

Head Movement as Feature Geometry Growth

Michael Barrie, Sogang University

Abstract: Head Movement (HM) is problematic in a number of ways. In this paper, I concentrate on the fact that HM is unformulable in Bare Phrase Structure (BPS). Rather than propose novel mechanisms or complicate the notion of Merge, I propose that HM be rethought of as Feature Geometry growth. Taking seriously the idea that syntactic heads are merely feature bundles organized into a geometry, I propose that these feature bundles can merge, forming larger feature geometries. The feature geometry thus formed must be consistent with the overall feature geometry of the language in question. Just as intervening features in a Feature Geometry cannot be missing, HM cannot skip intervening heads, thus deriving the Head Movement Constraint. This theory of HM is illustrated in depth with negation in English, resolving some thorny issues left unsettled under the traditional theory of HM. Some tentative remarks are made regarding noun incorporation and Long Head Movement. The paper concludes with a discussion of the advantages of the proposed theory of the traditional theory of HM and with some unresolved issues.

1 Introduction

This paper proposes a novel theory for head movement, addressing two problematic issues: the status of the Head Movement Constraint (HMC) as a primitive of grammar (Travis 1984), and Bare Phrase Structure (BPS) (Chomsky 1994). The HMC was originally constrained by Proper Government, a concept that is no longer tenable in light of Minimalism. In practice, the HMC is still regarded as a primitive of grammar, but one that no longer relies on Proper Government. Here, I seek to derive the HMC from a novel theory of HM, which is based on feature geometries. Regarding the second issue, BPS has no clear set theoretic notation to distinguish between the

formation of a phrase and the formation of a complex head. Again, the theory of HM proposed here addresses this lacuna using feature geometries.

Empirically, I address negation in English and verb movement in English and French. I show how the proposed theory of head movement accounts for some peculiar properties of English negation. I also illustrate the proposed theory of HM on the well-known difference in verb movement in English and French (Pollock 1989). This is important to show that no empirical ground is lost with the proposal of this new theory of HM. I further explore two issues of current interest to HM, noun incorporation and long head movement, but make only tentative remarks.

The gist of the proposal is as follows. I adopt the notion explored in much recent literature that the features of syntactic heads are arranged in a feature geometry (Cowper 2005; Harley & Ritter 2002; McGinnis 2005). I propose that head movement is simply the formation of a larger feature geometry combining compatible features. If two feature geometries contain incompatible features (i.e., features that are not part of the same overall feature geometry), then they cannot form a larger feature geometry, thereby excluding the possibility of HM. If they are compatible, then they simply form a larger feature geometry, still comprising a single syntactic head. Thus, HM boils down to feature geometry growth. The HMC constraint follows from this formulation (as do other properties of HM) by virtue of the fact that intermediate features in a feature geometry cannot be missing. For instance, *v* moving to C without stopping in T would result in a feature geometry with missing intermediate nodes and is thus impossible.

It is well known that HM has been controversial and problematic (Boeckx & Stjepanović 2001; Fanselow 2003; Mahajan 2003). Although several problems for HM have been discussed at length (Matushansky 2006; Roberts 2010), I focus on one particular problem here, which has not played a large role in these more recent discussions of head movement. Specifically, I focus

on the incompatibility of HM with BPS as developed by Chomsky (1994). In short, there is simply no way to express a complex head in BPS. Consider the following. Merge (a,b) results in a nested set $\{a, \{a,b\}\}$, in which a is the label of the phrase AP. Even if we come up with a set theoretic way to distinguish complex heads from phrases, in a situation where a and b are both heads how do we know if the result is a complex head or a phrase? Typically, if V merges with v , the result is typically assumed to be a complex head. If V merges with D, say where D is a clitic and just a single head, then the result is typically assumed to be a phrase, VP.

- (1) a. Merge (V, v) \rightarrow $v+V$ (complex head)
b. Merge (V, D) \rightarrow VP, where D is a clitic

I explore this problem in more detail below and suggest that features geometries offer a solution.

English negation has a number of interesting properties. The non-cliticized form *not* seems to block affix hopping but not head movement. The cliticized form *n't* also blocks affix hopping, but participates in HM if, for example, the auxiliary it cliticizes to raises to C. The adverbial negative *never* neither blocks affix hopping nor participates in HM in any way. The traditional explanation for these facts assumes that *n't* is the head of NegP and so participates in HM. Further, *not* is assumed to be in the Specifier of NegP and so does not block HM or participate in it. This is difficult to reconcile with the fact that it blocks affix hopping however, since other intervening non-heads (such as *never*) have no effect on affix hopping. We show here that these traditional assumptions are incorrect, however, and that the peculiar facts about negation and HM in English fall out naturally from the theory of HM as feature geometry growth proposed here.

The remainder of this paper is structured as follows. Section 2 covers the background for this proposal, namely it discusses the properties of HM and of feature geometries. Section 3 lays

out the proposal, namely that head movement proceeds through feature geometry growth.

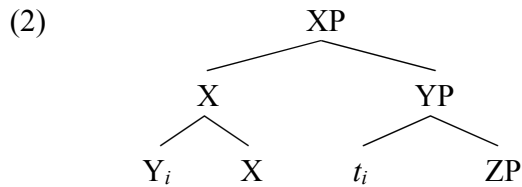
Section 3 also illustrates this proposal with a concrete example, namely verb movement from V to *v* in English and from V to T in French. Section 4 discusses the consequences of this proposal, concentrating on negation and HM in English. The main theoretical import of this section is that, while the HMC generally holds, exceptions to it can be found. In the proposed theory, occasional exceptions to the HMC are inevitable, though not unconstrained. Section 5 considers two other theoretical applications of the theory, namely noun incorporation and long head movement. It also considers some further theoretical aspects of the proposed theory. Section 6 is a brief conclusion, discussing some outstanding issues.

2 Background

This section provides the relevant background for the forthcoming proposal. First, I present the relevant properties of HM and then discuss some recent treatments of it in light of the problems that have been pointed out. In the second part of this section, I lay out the properties of feature geometries and how they have played a role in morphosyntax.

2.1 Head Movement

A number of survey papers exist providing an extended discussion of HM (Barrie & Mathieu 2016b; Harley 2013a; Roberts 2001; Roberts 2011). The reader is invited to consult these references for a fuller discussion than what is offered here. I concentrate only on those details that are relevant to the discussion. HM involves the movement of a head to a c-commanding head position, typically to the next higher head. In X-Bar Theory, it was not clearly established whether this movement took place by adjunction or by substitution, the point being moot after the abandonment of X-Bar Theory in favour of BPS. Nevertheless, the following hypothetical structure is a typical illustration of HM.



After HM, the phrase XP is headed by a complex head containing both X and Y. Most recent discussions that eschew technical details of HM still employ such formulations.¹

One of the most well-known properties of HM, adumbrated in the preceding paragraph, is accounted for by the Head Movement Constraint (HMC) originally formulated as follows.

(3) Head Movement Constraint: (Travis 1984)

An X^0 may only move into the Y^0 which properly governs it.

Without delving into the details of Proper Government, this constraint accounts for the following two properties.

- HM (generally) targets the immediately c-commanding head.
- HM cannot take place from a specifier.

In Travis' original discussion, she derived the HMC from the Empty Category Principle (ECP), a mechanism from Government & Binding Theory, which is no longer tenable under Minimalist assumptions. In brief, a head properly governs only the head of its complement. For example, in modern terms, if T takes vP as a complement, the T properly governs v . Subsequent treatments of the HMC sought to derive it from Relativized Minimality (Rizzi 1990), but this framework is also not tenable under the Minimalist Program. Roberts (2001) simply restates the HMC as

¹ This comment is not meant to criticize these other discussions for failing to address HM. It may very well be that such finer details does not matter for the analysis presented, so the absence of such a discussion is warranted.

follows, essentially leaving it as a primitive of grammar. This is virtually the current status of the HMC.

(4) Head Movement Constraint: (Roberts 2001)

Head movement of X to Y cannot “skip” an intervening head Z.

As mentioned in the introduction, HM has come under much scrutiny (Fanselow 2003; Mahajan 2003). This has led researchers to suggest several avenues to address these problems such as relegating HM to PF (Boeckx & Stjepanović 2001), recasting HM as remnant XP movement (Koopman & Szabolcsi 2000), or moving the head to the specifier position rather than the head position (Matushansky 2006; Toyoshima 2000). While these avenues may be the right solution for some phenomena, it is not clear that all forms of HM can be handled with these alternatives. Indeed Roberts (2010) outlines numerous problems with these approaches for canonical cases of HM in the literature.

Another property of HM that has not generated much attention in previous research is its incompatibility with BPS. As mentioned in the introduction, the operation for building phrase structure, Merge, takes two syntactic objects and produces a phrase. Consider the following derivations involving Merge between a head and XP, and two XP's, respectively.

- (5) a. Merge (X, YP) \rightarrow {X, {X, YP}}
b. Merge (XP, YP) \rightarrow {X, {XP, YP}}

In both cases, the output of Merge is a complex, nested set. The larger set contains the label and the smaller set composed of the two merged syntactic objects. Turning to HM, there is no set-theoretic way to distinguish between the following two structures.



Specifically, the first structure can be formed easily by Merge as follows.

$$(7) \quad \text{Merge}(X, Y) \rightarrow \{X, \{X, Y\}\}$$

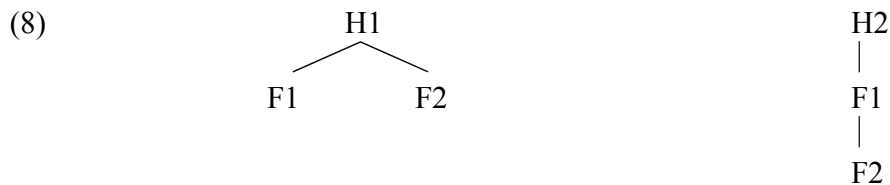
The second structure, however, has no set-theoretic correlate. One could propose another operation, call it HeadMerge, that somehow notationally distinguishes a phrase from a complex head. In essence, the current proposal actually falls along these lines; however, rather than proposing a new operation, I exploit already existing machinery to create a new formalism for HM. This, of course, is feature geometries, which I outline later.

Given the discussion above, one may wonder whether the distinction between heads and phrases (as well as the distinction between HM and XP movement) is needed within a minimalist approach to language. Indeed, dispensing with the notion, if possible, would be a welcome simplification of the theory. There are reasons to believe that the distinction is necessary to maintain, however. Unlike phrasal movement, HM is strictly clause-bound and highly local. First, phrasal movement, in particular A-Bar movement, can move long-distance, crossing many clauses. A-movement, of course, is typically does not exhibit long-distance movement; however, raising-to-subject verbs, ECM, and German long passives (Wurmbrand 2001) do show instances of longer movement, albeit in restructuring contexts. As mentioned, HM is constrained by the HMC, requiring very local movement. As mentioned, this is interpreted as movement to the immediately c-commanding head. Although we will see some exceptions to the HMC and account for them, phrasal movement and HM are fundamentally different in terms of locality. Finally, I note that several studies of aphasia suggest that HM and phrasal movement are processed differently in the brain; however, I do not discuss the details here (Grodzinsky & Finkel 1998; van der Meulen 2004). From these observations, I conclude that HM is a

significantly different operation from phrasal movement, necessitating a different mechanism that I propose below.

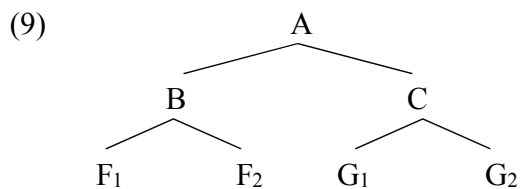
2.2 Feature Geometries

There has been much foundational work on syntactic feature geometries (Cowper 2005; Harley & Ritter 2002; McGinnis 2005), which has led to the analysis of complex agreement patterns such as Person-Case constraints, person hierarchies, and omnivorous number (Barrie 2016; Béjar 2003; Béjar & Rezac 2009; Preminger 2014). Feature geometries capture implicational relationships among the features of a syntactic head. In the following examples, the hypothetical head, H1, consists of two features in a sisterhood relation, meaning either F1 or F2 could be present. Also, both features could be present or both features could be absent. The hypothetical head, H2, however, contains the same two features in a mother daughter relationship. In this case again, both features could be present or absent. Also, F1 could appear alone, but F2 cannot appear alone. F2 entails the presence of F1.



3 Proposal

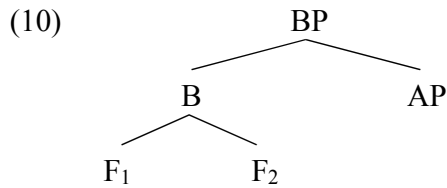
Assume the grammar for a language has the following (partial) feature geometry.



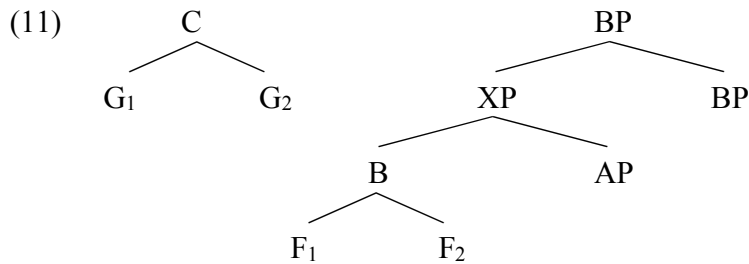
The gist of the proposal is that A can be formed by head movement of C to B (or B to C). If Merge (a,b) consists of the merger of two compatible pieces of feature geometry, then a new

larger feature geometry is formed, giving rise to a larger head. If not, then Merge (a, b) results in an XP.

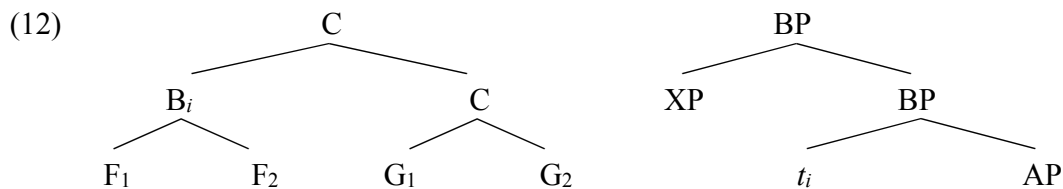
Consider the following hypothetical tree, in which the head B is part of the larger feature geometry in (9).



Next head, C, enters the workspace.

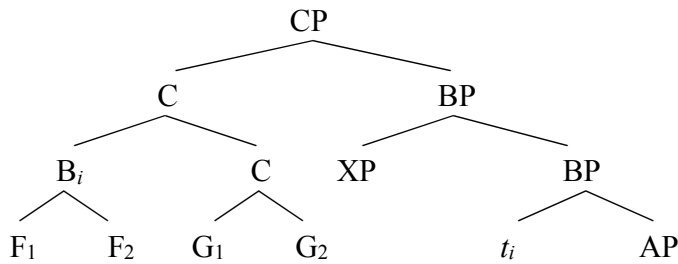


Finally, B undergoes HM to C. However, to avoid counter-cyclicity, we assume the mechanism proposed in Bobaljik and Brown (1997). Namely, before C merges with BP (forming CP), B undergoes sideward HM to C thereby satisfying the Extension Condition. As we saw in (9) above, the features of B and C are part of the same feature geometry, so the merger of the two simply creates a larger feature geometry, a single complex head.



Finally, C merges with BP to form CP.

(13)



If B and C were not part of the same feature geometry, then they would form a phrase upon merger.

In order for this system to work, the speaker must know the whole feature geometry to compute whether Merge (a, b) forms a head or an XP. Feature geometries are not universal (Cowper 2005: inter alia). In fact, Cowper derives many differences between the syntax of tense and aspect in English and Spanish by assuming slightly different feature geometric representations of Infl in the two languages. This proposal, then, makes the prediction that in early stages of acquisition, before children have worked out full feature geometries, head movement will be more limited. I leave this line of thought to future research.

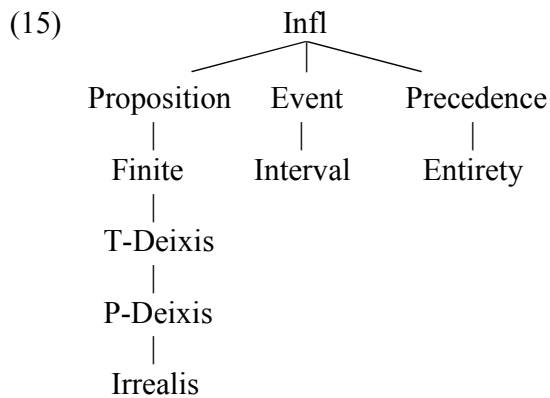
3.1 *Application to Verb Movement*

In this section, I make the proposal concrete, applying it to verb movement in English and French. Following Pollock (1989), the difference in word order between English and French boils down to verb movement. Updating his discussion to modern terms, English exhibits V-to-v movement, while French exhibits V-to-T movement. This is shown by the following contrast. Assuming, as is now standard, that the position of the adverb is constant, the difference in word order between English and French is accounted for by assuming that French has overt V-to-T movement.

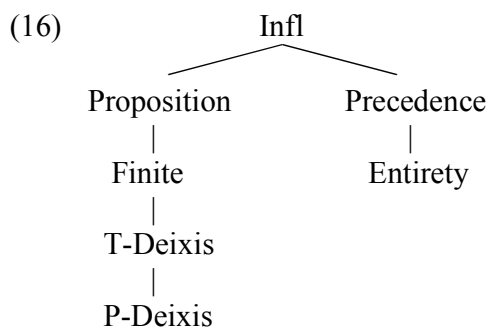
(14) a. John often eats anchovies.

- b. Jean mange souvent des anchois
 John eats often of.the anvovies
 ‘John often eats anchovies.’

Assume, following Cowper (2005), that the features of Infl are arranged in a feature geometry. Below is the feature geometry that Cowper assumes for English.

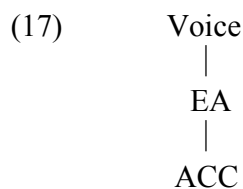


For convenience, I borrow Cowper’s structure for Infl in Spanish and apply it to French as the two languages have similar tense and aspect properties. Of course, the actual structure of Infl in French would need to be investigated further. The precise analysis of French Infl, however, is not crucial to the analysis here. Here is Cowper’s structure for Spanish Infl.



Although the features of Infl have been investigated in detail for a number of languages, the features of other functional projections have received less attention from the perspective of

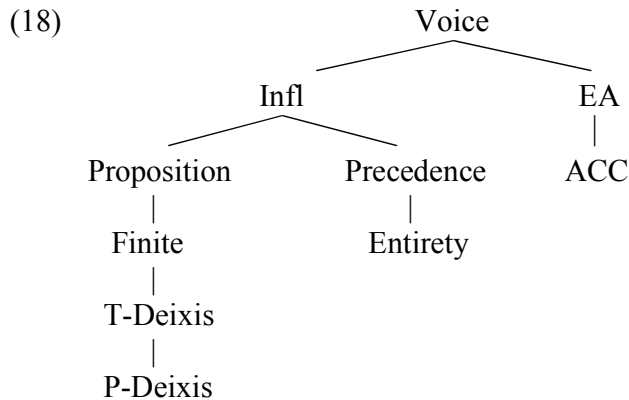
feature geometries.² Let us assume also that the features of v are also arranged in a feature geometry. Assume for convenience that there are two features, perhaps arranged as follows to allow for the distinction between transitive and unergative verbs. The feature EA is responsible for assigning a theta role to an external argument. The feature ACC is responsible for assigning accusative Case to the direct object. When ACC is absent, only an external argument is generated. This embryonic proposal is not meant to encompass a complete proposal for the organization of the features of the vP layer (see footnote 2). It is merely meant to provide a framework to illustrate the proposal of head movement.



The current proposal suggests that both geometries are actually part of a larger geometry that can remain separate in languages without V-to-T movement or can be united in languages with V-to-T movement.³ The structure after V-to-T movement is shown here.

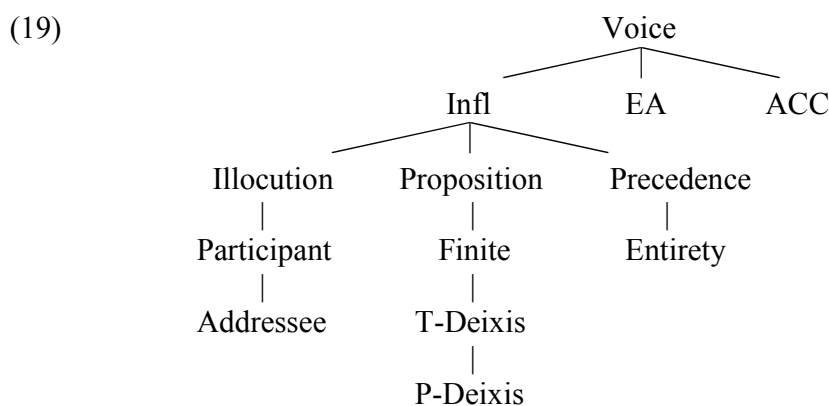
² In particular, most research into the structure of the vP layer has involved splitting the v head into separate projections (Alexiadou et al. 2015; Harley 2013b; Ramchand 2008). Pylkkänen (2008), among others, discuss the notion of feature bundling, in which the features of two distinct heads appear in a single syntactic head. Feature bundling clearly has much in common with the proposal presented here; however, further research into the feature bundles of the vP layer is necessary to determine the compatibility between the current proposal and the organization of the vP layer.

³ For now, I am using V as an abbreviation for $V+v$. I address verbal roots in the next section.



Let us now place this structure in a full clausal structure, again delaying the issue of how the verbal root enters the derivation. For ease of representation, I use the labels Voice and Infl in place of the full feature geometries. Consider first the derivation of the English sentence in (14)a.

Ultimately, V to C movement is possible as we have one single feature geometry that can be put together by the individual heads. The single, large feature geometry captures notion of extended projection. As with the *v*-layer, I do not give a full geometry for the C layer. Among the features encoded in the C layer are discourse participants and other discourse properties such as interrogation, topic and force (Heim et al. 2014; Isac 2012; Rizzi 1997; Wiltschko & Heim to appear). For ease of exposition, I represent only the discourse participant features here.

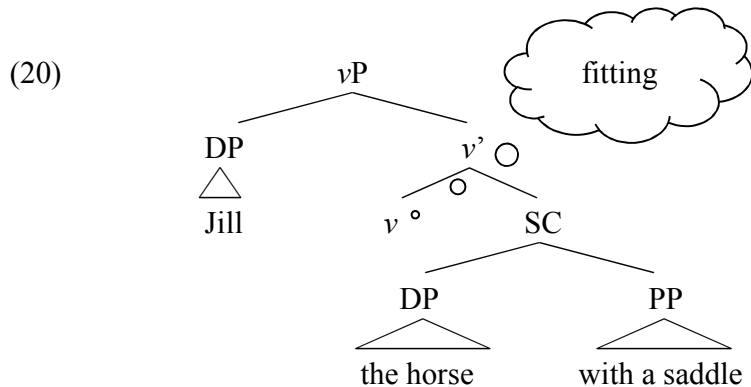


This section has shown how the feature geometry can “grow” thereby giving rise to the illusion of HM. In the next section I discuss how the lexical root comes to appear in the structure.

3.2 *Putting the Verb Root in*

The so-called Chomsky-Borer Conjecture holds that roots are devoid of all formal features (Borer 2005; Kayne 2009; Marantz 2001). Likewise, roots have been severed from categorial features, rendering the functional hierarchy independent of roots (Marantz, 2001), meaning that a given root can appear in any extended projection, as long as an interpretation can be assigned. Thus, it is not easy to see how roots fit into the theory of HM developed here. Specifically, since the root does not possess formal features of any kind it cannot form part of the feature geometry of the extended projection. I discuss here some current theorizing on roots and how they can be made amenable to the theory of HM proposed here.

Ever since the root was severed from its categorial information, it has typically been simply represented as $\sqrt{\quad}$, heading a \sqrt{P} (Harley 2013c). Another possibility that has been explored is that the root lexicalizes a set of functional projections (Harley 2005; Ramchand 2008). Ramchand has a well worked out proposal in which V is broken down into three functional projections. Ultimately, the choice of which root is chosen to lexicalize the relevant heads is arbitrary and does not play a role in the syntax. Harley (2005) provides the following structure, indicating that the concept of the root is not part of the syntactic structure at all, but is rather somehow loosely associated with it.



“Jill fit the horse with a saddle.”

I refer the reader to Harley’s and Ramchand’s respective discussions for further details on how this idea is implemented. What is important for the discussion here is that the lexical root is not part of the feature geometry. The lexical root is added at a later stage (cf. Haugen & Siddiqi 2013), the choice of which is largely non-linguistic.⁴ For ease of presentation I will continue to put the verb root in the tree diagrams below, although it should be remembered that they do not enter the derivation until later.

This section has presented the proposed theory of HM as Feature Geometry Growth. The core idea is based on the notion that HM is unformulable within BPS. Thus, the idea proposed here is that HM is simply the formation of a larger and larger feature geometry. The next section illustrates several problems with HM and English Negation and how HM as Feature Geometry growth resolves these problems.

4 English Negation

This section analyzes head movement and negation in English in light of the proposal for HM developed here. First, I go over properties of negation showing that analyzing *not* as a head is problematic for a strict version of the HMC. Second, I argue that *not* indeed is the head of NegP

⁴ Thus, the difference between *Mary read the book* and *Mary wrote the book* depends on how the speaker wishes to describe the real world rather than on any linguistic property.

(rather than a specifier or adverb). Finally, I show that the proposed theory of HM accounts for the facts presented in the first section in light of the fact that *not* is a head.

4.1 Auxiliary/Negation Interaction

The standard analysis for negation in English holds that *not* is in SpecNegP and *n't* is the head of NegP (Haegeman 1995). Consider the following data.

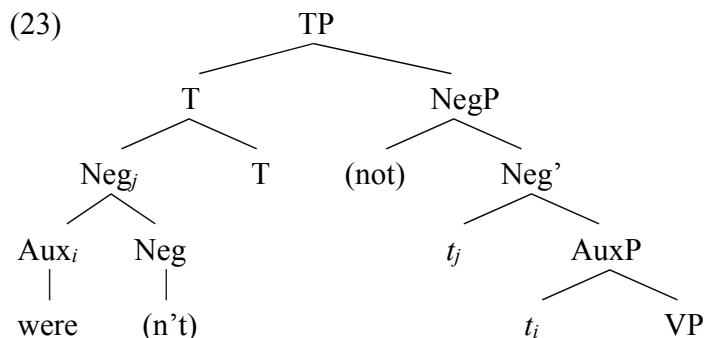
- (21) a. Mary is not reading.
b. Is Mary not reading?
c. Mary isn't reading.
d. Isn't Mary reading?

Assume, following Adger (2003) and Lasnik (Lasnik 1995), that auxiliaries are merged in specified functional projections below T and subsequently raise to T. The following pair of examples illustrate this clearly. The auxiliary *be* raises to T in the second example, but remains below Neg in the first example. Clearly, the presence of *not* is not an impediment to this movement.

- (22) a. The apples have not been washed.
b. The apples were_i not *t_i* washed.

Such observations led to the suggestion that *not* occupies SpecNegP.

Considering these examples further, observe that the auxiliary under T can then move on to C in interrogative constructions, (21)b. The contracted form of negation, *n't*, however, moves along with the auxiliary up to C in (21)d. Thus, the standard analysis holds that *n't* occupies the head of NegP, and the auxiliary picks up negation on its way to T. The following tree shows the respective locations of *not* and *n't* after HM to T.



Before moving on to the discussion of *not* as a head, I point out a problem with (23) under a model of syntax that assumes Late Insertion. In G&B Theory, lexical items were inserted fully inflected into the derivation. With the advent of Distributed Morphology, the concept of Late Insertion has become the predominant model of how words are ultimately built in the derivation (Halle & Marantz 1993; Harley & Noyer 2003; Haugen & Siddiqi 2013). In (23), an abstract auxiliary corresponding to passive voice merges with the VP as shown. Let's call this abstract auxiliary BE. At PF, the auxiliary must be adjacent to the features on T in order for the correct form of the auxiliary to be inserted. One could argue that the form of the auxiliary is simply grammatically-conditioned allomorphy, but this will not work for past tense forms with an overt /-ed/ morpheme such the perfect auxiliary *have-had*. The structure in (23) predicts that contracted negation should intervene between the auxiliary and the past tense marker resulting in **ha-n't-ed* rather than *ha-d-n't*.

4.2 'Not' as a Head

There are problems with the standard analysis, however. The negator *not* is different from other adverbs and negative adverbs. In particular, *not* blocks affix hopping and cannot be phrasal. Consider first affix hopping. As is well known, only the negator *not* requires *do*-support. Adverbs such as *never* do not. Consider the following examples.

- (24) a. Mary does not eat anchovies.
b. Mary never eats anchovies.

Clearly, some difference underlies the pattern in the examples above. Treating *not* as a head neatly captures the fact that affix hopping is blocked. Negative adverbs such as *never* can be phrasal, while *not* cannot. Consider the following examples. The constituency of the phrase *never in a million years* is confirmed by example (25)c.

- (25) a. Fred will never in a million years eat anchovies.
b. * Fred will not in a million years eat anchovies.
c. Never in a million years will Fred eat anchovies.

Again, these facts strongly suggest that *never* appears in SpecNegP while *not* is the head of NegP.

Williams (1994) provides further evidence that *not* is a head rather than a specifier. I review his evidence here. First, Williams notes that *not* acts as a head in that it appears to the left of the complement of Neg. Adverbs, however, typically can appear on either side of the phrase they adjoin to. Note, however, that this is not an argument against the claims that *not* occupies SpecNegP as Specifiers are universally on the left.

The second argument concerns the placement of adjectival modifiers with respect to the noun they modify. Adjectives and adjective phrases are typically pre-nominal in English; however, when the adjective has a complement, the AdjP appears post-nominally. Consider the following examples, based on Williams' original examples.⁵

⁵ In light of Abney's (1987) proposal that degree words do not appear in the Specifier of AdjP, but rather head a DegP that takes AdjP as a complement Williams' argument seems to be problematic. This can be remedied, however, by assuming that degree words in English occupy

- (26) a. the happy person.
 b. the very happy person
 c. the person happy with the results

Based on this paradigm *not* patterns with heads, as the following example shows.

- (27) The person not happy is Pete.

Williams' third argument concerns VP ellipsis. Potsdam (1997) makes the same argument with VP ellipsis in English subjunctives, so I cover both points here. Consider the following pairs of examples, based on Williams and Potsdam, respectively.

- (28) a. I consider Bill to be clever, and John not Δ .
 b. * I consider Bill often to be pleasant, but John never.
- (29) a. * Mary suggests that you attend the meeting tomorrow, and I also insist that you Δ .
 b. Mary suggests that you attend the meeting tomorrow, but I insist that you not Δ .

Assuming, as both authors above do, that VP ellipsis (or predicate ellipsis more generally) must be licensed by an overt head. Again, the data above fall into place if we assume (as Williams and Potsdam conclude) that *not* is a head.

In sum, there is considerable evidence that *not* occupies the head of NegP rather than the specifier. In the final section I discuss how the proposed theory of HM as the growth of feature geometries can overcome the problems from the previous two sections. First, there is the fact that, assuming Late Insertion, the standard analysis of HM predicts the wrong order for the auxiliary,

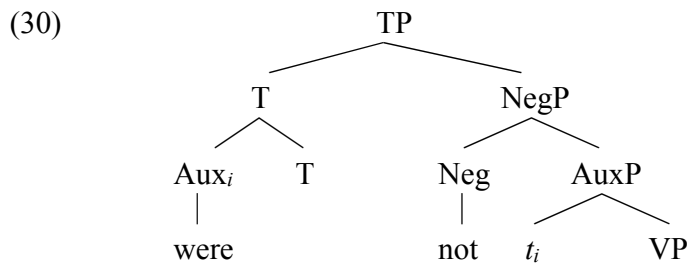
the Specifier of DegP, rather than the head. Given that degree expressions can be phrasal (as in the following bold-faced example), placing them in the Specifier seems unavoidable.

- i. a **most extremely** tall person

cliticized negation, and past tense. Second there is the paradox that *not* is a head, but does not block HM of the auxiliary to T.

4.3 Head Movement over Negation

Recall that the current proposal assumes that HM proceeds by building up compatible pieces of feature geometry. If an intervening head contains features that are not part of the part of the feature geometry under consideration, it is simply skipped. I propose here that that is exactly what is occurring here. Namely, the auxiliary skips the Neg head since the feature of Neg are not part of the feature geometry linking the auxiliary and T. The following tree diagram represents the proposed movement. Recall Cowper's proposal that tense and aspect features form a large feature geometry. Thus, the complex T head shown here is really a simplified representation of a much larger feature geometry.



This proposal appears to give rise to a new problem, however. Specifically, the original proposal for HM accounts nicely for *n't* cliticization (Late Insertion aside), but the proposed theory appears not to. In the tree in (30) cliticization does not appear to be a serious problem as the auxiliary and negation are adjacent. Cliticization of negation on to the auxiliary is problematic when T-to-C movement takes place, however. Consider the following example.

(31) Hasn't Mary read that book?

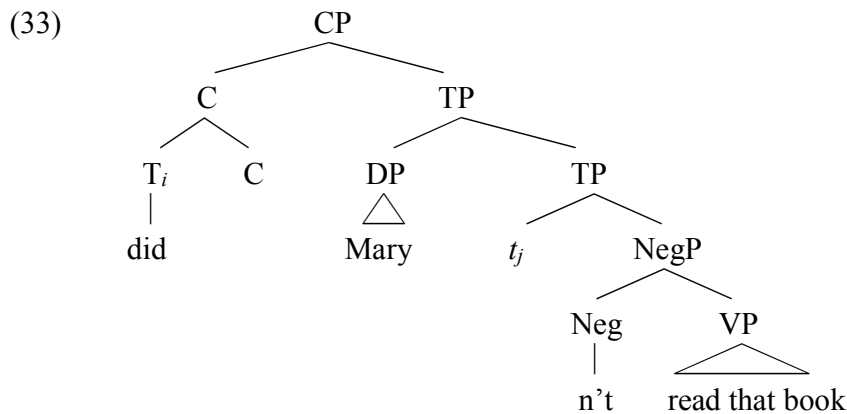
Under the the original account of HM the auxiliary moves up through Neg to T, then to C. While the the auxiliary passes through Neg, it picks up *n't* as a clitic. Under the theory of HM

proposed here, the auxiliary skips Neg. Assuming cliticization is a PF process, then a problem arises for the account proposed here. While I agree that this is a problem, I argue that it is not a new one, and in fact existed for the traditional account of HM, too.

While the traditional theory of HM accounts for negative cliticization in (31) it fails to account for negative cliticization in *do*-insertion contexts. Consider the following examples.

- (32) a. Did Mary not read that book?
 b. Didn't Mary read that book?

Do-insertion is alternatively analyzed as base generation of DO in T in the overt syntax or as PF insertion. Regardless of the correct analysis for *do*-insertion, the auxiliary *do* does not raise up through the head of NegP, as is required on the traditional analysis of HM. Consider the following example.



To be precise, at no point in the derivation are T and Neg joined by HM, So, we still need an account for cliticization both in the traditional analysis and the analysis proposed here.

Two possibilities for this problem present themselves. One possibility to this problem is that cliticization happens as soon as the host appears enters the derivation. Another is to assume some notion of head adjacency for cliticization at PF. I briefly discuss the need to consider PF cliticization next.

The need to consider post-syntactic cliticization is underscored by the following data from the following Colloquial Spoken English data (Johnson 1988). Contrary to Standard English, auxiliaries can cliticize to modals and negated modals and subsequently undergo subject/aux inversion.⁶

- (34) a. John could have done that, couldn't've he?
b. * John could have done that, couldn't have he?
c. What could've he possibly done about it?!?
d. * What could have he possibly done about it?!?

On the uncontroversial assumption that the modal *could* does not raise through a functional projection containing *have*, these data indicate that cliticization does not take place only by HM. Clearly, some notion of head adjacency is required. Since the proposed analysis fares no better or worse than the traditional analysis in this regard, I leave this point for future research.

This section has considered HM of auxiliaries to T in English and its interaction with negation. After showing that the English negator *not* is unequivocally a head rather than a specifier, I showed that the traditional analysis of HM, which entails a strict interpretation of the HMC cannot account for the data. I showed that the current proposal of HM as feature geometry growth can account for the data shown. This section ended with a brief discussion on cliticization and HM in English. It was suggested that the problems presented by these data are equally

⁶ Thanks to Moonhyun Sung for discussing these data with me in detail. I have discussed these data with several native speakers of English (both linguists and non-linguists). Many agree with the judgements as given here. Between the two grammatical sentences, some speakers accept only one, with variation among speakers. A few speakers do not accept any of these data. Clearly, more research is needed on the syntactic variation of this construction.

problematic for the current proposal and the traditional analysis of HM. Although an avenue of investigation was suggested, the analysis of the cliticization facts were left to future research. In the next section we discuss some other consequences of HM as feature geometry growth.

5 Other Consequences

This proposal has clear consequences beyond what I have explored in the previous section. Given the programmatic nature of this proposal, it is impossible to explore all avenues in detail. I only touch on some ideas here in order to stimulate future research. Specifically, I look at noun incorporation and long head movement, making some speculative remarks as they relate to the proposal.

5.1 Noun Incorporation

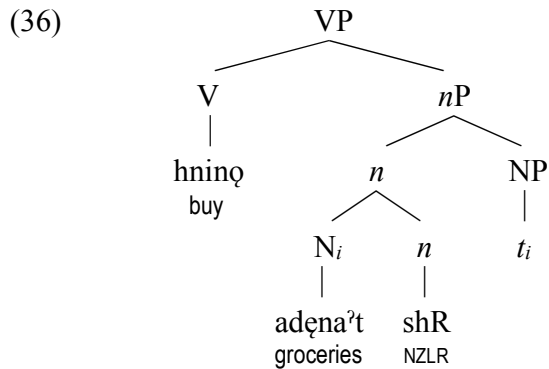
Noun incorporation is traditionally analysed as head movement (Baker 1988).⁷ As mentioned, the proposal here has a clear extension to N raising. The extended nominal projection has its own feature geometry, which can be built up the same we described above for the extended verbal projection. However, the features of the extended nominal and verbal projections comprise different feature geometries. They are not compatible. Although simple root merger may be possible for some forms of noun incorporation (discussed below), many instances of noun incorporation involve functional material from the extended nominal projection (Barrie & Mathieu 2016a). I discuss one example here. The incorporated noun is enclosed in square brackets in the following example.

⁷ I am considering generative accounts of noun incorporation that eschew Lexicalist accounts of the phenomenon. For further discussion see Massam (2009) and Barrie (2015).

- (35) agadənaʔtshähni:nóh [Onondaga, (Woodbury 2003: : 456)]
 ak- [atənaʔt- shR-] a- hniŋ -h
 1.SG.PAT- groceries- NZLR- EPEN- buy -STAT
 ‘I have bought groceries.’

In example (35), the incorporated noun consists of a noun root and a nominalizer. There is disagreement in the literature as to whether the incorporated noun is formed by head movement, forming a complex head (Baker 1988; Baker 2009), or by XP movement, forming a phrase (Barrie & Mathieu 2012; Barrie & Mathieu 2016a). If the incorporated noun is an XP, then it is clear that NI must proceed by XP movement. Let us explore the possibility that the incorporated noun is a complex head, formed by the same mechanism proposed here. The incorporated nouns consists of the nominal root and the nominalizer, *n*. At the relevant stage of the derivation, the verbal head merges with the complex nominal head. Given that the nominalizer hosts formal features that are incompatible with the formal features of the feature geometry of the extended verbal projection, Merge (V, n) results in $\{V, \{V, n\}\}$, a complex set, hence an XP.⁸ Consider the following example based on the Onondaga example above.

⁸ There is a complication at this point. Strictly speaking, the incorporated noun has merged only with the verbal head, a root. At this point there are no formal features of the extended verbal projection, so there should be no problem for head movement. One could assume that subsequent movement operations necessary to form the verbal complex will be ruled out because of the presence of nominal features in the verbal complex at this point. One could also adopt a neo-Davidsonian approach whereby all arguments are selected by a functional projection rather than by the root. Either approach assures that NI proceeds by XP movement rather than by head movement.



The complex head, *n*, in (36) contains features of the extended nominal projection. Given the current proposal of HM as feature geometry growth, the nominal features in the incorporated noun cannot combine with the verbal features in the verb. Thus, noun incorporation cannot arise by HM, in line with the XP movement analysis of (Barrie & Mathieu 2012; Barrie & Mathieu 2016a).

In this short section we conclude with a brief discussion on possibility of noun incorporation proceeding by head movement in the case that the incorporated noun is a bare nominal root. Salish languages provide a point of discussion here. Consider the following Upriver Halkomelem examples (Wiltschko 2009: 211).

- (37) a. th'έxw-wíl-t-es te ló:thel
 wash-dish-TR-3.SG DET dish
 'He washed the dish.'
- b. th'έxw-als-t-es te qwe'óp
 wash-fruit-TR-3.SG DET apple
 'He washed the apple.'

Unlike the forms of noun incorporation discussed for Onondaga above, noun incorporation in Upriver Halkomelem always consists of the incorporation of a bare root with no functional material (Wiltschko 2009). Given that no functional material from the extended nominal projection appears in the incorporated noun, noun incorporation in Upriver Halkomelem can, in

principle, proceed by head adjunction of roots as described in section 4. We leave these ideas for future research and move on to long head movement.

5.2 *Long Head Movement*

In a series of papers (Lema & Rivero 1990; Rivero 1993) a well-known exception to the HMC was explored in which the verb undergoes movement to C without stopping in T. This kind of movement is called Long Head Movement (LHM) and is found in various Romance and Balkan languages. Although subsequent proposals have dealt with this issue in various ways (Roberts 1994; Roberts 2010), it is clearly incompatible with the current proposal.

Let's consider an example of long head movement from Romanian (Alboiu 2000).

- (38) a. Mîncă-l-ar mama!
 eat-CL.3SG.ACC.M.-AUX.COND.3SG mother-the
 (affectionate idiom translated along the lines of '(He's so sweet) his mum could almost eat him.')
- b. Lua-te-ar dracul!
 take-CL.2SG.ACC-AUX.COND.3SG devil-the
 'Go to Hell!'

Again, only tentative remarks are made here as a full analysis is outside the confines of this paper. Notice that these examples have an affective or focussed reading. Let us propose that the verbal head (the V+v feature bundle) raises to a discourse position in the left periphery. For convenience let us assume it raises to SpecFocP. Thus, the LHM in these examples is only an illusion. In effect, the verb undergoes focus fronting as XP movement rather than as head movement. I leave these tentative remarks here and move on to some theoretical implications of this proposal.

5.3 *Theoretical Implications*

Any new theoretical proposal must be able to account for those facts covered by previous proposals. Indeed, it should ideally be able to cover more theoretical ground. In this section, I

discuss the HMC (Travis 1984) and the clause-boundedness of HM. I show that these two properties of HM fall out as a natural consequence of HM as feature geometry growth. First, I discuss the HMC and then move on to clause-boundedness.

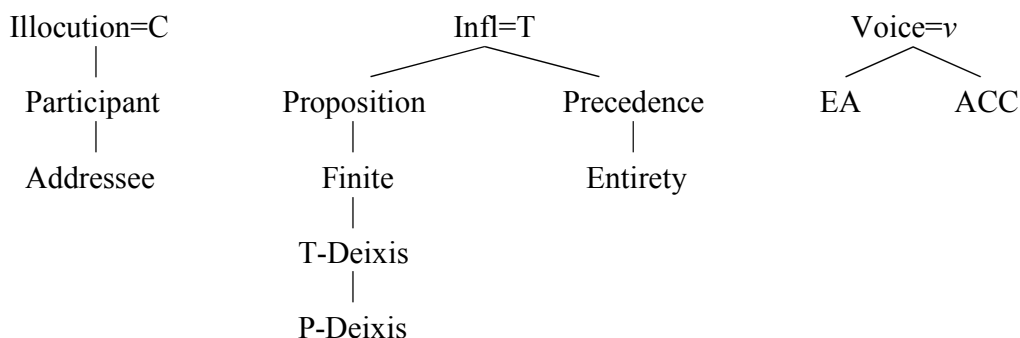
Travis' (1984: 131) original formulation of the HMC is given below.

(39) Head Movement Constraint

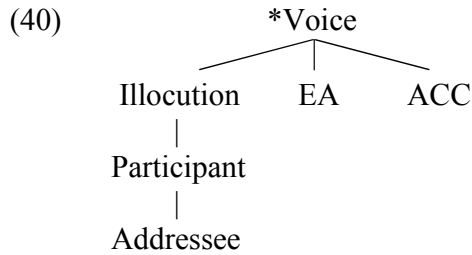
An X^0 may only move into the Y^0 which properly governs it.

Since all heads count as intervenors, that is as potential governors, HM is effectively constrained as movement to the immediately c-commanding head. Attempts have been made to reformulate the HMC in terms of Relativized Minimality (Rizzi 2001); however, bulky machinery was still required to capture the effects of the HMC. Under Minimalist formulations of the HMC it remains as a stipulation. Under HM as feature geometry growth, however, the effects of the HMC are captured without stipulation. Furthermore, as we saw in the previous section, it was able to account for occasional exceptions to the HMC as exemplified by negation in English.

In the proposed theory, the HMC is captured as follows. Complex heads cannot be formed that do not comply with the basic structure of the feature geometry. Consider the following hypothetical functional projections. Under the theory proposed here, v -to-T-to-C movement is possible as the resultant structure forms part of a large, coherent feature geometry.

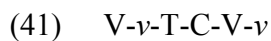


There is no ν -to-C HM because the following structure is not compatible with the overall feature geometry. The existence of the Illocution node dominated by the Voice node entails the existence of the Infl node, which is absent in (40).



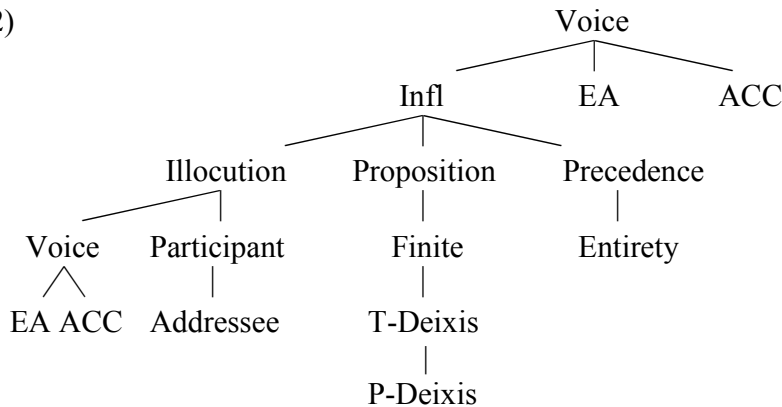
Thus, the basic effects of the HMC are captured under the current theory; however, it is loosened enough to allow exceptions as it is possible to skip a head that is not part of the overall feature geometry.

Another well-known property of HM is that it is clause-bound (see the discussion in Harley 2013a: 114f). While under the original GB formulation of the HMC the observation that HM is clause-bound could be captured by a complex definition of government, such an option is not permitted under the Minimalist formulation of the HMC. Under the proposed theory of HM, the clause-boundedness of HM falls out from that fact that two geometries cannot be combined. HM to C followed by HM to the higher clause. Under traditional HM this derives the following possible structure, assuming movement no higher than ν of the super-ordinate clause.



Under HM as feature growth, the complex head in (41) violates well-formedness constraints on the feature geometry as the same feature would dominate itself. Furthermore, the dominance relation between [Illocution] and the lower [Voice] node is reversed, as is the relation between [Infl] and the lower [Voice] node. I conclude that HM is clause bounded under the feature geometry grown theory of HM.

(42)



This section has investigated further empirical applications of the proposed theory of HM. Specifically, it has looked at noun incorporation and long head movement. It was proposed that noun incorporation typically cannot proceed by HM, but must rather take place as XP movement, as been argued recently in the literature (Barrie & Mathieu 2012; Barrie & Mathieu 2016a; Compton 2013). This section also considered long head movement briefly, arguing that some mechanism other than HM is required to account for cases of this phenomenon. In particular, the discussion here suggested that, at least for the data considered here, long head movement arises by focus movement to the CP layer and proceeds as XP movement. The last sub-section discussed some theoretical advantages of the proposed theory. Namely, HM as feature geometry growth is able to capture the effects of the HMC without stipulation and is also able to capture the observation that HM does not cross clauses.

6 Conclusion and Outstanding Issues

This paper has proposed a novel theory of HM as feature geometry growth. To recapitulate, if we take seriously the idea of feature geometries in syntactic structure, then HM can be thought of simply as the merger of two compatible pieces of feature geometry to form a larger feature geometry. This theory of HM was illustrated with negation in English, where it was shown that

some peculiar properties of the difference between *not* and *n't* fall out naturally from the proposed theory.

The chief problem of HM that was dealt with in this paper is the incompatibility of HM with BPS. Within X-Bar theory, the concept of HM was reasonably well understood, although it was never particularly clear whether HM was substitution or adjunction.⁹ Within BPS, Merge between two syntactic objects necessarily results in an XP. As discussed above, there is simply no way to represent HM in BPS. Rather than complicate BPS with novel set theoretic intricacies to capture the distinction between head merger and phrase merger, I have proposed that HM takes place as feature geometry growth as described here.

Additionally, several problems of HM that have appeared with the advent of the Minimalist Program have also been addressed here. Within a G&B framework, notions such as the HMC and the fact that HM is clause-bound were accounted for by appealing to Government. Government is no longer a valid concept under Minimalist theorizing, so the HMC was effectively treated as primitive of grammar. The theory proposed here captures both of these properties by appealing to well accepted properties of feature geometries.

Given the programmatic nature of this proposal, it is only natural that a number of issues remain outstanding. I consider these here and suggest directions for future research. The first issue is linear order. While the issue of linear order is far from settled, I note that the Linear Correspondence Axiom (LCA), (Kayne 1994) cannot be applied to a feature geometry with ternary and unary branching (but see Guimarães 2008). In some cases the observed word order is predicted by the structure given. Consider the feature geometry in (19). In this configuration, the Voice features (EA and ACC) asymmetrically command the other features. The LCA predicts

⁹ This point was made in a paper by Heidi Harley, but I forget the original source.

that the exponents corresponding to the Voice features appear to the left of the other exponents. This is in fact the observed order of the morphemes. I leave it to future research to investigate how linear order is established inside feature geometries and whether the LCA is even applicable to such structures. Harley (2013c), for instance, shows that affixes may be lexically specified as either a prefix or a suffix. Kayne's (1994) discussion of LCA compliant HM, on the other hand, predicts that all instances of HM result in left-adjunction of the raised head, resulting uniformly in suffixation. If Harley's results are accepted, then we may have to dispense with the notion of the LCA below the level of the head.

Despite some of the loose ends raised at the end of this section, the proposed theory of HM as feature geometry growth offers a novel avenue for our understanding of HM, a concept that has presented a number of thorny problems with the advent of the Minimalist Program. These issues were raised at the beginning of this section, so I do not repeat them here. Nevertheless, the proposed theory solves these problems, while still capturing the intended effects of HM. I leave the loose ends to future research.

References

- Abney, Stephen. 1987. *The English Noun Phrase and its Sentential Aspect*. Cambridge, MA: MIT Ph.D. Dissertation.
- Adger, David. 2003. *Core Syntax* Oxford: Oxford University Press.
- Alboiu, Gabriela. 2000. *The Features of Movement in Romanian*. Winnipeg, MB: University of Manitoba Dissertation.
- Alexiadou, Artemis, Alexiadou Anagnostopoulou & Florian Schäfer. 2015. *External Arguments in Transitivity Alternations: A Layering Approach* Oxford: Oxford University Press.

- Baker, Mark C. 1988. *Incorporation: A Theory of Grammatical Function Changing* Chicago, IL: University of Chicago Press.
- . 2009. Is head movement still needed for noun incorporation? *Lingua* 119.148-65.
- Barrie, Michael. 2015. *Generative Approaches to Noun Incorporation*. 언어 현상과 언어학적 분석 (*Linguistic Phenomena and Linguistic Analysis*), ed. by S.-Y. Shin, H.-S. Wang, H. Chung, J.-Y. Cheon, J.-U. Park, M. Barrie, I. Chung & B.-G. Kang. Seoul: Language and Information Society.
- . 2016. Another Note on Number. *Studies in Generative Grammar* 26.97-113.
- Barrie, Michael & Eric Mathieu. 2012. Head Movement and Noun Incorporation. *Linguistic Inquiry* 43.133-42.
- . 2016a. Noun Incorporation and Phrasal Movement. *Natural Language & Linguistic Theory* 34.1-51.
- Barrie, Michael & Éric Mathieu. 2016b. Head Movement. *The Routledge Handbook of Syntax*, ed. by A. Carnie, D. Siddiqi & Y. Sato, 133-49. Abingdon/New York: Routledge.
- Béjar, Susana. 2003. *Phi-Syntax: A Theory of Agreement*. Toronto: University of Toronto Dissertation.
- Béjar, Susana & Milan Rezac. 2009. Cyclic Agree. *Linguistic Inquiry* 40.35-73.
- Bobaljik, Jonathan David & Samuel Brown. 1997. Interarboreal Operations: Head Movement and the Extension Requirement. *Linguistic Inquiry* 28.345-56.
- Boeckx, Cedric & Sandra Stjepanović. 2001. Head-ing toward PF. *Linguistic Inquiry* 32.345-55.
- Borer, Hagit. 2005. *Structuring Sense, volume 1: In Name Only* Oxford: Oxford University Press.
- Chomsky, Noam. 1994. *Bare Phrase Structure* Cambridge, MA: MIT Department of Linguistics and Philosophy, MITWPL.

- Compton, Richard. 2013. Word-internal XPs and right-headedness in Inuit. Paper presented at the The 31st West Coast Conference on Formal Linguistics, Arizona State University.
- Cowper, Elizabeth. 2005. The Geometry of Interpretable Features: Infl in English and Spanish. *Language* 81.10-46.
- Fanselow, Gisbert. 2003. Münchhausen-Style Head Movement and the Analysis of Verb Second. *UCLA Working Papers in Linguistics* 13.40-76.
- Grodzinsky, Yosef & Lisa Finkel. 1998. The neurology of empty categories. *Journal of Cognitive Neuroscience* 10.281-92.
- Guimarães, Maximiliano. 2008. A note on the strong generative capacity of standard Antisymmetry Theory. *Snippets* 18.5-7.
- Haegeman, Liliane. 1995. *The Syntax of Negation* Cambridge: Cambridge University Press.
- Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. *The View from Building 20: Essays in Linguistic in Honor of Sylvain Bromberger*, ed. by K. Hale & S. J. Keyser, 111-76. Cambridge, MA: MIT Press.
- Harley, Heidi. 2005. How do verbs get their names? Denominal verbs, Manner incorporation and the ontology of verb roots in English. *The Syntax of Aspect: Deriving Thematic and Aspectual Interpretation*, ed. by N. Erteschik-Shir & T. Rapoport, 42-64. Oxford: Oxford University Press.
- . 2013a. Diagnosing Head Movement. *Diagnosing Syntax*, ed. by L.L.-S. Cheng & N. Corver. Oxford: Oxford University Press.
- . 2013b. External arguments and the Mirror Principle: On the distinctness of Voice and v. *Lingua* 125.34-57.

- . 2013c. Getting morphemes in order: Merger, Affixation and Head-movement. *Diagnosing Syntax*, ed. by L. Cheng & N. Corver, 44-74. Oxford: Oxford University Press.
- Harley, Heidi & Rolf Noyer. 2003. Distributed Morphology. *The Second Glot International State of the Article Book: The Latest in Linguistics*, ed. by L. Cheng, 463-96. Berlin, Germany: Mouton de Gruyter.
- Harley, Heidi & Elizabeth Ritter. 2002. Person and number in pronouns: a feature-geometric analysis. *Language* 78.482-526.
- Haugen, Jason D. & Daniel Siddiqi. 2013. Roots and the Derivation. *Linguistic Inquiry* 44.493-517.
- Heim, Johannes, Hermann Keupdijo, Zoe Wai-Man Lam, Adriana Osa-Gómez & Martina Wiltschko. 2014. How to Do Things with Particles. Paper presented to the Annual Meeting of the Canadian Linguistics Association, Brock University, 2014.
- Isac, Daniela. 2012. Decomposing force. *Towards a Biolinguistic Understanding of Grammar*, ed. by A.M. Di Sciullo, 89-115. Amsterdam: John Benjamins.
- Johnson, Kyle. 1988. Verb Raising and *Have*. *McGill Working Papers in Linguistics* 6.156-67.
- Kayne, Richard. 1994. *The Antisymmetry of Syntax* Cambridge, MA: MIT Press.
- . 2009. Antisymmetry and the Lexicon. *Linguistic Variation Yearbook* 8.1-31.
- Koopman, Hilda & Anna Szabolcsi. 2000. *Verbal Complexes* Cambridge, MA: MIT Press.
- Lasnik, Howard. 1995. Verbal Morphology: *Syntactic Structures* meets the Minimalist Program. *Evolution and Revolution in Linguistic Theory: Essays in Honor of Carlos Otero*, ed. by H. Campos & P. Kempchinsky, 251-75. Washington, D.C.: Georgetown University Press.
- Lema, José & Maria-Luisa Rivero. 1990. Long Head Movement: ECP vs. HMC. *Cahiers Linguistiques d'Ottawa* 18.

- Mahajan, Anoop. 2003. Word Order and (Remnant) VP Movement. *Word Order and Scrambling*, ed. by S. Karimi, 217-37. Malden, MA: Blackwell Publishers.
- Marantz, Alec. 2001. Words. Paper presented at the 20th West Coast Conference on Formal Linguistics, University of Southern California.
- Massam, Diane. 2009. Noun Incorporation: Essentials and Extensions. *Language and Linguistics Compass* 3.1076-96.
- Matushansky, Ora. 2006. Head Movement in Linguistic Theory. *Linguistic Inquiry* 37.69-109.
- McGinnis, Martha. 2005. On markedness asymmetries in person and number. *Language* 81.699-718.
- Pollock, Jean-Yves. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20.356-424.
- Potsdam, Eric. 1997. NegP and Subjunctive Complements in English. *Linguistic Inquiry* 28.533-41.
- Preminger, Omer. 2014. *Agreement and Its Failures* Cambridge, MA: MIT Press.
- Pylkkänen, Liina. 2008. *Introducing Arguments* Cambridge, MA: MIT Press.
- Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon: A First-Phase Syntax* Cambridge: Cambridge University Press.
- Rivero, Maria Luisa. 1993. Long Head Movement vs V2 and Null Subjects in Old Romance. *Lingua* 89.217-45.
- Rizzi, Luigi. 1990. *Relativized Minimality* Cambridge, MA: MIT Press.
- . 1997. The Fine Structure of the Left Periphery. *Elements of Grammar: A handbook in generative syntax*, ed. by L. Haegeman, 281-337. Dordrecht: Kluwer Academic Publishers.

- . 2001. Relativized Minimality Effects. *The Handbook of Contemporary Syntactic Theory*, ed. by M. Baltin & C. Collins, 89-110. Malden, MA: Wiley/Blackwell.
- Roberts, Ian. 1994. Two types of head movement in Romance. *Verb Movement*, ed. by D. Lightfoot & N. Hornstein. Cambridge: Cambridge University Press.
- . 2001. Head Movement. *The Handbook of Contemporary Syntactic Theory* ed. by M. Baltin & C. Collins, 113 - 47 Oxford: Blackwell.
- . 2010. *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals* Cambridge, MA: MIT Press.
- . 2011. Head-Movement and the Minimalist Program. *The Oxford Handbook of Linguistic Minimalism*, ed. by C. Boeckx, 195-219. Oxford: Oxford University Press.
- Toyoshima, Takashi. 2000. *Head-To-Spec Movement and Dynamic Economy*: Cornell University Ph.D. Dissertation.
- Travis, Lisa de Mena. 1984. *Parameters and Effects of Word Order Variation*. Cambridge, MA: Massachusetts Institute of Technology Ph.D. Dissertation.
- van der Meulen, Arendina Christina. 2004. *Syntactic movement and comprehension deficits in Broca's aphasia*: Universiteit Leiden Ph.D. Dissertation.
- Williams, Edwin. 1994. A Reinterpretation of Evidence for Verb Movement in French. *Verb Movement*, ed. by D.W. Lightfoot & N. Hornstein, 189-205. Cambridge: Cambridge University Press.
- Wiltschko, Martina. 2009. \checkmark Root Incorporation: Evidence from lexical suffixes in Halkomelem Salish. *Lingua* 119.199-223.
- Wiltschko, Martina & Johannes Heim. to appear. The syntax of confirmationals. A neo-performative analysis. *Outside the Clause*. Amsterdam/Philadelphia: John Benjamins.

Woodbury, Hanni. 2003. Onondaga-English/English-Onondaga Dictionary Toronto, ON:

University of Toronto Press.

Wurmbrand, Susi. 2001. Infinitives: restructuring and clause structure Berlin, New York:

Mouton de Gruyter.