

A morphological theory of head movement¹

Jeffrey K. Parrott
Georgetown University

1. Introduction

Chomsky (1998; 1999) and Boeckx & Stjepanovic (2001) have recently made what Lasnik (2001) calls the "tantalizing suggestion" that head movement is a PF operation. In large part, this suggestion results from a growing recognition that head movement does not fit comfortably into minimalist syntactic theory. The framework based on Chomsky (1995b) treats head movement as a generalized movement operation adjoining X^0 to Y^0 in order to check some feature of Y . Chomsky (1999) points out that it is not clear why a feature can be checked via adjunction, since features are assumed to be checked only in a spec-head relation. He notes that the moved head does not c-command its trace (at least without questionable modifications to the definition of c-command). Chomsky also questions the theoretical motivation for moving X^0 s versus XPs. In the typical case of XP movement, a feature F of a head Y was checked by the corresponding F of an XP in Y 's specifier. For example, the D feature of T was checked by a DP in T 's specifier. Given the mechanism of head movement, it is not obvious why, for example, the D feature of T could not be checked by raising and adjunction of D^0 to T^0 . Conversely, given the standard analysis of V to I movement, it is not obvious why the V feature of I could not be checked by movement of VP to the spec of IP .

Recent theoretical developments make a syntactic treatment of head movement even more problematic. According to Chomsky (1998; 1999), ϕ -feature and Case feature checking takes place in situ, via the operation Agree. Movement occurs only in order to 'check' an idiosyncratic feature, the EPP, that requires certain heads to have a DP specifier.² Within this framework, there is neither a motivation nor a mechanism for head movement. Stipulating some kind of EPP feature for heads resurrects all of the problems mentioned above (and probably more). Moreover, Chomsky's suggestion that phases consisting of vP and CP are Spelled-Out cyclically is a problem for V to I , a core case of head movement. Because a verb in situ is within vP , it should be Spelled-Out--hence frozen and unavailable for movement--before I is introduced into the derivation.

Empirical evidence for treating head movement in PF, while not as plentiful, is also beginning to appear. Both Chomsky (1999) and Boeckx & Stjepanovic (2001) cite a psychological study showing that aphasics react differently to XP and X^0 movement (Grodinsky & Finkel 1998). This finding casts doubt on the view that head movement and XP movement should be treated syntactically, as (near) identical instances of generalized feature checking movement. Rather, the aphasic study suggests a

¹ Thanks to Raffaella Zanuttini for her patience (which I do not deserve) and her many comments during the evolution of this paper. Thanks also to Kleantes Grohmann for many stimulating discussions, about these and other topics. All the shortcomings to be found in what follows are solely my own.

² As Lasnik (2001) points out, the EPP 'feature' really just restates Chomsky's (1981) original formulation of the EPP, that certain phrases must have subjects.

fundamental asymmetry between X^0 and XP movement, consistent with a PF analysis of head movement.

Boeckx & Stjepanovic (2001) present linguistic evidence for PF head movement. They point out that Lasnik's (1999a; 1999b) analysis of pseudo-gapping constructions relies crucially on the optionality of verb movement. According to Lasnik, English objects must raise overtly to Spec AgrOP. To derive pseudo-gapping constructions, Lasnik claims that the verb may either raise (1a), or remain in situ and be elided along with the entire VP at PF ((1b), Boeckx & Stjepanovic's example (2)):

- (1) a. Debbie ate the chocolate, and Kazuko ate_i [_{AgrOP} the cookie_j [_{VP} t_i t_j]].
b. Debbie ate the chocolate, and Kazuko did [_{AgrOP} the cookie_i [_{VP} ~~eat t_i~~]]

But this optionality must be stipulated for pseudo-gapping constructions, since verb movement cannot be optional in other constructions. Moreover, optional movement is not expected in a minimalist syntax, where feature checking movement is either required or prohibited. Boeckx & Stjepanovic conclude that a PF head movement approach provides a natural solution for this problem in Lasnik's analysis of pseudo-gapping. Movement of the object to AgrOP takes place in the syntactic component, and is therefore mandatory. After Spell-Out, the verb may either be moved or elided at PF. This will correctly rule out rule out (2), where the verb has raised and the in-situ object is elided along with VP (their (4)):

- (2) * Debbie got chocolate, and Kazuko got ~~chocolate~~ too.

Thus the literature to date contains both theoretical and empirical arguments for a PF treatment of head movement. However, an explicit theory of PF head movement is still missing. Until one is formulated many crucial questions remain unanswered. Where exactly does PF head movement take place--at some stage in the derivation from syntactic Spell-out to phonology, or in phonology itself? By what mechanisms are heads moved at PF? What are the properties of these mechanisms? Are they restricted in any way, and if so how? What is the motivation or trigger for head movement at PF?

This paper presents a theory of PF head movement that provides specific answers to these questions. Within the framework of Distributed Morphology (DM) (Halle & Marantz 1993), it is claimed that head movement takes place at the level of Morphological Structure (MS), between syntactic Spell-Out and phonology. Heads are moved by an MS operation Move Terminal (MT) which, like all other MS operations, is triggered in order to satisfy a morphological requirement such as affixation. After developing this theory and incorporating some of its consequences into DM, the paper turns to an analysis of V to I, a core case of head movement. It is argued that V to I's robust correlation with the phonological realization of agreement features cannot receive a satisfactory syntactic explanation. Within a MT account of V to I it possible to provide a non-stipulative explanation for the agreement correlation, in terms of independently motivated principles of morphological transparency. These principles make additional predictions beyond head movement, concerning for example the role of morphology in language change.

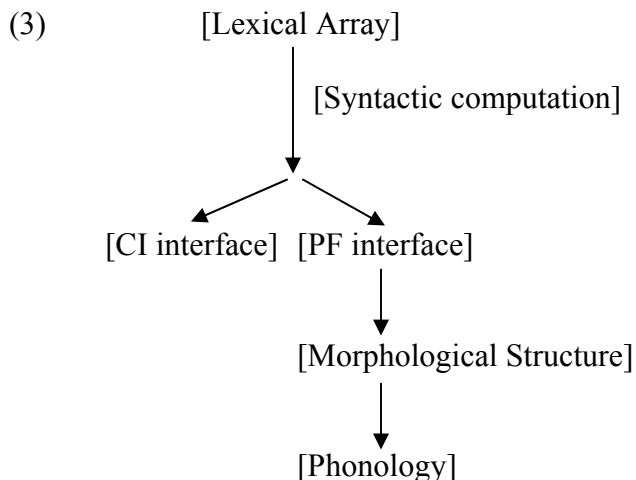
The structure of the paper is as follows. Section 2 briefly outlines some aspects of the theory of DM that are relevant for the analysis to follow. The head movement operation Move Terminal is introduced and discussed in Section 3, and the consequences for the theory of DM are explored in some detail. Sections 4 and 5 turn to the analysis of V to I. After a review of the basic facts of V to I in Section 4.1, Section 4.2 discusses the correlation of V to I and agreement, and Section 4.3 shows why syntactic analyses of V to I have not given a convincing explanation of this correlation. Sections 5.2 and 5.3 present the morphological mechanics of V to I, which are driven by the affixal requirements of Agr and T. It is claimed that when Agr is affixal, verb movement is obligatory. In Section 5.4 it is shown that the addition of independently motivated morphological transparency principals to the preceding theory accounts for the agreement correlation. In the paper's conclusion, it is suggested that the inability of syntax to move heads follows from recent theoretical developments.

2. Distributed Morphology

This section briefly summarizes some of the characteristics of Distributed Morphology that are crucial for the analysis to follow. This summary is based on the version of Distributed Morphology presented in Halle & Marantz (1993). Throughout the paper I argue for certain departures from Halle & Marantz's assumptions; below these departures are merely noted, but are not defended or discussed in detail.

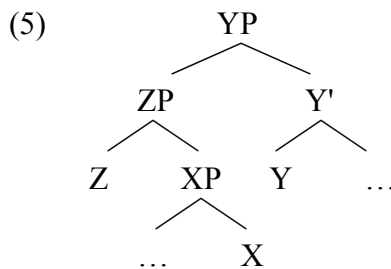
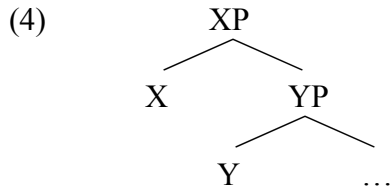
2.1 Morphological Structure

Distributed Morphology (DM) postulates a level of morphology that is independent from syntax, with its own distinct operations and properties. This level, called Morphological Structure (MS), is located between the syntactic computation and the phonological component:

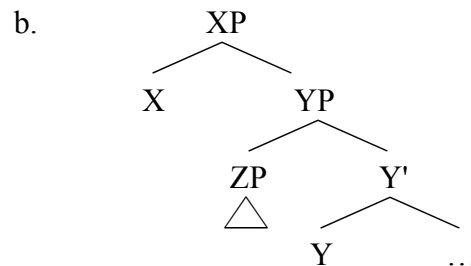
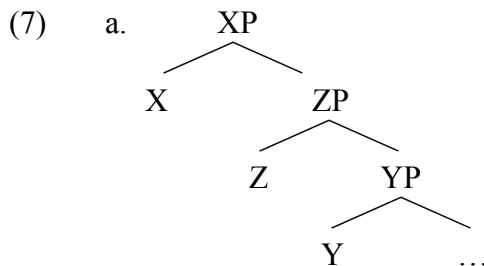
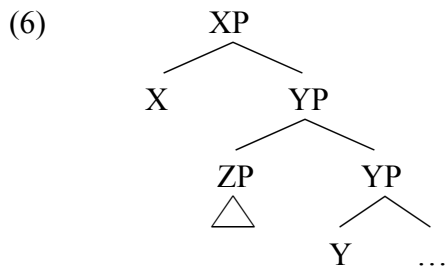


Like syntactic representations, MS representations consist of hierarchically structured bundles of features, usually represented in tree notation. However, representations at MS are also linearized, unlike syntactic representations. As a result,

one relation defined on morphological structures is Adjacency, which combines linear and hierarchical information. X and Y are Adjacent if they are next to each other in the tree, as when X immediately dominates Y and no non-adjunct (see below) intervenes between X and Y. Because Adjacency is linear, it can be defined either to the left or the right. In (4), X is left Adjacent to Y (or, Y is right Adjacent to X). X and Y are not Adjacent in (5) (although Y and ZP are):



Adjacency is not disrupted by intervening adjuncts, such as adverbs and floating quantifiers. Heads, such as negation, and argument XPs, such as objects, disrupt adjacency. X and Y remain adjacent in (6), but are not Adjacent in (7a-b):



Evidence for the adjunct/non-adjunct asymmetry in the disruption of Adjacency comes from independent analyses of diverse phenomena, for example English do-support and verbal inflection (Bobaljik 1995; Halle & Marantz 1993); Holmberg's Generalization and object shift (Bobaljik 1995); and case morphology in English and cross-linguistically

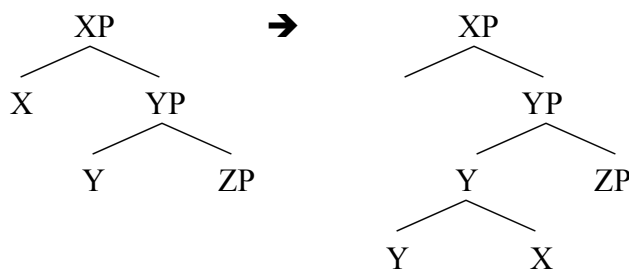
(Parrott 2001). The reason for the asymmetry and its effect on the Adjacency is not currently understood.³

2.2 Morphological Operations

Heads, also called terminals, and their features are subject to morphological operations at MS. These operations include Insertion, Fission, Merger, and Fusion. Only the latter two are discussed here, being relevant for what follows. In Section 3 below I propose the addition of an additional operation, Move Terminal (MT).

Merger and Fusion differ in their effects on terminal structure, their consequences for Vocabulary Insertion (VI), and the structural conditions under which they apply. Merger combines some head X and an adjacent head Y under a single terminal Y. Merger is taken to reverse the order of terminals (Bobaljik 1995; Halle & Marantz 1993). However, in Section 3.2 it is argued that the ordering of terminals after Merger is a consequence of the affixation requirements of terminals, and not a property of Merger itself. X and Y remain distinct terminals under Y. VI requires two separate Vocabulary Entries (VEs) to insert phonological material at X and Y. A language whose terminals are joined primarily by Merger will tend to have an agglutinative type morphology, where the mapping between morphosyntactic features and their phonological realization is one to one. Merger is schematized in (8):

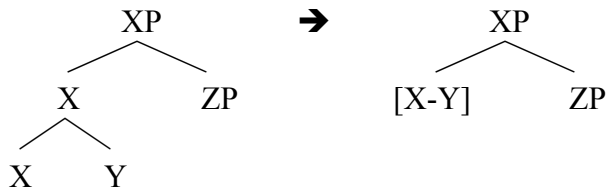
(8) Merger(X, Y)



Fusion combines two sister terminals X and Y into a single terminal [X-Y] containing the features of both X and Y. At VI, a single VE will insert phonological features at [X-Y]. Languages whose terminals are joined primarily by Fusion will have a fusional or analytic type morphology, where the mapping between morpho-syntactic features and their phonological realization is many to one. A structural condition on Fusion is that X and Y must be sisters under a single node. Therefore Fusion requires the prior application of some operation that creates sister terminals--either MT (i.e., head movement) or Merger. Fusion is schematized below:

³ Although it seems in keeping with adjuncts' generally unusual behavior.

(9) Fusion(X, Y)



2.3 Vocabulary Insertion

A crucial and characteristic claim of the theory of DM is that phonological features are not present prior to or during the syntactic derivation. Syntax operates only on bundles of syntactic and semantic features. When the syntactic derivation is complete, these bundles are subjected to further operations at MS. Phonological features are only introduced at MS, by an operation called Vocabulary Insertion (VI). According to Halle & Marantz (1993) VI applies to every terminal in the entire MS tree, after all the primary morphological operations have been performed. Then a special class of operations, called Readjustment Rules, may apply to the inserted phonological material.⁴

Vocabulary Entries (VEs) specify which phonological features are to be inserted at which terminals during VI. VEs have three parts. The substantive features of a VE identify a terminal where the phonological features of the VE will be inserted. Substantive features are at least a subset of the features of the insertion terminal. The contextual features of a VE specify any contextual conditions that must hold on the terminal identified in the substantive features. The phonological features of a VE are inserted at the specified terminal. Every VE must have at least one substantive feature; otherwise it has no terminal for insertion. A VE may or may not have contextual features, and its phonological features can consist of \emptyset . A VE is schematized in (10). The substantive features are on the left hand side of the arrow in square brackets; the phonological features are on the right hand side of the arrow in slash brackets; and the contextual features are to the right of the long slash. The VE in (10) inserts the phonological features /phono/ at a terminal containing the features [+F1, -F2], so long as that terminal is left adjacent to X:

(10) [+F1, -F2] \Leftrightarrow /phono/ / __ (Adjacent) X

The phonological features of a VE are inserted at some terminal if i.) no features of the insertion terminal contradict the substantive features of the VE, and ii.) the context of the insertion node is the same as that specified in the contextual features of the VE. VI is governed by the Elsewhere Principle. VEs that are more specific, in either their substantive or contextual featural descriptions, insert their features phonological features before more general VEs.

⁴ Below it is suggested that VI might be thought of as a kind of Spell-Out for MS, delivering phonological features to the phonological component. This would have the consequence of removing Readjustment Rules from MS. Other consequences might also follow, such as cyclic application of VI. Obviously, the entire line of thought constitutes a major departure from Halle & Marantz (1993).

Note that there is an interaction between VEs and morphological operations. For example, whether Fusion applies depends on a language's VEs. Suppose some language has separate VEs for X and Y:

- (11) a. [+X] ⇔ /phono/
b. [+Y] ⇔ /phono/

In such a language Merger (or MT) could apply to X and Y since X and Y will remain distinct terminals. Fusion must not apply to X and Y, or VI will fail. Now suppose that a language has only a single VE, whose substantive features include both X and Y:

- (12) [+X, +Y] ⇔ /phono/

In such a language Fusion must apply to X and Y at MS, or else VI will fail. Merger will not suffice. This point becomes more significant in Section 5.4.1 below.

2.4 Affixation

In DM, an affix may be defined as some terminal X that requires another terminal Y as a sister. Either Merger or MT will accomplish this.⁵ The terminal Y will be referred to as the 'host' terminal. In most cases the host terminal is specified. English T, for example, is an affix that requires Merger with a V element. However, in certain cases an affix may not require any specific host terminal. For example, certain Slavic clitics are satisfied by any left adjacent element (see Section 3.3). In what follows, an affix with a specified host terminal Y is referred to as a Y-affix.

3. Move Terminal

I propose that head movement⁶ is the result of a morphological operation called Move Terminal (MT). MT, like Insertion, Fission, Fusion, and Merger, applies at MS and is triggered by language specific morphological requirements. In this section we consider the specific properties of MT, and examine some of the consequences its inclusion has for the DM theory as a whole.

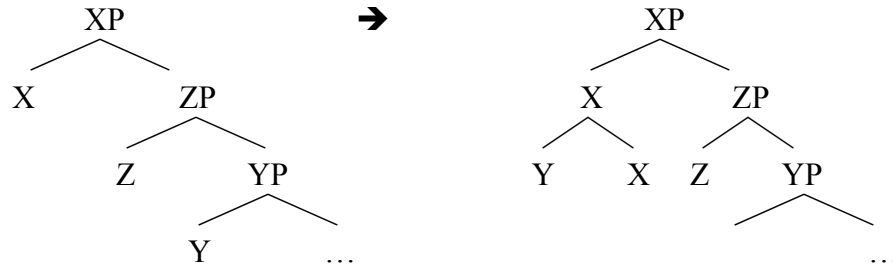
In essence, MT is Merger at a distance. Merger may apply only to Adjacent terminals, but MT has no such structural condition, applying to non-Adjacent as well as Adjacent terminals. MT is like Merger in that it combines two terminals X and Y under a single terminal; X and Y remain distinct, and require two VEs at VI. It is argued below that the ordering of terminals after application of Merger and MT results from the particular affixal requirements of terminals. Thus terminal reordering is not a property of the operations themselves, contrary to Halle & Marantz (1993). There are two differences

⁵ It is not obvious whether Fusion can satisfy an affixation requirement. The matter is obscured because the structural requirements of Fusion are such that it must be 'fed' by either Merger or MT. What is needed is a case in which Merger or MT create a morphological structure such as [X [Z [Z Y]]], where X is not sister to Y. If X's Y-affixation requirement is subsequently satisfied by Fusion that creates [X, Y, Z], then we will know that Fusion satisfies affixation requirements. Meanwhile, in the absence of this kind of evidence, we continue to assume that affixation requirements are met by Merger or MT.

⁶ I set aside completely, as does Chomsky (1999), the issue of incorporation (Baker 1988).

between Merger and MT. First, there is no adjacency condition on the terminals affected by MT, and second, MT never lowers terminals. MT is schematized in (13):

(13) Move Terminal(Y, X)



Since MT is a morphological operation, it must be triggered by some morphological requirement at MS. In the case of V to I, the trigger for MT argued to be the affixation requirement of the higher terminal Agr. Affixation requirements are probably the trigger in most cases of head movement, and below I claim that an affixation requirement also causes Long Head Movement. However, there may be other morphological requirements that can trigger MT. V2 might be such a case, where the requirement is that tense features must be in the second position of a clause. V2 will not be explored here. See Anderson (1992; 1993; 1996) and Boskovic (1995) for discussion of second position requirements, on different theoretical assumptions.

3.1 Lowering

A question about MT is why it cannot lower a terminal rather than just raising it, so that the affixal requirements of a lower terminal are met by MT of a higher host terminal. Put another way, can there be lowering without adjacency? In order to prohibit MT from lowering, we need two assumptions: (a) MS is derivational, and (b) the affixal requirements of a terminal must be met immediately after its introduction into the MS representation, or at least within some limited number of derivational cycles. On assumptions (a) and (b), a terminal could not be introduced into an MS derivation and then wait for its affixal requirements to be met by a host terminal introduced later in the derivation. Lowering without adjacency would thus be ruled out. We examine both assumptions in turn.

Assumption (a) is not implausible, if recent theories of syntax are on the right track. A representation is delivered to MS upon the application of Spell-Out in the syntax. If Spell-Out takes place only once during the syntactic computation, then MS receives a single, complete representation. In this case there would be no reason to expect a derivational MS. Morphological operations would be free to act globally, upon any part of the MS representation. A single application of Spell-Out was assumed in earlier versions of syntactic theory (e.g. Chomsky 1994), and this was the syntactic theory that formed the background to Distributed Morphology circa 1993. Thus, Halle & Marantz (1993) assume that morphological operations have available the entire MS representation. More recent theories of syntax, however, have for independent reasons postulated that Spell-Out applies multiple times throughout the derivational syntactic computation

(Chomsky 1998; Chomsky 1999; Uriagereka 2000). If this is correct, than MS also receives its representation derivationally, in multiple pieces. At any point in a derivation, MS operations have available only those subparts of the tree that are already Spelled-Out by syntax. MS and its operations would therefore apply derivationally.

Since the pieces delivered to MS would correspond to the pieces Spelled-Out by syntax, the question is just what kind of pieces the syntax Spells-Out. Does the syntax Spell-Out according to the algorithm proposed by Uriagereka (2000), or does it Spell-Out the 'phases' of vP and CP, as proposed by Chomsky (1998; 1999) Perhaps syntax Spells-Out a different kind of piece, as yet undetermined? The answer to this question will have real implications for the analysis of morphology, if anything like the proposals being entertained in this paper is correct; conversely, the analysis of morphology might be able to shed light on the multiple Spell-Out question in syntax.

Assumption (b) is more problematic. Even if MS is constructed derivationally, it does not follow that affixal requirements of terminals be satisfied immediately. The matter seems to hinge on how we conceive of the 'spell-out' from MS to phonology. When does VI apply at MS, and how are these phonological features delivered to the phonological component? Immediate satisfaction of affixal requirements might follow if MS spell-out (VI and delivery of phonological features to phonology) applies multiply, throughout the MS derivation, as in the syntax. Then, a terminal which failed to have its affixal requirements satisfied immediately would be spelled-out to phonology as such, causing a 'crash' in the phonological component.⁷

But it is not all clear that any such 'multiple spell-out' condition holds at MS. There is no obvious theoretical motivation for it. Furthermore, MS multiple spell-out would have to be somehow limited in its application, in order to avoid real problems in the derivation of morphological structures. For example, it could not be the case that host terminals are spelled out before the affixal terminals that require them are later introduced into the MS derivation. Similarly, we would not want to rule out Merger of a lower affixal terminal with an adjacent, higher host terminal. This kind of Merger seems necessary, as for example in the case of Korean negation (Section 3.2 below). Limitations on syntactic Spell-Out have been defined in various ways, as noted above. These limitations make sense in syntax, but they many not have an analogue in MS. Moreover, limitations on multiple Spell-Out in syntax have been proposed (at least partially) on the basis of empirical evidence. Evidence for these kinds of limitations on spell-out in MS is lacking, unless the (presumed) lack of lowering without adjacency is itself to be considered evidence.

The issues raised by assumptions (a) and (b) will have to be addressed in the future. In what follows, some version of these assumptions, and the consequent ban on MT lowering, is simply assumed.

3.2 Terminal (re)ordering

Next we address the assumption that Merger, thus MT, reorders the terminals they affect, so that the raised/lowered terminal always appears to the left of the target terminal. This

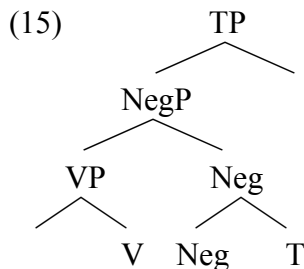
⁷ Notice that this is very different from a crash at the PF or LF interface, where certain features are simply illegible. Here, the features would be legible in phonological component, but the resulting phonological structure would fail to meet some well formedness condition. The issues cannot be explored here.

assumption is unnecessary if the order of terminals after Merger and MT is simply the consequence of a directionality parameter in affixal requirements. Terminals are either reordered or not, depending on whether the affixal terminal that triggers the operation requires its host to the left or to the right--that is, depending on whether the affixal terminal is a suffix or a prefix.⁸

Of course, affixal requirements are terminal specific and need not be uniform for all the terminals in a language. And thus it is no surprise that most languages have both prefixes and suffixes. It would be difficult to derive this fact if terminal ordering were set uniformly for Merger and MT. Consider Korean, where Tense is a verbal suffix, but negation is a prefix. In (14a), negation prefixes to the verb and T suffixes to the verb. In (14b), negation prefixes to T. The verb is spelled out as a non-finite participial form, and a light verb *hata* must be inserted to satisfy T's V-affixal requirement, reminiscent of do-support in English.⁹

- (14) a. Eric-ka Kenny-lul an-chuki-uss-ta.
 SUB OBJ NEG *kill* PAST DECL
 "Eric didn't kill Kenny"
 b. Eric-ka Kenny-lul chuki-ji anha-ss-ta.
 SUB OBJ *kill* PART NEG(+LIGHT VERB) PAST DECL
 "Eric didn't kill Kenny"

We assume that Korean has two negation terminals, one a V-affix (14a) and one a T affix (14b). Both negation terminals require a host to the right. (14b) can be derived as follows.¹⁰ When T is introduced into the MS derivation, it Merges with adjacent T-affixal Neg.¹¹ Because Neg is a prefix, the ordering of Neg and T must not be reversed, if Neg is to have its T host to the right as required:



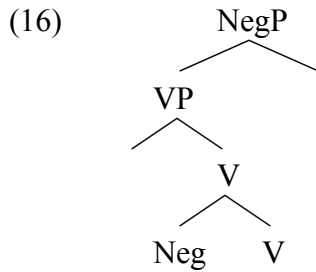
Turning to (14a), when V-affixal Neg is introduced into the MS derivation it Merges with adjacent host terminal V. But notice that here, ordering of Terminals must be reversed, if Neg is to have a host on the right as required:

⁸ It's not clear what this would mean for infixation.

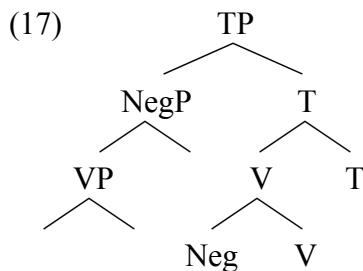
⁹ Kenny is a character who dies in every episode of the television series 'South Park'. He has been killed by the character Eric only once to my knowledge. Eric, who was suffering from the side effects of an improperly prescribed anti-ADD medication at the time, cannot be held entirely responsible.

¹⁰ Details, such as *hata* insertion, are omitted here and below.

¹¹ Notice that here, Merger lowers the host terminal to the affixal terminal.



When V-affixal T is introduced, MT raises V to T. Now terminal ordering must not be reversed, if T is to have its V host to the left as required:



In order to derive the morphology of negation in Korean, as well as the existence of both prefixes and suffixes generally, it is necessary that Merger and MT may either reorder or not reorder terminals. We conclude that the ordering of terminals after Merger and MT is a consequence of the affixation requirements of terminals, and not a property of the operations themselves.

3.3 Locality

A final issue concerns locality. It was first observed by Travis (1984) that head movement seems to obey locality restrictions. Locality of head movement has received various syntactic explanations, but how can it be expressed in a system where head movement is a morphological operation? Below we examine briefly some of the facts pertaining to locality of head movement, and reinterpret them according to the present morphological approach. From these facts we can formulate a general economy condition holding on affixation at MS. No locality property need be ascribed to MT itself.

Some of the first evidence for locality of head movement comes from English (Travis 1984). Modals and auxiliaries, which are presumably heads, are moved to sentence initial position for question formation:

- (18) a. Would Eric kill Kenny?
 b. Has Eric killed Kenny?

In sentences with a sequence of modals and auxiliaries, however, only the first head can move for question formation. Moving any of the other heads results in unacceptability:

- (19) a. Eric would have killed Kenny.

- b. Would Eric have killed Kenny?
 (20) a. * Have Eric would killed Kenny?
 b. * Killed Eric would have Kenny?

To account for the locality of head movement observed in (18-20), Travis proposed the syntactic Head Movement Constraint (HMC), which states that a head cannot move across another head. The HMC was later subsumed under the ECP and Relativized Minimality, according to which no element, head or XP, can move across another element of the same type (Rizzi 1990).

Expanding the data sample, we see that locality conditions on head movement are more complicated. Long Head Movement (LHM), a phenomena found in many Slavic languages and some varieties of Old Romance, is a clear counter example to the HMC, and a problem for the most straightforward formulation of Relativized Minimality. First discussed by Lema & Rivero (1989) and Rivero (1990; 1991), LHM involves the movement of one head over another head. Canonical LHM raises a participial verb out of the VP, crossing over a clitic auxiliary. The following Bulgarian examples are presented by Rivero (1991). (21) shows the position of the verb and auxiliary with an overt subject, and (22) illustrates LHM with a *pro* subject:

- (21) Petur e pročel knjigata.
 has_{CL} read_{PTC} the-book
 "Peter has read the book."
 (22) Pročel e knjigata.
 read_{PTC} has_{CL} the-book
 "(He) has read the book."

LHM is triggered by the clitic auxiliary, which must always have a host to its left:

- (23) * E pročel knjigata.
 has_{CL} read_{PTC} the-book

LHM does not occur when an overt subject to the left of the clitic auxiliary; however, LHM is possible if the overt subject is to the right (example from Wilder & Damir 1994):

- (24) Čel e Petur knjigata.
 read_{PTC} has_{CL} the-book

Where there is sequence of two auxiliaries, the second auxiliary moves over the first, clitic auxiliary. The following examples, from Croatian, are presented by Wilder & Damir (1994):

- (25) Marija i Ivan su bili čitali knjigu.
 M. and I. be_{3rdPL, CL} be_{PTC} read_{PTC} book
 "Maria and Ivan had read the book."

- (26) Bili su čitali knjigu.
be_{PTC} be_{3rdPL, CL} read_{PTC} book
 "(They) had read the book."

Now the main verb participial cannot be moved:

- (27) * Čitali su bili knjigu.
read_{PTC} be_{3rdPL, CL} be_{PTC} book

Syntactic attempts to account for the locality effects exhibited by LHM include modifications of Relativized Minimality (Lema & Rivero 1989; Rivero 1990; Rivero 1991), and the claim that the participial does not in fact cross the auxiliary head, but rather is adjoined to it (or both to C) (Boskovic 1995; Wilder & Damir 1994).¹² But none of these analyses can be straightforwardly incorporated into a morphological treatment of head movement, to which we now turn for the two cases presented.

First, consider the English data in (18-20). A full analysis of movement to C is unfortunately beyond the scope of this paper. Here it must merely be stated, without the necessary argumentation, that English root interrogative C is a T-affixal terminal. Therefore in (18-20) MT raises T to C in order to satisfy C's affixal requirements.

Next, consider the LHM data in (21-27). Again space reasons prohibit the presentation of a detailed morphological analysis, but suppose something roughly like the following. The clitic auxiliary is an affixal terminal. The identity of its host is not specified, but it must occur to the left of the clitic terminal.

Syntax may deliver to MS a representation where some element is left adjacent to the clitic, for instance a subject DP (21). Crucially, syntax is not motivated by any needs of the clitic in this case; the element occurs where it does for completely independent reasons. At MS, the clitic's affixal requirement can now be satisfied by the element that happens to be on its left. Nothing need raise from below. Note that there are significant questions regarding the ability of a phrase to satisfy affixal requirements. Can the clitic, or any terminal, be Merged to a phrase? Or, does MT raise the clitic to a terminal inside the phrase? If so, why is that 'closer' than raising something from below? These questions will have to be addressed in the future, but are deferred here.

LHM occurs precisely when no element occurs to the left of the clitic, and the reason for this is straightforward on the morphological account. Syntax has delivered to MS a representation with no element to the left of the clitic terminal. Again, the operations of syntax are completely independent of later morphological requirements. With nothing to its left, the affixal requirement of the clitic terminal triggers MT, which raises the participial verb to affix to the left of the clitic. Here we face another significant question: why does *pro* not suffice to satisfy the clitic's affixal requirement (23)? This question too must be deferred here.

¹² All of these analyses take for granted that the 'trigger' for LHM is the need to support the clitic auxiliary. But this raises the problem of why such a connection between morpho-phonology and syntax should exist. If head movement, like XP movement, is driven by abstract syntactic features, why does syntax care whether or how these features are pronounced in a later component PF? This general problem will be discussed below with respect to V to I.

Having reviewed some facts, and interpreted them in the present framework, we are now in a position to determine the observed locality effects on head movement can be accounted for in a morphology based analysis. First of all, recall that MT, like all other MS operations, is triggered by some morphological requirement. If this requirement is not met, then MT should result in unacceptability. This is quite independent of locality: all morphological operations, local or not, must satisfy a morphological requirement to be legitimate. On this view, the badness of (20a-b) in English does not involve head movement locality effects at all. Neither the perfect auxiliary *have* (20a) nor the participial verb *killed* (20b) are finite, and so in the current framework neither is affixed to T. Thus movement of these heads to C does not satisfy C's T affixal requirement.

Turning to the LHM facts, both the second auxiliary and the participial verb are suitable hosts: either will satisfy the clitic's affixal requirement, as we can see from (22) and (26). This means that the facts in (26-27) are due to locality, and not the result of an unsuitable host for the affixal terminal. This locality effect can be expressed straightforwardly by the condition in (28):

- (28) Economic Affixation Condition (EAC)
 An affixal requirement is satisfied by the closest suitable terminal.

Because the EAC holds at MS, the measurement of closeness can make use both linear and hierarchical relations. Thus the closest possible relation between terminals is sisterhood, the very relation under which an affixal requirement may be satisfied with no further morphological operations. The next closest relation between terminals is Adjacency, as defined above. Under Adjacency, a morphological operation may be satisfied by the application of Merger. Beyond the relations of sisterhood and Adjacency, closeness may be measured simply in terms of intervening terminals:

- (29) Closeness of terminals at MS:
 A terminal X is closer to a terminal Y than to a terminal Z, if fewer terminals intervene between X and Y than between X and Z.

The LHM locality facts follow immediately. In (25), no terminals intervene between the clitic and the second auxiliary (the two may even be adjacent). However, at least one terminal, the second auxiliary, intervenes between the clitic and the participial verb. Therefore by (29) the second auxiliary is closer, and it satisfies the clitic's affixal requirement according to the AEC.

3.4 Summary

Summarizing the conclusions of this section, head movement is the result of an MS operation MT. MT is triggered by affixal requirements of terminals. Like Merger, MT combines two terminals, which remain distinct under a single terminal. The order of terminals after MT (and Merger) depends on the affixal requirement of the terminal that triggered the operation. MT may apply to Adjacent or non-Adjacent terminals; however MT cannot lower a terminal.

The introduction of MT required modifications to Halle and Marantz's (1993) version of the theory of DM. MS is derivational, and affixal requirements must be satisfied as soon as possible during the MS derivation. Affixal requirements are met by the closest suitable terminal (the EAC), where closeness is calculated both in structural terms and by the number of intervening terminals.

4. V to I

In the preceding sections, we sketched the beginnings of a morphological theory of head movement. In what follows, this theory is applied, in more detail, to the core case V to I movement. V to I has received extensive syntactic analysis, but its correlation with agreement has not received any adequate explanation. Below it will be demonstrated that a morphological treatment of V to I allows a natural explanation of the agreement correlation, without recourse to the kinds of stipulation that characterize syntactic accounts.

4.1 Basic facts

V to I is among the best known phenomena in linguistic theory, having been extensively discussed in the literature for two decades. In certain languages, for example French, finite verbs appear to the left of negation, VP adverbs, and floated quantifiers. In other languages, for example English, finite verbs appear to the right of these elements (examples from Roberts (1999)):

Negation

- (30) a. Jean (ne) mange pas du chocolat.
b. * Jean (ne) pas mange du chocolat.
- (31) a. * John eats not chocolate.
b. John does not eat chocolate.

Adverbs

- (32) a. Jean embrasse souvent Marie.
b. * Jean embrasse souvent Marie.
- (33) a. * John kisses often Mary.
b. John often kisses Mary.

Floating quantifiers

- (34) a. Les enfants mangent tous le chocolat.
b. * Les enfants tous mangent le chocolat.
- (35) a. * The children eat all chocolate.
b. The children all eat chocolate.

Negation, VP adverbs, and floated quantifiers are assumed to be located at the edge of VP, but below inflection. Therefore, since Emonds (1978) the cross-linguistic difference illustrated in (30-35) has been attributed to verb movement (the analysis was extensively elaborated in Pollack (1989)). In raising languages, like French, the verb

moves from its generated position in VP to some inflectional position above VP. In non-raising languages, like English, verb movement does not occur, and the verb remains in situ within VP:

- (36) Raising: [IP V_i - I [VP t_i]]
 (37) Non-raising: [IP I [VP V]]

It is claimed (Roberts 1999) that movement to C depends on prior movement to I. Although French verbs raise to C for question formation, English main verbs may not (examples from Roberts (1999)):

- (38) a. Voit-il le cheval?
 (39) b. * Sees he the horse?

However, it is not the case that only raising languages may have raising to C. Danish, like English, is a non-raising language (examples from Vikner (1997; 2001)):¹³

- (40) At Johan ofte spiser tomatere...
that John often eats tomatoes
 "That John often eats tomatoes..."
 (41) ...fordi du ikke svarede på hans brev.
because you not replied to his letter
 "... because you didn't reply to his letter."

Unlike English, Danish main verbs raise to C for question formation (example from Vikner (2001)):

- (42) Svarede du nogensinde på hans brev?
replied you ever to his letter
 "Did you ever reply to his letter?"

English is unusual in that its verb system is mixed. In contrast to main verbs, English auxiliaries and the copula *be* all raise to I:¹⁴

- (43) a. Eric has not killed Kenny.
 b. Those bastards have brutally killed Kenny.
 c. Those bastards have all killed Kenny.
 (44) a. Kenny is not dead yet.
 b. Kenny is often dead.
 c. The three boys are all alive.

English auxiliaries and the copula also move to C for question formation:

¹³ V2 obscures V to I in Danish main clauses, but V2 does not apply in embedded finite clauses.

¹⁴ Every time Kenny is killed, the character Stan exclaims: "Oh my god! They killed Kenny! You bastards!"

- (45) a. Have those bastards killed Kenny again?
b. Is Kenny dead again?

For more detailed presentations of V to I, especially as concerns V's landing site and the split IP hypothesis, see for example Pollack (1989), Belletti (1990), and contributions to Hornstein & Lightfoot (1994).

4.2 The agreement correlation

V to I correlates with 'richness' of verbal agreement morphology, an important and central fact whose first observation is attributed to Roberts (1985). Typologically, modern Icelandic, Yiddish, and French have 'rich' verbal agreement morphology and verb movement. In modern English, Danish, and Faroese agreement is poor and verbs do not raise (Vikner 1997). Historically, both English and Danish lost verb movement soon after independent phonological changes largely eliminated verbal agreement morphology. Creole languages display the same pattern, as discussed by DeGraff (1997). Haitian Creole, for example, has no agreement morphology and also no verb movement, despite its French superstrate. French based Louisiana Creole also lacks agreement and verb raising. Mesolectal varieties of this creole, however, have more agreement morphology, accompanied by verb movement.

At first glance then, we might state the agreement correlation as follows:

- (46) The agreement correlation (first):
If agreement morphology is 'rich', then verbs raise;
If agreement morphology is 'poor', then verbs do not raise.

We can see immediately that the second part of (46) is incorrect: verb movement is possible in the absence of 'rich' agreement. Roberts (1999) mentions several cases. Northern and Scottish dialects of English ca. 1300-1500 had 'poor' agreement morphology and verb movement. The same is true of the Modern Kronoby dialect of Swedish.¹⁵ Southern (later 'standard') dialects of English retained verb movement for "at least 75 years" (p. 292) after losing 'rich' agreement; in Danish this gap lasted for about 200 years.¹⁶

In light of these facts, Roberts proposes a strictly one-way agreement correlation, so that:

- (47) The agreement correlation (second):
If agreement morphology is 'rich', then verbs raise.

However, it is evident that although verb movement is possible where agreement is 'poor', this state of affairs is quite uncommon. The examples mentioned by Roberts exhaust the cases to be found in the literature, to my knowledge. Historically, the trend is

¹⁵ And possibly Dutch and Afrikaans, although the existence of verb movement in these two languages is controversial.

¹⁶ Additionally, verb raising in embedded non-finite clauses occurs in Italian.

clearly for loss of verb movement to follow loss of 'rich' agreement. Furthermore, to my knowledge there are no known cases where a language with 'poor' agreement changes from non-verb raising to verb raising. The issue is discussed further below; meanwhile we may state the tendency:

- (48) The agreement correlation (third):
If agreement morphology is 'rich', then verbs raise;
If agreement morphology is 'poor', then verbs *tend* not to raise.

We have not yet defined 'richness' of agreement. Section 5.4 below presents an approach to this notion in terms of morphological transparency, following Emonds (1986) and Parrott (2001). The approach is somewhat different from previous attempts to define 'richness', such as Roberts (1992), Rohrbacher (1994) and Vikner (1997). Unfortunately, a proper comparison is beyond the scope of this paper. Postponing detailed discussion, we merely substitute morphological transparency for 'richness'. This gives the final statement of the agreement correlation:

- (49) The agreement correlation (final):
If agreement is morphologically transparent, then verbs raise;
If agreement is not morphologically transparent, then verbs *tend* not to raise.

4.3 The syntactic analysis

Although the agreement correlation is a central fact about V to I, it has not received an adequate explanation in current syntactic theory. Chomsky's (1995b) analysis of V to I, which has been generally adopted in the literature (e.g., Roberts 1999), holds that head movement is an instance of generalized movement. Movement, whether of heads or XPs, is motivated by the need to check abstract syntactic features. In a given language these features may either be pronounced or not, but the features themselves are part of UG and thus present in every language. Verb movement, then, is caused by a V feature on I. When the V feature is strong, as in French, verb movement is overt. When the feature is weak, as in English, movement takes place covertly at LF.

To account for the agreement correlation on this analysis, 'rich' agreement must be somehow connected with feature strength, so that the V feature of I is strong when agreement is 'rich' (and tends to be weak when agreement is 'poor'). The problem is that it is not clear how such a connection can be made. 'Richness' of agreement morphology really means the actual *pronunciation* of agreement morphemes on the verb. But the abstract syntactic features that correspond to pronounced agreement morphemes, including the V feature of I that causes movement, are present in the syntax whether they are pronounced or not. In order to explain the agreement correlation, then, a connection must be made between the syntactic and phonological components. The syntactic value of the V feature has to depend on its value at PF.

How can this be accomplished? We can dismiss the idea that syntax (perhaps the V feature itself?) 'looks ahead' to PF. The only alternative is to postulate an inherent lexical connection between syntactic feature strength and phonological value. Thus lexical items with strong features must have some phonological features. But there are

several serious problems with this account. The most obvious is that we have merely restated the same problem we began with. Why should there exist such a connection between the strength of a feature and its phonological value, even if this connection is expressed in the lexicon? A second problem is that only in the case of V to I movement is it true that strong features must be associated with phonological features. No phonological features accompany the strong feature that causes wh-movement, for example. Finally, it is well known that V to I is correlated not with just any phonological features, but rather with the phonological realization of a sufficient number of agreement distinctions. Thus French, with agreement 3 distinctions in each tense, has verb movement, but English, with only one distinction in one tense, lacks verb movement. This means that a lexical relation between strength and phonological features cannot be so straightforward as posited above. Its proper reformulation is far from obvious. Nor is it clear what mechanisms the lexicon would employ to keep track of the number and type of expressed agreement distinctions, phonologically expressed, that are required to accompany a strong feature.

It seems that the current syntactic analysis of head movement cannot explain the agreement correlation. The connection between V feature strength and phonological realizations of agreement follows from nothing, and must be stipulated. For example, Roberts (1999) states that:

- (50) If there is verbal agreement marking of the relevant type, then I has a strong V-feature. (p. 292)

It is not clear whether any modified syntactic analysis of V to I could avoid stipulations like (50). Certainly Chomsky's (1998; 1999) latest revisions of syntactic theory will not solve the problem, or any of the other problems associated with treating head movement in the syntax, as he notes. But an adequate explanation of the agreement correlation should be part of any theory of V to I. The morphological analysis offered below provides one.

5. A morphological analysis of V to I

This section presents a morphological analysis of V to I, within the modified DM framework outlined above. The first part of the analysis consists of the claim that verb movement is obligatory when Agr is a T-affix. When Agr is non-affixal, raising is possible but tends not to occur. This follows from the relative markedness of morphological operations, which is related to the EAC. The second part of the analysis introduces independently motivated morphological transparency principles. As a result of these, Agr can be an affix only when agreement features are phonologically realized to a sufficient degree.

5.1 On Agr

The analysis presented below depends on the existence of affixal Agr heads. But the existence of Agr heads has been questioned. First, Chomsky (1995a; 1998; 1999) claims that Agr heads should be eliminated from syntax. Agr heads are made up solely of

uninterpretable ϕ -features, and lack any semantic features. Agr heads, then, are completely uninterpretable at the LF interface. Chomsky argues that such lexical items, composed of nothing but uninterpretable features, violate (at least the spirit of) Full Interpretation, and therefore do not occur in a strictly minimalist grammar. The argument is conceptual, and Chomsky himself notes that an empirical need for Agr heads might well trump it.

The theory of DM offers an alternative approach that would avoid the problem Chomsky raises. Halle & Marantz (1993) claim that a morphological operation Insertion can add heads and morphological features to an MS representation in order to satisfy language specific morphological requirements. The heads and features inserted at MS must lack semantic features, which would be uninterpretable at phonology. As instances of inserted heads/features they mention case, noun class, and agreement. It seems that the Agr head is a good candidate for MS insertion, since it is precisely Agr's lack of interpretable semantic features that leads Chomsky to reject Agr heads in the syntax. If Agr were inserted post syntactically at MS as Halle & Marantz suggest, then concerns about Full Interpretation do not arise.

However, MS insertion of Agr cannot occur if the analysis of V to I presented below is correct. Halle & Marantz (1993) assume a non-derivational MS, allowing Insertion directly at the terminal where heads/features are required. On the present analysis, the Agr head is located at the root of the MS tree, above T. If Agr were inserted directly at T or V, the motivation for (obligatory) verb raising would disappear. Alternatively, suppose that MS is derivational, as suggested above, and direct Insertion is ruled out. Agr will be Inserted at the root of the MS tree, above T. This places Agr above the subject, which is already hosted in the spec of T. V to I is now predicted to raise the finite verb above the subject, contrary to fact.

We conclude, contra Chomsky, that Agr heads are present in the syntax (and not Inserted at MS). The conclusion is supported by empirical evidence for the existence of Agr heads in the syntax, presented by several researchers [(Grohmann 2000; Schütze 1997; Wexler 1998).

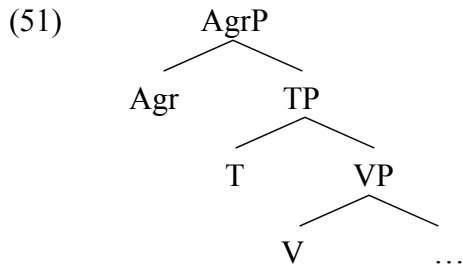
A second issue concerns whether Agr should be decomposed, for instance into a Person and a Number head. That would have little effect on the analysis below, so long as the head triggering V to I is higher than T, and contains features whose phonological realization is correlated with verb movement. If Vikner (1997) is correct that the feature correlated with V to I is simply person in every tense, then the Agr head could indeed be replaced with a Person head in the present analysis. Whatever the case may be, we will avoid notational complications here and keep the Agr head intact.

5.2 T-affixal Agr = Obligatory verb raising

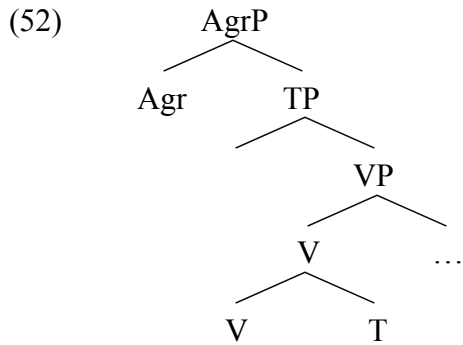
Verb raising is obligatory when Agr is a T-affix. This is because only the application of MT to V will result in a derivation where both Agr is affixed to T, and T is affixed to V. Consider an MS representation where Agr and T are affixal:^{17,18}

¹⁷ Irrelevant structure is not shown, here or below.

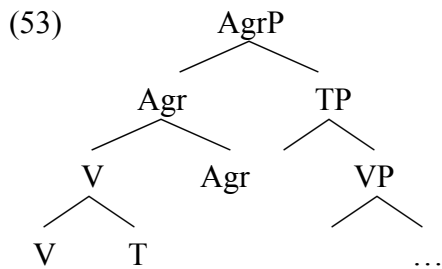
¹⁸ Note that the affixhood of T and Agr is not universal. While T and sometimes Agr are clearly affixal in the languages under consideration, they are sometimes just as clearly non-affixal, notably in creole languages. See Section 5.4 for some discussion; see also Halle & Marantz (1993).



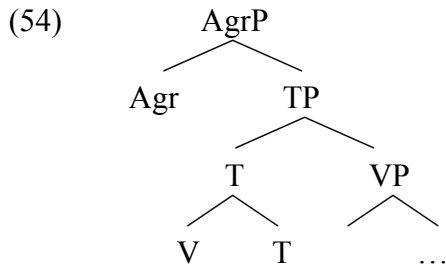
T can meet its V-affixal requirement either by Merger with V, or by MT of V to T. If Merger were applied to T and V, the resulting structure would be as in (52):



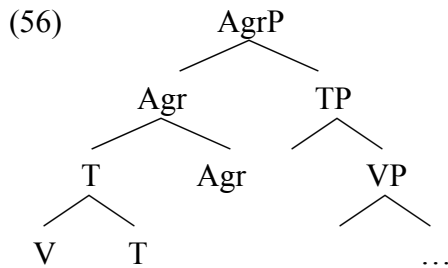
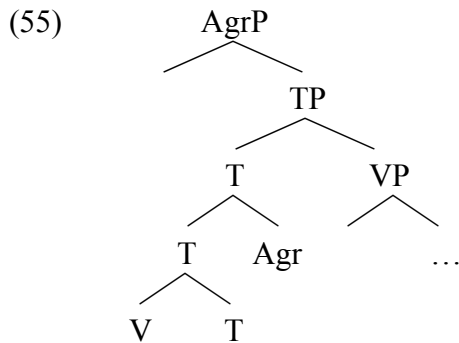
Agr's T-affixal requirements are unmet, and now cannot be met. Even if MT raises V to Agr, Agr is affixed V, and not to T as required:



Of course, the MT operation in (53) is illegitimate by definition, since it satisfies no morphological requirements. Therefore, any MS derivation that begins with Merger of V and T will be impossible when T-affixal Agr is present. The affixal requirements of Agr and T can only be met if the second option is chosen, and MT raises V to T:

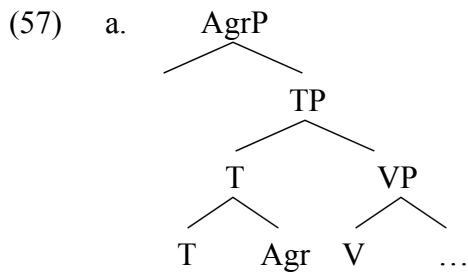


At this stage, Agr could either Merger with T, or T could MT to Agr. Either operation would result in Agr being affixed to T, satisfying Agr's affixal requirements:

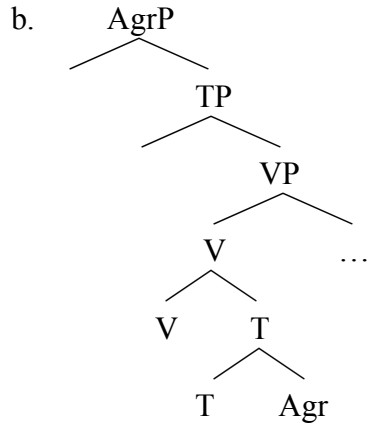


In order to distinguish the two options, we must consider movement to C. Languages with V to I move finite main verbs to C for question formation. If this is because of the T-affixal requirement of C, as suggested above, then (56) will be ruled out. Raising the Agr complex will not satisfy the affixal requirement of C. (And Raising T out of the Agr complex would disrupt T's affixation with either V or Agr.) Therefore Merger of Agr and the T complex (55) is the correct derivation.¹⁹

A question arises this point: what prevents a derivation where Agr Mergers with T, and then T Mergers with V? This would satisfy all affixal requirements:



¹⁹ This conclusion also lends support to the claim, made in Section 5.3.1 below, that there is preference for Merger over MT.

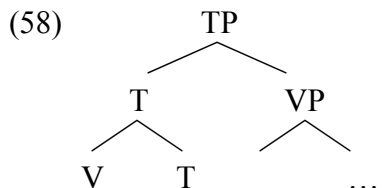


In order to rule out a derivation as in (57), we must invoke the two assumptions discussed in Section 3 above. Those were that MS is derivational, and that affixal requirements are satisfied as soon as possible. Terminals may not wait to satisfy their affixal requirements if a suitable host terminal is already available. Thus T must be affixed to V at the stage of the derivation when T is introduced. T's affixal requirements cannot wait until the later introduction of Agr. On these grounds, (57) is not a legitimate derivation.

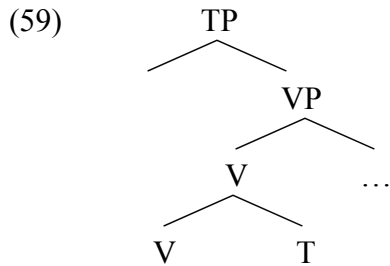
Summarizing, in an MS representation where Agr is a T-affix, MT must first raise V to T. Otherwise the affixal requirements of Agr and T will not be met. Therefore, where Agr is affixal verb raising is obligatory.

5.3 Non-affixal Agr = Verb raising

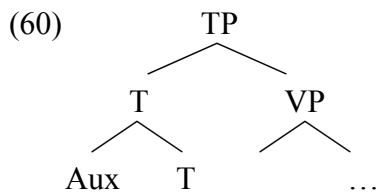
We now consider representations where Agr is not an affix, but T is. Verb raising is possible, but not obligatory in such representations. There are at least three ways to satisfy the V-affixal requirement of T where Agr is non affixal. Languages adopt either one of these strategies or a mixture of them. However, only one involves MT of V to T. This seems to be the least common strategy:



More common is simple Merger of T with Adjacent V:



A third possibility is that an auxiliary element, having verb features such that it can satisfy T's affixal requirement, affixes to T:²⁰



Below we examine in turn each strategy, its empirical effects, and the theoretical issues associated with it.

5.3.1 *MT of V*

In a language, or clause type, where Agr is non-affixal and MT raises V to T as in (58), finite verbs will appear to the left of negation, adverbs, and floated quantifiers. Thus the position of verbs in such languages is the same as in languages with T-affixal Agr. For reasons of morphological transparency discussed in Section 5.4, we know that raising is to T and not Agr only if agreement is 'poor'. Examples of this type of language were briefly discussed above; they include historical stages of English and Danish, modern dialects of Swedish, and older Northern dialects of English.

The question that arises immediately is why verb raising in the absence of affixal Agr is so uncommon, given that it is both empirically attested and theoretically possible. Typologically, there are only a handful of languages (or clause types) that have verb raising and no agreement. Looking at language change, the asymmetry is even more pronounced. There seems to be a strong tendency for languages that lose affixal Agr to lose verb raising as well, for example in English and Danish. Why should this be, given that verb raising is possible without affixal Agr? Indeed, both English and Danish retained verb raising for several generations after the loss of agreement (Roberts 1999). Why not simply continue raising verbs? Why should there be any subsequent change? And moreover, why is there no attested case of a change from non-raising to raising in a language lacking affixal Agr? How can we explain the second part of the agreement correlation?

Roberts (1999) argues that the relative tendency against verb raising in the absence of agreement is a result of feature markedness. Strong features, which cause movement, are marked. The child acquiring language will assume that all features are

²⁰ The auxiliary might head a phrase above T, or between TP and VP. Alternatively, it might be inserted at MS, as Halle & Marantz (1993) suggest for English *do*-support. Determining the status and initial position of this auxiliary element is well beyond the scope of this paper.

weak, unless presented with evidence to the contrary. Roberts claims that agreement morphology is evidence of a strong V-feature on I. If agreement is lost, the next generation of children reverts to the unmarked weak value of the feature, and verb raising is lost. The problem with this explanation, just as with all syntactic analyses of V to I, is that the connection between the phonological realization of agreement morphology and the strength of abstract syntactic features must be stipulated. As discussed above, it is hard to see why such a connection should hold. And in any case, an analysis that relies on feature strength is no longer an option in current syntactic theory (Chomsky 1998; Chomsky 1999).

A morphological explanation of this tendency relies not on feature markedness, but on the relative markedness of morphological operations. In Section 3.3 it was claimed that an economy condition, the EAC, holds on affixation at MS. The EAC requires that affixation requirements be met by the closest possible terminal. Sisterhood is the closest relation between terminals. Adjacency is the next closest relation, and non-Adjacent closeness is calculated in terms of intervening terminals. These three relations (Sister, Adjacent, non-Adjacent) constitute the structural conditions on three morphological operations: Fusion applies only to sister terminals, Merger only to Adjacent terminals, and MT to non-Adjacent terminals (also Adjacent ones).²¹ Given the EAC, then, the markedness of morphological operations corresponds to the closeness required in their structural conditions. Fusion will be the least marked, followed by Merger. MT is the most marked operation:

- (61) Markedness of morphological operations (from least to most marked):
Fusion > Merger > MT

Of course, markedness considerations will only come into play when there is a choice between operations that both satisfy an affixal requirement. (If no requirement is satisfied, then no operation is legitimate.) We saw above that MT--the most marked operation--must apply to V during the first step of the derivation when Agr is affixal. If Agr is rendered non-affixal by a phonological change, T's V-affixal requirement is still met by MT of V. But MT is no longer necessary in order to guarantee the affixation of T to Agr. Because T and V are Adjacent, the less marked operation Merger will suffice to satisfy all affixal requirements. Therefore Merger is preferred when Agr is non-affixal. In this way, markedness of morphological operations accounts for the second part of the agreement correlation.

5.3.2 *Merger with V*

In a language where Agr is non-affixal and T Mergers directly with V (59), verbs appear to the right of negation, adverbs, and floated quantifiers. Agreement morphology is 'poor'. Danish and (modern spoken) Faroese are examples of this type of language (Vikner 1997). For reasons of markedness, discussed above, Merger or V and T is preferred to MT of V, when Agr is non-affixal.

²¹ We leave aside Fission and Insertion. Since these operations have no structural requirements, their relative markedness cannot be evaluated with respect to closeness.

Negation poses an apparent problem for this analysis. In Danish, for example, finite verbs appear to the right of negation (Vikner 2001):²²

- (62) ...fordi du ikke svarede på hans brev.
because you not replied to his letter
"... because you didn't reply to his letter."

Intervening terminals are claimed to disrupt the Adjacency relation between two terminals, so that they cannot be Merged. If negation is a head located between T and V, as it seems to be the case in English, then Merger of T and V should be impossible.

There are at least two possible solutions to this problem. First, it could be the case that negation in these languages is not a head, but rather an adjunct. Because adjuncts do not disrupt negation, Merger of T to V would be possible despite the presence of intervening negation. Second, it could be that negation in these languages is located higher than T. If that were the case the problem disappears, since V and T would remain adjacent. Zanuttini (1997) argues for the existence of both adjoined negation and negation higher than T in some Romance varieties, so both options are at least possible in principle. But evidence is needed to determine the status of negation in the (Germanic) languages under consideration. Thus the problem of negation and adjacency must remain unresolved here, pending further investigation.

Languages of this type raise main verbs directly to C for question formation, for example in Danish (Vikner 2001):

- (63) Svarede du nogensinde på hans brev?
replied you ever to his letter
"Did you ever reply to his letter?"

If T is Merged to V as in (59), the resulting V complex cannot be raised, since it would not satisfy the T-affixal requirement of interrogative C. Therefore when interrogative C is present in the MS representation, MT applies to V in the first step of the derivation. The resulting T complex is then raised to interrogative C, satisfying all requirements. The system of morphological operations in Danish is thus mixed, in the sense that sometimes Merger satisfies the V-affixal requirement of T, and sometimes, when interrogative C is present, MT satisfies the requirement. This kind of mixture results from the tension between independent affixal requirements of terminals and the pressure to satisfy those requirements with less marked operations (Merger rather than MT).

5.3.3 Auxiliary Merger

The same tension produced a different kind of mixed morphological system in Modern English, where the affixal requirement of T may be satisfied in two ways.²³ Finite verbs appear to the right of adverbs (64a) and floated quantifiers (64b), which are adjuncts and

²² Danish root clauses are V2, obscuring the effects of V to I movement. For this reason all of the examples from Danish are embedded finite clauses, where V2 cannot occur.

²³ Actually three, since the copula raises. The copula is discussed below.

thus do not disrupt Adjacency. Thus, T merges directly with V when the two are Adjacent (59):

- (64) a. Those bastards brutally killed Kenny.
b. Those bastards all killed Kenny.

However, V and T cannot Merge when negation intervenes. This indicates clearly that English negation is located between T and V, and is a head:

- (65) * Eric not killed Kenny.

English might have developed a different kind of mixed system, where MT raises V to T only when negation is present. But that did not occur:

- (66) * Eric killed not Kenny.

Rather, English satisfies the V-affixal requirement of T by Merging T with an auxiliary (60), the semantically empty auxiliary verb *do*. The main verb, which has no affixal requirement, remains in situ, with its VE inserting the stem form:

- (67) Eric did not kill Kenny.

Merger with auxiliary *do* can take place in English only when T and V are not Adjacent.²⁴

- (68) a. * Eric did kill Kenny.
b. * Those bastards did brutally kill Kenny.
c. * Those bastards did all kill Kenny.

In Danish, and other languages where T is Merged directly with Adjacent V, MT raises V when interrogative C is present. However, this is not the case in English. MT does not apply in English:

- (69) * Killed Eric Kenny?

When interrogative C is present, T's affixal requirement is satisfied by auxiliary *do*. This results in a T complex that can be subsequently raised to C by MT, satisfying all requirements.²⁵

²⁴ This suggests that Insertion is more marked than Merger. But see fn. 21.

²⁵ The exception, of course, is subject wh-questions, where *do*-support does not occur:

- (i) a. Who killed Kenny?
b. * Did who kill Kenny?

Here, the subject is moved to spec, CP during the syntax, so that no argument XP intervenes between C and T at MS. We can follow Halle & Marantz (1993) in proposing that Merger of C and T is now possible only on the following two assumptions. First, neither the Agr head nor the subject trace interfere with Adjacency, because neither have an associated VE. This assumption makes predictions that must be tested. Second, V must MT to T so that a T complex is created. If T and V Merge, the resulting V complex cannot

(70) Did Eric kill Kenny?

Thus English, like Danish, has a mixed system of morphological operations. Sometimes the V-affixal requirement of T is satisfied by Merger of T with Adjacent V; sometimes, when either interrogative C or negation is present, this requirement is satisfied by Merger of T with auxiliary *do*.²⁶ A additional complication is *be*, which both raises and has morphologically transparent agreement. The issue of *be* is taken up in Section 5.4.1 below. In sum, then, the system of morphological operations found in English is quite mixed indeed, more so than Danish or typical V to I language such as French. Of course, the verbal and auxiliary system of English is more complex than presented here, with real questions about the various types of auxiliaries and modals, their position in the syntactic and MS representation, and their behavior when occurring together. A full treatment, though, is beyond the scope of this paper.

5.4 Agr and morphological transparency

So far, we have seen that affixal Agr causes obligatory verb movement. When Agr is non-affixal, verbs tend not to raise because of the greater markedness of MT. But it remains to be shown why the affixhood of Agr depends on the phonological realization of agreement. Below, this is argued to be the result of independently motivated principles of morphological transparency.

A principle of morphological transparency was first articulated by Emonds (1986). Emonds examined the unusual behavior of English pronouns in a number of syntactic environments. He observed that this unusual behavior did not occur in languages that, unlike English, mark case productively on all DPs. This observation led him to claim that the morphological form of English pronouns is not a realization of their syntactic case features. The reason, he argued, is that syntactic features can be the source of a morphological distinction only if those syntactic features are morphologically transparent. Otherwise, the distinction must be maintained by another mechanism, for him a local transformation.²⁷ Emonds defined morphological transparency as follows:

(71) **Morphological transparency (Emonds 1986)**

A syntactic category [formal feature] C is *morphologically transparent* on B iff a productive number of pairs of simple B which contrast with respect to C also differ phonologically.

satisfy C T-affixal requirement. The consequences of these assumptions are complicated, and furthermore likely interact with the assumptions, made in Section 3, that MS is derivational and affixal requirements are satisfied as soon as possible. The matter cannot be explored here.

²⁶ And the system is in fact more complicated, if the assumptions in fn. 25 are correct. If interrogative C is present, and C and T are adjacent, then MT raises V to T, and C Merges with the resulting T complex.

²⁷ On Emonds' analysis, when English lost productive case marking as result of independent phonological changes, syntactic case features became morphologically non-transparent in English. However, case marking on the closed set of pronouns was suppletive, and thus was unaffected by the phonological change. Syntactic case features, no longer morphological transparent, could not be responsible for the distribution of distinct pronoun forms in English. Emonds argued that a local transformation took over this function, and the interaction of this transformation with certain syntactic structures accounted for the unusual behavior of pronouns.

Parrott (2001), expanding Emonds' analysis within a DM framework, argues that morphological transparency affects the expression of formal features by constraining the content of VEs. Only morphologically transparent features can be contained in a VE's substantive feature set. Non-transparent morphological distinctions that remain in the wake of phonological changes can be maintained only by featural or structural descriptions in the contextual feature set of a VE.²⁸ A principle of Vocabulary Transparency is formulated as follows:²⁹

(72) **Vocabulary Transparency**

If some feature F of a terminal node N is not morphologically transparent, then F is not contained in the substantive feature set of a Vocabulary Entry for N.

The intuition behind both Emonds' and Parrott's analyses is an old one: "No one ever dreamed of a universal morphology."³⁰ Although UG provides the child with a set of primitive formal/syntactic features, the language particular phonological realizations of these features must be learned. A child can learn that some phonological features P are associated with some formal feature(s) F just when F is regularly associated with P in the child's input. If F is never associated with P in the input, then the child has no evidence to posit a relation P(F) in her grammar. But a phonological change might erase the productive association of P with the majority of instances of F, and leave P associated with a small non-productive subset F_N of F, so that P(F_N). Emond's and Parrott's claim is that a grammatical relation P(F) cannot be acquired solely on the basis of an input association P(F_N). In such cases, the child will associate P with some other formal feature, or with some property of the context in which P appears.

From this viewpoint, it seems clear that morphological transparency should also constrain the acquisition of morphological requirements. Consider the relevant case, affixation requirements on terminals. It cannot be that affixation requirements on particular terminals is part of UG, since languages can, and do, differ in this regard. For example, it seems to be the case that T heads are non-affixal in creole languages. Most creoles lack verbal tense or agreement morphology, but rather utilize non-affixal, phonologically independent particles to express tense (DeGraff 1997).³¹ On the present analysis, these particles can be regarded as the phonological realization of (i.e., the VE for) a non-affixal T.³²

²⁸ On Parrott's analysis, the substantive feature sets of English pronoun VEs do not contain syntactic case features. Rather, the contextual feature sets of pronoun VEs specify that a Subject Form (nominative) is inserted when the pronoun insertion node is left Adjacent to finite T. Object Forms (oblique) are inserted in all other contexts by the elsewhere VE. Social prescriptive pressure can result in the post-acquisition learning of supplemental, extra specific VEs for pronouns in certain environments. For example, an individual might learn a supplemental VE for 1st sing. pronouns that inserts the Subject Form in the context right adjacent to *and*. The over-specificity of these late learned VEs, and their interaction with the normal pronoun VEs, accounts for the unusual behavior of pronouns in English.

²⁹ Slightly modified from Parrott (2001).

³⁰ Who said this? Jespersen?

³¹ It has been claimed that some creoles also express agreement with particles. I have no reference available to confirm this, however.

³² Such particles might also express non-affixal mood or aspect heads.

The child must learn that an affixation requirement R holds of a particular terminal M; she must learn that her morphology contains the relation R(M). She can learn R(M) only if the affixation has a productive phonological realization in the input; that is, if it is the case that P(R(M)). Obviously, if M is a non-affix in the input, for example in creole languages, then a child will not learn that R holds of M. Nor can the child learn that R holds of M if M is not phonologically realized at all: M can never be affixal in the input if it is simply absent in the input. In such cases P(M) does not hold, so P(R(M)) cannot hold either. But a phonological change might erase the productive association of P with the majority of instances of M, but leave P(R) associated with a small non-productive subset M_N , so that P(R(M_N)). The claim made here, then, is exactly parallel to that above: a affixal requirement R(M) cannot be acquired solely on the basis of input P(R(M_N)). In such cases, the child will associate P with some other terminal, or with some property of the context in which P appears.

Therefore in order for a child to learn that R holds of M, M must be morphologically transparent. This is formulated below:

(73) **Affixation Transparency**

An affixal terminal must contain morphologically transparent features.

These principles explain why 'rich' agreement is correlated with V to I. Agreement is 'rich' when it is pronounced enough to be morphologically transparent. If so, Agr will be an affix and verb movement obligatory. Agreement is 'poor' when it is not pronounced sufficiently to be morphologically transparent. If so, Agr cannot be an affix according to Affixation Transparency. Then verb movement will tend not to occur, because MT is a more marked operation than Merger.

5.4.1 *Morphological transparency and agreement in English*

For an example of the application of these transparency principles in a particular case, consider agreement in English. The definition of morphological transparency asks us to count minimal pairs of items that differ only with respect to the feature in question. When the feature in question is agreement (person + number), the items to be compared are verbs. Minimal pairs cannot be constructed across tenses, since that would involve pairs that also differ with respect to the tense feature and not just the agreement feature. In English, there are 6 values of the agreement feature and so 15 possible agreement contrastive verb pairs within each tense of a given verb,³³ and thus 30 total pairs for both tenses of a given verb:³⁴

| | | |
|------|--|---|
| (74) | <p>[-past] I kill Kenny. You kill Kenny. He kills Kenny.</p> | <p>[+past] I killed Kenny. You killed Kenny. He killed Kenny.</p> |
|------|--|---|

³³ The number of distinct pairs possible in a given set of items is given by the equation $Y! / 2! \cdot (Y-2)!$, where Y is the number of items in the set. In this case, the equation is $6! / 2! \cdot (6-2)! = 6! / 2! \cdot 4! = 720 / 48 = 15$.

³⁴ This is just the sum of the number of pairs. We do not include both tenses in the set from which pairs are selected, since this would count non-minimal pairs that differ with regard to tense as well as agreement, as mentioned above.

We kill Kenny.
You (y'all) kill Kenny.
They kill Kenny.

We killed Kenny.
You (y'all) killed Kenny.
They killed Kenny.

There are zero pairs of [+past] verbs, out of a possible 15 pairs, that differ with respect to agreement and also differ phonologically. Among [-past] verbs, only one value of the agreement feature (3rd sing.) is phonologically distinguished. The other 5 values of the agreement feature (1st sing., 2nd sing., 1st plural, 2nd plural, 3rd plural) are phonologically identical. Thus in the [-past] tense there are 5 pairs of verbs, out of a possible 15, that differ with respect to agreement and also differ phonologically. The total number of pairs that differ with respect to agreement and also differ phonologically for a given verb is therefore 5 out of a possible 30, or 16%. This percent will remain the same no matter how many distinct verbs are counted.

Emonds' definition of morphological transparency requires a 'productive number of pairs.' Agreement contrastive pairs in English are certainly not productive, in the sense that newly added verbs will not have more agreement morphology. But Emonds is not specific about how many pairs are necessary for transparency. Recognizing that this needs to be made precise in the future, I want to leave the question open for research.³⁵ Here we will assume that 16% is not a sufficient number of contrastive pairs for morphological transparency. Therefore we conclude that agreement is not morphologically transparent in English.

The principles of Vocabulary Transparency and Affixation Transparency make precise the consequences of this conclusion. The Agr terminal is composed of agreement features. Agreement features are non-morphologically transparent in English. By Vocabulary Transparency, then, an English VE for Agr cannot contain agreement features in its substantive features. But there are no other morphologically transparent features of Agr that could be contained in the substantive features of a VE. Therefore Agr has no VE in English (assuming that a VE must have at least one substantive feature; otherwise its insertion site would be unspecified). Agr is either left with no phonological features, or perhaps a universal default \emptyset is inserted during VI.³⁶ By Affixation Transparency, Agr also cannot be an affixal terminal, since none of the features of Agr are morphologically transparent.

One question is how to account for English 3rd sing. -s on this theory. If -s is not a realization of Agr, what is it, and what features are involved? As discussed above, morphological distinctions that remain in the aftermath of phonological changes must be maintained by some mechanism other than by direct insertion according to the substantive features of VEs. Parrott (2001) argues that these distinctions are maintained by the contextual features of VEs. I propose that this is also the case here: English -s is inserted at $T_{[-past]}$, in the context of a 3rd sing. subject. The definition of 'subject' is

³⁵ Although see fn. 39.

³⁶ Of course, we would not want to allow lexical items to be filled in by a universal default \emptyset or anything similar. Indeed, there is good evidence that no such 'zero' lexical items exist, as opposed to the many documented instances of zero grammatical morphemes (see (Beard 1995). This difference is not reflected in DM, where both lexical items and grammatical morphemes lack phonological features before VI. (In Beard's system, by way of contrast, lexical items, but not grammatical morphemes, contain phonological features before and during the syntax.) If this characteristic of DM turns out to be a drawback, I think it can be changed without massive disruption to the rest of the theory.

deliberately left vague pending further research. The set of VEs for $T_{[-past]}$ is given below. (75a) inserts /s/ at a terminal containing the features $T, [-past]$ just in case the subject has the features +3rd sing. (75b), the elsewhere VE, inserts \emptyset at a terminal containing the features $T, [-past]$ whenever the first VE does not apply.

- (75) VE for English $T_{[-past]}$
- a. $[T_{[-past]}] \Leftrightarrow /s/ / \text{Subject}_{[3rd\ sing.]}$
 - b. $[T_{[-past]}] \Leftrightarrow / \emptyset /$

This analysis of English -s makes at least two testable empirical predictions. First, Parrott (2001) observes that non-transparent morphological distinctions often exhibit behavior that is not observed when the distinction is transparent. This is because, unlike VE insertion according to substantive features, the insertion of a VE according to context can be disrupted. One possibility for disruption is that some other part of the environment can change the context of a terminal so that it no longer matches the contextual features of its VE. Then an unexpected VE, usually the elsewhere VE, will be inserted. Parrott (2001) argues that this is the case for English pronouns in coordinates. A second disruption possibility is that the contextual features of a VE are miscalculated, especially when the environment they specify is ambiguous. This second possibility seems more likely in the present case, since the definition of subject in the VE for -s is ambiguous. Thus, where there is more than one candidate subject, for instance where a complex subject includes embedded subjects, we might expect agreement 'errors.'

A second prediction involves language change. VEs inserted according to contextual features are subject to disruption, as noted; moreover, they have significantly more computational complexity than substantive feature VEs. Thus it is reasonable to assume a tendency toward transparent grammars, here defined as a vocabulary made up of VEs inserted according to substantive features and not contextual ones.³⁷ If this is the case, we expect that non-transparent morphological distinctions--such as English -s--will be especially susceptible to re-analysis, leveling, and other forms of reorganization. These reorganization processes are predicted to result in morphological transparency, thereby replacing a contextually inserted VE with a substantively inserted one.³⁸

The present analysis of agreement raises the additional issue of how to account for the English verb *be*. Consider *be* in present and past tense:

- | | |
|----------------------|-----------------------|
| (76) [-past] | [+past] |
| I am sad. | I was sad. |
| You are sad. | You were sad. |
| He is sad. | He was sad. |
| We are sad. | We were sad. |
| You (y'all) are sad. | You (y'all) were sad. |
| They are sad. | They were sad. |

³⁷ Kiparsky (1971) and Peinovich (1979) both make analogous claims with regard to phonological rules.

³⁸ -s seems to be undergoing leveling in some English dialects, including African American English (AAE) and Smith Island English (SIE). For AAE I have no references at hand. Regarding SIE, -s leveling has not been researched, but is given brief mention in Schilling-Estes (2000).

In the present tense, there are 9 agreement contrastive pairs out of a possible 15 (60%); in the past tense there are 8 agreement contrastive pairs out of a possible 15 (53%). For *be*, then, there are a sum total of 17 agreement contrastive pairs out of a possible 30 (57%). It seems reasonable to conclude that agreement is morphologically transparent on *be* in English.³⁹

Besides displaying morphological transparent agreement, *be* has another distinguishing characteristic. Unlike other English verbs, *be* raises:

- (77) a. Kenny is not dead yet.
 b. Kenny is often dead.
 c. The three boys are all alive.
 (78) Is Kenny dead yet?

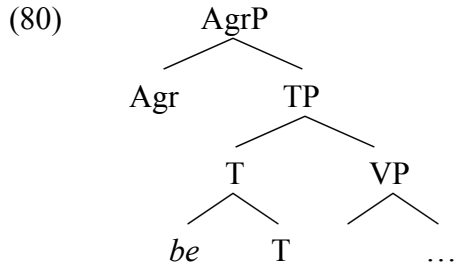
It was claimed above that English has no VE for Agr, and that in English Agr is not an affixal terminal. Both claims seem to be contradicted in the case of *be*. The transparency of agreement on *be* suggests the existence of a VE for Agr. This, and the unique raising behavior of *be*, also suggests that Agr is an affixal terminal. It might be possible to stipulate that English in fact contains an affixal Agr terminal and a VE for Agr, both somehow linked to the co-occurrence of *be*. But this move is obviously ad hoc, and in any case it is not clear by what mechanism the verb *be* could be linked with Agr in the necessary way.

A more plausible solution is rooted in the observation that *be* is suppletive: verbal, tense, and agreement features are realized in a single morpheme. On the theory of DM, this indicates that the operation Fusion has taken place, combining the features of *be*, T, and Agr into a single terminal. The transparency of agreement on *be* indicates to the child only that agreement features must be associated with the VE for *be*, and not that Agr itself has a VE. Thus there is no VE for Agr in English, but only VEs whose substantive features refer to terminals containing the complex of V-T-Agr features resulting from Fusion. An example is given below:

- (79) [V_[be], T_[-past], Agr_[3rd Sing.]] ⇔ /Iz/

The problem of affixation is resolved in the same manner. Agreement is transparently affixal only on suppletive *be*, and this is not sufficient evidence for the child to learn that Agr is an affixal terminal. Rather, because the VE for *be* includes agreement features contained in Agr, it cannot be inserted if *be* is not Fused with Agr (and of course T) during the MS derivation, before the application of VI. Thus the morphological requirement here holds of *be* and not of Agr. Note that this is an example of a morphological requirement other than affixation, motivated by VEs. The VE of *be* requires Fusion with Agr, and not just Merger or MT with Agr. But of course, combination via Merger or MT is a prerequisite for the application of Fusion. The observed raising of *be* in English indicates that MT must raise *be* to T, at least:

³⁹ Notice that this suggests 50% as the number of pairs required to establish morphologically transparency.



At this stage, either MT or Merger could combine Agr with the T complex. The question is left unresolved, although the EAT suggests Merger (despite the fact that there is no affixal requirement here).

Notice that raising is not obligatory under conditions of the kind just outlined, since Agr is still non-affixal. The morphological system of English, in the case of *be*, fulfills its requirements by applying MT of V to T. Some other language might fulfill the same requirements with another combination of operations. And indeed, in the case of main verbs other than *be* English does not fulfill morphological requirements with MT.

5.5.2 Summary

Summarizing, according to the principle of Affixation Transparency, affixal terminals must have morphologically transparent features. When agreement is not regularly pronounced, it is not morphologically transparent. The result is that Agr has no morphologically transparent features, and cannot be an affix. Verb movement will tend not to occur, for reasons discussed above.

Applying these principles to English, it was shown that agreement features are not morphologically transparent in English. According to the principles of Vocabulary Transparency and Affixation Transparency, English Agr has no VE, and Agr is not an affixal terminal in English. Whatever role Agr plays in the syntactic computation, it is ignored in the MS derivation, triggering no operations and no VI (apart from a possible universal default \emptyset). English *-s* is not inserted by a VE for Agr, but by a VE for $T_{[-past]}$, in the context of a 3rd sing. subject. Agreement is transparent on English suppletive *be*, which furthermore raises. The VE for *be*, whose substantive features consist of the features for *be*, tense, and agreement, requires the application of Fusion to *be*, T, and Agr prior to VI. English uses MT to assemble these heads for Fusion, resulting in the raising of *be*, but not other verbs.

6 Conclusion

This paper has presented a morphological theory of head movement, within a modified version of DM. It was argued that this theory can provide a good account of V to I, and especially the agreement correlation, whose adequate explanation has escaped syntactic analysis. Predictions were made, and avenues for further research were pointed out. A serious theoretical question remains to be addressed, however: what is the source of the XP/X^0 movement asymmetry? Why can the syntactic computation move XPs, but not heads?

I would like to suggest in conclusion that this asymmetry follows quite naturally from recent syntactic theory. According to Chomsky (1995a; 1998), movement is no

longer motivated by the need to check inflectional features. Rather, movement is caused by the EPP feature, which requires certain heads to have an XP as their specifier. As Lasnik (2001) points out, this is nothing more than a restatement of an earlier version of the EPP: certain phrases must have subjects. But on this new (old) conception of movement, head movement is ruled out in principle. On the theory of bare phrase structure (Chomsky 1995a), movement of a head to a specifier position violates the chain uniformity condition. The tail of the chain is head, but its head is an XP. It follows, then, that the syntax cannot move heads. To the extent that this view of movement is correct, the morphological theory of head movement is supported.

References

- Anderson, S. J. (1996). How to put your clitics in their place, or why the best account of second position phenomena may be something like the optimal one. *The Linguistic Review*, 13, 165-191.
- Anderson, S. R. (1992). *A-Morphous morphology*. Cambridge: Cambridge University Press.
- Anderson, S. R. (1993). Wackernagel's revenge: Clitics, morphology, and the syntax of second position. *Language*, 69(1), 68-98.
- Baker, M. (1988). *Incorporation: A theory of grammatical-function changing*. Amsterdam: John Benjamins.
- Beard, R. (1995). *Lexeme-morpheme base morphology : a general theory of inflection and word formation*. Oxford: Oxford University Press.
- Belletti, A. (1990). *Generalized verb movement*. Turin: Rosenberg and Sellier.
- Bobaljik, J. D. (1995). *Morphosyntax: the syntax of verbal inflection*. Unpublished Ph.D., MIT.
- Boeckx, C., & Stjepanovic, S. (2001). Head-ing toward PF. *Linguistic Inquiry*, 32(2), 345-355.
- Boskovic, Z. (1995). Participial movement and second position cliticization in Serbo-Croatian. *Lingua*, 96, 245-266.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, N. (1995a). Categories and transformations, *The Minimalist program* . Cambridge: MIT Press.
- Chomsky, N. (1995b). *The minimalist program*. Cambridge: MIT Press.
- Chomsky, N. (1998). Minimalist inquiries: the framework : MIT Working Papers in Linguistics.
- Chomsky, N. (1999). Derivation by phase : MIT Working Papers in Linguistics.
- DeGraff, M. (1997). Verb syntax in, and beyond, creolization. In L. Haegeman (Ed.), *The New comparative syntax* . New York: Longman.
- Emonds, J. (1978). The complex V-V' in French. *Linguistic Inquiry*, 9, 151-175.
- Emonds, J. (1986). Grammatically deviant prestige constructions. In M. Brame, H. Contreras, & F. J. Newmeyer (Eds.), *A festschrift for Sol Saporta* . Seattle: Noit Amrofer.
- Grodinsky, Y., & Finkel, L. (1998). The Neurology of empty categories: Aphasics' failure to detect ungrammaticality. *Journal of Cognitive Neuroscience*, 10, 281-292.
- Grohmann, K. K. (2000). *Prolific peripheries: A Radical view from the left*. Unpublished Ph.D., University of Maryland, College Park.
- Halle, M., & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale & S. J. Keyser (Eds.), *The view from Building 20* . Cambridge: MIT Press.
- Hornstein, N., & Lightfoot, D. (Eds.). (1994). *Verb Movement*. Cambridge: Cambridge University Press.
- Kiparsky, P. (1971). Historical Linguistics. In W. O. Dingwall (Ed.), *A Survey of linguistic science* . College Park, MD: University of Maryland.
- Lasnik, H. (1999a). *Minimalist analysis*. Oxford: Blackwell.

- Lasnik, H. (1999b). On feature strength: Three minimalist approaches to overt movement. *Linguistic Inquiry*, 30, 197-217.
- Lasnik, H. (2001). A Note on the EPP. *Linguistic Inquiry*, 32(2), 356-362.
- Lema, J., & Rivero, M.-L. (1989). *Long Head Movement: ECP vs. HMC*. Paper presented at the NELS 20, Pittsburgh.
- Parrott, J. K. (2001). *Explaining linear ordering asymmetries in English: A Morphological alternative*. Paper presented at the Asymmetry Conference, Montreal, Canada.
- Peinovich, M. P. (1979). *Old English noun morphology*. Amsterdam: North Holland.
- Pollack, J. I. (1989). Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry*, 20, 365-424.
- Rivero, M.-L. (1990). Patterns of V0 raising in Long Head Movement, and negation: Serbo-Croatian vs. Slovak. *International Journal of Basque Linguistics and Philology*.
- Rivero, M.-L. (1991). Long Head Movement and negation: Serbo-Croatian vs. Slovak and Czech. *The Linguistic Review*, 8, 319-351.
- Rizzi, L. (1990). *Relativized minimality*. Cambridge: MIT Press.
- Roberts, I. (1985). Agreement parameters and the development of English modal auxiliaries. *Natural Language and Linguistic Theory*, 3(1), 16-25.
- Roberts, I. (1992). *Verbs and diachronic syntax*. Dordrecht: Kluwer.
- Roberts, I. (1999). Verb movement and markedness. In M. DeGraff (Ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development*. Cambridge: MIT Press.
- Rohrbacher, B. (1994). *The Germanic languages and the full paradigm: A Theory of V to I raising*. Unpublished Ph.D., University of Massachusetts.
- Schilling-Estes, N. (2000). *Exploring morphological change: The was/weren't pattern in Smith Island English*. Paper presented at the NWAV 29, Michigan State University.
- Schütze, C. (1997). *Infl in child and adult language: Agreement, case and licensing*. Unpublished Ph.D., MIT.
- Travis, L. (1984). *Parameters and effects of word order variation*. Unpublished Ph.D. dissertation, MIT.
- Uriagereka, J. (2000). Multiple Spell-Out. In S. Epstein & N. Hornstein (Eds.), *Working Minimalism*. Cambridge: MIT Press.
- Vikner, S. (1997). V-to-I movement and inflection for person in all tenses. In L. Haegeman (Ed.), *The New comparative syntax*. New York: Longman.
- Vikner, S. (2001). *The Syntax of English auxiliary verbs*. Paper presented at the Stuttgart University, Stuttgart, Germany.
- Wexler, K. (1998). Very early parameter setting and the unique checking constraint: a new explanation of the optional infinitive stage. *Lingua*(106), 23-79.
- Wilder, C., & Damir, C. (1994). Long Head Movement? Verb movement and cliticization in Croatian. *Lingua*, 93, 1-58.
- Zanuttini, R. (1997). Negation and verb movement. In L. Haegeman (Ed.), *The New comparative syntax*. New York: Longman.