View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by idUS. Depósito de Investigación Universidad de Sevilla

IBERIA
An International Journal of Theoretical Linguistics

Merge Over Move and the Extended Projection Principle: MOM and the EPP Revisited

Juan Carlos Castillo, John E. Drury & Kleanthes K. Grohmann
juan.castillo@uni.edu, john.drury@mcgill.ca, kleanthi@ucy.ac.cy
University of Northern Iowa, McGill University & University of Cyprus

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.

Keywords: A-movement, Agree, comparison, deficiency, derivation, economy, EPP, Merge, Move, Spell-Out/Transfer, T-domain.

Resumen: En este trabajo se examina una serie de propuestas destinadas a evitar problemas relacionados con varios casos de la distinción Ensamble sobre Movimiento (EsM) tomando como base el coste de la operación. Se concluye que mientras que los mecanismos introducidos en estas propuestas muestran rasgos que pueden ser interesantes de manera independiente, estos mismos se apartan sustancialmente del interés de las explicaciones derivacionales del marco EsM, basado en la economía. Será entonces cuando se considerará la eliminación de una asunción que convierta los casos básicos que implican construcciones-*there* en un problema con el que empezar: que el T no-finito ha de contener una posición de especificador (para cotejar un rasgo PPE/D). Negar la existencia de dichos rasgos elimina el problema para el que la distinción EsM se introdujo, permitiendo que los casos esenciales de movimiento del asociado frente a la inserción directa del expletivo se conviertan en un caso de verdadera opcionalidad. Se examinan consecuencias para otros fenómenos y se concluye que la propuesta es compatible con gran parte de las investigaciones recientes que estudian estos fenómenos.

Palabras clave: Movimiento argumental, Acuerdo, comparación, deficiencia, derivación, economía, PPE, Ensamble, Movimiento, Materialización/Transferencia, dominio-T.

Resumen: É analisada uma classe de propostas que procuram evitar problemas que emergem em várias instanciações da distinção de '*Merge* sobre *Move*' (MOM). Conclui-se que, enquanto os mecanismos aí introduzidos exibem independentemente propriedades interessantes, eles subtraem substancialmente ao interesse das explicações da economia de derivação de MOM. A remoção de uma assunção será assim considerada, aquela que faz com que os casos centrais envolvendo construções com *there* constituam um problema: T não-finito deve abrigar uma posição *Spec* (satisfazendo uma propriedade do PPE/D). Negar a existência de tais propriedades remove o problema que a distinção MOM procurou resolver, permitindo aos casos centrais de movimento-associado *vs* inserção de expletivo emergir como um caso de verdadeira opcionalidade. São analisadas consequências para outros fenómenos e a proposta demonstra ser consistente com muita da investigação recente sobre os fenómenos.

Palavras chave: Movimento-A, Concordância, comparação, deficiência, derivação, economia, PPE, *Merge*, *Move*, *Spell-Out/Transfer*, domínio-T. [1]

---

[1] The last-named author feels very honored to have been considered as an invited contributor to the inaugural issue of *Iberia* by the editors. Collectively, we would like to express our deepest gratitude to the *Iberia* editors, and Ángel Jiménez Fernández in particular, for accepting our original musings on the EPP for (re)publication. We thus celebrate a double inauguration: The launch of a new, exciting journal on theoretical linguistics (for which we wish the best of luck and a glorious future) and the marking of the 10th anniversary of our first attempt to critically inspect the EPP (which now is available widely and accessibly at last). The current article is our original essay from 1999, appended by a preface (section 0) and a postscript (section 6). Everything couched in between (apart from some minor changes as laid out in the preface below) appeared originally as: Castillo, Juan Carlos, John E. Drury & Kleanthes K. Grohmann. 1999. Merge over Move and the Extended Projection Principle. *University of Maryland Working Papers in Linguistics* 8, 63–103. In addition to the thankees originally acknowledged, we wish to thank a number of colleagues who've been interested in this work, discussed aspects with us, and kept pushing us for publication, over the years, in particular Željko Bošković, Sam Epstein, Gereon Müller, Anna Roussou, and Joachim Sabel.

Original acknowledgements: 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 [Uriagereka (1999a)]) 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

0. Preface: Setting the Stage

The present essay was initially published as a working paper ten years ago (see fn. 1 for full reference) that included a discussion of some proposals we had put forward in two earlier manuscripts which contained material that didn't make it into the final version (Castillo, Drury & Grohmann 1997, 1999). Across these efforts (small parts of which surfaced in Grohmann, Drury & Castillo 2000) we had sketched some of our joint thinking about interactions between certain conceptions of the nature of syntactic derivations, economy, local/cyclic convergence evaluation, and the inventory features/properties that motivate or 'drive' displacement operations (in particular, so-called 'EPP-features'). The earlier two essays were, to our knowledge, among the first discussions circulated in the early days of the Minimalist Program (Chomsky 1993, 1994, 1995) that considered seriously the notion of cyclic convergence evaluation in a derivational approach to syntax (see also Wilder & Gärtner 1997, an important contribution which we failed to cite in the original essay), a notion which has since become fairly orthodox in recent years under the label of 'Phase Theory' (Chomsky 1998, published in 2000, and subsequent work, in particular Chomsky 2001, 2004, 2005, 2007, 2008). Thus, as a word of caution, what follows is, even a decade after the rise of Phase Theory, soundly couched within the frame of discussion preceding (i) phases as local evaluation metrics, (ii) generalizing the EPP as the only property of the grammar bringing about displacement, and (iii) collapsing Merge and Move into the distinction between External Merge and Internal Merge, respectively. This said, we do address (i) in offering an alternative metric for local computations, more in line with what one may call 'kernel sentences' (Uriagereka 1999a);[2] concerning (ii), we reserve

---

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.

[2] At least with respect to (i), please also bear in mind that the original essay stems from 1999, after which we all individually concentrated on different formulations of evaluation metrics for our respective doctoral research, the result of which, very different from one another and focusing on different empirical domains of language as well, can be read up in Grohmann (2000), Castillo (2001), and Drury (2005), respectively.

the term for the 'subject property' that requires filling of the specifier of T (cf. Lasnik 2001) and regarding (iii) we agree in so far as the apparent economic advantage of one over the other is flat out denied (see also Epstein & Seely 2006 and a host of work over the past decade).

In what follows, we have left the original text more or less intact (starting with section 1), making only minor corrections, updating references, and adding a few new footnotes with pointers to more recent work.[3] We have, however, added a new concluding section in the form of a postscript which attempts to situate this (now decade-old) discussion relative to more recent developments (section 6).

1. Introduction

Notions of economy and optimality have been much discussed in the Minimalist Program (MP; Chomsky 1993, 1995, 1998). The present discussion begins with a consideration of a particular kind of derivational economy condition: the 'Merge over Move' preference (henceforth, MOM). The idea is familiar from Chomsky (1995, 1998): 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. Such a condition is transderivational as it requires reference to a set of competing routes through a combinatoric possibility space. While transderivational conditions are not new in linguistic theory, this cost-of-

[3] Our original paper was based on Noam Chomsky's manuscript circulated before it even appeared as Chomsky (1998). All quotes here are sourced to the 1998 MITWPL paper and updated in brackets to the corresponding page in the 2000 book chapter, from which the particular formulations and expositions are taken and which hence supersedes any (minor) discrepancies with the 1998 version.

Note also that, as a consequence, our discussion of Agree and related technical matters may read a bit dated and perhaps even obsolete at times, in part due to the fact that our knowledge of these concepts has constantly been enriched since, in particular through Chomsky's (2001 *et seq.*) own work, Pesetsky & Torrego's (2001) important contribution, and many others.

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. The connection between EPP-features and MOM will be unpacked as we proceed.

This essay is organized as follows. The remainder of section 1 outlines the initial formulation of MOM and points out the nature of the problems.[4] 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

[4] This essay focuses only on the INCONSISTENCY problem for three reasons. First, the complexity concerns have been thoroughly discussed elsewhere; notably Collins (1997), Johnson & Lappin (1997), Yang (1997), Chomsky (1998), and Frampton & Gutmann (1999). 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).

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. 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 EPP properties are not to be trivially dispensed with.

Section 3 thus takes on the burden of sketching some reasons for thinking that the cost to analysis is not disastrous. In fact, we find an eliminativist approach to EPP properties to be in concert with independently proposed conceptions of control and raising.[5] We close with a discussion of the consequences our conclusions and suggestions have for the architecture of the grammar. The newly added section 6 puts the original contribution in a more contemporary setting.

[5] See, for example, Hornstein (1999, 2001) and Manzini & Roussou (2000). 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); see Epstein & Seely (2006) for a full-fledged proposal very much in line with ours. 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 (1999) also cite an unpublished manuscript by Martin (1999) which aims to eliminate the EPP.

2. 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 *et seq.*, among others; but see Nunes 2004 for recent discussion and its historical context within MP and e.g., Hornstein 2001 for additional criticism).

Deploying this cost distinction minimally requires 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 (i.e. PF/LF) at which the output of a given derivation is evaluated for convergence. 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 (which should always result in lack of convergence).

This demands a comparison of candidate convergent derivations 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 (see also Sternefeld 1997). 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*). Such an augmented array is called the 'numeration' (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 e.g. Johnson & Lappin 1997, 1998). 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 contrast 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 the stage depicted in (2), the derivations of (1a) and (1b) are identical, and thus the derivations are comparable:[6]

(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) below, 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) further down).

(3) (a) [there [to be [[a man] [in the room]]]]
(b) [[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

[6] Here and below we will abbreviate numerations in this way (leaving out items whose indices have been reduced to zero).

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.[7]

(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}] [$_{PP}$ …]]]]]

However, we now encounter an analytical quandary that highlights the nature of the INCONSISTENCY problem alluded to earlier.[8] 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.

[7] 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.

[8] 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.

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

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

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 evaluation — if it is to work properly — seems to require sensitivity to distinctions stated over the outputs of derivations (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.[9]

[9] But perhaps we have the wrong notion of convergence in mind. Chomsky

Matters are complicated further by the following type of example:[10]

(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).

(1998: 8 [2000: 95]) has the following to say on this matter: '[A]n 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*' (emphasis added — CDG). Keep this comment in mind. We will return to this at various points below.

[10] (9) is attributed to Juan Romero and Alec Marantz in Uriagereka (1999b) and also discussed by Frampton & Gutmann (1999). In our original essay, we failed to mention the first published discussion of these cases, which can now be retrospectively credited to Wilder & Gärtner (1997).

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; see also fn. 1).

3. 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 locally evaluate. This second strategy is the one that we pursued in CDG. We review this next.

3.1. 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) 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').[11]

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 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).

[11] 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. (The notion of T-Domain was extended to other parts of the clause, albeit on a different theoretical basis, in Grohmann (2000, 2003) to partition the clause, and hence local computations, into three major domains.)

(10) 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. 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) 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 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]]]][12]

(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]

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

[12] 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 (17a) is evaluated.

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 either not be comparable or 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.

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 sub-trees 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,[13]

(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

[13] Chomsky (1998: 25 [2000: 110]) 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.

before anything else, there is thus a kind of optionality that is built-in to this system.[14] The adjunct status of the *that*-clause in (14a–b) is suggested by the impossibility of extraction shown in (19):[15]

(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 sub-tree thus far established. So whichever of the two is targeted, movement will have to occur in the other, thus yielding both derivations.[16]

To sum, in CDG we propose retaining the MOM distinction and the economy explanation for (1) by claiming that convergence is evaluated for TPs

[14] 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 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.

[15] 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) in particular that these are complements based on alleged reconstruction contrasts.

[16] 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.

in a cyclic fashion.[17] 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; see fn. 4); (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, including Chomsky (1998) and Uriagereka (1999a, 1999b). Chomsky's approach, as we will indicate, is conceptually 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). Uriagereka's suggestion, in contrast, takes a subtly different take on the 'locus' of the conditions governing what can constitute a legitimate sub-derivation. We turn first to Uriagereka, and then to Chomsky.

3.2. 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

[17] This logic may be subsumed by a more general or differently articulated understanding of cyclicity (as in Chomsky 1998, Epstein *et al.* 1998, Frampton & Gutmann 1999,). We will return to this possibility.

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*).

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* (emphasis added — 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.

Apart from the lingering mystery, however, we have two 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). '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 (cf. Wilder & Gärtner 1997), 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 that this story trades on what must be a very global kind of economy of derivations. The notion of 'derivational horizon', *however* it is construed,[18] involves going through potentially non-trivial stretches 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'.

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

[18] 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.

eventual 'outputs'. We will return to this issue again in the discussion of Chomsky's (1998) proposals, to which we turn directly.

3.3. Agree, Phases and EPP-features

In a continuation of the project outlined in Chomsky (1995), Chomsky (1998) reconsiders many of the issues and much of the technical details. We will not be exhaustive in our summary 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 is that 'feature-movement' is abandoned. Although 'Merge' is thought of more or less as introduced in Chomsky (1995), he introduces two additional operations, including the one dubbed 'Agree'. Consider:

> A second is an operation *Agree*, which 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 the specifier (Spec) of α ([Spec, α]).
>
> (Chomsky 1998: 14 [2000: 101])

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:

> 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 (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.
>
> (Chomsky 1998: 14 [2000: 101–102]; emphasis added — 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 element 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) [*u*F] (the Probe) probes its search space (the Probe's c-command domain) for an item β with matching feature [*i*F] (the Goal). The Goal must be active in that it contains illegible features as well.[19] The Goal must furthermore be local, where this notion is understood in terms of 'closest in the command path of the Probe'.[20] 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.

[19] To update our original presentation, an illegible feature F on the Goal is an uninterpretable feature F, often represented as [*u*F] (cf. Pesetsky & Torrego 2001), whereas a matching active feature F on the Goal is interpretable, or [*i*F]. No [*u*F] may reach LF but must be eliminated prior, whereas any [*i*F] is interpretable at LF and may remain active throughout, meaning that it could also engage in more than one Agree relation (Hiraiwa 2005). See also Hornstein, Nunes & Grohmann (2005) for exposition.

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

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 (simplified representation; cf. fn. 19):

(25) (a) [T[per, EPP] be [[a man]$_{[\phi, \text{case}]}$ [in the room]]]

(b) [T[~~per~~, EPP] be [[a man]$_{[\phi, \text{case}]}$ [in the room]]]

(c) there[per] [T[~~EPP~~] be [[a man]$_{[\phi, \text{case}]}$ [in the room]]]

(d) T$_{[\phi, \text{EPP}]}$ seem [there[per] [T be [[a man]$_{[\phi, \text{case}]}$ [in the room]]]

(e) there[~~per~~] [T$_{[\phi, \text{~~EPP~~}]}$ seem [~~there~~[per] [T be [[a man]$_{[\phi, \text{case}]}$ [in …]]]

(f) there [T$_{[\phi]}$ seem [~~there~~[per] [T be [[a man]$_{[\phi, \text{~~case~~}]}$ [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 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 [Spec, T], 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).[21] Lastly, with the EPP checked and the illegible

[21] 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

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. 21) 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:

(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

there have a [person] feature? Why should T_{def}?) We will return to this momentarily.

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 'Enlightened Self-Interest' from Lasnik 1995a, the notion that it is the properties of the Probes and not the Goals that 'drive' derivations; see also Bošković 1997). 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 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 *primitive* cost-distinction, then it would loose the 'less vs. more' economy 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 our present 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 a world without expletives, 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 (*there*-)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', a locally specified version of the original Numeration NUM — 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 "workspace"), 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.
>
> (Chomsky 1998: 19–20 [2000: 106], footnote omitted)

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 θ-roles are assigned or a full clause including tense and force. Call these objects *propositional*.
>
> (Chomsky 1998: 20 [2000: 106])

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, 2006: chap. 2). 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 CHL 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 (also Wilder & Gärtner 1997) 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, 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.

3.4. 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 derivations are evaluated in a chunk-by-chunk fashion, then the MOM distinction falls out

straightforwardly as a consequence of the more general condition on derivations (now localized to subarrays) that the set of (local) choices must be exhausted (though see below regarding the initially problematic case, (1a–b)). Subarrays containing an expletive will simply fail this condition if the expletive is not inserted. 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* we have a way to systematically factor derivations into local sub-derivations. Otherwise, we are back where we started with respect to the more complicated cases. The structure of the account Chomsky or CDG offer 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 subarray, it must be implicated in the corresponding sub-derivation or else the array will not be exhausted. If the expletive is not in the subarray, 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). That is, we think that there is something else which is causing trouble: the EPP.

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

This section entertains an eliminativist agenda for the 'EPP'. Our motivation flows from the attractiveness of an 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, also discussed by Wilder & Gärtner 1997) 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 as originally conceived). 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 independently proposed ideas within minimalist research about each of the relevant phenomena.

4.1. 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).[22]

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

(i) *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.

(ii) *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.

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 subarrays, or 'phases'), then the much more general condition '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. Note that the removal of these distinctions causes the INCONSISTENCY problem that arose in discussion of (6) in section 1 to simply vanish.

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]]]]]]

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.

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.[23]

Chomsky's discussion of Case, we believe, sheds some light on this matter. Consider:

> Manifestation of structural Case depends on 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 ([±plural], etc.). For both 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 assigned*.
>
> (Chomsky (1998: 39 [2000: 123–124]); emphasis added — CDG)

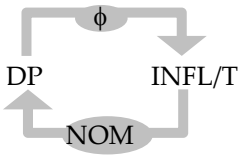
And a little later he mentions that:

> With this shift of 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*.
>
> (Chomsky (1998: 43 [2000: 127]); emphasis added — CDG)

[23] Our suggestions below turn out to be very close to what is proposed by Frampton & Gutmann (1999). 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.

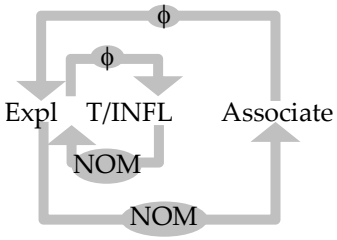
The general picture, abstracting from Chomsky's particular implementation, is as he mentions, a traditional one. In the case of an ordinary nominal and finite T, the idea is that these items 'swap' Case for ϕ-feature specification. Schematically things are as in (29):

(29)



The 'transmission' idea about expletive-associate relations could then be depicted as in (30):[24]

(30)



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 *empty* 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

[24] Note that it cannot be, strictly speaking, 'nominative' case that is realized on the associate. This can be made clear by switching in a 3rd person pronoun in the existential construction (which can be made to obviate the definiteness restriction with the addition of focus, as with the use of *only* in (i)):

(i) There only was him/he* in the room.

The 'transmission' logic we are reviving in (30) has to be conceived as dissociating structural from morphological case, well known from Distributed Morphology (Marantz 1991, Halle & Marantz 1993, Embick 1997), but also elsewhere (see e.g. Bobaljik 2008, Bobaljik & Wurmbrand 2008, and Legate 2008 for recent discussion).

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

mentioned by Chomsky.[26] Either way, we can reformulate Chomsky's Agree operation as 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). 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.

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

(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).[27]

(32)	**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

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

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 (?)[28]
	(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 easy to make this follow from anything. What we want is to have empty ϕ-sets induce movement, and then subsequently have the Agree operation 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

[28] 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. See Drury (2005) for discussion.

Case. We will assume then that ϕ_{empty} cannot be a Probe, though we recognize that more needs to be said.[29]

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

[29] 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.

have no 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.[30] 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.[31] From the perspective of what we might very generally refer 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 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.

[30] See Platzack (1983), Maling (1988), Zwart (1992), Jonas & Bobaljik (1993), and Vikner (1995), among others.

[31] 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.

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 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.

4.2. 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.

4.2.1. 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.[32]

Second, there are cases that seem to be counterexamples for our proposal. David Pesetsky (p.c.) 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.

[32] 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.

(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]].

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.[33] 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.

[33] 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 Epstein & Seely (2006) for extended discussion.

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 since it appeared to himself$_i$ that this would be the last chance.

(b) Mary kissed John$_i$ since it appeared to himself$_i$ that this would be the last chance.

(43) (a) [That he/John would lose] appeared to himself to be plausible.

(b) [That Mary would lose] appeared to John to seem to himself to be plausible.

(c) *[That John would lose] appeared to Mary to seem to himself to be plausible.

In (42a–b), *himself* appears in an adjunct, yet *John* seems to be accessible as an antecedent in either subject (42a) or object (42b) position (perhaps slightly less so in the latter, but the case is not entirely deviant). 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. 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.[34]

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 should mention one other sort of evidence that is sometimes mentioned with respect to intermediate traces of A-movement — so-called Quantifier Float.[35] 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.

There is a view on examples such as these, due to Sportiche (1988), which suggests 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.

[34] See Sells (1987) and Reinhart & Reuland (1993), among others. See also Drury (2005) for further discussion of these and related cases.

[35] Thanks to Tim Stowell (p.c.) for bringing this up at a presentation of this paper.

However, Sportiche's Q-Stranding analysis is not the only game in town — there are (at least) two other 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.[36] 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-style Q-stranding analysis simply cannot be correct.

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.

[36] Proponents of the 'Q-Float' view include Postal (1974), Kayne (1975), Fiengo & Lasnik (1976), Maling (1976), and Baltin (1982). Defenders of the 'Base Generation' view include Klein (1976), Williams (1980, 1994), Dowty & Brodie (1984), and Bobaljik (1995, 1998). In the present context, a very enlightening discussion can be found in Bošković (2002), based on Takahashi's (1994) approach to successive cyclicity, which can account for the occurrence of 'stranded' quantifiers without assuming a Sportiche-style derivation.

4.2.2. Control

Our proposal with respect to control should by this point be obvious, and we will consequently 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 existence of the EPP, Hornstein (1999, 2001) and Manzini & Roussou (2000).[37] 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' an additional 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]

[37] Over the past decade, a lot of work has reappraised the standard, and more recent, analyses of control. Both Manzini & Roussou (2000) and Hornstein (2001) were only available to us then in their pre-published forms, but a comprehensive overview of control constructions, various theoretical approaches to them, and arguments in favor of the movement analysis can be found in Boeckx *et al.* (forthcoming).

[38] In fact, Hornstein (2001) adopts our present proposal and does away with the

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 do 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 expletive insertion that traded on the (im)possibility 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 Manzini & Roussou 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.

4.2.3. ECM

Exceptional Case Marking (ECM) 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.

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.

However, 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 denial of the EPP (in non-finite [Spec, T]), it must be the case that ECMed expressions have to be, in fact, in some functional projection of the matrix clause. This in turn suggests 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 (1999, 2006) 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 (*believe* itself or some functional head) and the associate in the embedded clause. Again, these cases will differ from, for example, objects in simple transitives, in that the latter will involve Case and θ-assignment in a single position.[41]

[40] Notably, Johnson (1991), Koizumi (1993), Lasnik & Saito (1992, 1993), Lasnik (1995a, 1995b, 1997), Bošković (1997), and Epstein & Seely (1999).

[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 antecedent-contained deletion constructions. Such an approach might suggest assimilating definiteness islands to 'left branch'/CED violations, as Mahajan (1992) has proposed for Turkish.

A complication arises in connection with so-called *wager*-class verbs (Postal 1974, Pesetsky 1991, and Bošković 1997). These 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]

5. Architectural Concerns as Concluding Remarks

Through a study of various accounts centering around formulating a consistent implementation of the MOM cost-of-operations distinction, we concluded that the array of cyclicity-related mechanisms introduced to partition derivations into sub-parts within which MOM could properly function lead instead towards 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.

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

[42] It is worth noting that (52a) may in fact be deviant for processing reasons. Note that *wager* takes four arguments (*e.g.*, [*Bill*] wagered [*John*] [*ten dollars*] [*that 'such-and-such'*]). Realizing the argument picking out 'what the wager was about' as an infinitival clause as in (52a) could very likely result in an unrecoverable garden-path in which [*a man*] is parsed an argument picking out a *wager*-participant (e.g., [*John*] in the example above). If this is correct, then this particular set of '*wager*-class' cases do not require any further special treatment.

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).

6. Postscript: Retrospective Final Notes

The central themes explored above have continued to undergo extensive scrutiny over the 10+ years since we first circulated (versions of) the present essay. These themes are technical manifestations of old issues, and the depth of each is reflected in the degree of research attention which they continue to attract. One of our main motivations for getting this piece into print at this point in time is to allow the opportunity to highlight the extent to which our main concerns have been addressed. Of course, we cannot here attempt an adequate summary of the current state-of-the-art. To be sure, the landscape is now occupied by a range of intricate and rigorously defended theoretical positions, and in this respect it seems reasonable to say that the issues which concerned us here have been considerably advanced. On the other hand, the *diversity* of current views on the key issues which we only briefly canvassed in the present piece also makes it quite clear that these matters remain open and unresolved. Thus, we would like to close with a few brief remarks on how we view the present-day theoretical terrain regarding: (i) the status of the EPP, (ii) localizing convergence evaluation (and related ideas, e.g., phases, multiple Spell-Out, etc.), and (iii) derivational 'cost of operations' economy principles (e.g., MOM).

6.1. EPP Today

If we hold constant the notion of *Last Resort* in minimalist approaches to the organization of narrow syntax, then in our view, the inventory and distribution of features/properties that can serve as the driving forces of displacement should be principled. It is difficult to see the content of a notion such as *Last Resort* if the features driving displacement are obligatorily present where movement can be detected, obligatorily absent where movement operations cannot be motivated, and optionally present where human language grammars appear to permit more than one possible derivational route to convergence.

Many theoreticians have concurred with this general outlook with respect to the notion of EPP features, and there has accordingly been a number of efforts which aim to show that reference to unprincipled 'movement features' can be avoided.[43] What replaces EPP properties in such proposals varies widely, consonant with the intuition that there is no single underlying generalization to be had, but rather a conspiracy of many different sub-components of the grammar. Epstein & Seely (2006), for example, suggest that the work of EPP properties can be handled by appeal to other notions variously relating to our sub-theories regarding: (i) predication, (ii) locality, (iii) morphology, (iv) Case discharge/valuation, and (v) null complementizers. Authors that have taken this approach, which we may refer to as the 'eliminativist approach' (cf. fn. 5 and section 4), include also Manzini & Roussou (2000) as well as Bošković (2002), and, to some extent, Hornstein (2001, 2009).

In contrast, other work has instead pursued a 'reductivist approach', which attempts to retain the notion of EPP properties but aims to situate them in the architecture of the human language faculty in a principled way. In these terms, EPP properties have typically been suggested to be externally imposed on narrow syntax as interface conditions. These have included formulations or explanations of the EPP in terms of phonological properties and/or PF-conditions, semantic properties and/or LF-conditions, pure PF and/or LF

[43] We mentioned some in the text above; early minimalist accounts include Takahashi (1994), more recently the issue has been picked up by Boeckx (2008) and Hornstein (2009), for example.

interface conditions, or a combination of narrow-syntactic and interface properties of the grammar. See, for example, Landau (2007) for a recent overview of such attempts and his own take to combine syntactic and PF-driven motivation: Following Chomsky (2000), he takes the EPP to instantiate selectional requirements of functional heads, but he narrows this requirement down to p-selection by [P], 'a selectional feature that must be locally satisfied by some element with phonological content' (Landau 2007: 487, following Holmberg 2000; see also Roberts & Roussou 2002, Adger 2003, Haeberli 2003).

As a final point, the eliminativist and reductivist approaches are useful to keep in mind in parallel, examining the give-and-take between the two in terms of descriptive and explanatory adequacy. However, everybody agrees that unmotivated 'movement features' should have no place in the theory. This leads us to a stronger conclusion about one aspect of this paper: Where eliminativist approaches fail, reductivist approaches *have to* succeed.

6.2. Phases Today

A point we didn't make in this essay, but which we should have, concerns the relationship between the notion of phases marked by C/*v* (Chomsky 2000 *et seq.*, see also Richards 2004), and the notion of local convergence domains marked by T (as we initially proposed in Castillo, Drury & Grohmann 1997, as mentioned above). One of the conceptual oddities (in our view) of Chomsky's proposals regarding phases as marked by C/*v* is the way these phases are understood to relate to 'Spell-Out/Transfer'; that is, the idea that the *complement domain* of (strong) phase heads must spell-out at the point the next (strong) phase head is introduced. Setting aside the problems with the general perspective (see e.g. Boeckx & Grohmann 2007 for critical discussion), the proposed derivational timing of 'what spells-out when and why' doesn't really hang together clearly. For example, if phases are understood to be individuated on the basis of whether they correspond to units that can be regarded as propositional in some sense at the Conceptual-Intentional interface, why should only the *complement domain* spell-out? And why should it be that this does not happen until the next phase-inducing head is introduced? And what is the status of potential phase-heads in the nominal domain, and perhaps

even within prepositional phrases (see e.g. Abels 2003, Svenonius 2004, and Lee-Schoenfeld 2005, but also Chomsky 2005: 17–18).

The notion of T-domains as the relevant units which must spell-out picks out — at least with respect to C-phases — exactly the same 'Spell-Out domains' as in Phase Theory. But although the T-domain view of local convergence evaluation has to bear the burden of explaining *why* these domains should be special at the interface (see Grohmann 2003 for one possible explanation), no appeal to the properties of the heads that *select* these domains (e.g., C) is required — nor is there any need to arbitrarily slice up the dependents of heads into separate Spell-Out domains (complement domain vs. the rest). Just the stipulation regarding the Spell-Out domain(s) suffices, and this perspective refocuses the attention then on what's special about the 'T'-layer in clauses (rather than what may be special about the heads that select them).

Of course, it might seem that the T-domain view would, at least at first blush, have nothing to say about phases induced by *v*. However, if Pesetsky & Torrego (2004) are more generally correct in positing 'object-related' (lower) T, then it seems to us that the view of these domains as the locus of 'Spell-Out/Transfer' can be generalized in ways that should handle whatever the postulation of C/*v*-phases can handle, given that what's crucial in the broader set of efforts to connect locality of dependencies and the dynamics of interface evaluation is which parts of structural descriptions are actually transferred and hence unavailable to derivationally 'downstream' operations. Shifting attention to Spell-Out/Transfer domains makes phase-inducing heads 'special' only in that they happen to be the elements that unambiguously signal to the system that a domain which can/should be evaluated for convergence is (or should be) 'complete'.[44]

[44] This also squares more clearly than the C/*v*-phase conception, in our view, with the spirit of Last Resort concerning displacement operations more generally, in particular concerning when and where elements within the complement domain of a strong phase head can/must 'move' (see Drury 2005: 157–162 for an outline of the problem). On the T-domain view, in contrast, convergence is evaluated when these units are complete, which will be signaled by the introduction of the selecting head (whatever it is). This makes it possible to handle successive-cyclic movement as local-convergence-driven and subject in the strictest possible way to Last Resort.

However, there are presently lots of different conceptions of the notion of 'phase' and the mechanics of 'Spell-Out/Transfer' on the market (see e.g. Grohmann 2009 for a recent overview). Our sole point here with respect to this larger body of ongoing work is to take note of two different ways of thinking about Spell-Out/Transfer domains: (i) characterizing these domains directly (e.g., XPs of certain types 'spell-out') vs. (ii) characterizing Spell-Out domains in contextual/relational terms (e.g., complement domains of X-heads 'spell-out').[45] It seems to us that interface motivations for the (ii)-type approach taking C/*v* phases as 'propositional' don't relate in any clear way to the ideas about which units are alleged to be the Spell-Out domains (see also Boeckx & Grohmann 2007).

6.3. MOM Today

The MOM constraint was always connected with *Last Resort* logic: *ceteris paribus*, the system should 'do as little as possible'. MOM thus found a rationale, for example, in approaches which decomposed displacement operations into principled inventories of independent sub-operations (e.g., 'Move = Copy + Merge'). However, the manner in which this particular cost-of-operations view has to work turns out to require radically restricting the derivational alternatives that can be compared in ways that, in the end, obviates the need to refer to comparison at all. Similar conclusions have been independently reached by a number of other authors (see, among many others, the contributions collected in Epstein & Seely 2002).

However, the notion of *comparison* of derivations (or representations) with respect to 'do as little as possible' is of course alive and well in minimalist approaches as well as in other frameworks (e.g., Optimality Theory). The route we pursued in the present essay has a natural (though not inevitable) continuation in approaches which take syntactic derivations to be both 'crash-proof' (in the sense of Frampton & Gutmann 2002; see also the volume put together in Putnam, to appear) and deterministic in the sense of obviating any

[45] See Bobaljik & Wurmbrand (2005) on some reasons for favoring a contextual approach to demarcating phases, which in the terms discussed here can allow given sub-structures of type XP to either be a Spell-Out domain or not — depending on the selecting head. See also den Dikken (2007).

need to include the deployment of economy metrics in comparisons between otherwise convergent derivations. In approaches which adopt one or another of the available views that go under the label of 'Phase Theory', it remains to be established whether 'transderivational comparison' needs to be part of the theory.

We would like to close with pretty much the same conclusion we reached some ten years ago: Once (some form of) local convergence evaluation is introduced into the system, nothing needs to be said explicitly on '(transderivational) comparison'; in addition, it is not at all obvious that there is any real 'cost' distinction, or that such even could be made, with respect to narrow syntactic operations.

References

Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Ph.D. dissertation, University of Connecticut, Storrs.

Adger, David 2003. *Core Syntax: A Minimalist Approach*. Oxford: Oxford University Press.

Aoun, Josef. 1985. *The Grammar of Anaphora*. Cambridge, MA: MIT Press.

Baltin, Mark R. 1982. A landing site theory of movement rules, *Linguistic Inquiry* 13, 1–38.

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

Bobaljik, Jonathan David. 1995. Floating quantifiers: Handle with care, *GLOT International* 3, 3–10. [Reprinted in Lisa L.-S. Cheng & Rint Sybesma (eds.). 2003. *The Second Glot International State-of-the-Article Book*. 107–148. Berlin: Mouton de Gruyter.]

Bobaljik, Jonathan David. 2008. Where's Phi? Agreement as a post-syntactic operation. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi-Theory: Phi Features across Interfaces and Modules*. 295–328. Oxford: Oxford University Press.

Bobaljik, Jonathan David & Höskuldur Thráinsson. 1998. Two heads aren't always better than one, *Syntax* 1, 37–71.

Bobaljik, Jonathan David & Susi Wurmbrand. 2005. The domain of agreement, *Natural Language and Linguistic Theory* 23, 809–865.

Bobaljik, Jonathan David & Susi Wurmbrand. 2008. Case in GB/minimalism. In Andrej Malchukov & Andrew Spencer (eds.), *The Oxford Handbook of Case*. 44–58. Oxford: Oxford University Press.

Boeckx, Cedric. 2008. *Bare Syntax*. Oxford: Oxford University Press.

Boeckx, Cedric & Kleanthes K. Grohmann. Putting phases in perspective, *Syntax* 10, 204–222.

Boeckx, Cedric, Norbert Hornstein & Jairo Nunes. Forthcoming. *Control as Movement*. Cambridge: Cambridge University Press.

Bošković, Željko. 1997. *The Syntax of Non-Finite Complementation: An Economy Approach*. Cambridge, MA: MIT Press.

Bošković, Željko. 2002. A-movement and the EPP, *Syntax* 5, 167–218.

Castillo, Juan Carlos. 2001. Thematic relations between nouns. Ph.D. dissertation, University of Maryland, College Park.

Castillo, Juan Carlos, John E. Drury & Kleanthes K. Grohmann. 1997. Cyclic evaluation as a way to achieve local economy. Ms., University of Maryland, College Park.

Castillo, Juan Carlos, John E. Drury & Kleanthes K. Grohmann. 1999. The status of the Merge over Move preference. Ms., University of Maryland, College Park.

Chomsky, Noam. 1955. The logical structure of linguistic theory. Ms., Harvard University/Massachusetts Institute of Technology, Cambridge. [Published in part as *The Logical Structure of Linguistic Theory*, New York: Plenum, 1975.]

Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*. 232–286. New York: Holt, Reinhart, & Winston.

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use.* New York: Praeger.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale & Samuel Jay Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. 1–52. Cambridge, MA: MIT Press. [Reprinted in Chomsky (1995), 167–217.]

Chomsky, Noam. 1994. Bare phrase structure, *MIT Occasional Papers in Linguistics* 5. [Reprinted in Gert Webelhuth (ed.), *Government and Binding Theory and the Minimalist Program*. 383–439. Malden, MA: Blackwell, 1995.]

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

Chomsky, Noam. 1998. Minimalist inquiries: The framework. *MIT Occasional Papers in Linguistics* 15. Cambridge, MA: MITWPL. [Reprinted as Chomsky (2000).]

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. 89–155. Cambridge, MA: MIT Press.]

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*. 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and Beyond — The Cartography of Syntactic Structures*, vol. 3. 104–131. New York: Oxford University Press.

Chomsky, Noam. 2005. Three factors in language design, *Linguistic Inquiry* 36, 1–22.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans–Martin Gärtner (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Semantics*. 1–29. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*. 133–166. Cambridge, MA: MIT Press.

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

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

Dowty, David & Belinda Brodie. 1984. A semantic analysis of 'floated' quantifiers in a transformationless grammar, *WCCFL 3: Proceedings of the Third West Coast Conference on Formal Linguistics*, 75–90.

den Dikken, Marcel. 2007. Phase extension: Contours of a theory of the role of head movement in phrasal extraction, *Theoretical Linguistics* 33, 1–41.

Drury, John E. 2005. Alternative directions for minimalist inquiry: Expanding and contracting phases of derivation. Ph.D. dissertation, University of Maryland, College Park.

Embick, David. 1997. Voice and the interfaces of syntax. Ph.D. dissertation, University of Pennsylvania

Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima & Hisatsugu Kitahara. 1998. *A Derivational Approach to Syntactic Relations*. New York: Oxford University Press.

Epstein, Samuel David & T. Daniel Seely. 1999. On the non-existence of the EPP, A-chains, and successive cyclic A-movement. Ms., University of Michigan, Ann Arbor & Eastern Michigan University, Ypsilanti.

Epstein, Samuel David & T. Daniel Seely (eds.). 2002. *Derivation and Explanation in the Minimalist Program*. Malden, MA: Blackwell.

Epstein, Samuel David & T. Daniel Seely. 2006. *Derivations in Minimalism*. Cambridge: Cambridge University Press.

Fiengo, Robert & Howard Lasnik. 1976. Some issues in the theory of transformations, *Linguistic Inquiry* 7, 182–191.

Frampton, John & Sam Gutmann. 1999. Cyclic computation: A computationally efficient minimalist syntax, *Syntax* 1, 1–27.

Frampton, John & Sam Gutmann. 2002. Crash-proof syntax. In Epstein & Seely (eds.), 90–105.

Freidin, Robert F. 1978. Cyclicity and the theory of grammar, *Linguistic Inquiry* 9, 519–549.

Freidin, Robert. 1986. Fundamental issues in the theory of binding. In Barbara Lust (ed.), *Studies in the Acquisition of Anaphora*. 151–181. Dordrecht: Reidel.

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

Grohmann, Kleanthes K. 2000. Prolific peripheries: A radical view from the left. Ph.D. dissertation, University of Maryland, College Park.

Grohmann, Kleanthes K. 2003. *Prolific Domains: On the Anti-Locality of Movement Dependencies*. Amsterdam: John Benjamins.

Grohmann, Kleanthes K. 2009. Phases and interfaces. In Kleanthes K. Grohmann (ed.), *InterPhases: Phase-Theoretic Investigations of Linguistic Interfaces*. 1–22. Oxford: Oxford University Press.

Grohmann, Kleanthes K., John E. Drury & Juan Carlos Castillo. 2000. No more EPP, *WCCFL 19: Proceedings of the Nineteenth West Coast Conference on Formal Linguistics*, 153–166.

Haeberli, Eric 2003. Categorial features as the source of EPP and abstract case phenomena. In Ellen Brandner & Heike Zinsmeister (eds.), *New Perspectives on Case Theory*. 89–126. Stanford, CA: CSLI Publications.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. 111–176. Cambridge, MA: MIT Press.

Hiraiwa, Ken. 2005. Dimensions of symmetry in syntax: Agreement and clausal architecture. Ph.D. dissertation, MIT.

Holmberg, Anders. 2000. Scandinavian stylistic fronting: How any category can become an expletive, *Linguistic Inquiry* 31, 445–483.

Hornstein, Norbert. 1994. An argument for minimalism: The case of antecedent-contained deletion, *Linguistic Inquiry* 25, 455–480.

Hornstein, Norbert. 1999. Movement and control, *Linguistic Inquiry* 30, 69–96.

Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal*. Malden, MA: Blackwell.

Hornstein, Norbert. 2009. *A Theory of Syntax*. Cambridge: Cambridge University Press.

Hornstein, Norbert, Jairo Nunes & Kleanthes K. Grohmann (2005). *Understanding Minimalism*. Cambridge: Cambridge University Press.

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

Johnson, David & Shalom Lappin. 1997. A critique of the minimalist program, *Linguistics & Philosophy* 20, 272–333.

Johnson, David & Shalom Lappin. 1998. *Local Constraints vs. Economy*. Stanford, CA: CSLI Publications.

Johnson, Kyle. 1991. Object positions, *Natural Language and Linguistic Theory* 9, 577–636.

Jonas, Dianne & Jonathan David Bobaljik. 1993. Specs for subjects: The role of TP in Icelandic, *Papers on Case and Agreement I: MIT Working Papers in Linguistics* 18, 59–98.

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

Klein, S. 1976. A base analysis of the floating quantifier in French, *NELS VII: Proceedings of the Seventh Annual Meeting of the North Eastern Linguistics Society*, 147–163.

Koizumi, Masatoshi. 1993. Object agreement and the spilt-VP hypothesis, *Papers on Case and Agreement I: MIT Working Papers in Linguistics* 18, 99–148.

Landau, Idan. 2007. EPP extensions, *Linguistic Inquiry* 38, 485–523.

Lasnik, Howard. 1995a. Verbal morphology: *Syntactic Structures* meets the minimalist program. In Paula Kempchinsky & Héctor Campos (eds.),

Evolution and Revolution in Linguistic Theory: Essays in Honor of Carlos Otero. 251–275. Washington, DC: Georgetown University Press.

Lasnik, Howard. 1995b. Last resort. In Shosuke Haraguchi & Michio Funaki (eds.), *Minimalism and Linguistic Theory*. 1–32. Tokyo: Hituzi Syobo. [Reprinted as 'Last resort and attract F' in Lasnik, Howard. 2003. *Minimalist Investigations in Linguistics Theory*. 22–41. London: Routledge.]

Lasnik, Howard. 1997. Levels of representation and the elements of anaphora. In Hans Bennis, Pierre Pica & Johan Rooryck (eds.), *Atomism and Binding*. 251–268. Dordrecht: Foris. [Reprinted in Lasnik, Howard. 2003. *Minimalist Investigations in Linguistics Theory*. 42–54. London: Routledge.]

Lasnik, Howard. 2001. A note on the EPP, *Linguistic Inquiry* 32, 356–362.

Lasnik, Howard & Mamoru Saito. 1992. *Move α: Conditions on Its Application and Output*. Cambridge, MA: MIT Press.

Lasnik, Howard & Mamoru Saito. 1993. On the subject of infinitives, *CLS 27: Papers from the 27th Regional Meeting of the Chicago Linguistic Society*, 324–343.

Lebeaux, David. 1988. Language acquisition and the form of the grammar. Ph.D. dissertation, University of Massachusetts. [Published as *Language Acquisition and the Form of the Grammar*, Amsterdam: John Benjamins, 2000.]

Lee-Schoenfeld, Vera. 2005. Beyond coherence: The syntax of opacity in German. Ph.D. dissertation, University of California, Santa Cruz.

Legate, Julie Anne. 2008. Morphological and abstract case, *Linguistic Inquiry* 39, 55–101.

Mahajan, Anoop. 1992. The specificity condition and the CED, *Linguistic Inquiry* 23, 510–516.

Maling, Joan. 1975. Notes on quantifier-postposing, *Linguistic Inquiry* 7, 708–718.

Maling, Joan. 1988. Variations on a theme: Existential sentences in Swedish and Icelandic, *McGill Working Papers in Linguistics: Special Issue on Comparative Germanic Syntax*, 168–191.

Manzini, M. Rita & Anna Roussou. 2000. A minimalist theory of A-movement and control, *Lingua* 110: 409–447.

Marantz, Alec. 1991. Case and licensing, *ESCOL '91: Proceedings of the Eastern States Conference on Linguistics 1991*, 234–253.

Martin, Roger. 1999. Features and minimalism. Ms., Tsukuba University.

Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.

Pesetsky, David. 1991. Zero syntax II: An essay on infinitives. Ms., MIT.

Pesetsky, David & Esther Torrego. 2001. T-to-C: Causes and consequences. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*. 355–426. Cambridge, MA: MIT Press.

Pesetsky, David & Esther Torrego. 2004. Tense, case, and the nature of syntactic categories. In Jacqueline Guéron & Jacqueline Lecarme (eds.), *The Syntax of Time*. 495–537. Cambridge, MA: MIT Press.

Platzack, Christer. 1983. Existential sentences in English, Swedish, German and Icelandic, *Papers from the Seventh Scandinavian Conference of Linguistics*, 80–100.

Postal, Paul M. 1974. *On Raising: One Rule of English Grammar and Its Theoretical Implications*. Cambridge, MA: MIT Press.

Putnam, Michael T. (ed.). To appear. *Exploring Crash-Proof Grammars*. Amsterdam: John Benjamins.

Reinhart, Tanja & Eric Reuland. 1993. Reflexivity, *Linguistic Inquiry* 24, 657–720.

Richards, Marc D. 2004. Object shift and scrambling in North and West Germanic: A case study in symmetrical syntax. Ph.D. dissertation, University of Cambridge.

Roberts, Ian & Anna Roussou. 2002. The Extended Projection Principle as a condition on the tense dependency. In Peter Svenonius (ed.), *Subjects, Expletives, and the EPP*. 125–155. Oxford: Oxford University Press.

Sells, Peter. 1987. Aspects of logophoricity, *Linguistic Inquiry* 18, 445–479.

Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure, *Linguistic Inquiry* 19, 425–49.

Sternefeld, Wolfgang. 1997. Comparing reference sets. In Chris Wilder, Hans-Martin Gärtner & Manfred Bierwisch (eds.), *The Role of Economy Principles in Linguistic Theory*. 81–114. Berlin: Akademie-Verlag.

Stowell, Tim. 1981. Origins of phrase structure. Ph.D. dissertation, MIT.

Svenonius, Peter. 2004. On the edge. In David Adger, Cécile de Cat & George Tsoulas (eds.), *Peripheries: Syntactic Edges and Their Effects*. 259–287. Dordrecht: Kluwer.

Takahashi, Daiko. 1994. Minimality of movement. Ph.D. dissertation, University

of Connecticut, Storrs.

Uriagereka, Juan. 1998. *Rhyme and Reason: An Introduction to Minimalist Syntax.* Cambridge, MA: MIT Press.

Uriagereka, Juan. 1999a. 'Minimalist Issues'. Class lectures, University of Maryland, College Park.

Uriagereka, Juan. 1999b. Multiple Spell-Out. In Samuel David Epstein & Norbert Hornstein (eds.), *Working Minimalism*. 251–82. Cambridge, MA: MIT Press.

Vikner, Sten. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. New York: Oxford University Press.

Wilder, Chris & Hans-Martin Gärtner. 1997. Introduction. In Chris Wilder, Hans-Martin Gärtner & Manfred Bierwisch (eds.), *The Role of Economy Principles in Linguistic Theory*. 1–35. Berlin: Akademie-Verlag.

Williams, Edwin. 1980. Predication, *Linguistic Inquiry* 11, 208–238.

Williams, Edwin. 1994. *Thematic Structure in Syntax*. Cambridge, MA: MIT Press.

Yang, Charles D. 1997. Minimal computation: Derivation of syntactic structures. M.Sc. thesis, MIT.

Zwart, C. Jan-Wouter. 1992. Dutch expletives and small clause predicate raising. *NELS 22: Proceedings of the Twenty-Second Annual Meeting of the North Eastern Linguistics Society*, 477–491.

Juan Carlos Castillo
john.drury@mcgill.ca
University of Northern Iowa

John E. Drury
juan.castillo@uni.edu
McGill University

Kleanthes K. Grohmann
kleanthi@ucy.ac.cy
University of Cyprus