Introducing Exxtension

Ken Safir

Rutgers University

1. Introduction

The essential goal of this short paper* is to preserve a key intuition behind Chomsky's (1995) notion of Extension, informally stated in (1).

(1) **Extension**: In the course of a derivation, all Merge or Move attaches to the undominated node.

Extension embodies an intuition that appears worth preserving, namely, the notion that at any point in a derivation, what counts as the top of the tree *at that moment* restricts the class of possible operations that can apply. This principle, if a version of it is correct, propels an aggressively derivational view of syntax.

However, Chomsky's proposal is too restrictive to permit a wide range of syntactic analyses that appear to be worth preserving, all of which violate Extension by merging constituents *near* the top of the tree. After some 'fixes' of Extension are rejected, including interarboreal movement, I then propose to revise Extension as Exxtension. The essential idea of Exxtension is to loosen, yet make precise and restrictive, the definition of what counts as (*near*) the top of the tree. The new definition, to be introduced after some structural assumptions have been established, has interesting consequences for the role of thematic selection and modification once it has been enhanced by the Selection Principle, also to be introduced below. Then all of the analyses discussed that are worth preserving will be shown to be viable within the Exxtension approach, except one, for which a different sort of analysis will be sketched.

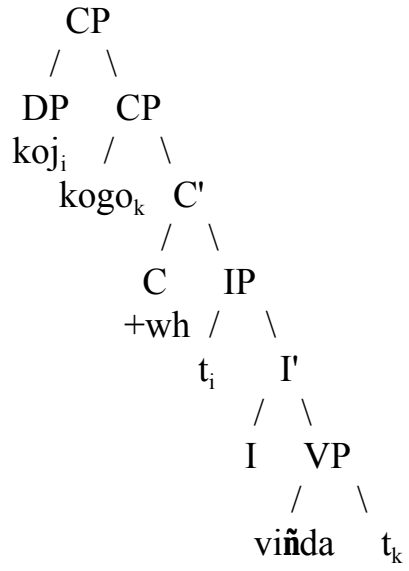
2. Extension Is Too Restrictive

Many analyses that have played a major role in both minimalist and principles and parameters syntax violate Extension. Four such analyses are *late attachment*, *tucking in*, *head adjunction to heads*, and *LF-movement*, as well as analyses that employ these analyses as component parts, such as Fox's (2002) of antecedent contained deletion. Let us consider all of these in turn.

Richards (1997, 1999) proposes a version of Merge he calls *tucking in* which is part of an explanation for the distribution of superiority restrictions or their absence, particularly for languages, like Bulgarian, where more than one wh-phrase can be fronted in a question. He argues that sentences like (2) are represented by trees like that in (3).

(2) Koj kogo vińda (Bulgarian from Rudin 1988:472-3)
who whom sees "Who sees whom?"

(3)

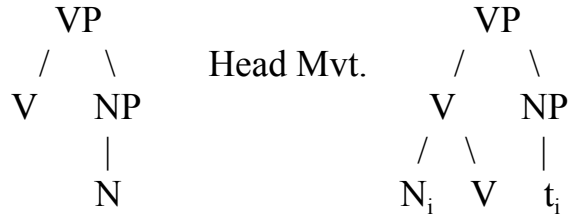


In Richards' account, *koj*$_i$ moves first to Spec CP to satisfy the +wh feature of C, then *kogo*$_k$ moves and *tucks in* under *koj*$_i$ to satisfy Shortest Move. This predicts the ill-formedness of the opposite wh-phrase ordering (**Kogo koj vińda*). In

other words, after an initial wh-movement satisfies the needs of a wh-feature on C, subsequent movement attracted by the wh-feature on C will not attach to the undominated node, as Extension would require, but rather adjoins to extend the sister node of the first fronted wh-phrase, as illustrated in (3). If he is right, then Merge must be able to adjoin to a position below the undominated node, violating Extension.

Any number of head movement analyses assume that one head can adjoin to another. Baker's (1988) use of head adjunction as the mechanical instantiation of incorporation, illustrated in (4), is one such widely appealed to analysis.

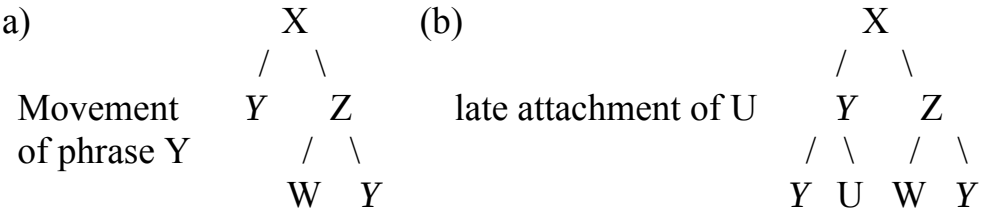
(4)



Adjunction to the head of a phrase by another head is not adjunction to the undominated node, and therefore is a violation of Extension.

Late attachment of adjuncts has been appealed to to account for adjunct/complement asymmetries in the distribution of reconstruction effects (e.g., Lebeaux , 1990, Chomsky, 1995, Safir, 1999, 2004). The essential distinction between adjuncts and complements is the assumption that complements must be introduced as sisters to the heads that select them, whereas adjuncts are attached higher than complements, such that they are not sisters to the selecting head. Lebeaux and Chomsky exploit this distinction to suggest that adjuncts do not have to be added to the tree when a nominal is first constructed, but can be added to the nominal at a later point in the derivation by adjoining to the whole nominal. A complement could be formed this way since adjunction is of a head to a head or an XP to an XP only. Now consider

movement that leaves a copy, as in (5a), followed by late attachment, as in (5b).



The late attachment of U means that U is not c-commanded by W at any point in the derivation (5a,b). If, however, U were a complement of Y and moved as part of Y, then a copy of U would be c-commanded by W because W would c-command the copy of Y. Attachment to a moved constituent (Y, in this case) is not adjunction tothe undominated node, and therefore Extension is violated.

Contrasts that have been exploited as evidence for late attachment are presented in (6). For reasons that I will make clear further below, not all tests that have been claimed to be evidence for reconstruction are indeed good evidence, and so I limit my support for this contrast to the case of quantifier bound pronouns as they participate in crossover effects.

(6a) *[which of the [pictures of *each former nazi*]][is *he* supposed to destroy t]?
Ans. The picture of *him* in *his* uniform.

(b) [which of the [pictures of *each former nazi*]][t is most likely to expose *him*]
Ans. The picture of *him* in *his* uniform.

(c) [[[which of the women] [near *each dancer*]] [is *he* supposed to turn to t]]
Ans. The woman on *his* left.

The late attachment analysis explains these contrasts as follows: Since *(of) each former nazi* is thematically related to *pictures* in (6a), it must be a sister to *pictures* when *pictures* is merged. Thus *pictures of each former nazi* must leave

a copy (in the position marked with "t") containing the quantifier (or rather its variable) and strong crossover ensues for (6a) because *he* c-commands the variable it is supposed to depend on (see Safir, 2004). In (6b), the position of "t" is not c-commanded by the pronoun, so the pronoun can depend on the quantifier (i.e., no crossover ensues). Now consider (6c). Since *near each dancer* can be adjoined to *which of the women* by late attachment, no copy of *each dancer* is left in the position of "t", hence no crossover effect ensues in (6c). If late attachment is banned by Extension, then this attractive explanation for othwerwise puzzling crossover effects would be lost.

LF-movement in general, and quantifier raising (QR) in particular, also violate Extension if it applies to trees already built (to yield further structural effects relevant to interpretation - primarily, but not exclusively, scope). However, there are ways, to be addressed later, in which most of the LF-movements that have been discussed can be rendered consistent with Extension without changing the notion at all. For this reason, I will have a lot less to say about this particular class of violations, once my commitment the sorts of analyses that render them harmless to Extension is acknowledged below.

Another analysis that violates Extension is Fox's account of antecedent contained deletion (ACD) of elided VP's. I postpone my discussion of it until section 4, but it suffices here to say that it includes both late attachment and QR as component parts and thus violates Extension.

How then can we render these analyses consistent with fundamental structural principles?

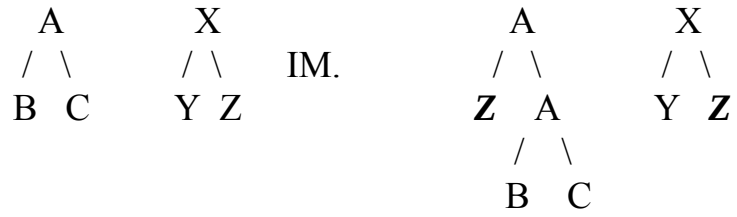
3. How Not To Fix Extension

That Extension is too restrictive to permit some standard analyses has been understood and addressed for some of the cases above. Chomsky (1995:

191) suggests that Extension only applies to substitution, hence not to head movement or LF-movement, presumably. Movement to Spec CP, by contrast, is then taken to be substitution, but if Richards is right, then we need tucking in analyses which treat movement into the C-domain as adjunction. As we shall see, evidence from late attachment shows that a form of adjunction that cannot be treated as substitution must also be tree-top sensitive. Chomsky (2000:137) suggests that an operation called *local merge* can outrank Extension in certain configurations, forcing the moved constituent to be as close as possible to the attracting head for cases like head movement or tucking in, but even this ad hoc fix also fails to apply naturally to cases of late attachment discussed below.

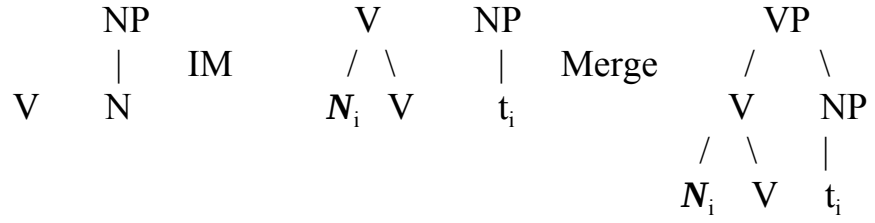
Bobaljik and Brown (1997) attempt to accommodate some of the analyses in question without compromising the definition of Extension by positing the existence of *interarboreal movement* (henceforth *IM*) (see also Nunes, 2001 and Hornstein, 2001). Though Extension requires Move and Merge to join to an undominated node, the argue, it does not require that a node like Z in (7) must attach to the tree that dominates its launching site.

(7)



Thus IM respects Extension in (7) because Z attaches to the undominated node A. These authors analyze head movement accordingly, as in the schematic representation of N-incorporation (see Baker, 1988) represented in (8).

(8)



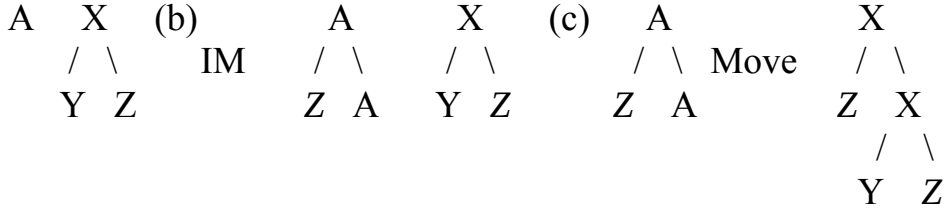
The problem with IM, however, is that it permits many more derivations than we need to justify head movement, and at the same time it cannot capture the late attachment analysis.

IM has been extended to Across-the-board extraction and parasitic gaps by Nunes (2001), and to a general theory of coconstrual (anaphora) by Kayne (2002) and by Hornstein (2001). However, these approaches permit a wide variety of structures to be derived that we would probably want to exclude in principle. For example:

(9a) Without unnatural restrictions, IM negates island restrictions in a wide variety of cases where such restrictions seem (empirically) to be in force.
(b) IM is neither left nor right, so the claim that there is no "rightward" movement is lost in the formation of trees by sideward movement.
(c) IM allows the same copy to be moved twice or more from the same position, e.g., once to a separate tree, and then again to a position within the same tree.

The last of these problems is schematically represented in (10)

(10a)



Notice that it is not possible to move twice from the same position in a single tree built respecting Extension and excluding IM, since Attract Closest will

always select the higher of two copies. Movement from separate trees to an attractor cannot express the notion "closest".[1]

The obliteration of island distinctions and left/right distinctions can be illustrated with a single, somewhat complicated, example. Consider (11).

(11) When did the man who said that Mary had died reveal that Joe had left?

This example is a perfectly well-formed if it asks about the time of revelation or the time of Joe's leaving, but it certainly cannot be a question about the time of Mary's death. Extraction of *when* from the clause *Mary had died when* would violate the Complex NP Constraint and the Subject Constraint, strong islands especially fatal to adjunct-extraction. Unfortunately, IM easily generates (11) as a question about the time of Mary's death, as illustrated in (12).

(12) Tree A Tree B

(a) [$_{C'}$ that [Joe had left]] [$_{IP}$ Mary had died when]

- *Trees A and B built respecting Extension*

(b) [$_{CP}$ *when* that [Joe had left]]] [$_{IP}$ Mary had died *when*]

- *IM extends tree A with* ***when*** *moved from tree B to tree A*

(c) [$_{I'}$ reveal [$_{CP}$ *when* that Joe had left] - *Merger to increase tree A*

(d) [$_{DP}$ the man who said [that [Mary had died *when*]]] -*Tree B further extended*

(e) [$_{IP}$ [$_{DP}$ the man who said that Mary had died *when*][$_{I'}$ reveal [$_{CP}$ *when* that Joe had left]]] - *Merger of tree B (the subject) to tree A (the predicate)*

(f) [$_{CP}$ *when* did [$_{IP}$ [$_{DP}$ the man who said that Mary had died *when*][$_{I'}$ reveal [$_{CP}$ *when* that Joe had left]]]] *Mvt. of <u>when</u> in Spec-CP to matrix Spec-CP*

(g) Two lower *when* copies delete to leave (11).

After copy deletion, (11) is (falsely) predicted to be a well-formed question about when Mary died. The reason *when* ends up to the right of the site it originates in is that it moves from tree B to tree A before there is a left/right relation between these trees. Moreover, by moving *when* back and forth between trees, no instance of movement violates any island, but the wh-phrase can return to a tree after the island has been built over its last copy in that tree. Thus surface outputs can violate islands and permit the launching site of movement to appear to the left of its landing site.[2]

Attempts have been made (e.g. by Nunes and by Bobaljik and Brown) to limit the scope of IM, but a variety of ad hoc assumptions about chains and copy sets are necessary. Moreover, even some of the more intractable properties of IM are embraced in approaches championed by Kayne (2002) and by Hornstein (2001) that are intended to account for the distribution of anaphora. I have argued in Safir (2003) that IM-based accounts for the distribution of anaphora, such as those proposed by Hornstein and Kayne, face problems so severe that they fail on their own terms. In Kayne's approach, for example, the obliteration of left/right distinctions in movement seem particularly poignant, given his general commitment to antisymmetry (see Kayne, 1994). The notion "left/right" can only be restored by a global requirement that checks fully built trees to insure that all previous operations were 'ultimately' leftward or not. This seems a high price to pay. Further problems of IM accounts I skip due to space limits.[3]

In spite of all the new derivations that IM makes possible, Extension in the IM theory still excludes both the tucking in analysis and the late attachment analysis. A brief reinspection of (5a,b), for example, reveals that late attachment cannot be generated, since a moved constituent is never undominated (every instance of Move embeds Merge), hence Extension blocks attachment to it.

From this much we can conclude that IM erases linear and island effects and permits a derivational theory of coconstrual with undesirable consequences, yet it still does not permit us to generate all of the analyses we need. Since IM is not useful and diminishes the restrictiveness of the theory, it seems best to make this class of theoretically undesirable derivations impossible. The provisional version of what I call *Insularity* in (13) suffices to achieve this effect, although other researchers have put it differently.[4]

(13) **Insularity:** If B is a copy that extends A, then A dominates every copy of B.

Let us assume, as is generally done in minimalist theorizing, that we know copies by virtue of their identical numeration indices.

As an immediate consequence of Insularity, the excesses of IM are now impossible, but now the full force of Extension returns, and all of the analyses discussed in section 2 are still excluded. We are back to square one: How can we preserve the salutary effect of Insularity and still permit the analyses we require?

At least one of the Extension-violating analyses does not require a reformulation of Extension, however. Covert movement, such as QR, does not have to be understood as post-cyclic movement. Rather we can assume that covert movement is unpronounced overt movement (i.e., all covert movements must extend the tree at the moment of their application, hence they are interspersed with pronounced movements). This is a conclusion reached by Bobaljik (1995, 2002), Pesetsky, (1998) and antecedents cited there, as well as many others since. An instance of QR would be derived as in (14).

(14a) [$_{IP}$ *someone* [$_{IP}$ everyone loves *someone*]] *After quantifier raising*

(b) [$_{IP}$ *someone* [$_{IP}$ everyone loves *x*]] *As interpreted at LF*

(c) [$_{IP}$ [$_{IP}$ everyone loves *someone*]] *Only lower copy pronounced at PF*

Rather than arguing this point further, I will simply henceforth assume that covert phrasal movement is regulated by Extension and that QR forms a chain that is pronounced at its lowest link. From this perspective, differences between languages purported to contrast with respect to whether a given movement is covert or not, contrast instead with respect to which copy in the chain formed by movement is pronounced. This fix, however, does not address the other analyses that violate Extension as enforced in a theory restricted by Insularity.

4. Introducing Exxtension

The key intuition, that operations are oriented toward the top of the tree at any point in the course of a derivation, can be respected once it is observed that all of the analyses in question involve Merge to a position *near* the top of the tree. The late attachment case is instructive, because late attachment, the addition of a modifying adjunct, is an operation that is not feature driven (hence not attracted by a strong feature) nor an instance of Move, yet it is still sensitive to the top of the tree. Consider (15a,b) (and compare with (6a,b)).

(15a) [[which picture of the women next to *each dancer*][should *he* turn to t]]?

Ans. The picture (of the women) on *his* left.

Ans.*The picture of the women who were *his* partners.

(b) [[which picture of the women next to *each dancer*][t are part of *his* personal collection]?

Ans. The picture (of the women) on his left.

Ans. The picture of the women who were his partners.

Where the adjunct describes the location of the picture, *next to each dancer* is modifying the whole phrase, *the picture of the women* or at least *picture of the women* (i.e., which is the *x* such that *x* is a picture of the women and *x* is next to each dancer). However, when a relative clause describing the women is employed in the answer, *next to each dancer* modifies *women* and ought to be able to receive a functional interpretation (e.g., what is the *x* such that *x* is a picture of *each dancer* standing next to *his* partners). In the latter case, late attachment, if it could attach so low within the fronted constituent, should not be part of the copy left in the trace position, but if that were true, then the answer describing a picture of the dancer standing next to his partners ought to be acceptable. Rather it seems that late attachment is sensitive to being nearer to the top of the tree, even if it does not adjoin at the very top. Example (15b) is a control that shows this distinction does not matter where the extraction is from a position which would not induce a crossover violation; In (15b) the pronoun *his* in the question does not c-command the trace. Hence attachment does not have to be late for the 'low' adjunct to be attached to *women* (i.e., *next to each dancer* could adjoin to *the women* before *of* and *each* are merged higher).

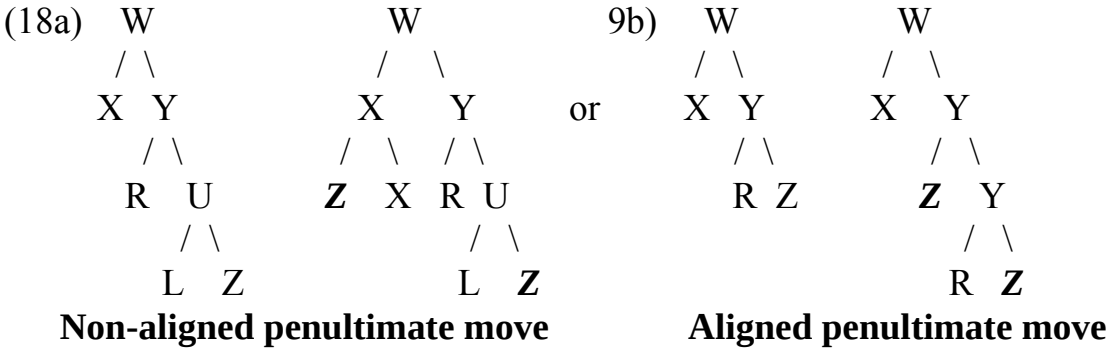
We now have enough of an idea of what the top of the tree is to reformulate Extension as *Exxtension*.

(16) **Exxtension:** Merge only to a crest node of phrasemarker P.

(17) **Crest nodes of phrasemarker P**: The undominated node P and its immediate daughters (i.e., assuming binary trees, a set of three nodes).

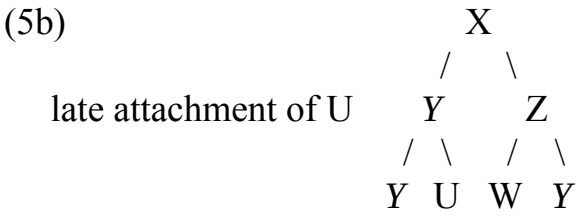
In what follows I distinguish cases where Merge adjoins to the undominated

node, which I refer to as ultimate merge, and cases when Merge applies to a dominated crest node, which I call penultimate merge. When Move is to a penultimate crest node, we see structures like those illustrated in (18).



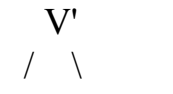
Non-aligned penultimate move **Aligned penultimate move**

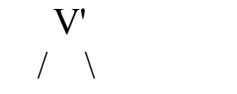
Incorporation is an instance of non-aligned penultimate move (18a) and tucking in is an instance of aligned penultimate move (18b). Late attachment is an instance of penultimate merge to a moved constituent that is a crest node



Without mitigation, penultimate merge (henceforth, *p-merge*) permits a wide variety of derivations that we might hope to exclude. For example, (19d) could be derived by recursive p-merge.

(19a) 

b) 

c) 

d) 

Derivations like that in (19) are possible because p-merge is building in new selectional relations that can result in well-formed subordinate structures if selectional relations can be established after p-merge, a possibility we can block by positing the *Selection Principle* (henceforth, *SP*).

(20) **The Selection Principle:** X enters into a selectional relation with Y iff X is merged to Y by initial ultimate merge.

By "initial ultimate merge" I mean that the first time that F_i (a formative with numeration index *i* or a phrase containing a string of such formatives with distinct indices, F_j, F_{j+1}, F_{j+2}, etc.) is introduced by ultimate merge is the only point where F_i establishes a selectional relation. Since movement is never *initial* ultimate merge, it can never create new selectional relations. Insofar as the SP regulates theta-assignment/selection based on how a form or phrase enters a derivation, it is a fundamentally derivational approach to thematic relations, as opposed to a theory that regulates such relations based on the output of structure-building.[5] Moreover, the SP derives the biuniqueness of the theta criterion, since no argument can be merged in more than one position and receive more than one theta-role in this system.[6]

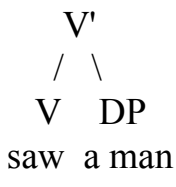
In this theoretical architecture, p-merge only permits modifiers to be

introduced when a selectional relation is not required. For example, any head p-merged onto an object will fail in its own thematic assignment.

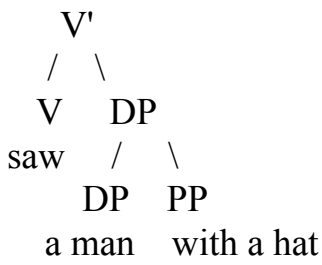
(21a) [saw [a man]]
(b) *[saw [with [a man]]

The penultimate merger of *with* with *a man* cannot establish a selectional relation between *with* and *a man* - thus, whatever selectional requirements *with* has (i.e., it must take a prepositional object) are not met. On the other hand, if *a man* is the complement of *saw*, then nothing merged to *a man* by p-merge could receive a distinct theta role from that assigned to *a man* and nothing subsequently p-merged with the verb would change the theta-role that the verb assigns to *a man*. By the same token, p-merge permits the introduction of modifiers.

(22a) [saw [a man]]



(b) [saw [[a man] [with a hat]]] - p-merge of PP



The PP modifies what it attaches to but cannot change the thematic status of what it attaches to. If meaningful material can only be introduced by p-merge if it modifies what it adjoins to or if it is thematically empty, then p-merge can be employed to introduce modifiers that must be prevented from entering into a selectional relation. For example, suppose some prepositions are introduced by penultimate merge only to satisfy Case requirements as in the case of *of* in

nominals. Thus p-merge of *of* does not alter the thematic relation between *picture* and *Bill* in (23b).

(23a) [picture Bill] - *Selection of Bill by picture*
(b) [[picture [of Bill]] - *P-merge, no new theta-assignment*.

P-merge can also be used to introduce adjectives to nominals and adverbs to VPs. If these are introduced by p-merge, then the fact that they do not head APs and AdvPs follows from the Selection Principle, since they cannot head any constituent they adjoin to - in contrast to selecting heads introduced by initial ultimate merge.

(24a) [$_{VP}$ snared a fish]
(b) [$_{IP}$ Past [$_{VP}$ snared a fish]] *Ultimate merge of T*
(c) [$_{IP}$ Past [$_{VP}$ deftly [$_{VP}$ snared a fish]]] *P-merge of adverb*
(d) [$_{IP}$ Past [$_{VP}$ [$_{VP}$ deftly [$_{VP}$ snared a fish]] [with the net]] *P-merge of PP*

Introducing T by ultimate merge permits T to establish a relationship of selection with VP (or whatever complement it takes, perhaps AgrP). Penultimate merge then "tucks in" adverbs without requiring that Adverbs must be generated in Spec positions, for example.[7]

Exxtension as it interacts with the SP is consistent with what we know of incorporation. Consider the incorporation of prepositions (applicative constructions in Baker, 1988) where the P of the PP head-adjoins to the V, as schematically represented in (25). The DP complement of P remains the complement of a copy of P in the trace position, while the PP itself remains the complement of V, such that the transparency of selectional relations is

undisturbed. Thus the interaction of Exxtension and SP derives the effect of Baker's Government Transparency Corollary, which was stipulated to preserve such transparency.[8]

(25) $[_{V'}\ [_{V}\ P\ V]\ [_{PP}\ t\ DP]]$

Following Baker, the same sort of relation could extend to *noun-incorporation* and *possessor-raising*, as well as various instances of head-movement that adjoin to functional categories, such as V-to-I or I-to-C. Thus penultimate move, given the Selection Principle, achieves just the right results for head movements.

4. Exxtension, Promotion and ACD

As mentioned in the introduction, Fox (2002) has proposed an analysis of ACD that, while attractive in its goals, is not consistent with Extension, as it involves late attachment. Although Fox also employs QR in his account of ACD, he assumes a phonological account of QR like that advocated by Bobaljik (2002) and others, hence his use of QR does not violate Extension. Although late attachment is now consistent with Exxtension, Fox's analysis also has the unfortunate consequence that the promotion analysis of relative clauses must be abandoned, or else he must allow for two rather different relative clause formation derivations, often producing what appears to be the same outward result. Rather than simply rendering his account consistent with Exxtension, then, it seems worthwhile to try to preserve the promotion analysis of relatives while still meeting the goals of Fox's analysis. What connects this effort to the goals of this paper is that certain properties of penultimate move are crucial to the result I seek to achieve.

Any account of ACD faces two basic problems. First, an antecedent for VP-deletion must be created and second, the antecedent should not introduce infinite regress. Examples like (26b), as opposed to (26a), permit elision without a discourse antecedent, which is possible if the analysis yields an antecedent for the ellipsis (marked with a dash) that amounts to some sort of VP that includes the verb *criticize* and its object.

(26a) Did John leave?

No, but Bill did /leave/

(b) I will criticize anyone you do /criticize *t*/.

The second problem is to show how this is possible without positing an infinite regress (if *criticize anyone you do* is the missing VP, then filling in *criticize anyone you do* in the position of the elision reintroduces an elided VP).

May's (1985) theory exploited QR, an adjunction of the direct object to VP, to move the ellipsis site out of VP and at the same time create a VP, [hire *t*], that could serve as an antecedent for the ellipsis, solving both problems at once.

(27a) I will hire anyone you do.

(b) I will $[_{VP}\ [_{DP}$ *anyone you do*__] $[_{VP}$ hire *t*]]

(c) I will $[_{VP}\ [_{DP}$ *anyone you do*__] $[_{VP}$ hire *anyone you do* __]]

With the advent of copy theory, however, the trace of QR still contains the ellipsis as in (27c), and infinite regress results again, as Hornstein (1995) points out. Fox avoids this problem by assuming:

(28a) Silent QR applying in narrow syntax creates a VP with an object gap.

(b) The relative clause coda adjoins to the quantified phrase raised by QR
(c) The launching site of QR (bolded) is pronounced.

The trick is to assume that the relative clause is not attached to *anyone* before QR, hence the copy that would otherwise cause regress is not introduced until the antecedent for the ellipsis has been created. Thus the derivation of (27a) proceeds as in (29), where both trees and bracketed representations are provided.

(29a) [$_{VP}$ [$_{VP}$ hire ***anyone***] [$_{DP}$ *anyone*]] - *Right-adjoined QR*
(b) [$_{VP}$ [$_{VP}$ hire ***anyone***] [$_{DP}$ [$_{DP}$ *anyone*] [$_{CP}$ *anyone* [$_{IP}$ you do _]]]]
-After late attachment to QR
(c) [$_{VP}$ [$_{VP}$ hire *x*] [$_{DP}$ [$_{DP}$ *anyone*] [$_{CP}$ *x* [$_{IP}$ you do _]]]] *-Trace interpretation*
(d) [$_{VP}$ [$_{VP}$ hire *x*] [$_{DP}$ [$_{DP}$ *anyone*] [$_{CP}$ *x* [$_{IP}$ you hire x]]]] - *Reconstructed elision*

(a) VP *QR before late attachment*
/ \ (*ultimate merge)*
VP DP
/ \ *anyone*
V DP
hire ***anyone***

(b) VP *After late attachment*
/ \
VP DP
/ \ / \
V DP DP CP
hire ***anyone*** *anyone* / \
anyone IP
[you do ---]

(c) *Trace interpretation*
VP
/ \
VP DP
/ \ / \
V DP DP CP
hire x *anyone* / \
x IP
[you do ---]

(d) *Reconstructed or licensed VP*
VP
/ \
VP DP
/ \ / \
V DP DP CP
hire x *anyone* / \
x IP
[you do hire x]

Fox's analysis faces the problems in (30).

(30a) QR is crucially right adjunction.
(b) Stipulates identity of relative clause nucleus with raised Spec-CP of relative clause coda (trace interpretation).
(c) PF-deletion is permitted outside the copy set, i.e., the *anyone* in the CP.
(d) The analysis is incompatible with the promotion theory of relative clauses.

The reason that QR must be right adjunction is that a relative clause attached to it surfaces in the phonology to the right of the quantified phrase, where it would be if it is attached to the silent raised quantified DP. Since the nucleus of the relative (*anyone*) arises from the object position of *hire*, it is not in a copy set relation with the raised *anyone* in the Spec CP of the relative clause coda, yet the *anyone* that is the head of the relative clause must establish an identity relation with the *anyone* in Spec CP. The promotion analysis requires no such device. Moreover, the deletion under identity at the point of trace interpretation must treat the Spec CP as part of the same copy set as the head of the relative, even though these two instances of *anyone* cannot be copies if they originate in different theta-positions (the *anyone* in Spec CP must have originated within the ellipsis site, and so I am assuming that *VP-deletion* means the licensing of null, but structurally present VPs). Once again, stipulating a copy set relation between the head of a relative and its other occurrences in the relative is unnecessary in the promotion account, since the relation between these forms is the copy relation by hypothesis. In fact, (30b) and (30c) are a consequence of (30d), the abandonment of the promotion account of relative clauses that this analysis requires.

The more general issue this raises is that the late attachment of relative

clauses to the relative clause nucleus is quite essentially incompatible with the promotion analysis of relative clauses. By contrast, late attachment of PPs has been shown to be empirically supported on subtle and convincing data, but late attachment of PPs introduces no such dilemma, since nothing the late attached PP adjoins to is posited to have its origin in the late attached constituent. Thus it is not late attachment per se that is at issue, but late attachment of relative clauses. On the other hand, there is also persuasive evidence that relative clause nuclei, especially for restrictive relatives, must be promoted, as I briefly review below. If this is true, then we must abandon the late attachment analysis of relative clauses and address in some other way the results achieved by any proposal that depends on this combination of analyses.

The promotion theory of relative clauses exploits copy theory to account for reconstruction effects that relate the heads of relative clauses to the gap position in relatives. Relative clause nuclei originate in relative clause codas (the CP) and are promoted to their surface positions (e.g., Schacter, 1973, Vergnaud, 1974, Kayne, 1994, amongst many others). This analysis is consistent with E(x)xtension.

Evidence for the promotion analysis and late attachment, where they do not conflict, is presented in (31) and (32) (based on Safir, 1999, 2004).

(31a)*Mary will select the picture of *each man* that *he* has just finished painting
(b) Mary will select the picture near *each man* that *he* has just finished painting

In (31a) we see a strong crossover (SCO) effect, such that Mary cannot be selecting the picture such that for each x, x a man, the picture is of x and x has just finished painting it. The SCO effect arise because *he* c-commands the object position of *painting* where a copy of *picture of each man* would reside.

When *each man* is embedded in a late-attached adjunct, as in (31b), the adjunct is not part of the trace and so *he* does not c-command *each man* at any point in the derivation. The same contrast is presented for weak crossover in (32).

(32a)*?The new warden found a picture of *each nazi* that *his* guard had should have confiscated.
(b)?The new warden found a picture next to the bed of *each nazi* that *his* guard should have confiscated.

Finally, there is evidence from the earlier literature that pronouns in the nuclei of relative clauses can be bound as if they occupy the positions they would have originated in (a classic reconstruction effect). In (33), *every soldier* c-commands *his* in its reconstructed position.

(33) The guards inevitably confiscate the photo of *his* family that *every soldier* treasures most.

The evidence for late attachment of relative clauses depends exclusively on the purported contrast reported by Freidin (1986) and Lebeaux (1990), which, though real for some examples, is undetectable for a wide range of speakers both for the cases in (34) (noted by Kuno, 1997) and others (see Safir 1999, 2004 for discussion and references).

(34a) Which criticism of *Lee* did *he* choose to ignore.
(b) Which evaluation of *Lee's* physical fitness did *he* use when *he* applied to NASA for space training?
(c) Whose allegation that *Lee* was less than truthful did *he* refute vehemently?

It appears that even for the complement PP case in (34a), the Principle C effect that is supposed to be forced by reconstruction (i.e., leaving a copy in the position of the trace such that the subject *he* could c-command it) is not empirically robust at all. In the reference cited, I argue that vehicle change of names to pronouns (first proposed for VP ellipsis by Fiengo and May,1994) within the unpronounced copy disarms the purported Principle C effect and so it cannot be used as evidence to support reconstruction for either complement PPs or complement relative clauses. If so, then the evidence of late attachment for relatives, that it suppresses the Principle C effect, disappears. Unlike the quantifier evidence for late attachment of PPs, however, there is no other evidence for late attachment or relatives, and hence no justification for it.

If we reject late attachment of relative clauses, however, then we must reject Fox's analysis of ACD. In the short space remaining, I just outline how an account of ACD based on p-merge can be fashioned to achieve the same results without the same drawbacks. Recall that we have two analytic desiderata, in line with May's original attack on the problem: an antecedent [V t] must be created and "t" cannot contain the ellipsis site or infinite regress occurs.

Suppose that relative clause promotion does indeed move the relative clause nucleus to Spec CP of the clause, such that the CP in (27a) is a complement of a D (silent in this case) as illustrated in (35a).[9] The whole relative clause is then a complement to *hire*, as in (35b). Suppose then that the relative clause nucleus in Spec CP is adjoined by non-aligned penultimate move to the V that selects the relative clause as its object, creating the VP [[hire *anyone*] [*anyone* that you do __]]. We know this is structurally possible for the case of head movement, but what is novel here is that a phrase (NP in Kayne's analysis) has adjoined to a head. Since the merge portion of penultimate move is p-merge in (35c), adjunction of *anyone* to *hire* does not affect the selectional

relations of either. Subsequent QR of *anyone* in (35d) (adjunction to VP, left or right doesn't matter, since it is not pronounced and not attached to the clause) will create a larger VP that embeds the one just described, including the subpart [hire *anyone*] which will correspond to, and license, the extraction site (marked between slashes).

(35a) $[_{DP}$ D $[_{CP}$ *anyone* $[_{C'}$ that you do /hire *anyone*/]]]

-Promotion derivation of relative clause

(b) $[_{VP}$ hire $[_{DP}$ D $[_{CP}$ *anyone* $[_{C'}$ that you do /hire *anyone*/]]]]

-Merger of relative CP with V

(c) $[_{VP}$ $[_{VP}$ hire ***anyone***] $[_{DP}$ D $[_{CP}$ *anyone* $[_{C'}$ that you do /hire *anyone*/]]]]

- Penultimate adjunction to <u>hire</u> of <u>anyone</u>

(d) $[_{VP}$ *anyone* $[_{VP}$ $[_{VP}$ hire ***anyone***] $[_{DP}$ D $[_{CP}$ *anyone* you do /hire *anyone*/]]]]

- QR of <u>anyone</u>

(e) $[_{VP}$ *anyone* $[_{VP}$ $[_{VP}$ hire *x*] $[_{DP}$ D $[_{CP}$ *x* you do /hire *x*/]]]]

- Q-variable creation (LF)

(f) $[_{VP}$ $[_{VP}$ $[_{VP}$ hire *anyone*] $[_{DP}$ D $[_{CP}$ *x* you do]]]] - copy deletion (PF)

If this is possible, then ACD is reconciled with the promotion analysis of relative clauses and all instances of *anyone* are in the same copy set, consistent with the fact that only one member of this copy set is pronounced (however that is determined - an issue in both Fox's theory and mine).

We can now conclude that if derivations like that in (35) can be justified, then ACD can be explained by non-aligned p-move without sacrificing the promotion analysis of relative clauses and without stipulating that QR must be right-adjunction.

Some of the moves made here will need further justification, however.

Not least, there is the question of what motivates the phrasal p-move to V. It would also appear that adjunction to V is a step that must be followed by QR of the adjoined head, as if QR has to move in two steps for some reason. Moreover, this sort of adjunction requires us to sacrifice the notion that like adjoins to like, in that I permit a phrase to adjoin to a head.

There are certain phenomena, however, for which analyses based on phrasal p-move to V might be attractive, such as Hindi focus movement, for which any account must explain why a focused phrase extracted from a postverbal tensed clause must occur strictly left-adjacent (for most speakers) to the matrix verb. Since Hindi is essentially SOV, apart from tensed clause complements, the order S-O-FOC-V is an embarrassment, especially if V assigns its theta-role to the object as its left sister.[10] The intrusion of a focused phrase to the left of V is immediately explicable from the perspective of p-move, which does not affect selectional relations.

There are also interesting questions to be raised concerning the right version of the promotion analysis for relative clauses, though these, as well as a wider range of ACD cases that need to be considered, I set aside for reasons of space.

5. Conclusion

Exxtension and Insularity work together with the Selection Principle to restrict all structure-building (any instance of Move or Merge) to the top of the tree, but with an extended notion of the top of the tree that permits many useful structural analyses, such as modifier merging, head adjunction, tucking in, late attachment, relative nucleus promotion, and antecedent creation for ACD. On the other hand, a wide range of IM analyses permitted by theories that lack Insularity are excluded. Additional consequences for copy sets, chain relations

and c-command definitions go beyond what can be presented in the space available, but also form part of the research program.[11] From the larger perspective, insofar as Exxtension is sensitive to moments in a derivation, any evidence for Exxtension is evidence for a derivational approach to syntactic relations.

Notes

*This is a preliminary report on work that I intend to present in a larger format eventually. I would like to thank audiences at Rutgers, MIT, the University of Connecticut at Storrs, and the audience at the Tokyo Conference on Psycholinguistics at Keio University for their questions and suggestions that have helped propel this research, though not all of their comments and questions are responded to here. My special thanks to Yukio Otsu for making my presence at TCP and my participation in this volume possible.

[1] Hisatsugu Kitahara points out (personal communication) that IM accounts cannot employ Chomsky's Agree, since Agree looks down only from the sister of the probe. Agree does not derive Extension, however, for cases where adjunction is not feature driven.

[2] Notice that intermediate movement to Spec CP that lacks an interrogative feature is necessary for any cyclic wh-movement. Thus, however the intermediate trigger arises, it would seem to be necessary in the same way for the intermediate CP in sentences like *Who did Sarah say that she saw?*

[3] Kayne (2002) proposes that all coconstrual should be effected by movement, but he claims that the ban on movement rightward and downward can then be used to predict Principle C effects. He suggests that pronouns always get their antecedency from referring expressions that originate in doubled constructions, such as [*Sol he*]. Referring expressions, if they originate low, could not move

from a c-commanding position to a c-commanded one for principled reasons, Kayne argues. Thus *Sol said Sue saw him* is derived from [*said Sue saw* [*Sol him*]] in the lowest position and movement of *Sol* into a higher subject position, perhaps already inhabited by *he* which may or may not be pronounced (depending on whether or not the name moves higher subsequently). Unfortunately, these assumptions permit the derivation in i. which would fail to capture a typical Principle C effect for a sentence like "*He* said Sue saw *Sol.*"

	Tree A	Tree B	
i.	[him]	[*Sol* he]	
ii.	[*Sol* him]	[*x* he]]	*-IM for coconstrual*
iii.	[said Sue saw [*Sol* him]]		*-Continued building of tree A*
iv.	[[*x* ***he***] [said Sue saw [***Sol*** (him)]]		*-Merger of Tree B with Tree A*

[4] Insularity is essentially Epstein et. al's (1998) "First Law" also discussed in Epstein (2001:334). They needed the First Law to insure that movement is always to a position that can be derivationally defined as c-commanding its trace. In the theory presented here, c-command is not defined derivationally because non-aligned p-move seems necessary and would not c-command its trace as a result of Merge (see note 11 for a formulation of c-command).

[5] This recalls Chomsky's (2000:103) less general proposal that pure merge in a θ-position is required of (and restricted to) arguments.

[6] Full interpretation still requires that every argument be interpreted and that every theta-role be assigned which needs to be assigned. Note also that Spec XP in this theory could be defined as initial ultimate merge to a nonhead, but there are some multiple specifier cases for which this might be too restrictive.

[7] Cinque (1999) argues that adverbs *should* be generated in Spec positions

associated with related heads, a position I am then committed to reject, though I have no refutation to offer here.

[8] See Baker (1988:450 fn.17) where the Government Transparency Corollary is speculated to be a potential consequence of some general property of movement. Although this transparency principle was stated in terms of government, it depended more on selection than government in practice.

[9] The relative clause analysis used in the text is roughly Kayne's, which has certain imperfections I shall not address here. See also Bianchi (1999) for extensive discussion. In future work I will argue for a more articulated left periphery of the relative clause.

[10] The assumption that theta-assignment is necessarily to the left in SOV languages is in dispute. See Zwart (1997) and Haider (1997) for discussion.

[11] For example, c-command can now be defined as in i., and the similarity to Insularity is then apparent in the definition in ii.

i. C-command: The c-command domain for x is the minimal node that dominates every copy of x.

This formulation insures that an adverb has scope over all it is a sister to, but only within the crest node it attaches to. For non-aligned p-move, the antecedent c-commands its trace because both crest nodes A and B dominated by undominated C will dominate copies of x, hence C will be the first node that dominates every copy of x. C-command in i. is reminiscent of the second part of (13), a similarity that can be exploited by reformulating Insularity as ii.

ii. Revised Insularity: Every lexical item in the numeration is c-commanded in X.

Revised insularity then derives Chomsky's (1995:189) condition that a single phonological object must be given to PF, rather than several phrases built from the same lexical array, such that each phrase converges individually.

References

Baker, Mark (1988) *Incorporation*. Chicago: University of Chicago Press.

Bobaljik, Jonathan (1995) Morphosyntax: The syntax of verbal inflection. Unpublished Ph.D. dissertation, MIT.

Bobaljik, Jonathan (2002) A-chains at the PF interface: Copies and "covert" movement. *Natural Language and Linguistic Theory* 197-267.

Bobaljik, Jonathan and Samuel Brown (1997) Interarboreal operations: Head movement and the Extension requirement. *Linguistic Inquiry* 28:345-356.

Bianchi, Valentina (1999) *Consequences of Antisymmetry: Headed relative clauses*. Berlin and New York: Mouton de Gruyter.

Chomsky, Noam (1995) *The Minimalist Program*. Cambridge: MIT Press.

Chomsky, Noam (2000) Minimalist inquiries: The framework. In Roger Martin, David Michaels and Juan Uriagereka (Eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik* (pp. 89-155). Cambridge: MIT Press.

Cinque, Guglielmo (1999) *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.

Epstein, Samuel David (2001) Un-principled syntax. In Samuel David Epstein and Norbert Hornstein (Eds.), *Working Minimalism*, (pp. 317-345). Cambridge: MIT Press.

Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima and Hisatsugu Kitihara (1998) *A Derivational Approach to Syntactic Relations*. Oxford: Oxford University Press.

Fiengo, Robert and Robert May (1994) *Indices and Identity*. Cambridge: MIT Press.

Fox, Danny (2002) ACD and the copy theory of movement. *Linguistic Inquiry* 33:64-96.

Freidin, Robert (1986) Fundamental issues in the theory of binding. In *Studies in the acquisition of anaphora* Vol. 1. 151-188. Dordrecht: Reidel.

Haider, Hubert (1997) Precedence among predicates. *Journal of Comparative Germanic Linguistics* 1: 3-41.

Hornstein, Norbert (1995) *Logical Form: From GB to Minimalism*. Oxford: Blackwell.

Hornstein, Norbert (2001) *Move! A Minimalist Theory of Construal*. Oxford: Blackwell.

Kayne, Richard (1994) *The Antisymmetry of Syntax*. Cambridge: MIT Press.

Kayne, Richard (2002) Pronouns and their antecedents. In Samuel David Epstein and T Daniel Seely (Eds.), *Derivation and Explanation in the Minimalist Program* (pp. 133-166). Oxford: Blackwell.

Kuno, Susumu (1997) Binding theory and the minimalist program. Unpublished ms., Harvard University, Cambridge.

Lebeaux, David (1990) Relative clauses, licensing, and the nature of derivation. In *NELS 20*: 318-332. GLSA, University of Massachusetts, Amherst.

May, Robert (1985) *Logical Form: Its Structure and Derivation*. Cambridge: MIT Press.

Nunes, Jairo (2001) Sideward movement. *Linguistic Inquiry* 32:303-344.

Pesetsky, David (1998) Some optimality principles of sentence pronunciation. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky (Eds.), *Is the Best Good Enough?* (pp. 337-383). Cambridge: MIT Press.

Richards, Norvin (1997) What moves where when in which language? Unpublished Ph.D. dissertation, MIT.

Richards, Norvin (1999). Featural cyclicity and the ordering of multiple specifiiers. In Samuel David Epstein and Norbert Hornstein (Eds.), *Working Minimalism*, (pp. 127-158).Cambridge: MIT Press.

Ross, John (1973) Nouniness. In O. Fujimura (Ed.), *Three Dimensions of Linguistic Theory*,. Tokyo: Tokyo Institute for Advanced Studies of Language..

van Riemsdijk, Henk and Edwin Williams (1981) NP-structure. *The Linguistic Review* 1: 171-217.

Rudin, Catherine (1998) On multiple questions and multiple wh-fronting. *Natural Language and Linguistic Theory* 6:445-501.

Safir, Ken (1999) Vehicle change and reconstruction in A'-chains. *Linguistic Inquiry* 30: 587-620.

Safir, Ken (2003) Anaphors, movement and coconstrual. In Lars-Olaf Delsing, Cecilia Falk, Gunlög Josefsson, and Halldór Sigurðsson (Eds), *Grammar in Focus Volume II: Festschrift for Christer Platzack* (pp. 283-293). Department of Scandinavian Languages, Lund University, Lund.

Safir, Ken (2004) *The Syntax of (In)dependence*. Cambridge: MIT Press.

Schacter, Paul (1973) Focus and relativization. *Language* 49: 19-46.

Vergnaud, Jean-Roger (1974) French relative clauses. Unpublished Ph.D. dissertation, MIT.

Zwart, C. Jan-Wouter (1997) *Morphosyntax of Verb Movement*. Dordrecht: Kluwer.

(safir@ruccs.rutgers.edu)