In Sachiko Aoshima, John Drury and Tuomo Neuvonen, eds. 1999.
University of Maryland Working Papers in Linguistics 8, 63-103.

Merge Over Move and the Extended Projection Principle

JUAN CARLOS CASTILLO
JOHN DRURY
KLEANTHES K. GROHMANN

Abstract

A class of proposals are examined that aim to avoid problems that arise in various instantiations of the "Merge over Move" (MOM) cost-of-operation distinction. It is concluded that while the mechanisms introduced there exhibit independently interesting features, they subtract substantially from the interest of the MOM economy of derivation explanations. The removal of an assumption will then be considered that makes the core cases involving *there*-constructions a problem to begin with: that non-finite T must host a specifier position (checking an EPP/D-feature). Denying the existence of such features removes the problem that the MOM distinction was introduced to solve, allowing the core cases of associate-movement vs. expletive-insertion to arise as a case of true optionality. Consequences for other phenomena are examined and the proposal is found to be consistent with much recent research investigating these phenomena.

Introduction

Notions of economy and optimality have been much discussed in recent linguistic theory. Our topic regards a particular kind of derivational economy condition. Specifically, we examine the "Merge over Move" preference (henceforth, MOM). The idea is familiar within Chomsky's (1995; 1998) Minimalist Program: Move is a more expensive operation than Merge and derivational steps at which both operations are possible are required to choose the 'cheaper' Merge operation as the next step. The function of such a condition is to force derivations in a certain direction when indeterminism arises. The view advanced in Chomsky (1995) has it that such a condition is inter-/transderivational as it requires reference to a set of competing routes through a space of possibilities. While transderivational conditions are not new, this cost-of-operations distinction is perhaps one of the most interesting novelties of minimalist investigations.

There has been much discussion around the issue of what the proper formulation of such a condition should look like (or whether there should even be such a condition) and some concerns have arisen. The view that MOM presides over transderivational 'cost' comparisons has been argued to introduce potentially serious complexity. Furthermore, the original conception demands inconsistent interpretations of the putative comparison principles across a range of standard cases involving expletives. We will refer to these as the COMPLEXITY and INCONSISTENCY problems. Solutions to these problems have been offered but we believe that the mechanisms they evoke subtract from the interest of the MOM distinction. We explore another route: eliminating so-called EPP-features.

This essay is organized as follows. Section 1 outlines the initial formulation of MOM and points out the nature of the problems.[1] With these on the table, section 2 critically examines various proposed responses to the difficulties. Several of these have accumulated in the recent literature, some of which formulate derivational restrictions relating to the familiar and general notion of cyclicity. We conclude that while each of these approaches introduce novelties of considerable interest, collectively they suggest that the intended salvage of the MOM distinction is a mistake (a conclusion that we share with some of these other researchers whose work we mention). In most cases the additional mechanisms over-determine the data if MOM is retained. Our discussion thus points us back to MOM's evidential base, which, as has been pointedly observed by others, involves really just one core type of contrast:

(1) a. There seems to be a man in the room.
    b. *There seems a man to be in the room.

We deny the need of a derivational economy explanation for (1a,b) by exploring the possibility of eliminating the EPP. If, in this particular case, the embedded non-finite T(ense) does not host a feature demanding a specifier, then the contrast above is immediately explained away (i.e. non-finite T cannot license a specifier). Of course it is required that we examine this potentially eliminable component of the theory and what role it in fact plays. "EPP-features," we submit, do nothing but 'code' a mysterious residue from antecedent frameworks. Thus, we aim to try living without it. But it turns out that the EPP, ill-motivated though it may be on conceptual grounds, is not to be trivially dispensed with. Section 3 thus takes on the burden of showing that the cost to analysis is not disastrous. While there are undeniably problematic cases which arise, we find the general direction of inquiry to be in concert with some recent studies that pursue novel conceptions of Control and Raising.[2] We close with a discussion of the consequences our conclusions and suggestions have for the architecture of the grammar.

---

[1] This essay focuses only on the INCONSISTENCY problem for three reasons. First, the complexity concerns have been thoroughly discussed elsewhere; notably Chomsky (1998), Collins (1997), Frampton & Guttman (1998), Johnson & Lappin (1997), and Yang (1996). Second, the discussion that we have seen on these issues simply has not convinced us that the formulations of economy principles that involve global sorts of optimization (with whatever attendant complexity) are conceptually undesirable. We are unmoved by the suggestion that a competence-level account just should not have such properties. Third, the questions about complexity turn out to be moot for us since we believe that a proper resolution of the INCONSISTENCY problem leads to a conception under which the COMPLEXITY problem does not arise (and which lead to our discussion of the EPP).

[2] For example, Hornstein (1999, forthcoming) and Manzini & Roussou (forthcoming). In addition, independent conceptual arguments offered by Epstein & Seely (1999) regarding the status of chains converge on the same conclusions that we reach here through our examination of the MOM distinction (and the proposals designed to consistently retain it). We find this convergence (from two relatively distant conceptual starting points) to suggest that the elimination of the EPP is the correct direction in which to advance inquiry and conclude that, at least on a first pass of some relevant phenomena, that it seems to be an empirically sustainable course. Epstein & Seely also cite a manuscript by Martin (1999) which aims to eliminate the EPP.

1 Motivations and Problems for MOM

Take MOM to be jointly constituted by the following: (i) Merge is cheaper than Move and (ii) at any point in a derivation where both Merge and Move are both applicable, the cheaper operation is chosen. The cost distinction can be motivated by understanding Move as a composite or conspiracy of operations that includes Merge as a sub-part (e.g. Move = Copy + Merge (+Delete) (+Form Chain)) so it is simply a matter of 'more' vs. 'less' (Chomsky, 1998 and others).

Deploying this cost distinction requires some formulation of a notion of comparison of derivations (or, minimally, a notion of comparison of possible operations at a given derivational stage) so that there is a way of assessing what counts as an optimal computation vis-à-vis its (sub-optimal) alternatives. To start, Chomsky (1995:220-221) proposes that economy conditions only compare convergent derivations. The reason appears straightforward: Chomsky is operating within the assumptions of a "weak" derivational system that relies on both a series of structure building operations (i.e. a derivation) and levels of representation which evaluate the output of a given derivation to ensure that it is convergent. Thus, if the operations of the computational system (C_{HL}) are optimal in the sense that fewer, less costly procedures are favored, then the maximally economic derivation is one in which no operations apply.

This demands a comparison of candidate derivations that are convergent output structures which inspects their respective derivational histories to determine which one(s) were the least costly. But surely we do not want derivations corresponding to shorter sentences with less words in them to rule out longer ones. So, a way to begin to restrict the class of comparable convergent derivations—call this class the Reference Set ($\mathfrak{R}$)—is to only allow comparison between derivations that involve identical sets of lexical choices. Call the set of choices implicated in a given derivation an ARRAY. Chomsky (1995) augments this notion of ARRAY with subscripted integers for indicating the number of tokens of each item which will participate in a derivation (e.g. the_2 = *the* will be used two times, as in *the boy saw the girl*). These augmented arrays he calls NUMERATIONS (NUM).

Further restrictions are plausible. For example, the possibility that the system could compare *John loves Mary* and *Mary loves John* for economy purposes might be excluded. Note also that a NUM with k lexical tokens (i.e. k = the sum of the subscripted integers) will in the limit correspond to $k!$ possible Select & Merge orders and therefore at least $k!$ possible derivations to compare. This becomes even more intractable if Move and Merge operations interleave (as pointed out by Johnson & Lappin, 1997 and others). To reduce the complexity that follows, Chomsky proposes to further constrain $\mathfrak{R}$ by only allowing comparisons between derivations that, at each timestep t, have identical numerations. This is best explained by way of tracing the derivation for the crucial case in (1), repeated below.

(1) a. There seems to be a man in the room.
b. *There seems a man to be in the room.

Up to a point, the derivations of (1a) and (1b) are identical, and thus the derivations are (up to that point) comparable. Consider the relevant point in (2):[3]

(2) $\text{NUM}_{(1a,b)}$ = {there_1, T_1, seems_1}
[to be [[a man] [in the room]]]

At this point, there are two options: Merge of the expletive or Move of *a man* ((3a) and (3b), respectively). If MOM is enforced at this stage, then the less expensive option (3a) will outrank (3b). Although both derivations arguably converge if viewed independently, this local comparison has the outcome that only the continuation of (3a) will be optimally well-formed (shown in (5) below).

(3) a. [there [to be [[a man] [in the room]]]]
b. $[[\text{a man}]_i$ [to be [t_i [in the room]]]]

Also note that this requires that MOM to be enforced at the particular stage of the derivation pictured in (2). That is, it must be locally enforced. Again, although it is true that under these assumptions both derivations converge (viewing each independently), and that convergence is established at the interfaces (at the end of the derivation), it is also the case that following the stage in (3a,b) these derivations will be non-comparable in virtue of having different numerations:

(4) a. $\text{NUM}_{(1a),(3a)}$ = {T_1, seems_1}
b. $\text{NUM}_{(1b),(3b)}$ = {there_1, T_1, seems_1}

This creates a problem—convergence is by its very nature a property of derivations stated over outputs, whereas the comparison metric must decide locally. Further, the final outputs in (1a,b) both involve an instance of overt movement (*there* in (1a) and *a man* in (1b)) and thus are equally costly viewed globally/as outputs. This suggests understanding comparison as being about potential next-step continuations of a single derivation.[4]

(5) a. [seems-T [there [to be [[a man] [in the room]]]]]]
b. $[_{TP}$ there [seems-T [there [to be [[a man] $[_{PP}$ in the room]]]]]]
c. $[_{TP}$ there_i [seems-[T-$[\text{FF}_{\text{a man}}]$] [$t_i$ [to be [[a man-t_{FF}] $[_{PP}$ ...]]]]]

[3] Here and below we will abbreviate numerations in this way (leaving out items whose indices have been reduced to zero).
[4] We include a step (5c) in which matrix T is linked to the associate via feature movement as in Chomsky (1995), but the details of this do not matter for this part of the discussion. In section 3 we will be forced to reconsider the nature of the expletive-associate relation following our proposed abandonment of the EPP.

However, we now encounter an analytical quandary that highlights the nature of the INCONSISTENCY problem alluded to earlier.[5] Consider the following cases involving expletive *it*:

(6) a. *John seems that it was in the room.
b. It seems that John was in the room.

The relevant timestep of the derivation is depicted in (7), which we can view as the source derivation for both (6a) and (6b). And, like the case illustrated in (1)-(5) above, there is the option of applying either Merge (of *it*) or Move (of *John*).

(7) $\text{NUM}_{(6a,b)}$ = {it_1, T_1, seems_1, that_1}
[T-was [John [in the room]]]

(8) a. [$_{\text{TP}}$ it [T-was [John [in the room]]]]
b. [$_{\text{TP}}$ John_i [T-was [t_i [in the room]]]]

But locally enforcing MOM makes the wrong prediction here. As we saw above, there is a conflict between the notion of only comparing convergent derivations and the fact that MOM is relevant to particular derivational steps. The contrast in (6) makes this conflict become more than just a conceptual worry. The cheaper alternative in (8a) involving the Merge of the expletive leads to a crash in (6a)—a case of Superraising/Minimal Link Condition violation. But the fact that (6a) crashes cannot be relevant for the contrast in (6) if MOM is forced to act locally to account for (1a,b) above. And we cannot respond to this case by globalizing the evaluation so as to take into account the eventual non-convergence of the locally cheaper alternative (8a), because we then lose the account of (1a) vs. (1b). (Recall from above that (1a,b) both converge and are globally equally costly.[6])

Chomsky (1995:344-348) notes the difference between these cases (the *there* vs. *it* examples of the type in (1) vs. (6)) but does not seem to consider the tension we have noted to be an issue. He mentions that "we select Merge over Attract/Move if that yields a convergent derivation, irrespective of the consequences down the road as long as the derivation converges." But convergence is exactly the sort of thing that happens "down the road." Local

[5] Epstein et al. (1998:11-12) make the same kind of conceptual point that we are about to raise—what we are referring to as the INCONSISTENCY problem. They note that the reference to legitimate vs. illegitimate interface objects is paradoxical when paired with the usual conception of Last Resort which characterizes a notion of "licit rule-application" in terms of the eventual output object conceived as a level of representation. As they put it: "the requirements of the interfaces [...] apply exclusively at the interfaces, and not internal to a derivation—yet Last Resort, requiring that rules apply only to yield 'legitimacy', entails that legitimacy must be defined [...] within a derivation." How we understand the notion of "convergence" will be an issue throughout this essay.

[6] Again, keep in mind that we will just deny that both (1a,b) converge in section 3.

evaluation—if it is to work properly—seems to require sensitivity to distinctions stated over the outputs of derivations/at the interfaces (i.e. whether or not a derivation converges), in order to decide whether a more costly option can be pursued. Put another way, whichever way one goes on the COMPLEXITY issue (local vs. global evaluation) the INCONSISTENCY problem arises.[7]

Matters are complicated further by the following type of example:[8]

(9) a. There was [a rumor [that a man$_i$ was t$_i$ in the room]] in the air.
b. [A rumor [that there was a man in the room]]$_i$ was t$_i$ in the air.

These pairs of examples offer apparent cases of optionality that are puzzling for a local view of MOM (even if we resolve the INCONSISTENCY problem somehow). Take (9a,b): these are derived from the same numeration and are both perfectly acceptable. But (9b) should block (9a) because insertion of *there* in the *that*-clause associated with *a rumor* should be forced by MOM (yielding (9b)) if this condition involves local assessment. Thus, the movement of *a man* in the *that*-clause in (9a) should be locally vetoed by MOM.

Summarizing then, the initial cases involving *there*-constructions insist that MOM is enforced at the relevant intermediate stages of the derivations where there is a choice between operations. But if we take seriously the apparently sensible idea that only convergent derivations enter into comparison and if convergence is understood to be a property of the eventual outputs (i.e. at the interfaces) then we have got a conflict on our hands. This conceptual worry is rendered concrete by the examples involving expletive *it* which seem to require exactly this sensitivity to the eventual interface status of the outputs of derivations. And, if we adjust our understanding of these principles to allow this sort of global sensitivity (look-ahead) that simultaneously takes into account the interface/convergence status of derivations as well as their respective derivational histories, then we lose the account of the initially problematic cases (1a,b). On top of all this, even if these worries are resolved, we still have the apparent optionality of the more complicated examples of the sort in (9) to contend with. These are not new problems. As noted earlier, these are exactly the kinds of difficulties that have driven various alternative accounts. The kinds of systems that have been devised to deal with the problems are of some interest to us, because we think that they all suggest that the cost-of-operations conditions are unnecessary and in fact suggest an approach which eliminates the EPP. We turn next to a class of solutions to these troubles beginning with our own previous work (Castillo, Drury & Grohmann, 1997, hence CDG).

[7] But perhaps we have the wrong notion of convergence in mind. Chomsky (1998:12) has the following to say on this matter: "[…] an embedded clause may converge, for example, the bracketed subpart of 'John thinks [it is raining].' The phrase 'converge at an interface' should not mislead: *convergence is an internal property of an expression, detectable by inspection* [italics ours—CDG]." Keep this comment in mind. We will return to this at various points below.

[8] (9) is attributed to Juan Romero and Alec Marantz in Uriagereka (1999b) and also discussed by Frampton & Guttman (1998).

2 Some Solutions that Retain the MOM Cost Distinction

All of the responses to the problems with the MOM distinction that we are aware of involve one of two basic strategies. They either reject altogether the comparison technology involving numerations for the core case in (1) resorting instead to other mechanisms, or they introduce constraints which effectively partition derivations into sub-parts in order to reduce the more complicated cases to the simpler ones. This latter strategy has itself been approached in two ways. The first amounts to conditions on the formation of numerations that has a reflex in the establishment of comparison possibilities. The second strategy attempts to motivate a partition of derivations based on plausible (LF) interface concerns and asserts that certain structural sub-domains constitute objects that the interface is forced to (cyclically) evaluate. This second strategy is the one that we pursued in CDG. We review this next.

2.1 Cyclic Convergence Evaluation The starting point of the inquiry in CDG was the following observation about the cases in (1) and (6). Firstly, (1a) and (1b) both converge as outputs under standard assumptions, whereas in (6) only (6b) converges. Secondly, the choice-points for (1) vs. (6) involve non-finite T versus finite T, respectively. These observations point the way to a solution that allows the retention of both the local character of economy comparison as well as the notion that only convergent derivations are compared. The proposal evokes something akin early notions like "kernel sentences" (Chomsky, 1955/1975) or cyclic nodes (Chomsky, 1973; Freidin, 1978 inter alia). Specifically, CDG understood convergence evaluation to take place at each derivational occurrence of (satisfied) finite T.

The view has a kind of intermediate status between what Chomsky (1995) identifies as "strong" vs. "weak" derivational systems. The former constitutes an approach that involves only "pure" derivation and direct step-by-step feeding of the grammar-external systems. "Weak" derivational systems involve the notion of "level of representation" which assumes that the syntax interfaces with the grammar-external systems at a point conceived as a unified syntactic object that is evaluated, as it were, "all at once." CDG's view was (following Uriagereka, 1999b) that evaluation proceeds in sub-parts—basically tensed-S by tensed-S (what CDG dubbed "T-Domains").[9]

CDG thus claimed that sub-parts of derivations form units of a type that the LF-interface cares about, thus forcing a requirement to evaluate their well-formedness as these units are established. We argued that one such sub-part that forces this kind of cyclic convergence evaluation corresponds to derivational points at which a TP headed by finite T has been constructed; these can be

[9] Here we follow Uriagereka (1999b), but with a different kind of motivation. Uriagereka abandons the once-per-derivation view of Spell-Out in the course of simplifying the PF-driven linearization procedure. Our motivation comes from the LF-interface—it evaluates T-Domains. The resulting model is thus similar to that in Jackendoff (1972), where the interpretive component(s) are 'fed' from cycle to cycle. See also section 2.3.

understood as domains that correspond to earlier assumptions about nodes defining cycles or to even earlier notions of kernels. CDG assume that when derivations have constructed a finitely-headed TP, the LF-interface insists that convergence is checked.

Consider how this works for the simple cases with the *there*-expletive in (1), where the expletive is merged, and the *it*-expletive in (6), where the subject moves. For (1), at the derivational choice-point (repeated here as (10)), the derivation has not reached an occurrence of finite T, and MOM works exactly as it is proposed in Chomsky (1995) (it locally steps in to enforce the ranking of (11a) over (11b)).

(10) $\text{NUM}_{(1a,b)}$ = {there$_1$, T$_1$, seems$_1$}
[to be [[a man] [in the room]]]

(11) a. [there [to be [[a man] [in the room]]]]
b. [[a man]$_i$ [to be [[~~a man~~]$_i$ [in the room]]]]

This remains the unproblematic case since, viewed separately, both (11a) and (11b) could each continue to convergence.[10] Thus the ill-formedness of (1b) (the continuation of (11b)) results from a derivational cancellation (or alternatively, results from there being no route to this output given that the choice in (11a) is locally enforced). The more expensive (11b) is simply not allowed to continue to a point where convergence would be evaluated (i.e. it does not reach the interface/cannot become a T-Domain).

Now consider the case in (6), with the local choice point repeated here in (12) and the possible continuations in (13):

(12) $\text{NUM}_{(6a,b)}$ = {it$_1$, T$_1$, seems$_1$, that$_1$}
[T-was [[John] [in the room]]]

(13) a. [$_{TP}$ it [T-was [[John] [in the room]]]]
b. [$_{TP}$ [John]$_i$ [T-was [[~~John~~]$_i$ [in the room]]]]

In contrast to the *there*-cases above, here finite T has been introduced and according to CDG, convergence should be evaluated at the point in (13). Assuming that in (13a) *John* will not have checked its Case-features (since the expletive will have served to check Nominative Case, making even FF-movement impossible), this derivation fails as a non-convergent T-Domain. The more expensive alternative (13b), on the other hand, is well-formed as it is and thus this derivation can continue. Note that by localizing the notion of convergence evaluation, we enable the system to make use of the sensible restriction that derivations which do not converge do not count for comparison. Crucially this

[10] Again, please bear in mind that we will reevaluate this assumption in section 3.

happens without the need for any look-ahead since, under this proposal, convergence is not decided all at once over eventual outputs, but over sub-parts (T-Domains).

In section 1 we noted a conflict, namely that on the one hand, MOM is local to derivational steps as is necessary for the account of (1) and, on the other hand, we must somehow capture the notion that only convergent derivations are compared. This dilemma is resolved in CDG.

Now consider (9) from section 1, repeated here as (14):

(14) a. There was [a rumor [that a man$_i$ was t$_i$ in the room]] in the air.
b. [A rumor [that there was a man in the room]]$_i$ was t$_i$ in the air.

If convergence is derivationally assessed for T-Domains, (14) constitutes a puzzle. Consider them at the timestep depicted in (15):

(15) NUM$_{(15a,b)}$ = {there$_1$, was$_1$, T$_1$, a$_1$, rumor$_1$, that$_1$}
[was-T [$_{SC}$ [a man] [in the room]]]

(16) a. [[a man] [was-T [$_{SC}$ [a man] [in the room]]]]
b. [there [was-T [$_{SC}$ [a man] [in the room]]]]

At the point in (15) we have introduced finite T and thus, once it has been satisfied (i.e. has a specifier), we have established a domain at which convergence must be evaluated. And, as in previous cases, there is the option of applying Merge (or the expletive) or Move (or the associate). But given the considerations just discussed, the locally cheaper option will converge as a licit T-Domain at the point depicted in (16b), and the derivation will continue. Here is the apparent problem: this means that (14b) should cause the derivation leading to (14a) to cancel at the point shown in (16a), since that derivation requires the locally more expensive choice of moving *a man*. But both (14a) and (14b) are acceptable; in fact, they are both locally grammatical (i.e. they both involve locally well-formed T-Domains at the points depicted in (16a,b)). Thus, the continuation of the derivation (16b) will go as in (17), while the derivation from the point in (16a) is forced to halt according to our proposal—that route is (contrary to fact) locally cancelled.

(17) a. [there [was-T [$_{SC}$ [a man] [in the room]]]][11]
b. [that [$_{TP}$ there [was-T [$_{SC}$ [a man] [in the room]]]]]
c. [$_{NP}$ [a rumor] [that [$_{TP}$ there [was-T [$_{SC}$ [a man] [in the room]]]]]]
d. [in the air]

[11] The movement of the formal features of the associate (or the application of Chomsky's Agree) is omitted from this derivation, but it occurs before the T-Domain in (19a) is evaluated.

e. [$_{SC}$ [$_{NP}$ [a rumor][that [there [was-T [[a man] [in the room]]]]]] [in the air]]

Note: (17d) is constructed separately, then merged with (17c) to form (17e).

f. [was-T [[[a rumor] [that [there [was-T [[a man][in the room]]]]]]] [in the air]]

g. [[[a rumor] [that there was-T a man in the room]]$_i$ was-T [t_i [in the air]]]

Note: [$_{NP}$ [a rumor] [that...]] *raises from the small clause to become matrix subject.*

The comparison logic forces a conclusion about this kind of puzzle: these two derivations must not be comparable (or, if comparable, must somehow be equally costly). The point is quite general. If derivations are locally economical and if local economy if enforced as we have suggested (within a system that relativizes convergence to T-Domains), then the apparent step-by-step optionality exhibited by these cases has to be an illusion—some independent considerations must require that these two derivations take divergent routes and thus are partitioned into distinct reference sets. Given the comparison reasoning and the cost of operations idea, these are really the only coherent ways to think about this kind of optionality.

In CDG we offered an argument that these problematic (a)/(b) pairs are in fact non-comparable (or equally costly—see below), based on the adjunct status of the relevant *that*-clauses in each. The argument is summed up in (18):

(18) a. Given the bottom-up nature of structure building, complex adjuncts must be assembled as separate subtrees before being integrated with other structure

b. and nothing forces adjuncts to be constructed before or after other material is combined; these different orders of construction split derivations into distinct reference sets,[12]

c. so, if the adjunct is built first, MOM forces expletive insertion in the adjunct; if the matrix is constructed first, MOM forces expletive insertion in the matrix.

That is, one obvious technical way to make the troublesome derivations for (14) non-comparable is to factor their major parts into sub-derivations. This is independently required given the strict Extension Condition—complex non-complements must be separately constructed before they are merged/adjoined to other structure. And, since nothing forces the construction of the adjunct before anything else, there is thus a kind of optionality that is built-in to this system.[13]

[12] Chomsky (1998:38) notes this property of the system in his discussion of phases, noting that, for example, the construction of a DP and a modifying relative clause is essentially "unordered," but does not explore the effects of this fact for the comparison logic for economy.

[13] Basically this is just the same (boring) optionality as there is, for example, in the assembly of the DP *the man* where the determiner can be Selected first, or the noun can be. In the case we are examining, the

The adjunct status of the *that*-clause in (14a,b) is suggested by the impossibility of extraction shown in (19):[14]

(19) a. John heard a rumor/a claim that Bill kissed Monica.
b. *Who$_i$ did John hear a rumor/a claim that Bill kissed t_i.

To make the logic perfectly clear, suppose that we have reached the following stage of the derivation which could continue to form either (14a) or (14b):

(20) NUM$_{(15a,b)}$ = {there$_1$, that$_1$}
[was-T [$_{SC}$ [a rumor] [in the air]]] — [was-T [$_{SC}$ [a man][in the room]]]

At this point both the adjunct (to be) and the matrix clause have been separately constructed (as mandated by the Extension Condition) and have both reached the point at which there is an occurrence of finite T. Now MOM forces the system to Select and Merge the expletive *there*, but no principle forces the targeting of one or the other subtree thus far established. So whichever of the two is targeted, movement will have to occur in the other, thus yielding both derivations.[15]

To sum, in CDG we propose retaining the MOM distinction and the economy explanation for (1) by claiming that convergence is evaluated for TPs in a cyclic fashion.[16] This allowed us (i) a straightforward way to maintain that only convergent derivations are compared (by "bringing convergence to the relevant choice points"); (ii) to avoid the COMPLEXITY concerns (whatever their ultimate force); (iii) to maintain a consistent interpretation of the evaluation mechanisms.

Further, the apparent counterexamples turn out in the end to be consistent with our story following a closer inspection of their derivations and the effects of the Extension Condition on the construction and integration of complex adjuncts.

As mentioned earlier, there are other very similar approaches. Setting aside proposals that simply abandon the economy explanations (or that attempt to get their effects in other ways), there are two other notable attempts that are only subtly different from CDG, Chomsky (1998) and Uriagereka (1999a; 1999b).

Chomsky's approach, as we will indicate, seems to us to be conceptually

---

material corresponding to the *that*-clauses in the problematic examples can be built first or second, the system does not have any reason to choose one or the other route. This means that MOM can force the expletive insertion in the sub-structure that is (optionally) built first, then the more complex movement operation is the only option when the second sub-structure is built.

[14] These have been proposed to be adjuncts by Stowell (1981), Grimshaw (1990) and others. This of course goes against the conclusions of Freidin (1986) and Lebeaux (1988) that these are complements based on alleged reconstruction contrasts. This is beside the point at the moment since the worry in the text is about the logic of economy comparison and not about reconstruction.

[15] There is a wrinkle here regarding the Extension Condition, which is not relevant to this paper since we will be introducing a different analysis in section 3 that does not require these cyclicity-type conditions on convergence evaluation.

[16] This logic may be subsumed by a more general or differently articulated understanding of cyclicity (as in Epstein et al., 1998; Frampton & Guttman, 1998; Chomsky, 1998). We will return to this possibility.

similar to CDG, but differs in the claims about what domains constitute "special" sub-parts of derivations (for CDG these were finite TPs; for Chomsky, these are *v*P and CP). We turn first to Uriagereka, and then to Chomsky.

2.2 Kernel Numerations Uriagereka (1999a) introduces the following notion:

(22) Kernel Numeration:
A kernel-numeration is the minimal multi-set of formal features all of whose grammatical properties are satisfied within its derivational horizon.

Broadly, the idea is that numerations are 'minimally' formed. There are two important details that must be added. First, it must be the case that

(23) Syntactic derivations 'start' in kernel numerations.

That is, there must be some statement that kernel numerations are the only sort of numerations that can correspond to derivations for economy evaluation purposes. Second, it is important that it is the multi-set of formal features and not "lexical items" that is being 'minimized' in the above formulation. This is what allows the economy reference to "convergence" to work without prohibiting expletives from *ever* entering a numeration. After all, expletives are not necessary for convergence. The problem (that arises under a version of (22) with "lexical items" substituted for "formal features") was pointed out by one of the present authors, and for ease of reference we will follow Uriagereka calling this "Castillo's Problem."

But consider the following. It has to be the case under this convergence-sensitive economy view of numeration formation that the expletive *there* has no features. If it had Case and/or agreement features, as Uriagereka is careful to point out, then it would run straight away into Castillo's Problem, since *there*'s inclusion in any numeration would not be "minimal" (i.e. any derivation can in fact do with 'less'). This also entails that *there* cannot have any categorial features either, contra many standard accounts that assume it to correspond to (just) a D-feature (which of course goes along with coding the EPP as an illegible D-feature on T). But Uriagereka is forced to deny this standard view, so now a pretty serious question arises: what *is* this expletive? Uriagereka (1999a; 1999b) unfortunately does not tell us, though he does mention Chomsky's desire to "eliminate categorial features and principles that make reference to them (e.g., the EPP), at least as syntactic elements." So we are left with something of a mystery, which, in fairness, was really there all along and certainly not specific to Uriagereka's discussion. The mystery is, of course, the EPP ('coded' in the

standard view by the introduction of the mysterious D-feature, which does not seem to do anything but 'check' the features on its conceptual twin, *there*).[17]

Uriagereka (1999a) also mentions that his formulation of "kernel numeration"

> doesn't establish that a derivation with and without there are the same; that couldn't be, or there is a man here would outrank a man is here (with overt movement), contrary to fact. Rather [the formulation of kernel numerations] says that the formation of a derivation with an item like there, which by hypothesis has no formal features, is *as minimal* (hence, legitimate) as the formation of a derivation without there [italics ours—CDG].

There is an accompanying suggestion (though somewhat oblique) elsewhere in his text that categorial features are perhaps "not primitive" and might be "derivable from syntactic functions." So there is a promissory note here regarding the nature of *there* and of the EPP, but as already mentioned, Uriagereka inherited this note as does every approach to these issues. The virtue of his discussion, as we see it, is the explicitness with which the mystery is stated (so, although the mystery is passed on, this is done with clear recognition of the mystery as such).

Apart from the lingering mystery, however, we have two serious worries about this general approach. The first relates to the notion of "equally minimal" numerations. We find this notion difficult to understand given the troubles it is introduced to handle (which are the same as those we discussed in section 1 and above).[18] 'Equally minimal' makes the presence or absence of *there* an option (that is, no principles are insisting on it being in a numeration or not). While this seems to us to be correct in spirit, if *there*-insertion is indeed an option, and given that the formulation of "kernel numeration" was designed to deal with the apparent optionality exhibited by the Romero/Marantz cases, why state it in "economy of convergence" terms only to nullify this aspect of the formulation for exactly the crucial case? We can stress this point by asking the following question: if the introduction of 'economy of convergence' in the formation of kernel numerations is not important for this case (and it cannot be, given the way things are stated), then what is it for?

The second worry is especially sharp if the attempt to introduce the economy notion here is spurious—technically, this story trades on what must be a very global kind of economy of derivations. The notion of "derivational horizon," *however* it is construed,[19] involves going through potentially non-trivial stretches

[17] Conceptually, this reminds us a bit of so-called 'null' case which exists solely to license a formative whose uniqueness is for the most part determined by it being the only item that can bear this case. Actually, it occurs to us that lots of notions enter the syntax in this manner—that is, as 'twins'. Consider, for other ubiquitous examples, the often inter-motivated introduction of (specific) instances of movement and the features of categories that 'drive' them. This, one might think, is undoubtedly a methodological artifact, but one never knows—perhaps All is Two.

[18] Though it occurs in Uriagereka's (1999b) larger discussion about Spell-Out and linearization, which is independent of the success or failure of the notion of kernel numerations.

[19] See Uriagereka (1998) for the specifics and much other interesting discussion relating to these matters. But, our point holds, as stressed in the text, *however* "derivational horizon" is to be cashed out.

of computation or sequences of operations. If we have understood this formulation correctly, it necessarily makes reference to a potentially large space of derivational possibilities to determine whether the "grammatical properties are satisfied." This is problematic enough if we take the COMPLEXITY problem seriously (which as mentioned in the introduction, we are not so moved about), but it is doubly problematic if this problem has been spuriously introduced.

Note that the notion of "legitimate interface object" as it plays out in attempting to constrain the inner workings of derivations is once again raising its head. This is again an instance of the 'paradox' noted by Epstein et al. (1998) which arises when trying to understand "licit rule-application" in terms of eventual "outputs." We will return to this issue again in the discussion of Chomsky's (1998) proposals, to which we turn directly.

2.3 Agree, Phases and EPP-features: In a continuation of the project outlined in Chomsky (1995), this most recent paper reconsiders many of the issues and much of the technical details. We will not be exhaustive in our summary of this most recent installment in the minimalist series for reasons of space, but we will sketch the basics and the particular facets of the proposals that are relevant to this discussion.

The first thing to note about Chomsky's recent paper is that 'feature-movement' is abandoned. Although 'Merge' is thought of more or less as introduced in Chomsky (1995),[20] he introduces a second operation dubbed 'Agree'. Consider the following quote (p. 22):

> […] Agree [...] establishes a relation (agreement, Case-checking) between an LI α and a feature F in some restricted search space (its domain). [...] A third operation is Move, combining Merge and Agree. The operation Move establishes agreement between α and F and merges P(F) to αP, where P(F) is a phrase determined by F (perhaps but not necessarily its maximal projection) and αP is a projection headed by α. P(F) becomes SPEC-α.

This is relevant to the current discussion because reconceptualizing the character of the basic operations has an impact on how we understand the relative cost-of-operations distinction. However, despite this new view of the basic operations, Chomsky does say that Move "combines" Merge and Agree, and he goes on to make the point about the combination of operations and "cost" explicit (pp. 22-23):

> Plainly Move is more complex than its subcomponents Merge and Agree, *or even the combination of the two*, since it involves the extra step of determining P(F) (generalized "pied-piping"). Good design conditions would lead us to expect that simpler operations are preferred to more complex ones, so that Merge or Agree

[20] With the exception of coding adjunction as a kind of ordered merge (his "pair" vs. "set" Merge) and his rejection of strict Extension. We will not address this here (but see Drury, this volume).

> (or their combination) preempt Move, which is a "last resort," chosen when nothing else is possible. Preference of Agree over Move yields much of the empirical basis for Procrastinate and has other consequences, as do the other preferences [italics ours—CDG].

That is: Move = Merge + Agree + the "extra step of determining P(F)." Note that while the MOM distinction remains intact, there is the complication of having a third independent operation in the mix. We'll return to this below.

Chomsky also states a θ-theoretic restriction that is relevant to this discussion, reproduced here in (24) (his (6)):

(24) Pure merge in theta positions is required of and restricted to arguments

"Pure" merge, for Chomsky is merge "that is not part of move." One effect of (24) is that the only thing that can (purely) merge to a non-θ-assigner is an expletive. This assumption, coupled with cost-of-operations distinctions, handles the basic cases. But first we need to consider the Agree operation in some more detail to illustrate how this is supposed to work.

Agree "establishes a relation" between one thing and another in "some restricted search space." Chomsky introduces some new terminology and technology in his explication of this "relation." Call this the Probe/Goal System. The assumptions are as follows. An item α with some illegible feature(s)—the Probe—searches its domain (the things the Probe c-commands) for an item with matching features—the Goal. The Goal must be active in that it contains illegible features as well. The Goal must furthermore be local, where this notion is understood in terms of "closest in the command path of the Probe."[21] A Probe that finds a matching Goal (that is active and local) results in the application of Agree which deletes the illegible features on the Goal. Features must match one-to-one and Agree cannot selectively delete features on a Probe (deletion of features is "all or nothing").

With these assumptions in place, consider (1) again. The derivation of (1a) could be seen to go as follows:

(25) a. $[T_{[per,\ EPP]}$ be $[[a\ man]_{[\phi,\ case]}$ [in the room]]]
b. $[T_{[\cancel{per},\ EPP]}$ be $[[a\ man]_{[\phi,\ case]}$ [in the room]]]
c. $there_{[per]}$ $[T_{[\cancel{EPP}]}$ be $[[a\ man]_{[\phi,\ case]}$ [in the room]]]
d. $T_{[\phi,\ EPP]}$ seem [$there_{[per]}$ [T be $[[a\ man]_{[\phi,\ case]}$ [in the room]]]
e. $there_{[\cancel{per}]}$ $[T_{[\phi,\ \cancel{EPP}]}$ seem [$\cancel{there}_{[per]}$ [T be $[[a\ man]_{[\phi,\ case]}$ [in ...]]]
f. there $[T_{\cancel{[\phi]}}$ seem [$\cancel{there}_{[per]}$ [T be $[[a\ man]_{[\phi,\ \cancel{case}]}$ [in ...]]]

Chomsky assumes that both *there* and non-finite T (T_{def}) host a [person] feature and that T generally hosts an EPP feature. In (25a) T_{def} is the Probe and

[21] "Closest" c-command is later qualified to also involve the notion of "equidistance" familiar from Chomsky (1995). This is not important for now.

the associate *a man* is the Goal. The person feature of T_{def} matches that of the associate and deletes (25b), though the features of the associate remain unchanged (deletion is all or nothing). If there were no expletive in the array, *a man* would raise at this stage. Assuming there is an expletive available, and given that Merge is cheaper than Merge + Agree + the "extra step" of determining the PF properties of the displaced expression, the expletive is inserted in (25c), causing the EPP feature to delete. The derivation continues to the point in (25d) where matrix T has been introduced. The closest Goal for matrix T is the expletive (which is 'active' in virtue of its uninterpretable [person] feature). Thus *there* raises to the matrix SpecTP, the EPP feature is checked and the illegible [person] feature of *there* is deleted under matching with matrix T. Matrix T remains unchanged (again because deletion is all or nothing), shown in (25e).[22] Lastly, with the EPP checked and the illegible features of the expletive deleted, matrix T is now the Probe and the associate the Goal. The ϕ-features of T delete as does the Case feature of the associate, and the derivation converges. The ill-formed (1b) cannot be generated since at the point in (25a), if *there* is available it must be Merged.

There is a glitch here (partly mentioned in fn. 22) which brings the cost-of-operations distinction into sharper relief. If Merge and Agree do not differ in relative "cost", the steps (25b) and (25c) could be reversed. That is, the Merge of the expletive to satisfy the EPP could occur before the matching of the [person] features of T_{def} with the associate. But then we apparently have two options. Under Chomsky's assumption that the expletive is an X^0 which, when Merged, is a Probe that finds the illegible [person] feature of T_{def} as its Goal, this illegible feature of T_{def} should delete (this, in fact, is how successive movement of an expletive through sequences of defective T's is implemented, with *there* serving to both check the EPP and to delete the [person] feature on each T_{def}). But when there is no expletive, the [person] feature on T_{def} finds the illegible features of the associate as its Goal, forcing the deletion of [person] for T_{def}. Then the EPP drives the movement of the associate to Spec-T_{def}, leaving the associate's illegible Case intact (remember deletion is all or nothing, and T_{def} has just a [person] feature). So here is the glitch: in the case where the expletive is inserted, why can't the expletive delete [person] on T_{def} *and* then delete its own [person] feature by finding the associate as its Goal (just as it found T_{def})? This would render the expletive 'inactive' as a potential Goal for later Probes, and allow derivations of the following impossible examples:

[22] There is a trick here that Chomsky introduces that we will gloss over for now. When *there* raises to matrix T, it does not delete its [person] feature under Merge. Rather, Chomsky says that *there* is an X^0, which, given its illegible [person] feature, acts as a Probe to find the closet Goal. This happens to be matrix T (which it just Merged to) and because of the all or nothing requirement on deletion, matrix T's features remain unchanged while the [person] feature is deleted. A number of questions arise, such as about the X^0 status of *there* (e.g., does this mean that it projects, or does it not? Is this not trivially the case given the relational definitions of phrasal status? Why should *there* have a [person] feature? Why should T_{def}?) We will return to this momentarily.

(26) a. *There seems there to be a man in the room.
b. *A man seems there to be in the room.

(26a) would arise by having two expletives in the array. Nothing would stop this if the embedded *there* managed to render itself 'inactive'. Similarly with (26b): if *there* renders itself inactive then nothing prevents matrix T from finding *a man* as its Goal, deleting the relevant illegible features and forcing it to raise to check the EPP. If we tried to say that T_{def} cannot, for some reason, delete its [person] feature under Agree with the associate, then how does it work in raising constructions? For raising, Chomsky assumes that Agree applies, between T_{def} and the associate, and coupled with the unavailability of an expletive, forces the associate to raise to Spec-T_{def} to check the EPP. And in the situation where *there* moves more than once there is the peculiarity that T_{def} and *there* in a lower specifier do not enter into a Probe/Goal relation until after *there* has moved, so that *there* is the Probe and T_{def} is the Goal (so the latter's [person] feature deletes, but not the former's).

The reply to these worries could be that it is impossible because Merge is cheaper than Agree. After *there* has been inserted, even if Agree relates *there*-T_{def} instead of T_{def}-associate, it will be cheaper to Merge the next element from the array than to apply Agree again. But this does not seem so straightforward. What has to happen is that an illegible feature ([person] on *there*) has to be prohibited from falling under this "Suicidal Greed" (Chomsky's rendering of Lasnik's "Enlightened Self-Interest"—the notion that it is the properties of the Probes and not the Goals that 'drive' derivations). Note that it cannot in general be the case that Merging the next item from the array is preferred. The presence of illegible features on an item forces something to happen roughly as soon as possible, but the system does not, for instance, continue on with the derivation after (25c) without some instance of Agree applying just because Merge is cheaper. Further, if Merge vs. Agree was introduced as a cost-distinction then it would loose the 'less vs. more' rationalization that MOM has.

Further, consider the case of *there* as the specifier of matrix T under Merge (not Move). What happens in this case is that the expletive (Probe) locates the closest Goal (T) and Agree applies to eliminate the [person] features of *there* (though, again, not the ϕ-features of T, since it has a full set and deletion is all or nothing/one-to-one). The EPP is also checked. Then matrix T (now the Probe) finds the associate (Goal) and Agree applies to eliminate the illegible features. The introduction of [person] features along with the EPP seems to introduce an undesirable redundancy (which seems to be the source of the above noted trouble). Further, the EPP and the impoverished ϕ-set of *there* and T_{def} seem to be required because the system has the Agree operation. Since things can Agree without moving, something else must force overt subjects in finite clauses and successive cyclic movement. We will not pursue this any further here, since we are going to attempt to deny the existence of intermediate specifiers in the next

section and rethink the nature of *there*, but it is important to point out that the mechanics are not exactly clear.

Observe that in addition to the above concerns that the cost of operations hierarchy and the θ-theoretic condition are both required to handle the core case, the θ-theoretic condition only ensures that *a man* will not be Merged directly to non-finite T; it does not prohibit the movement to this position.

But, a further question arises, pointed out by Chomsky himself. Namely, with this conception of cost distinction in place, why does Agree not always prohibit the "extra step of determining P(F)?" That is, how does movement ever occur since the cheaper Agree operation should suffice? This is just the original worry about (1a,b) which started this discussion.

Chomsky's answer to this question is three-fold. First, the θ-theoretic restriction, as noted above, blocks arguments from Merging anywhere where they would not immediately be identified/marked as such (i.e. θ-marked). Second, imagine for just a moment that expletives do not exist in natural languages (despite the consequences for the existence of this very essay). In this counterfactual world it would then follow from the combination of the θ-theoretic condition and the existence of EPP-features that there must be movement. That is, arguments cannot Merge to check the EPP, and these positions appear to host nothing but this feature and perhaps an impoverished ϕ-set (i.e. just [person], so Agree will not help), so, therefore, there must be movement.

However, there are indeed expletives in natural languages. So there is still the issue of why the Merge of an expletive will not always block Move (i.e. the original problem). This leads us to the third part of Chomsky's answer: the introduction of "phases." Chomsky considers "a straightforward solution" which

> […] would be to take the derivational approach still more seriously and extend further the procedures...that reduce access to the domain of L. Suppose we select LA ["Lexical Array" = "Numeration"—CDG] as before […] the computation need no longer access the lexicon. Suppose further that at each stage of the derivation a subset LA_i is extracted, placed in active memory (the "work space"), and submitted to the procedure L. When LA_i is exhausted, the computation may proceed if possible. Or it may return to LA and extract LA_j, proceeding as before. The process continues until it terminates. Operative complexity is some natural sense is reduced, with each stage of the derivation accessing only part of LA. If the subarray in active memory does not contain EXPL, then Move can take place in the corresponding stage; if it does, Merge of EXPL preempts Move.

This may sound the same as views we have already discussed, but it is not. It differs from Uriagereka in alleging not to require reference to "convergence" and from CDG in its choice for which sub-domains are going to be considered "special." Like the view in CDG, Chomsky seems to have in mind an LF-interface motivated notion of 'subarray'. Consider:

> […] the subarrays LA_i that can be selected for active memory […] should determine a natural syntactic object SO, an object that is relatively independent in terms of interface properties. On the "meaning side," perhaps the simplest and most principled choice is to take SO to be the closest syntactic counterpart to a proposition: either a verb phrase in which all theta roles are assigned or a full clause including tense and force. Call these objects "propositional."

Syntactically, Chomsky takes the correspondents to the notion "proposition" to be C and *v*. The satisfied occurrences of these formatives signal a completed 'phase'. Chomsky briefly considers another possible way to understand subarrays—in terms of convergence. But he observes that this route involves the inclusion of 'look-ahead' properties and regards this as problematic.

A question arises, however, regarding how we are to differentiate between phases being "propositional" vs. their being "convergent." This matter is cleared up by the recognition of the fact that while Chomsky is attempting to motivate the existence of phases in terms of a grammar-external notion ("proposition"), he is explicitly introducing syntactic proxies (*v*/C) and not requiring that the notion "proposition" play a technical role within the proposed mechanisms (this is the move of CDG as well, except the 'proxy' was T and not *v*/C). Thus we do not have to worry about having to establish convergence for these objects to determine whether they are propositional which consequently would let us know that these objects are legitimate phases (pace Epstein & Seely, 1999:42-43). This is simply an instance of the minimalist sort of reasoning which considers plausible aspects of the grammar-external systems and examines the possibility that C_{HL} might be 'responsive' to such grammar-external properties in a direct and perhaps optimal way. On the other hand, we agree with Epstein & Seely that if we are to technically understand "phases" as crucially referring to (as opposed to merely "being motivated by") the (eventual) propositional status of *v*P/CP as grammar-external objects, then this requires the convergence notion that Chomsky aims to have rejected (i.e. phases are convergent) and thus the 'look-ahead' that he is similarly attempting to dispense with.

It should be obvious that Chomsky's introduction of phases and the other collection of conditions serves to handle the Romero/Marantz cases mentioned at various points above (in roughly the manner in which T-domains work for CDG). An expletive can either be part of a particular phase or not. If it is, then the MOM preference forces its insertion (to satisfy the EPP). If it is not, then the EPP forces raising of the associate. And since numerations have effectively been parceled into smaller objects, the situation where an expletive is potentially available for more than one clause never arises, yielding the optionality of the Romero/Marantz facts.

There are some more details of Chomsky's account that we might consider, (and we will have to do so in section 3) but we have enough on the table to make a point about his system that bears on our topic. That is, Chomsky's system suggests that MOM is dispensable. We turn to this next.

2.4 Salvaging the Relative Cost of Operations The additional mechanisms (phases etc.) discussed above, to the extent that they have independent plausibility, suggest to us that we should dispense with the cost-of-operations distinctions. The same goes for the "T-Domains" of CDG and Uriagereka's "Kernel Numerations." That is, if the lexicon is accessed via a numeration which is then in turn accessed in sub-parts then the MOM distinction falls out straightforwardly as a consequence of the more general condition on derivations (now localized to sub-arrays) that the set of choices must be exhausted. Sub-arrays containing an expletive will simply fail this condition if the expletive is not inserted. Assuming that we should in general prefer things not to be overdetermined by our putative laws, we are led to an immediate conclusion. In this case, the more general notion is clearly the EXHAUST ARRAY condition. One might think that under some of the views that we have discussed the Exhaust Array condition just *is* the MOM condition. But this is not so. Exhaust Array holds whether there is movement or not, it does not matter what operations are deployed, so long as all the items that have been selected are used. Importantly, this only will allow MOM to be dispensed with *if* the sub-array view is assumed. Otherwise, we are back where we started with respect to the more complicated cases. The structure of the account Chomsky offers allows us to dispense with the relative cost-of-operations distinction for the previously troublesome Romero/Marantz cases. If the expletive is present in a sub-array, it must be implicated in the corresponding sub-derivation or else the array will not be exhausted. If the expletive is not in the sub-array, then movement is forced to occur to license the relevant features.

However, taking this line once again leaves our core case (1) out in the cold. Without the cost distinctions, the account of these is lost again. But with the additional (cyclic) constraints on derivations/lexicon access that handle the problematic cases, this contrast seems (again) to stick out as the only empirical motivation for the cost-distinctions. We propose trying to handle this in a different way—by denying that both (1a) and (1b) are both well-formed (thus nullifying any potential comparison between them and thus removing the motivation for comparison). It will turn out that our route also does not require that anything special be said about cyclicity/lexicon access. Our approach to the simple cases extends straightforwardly to the complicated ones. That is, we think that there is something else lurking in the background which is causing trouble. That 'something else' is the EPP.

3 A Non-Economy Story: The Non-Existence of [Spec, Non-Finite T]

The GB and MP literature has a 'secret' that everyone acknowledges: we do not know what the EPP is, we just know we need it. This section entertains the possibility that we do not in fact need it. Our motivation flows from the discussion in the previous section. It is difficult not be attracted to the analytical strategy of simply denying that the case in (1) presents a real problem in the first

place, and exploring a non-economy alternative and its consequences. Further, to the extent that the various mechanisms introduced in CDG, Chomsky (1998), and Uriagereka (1999a; 1999b) have independent plausibility, we suggested above that these subtract from the interest of the MOM distinction for the following reason. Once the additional mechanisms are in place, the cases that seem difficult to initially account for (e.g. the Romero/Marantz examples) come out right even if there is no cost distinction between operations. This reinforces the interest of examining alternatives. Suppose, then, that non-finite T cannot host a specifier position (i.e. deny the EPP). What are the consequences? We need to minimally address: (i) the core cases involving expletives (and discuss the nature of expletive *there*), (ii) A-movement in raising constructions, (iii) Control, and (iv) exceptionally case-marked nominals. We consider each of these in turn and find that the consequences of eliminating the EPP are surprisingly far from disastrous. They in fact seem to fit rather well with a number of emerging ideas within minimalist research about each of the relevant phenomena. We turn first to the core cases.

3.1 Expletive-*There* The first obvious consequence of denying the existence of EPP-features is that the contrast in (1) follows directly. If non-finite T cannot host a specifier position, (1b) is impossible. Thus, *a man* can move or an expletive can be inserted as options when matrix T is introduced. This allows us, following Collins (1997), to dispense with numerations in favor of simply Copying from the lexicon. But, given the assumption that non-finite T cannot host a specifier, we are able to maintain this view without the cost distinctions and without the introduction of notions like "incomplete chain" which are necessary for Collins).[23]

Further, as noted above (section 2.4), if there is any independent need for numerations/arrays or sub-divisions of these in terms of relativized convergence domains (or sub-arrays, or 'phases') then the much more general condition

[23] Collins (1997:123-124) introduces the following 'chain formation principle' and the ancillary notion of 'incomplete chain':

> Chain Formation Principle 4: If there are two operations OP_1 and OP_2 applicable to a set of representations Σ (both satisfying Last Resort and Minimality), then choose the operation that extends an incomplete chain.
> Incomplete Chain: Let Ch be a (nontrivial) chain of the form (α,...,t), where α has D feature that has entered into a checking relation with a EPP feature and has an unchecked Case feature. Then Ch is an incomplete chain.

Despite Collins' denial of MOM and the abandonment of numerations, MOM is nonetheless hiding here. The idea that a choice of Move over Merge is required when an incomplete chain has been formed. For the core cases in (1), Move of the associate to the Spec-T_{def} vs. expletive insertion are options. But if the associate has moved, then, when the matrix T is introduced and the option again arises, the above definitions step in to force movement of the associate (blocking expletive insertion) since the associate together with its 'base' position constitutes an 'incomplete chain'. We do recognize, however, that we owe Collins an account of phenomena he takes as requiring the EPP independent of Case/Agreement (quotative and locative inversion). We have not dealt with these phenomena in this essay, but should our proposed elimination of the EPP prove to be empirically sustainable, we will have to.

'Exhaust Array' will suffice to introduce the expletive if it is present in the set of initial lexicon choices. Thus the effects of the MOM can be indirectly maintained should it prove necessary to appeal to this distinction for other reasons yet to be uncovered. But, crucially, this will be the case without any need to provide theoretical justification for the cost-distinctions since under this view they are not needed. Therefore, there are not any. Note that the removal of these distinctions causes the INCONSISTENCY problem that arose in discussion of (6) in section 1 to simply vanish (as the reader can verify).

But now we face a question that we raised in section 2.2 for Uriagereka's approach: what does the expletive do? If EPP-features are denied, what properties license expletives when they occur? Take a simple case:

(27) There arrived a man.

With the EPP eliminated, what is the expletive doing in constructions like (27)? There is a general intuition about *there* that has been lost, we think, in all the technical apparatus that has emerged through various re-analyses. The intuition is that *there* serves to "transmit" both Case to the associate and the ϕ-features of the associate to T/INFL. This was the intuition behind the 'coindexing' analyses offered in Chomsky (1986), hence KOL. In KOL Chomsky supposes that *there* and its associate are coindexed at D-structure, so that the initial representation of *there seems to be a man in the room* is as in (28):

(28) [*e* [Infl [seem [there$_i$ [to be [a man$_i$ in the room]]]]]]

This is buttressed in KOL with a technical discussion about CHAINS and Case/θ-relations so that when *there* raises to the subject position, the CHAIN relation it instantiates serves to "transmit" Case to the associate-nominal thus complying with the Visibility Condition (Aoun, 1985). We want to see how this older idea might be understood within current substantive and technical assumptions.[24]

Chomsky's (1998:60-61) discussion of Case, we believe, sheds some light on this matter. Consider:

> Manifestation of structural Case depends on the interpretable features of the probe: finite T (nominative), *v* (accusative), control T (null), on earlier assumptions. We may therefore regard structural Case as a single undifferentiated feature […] Its manifestation depends on the interpretable features (namely ϕ-features) of the goal, so that it too can be taken to be undifferentiated as to the value of the individual features of the ϕ-set. For both

[24] Our suggestions below turn out to be very close to what is proposed by Frampton & Guttman (1998). There are some differences, but they are narrow enough to require too detailed a comparison to fit in the present essay. The main difference is that they retain the EPP and successive cyclic A-movement/intermediate traces, which we reject here. But the ideas about Case/Agreement relations generally are the same as they propose (though they have a different take on expletives). We refer the reader to their paper on these points.

> probe and goal, the form of the uninterpretable features is determined by Agree. To rephrase in traditional terms, *verbs agree with nouns, not conversely, and Case is* <u>*assigned*</u> [italics ours—CDG].

And later (p. 66) he mentions that:

> With this shift in perspective, structural Case is demoted in significance. The Case Filter still functions indirectly […] But what matters primarily are the probes, including ϕ-features of T, *v*. That reverses much of the recent history of inquiry into these topics and also brings out more clearly the question of why Case exists at all. The question arises still more sharply if matching is just identity, so that case *can never be attracted*; *operations are not induced by Case-checking requirements* [italics ours—CDG].

The general picture, abstracting from Chomsky's particular implementation, is as he mentions, a traditional one. In the case of an ordinary nominal and finite T, the idea is that these items, as it were, 'swap' Case for ϕ-feature specification. Schematically things are as in (29):

(29)
```
  ϕ ———〉
DP     INFL/T
  〈—— NOM
```

The 'transmission' idea about expletive-associate relations could then be depicted as in (30):

(30)
```
  〈————————— ϕ
  ϕ ———〉
Expl   T/INFL     Associate
  〈—— NOM
NOM ——————————〉
```

Intuitively: T/INFL needs ϕ-features, in exchange for which it discharges nominative Case. The associate has ϕ-features, and needs Case. What the expletive is able to do is mediate this exchange. Suppose that, contra Chomsky (1998), T/INFL has an <u>empty</u> ϕ-set. A "reflex" of this empty ϕ-set is the ability to discharge a Case (nominative on T, accusative on *v*).[25] With respect to Case and nominals, one of two technical routes are available. We can either (i) stick to the idea that Case is an illegible feature, or (ii) assume that nominals have no Case and straightforwardly implement the traditional "assignment" intuition mentioned by Chomsky.[26] Either way, we can reformulate Chomsky's Agree operation as

[25] We will come back to the status of object case and *v* below.

[26] These two views are after all, in a sort of 'figure/ground' relation. That is, either nominals lack something that they must receive in order to be Visible (in Aoun's sense) or they have something they must get rid of in order not to be illegible. These are the 'checking' vs. 'assignment' alternatives. For the proposal here, it does

follows. Assume that T/*v* have empty ϕ-sets and that these items thus cannot on their own be a Probe that matches the (Goal) associate's ϕ-features. Empty ϕ-sets, then, are what trigger the overt movement where the expletive is not available (e.g., *A man arrived*) and empty ϕ-sets are, as mentioned above, associated with the 'reflex' of structural Case (dependent on the particular functional item).

We are in a way taking Chomsky's suggestion a step further. He claims that T/*v* are associated with ϕ-sets that are undifferentiated with respect to their values, but that they nonetheless contain a "full-complement" of ϕ-feature attributes (i.e. person, number, gender etc., all with unspecified/'zero' values). We are supposing that T/*v* have ϕ-sets that are more radically underspecified—ϕ-sets with no attribute- and thus no value-specifications. The difference, then, between T and T_{def} can be cast as the presence or absence of such a radically underspecified ϕ-set (T has one, T_{def} does not).

Now consider expletive *there*. Its crucial property, as we noted, is that it can somehow mediate the T-associate relation. This could be seen to follow if expletives have essentially the ϕ-set that Chomsky proposes for T/*v*. That is, a ϕ-set specified for attributes (person, number, gender) but not including values for these attributes (e.g., 3rd person, plural, masculine for *a man*). Now the Agree operation takes on a flavor of much work in unification-based approaches (which is quite similar to the kind of matching mechanism Chomsky has introduced anyway, and that was close to implicit, if not formalized, in 'checking' theory).[27] Again, the radically underspecified ϕ-sets of T/*v* must be unable to serve as a Probe in the T-associate relation, but must be suitable Goals (e.g. when a subject raises it must be able to target T as a Goal). Thus, we get overt movement of *a man* in (31a), followed by the Agree operation which "fills-in" the ϕ-set of T.

(31) a. A man arrived.
b. There arrived a man.

The expletive case in (31b) involves the insertion of *there* (again, under our view, an option of Copying from the lexicon) licensed by its ability to enter into Agree with the fully specified ϕ-set of the Associate. Now the expletive contains a ϕ-set with values for the ϕ-attributes. Suppose now that the expletive targets T as its Goal, and fills-in missing attributes and values of T. And, following Chomsky, we view structural Case assignment as a reflex of this process (see below).

The typology of items with respect to ϕ-sets is as in (32) (where a random particular instance of ϕ-attribute values is provided for the nominal example).[28]

---

not seem to matter which way we go on this issue with respect to *Case*, though it may have other consequences.

[27] We have not checked yet, but we suspect that expletive/associate relations have been discussed in something very much like the way that we are suggesting in the Head-Driven Phrase Structure Grammar literature (minus the discussion of Last Resort, below). The notation we have adopted here reflects this intuition. We would welcome comments and references on this.

[28] And we would assume that expletive *it* has an intrinsic legible ϕ-set (i.e. with values for whatever ϕ-attributes it has).

(32)	$\mathbf{T_{def}}$	**T/*v***	***there***	**nominals**
	∅	ϕ: ∅	ϕ:	ϕ:
			per: ∅	per: (3rd)
			num: ∅	num: (sing)
			gen: ∅	gen: (masc)
			etc.	etc.

But we have not yet looked carefully at the array of logical possibilities that our suggestions make available with respect to potential Probe/Goal relations between our hypothetically specified formatives. There are nine possibilities, listed in (33) (ϕ_{empty} = radically underspecified; ϕ_{attr} = just attributes; ϕ_{val} = includes specific values, e.g. "plural," etc.).

(33)				
	a.	ϕ_{empty}	ϕ_{val}	disallow (?)
	b.	ϕ_{empty}	ϕ_{attr}	disallow (Last Resort)
	c.	ϕ_{empty}	ϕ_{empty}	disallow (Last Resort)
	d.	ϕ_{attr}	ϕ_{val}	expletive-associate
	e.	ϕ_{attr}	ϕ_{attr}	disallow (Last Resort)
	f.	ϕ_{attr}	ϕ_{empty}	disallow (Last Resort)
	g.	ϕ_{val}	ϕ_{val}	conflict/blocking effect (Last Resort)
	h.	ϕ_{val}	ϕ_{attr}	reflexives (?)
	i.	ϕ_{val}	ϕ_{empty}	NP-Case Assigner

Of the logical possibilities, only four look to be potentially useful (and one of those is problematic, see below). Interestingly, the five others are plausibly ruled out by Last Resort. The reasoning would be that each step of a derivation has to be such that it does something to increase legibility. Assuming that anything but a fully specified ϕ-set constitutes an illegible object, instances of Agree that unified pairs like those in (33b,c,e,f) would all involve no increase (in fact, no change) in legibility. Additionally, the case in (33g) would be an instance of two items that are already fully specified and hence both legible, so Agree could not pair them.

Consider just the options that are allowed by Last Resort:

(33')				
	a.	ϕ_{empty}	ϕ_{val}	(???)
	d.	ϕ_{attr}	ϕ_{val}	expletive-associate
	h.	ϕ_{val}	ϕ_{attr}	reflexives (?)[29]
	i.	ϕ_{val}	ϕ_{empty}	NP-Case Assigner

We probably would want to rule out (33a). If permitted this would enable T and some ϕ-set it c-commands to enter into Agree (as in Chomsky's system). We stipulated above that an item with ϕ_{empty} cannot be a Probe. But it does not seem

[29] It occurs to us that (33h) might be a useful way to think about the agreement properties of local anaphora, but we are not going to pursue this here.

easy to make this follow from anything. What we want is to have empty ϕ-sets induce movement, and to then subsequently have the Agree operation to fill in the empty ϕ-set with the ϕ-attributes/values of the moved item. Further, we want the Agree operation to apply when ϕ_{attr} commands ϕ_{full} in the expletive-associate case, so we cannot introduce a general directionality constraint which insists on ϕ_{full} commanding the ϕ-sets that it renders legible under Agree. The most general thing that we can say is that Probes must have ϕ-attributes, so that both ϕ_{full} and ϕ_{attr} are permitted to be Probes. Thus, prior to Agree, expletives and full nominals have properties which form a subset relation with respect to illegibility. Expletives are doubly deficient in that they have zero Case value and zero ϕ-values, while regular nominals only have zero Case. We will assume then that ϕ_{empty} cannot be a Probe, though we recognize that more needs to be said.[30]

There is an additional worry in connection with (33a) if we consider the status of direct objects in English. Under our assumptions, the derivation of simple transitive like *The man saw the woman* would have a derivational stage like the following:

(34) [*v* [V [$_{DP}$ D N]]]

According to our story thus far, *v* would have an empty ϕ-set, and the DP would have a 'full' ϕ-set with values specified. This is exactly the situation in (33a). With respect to direct objects, there seem to be roughly four options (all of which have been defended in the literature): (A) objects move "covertly"; (B) objects move "overtly"; (C) objects do not move, but their features do (for us, following Chomsky, this = "Agree"); or (D) objects do not move period (their Case- and θ-positions are one and the same). Note that we are predisposed here to reject (C) since we have followed Chomsky in rejecting feature movement and the technical problems it raises. And, if we were to permit Agree to apply in this case we would be forced to accept (33a) as a general possibility and would have to introduce some other extra assumption to rule it out for the T-associate relation we have just been discussing above.

We are also inclined to reject option (A) since it involves a pre- vs. post-Spell-Out movement distinction and an attendant notion of "strong" vs. "weak" features. The notion of "strong" features seems easy enough to motivate on PF-grounds—those languages with overt agreement require overt movement. Under the view of the categories T and *v* that we have been pursuing, lack of movement to allow Agree to fill-in the ϕ-sets of such items would result in an item being sent to PF with no specification for its inflection. But "weak" features have no

[30] A full answer to these worries, we think, involves cross-linguistic work, which we have not the space to undertake in this essay. For example, the idea that ϕ_{empty} cannot be a Probe might run afoul of situations in other languages (e.g., Icelandic) where post-verbal nominatively Case-marked subjects show up. But the virtue of stating things as we have done here is that we can vary what the ϕ-set of T looks like across languages (which seems right). So for cases in other languages which seem to require something like (33a), the obvious analytical move that is available is to posit different ϕ-sets allowing T to be a suitable Probe.

motivation under this view. There is of course theory-internal motivation (i.e. to have everything work in the same way with only a Spell-Out difference involved) but it seems to us to miss the crucial aspect of the intuition about "strong" features which is tied to overt-/PF-manifestation of inflectional morphology. "Weak" features simply force a spurious kind of generality into the system. Rejecting this generality leads to the following view with respect to "morphologically motivated" movement: either there is overt movement (i.e. movement is morphologically motivated) or there is no movement (i.e. movement is *not* morphologically motivated, so it cannot happen). This view would take the lack of object agreement in a language like English to signal the impossibility of object movement. That is, in simple transitive constructions like *the man saw the woman* the direct object *the woman* is *in situ*. This means that this object must receive its Case and θ-role in one and the same syntactic position (i.e. head-complement relation). Put another way, direct objects are inherently Case-marked. This would, incidentally, explain the impossibility of having an expletive associating with the object of simple transitives:

(35) a. *The man saw there a woman
b. *The man there saw a woman

In fact, we might generally take the possibility of expletive-insertion to correspond only to situations in which Case and θ are not assigned in the same position. Thus, in raising, passives, unaccusatives, ECM-constructions, and so on we would expect to see expletives, but not elsewhere.

This leads straight to another consequence—the denial of the VP-internal subject hypothesis. Assuming subjects are generated within the VP-shell and move out to their surface position (in which they enter into the Case-Agreement relations), English should allow Transitive Expletive Constructions with the expletive associated with the subject:

(36) *There the man saw the woman.
(37) *There a man jumped.

These are, of course, possible in other languages.[31] But it would be perfectly consistent to deny the VP-internal subject hypothesis and suggest that languages can differ with respect to how many positions are licensed outside the VP (e.g., see Bobaljik & Thráinsson, 1998 on the Split-Infl Parameter). The rejection of the VP-internal subject hypothesis does not worry us very much, given how close it actually is to other versions of subject/external θ-role assignment that are currently on offer.[32] From the perspective of what we might very generally refer

[31] See among others, Jonas & Bobaljik (1993); Maling (1988); Platzack (1983); Vikner (1995); Zwart (1992).
[32] For example, see Williams (1994) on this matter and for a criticism of arguments for VP-internal subjects. Williams actually points out that VP-internal vs. –external subject analyses are actually only subtly different, despite superficial appearances to the contrary.

to as a VP-external subject theory, our earlier worries above about (33a) dissolve. This is so because subjects are always base-generated (in English) in their Case/Agreement positions (excluding the expletive/associate situation). So the situation of the empty ϕ-set c-commanding the full one can only be resolved by the expletive-insertion strategy, and we can maintain generally that empty ϕ-sets cannot be Probes. We will examine these issues in some more detail below in the discussions of Control, Raising, and ECM. A full statement of the Agree operation and the consequences for clause English clause structure will be given at the end of this section once the full range of cases are on the table.

To sum up, we have adopted and seriously modified Chomsky's Agree operation, and denied that T_{def} is able to host a specifier position in virtue of not having a ϕ-set (not even an underspecified one). Our reformulation reinvigorates the "transmission" hypothesis in a technical way that captures the underlying theoretical intuition rather directly.

It might be objected that this conception merely replaces the EPP with another mysterious object—empty ϕ-sets. But we consider this to be an advance. The notion of an empty ϕ-set makes conceptual sense under the 'traditional' view that T agrees with NP and not vice versa. What this view says is that T is capable of having ϕ-values, and furthermore must be so specified. The novelty is the attribution of the 'full-complement' of (valueless) ϕ-attributes to expletive-there. But this also seems to be the right way to think about this item. Note that in Chomsky's account, there has just a [person] feature and is able to check EPP, and the T-associate relation is independent of this fact. Agree applies to the T-associate pair regardless of whether or not the expletive is inserted or whether something moves to check the EPP. The need for the EPP and the postulation of the lone [person] feature follow directly.

For our view, in contrast, it follows directly that (i) the presence of *there* allows the 'long-distance' agreement, and (ii) *there*'s absence requires movement. Importantly, this follows without the EPP. For Chomsky, the EPP is a necessary because he has rejected the Case-transmission intuition and therefore parceled out the licensing of the expletive and the T-associate relation into independently functioning components in his analysis. The presence or absence of movement, under his view, is understood exactly to the extent that we understand what EPP-features are. We take this as extremely strong conceptual support for our alternative conception of these matters. But, our claim that T_{def} cannot host a specifier has other potential consequences which we examine next.

3.2 Consequences for other Domains Aside from the core case of interest in (1), denying specifiers for non-finite T has consequences for (at least) analyses of (i) raising, (ii) control, and (iii) ECM. We address each of these in turn.

3.2.1 Raising Our claim about the non-existence of specifiers for T_{def} has immediate consequences for raising constructions. Consider (38):

(38) a. John seems to be likely to appear to be happy.

b. [John$_i$ [seems [t$_i$ [to be likely [t$_i$ [to appear [t$_i$ [to be[t happy]]]]]]]]]
c. [John$_i$ [seems [to be likely [to appear [to be [t happy]]]]]]

(38b) is the standardly assumed structure, including 'intermediate' traces in each specifier of each instance of non-finite T. In the corresponding derivation each movement is taken to occur to check the EPP-feature hosted by each intermediate T. Removal of the EPP yields a derivation corresponding to (38c) under our account. There are a few things worth noting about the motivations for the standard view.

First, the reality of these intermediate movements is suggested by data such those in (39).

(39) a. *A man is likely there to be in the room.
b. *A man is likely it is in the room.

These kinds of violations are explained on standard assumptions if the movements in raising constructions are required to be strictly local. Thus the presence of *it/there* serves to block the obligatory intermediate landing sites for A-movement of *a man*, resulting in ill-formedness. This general restriction has received many different sorts of technical implementations, but the basic idea is constant throughout—the NP *a man* has exactly one route to matrix T and that is through the intervening specifiers. If any of these are occupied, the movement fails (or the Chain is ill-formed, or the trace is not properly bound etc.). For us, the ill-formedness of these examples follows straightaway from the total inertness of the intermediate T_{def} projections. Nothing can occupy these positions, period. Thus (39a,b) are not 'movement' violations for us. A derivational stage including Merge of anything to this position violates Last Resort.[33]

Second, there are cases that seem to be pretty strong counterexamples for our proposal. David Pesetsky (personal communication) credits Danny Fox with pointing out the following problem for a denial of intermediate A-movements. Consider:

(40) a. John$_i$ seems to Mary to appear to himself$_i$ to be happy.
b. *Mary seems to John$_i$ to appear to himself$_i$ to be happy.
c. *Mary$_k$ seems to John$_i$ [t$_k$ [to appear to himself$_i$ t$_k$ to be t$_k$ happy]]

[33] There is, however, an interesting consequence if we consider (39) together with the view on Case and Agree(ment) suggested above. Together, these *require* that the system be derivational. If the system were representational, there would be no reason why *there* could not be licensed by matrix T under our version of Agree. We have not constrained Agree to prohibit potential 'multiple' applications as it seems unnecessary to do so within the derivational conception we have assumed. So, in (39a) for example, Agree could apply between matrix-T and *there*, serving to fill-in the ϕ-attribute values and render the expletive legible. We do not place much weight on this consideration, as our proposal was developed with a derivational system in mind, but we find it interesting nonetheless that the implementation seems to insist on the derivational view.

If T_{def} cannot host a specifier as we claim, then typical cases of raising must involve "one-fell-swoop" movement over the 'inert' T_{def} projections and directly to the matrix T.[34] Under the standard assumptions, the binding of the reflexive is unproblematic in (40a) since *John* has raised from its base position *over* the reflexive to the specifier of *to appear* and then subsequently raised to its surface position. Thus we understand the reflexive to be locally bound in virtue of the trace/copy in the intermediate position. (40b), on the other hand, is ruled out as ill-formed in virtue of a kind of 'blocking effect' since *Mary*, by hypothesis, has raised through the specifier of *to appear* as in (40c). Thus typical binding requirements could be seen to rule out (40b) on the assumption that the intermediate movement really takes place (which we are now problematically denying).

While we do not have time and space in this essay to develop a full story about binding and reconstruction/connectivity (which is what we think is necessary anyway given minimalist rejection of government), there are some things to consider about these cases. First, it does seem right to us that these datives can indeed bind out of their PPs (if they could not we would have a straightforward answer to this problem). Variable binding by a quantifier (41a), negative polarity licensing (41b), Condition-C violations and their absence (41c,d), and reflexive binding (41d) all seem to point towards that conclusion (despite the fact that there is no c-command under usual definitions):

(41) a. It seems to every boy to appear to his mother that the earth is flat.
b. It seems to no man to appear to any woman that the earth is flat.
c. *It seems to him_i to appear to $John_i$ that the earth is flat.
d. It seems to his_i mother to appear to $John_i$ that the earth is flat.
e. It seems to $John_i$ to appear to him_i/$himself_i$ that the earth is flat.

However, the non-complementary distribution of the pronoun and reflexive shown in the last example is suggestive of the need to examine this situation more closely. There is a pretty compelling case to be made that *himself* in these constructions is actually a logophor. Consider:

(42) a. $John_i$ kissed Mary after/before/because/since it appeared to $himself_i$ that the earth was flat.
b. Mary kissed $John_i$ after/before/because/since it appeared to $himself_i$ that the earth was flat.

(43) a. [The possibility that he/John would lose] appeared to himself to be plausible
b. [The possibility that Mary would lose] appeared to John to seem to himself to be plausible

[34] Epstein & Seely (1999) reach this same conclusion as a consequence of their arguments against the notion of "Chain" (which leads them to also reject the EPP). We refer the reader to their paper for the discussion.

c. *[The possibility that John would lose] appeared to Mary to seem to himself to be plausible

In (42a,b), *himself* appears in the *to*-phrase in an adjunct, yet *John* seems to be perfectly accessible as an antecedent in either subject or object position. In (43a) the antecedent for *himself* is buried inside the subject NP but is nonetheless accessible. As in the case of the datives in the raising constructions, command seems to be completely unnecessary. It is plausible that the 'blocking' effect in the Pesetsky/Fox cases is a matter of a preference for a c-commanding antecedent where one is available, and does not require the postulation of an intermediate movement to the Spec-T_{def} as it seemed at first blush. On such a view, it must be the case that when the preferred (c-commanding) antecedent does not agree, this is cannot be overridden in any way to access some other potential antecedent. This would make the non-complementary distribution pronouns (noted above) intelligible since the evidence seems to point to this not being a case of ordinary anaphora. Rather, *himself* in these cases appears to be logophoric. Along these lines, note the contrast between (43b) and (43c). This suggests something of a hierarchy of accessibility preferences: c-commanding antecedents are demanded over m-commanding antecedents which are demanded over "other" antecedents. (43a) shows that even m-command is unnecessary for *himself* to be bound. Further, the (43b) versus (43c) contrast shows the same kind of "blocking effect" that we saw in the original apparent counterexamples, but in a way that suggest that this may have nothing at all to do with intermediate traces. In (43b), of the available antecedents, the 'better one', structurally speaking, is the one that happens to also have matching ϕ-features, so all is well. But in (43c), the 'better' antecedent, structurally speaking, happens to be one that does not match in ϕ-features, and coreference is blocked.

We have absolutely nothing to say about such an apparent structural preference hierarchy, though saying that such exists does seem to us to appear to ourselves to explain the facts (and it has a ring of plausibility to it as well). But for certain we take the above considerations to alleviate the concern raised by the cases in (40) which appear to be instances of logophoricity, not anaphoricity.[35]

This (almost) concludes our discussion of raising. On a first pass, we find no serious reasons to think that EPP-driven intermediate traces exist in these constructions. But before we turn to a brief note about Control, we are obligated to mention one other sort of evidence that is sometimes mentioned with respect to intermediate traces of A-movement—so-called Quantifier Float.[36] Consider:

(44) a. The boys seem **all** to appear to like ice cream.
b. The boys seem to **all** appear to like ice cream.
c. The boys seem to appear **all** to like ice cream
d. The boys seem to appear to **all** like ice cream

[35] See Reinhart & Reuland (1993), Sells (1987) among others.
[36] Thanks to Tim Stowell for bringing this up at a presentation of this paper.

There is a view on examples such as these, due to Sportiche (1988), which suggest that renaming this phenomena "Quantifier Stranding" would be technically more accurate. The idea is that *all* in these examples is initially associated with the DP (*all the boys*) in its lowest (i.e. θ-) position, and that *the boys* moves independently thereby 'stranding' *all* in its base position. Since the DP can obviously move together with the quantifier (*all the boys seem to like ice cream*), there is the possibility that any one of the moves the DP makes could strand *all*. So, under these assumptions, examining the positions in which *all* may surface serves as a probe into the nature of such movement relations. The relevance of the cases in (44) should be obvious: if Sportiche is correct then we cannot be since *all* appears in exactly the places we deny there is movement to.

Thankfully, Sportiche's Q-Stranding analysis is not the only game in town—there are (at least) two others trains of thought on these matters. One we may refer to (this time accurately) as the "Q-Float" theory which posits independent movement of *all* to the surface positions in which it appears, and the other we can call the "Base Generation" view, which understands *all* to simply adjoin where it appears without it being involved in any movement relation.[37] We have nothing to add to the current debate on the proper account of these cases, but we take the availability of these other analytical options to suggest that this is not a serious worry for our denial of the EPP. In fact, to the extent that what we have proposed here is correct, our results could help decide some of these related issues. Put another way, if there is no EPP, then the Sportiche analysis simply cannot be correct. We save this for a future investigation.

By way of moving the discussion to Control, consider the following:

(45) a. The men seem to appear to want to leave.
b. The men 'seemta' 'appearta' 'wanna' leave.

If there are no traces in the specifiers of the intermediate non-finite T's (since there are, we claim, no such specifiers) then the contraction possibilities follow directly without having to make any stipulations about traces vs. PRO vs. Case-marked traces of *wh*-movement (e.g. **who do you wanna vanish?*). We can simply assert that in (45) the contraction is possible because nothing intervenes (surely a preferable story). But this means that for the last contraction in the series in (45b) (i.e. *want to* → *wanna*) that there better not be a PRO in Spec-T. We turn to this directly.

3.2.2 Control Our proposal with respect to Control should by this point be obvious and we will not have much to say on this matter (though this will clear up some details we left dangling above in the discussion of Agree). There are a couple of proposals currently on the market that are compatible with denying the

[37] Proponents of the "Q-Float" view include Baltin (1982), Fiengo & Lasnik (1976), Kayne (1975), Maling (1976), and Postal (1974). Defenders of the "Base Generation" view include Bobaljik (1995), Dowty & Brodie (1984), Klein (1976), and Williams (1980; 1994).

existence of the EPP, Hornstein (1999; forthcoming) and Manzini & Roussou (forthcoming). Both approaches aim to eliminate the Control Module by assimilating Control and Raising, and both require a rethinking of the θ-criterion. Consider (46) and Hornstein's (47) and Manzini & Roussou's (48) derivations:

(46) John expects to want to leave.

(47) a. [John$_{\theta}$ leave]
b. [John$_{\theta}$ [to [**t**$_{\theta}$ leave]]]
c. [John$_{\theta\theta}$ [want [**t**$_{\theta}$ [to [**t**$_{\theta}$ leave]]]]]
d. [John$_{\theta\theta}$ [to [**t**$_{\theta\theta}$ [want [**t**$_{\theta}$ [to [**t**$_{\theta}$ leave]]]]]]]
e. [John$_{\theta\theta\theta}$ [expects [**t**$_{\theta\theta}$ [to [**t**$_{\theta\theta}$ [want [**t**$_{\theta}$ [to [**t**$_{\theta}$ leave]]]]]]]]]
f. [John$_{\theta\theta\theta}$ [**t**$_{\theta\theta\theta}$ [expects [**t**$_{\theta\theta}$ [to [**t**$_{\theta\theta}$ [want [**t**$_{\theta}$ [to [**t**$_{\theta}$ leave]]]]]]]]]]]

(48) a. [to $^{\theta}$leave]
b. [to [$^{\theta}$want [to $^{\theta}$leave]]]
c. [T [$^{\theta}$expect [to [$^{\theta}$want [to $^{\theta}$leave]]]]]
d. [John [T [$^{\theta}$expect [to [$^{\theta}$want [to $^{\theta}$leave]]]]]]
e. [John$^{\theta\theta\theta}$ [T [expect [to [want [to leave]]]]]]

Hornstein assumes that Control and Raising are just the same thing really; the difference between them is only whether or not the DP which raises through a position 'picks-up' a θ-role. For Hornstein, θ-roles are features which can be 'assigned to' DPs. The derivations he proposes include movement through the specifiers of non-finite T, but there is nothing necessary about this.[38] And, if we wanted to account for the *wanna*-contraction cases by claiming that there is simply no formative intervening between *want* and *to* in the cases where contraction is permitted, we would want to deny this aspect of Hornstein's derivations.

Manzini & Roussou view the overt DP (the 'controller') as being base-generated in its surface position and 'attracting' all the unassigned θ-roles that it can. So, as with Hornstein's view, the difference between Control and Raising is simply a matter of how many θ-roles the subject DP comes to be associated with. Their view also makes the denial of the EPP possible (they themselves point out this feature of their system with respect to embedded clauses).[39]

Both of these views suggest to us that Control phenomena does not present any difficulty for denying the existence of the EPP. However, given our remarks about the Agree operation discussed above (in 3.1), we actually have reason to favor the Manzini & Roussou account. Recall that we suggested abandoning the VP-internal subject hypothesis in order to head towards an understanding of

[38] In fact, Hornstein (forthcoming) adopts our present proposal and does away with the EPP (at least for these embedded cases).

[39] Manzini & Roussou also make the same point that we have here about the immediately available explanation for *wanna*-contraction.

expletive insertion that traded on the possibility/impossibility of structurally separating Case and θ-assignment—where these occur in different positions, we expect expletives, otherwise not. This comes down to the claim that Case and θ are generally not realized in structurally distinct configurations connected by movement (at least not in English). The view offered by M&R allows us to maintain this. We can say that in the general case items 'enter' the syntax through their Case position and 'attract' θ-roles.

3.2.3 ECM Exceptional Case Marking presents us with another sort of puzzle since *a man* and *there* in (49a,b) appear to be in exactly the position that we claim must be empty.

(49) a. John believes a man to be in the room.
b. John believes there to be a man in the room.

Although we might note that the EPP-based understanding of the structural position of these ECMed expressions runs into trouble with cases where the expression is not in the specifier of any T, but rather in a small clause:

(50) a. John believes [$_{SC}$ a man likely to be in the room]
b. John believes [$_{SC}$ there likely to be a man in the room].

Given our the denial of the EPP, it must be the case that ECMed expressions must in fact be in some functional projection of the matrix clause. This in turn suggest that there is overt verb movement in these cases, since *believe* appears in front of the embedded subject. That there might be cases of overt verb and object movement in English has been suggested by a number of researchers.[40] Consider then a possible derivation for (49a):

(51) a. [John [T [believe a man [~~believe~~ [to be in the room]]]]]
b [John [T [believe there [~~believe~~ [to be a man in the room]]]]]

The conclusion is forced on our view (as in Epstein & Seely's similar rejection of the EPP) since *a man*/*there* cannot occupy the specifier of T_{def}. Assuming the suggestions of previous sections are correct, the ECMed expressions can be viewed as being inserted/base-generated as the specifier of *believe* or some functional projection above it. In (51a) *a man* will thus receive structural accusative (*believe* does not have a θ-role to assign, so it will not 'inherently' mark this expression and Case thus works here like it does with nominative and T—i.e. Spec-Head). In (51b) *there* is inserted and Agree takes care of the Case/agreement swap between the accusative Case assigning object

[40] Notably, Boskovic (1997), Epstein & Seely (1999), Johnson (1991), Koizumi (1993), Lasnik (1995; 1997), Lasnik & Saito (1993).

(*believe* itself or some functional head) and the associate in the embedded clause. Again, these cases will differ from, for example, the objects in simple transitives, in that the latter will involve Case and θ-assignment in a single position.[41]

A complication arises which we will just mention for now. One involves so-called *wager*-class verbs (Boskovic, 1997; Pesetsky, 1991; Postal, 1974). These strangely allow ECMed *there*, but not a regular nominal:

(52) a. *John wagered a man to be in the room.
b. John wagered there to be a man in the room.

Time and space preclude a closer examination of these cases so we leave them aside for now.[42]

4 Architectural Concerns and Conclusion

This essay has introduced a shift in approach. Through a study of various accounts centering around a consistent implementation of the MOM cost-of-operations distinction, we concluded that the array of cyclicity-related mechanisms which were introduced to partition derivations into sub-parts within which MOM could properly function suggested dispensing with MOM. It is worth noting that it is unsurprising given current architectures that the range of MOM should be so narrowly confined to these expletive cases. Given standard views within minimalist investigations about Case/θ-interaction, there is not really any room for the kind of indeterminacy that MOM was introduced to handle. Normally constraints on movement operations (or chains) serve to ensure a deterministic relation between θ-positions and Case-positions, so the possibility of the appearance of expletive *there*, assuming a denial of the kind of Case Transmission story we've reintroduced, should be limited to all and only those syntactic positions in which minimalist investigations have posited EPP-features.

[41] Though perhaps some objects move for other reasons. Definite objects might have to get out of the VP for scope reasons along roughly the lines of Hornstein's (1994) implementation of Diesing's (1992) in a discussion of ACD constructions. Such an approach might suggest assimilating definiteness islands to 'left branch'/CED violations as Mahajan (1992) has proposed for Turkish.

[42] There is also some curious agreement properties that emerge in the following:

(i) a. There was/were believed to be many/several/two men in the room
b. There is/are believed to be many/several/two men in the room
c. There was/were likely to be many/several/two men in the room.
d. There is/are likely to be many/several/two men in the room
(ii) a. There *was/were many/several/two men believed to be in the room
b. There *is/are many/several/two men believed to be in the room
c. There *was/were many/several/two men likely to be in the room.
d. There *is/are many/several/two men likely to be in the room

Chomsky (1998) observes and 'abstracts away' from the unexpected order in cases like *there was a proof discovered* or *there were three men arriving*. The properties of Case/agreement in these examples requires further study.

The problem which MOM was designed to solve, we suggest, was an artifact of the importation of the EPP into the minimalist framework under 'feature-checking'. The EPP makes it the case that Case/agreement will be independent of the possibility of having an overt subject, and this yields the optionality between Merge and Move that MOM was designed to rule out. We have suggested through this essay that these 'conceptual twins' are dispensable. This requires a rethinking of what expletive *there* does. With the EPP eliminated, it must have a crucial role in the 'transmission'/exchange of Case and agreement properties. We've sketched here a beginning route into understanding how this might work with some alterations of recent technical assumptions of Chomsky (1998).

We have left to the side the notions of cyclicity and convergence which formed the main part of the discussion in section 2. Given the shift we have introduced here (abandoning MOM and the EPP), these specific mechanisms are not necessary (though they may be for other reasons).

Acknowledgments
Thanks to Mark Arnold, Cedric Boeckx, Noam Chomsky, Marcel den Dikken, Norbert Hornstein, Matt Kaiser, Jairo Nunes, David Pesetsky, and Juan Uriagereka (and other participants in Juan Uriagereka's Spring 1999 Seminar on Minimalism) for discussions, comments, counterexamples, and criticism regarding various drafts of this project. The core ideas of this paper were presented at the 1999 LSA Annual Meeting in Los Angeles; thanks to the audience there, especially Martha McGinnis, David Pesetsky, and Tim Stowell. Remaining errors of fact or interpretation or faulty reasoning can be blamed on a proper subset of the present authors—to be defined by the author(s) you confront.

References

Aoun, J. 1985. *The Grammar of Anaphora*. Cambridge, MA: MIT Press.
Baltin, M. 1982. A Landing Site Theory of Movement Rules. *Linguistic Inquiry* 13, 1-38.
Bobaljik, J.D. 1995. *Morphosyntax: The Syntax of Verbal Inflection*. Doctoral dissertation, MIT.
Bobaljik, J.D. & H. Thráinsson. 1998. Two Heads Aren't Always Better then One. *Syntax* 1, 37-71.
Boskcovic, Z. 1997. *The Syntax of Non-Finite Complementation: An Economy Approach*. Cambridge, MA: MIT Press.
Castillo, J.C., J. Drury & K. K. Grohmann. 1997. Cyclic Evaluation as a Way to Achieve Local Economy. Manuscript, University of Maryland.
Chomsky, N. 1955/1975. *The Logical Structure of Linguistic Theory*. New York: Plenum.
Chomsky, N. 1973. Conditions on Transformations. In S. Anderson & P. Kiparsky (eds.), *A Festschrift for Morris Halle*. New York: Holt, Reinhart, & Winston, 232-286.
Chomsky, N. 1986. *Knowledge of Language: Its Nature, Origin, and Use.* New York: Praeger.
Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 1998. Minimalist Inquiries: The Framework. *MIT Occasional Papers in Linguistics* 15. Cambridge, MA: MITWPL. To appear in R. Martin, D. Michaels & J. Uriagereka (eds.), *Step by Step*. Cambridge, MA: MIT Press.

Collins, C. 1997. *Local Economy*. Cambridge, MA: MIT Press.

Diesing, M. 1992. *Indefinites*. Cambridge, MA: MIT Press.

Dowty, D. & B. Brodie. 1984. A Semantic Analysis of "Floated" Quantifiers in a Transformationless Grammar. In M. Cobler, S. MacKaye & M.T. Wescoat (eds.) *Proceedings of WCCFL*. Stanford Linguistics Association, 75-90.

Epstein, S.D., E.M. Groat, R. Kawashima & H. Kitahara. 1998. *A Derivational Approach to Syntactic Relations*. Oxford: Oxford University Press.

Epstein, S.D. & D. Seely. 1999. On the Non-Existence of the EPP, A-Chains, and Successive Cyclic A-Movement. Manuscript, University of Michigan.

Fiengo, R. & H. Lasnik. 1976. Some Issues in the Theory of Transformations. *Linguistic Inquiry* 7, 182-191.

Frampton, J. & S. Guttman. 1998. Cyclic Computation: A Computationally Efficient Minimalist Syntax. Manuscript, Northeastern University.

Freidin, R. 1978. Cyclicity and the Theory of Grammar. *Linguistic Inquiry* 9, 519-549.

Freidin, R. 1986. Fundamental Issues in the Theory of Binding. In B. Lust (ed.) *Studies in the Acquisition of Anaphora*. Dordrecht: Reidel.

Grimshaw, J. 1990. *Argument Structure*. Cambridge, MA: MIT Press.

Hornstein, N. 1994. An Argument for Minimalism: The Case of Antecedent-Contained Deletion. *Linguistic Inquiry* 25, 455-480.

Hornstein, N. 1999. Movement and Control. *Linguistic Inquiry* 30, 69-96.

Hornstein, N. Forthcoming. *Move!* Oxford: Blackwell.

Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Johnson, D. & S. Lappin. 1997. A Critique of the Minimalist Program. *Linguistics & Philosophy* 20, 272-333.

Johnson, K. 1991. Object Positions. *Natural Language and Linguistic Theory* 9, 577-636.

Jonas, D. & J.D. Bobaljik. 1993. Specs for Subjects: The Role of TP in Icelandic. In J.D. Bobaljik & C. Phillips (eds.) *Papers on Case and Agreement I. MIT Working Papers in Linguistics* 18. Cambridge, MA: MITWPL, 59-98.

Kayne, R. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

Klein, S. 1976. A Base Analysis of the Floating Quantifier in French. In *Proceedings of NELS VII*, MIT.

Koizumi, M. 1993. Object Agreement and the Spilt-VP Hypothesis. In J.D. Bobaljik & C. Phillips (eds.) *Papers on Case and Agreement I. MIT Working Papers in Linguistics* 18. Cambridge, MA: MITWPL, 99-148.

Lasnik, H. 1995. Last Resort. In S. Haraguchi & M. Funaki (eds.) *Minimalism and Linguistic Theory*. Tokyo: Hituzi Syobo, 1-32.

Lasnik, H. 1997. Levels of Representation and the Elements of Anaphora. In H. Bennis, P. Pica, & J. Rooryck (eds.) *Atomism and Binding*. Dordrecht: Foris, 251-268.

Lasnik, H. & M. Saito. 1993. On the Subject of Infinitives. *Papers from the 27th Regional Meeting of the Chicago Linguistic Society*, 324-343.

Lebeaux, D. 1988. *Language Acquisition and the Form of the Grammar*. Doctoral dissertation, University of Massachusetts.

Mahajan, A. 1992. The Specificity Condition and the CED. *Linguistic Inquiry* 23, 510-516.
Maling, J. 1975. Notes on Quantifier-Postposing. *Lingusitic Inquiry* 7, 708-718.
Maling, J. 1988. Variations on a Theme: Existential Sentences in Swedish and Icelandic. *McGill Working Papers in Linguistics: Special Issue on Comparative Germanic Syntax*. Montreal: McGill University, 168-191.
Manzini, R. & A. Roussou. Forthcoming. A Minimalist Theory of A-Movement and Control. *Lingua*.
Martin, R. 1999. Features and Minimalism. Manuscript, Tsukuba University.
Pesetsky, D. 1991. Zero Syntax II: An Essay on Infinitives. Manuscript, MIT.
Platzack, C. 1983. Existential Sentences in English, Swedish, German and Icelandic. In F. Karlsson (ed.), *Papers from the Seventh Scandinavian Conference of Linguistics*. Helsinki: University of Helsinki,80-100.
Postal, P. 1974. *On Raising*. Cambridge, MA: MIT Press.
Reinhart, T. & E. Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24, 657-720.
Sells, P. 1987. Aspects of Logophoricity. *Linguistic Inquiry* 18, 445-479.
Sportiche, D. 1988. A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure. *Linguistic Inquiry* 19, 425-49
Stowell, T. 1981. *Origins of Phrase Structure*. Doctoral dissertation, MIT.
Uriagereka, J. 1998. *Rhyme and Reason: An Introduction to Minimalist Syntax.* Cambridge, MA: MIT Press.
Uriagereka, J. 1999a. Minimalist Issues. Seminar, University of Maryland, College Park. Class lectures, papers and other material available on-line at: http://www.umd.edu/Linguistics/People/Faculty/JUriagereka/home.html
Uriagereka, J. 1999b. Multiple Spell-Out. in S.D. Epstein & N. Hornstein (eds.) *Working Minimalism*. Cambridge, MA: MIT Press, 251-82.
Vikner, S. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford: Oxford University Press.
Williams, E. 1980. Predication. *Linguistic Inquiry* 11, 208-238.
Williams, E. 1994. *Thematic Structure in Syntax*. Cambridge, MA: MIT Press.
Yang, C. 1997. Minimal Computation: Derivation of Syntactic Structures. Manuscript, MIT.
Zwart, C.J.-W. 1992. Dutch Expletives and Small Clause Predicate Raising. In *Proceedings of NELS 22*, University of Mass., Amherst: GLSA, 477-491.

Juan Carlos Castillo
Dept. of Spanish and Portuguese
University of Iowa
111 Phillips Hall
Iowa City, IA 52242
berasteg@blue.weeg.uiowa.edu

John Drury, Kleanthes K. Grohmann
Dept. of Linguistics
University of Maryland
1401 Marie Mount Hall
College Park, MD 20742
jdrury/grohmann@wam.umd.edu