *John_i seems that it is likely t_i to win.
- (30) a. They could have left town.
 b. Could they t_i have left town?
 c. *Have_i they could t_i left town?

Both sets of data contain an ungrammatical sentence. In (29c), the DP *John* has undergone superraising, a phenomenon usually handled by Case Theory. In (30c), the auxiliary *have* has raised above the modal *could*, in violation of the HMC. What the researchers found was that Broca's patients were able to detect the ungrammaticality in (29c), the case involving XP movement, far more easily than in (30c), the case involving HM. Grodzinsky and Finkel take these results to mean that comprehension in Broca's patients is not affected by HM.

Whatever the explanation, these results indicate a fundamental difference between XP movement and HM. More recently, Friedmann *et al.* (2006) have examined the comprehension of verb movement in Hebrew in aphasic patients. Their results indicate that aphasics *are* sensitive to comprehension of verb movement, in direct contrast to the previous experiment. Van der Meulen examined the comprehension of HM in French *wh*-questions, but the question as to whether HM affects comprehension in Broca's aphasia patients was not able to be conclusively answered, although she does ultimately conclude (on the basis of other experiments reported in her dissertation) that XP movement does affect comprehension more than does HM in individuals with Broca's aphasia.

We have just scratched the surface with respect to the relationship between HM and aphasia. What is clear from the discussion above is that more research into the neurological processing of HM versus XP movement is necessary. Important research such as this has the potential to address many of the issues related to HM discussed above. Recall that several proposals have been put forth to amend HM, including minor alterations keeping HM basically intact, relegating HM to PF, and recasting HM as remnant XP movement. It is entirely possible that more than one of these mechanisms is at play. That is, some instances of HM are really PF movement, while other instances of HM are really remnant XP movement. In this light, we do not expect uniform experimental results in the examination of HM and aphasia. Thus, we tentatively end this section with the suggestion that the contradictory experimental results above arise from differences in how HM manifests in the languages under consideration.

7 Conclusion

In this chapter we have surveyed the rise of HM from its earliest incarnations in GB Theory to our most recent understanding of it within a Minimalist perspective. Throughout the GB era, the operation of HM began to play a larger and larger role in syntax. Specifically, early work by Baker and Pollock led to an enormous research program on deriving word structure syntactically – a research program that remains active to this day. The empirical underpinning of this line of research was the realization that morphological structure within a word and syntactic structure bear too many striking similarities to be a mere accident. As the properties of HM were being investigated, a clear property came to light that still figures in syntactic theorizing. The Head Movement Constraint of Travis states that HM is confined to targeting the immediately c-commanding head. Long Head Movement was shown to be a counter-example to this generalization, prompting a re-thinking of the HMC.

Minimalism adopts Bare Phrase Structure, which does not provide any formal mechanism for HM. Furthermore, several other problems with HM were identified early on by researchers, leading to numerous discussions and reformulations of this operation. These include the following: (i) relegating HM to the PF component, (ii) recasting HM as massive remnant movement, and (iii) reprojection. We discussed the theoretical and empirical underpinnings of these three approaches. Despite the theory-internal problems with HM, many researchers to this day eschew discussions on this matter and simply assume there is some technical solution that makes HM available as an overt syntactic operation. We then presented Roberts' current proposal on HM within a Minimalist framework. As the most worked-out proposal on HM, we have presented, we hope, enough of the core details of his mechanism to allow the reader to apply these to other instances of HM.

In the final section we reviewed briefly some current research on how HM is affected in patients with aphasia. As the research presented indicates, there is a lack of consensus on

the effect of aphasia on HM. We suggested tentatively that some of the observed variation in how HM is affected may be due to the lack of a uniform operation of HM. Although an elegant solution to the problems of HM is certainly desired, the studies on alternatives to HM (PF movement, remnant XP movement, and reprojection) suggest that HM may not be a unitary operation, but that some of instances of HM may indeed be better understood as one of these other kinds of movement, crucially with a difference in how they are affected by aphasia. In brief, the issue of HM is far from settled. How it is to be understood and whether it should be treated as a single operation remain open questions.

Further reading

An excellent discussion of the empirical issues of HM in Germanic, Romance, and Celtic can be found in Roberts (2001), along with a general discussion on the properties of HM. Contemporary discussions on the properties of HM may be found in Harley (2011 and 2013). Finally, a recent survey that compares the empirical strengths of recent alternatives to HM (HM at PF, remnant XP movement, and reprojection) can be found in Roberts (2011).

References

- Adger, D. 2003. *Core Syntax*. Oxford: Oxford University Press.
- Baker, M.C. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16:373–416.
- Baker, M.C. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago, IL: University of Chicago Press.
- Baker, M.C. 2009. Is head movement still needed for noun incorporation? *Lingua* 119:148–165.
- Barrie, M., and E. Mathieu. 2012. Head movement and noun incorporation. *Linguistic Inquiry* 43:133–142.
- Belletti, A. 1990. *Generalized Verb Movement: Aspects of Verb Syntax*. Turin: Rosenberg and Sellier.
- Benedicto, E.E. 1998. Verb movement and its effects on determinerless plural subjects. In *Romance Linguistics: Theoretical Perspectives*, ed. A. Schwegler, B. Tranel, and M. Uribe-Etxebarria, 25–40. Amsterdam/Philadelphia: John Benjamins.
- Bobaljik, J.D., and S. Brown. 1997. Interarboreal operations: Head movement and the extension requirement. *Linguistic Inquiry* 28:345–356.
- Boeckx, C. and S. Stjepanovic. 2001. Head-ing towards PF. *Linguistic Inquiry* 32:345–355.
- Brody, M. 1997. Mirror theory. *UCL Working Papers in Linguistics* 9:179–122.
- Brody, M. 2000. Mirror theory: Syntactic representation in perfect syntax. *Linguistic Inquiry* 31: 29–56.
- Bury, D. 2003. Phrase structure and derived heads. PhD dissertation, University College London.
- Bury, D. 2007. Verb movement and VSO-VOS alternations. *UCL Working Papers in Linguistics* 19:77–91.
- Cardinaletti, A., and M. Starke. 1999. The typology of structural deficiency: On the three grammatical classes. In *Clitics in the Languages of Europe*, ed. H. van Riemsdijk, 145–233. Berlin: Mouton de Gruyter.
- Chomsky, N. 1970. Remarks on nominalization. In *Readings in English Transformational Grammar*, ed. R. Jacobs and P. Rosenbaum, 184–221. Washington: Georgetown University Press.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. 1993. A minimalist program for linguistic theory. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. K. Hale and S.J. Keyser, 1–52. Cambridge, MA: MIT Press.
- Chomsky, N. 1994. *Bare Phrase Structure*. MIT Occasional Papers in Linguistics, 5.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. 2000. Minimalist inquiries: The framework. In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. R. Martin, D. Michaels, and J. Uriagereka, 89–156. Cambridge, MA: MIT Press.

- Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, ed. M. Kenstowicz, 1–52. Cambridge, MA: MIT Press.
- Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. New York, NY: Oxford University Press.
- Déchaine, R.M., and M. Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33:409–442.
- Den Besten, H. 1977. On the interaction of root transformations and lexical deletive rules. MIT and University of Amsterdam.
- Den Besten, H. 1983. On the interaction of root transformations and lexical deletive rules. In *On the Formal Syntax of the Westgermania*, ed. W. Abraham, 47–131. Amsterdam: John Benjamins.
- Di Sciullo, A.M., and E. Williams. 1987. *On the Definition of Word*. Cambridge, MA: MIT Press.
- Donati, C. 2006. On Wh-head movement. In *Wh-movement: Moving On*, ed. L. Cheng and N. Corver, 21–46. Cambridge, MA: MIT Press.
- Fanselow, G. 2003. Münchhausen-style head movement and the analysis of verb second. *UCLA Working Papers in Linguistics* 13:40–76.
- Friedmann, N., A. Gvion, M. Biran, and R. Novogrodsky. 2006. Do people with agrammatic aphasic understand verb movement? *Aphasiology* 20:136–153.
- Grodzinsky, Y., and L. Finkel. 1998. The neurology of empty categories. *Journal of Cognitive Neuroscience* 10:281–292.
- Hale, K.L., and S.J. Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. Cambridge, MA: MIT Press.
- Harley, H. 2004. Merge, conflation and head movement: The first sister principle revisited. In *NELS 34*, ed. K. Moulton and M. Wolf, 239–254. Amherst, MA: GLSA.
- Harley, H. 2011. Affixation and the mirror principle. In *Interfaces in Linguistics*, ed. R. Folli and C. Ulbrich, 166–186. Oxford: Oxford University Press.
- Harley, H. 2013. Getting morphemes in order: Merger, affixation and head-movement. In *Diagnosing Syntax*, ed. L. Cheng and N. Corver, 44–74. Oxford: Oxford University Press.
- Jackendoff, R. 1977. *X-bar Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.
- Julien, M. 2002. *Syntactic Heads and Word Formation*. Oxford: Oxford University Press.
- Kayne, R. 1998. Overt vs. covert movements. *Syntax* 1:128–191.
- Kishimoto, H. 2007. Negative scope and head raising in Japanese. *Lingua* 117:247–288.
- Kishimoto, H. 2008. On the variability of negative scope in Japanese. *Journal of Linguistics* 44: 379–435.
- Kiss, T., and B. Wesche. 1991. Verb order and head movement. In *Text Understanding in LILOG*, ed. H. Otthein and R. Claus-Reiner, 216–240. Dordrecht: Springer.
- Koenenman, O. 2000. *The Flexible Nature of Verb Movement*. Utrecht University.
- Koopman, H. 1984. *The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar*. Dordrecht: Foris.
- Koopman, H., and A. Szabolcsi. 2000. *Verbal Complexes*. Cambridge, MA: MIT Press.
- Lasnik, H. 1995. A note on pseudogapping. In *MIT Working Papers in Linguistics 27: Papers in Minimalist Syntax*, 143–163. Cambridge, MA: MITPLS, Department of Linguistics and Philosophy, MIT.
- Lechner, W. 2006. An interpretive effect of head movement. In *Phases of Interpretation*, ed. M. Frascarelli, 45–71. Berlin: Mouton de Gruyter.
- Lechner, W. 2007. Interpretive effects of head movement. *LingBuzz*. Downloadable at <http://ling.auf.net/lingbuzz/000178>
- Lees, R. 1960. *The Grammar of English Nominalizations*. The Hague: Mouton.
- Lema, J., and M.-L. Rivero. 1990a. Long head movement: ECP vs. HMC. *Cahiers Linguistiques d'Ottawa* 18:61–78.
- Lema, J., and M.-L. Rivero. 1990b. Long head-movement: ECO vs. HMC. In *NELS 20*, ed. J. Carter, R.-M. Déchaine, W. Philip, and T. Sherer, 333–347. Amherst, MA: University of Massachusetts, GLSA.
- Mahajan, A. 2003. Word order and (remnant) VP movement. In *Word Order and Scrambling*, ed. S. Karimi, 217–237. Malden, MA: Blackwell Publishers.
- Marantz, A. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4:201–225.
- Marantz, A. 2001. Words. In *20th West Coast Conference on Formal Linguistics*. University of Southern California.

- Matushansky, O. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37:69–109.
- Pollock, J.-Y. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20:356–424.
- Rivero, M.-L. 1993. Long head movement vs V2 and null subjects in Old Romance. *Lingua* 89: 217–245.
- Rivero, M.-L. 1994. Clause structure and V-movement in the languages of the Balkans. *Natural Language & Linguistic Theory* 12:63–120.
- Rizzi, L. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.
- Roberts, I. 1991. Excorporation and minimality. *Linguistic Inquiry* 22:209–218.
- Roberts, I. 2001. Head movement. In *The Handbook of Contemporary Syntactic Theory*, ed. M. Baltin and C. Collins, 113–147. Oxford: Blackwell.
- Roberts, I. 2010. *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. Cambridge, MA: MIT Press.
- Roberts, I. 2011. Head-movement and the minimalist program. In *The Oxford Handbook of Linguistic Minimalism*, ed. C. Boeckx, 195–219. Oxford: Oxford University Press.
- Surányi, B. 2005. Head movement and reprojection. In *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio Linguistica. Tomus XXVI*, 313–342. Budapest: ELTE.
- Surányi, B. 2007. On phase extension and head movement. *Theoretical Linguistics* 33:121–132.
- Thiersch, C. 1978. Topics in German syntax. PhD dissertation, MIT.
- Toyoshima, T. 2000. Head-to-spec movement and dynamic economy. PhD dissertation, Cornell University.
- Travis, L.D.M. 1984. Parameters and effects of word order variation. PhD dissertation, MIT.
- Van der Meulen, A.C. 2004. Syntactic movement and comprehension deficits in Broca's aphasia. PhD dissertation, Universiteit Leiden.
- Wojdak, R. 2008. *The Linearization of Affixes: Evidence from Nuu-Chah-Nulth*. New York, NY/Berlin: Springer.