PHASING OUT PHASES AND RE-RELATIVIZING RELATIVIZED MINIMALITY[*]

J O H N   T O R R
University of Cambridge

ABSTRACT   There is considerable redundancy between Phase Theory (PT) (Chomsky 2000, 2001, 2005a, b, 2007) and the theory of Relativized Minimality (RM) (Rizzi 1990, 2002; the Minimal Link Condition of Chomsky 1995a), an impoverished version of which is retained in the form of 'intervention effects' in recent work on phases by Chomsky. While 'rich' RM offers the potential for a fully unified account of locality, which can be grounded evolutionarily in processing and 'third factor' (Chomsky 2005a) considerations, phases appear capable of accounting only stipulatively for a narrow subset of the same phenomena. Adopting Chomsky's (2005a: 10) 'guiding intuition that redundancy in computational structure is a hint of error,' I argue that phases can very profitably be eliminated from the theory entirely, with a richer and strongly re-relativized version of RM reinstated as the principal account of locality in Generative Grammar.

1   INTRODUCTION

The Minimalist Program explores the possibility that language, as a biological 'organ of the body' (Chomsky 2005a: 1), may be in some sense 'perfect,' an 'optimal solution to legibility conditions' (The Strong Minimalist Thesis – Chomsky (1998: 6-7)). Chomsky (2005a: 1) argues that the Biolinguistic perspective on the human language faculty (FL) implicates three factors in the shaping of an individual's eventual linguistic 'steady-state' or I-language: genetic endowment (UG), linguistic experience (PLD), and 'principles that are language- or even organism-independent' – what he refers to as 'third factor' considerations, including general principles of data analysis and computational efficiency. In order to be considered 'principled,' any explanation of a given linguistic phenomenon will then ideally meet the following two criteria: first, it

---

should ultimately be 'reduc[ible] to properties of the interface systems' (Chomsky 2005a: 10), that is, to the requirements of bare output (or 'legibility') conditions imposed by the sensorimotor (SM) and (especially) the conceptual-intentional (C-I) performance systems; second, it should be 'optimal' in the sense of maximally computationally efficient and non-redundant.

I argue here that despite first appearances, phases constitute non-optimal solutions in several critical respects. Symptomatic is the fact that both phases and (rich) RM appear to cover much of the same conceptual and empirical ground – each apparently reducing the search spaces between probes and goals, and each enforcing successive cyclic A'-movement, for example. Such redundancy is highly suspect from a Minimalist perspective, and should urge us to make do with just one of these components.

One of the principal empirical arguments for phases is based on the Phase Impenetrability Condition (PIC; Chomsky 2000), which offers an account of various locality phenomena, from successive cyclic A'-movement to CED effects (Huang 1982); these were dealt with in GB by bounding nodes and the ECP, and later by barriers (Chomsky 1986). However, phases were not the only, nor were they the first, alternative to the ECP and barriers of late GB. Until Chomsky (2000), Relativized Minimality had enjoyed over a decade of scholarly acceptance as the principal account of locality; indeed, Chomsky (2001, 2005a, b, 2007) continues to make reference to 'intervention effects.' Nevertheless, while he does not seek to supplant RM entirely, in Chomsky (2001: 26) he does attempt to eliminate the concomitant notion of Equidistance from the theory. I argue here that this move was misconceived, and that instead it is phases themselves which should be abandoned.

The paper is arranged as follows: section 2.1 begins by considering the arguments for phases based on computational efficiency; section 2.2 then outlines Hornstein's (2001, 2009) radically Minimalist and RM-based approach to locality, which I adopt here; section 2.3 discusses some support for RM from processing, third factors and evolution; in section 3 we move on to looking at successive cyclic wh-movement, including a discussion of the motivation for intermediate EPPs; section 4 does the same for successive cyclic A-movement, and tackles Chomsky's (2001) arguments against Equidistance; finally, a conclusion is given in section 5.

2 General Conceptual Arguments

2.1 Computational Efficiency

Phases initially appear to offer a number of important computational advantages: the search spaces between probes and goals are apparently narrowed

considerably, certain look-ahead and look-back procedures obviated, and the amount of information which must be held in 'active memory' substantially reduced. However, search spaces and 'active memory' load are only reduced in clauses containing CPs and transitive vPs (Chomsky's (2005b) 'v*P' (a convention which I henceforth adopt)). This is because in Chomsky's (2001, 2005b) system TP cannot be a phase[1], and neither can unaccusative vP[2]. However, this means that for cases of successive cyclic A-movement across (a potentially unlimited number of) TP clauses and unaccusative vPs, these computational gains are entirely negated. Worse still, even transitive structures must allow the Binding Principles A (for anaphors and OC PRO), B and C to operate across phase boundaries.[3]

Chomsky also argues that on optimal assumptions Spell-Out must be assumed not to 'look ahead' to determine which features are interpretable at the interfaces; nor should it be able to 'look back' in order to ascertain which features entered the derivation unvalued. Chomsky assumes that although Spell-Out is not sensitive to feature interpretability (not being 'semantic'), it is sensitive to the valued/unvalued contrast which mirrors it. Hence he concludes that in order to 'know' which features to strip away from the structure to be sent to LF, Spell-Out must occur 'shortly after the uninterpretable features have been assigned values' (Chomsky 2001: 5). However, Epstein and Seely (2002: 85) point out that the 'shortly after' argument doesn't really stack up: If Spell-Out is blind to the interpretability distinction between those features which are inherently valued and those which receive their valuation during the course of the derivation, then it will be blind as soon as the valuation has taken place; whether Spell-Out occurs immediately after the valuation, or at some later point, should therefore make no difference.

2.2   A Radically Minimalist Theory of Locality

Perhaps recognizing the problems posed by binding for PT, Chomsky argues that 'BT is at the outer edge of the C-I interface' (Chomsky 2005b: 8). Yet the fact that all three of BT's Principles make use of c-command relations argues strongly for a syntactic treatment. Hornstein (2001, 2009) therefore proposes

---

1 Otherwise wh-movement would have to proceed through spec TP as well as spec CP.

2 Or at least, it is not a strong phase and hence does not transfer its complement to Spell Out; otherwise logical objects could not raise to spec TP. However, phase strength appears to be little more than a stipulation intended to get around observations by Legate (2003) regarding the phase-like behaviour of unaccusative vP.

3 That Principle C can operate across an unlimited number of phase boundaries is obvious. For cases where Principles A and B operate across one phase boundary, consider *for*-infinitive structures such as the following: $Jack_i$ wants more than anything [$_{CP}$ for *him_i/$himself_i$ to win]

subsuming BT locality effects under general Minimality principles. He argues that anaphors are the overt residue of A-movement, and therefore fall within the remit of RM rather than Principle A. Control[4] too, argues Hornstein, is more parsimoniously viewed as an instance of A-movement.[5] Principle B, he suggests, is simply the elsewhere case which results in bound pronominalization: what the grammar does when movement is not an option. Finally, Principle C can be accounted for by supposing that where bound pronominalization can occur, it must occur. The clear parallels between binding and A-movement

---

4 Also handled by BT since the abandonment of the PRO theorem.

5 Note that this removes the arguments for CP being a phase based on ECM/raising vs. control phenomena: in Chomsky's (2000) PT system the distinction between these constructions relies on the contents of 'defective' TP clauses being accessible to higher probes, in contrast to those of their CP counterparts; in the latter, PRO is sheltered by a CP phase and has its Case feature valued as null and deleted by φ-complete non-finite control T. Non-finite ECM/raising T, meanwhile, lacking a [number] feature, is unable to value and delete its specifier's Case, that specifier nevertheless able (rather fortuitously) to enter into Agree relations with matrix v (for ECM) or matrix T (for raising), owing to the lack of a CP layer here. This is certainly an elegant solution. But note the unappealingly stipulative nature of the claim that raising T lacks [number], as well as the arbitrariness (noted by Hornstein et al. 2010: 20) of having an entire subcategory of Case reserved for a single lexical item, PRO. In later work, Chomsky (2005b) assumes raising and ECM T to lack φ-features altogether, and derives this property from the unavailability of inheritable features from a governing C, unlike in the control case. But if, as Chomsky claims, T cannot act as a probe until C is introduced into the derivation, then one wonders why the complements of ECM predicates can have expletive 'there' subjects (*Jack believes there to be a problem*), which even allowing for a raising to object (spec VP) analysis (Chomsky 2009) can surely not have been base generated either within the embedded v(*)P or in spec of matrix VP, for thematic reasons. This suggests that ECM T can carry its own inherent EPP feature, and given that EPP appears to be associated very intimately with probehood, this in turn would seem to call into question any analysis of the differences between raising/ECM and control which relies in part on the availability of probe inheritance.

   Alternatively, if we adopt the Movement Theory of Control, we no longer require these particular distinctions between ECM/raising and control, which appear to be something of a hangover from GB's overly baroque defective T and blocking category vs. barrier (by inheritance) stipulations. These were designed to accommodate the antagonistic requirements of the ECP, the Case filter and the PRO theorem: though each appearing in spec of non-finite TP, traces, overt nominals and PRO had respectively to be antecedent governed, assigned Case under government, and ungoverned; with non-finite T itself arbitrarily stipulated as unable to govern its specifier (hence unable to assign Case and prevent raising, and unable to govern PRO), non-finite TP had paradoxically to be at once opaque *and* transparent to outside government. Under the MTC, however, the need for either GB's or Phase-Theoretic Minimalism's highly complex and stipulative technology in this domain, along with the above arguments for phases, immediately evaporates: PRO is now treated as an A-trace, its antecedent having A-raised out of the CP 'phase' - crucially, without having first passed through its edge. (And see Hornstein et al. (2010: 125-130) for an interesting RM-based solution to Landau's (2003) claim that the MTC wrongly predicts that subject control constructions should be passivizable (if CP is indeed not a phase), just as ECMs are; i.e. Landau contends that e.g. *$Jack_i$ was tried t_i to kiss Mary* is incorrectly predicted by the MTC to be on a par with $Jack_i$ was expected t_i to kiss Mary).

were of course explicitly acknowledged in Chomsky (1981), where A-traces are regarded as null anaphors[6]; Rizzi (1990: 7) also recognizes that his incipient theory of RM bares 'a certain similarity with the Theory of Binding' in that antecedent governors as interveners are very much like the Specified Subjects or Accessible Subjects of Chomsky's (1973, 1981) definitions of Binding Domains: 'Relativized Minimality, in a sense, generalizes this idea to government relations.' (Rizzi 1990: 8)

However, Hornstein argues for the reverse generalization: Move is clearly needed independently anyway, and so its principles should be extended so as to account for all types of construal. He points out the striking number of distributional characteristics shared by OC PRO, locally bound anaphors and A-traces on the one hand, and NOC PRO and pronouns/non-locally bound anaphors on the other. Members of the former group must have an antecedent by which they are locally bound in a c-command configuration, while in the latter group the constituents need not have an antecedent; and even if one is present, it need be neither local nor c-commanding. Furthermore, structures involving the former group members can only give rise to sloppy interpretations under ellipsis, must have obligatory 'de se' readings in 'unfortunate' contexts, and do not admit split antecedents; by contrast the latter group are found in structures admitting both sloppy and strict readings under ellipsis, may give rise to non-'de se' readings and can take split antecedents (see Hornstein et al. (2010: 47); and Hornstein (2001: 155) for discussion).

Such parallels are simply too strong to be ignored, and point inexorably towards the need and very real possibility of a fully unified account of locality. Phases are simply too rigid to be of any use here; RM, on the other hand, can straightforwardly be pressed into service.

2.3   Third Factors, Processing and Evolution

Consider the fact that, like phases, RM is also aimed at reducing computational complexity, by minimizing the distance between antecedents and traces. The father of RM, Luigi Rizzi (1990, 2002), summarizes its conceptual benefits in the following terms:

---

6 A problem with this approach was that Chomsky (1981, 1986) continued to assume that all movement is constrained by a number of separate locality modules in addition to the Binding Theory, including Bounding Theory and the ECP. Yet the fact that, for example, it is difficult if not impossible to find a single ECP or subjacency violation by A-traces which is not also a Binding Theory violation, argues for some considerable redundancy here. Even Chomsky's (1986) complex Barriers system only seeks to conflate Bounding Theory and the ECP, still assuming Binding Theory to constitute a separate locality module. With RM, PT and the BT each handling various overlapping aspects of locality in current theory, I would argue that such redundancy still very much haunts us.

> RM can be intuitively construed as an economy principle in
> that it severely limits the portion of structure within which a
> given local relation is computed...the principle reduces ambigu-
> ity in a number of cases: whenever two elements compete for
> entering into a given local relation with a third element, the
> closest always wins.
> (Rizzi 2002: 224)

As with phases then, search spaces are delimited, but this time the desired results also apply to A-movement (and even head-movement; cf Travis' (1984) Head Movement Constraint), and need not be arbitrarily stipulated in terms of the phase heads v* and C; rather, RM as part of competence (call this RM-C) plausibly receives direct motivation from processing concerns, since in this case the link can quite naturally be derived via an evolutionary response of the competence systems to the demands of performance, the parser naturally preferring minimal ambiguity with respect to filler-gap dependencies (call this RM-P). In this crucial respect, then, as Rizzi (2002: 224) points out, RM 'appears to be a *natural* principle of mental computation' (my emphasis; JT).

On the natural link between RM and processing, Ortega-Santos (2011) has recently presented a paper in which he argues that 'Relativized Minimality is a conventionalized property of the grammar that is functionally grounded as a response to memory [and the exigencies of] a cue-based retrieval parser [constrained by] similarity-based interference' (Santos 2011: 35). Van Dyke & Lewis (2003; cited in Santos 2011: 41) present experimental results from a self-paced reading and grammaticality judgement task involving garden path sentences. They showed that pairs of sentences containing identical numbers of words, and identical numbers of nouns in similar ambiguous regions of each sentence, nevertheless exhibited differential rates of interference depending on how closely the linguistic features – or 'cues' – of the interfering word matched those of the word to be retrieved. For example, a copula verb like 'be' interfered less in the retrieval of a factive verb like 'forget' than did another factive verb like 'know'; additionally, a Nominative singular noun interfered to a lesser extent with the retrieval of an Accusative plural noun than another Nominative singular noun.

As Santos points out, ascribing to the parser the properties of cue-based retrieval and decay has the considerable advantage that these are both found extensively in other cognitive domains; interference effects of this kind have been observed with respect to people's memory for motor skills (Adams 1987; cited in Santos 2011: 40) and visual stimuli (Chandler 1991; cited in Santos 2011: 40) as well as in the retrieval of information accessed during exercises in mental arithmetic (Campbell 1991; cited in Santos 2011: 40). Such do-

main general effects therefore begin to look suspiciously like prototypical third factor constraints, opening up the further possibility that RM may not only be evolutionarily derivable as a response to the demands of parsing, but may itself simply be a reflection of how the world – or at least cognition – necessarily works (thereby taking us 'beyond explanatory adequacy' in the sense of Chomsky 2004).

If this line of reasoning is correct, then we would predict that there should also be RM-like effects in other areas of the linguistic system, not just in syntax. Rizzi (2004: 227) points out that interference effects are indeed found pervasively in phonology, and cites an example presented by Morris Halle of assimilation blocking in Sanskrit:

> "A *Coronal* nasal assimilates the *Coronal* features from a retroflex consonant that precedes it...The nasal can be arbitrarily far away from the retroflex, provided that *no Coronal consonant intervenes.*"
> (Halle 1995: 22; cited in Rizzi 2002: 227; my emphasis—JT)

The critical point here is that all of these effects display evidence not of rigid locality, but of relativized locality, what Santos (2011: 35) calls 'similarity-based interference.' But he also argues that despite the obvious parallels between RM-C and RM-P/third factor considerations, the former cannot simply be reduced to the latter, contra certain attempts in the literature to do just that (e.g. Pritchett (1991), Kluender (2004)). The reason is that there are certain properties inherent to RM-C which are found in neither RM-P nor other cognitive domains, nor even in some cases in other linguistic domains; these include, for instance, the role of c-command (i.e. hierarchical as opposed to simple linear ordering as found in phonological RM-like effects), the existence of Minimal Domains and Equidistance, cross-linguistic variation with respect to island phenomena etc.

Now consider phases. Although these encode a structurally defined notion of locality – i.e. in terms of predetermined v*P and CP phases – such rigid locality appears to have no obvious correlate either elsewhere in the linguistic system or in cognition more generally. Nor does an evolutionary explanation appear available here: despite the various processing-related metaphors of PT such as 'active memory,' 'transfer,' and of course 'phases' themselves, in this case a link with processing appears untenable given that phases (like Minimalist derivations more generally) are constructed in bottom up fashion, while (in processing) the right branching structures of a head initial language such as English are clearly parsed from the top down. That is, while phases (informally speaking) reduce 'active memory' load from 'right to left' in a

right branching language (which, if we follow Kayne (1994) would actually include all languages), this does not mirror the parser, which operates from 'left to right.' Although such concerns in no way invalidate a bottom-up-based derivational model of competence, I would argue that they do seem to deny phases the sort of appealingly natural evolutionary link with processing that is open to RM.

3   SUCCESSIVE CYCLIC WH-MOVEMENT

There is considerable evidence presented in the literature (see e.g. Radford 2004: 394-407) for successive cyclic wh-movement through both spec CP and spec v*P (and even spec vP, see e.g. Legate (2003)), from morphological facts concerning participle agreement in French, to multiple spellout of lower copies in intermediate spec CPs in colloquial German/Dutch/Afrikaans, to various reconstruction effects on Binding. However, teasing out precisely how and why such intermediate movement takes place has proved highly problematic.

In recent theory (e.g. Chomsky 2005b, 2007), movement has been partially divorced from features such as Case, and reassigned instead to EPP or Edge features. This appears necessary once the covert movement cycle is dropped, in order to account for e.g. Nominative Case assignment to in situ logical objects in unaccusative expletive 'there' constructions such as 'there arrived several men.' Chomsky (2005: 7) notes that 'to a large extent, External Merge yields generalized argument structure...; and Internal Merge yields discourse-related properties such as old information and specificity, along with scopal effects.' EPP, then, to a first approximation, allows for the instantiation of a duality of semantics. And it may be that the highest EPP in a chain is always motivated by discourse/scopal considerations[7], or at least it was when speakers first innovated its assignment to a given head, its discourse significance perhaps becoming bleached over time.

However, what is surely not the case is that each individual successive cyclic movement through intermediate spec v(*)P and spec CP positions is motivated by discourse/scopal considerations. As Rizzi (2004: 9) notes, 'the paradox of these intermediate positions is that, on the one hand they must autonomously cause a movement step (if we want to take seriously the idea that each step is locally determined, with no "look-ahead" to subsequent derivational steps), and at the same time we should make sure that movement should not stop there.'

PT attempts to motivate such intermediate movement by assuming that once a phase has been constructed, the complement of the phase head (i.e. TP

---

7 Rizzi (2004) argues, for instance, that both fronted topics and subjects raised to spec TP share a general 'aboutness' property.

or VP) is handed over to the interfaces by Spell Out, thus becoming immune to further computations. As a result, only the phase 'edge' (including the phase heads v* or C and their specifiers), will be accessible to a higher probe.

However, RM also provides motivation for successive cyclic wh-movement through both spec v(*)P and spec CP. We will first look at the motivation for constituents moving through filled spec v*P and spec CP positions. This is important because I will argue here that what goes for one spec XP position in one language with respect to successive cyclicity, must necessarily go for all spec XP positions in all languages, if C_{HL} is assumed to be unable to make use of look-ahead procedures, i.e. moving constituents selectively through certain spec XP positions 'in order to' avoid Minimality effects. Where languages can vary parametrically is in the number of (non-base generated) specifiers they permit in a given XP position, or, equivalently, the number of EPPs which are permitted on a given head (a factor which one would ultimately hope to derive via a suitably refined feature geometric theory of heads). If this number is one or greater, escape hatches will become available through Equidistance (see below), if it is less than one they will not.

The version of Relativized Minimality adopted here is essentially the derivational version formulated as The Minimal Link/Shortest Move Condition presented in Chomsky (1995a: 355) and Hornstein et al (2005: 149-163):

> K attracts α only if there is no β, β closer to K than α, such that K attracts β.
> (Chomsky 1995a: 311)

On this account, then, in making the shortest movement possible, a constituent α must not cross another constituent β which c-commands the trace of α and is of the same 'type' as α (in the GB terms of Rizzi (1990) a constituent β which is a potential antecedent governor for the trace of α). There is a special exemption, however, if the moved and intervening elements are 'Equidistant' from the target or source of the movement, where Equidistance and the concomitant notion Minimal Domain can be defined as follows:

Equidistance
> If two positions α and β are in the same Minimal Domain, they are Equidistant from any other position.

Minimal Domain
> The Minimal Domain of α, or MinD(α), is the set of categories immediately contained or immediately dominated by projections of the head α, excluding projections of α.
> (Hornstein et al. 2005: 149, 163)

With these definitions in mind, consider the following examples illustrating superiority effects:

(1)   a. * What did who buy?

          b.   What did she buy?

          c.   Who bought what?

Only 1(a) is deviant (unless D-linking features are introduced by adding emphatic stress to 'what' to indicate an echo question). To see why, consider the derivations of these three structures:
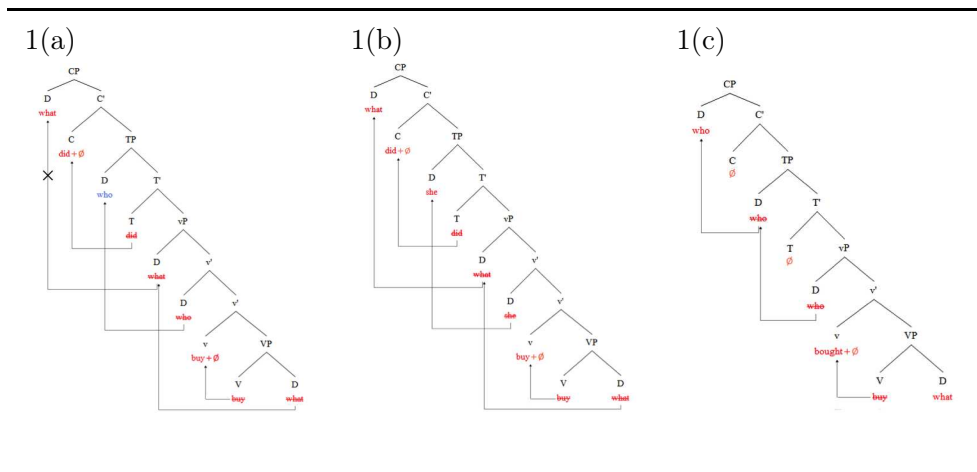
Consider first the structure in 1(a). According to the theory of RM, the illicit move here comes when the wh-object moves from the outer spec v*P position to spec CP. The reason is that in so doing it must cross the wh-subject in spec TP. The wh-object's extraction site is in MinD(v*) and its landing site is in MinD(C). Since the wh-subject is in neither of these – but is in MinD(T) – the move is illicit. The full acceptability of 1(b) then follows from the fact that the D in spec TP in this case carries no uninterpretable wh-feature to trigger the intervention effect. The full acceptability of example 1(c), meanwhile, follows from the fact that the wh-object remains in its base generated position, with the wh-subject raising first to spec TP and then to spec CP; since no crossing of wh-elements occurs here, there can be no RM violation.

Returning now to the question of why successive cyclic movement proceeds through spec v*P, on the current analysis the initial movement in 1(a), by which the wh-object moves to outer spec v*P, crossing the (at this point) in situ wh-subject in inner spec v*P, is perfectly licit, since both the object's landing site and the subject are located in MinD(v*), hence both positions are

Equidistant from the object's extraction site. Similarly, since Equidistance is here defined with respect to the sources as well as the targets of movement, both the subject and the object in the spec v*P positions are Equidistant from any other position at this point; hence when the subject moves to spec TP, crossing the object in the outer spec v*P position, there is again no Minimality violation. Clearly then, movement through spec v*P might be motivated on the grounds that it enables the circumvention of Minimality effects in the v*P domain.

Two questions arise at this point, however: 1. what good is such circumvention, if subsequent movement of a wh-element past spec TP will in any case induce the superiority effects seen in 1(a) above? And 2. given that in English spec of *unaccusative* vP and spec CP will never be host to more than one wh-element (and therefore will never serve as escape hatches by Equidistance), what motivates obligatory successive cyclic movement through these positions?[8]

The answer to both questions is that the situation in English is far from universal. Consider first question 1: there are many languages in the world in which subjects appear to remain in their base-generated spec v*P position rather than raising to spec TP from where they can induce Minimality violations on passing wh-constituents (provided these subjects are themselves wh-elements). McCloskey and Chung (1987), for instance, argue that in VSO languages like Irish the verb raises to T while the subject remains in its base generated position internal to the verb phrase. Assuming successive cyclic wh-movement through spec v*P, we can then straightforwardly account for the lack of superiority effects seen in Irish:[9]

(2)   ENGLISH:

       a. * What$_i$ does John believe that who bought t_i?

       b. * What$_i$ does who believe that Mary bought t_i?

---

8 See Legate (2003) for evidence of successive cyclic wh-movement through spec of unaccusative vP in addition to transitive v*P.

9 A complicating factor is that other VSO languages, such as closely related Welsh, do show superiority effects. Clearly there are other factors in play here therefore. For instance, there is some evidence from adverb placement that in Celtic languages subjects do in fact move to a head intermediate between spec v*P and spec TP, i.e. to an AspP (McCloskey 1996). Given my proposal that cross linguistic parametric differences with respect to locality arise as a result of the ability/inability of certain heads to host the requisite EPPs, together with the suggestion I make below that the wh-EPPs involved in successive cyclicity are in fact inherited from a higher C head, one possibility is that Welsh Aspect heads are simply unable to inherent wh-EPPs, whereas Irish Aspect heads possess this ability, allowing an object to pass through MinD(Asp) and thereby escape superiority by Equidistance. The main point here is that in principle, wh-movement through spec v*P can be motivated by RM considerations.

(3) IRISH:

a. *Cad$_i$ é a chreideann Seán a cheannaigh cé t$_i$?*
what aL believe John aL bought who
"what does John believe that who bought?"

b. *Cad$_i$ é a chreideann cé a cheannaigh Máire t$_i$?*
what aL believe who aL bought Mary
"What does who believe [that Mary bought t]?" (Maki and Baoill 2005: 10)

In the English examples, the intervening subjects are located in spec TP, meaning that if they carry a wh-feature they will induce Minimality violations as any passing wh-elements attempt to raise from a spec v*P to a spec CP position. In Irish, however, all intervening subjects remain in spec v*P (explaining the word order differences), with the result that these sentences are entirely acceptable. In such languages, moving an object into the same MinD(v*) as the subject is therefore a crucial intermediate step to ensuring eventual grammaticality by Equidistance. Note that Phase Theory can provide no inherent explanation for the above contrasts, and would arguably have to rely on RM to provide the requisite story.

Turning now to question 2 above, consider the fact that unlike English, some languages allow multiple wh-elements to be fronted to the edge of CP or TP, presumably passing through the intermediate spec v(*)P positions along the way. In fact Rudin (1988) identifies two types of such multiple wh-fronting languages; in the first [+MFS] (Multiply Filled Spec) type, all wh-elements move to spec CP; in the second [−MFS] type, however, only one of these elements moves to spec CP, the rest adjoining to TP. Rudin thus postulates the following two configurations:

$$\left[_{\text{CP}} \text{WH}_i \text{ WH}_j \text{ WH}_n \text{ ...} \left[_{\text{IP}} \text{ ... } t_i \text{ } t_j \text{ } t_n \text{ ... }\right]\right]$$
$$\left[_{\text{CP}} \text{WH}_i \left[_{\text{IP}} \text{ WH}_j \text{ WH}_n \left[_{\text{IP}} t_i \text{ } t_j \text{ } t_n \text{ ... }\right]\right]\right]$$

Some examples illustrating these two types of language are given below:

(4) BULGARIAN: [+MFS]

*koj (*e) kakvo (*e) na kogo e dal*
who what to whom HAS given

(5) SERBO-CROATIAN: [−MFS]

ko	*je*	*što*	*(*je)*	*kome*	*(*je)*	*dao*
Who	HAS	what		to whom		given

"Who gave what to whom?"   (Dayal 2003: 16)

Note that only in [−MFS] languages may any non-wh-constituents, such as adverbs or an auxiliary like 'have,' intervene between the first wh-element and the other fronted wh-elements. This follows if the first wh-element is in spec CP while the others are contained in the TP domain. Rudin also points out that only [−MFS] languages allow the wh-elements to be in any order, which would follow if all but one of these elements occupy adjoined rather than spec positions, adjunction being typically much freer with respect to word order than substitution. A further interesting difference between these two types of multiple wh-fronting language is that only the [+MFS] languages allow for wh-island violations, as can be seen from the following examples:

(6)   BULGARIAN:

   Vidjah   edna   kniga [$_{CP}$ **kojato$_i$**   *se čudja*   [$_{CP}$ **koj**   *znae*
   Saw-1s   a   book   which   wonder-1s   who   knows
   [$_{CP}$ *koj*   *prodva*   **t$_i$**]]
      who   sells

   "I saw a book which I wonder who knows who sells (it)."

(7)   SERBO-CROATIAN:

   *...   osoba,   koja sam   ti   rekao [$_{CP}$ *gde   (on)   živi*]
      individual   who have-1s   you   told   where   he   lives

   "... the individual who you asked me where (he) lives."

                                    (Rudin 1998: 457–459)

In 6 the relative element kojatoa 'which' has been extracted across two intervening spec CP positions filled with the wh-pronoun koj 'who'; this, claims Rudin, is allowed because Bulgarian is [+MFS]. By contrast, as illustrated in 7, such extraction is disallowed even across a single spec CP position in Serbo-Croatian, a [−MFS] language.

Both RM and PT are equally capable of accounting for these facts: multiple spec CP positions in the [+MFS] languages allow for more than one wh-element to move to the phase edge and hence escape being Spelt Out on the current cycle in PT, while such positions become MinD(C) escape hatches by Equidistance within an RM-based theory. Conversely, in the [−MFS] languages, such escape hatches as defined by either theory remain unavailable in the CP domain and wh-islands consequently remain inviolable.

So far then, such phenomena do not allow us to choose between PT and rich RM; but they do highlight the inherent redundancy with respect to successive cyclic wh-movement through spec CP (and spec v*P positions) between the two. Furthermore, arguably RM provides a natural account for why in a [−MFS] language like Serbo-Croatian, movement of a wh-element to spec CP past wh-elements in the TP domain does not give rise to superiority effects: it would seem that there are unlimited intermediate TP adjunction sites available, thus allowing for successive cyclic wh-movement through MinD(T) in such languages, and circumventing intervention by Equidistance. By contrast in English, where only a single spec TP position is available, such a strategy is unavailable, with the result that superiority effects are found in examples such as 1(a). Since TP is not itself a Phase, PT arguably cannot account for the English/Serbo-Croatian contrast in this respect, and would again have to rely on some sort of RM component.[10]

Having seen why successive cyclic wh-movement through both intermediate spec v*P and spec CP positions is very useful in languages other than English, we can now return to question 2 above: why is wh-movement compelled to proceed successive cyclically through spec of CP and *unaccusative* vP in English, even though, as we have seen, such precautions do not lead to a greater number of convergent derivations here? That is, successive cyclic movement through spec vP does not avoid ungrammaticality, since there will never be a potentially intervening base-generated wh-element in its specifier position anyway (given that unaccusative v does not assign an external theta role); and since spec CP cannot be host to more than one wh-specifier, such positions do not seem to be relevant as Minimal Domain 'escape hatches' in English.

Assuming that a certain degree of movement in syntax is indispensable if language is to be capable of expressing discourse and scopal, as well as propositional, meaning effectively, one can well imagine a tension arising between the (RM-P) demands of the parsing systems (which would ideally prefer no movement) on the one hand, and those of the I-language syntactico-semantics (which demand it) on the other. Interestingly, this inter-modular conflict appears to transmute into an intra-modular tension within the grammar itself: on the one hand we have something like the Minimal Link Condition (i.e. RM-C) which essentially boils down to the view that a probe K is lazy and will only search as far down the structure as it needs to, and on the other we have compensatory successive cyclic movement, which brings a potential goal α that is initially lower down the tree than another goal β of the same type,

---

10 An account in terms of Late Adjunction (see e.g. Stepanov 2001) to TP (i.e. after wh-movement to spec CP has already taken place) may be possible in PT, however.

into the same MinD(X) as β, thus making both α and β equally viable goals for K in structural terms. And note that although the notion of a Minimal Domain may at first appear as stipulative as Chomsky's Phases, I would argue this can, along with Equidistance, in fact be quite straightforwardly derived under what are arguably the most natural assumptions regarding the search algorithm initiated by K: namely, that it proceeds top down and *categorially cyclically*, automatically searching on each obligatorily full cycle *all and only* the labels and head of a given category X along with their immediate constituents, before initiating Move/Agree operations with *any* suitable goals encountered on that particular cycle and/or proceeding downwards into the complement of X should some or all of the required features not have been encountered[11]. For example, the initial cycle of the search algorithm for a probe K would follow the path 1-9 as follows:
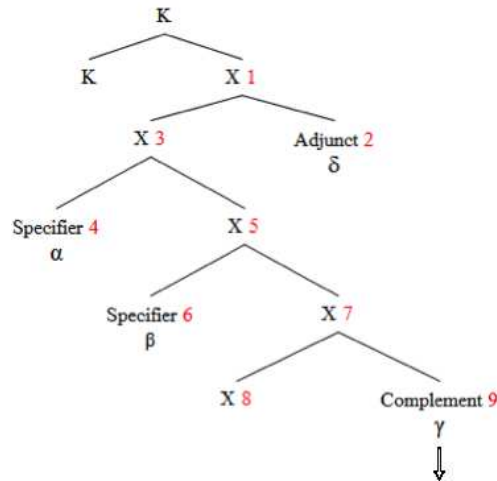
Since both α and β are encountered on the same cycle, both are viable goals (as are δ, γ, the head of X and perhaps any of X's projections *in principle*).

11 CED effects (Huang 1982) may also be derivable without reference to Chomsky's phase-based and somewhat arbitrary (2005b: 20) 'deep search' stipulation (which states that although constituents at the 'phase edge' are accessible to probes, their contents are not): if we assume that the search algorithm proceeds categorially cyclically, but always *downwards* through the tree, then clearly proceeding into complements, rather than backtracking to specifiers or adjuncts, would be the optimal route should none of the immediate constituents of the probed phrase possess all or any of the required features. This explanation of CED effects also accounts for why in general a constituent β must c-command the trace of another constituent α in order to act as an intervener to α; given strict binary branching (and perhaps additionally Pair Merge), if β does not c-command the trace of α, then β must be contained within a specifier or adjunct of the main tree, not within a complement; extracting β would therefore require non-optimal search backtracking.

However, whether or not circumvention of Minimality by Equidistance is actually possible in any given derivation in any given language will depend on the specific featural composition of the head X. If X is able to host multiple EPPs, i.e. multiple specifiers, as C is in Bulgarian, for instance, then MinD(X) will make available multiple escape hatches, allowing for e.g. wh-island violations; otherwise it will not. If, on the other hand, X is able to host a *single* inherited EPP, then this will be enough to open up a single MinD(X) escape hatch whereby a lower constituent can safely pass a constituent of the same type which has been base generated in spec XP, as we saw with the licit Irish violations of superiority.

Note, however, that on optimal assumptions, the C_{HL} cannot 'know' in advance whether or not the heads employed by the specific language and derivation in question are capable of hosting the requisite (multiple) EPPs, nor even whether there will be any interveners present in any given derivation; that is, the competence systems cannot be expected to 'react' in synchronically teleological fashion, i.e. to make use of 'look-ahead procedures,' even on a single cycle/movement operation[12]. Therefore, arguably the only way for UG to ensure that Equidistance is exploited wherever possible is to adopt a universal strategy whereby movement diacritics are blindly distributed across all suitable (i.e. compatible) heads *potentially* hosting an intervening specifier between a base-generated constituent and its eventual landing site; this could perhaps be implemented via something like Chomsky's (2005b) feature spreading inheritance device[13]. Assuming C_{HL} not to discriminate between v and v* heads as far as such feature spreading is concerned, the former will inherit wh-EPPs, despite never hosting base generated potential interveners in its spec.

Cross-linguistically ubiquitous, successive cyclic wh-movement, then, can be viewed as a non-teleological remedy to the effects of RM-C, which may itself have originated as an evolutionary adaptation to the constraints of RM-P, and/or perhaps as the result of third factor constraints that shaped both the competence and performance systems to some extent independently but

---

12 See Biberauer & Richards (2006) for discussion.

13 Though crucially this long-distance spreading would apply to *all* intermediate probe heads (perhaps only spreading EPP features), not just arbitrarily to those initiating successive cyclic A-movement as in Chomsky's system. Perhaps we could circumvent the countercyclic nature of such spreading by allowing it to take place in the numeration before lexical items are divided into their various subarrays. Note, incidentally, that the notion 'subarray' need not necessarily be associated with phases; in fact, with the early 'Merge over Move' arguments for phases in Chomsky (2000) now rendered irrelevant by Chomsky's (2005b: 7) claim that 'IM is as free as EM,' it would seem more natural to associate subarrays only with clausal heads (a clause being a full propositional unit), rather than arbitrarily defining them over v(*)P and CP. Indeed, retaining subarrays in this way allows us to account for the MOM data while also eliminating phases.

in similar directions. Assuming, furthermore, the parser to be to some extent parasitic on the grammar (perhaps on some level the two may even be considered synonymous – see Phillips (1995) and Mobbs (2011)) – it may be that the processing systems are aware of Equidistance and the (im)possibility of a given intermediate trace in a given language, and that this also helps to ameliorate the computational burden imposed by RM-P filler-gap ambiguities.

Thus, by liberally distributing movement diacritics wherever they are both compatible and potentially (but not necessarily) needed, the grammar appears to adopt something of a blind 'Sledgehammer' solution to the RM-C/P/movement clash problem, or in Rizzi's (2004: 2) terms, 'the opposition between locality and movement as Last Resort':

> **The Sledgehammer Solution**
> An inherent wh-attracting EPP feature on a C head will universally be copied onto all compatible lower C and v(*) heads intermediate between the initial extraction point and the final landing site.

A question which arises here is why the Sledgehammer Solution has not led to universal successive cyclic wh-movement through spec TP in order to try to circumvent the sorts of superiority effects seen in English. One answer to this is that it has: Boeckx et al. (2005: 8) claim that 'there is a growing consensus that movement proceeds through adjunction to each maximal projection along the way.' A language would then also need to host *multiple* specifiers/adjunctions in TP in order to circumvent superiority, of course. This explanation is highly appealing in that it would allow for an essentially identical treatment of successive cyclic A'- and A-movement, with many cross-linguistic parametric differences in locality reducible to the (in)ability of certain functional heads to host various inherited EPPs (in accordance with the Borer Chomsky Conjecture and Fukui's (1988) Functional Parameterisation Hypothesis—see also Boeckx 2008 for extensive arguments for XP movement through each maximal projection, though his account dispenses with EPPs). This would, however, create problems for the Irish/Welsh contrast in superiority effects and the potential solution suggested in note 9.

Another answer, which adheres to the traditional view that in general only A-movement proceeds through intermediate spec TP positions, while only A'-movement proceeds through spec CP positions, is that the grammar may have only evolved so as to spread EPPs to those positions which are most likely to host interveners to the relevant movement 'type' in question. Since spec v*P is an initial base-generated position for both Case and wh-feature bearing constituents, it is a very common potential intervening position for both A

and A'-movement. However, of spec TP and spec CP, only the former is an eventual landing site for A-movement, while only the latter is so for wh-movement. Given that the frequency with which wh-elements are found in spec TP positions is clearly relatively low (given that, as noted above, in many of the world's languages subjects appear to either remain in spec v*P or to move to an intermediate Aspect phrase rather than all the way to spec TP from where they can be wh-extracted and leave a wh-trace), therefore, FL may not have evolved so as to enforce successive cyclic wh-movement through spec TP. The likelihood of a Case-bearing argument appearing in spec CP, meanwhile, is arguably nil, owing to the A/A'-movement chain asymmetry. (i.e. the apparent ban on 'IMPROPER MOVEMENT' discussed in Chomsky (1973, 1981)) Therefore FL has even less motivation to enforce intermediate A-movement through spec CP. Note that this approach predicts that A-movement, as well as A'-movement, will proceed successive cyclically through all intermediate spec v(*)P positions, and in the next section we will see evidence that this is indeed the case.

One final question which needs answering before closing this section is how well Chomsky's PT handles the English superiority effects in 1(a). The first point to note is that Chomsky's (2001) PT is forced to rely on an (albeit impoverished) RM component in order to deal with such phenomena. Chomsky's culling of RM clearly indicates that he recognizes the inherent redundancy in the system here; but I maintain that his efforts towards theoretical parsimony are directed against the wrong component.

In Chomsky's system Equidistance is eliminated, traces do not induce intervention effects at all, and only the final movement in a chain counts for the purposes of Minimality, the latter now being computed only at the 'next higher strong phase' (Chomsky 2001: 28); given the PIC, if a constituent wishes to move on further, it will have to be on the edge of the current phase, and hence outside the remit of RM. For this reason, the only movement steps which will be evaluated for Minimality in 1(a) will be the ones by which the subject raises from inner spec v*P to spec TP, and where the object then raises from outer spec v*P to spec CP. As far as the former case is concerned, by the time the phase level is reached and Minimality assessed, the only wh-element intervening between the wh-subject's extraction and landing sites will be the wh-object trace in spec v*P, and since traces do not induce intervention in Chomsky's system, even with Equidistance gone this move will be correctly licensed. As for the (we have concluded) offending movement by which the wh-object moves to spec CP, crossing the wh-subject in spec TP, this will, as in the rich RM account, not be deemed licit, since the intervening subject is overt. So far, then, both Chomsky's impoverished version, and our 'rich'

version of RM are on equal footing empirically here.

However, the situation is somewhat different in the context of his 2005b paper. Here, Chomsky still takes only the head of a chain to be a potential intervener, but now dispenses (Chomsky 2005b: 9) with the 'next higher phase' stipulation (which was, it must be said, suspiciously arbitrary), so that Minimality will presumably be computed for the complement of each phase head as soon as it has been merged. However, he also suggests that A- and A'- movement 'proceed in parallel' (Chomsky 2005b: 14), and appears to regard this as effectively nullifying the ability of A-moved constituents to induce intervention effects in A'-moved elements and vice versa:

> If T were a phase head, or an independent probe for some other reason (as assumed in earlier work, mine in particular), then raising of subject to SPEC-T would be blocked by intervention of the ϕ-features of 'who' in the outer SPEC of v*. But since it is not a phase head, and both operations are driven by the phase head C in parallel, the problem does not arise.
> (Chomsky 2005b: 19)

On this revised account then, example 1(a) is predicted to be perfectly acceptable, contrary to fact. Indeed, Chomsky himself notes that 'there should be no superiority effect for multiple wh-phrases; any can be targeted for movement;' but, as he then observes, 'that leaves the problem of explaining the superiority phenomena in the languages in which they appear' (Chomsky (2005b: 18)). Clearly the fact that the rich RM theory predicts the superiority effects, while Chomsky's PT does not, offers some important evidence favouring the former.

4   Successive Cyclic A-movement

When it comes to successive cyclic A-movement, as in the following examples in 8, things appear to be even worse for phase-theoretic approaches:

(8)   a. There$_i$ seem [$_{\text{TP}}$ t_i to be believed [$_{\text{TP}}$ t_i to have been discussed several problems]]

        b. Several problems$_i$ seem [$_{\text{TP}}$ t_i to be believed [$_{\text{TP}}$ t_i to have been discussed t_i]]

Whereas RM can naturally motivate movement of arguments through multiple spec TP positions (since once again these are potential hosts to interveners), PT cannot since, as noted, TP is not a phase, and hence by the PIC constituents need not move through its edge.

Chomsky (2005b: 23) attempts to derive successive cyclic A-movement via spreading of the edge feature of matrix C to all lower Ts in the structure. Inheritance between C and the head of its TP complement certainly seems plausible enough in view of their shared tense/ϕ- properties. But, as formulated by Chomsky, this additional long-distance spreading down the tree *exclusively for A-movement* seems at best a cumbersome further stipulation which appears to lack any independent motivation. The advantage of an RM alternative here is that we potentially derive a natural and *unified* account for *all* successive cyclic movement.

The motivation discussed above for successive cyclic wh-movement through spec v(*)P carries over to the A-movement case: this is because spec v*P is both a base-generated wh- and A-position, hence a potential wh- *and* potential A-intervener. However, Sauerland (2003: 308) notes that 'it is widely assumed – often tacitly – that A-movement does not move through intermediate positions where it does not check morphological features.' Unlike A'-movement, in other words, A-movement is standardly assumed not to move through intermediate spec v(*)P positions, resulting in a rather awkward asymmetry.

Note that the exclusivity in English of *unaccusative* verbs in the intermediate clauses through which A-movement proceeds in no way constitutes evidence[14] that A-moved constituents cannot move through intermediate spec v(*)P positions, since even in sentences where no argument/trace appears in spec vP, an argument in spec TP induces Minimality violations anyway, as seen in the following super-raising examples:

(9) a. * Jack seems that *it* was expected to win the race
        [$_{TP}$ Jack$_i$ [$_{vP}$ seems [$_{CP}$ that it was [$_{vP}$ expected [$_{TP}$ t_i to [$_{v*P}$ t_i win the race]]]]]]

    b. ** Pete$_i$ seemed that Jack$_j$ was told t_j t_i to help us[15]
        [$_{TP}$ Pete$_i$ [$_{vP}$ seemed [$_{CP}$ that [$_{TP}$ Jack$_j$ was [$_{vP}$ told t_j [$_{CP}$ [$_{TP}$ t_i to [$_{v*P}$ t_i help us]]]]]]]]]
        With the meaning: "It seemed that Jack was told Pete would help us."

---

14 Nor does the fact that such unaccusative structures involve no intervening CP or v*P 'phases' necessarily constitute an argument for the PIC (see note 5).

15 The strong unacceptability of this example could be taken to constitute evidence that illicit extraction out of two CP phases has taken place. However, these effects are just as easily accounted for in an RM-based theory by observing that the extracted element 'Pete' incurs not one, but two Minimality violations, given that it must cross both the null copy of 'Jack' in spec VP, and its antecedent in the intermediate spec TP, without entering the Minimal Domains (MinD(V) and MinD(T)) containing either.

Hence A-movement across *transitive* clauses can also presumably be ruled out by the overt subject in spec TP, rather than by its trace in spec v*P: like T, v(*) could therefore in principle host a single inherited A-EPP in English.

A'-movement has often been described as the freest type of phrasal movement, owing to its ability (in certain contexts) to operate across 'unbounded' domains. However, it seems more likely that A- and A'-movement (keeping here to wh-movement) are in principle equally 'unbounded'; the only difference is that virtually all verbs must select at least one argument, whereas there are no verbs which obligatorily select for a certain number of wh-arguments (though some select for interrogative CPs). Therefore, A-movement, which involves arguments, is naturally expected to exhibit far more restrictive behaviour than wh-movement: A-interveners are simply far more prolific.

With these considerations in mind, I will attempt to eliminate this asymmetry between A'- and A-movement by arguing that in fact A-movement *does* proceed through all intervening spec v(*)P positions (in addition to TP positions as standardly assumed); the empirical evidence supporting this directly parallels that for wh-movement.

Consider first the following examples:

(10)   a. Every child$_i$ doesn't seem to his$_i$ father [t_i to be smart]
          With the interpretation: "it's not the case that for every child,
          the child seems to his father to be smart."

       b. [A boy]$_i$ doesn't seem to his$_i$ father t_i to be a loser
          With the interpretation: "no boy seems to his father to be a
          loser."   (Sauerland 2003: 310-311)

Sauerland (2003: 310-311) argues that in order for 10(a, b) to have the interpretations indicated, it is necessary for the matrix subject to reconstruct to a position *below* the negation ('every child' and 'a boy' having only narrow scope with respect to 'n't' on these interpretations), but above the EXPERI-ENCER object 'his father' in spec VP, into which it binds. If we assume that the intermediate position is the spec of the v occupied by the unaccusative verb 'seem,' these facts can be straightforwardly accounted for.

Further evidence comes from French past participle agreement. It has often been argued (e.g. Kayne 1989, Sportiche 1998, Richards 1997) that the fact that only *preposed* wh-objects trigger overt agreement marking on participial verbs taking *avoir* 'have' as the perfect auxiliary indicates that the wh-object must have moved to spec CP via the spec v*P position:

(11)   FRENCH:

a. *quelle*   *bêtise*   [$_{TP}$ *a-t-il$_i$* [$_{v*P}$ *t$_i$*
   which-FEM   blunder-FEM   has he
   commise   *t$_i$?*]]
   committed-FEM?

b. [$_{TP}$ *Il$_i$* *a*   [$_{v*P}$ *t$_i$* *commis*   *t$_i$?* *quelle*
      He   has   committed-MASC   which-FEM
   bêtise   ]]
   blunder-FEM? (Radford   2004:   403)

By contrast, those verbs taking *être* 'be' as the perfect auxiliary agree with their subjects. Yet given that such verbs are unaccusative, their surface subjects must be considered their logical objects. That being so, it seems reasonable to conclude that the agreement features on these participles are similarly the result of intermediate movement – only this time A-movement – of a base generated object to the spec vP position, before moving on to spec TP in order to pick up Nominative Case:

(12) FRENCH:
   a. [$_{TP}$ *Elle$_i$* *est* [$_{vP}$ *t$_i$* *morte*   *t$_i$*]]
         She   has   died-FEM

   b. [$_{TP}$ *Il$_i$* *est* [$_{vP}$ *t$_i$* *mort*   *t$_i$*]]
         he   has   died-MASC

There thus appears to be evidence from both reconstruction and morphology which supports the view that A-movement proceeds successive cyclically through intermediate spec vP positions. Interestingly, the same arguments which I presented in order to *motivate* successive cyclic A'-movement through intermediate spec CP and v(*)P positions appear to exist for A-movement through spec v(*)P and TP. What seems to dictate the specific intermediate positions through which A'- and A-movement will proceed is simply the categorial nature of the initial extraction points and prototypical final landing sites, since these are the sites most likely to host interveners of the relevant type. For both wh- and Case-feature bearing DPs, spec v*P is universally an initial extraction point, (and final landing site for A-movement[16] in languages with OS); but only in A'-movement is the landing site typically spec

16 As is spec VP, assuming a subject to object raising analysis of ECMs (Chomsky 2009) and object control constructions (Hornstein et al. 2010). It may be, then, that wh-movement also proceeds through intermediate spec VP positions in languages which allow both clause mate and non-clause mate superiority violations.

CP, while, conversely, only A-movement targets spec TP. Spec CP and TP are therefore equivalent intervening 'positions' in the A'- and A-movement cases respectively, as illustrated by the super-raising examples discussed in 9 above, which directly mirror familiar wh-island violations in English such as 13 below:

(13)   * When$_i$ did you wonder [$_{CP}$ who [Jack invited to the ball t_i]]?

If A-movement operates essentially identically to A'-movement, then in order to circumvent the sorts of Minimality violations seen in the super-raising examples in 9, we predict that a language will have to permit multiple A-specifiers in both v(*)P and TP (just as languages like Romanian, which violate wh-islands, turn out to feature multiple wh-fronting), thereby bringing Equidistance once again into play.

There certainly seems to be some plausible evidence for the existence of such Multiple Subject Constructions, though admittedly this is not uncontroversial. Consider the following examples from Japanese and Mandarin Chinese:

(14)   a. JAPANESE:
       Syusyoo-ga(-no)*   *saikin*   *byooki-ga*   *omo-i*
       Prime Minister-NOM(*-GEN)   recently   illness-NOM   serious

       "The Prime Minister is seriously ill."   (Nakamura 2010: 358)

       b. MANDARIN:
       xiang   *bizi*   *chang*
       elephant   nose   long

       "Lit. Elephants, noses are long."   (Ura 1994: 32-40 citing Teng 1974 and Li and Thompson 1981; cited in Zwart 1997: 16)

In 14(a) both 'subjects' bear Nominative Case, suggesting that they have been licensed by the same head; note too that when both Nominative, they can be separated by an adverb, indicating that they do not together form a constituent. However, this contrasts with a situation where the 'possessor' noun is Genitive, in which case separating the elements in this way yields an ungrammatical sentence, as indicated. In example 14(b), the two DPs are again both analyzed as subjects by Ura (1994: 49) in an Agr-based system making use of both the spec AgrSP and spec TP positions. However, from example 14(a), in which both DPs bear Nominative Case, a reasonable hypothesis in the context of an Agr-less system employing v(*)P shells would be that in such MSC constructions the two subjects have passed through the same projection, namely TP; if this analysis is correct, (i.e. if T can here bear

multiple A-EPPs, unlike in English) then such languages are predicted to license the sort of super-raising examples which are barred in English, MinD(T) now able to act as an escape hatch by Equidistance.

Ura maintains that this is precisely what we find, and formulates the following typological generalization:

> If a language allows the so-called "Multiple Subject Construction," then it also allows super-raising to take place.
> (Ura 1994: 5; cited in Zwart 1997: 19)

An example of such a licit super-raising construction in Mandarin Chinese appears in 15(a) below where, in contrast to 15(b), the logical object raises across a transitive verb 'reng' and its thematic subject to the matrix subject position:

(15) MANDARIN
    a. Super-raising:
      ta keneng Zhangsan reng le   nei kuai rou
      he possible Zhangsan throw ASP that piece meat
      "Lit. He$_i$ is possible that Zhangsan tossed him$_i$ that piece of meat."

    b. Non-super-raising variant:
      keneng Zhangsan reng le   nei kuai rou gei la
      possible Zhangsan throw ASP that piece meat to he
      "It is possible that Zhangsan tossed that piece of meat to him."

    (Ura 1994: 10 citing Li 1990 and Shi 1990; cited in Zwart 1997)

Assuming the Sledgehammer Solution now to apply to all XP movement (though of course A-EPPs will spread to all lower Ts rather than Cs as in the A' case), nothing prevents the lower v* in 15(a) from having inherited a Case-attracting (i.e. tense-associated[17]) EPP in the numeration. This being so, the logical object will crucially move first to spec v*P, where it will be in the same MinD as the subject, i.e. MinD(v*), thus ensuring that the move is licit; the subsequent movement of 'he' from the embedded outer spec v*P position to the matrix spec TP (via matrix spec predP), will also be problematic unless we allow this pronoun to move first to the outer specifier position of the embedded TP, where it will enter the same MinD(T) as the subject (moved by this point to inner spec TP); otherwise the latter will induce an intervention effect as the

---

17 Assuming, with Pesetsky and Torrego (2001), that finite T values a Nominative argument's Case feature as [+finite] (i.e. Nominative = Finite).

logical object crosses it on its way to the upper clause. The derivation I am therefore proposing for 15(a) is as follows:

(15) a.
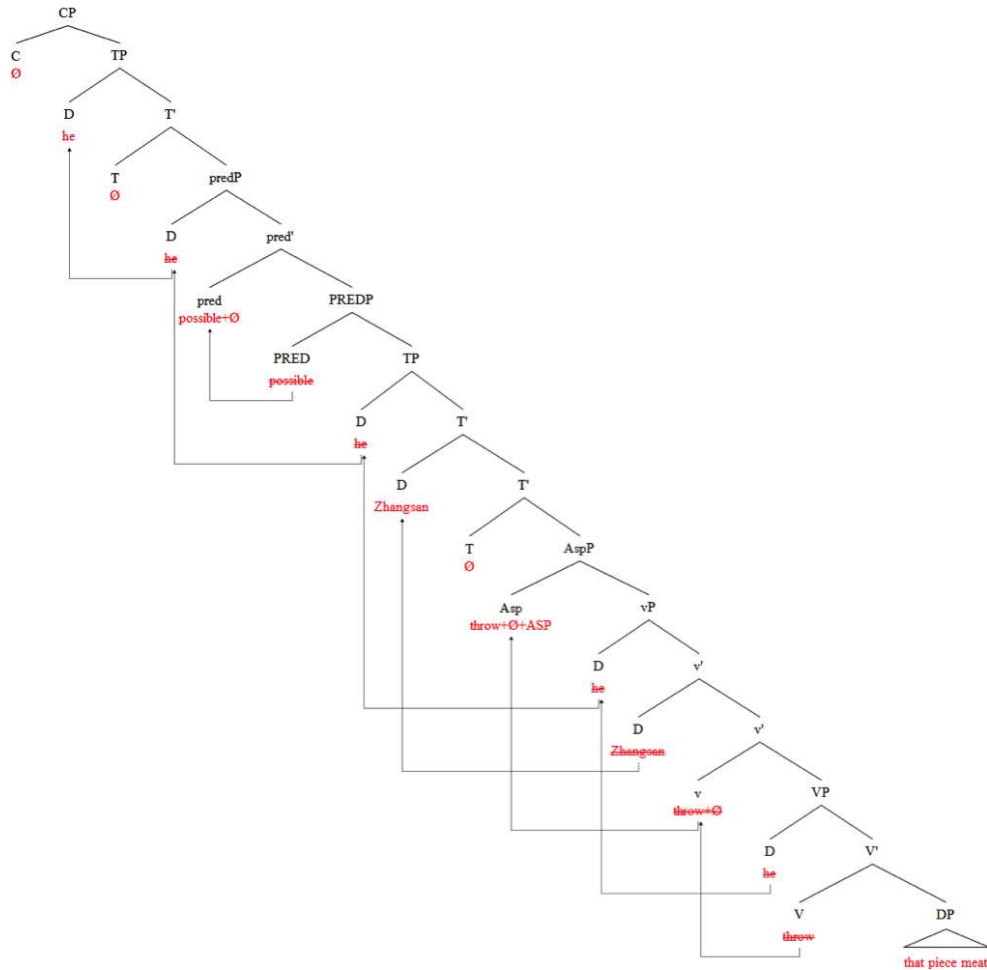
However, while Mandarin Chinese seems to permit super-raising, it turns out that Japanese apparently does not (Takae Tsujioka 2002: 48). This may be either because multiple Japanese Nominative subjects are not, in fact, licensed by the same head, Nominative Case simply being a default Case in Japanese and therefore not necessarily assigned by Infl as argued for by Saito (1983, 1985). Alternatively, it may turn out that although the ability of TP and v(*)P to host multiple specifiers is a necessary condition to allowing cases of super-raising, it is not, by itself, sufficient; Japanese might simply lack the ability to host *multiple* inherited[18], rather than inherent, A-EPPs in spec v(*)P/TP, for instance, such that its arguments would be unable to e.g. pass through intermediate spec TP positions on their way to a higher TP – they

---

18 i.e. those which have spread from a higher head.

can move there only for Case checking purposes. This suggests that perhaps we should invert and weaken Ura's generalization to the following:

> If a language allows super-raising, then there is a good chance it will allow Multiple Subject Constructions.

In any case, the Mandarin example in 15 (and Ura 1994 provides further examples from other languages) provides important further empirical evidence as well as critical motivation for successive cyclic movement through spec TP and spec of v*P, though crucially not owing to an inherent Case-EPP on v*, since the matrix subject ends up with Nominative.

On this view, then, with respect to locality constraints, the only parametric difference between a language like English, and languages like Romanian and Mandarin Chinese, is that these latter two languages permit multiple inherited/inherent A'- and A-EPPs respectively on v(*), and C and T (again respectively), whereas English only allows singleton EPPs (either inherent or inherited[19]) in either case; this is what ultimately forces English to make do with argumentally 'empty'[20] unaccusative clauses as the only permissible intermediate structures in successive cyclic A-movement (explaining why such examples are always a little awkward): escape hatches are simply not available in argument rich clauses owing to the paucity of EPPs on T heads.

Note that, crucially, all of the above argumentation relies on an enriched version of RM which reincorporates the notions of Equidistance and Minimal Domain. Phase Theory, as discussed above, is unable to provide any intrinsic account of successive cyclic A-movement, even through multiple TPs, since these are not phases. And even if we were to accept Chomsky's feature spreading mechanism as descriptively adequate here, by itself it certainly does not motivate/explain the phenomenon of successive cyclic A-movement (it merely restates it), nor does it offer the prospect of a unified account of all successive cyclic XP movement, the A' case being dealt with entirely separately by the PIC.
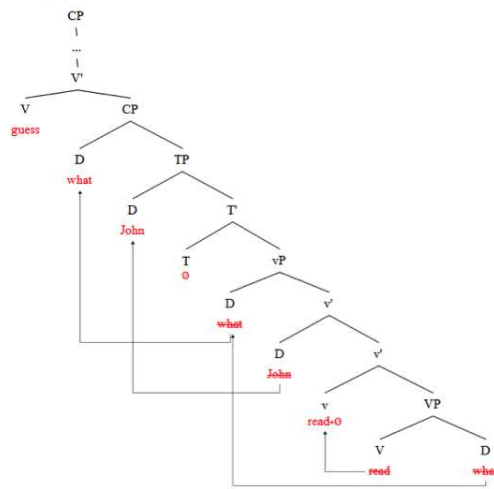
Nevertheless, Chomsky (2001: 26) proposes abandoning Equidistance. He discusses the following two examples, which he claims both involve object shift (OS):

(16)   a. (guess) what$_{\text{obj}}$ [John$_{\text{subj}}$ T [$_{\text{v*P}}$ $t_{\text{wh-obj}}$ [t_{subj} read $t_{\text{wh-obj}}$]]]

b. *John$_{\text{subj}}$ T [$_{\text{v*P}}$ that$_{\text{obj}}$ [t_{subj} read t_{obj}]]

(Chomsky 2001: 26)
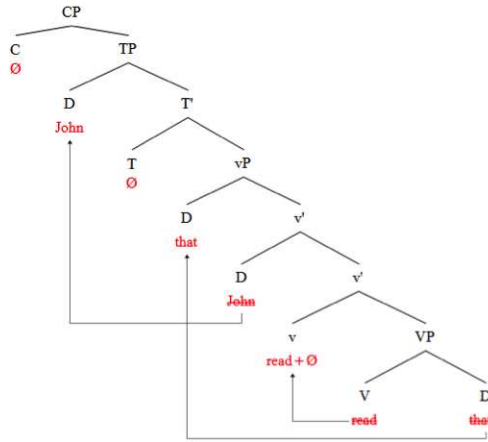
---

19 Only inherited for v(*)

20 Empty of arguments c-commanding the lower trace, more specifically.

16(a)   16(b)

Chomsky argues that in English 'OS' is only allowed if the shifted object subsequently vacates the 'phonological edge' of v*P, which allows the subject to raise to spec TP as in 16(a), without intervention. He attributes the ungrammaticality of 16(b) in English to the fact that the overt object in spec v*P induces a Minimality effect when the subject is raised to spec-TP, despite the fact that both the subject and shifted object are located within MinD(v*) at the point when the subject is moved. Given that Equidistance apparently fails to apply here, Chomsky (2001: 26) proposes that it be abandoned and that instead 'inactive traces' (i.e. all traces) should be exempted from inducing Minimality effects.

In order to avoid the necessity of counter-cyclic operations – given the problematic fact for his proposal that in 16(a) the subject raises to spec TP before the object has vacated the outer edge of v*P – Chomsky stipulates that Minimality only be assessed at the 'next higher phase level' (Chomsky 2001: 28), at which point this position will indeed be empty. And even in the 2005b paper, where the 'next higher phase' condition is dropped (Chomsky 2005b: 9), intervention effects are only considered once the level of the current phase head is reached, and again only (what are at this point) overt constituents may be interveners.

However, Chomsky's proposal to eliminate Equidistance from the theory rests on two claims: 1. that what is wrong with sentence 16(b) is a Minimality

violation; and 2. that sentence 16(a) is licit precisely because the object has raised further, vacating the 'phonological edge' (a dubious concept, I would suggest, insofar as it impacts on narrow syntactic computations). Yet both of these points seem to me far from clear.

In 16(a) there is only reason to postulate intermediate A'-movement, not A-movement, of 'what' via spec v*P before moving on to spec CP; therefore only its lowest trace in comp VP will bear the all-important Case feature needed to invoke A-intervention anyway[21]. Yet Chomsky claims that no Minimality effect is induced here because the subject (or rather its trace, once the phase level is reached and Minimality is assessed) is at the all-important and accessible phonological edge of the phase, the phonetically null wh-object trace in outer spec v*P (rather dubiously) not 'counting' for intervention purposes as noted.

Given that the subject and object are, prior to the subject raising, both within MinD(v*), in an Equidistance-based theory we would expect no Minimality violation in 16(a) anyway – even if the wh-object in outer spec v*P had Case; nor should we expect Minimality effects when the object subsequently raises to spec-CP, crossing the subject in spec TP. The reason for this is that the constituent located in spec TP does not carry a wh-feature (the uninterpretable featural type I assume to be driving this movement operation); therefore (vacuously) in accordance with (rather than despite) Equidistance, the move is licit; note too that changing 'John' to 'who' in 16(a) (so that the subject does now have a wh-feature) derives an ungrammatical sentence, just as a strongly re-relativized version of RM tuned to the uninterpretable features of goals predicts.

Now consider the all important 16(b); all we need say here is that in English v* does not inherently contain an EPP feature inducing A-movement (though it may inherit a single EPP for the purposes of successive cyclic A-movement, as discussed); hence, in the case of objects, simple in situ Agree must take place without concomitant movement when Accusative Case-/ϕ-feature valuation is at stake, though subsequent wh-movement through spec v*P, as seen in 16(a), is of course permitted. 16(b) is therefore not ungrammatical because the subject has crossed the shifted object (Equidistance is therefore saved), but simply because English does not make available the inherent A-EPP on v* which leads to the availability of true OS. And yet, in this example, the object appears precisely in the unlicensed preverbal position, and the sentence is accordingly ungrammatical.

---

21 I follow the GB wisdom that only the lowest wh-traces (along with 'wh-in situ' elements, including wh-subjects in spec TP) bear Case features.

The point of all this is that Chomsky's argument here (the only one offered for the elimination of Equidistance to my knowledge) hinges on the assumption of a weakly relativized version of RM, in which a non-Case bearing wh-object in spec v*P might be expected to induce intervention on a non-wh A-moved subject, were the wh-object to remain in this position. The proposal being put forward here, however, is that RM should be strongly re-relativized,[22] so that Case-features must be present to induce intervention effects on A-movement (and wh-features on wh-movement). It is curious that Chomsky chooses to make his case against such a weak version of RM here, given that his 1995a version of the MLC enforced strict featural identity between target and intervener:

> **Minimal Link Condition**
> K attracts α only if there is no β, β closer to K than α, such
> that K attracts β
> (Chomsky 1995: 311)

Since T does not probe for wh-features, but for ϕ- and Case features, the Caseless wh-object in outer spec v*P is never a potential attractee for T, and hence is not a potential A-intervener; ϕ-features are necessary for intervention here, but not sufficient.

Chomsky's arguments against Equidistance are therefore less than convincing; but if we retain the rich version of RM, we are left with considerable overlap between this component and PT as we have seen. In fact, Chomsky (2005: 9-10) himself appears at one point to acknowledge the redundancy between phases and RM; he concedes that the PIC may not, in the end, apply in the narrow syntax at all, given that in almost all cases even his impoverished version of RM ensures that search cannot proceed into a lower phase:

> Note that for narrow syntax, probe into an earlier phase will
> almost always be blocked by intervention effects..It may be,
> then, that PIC holds only for the mappings to the interface,
> with the effects for narrow syntax automatic.
> (Chomsky 2005: 10)

But, of course, if the PIC fails to apply in the narrow syntax, then the crucial arguments for phases regarding the reduction of narrow syntactic computational burden are lost, even for the transitive (non-Binding/Control) structures

---

22 *Pace* Rizzi's (2002) arguments against strict featural identity as a definition of 'sameness' in RM, for which I believe there to be adequate solutions, though these lie beyond the scope of this paper.

for which they were arguably even formulable: search spaces between probes and goals are now regulated once again solely by RM.

5   Conclusion

Three components (BT, RM, PT) regulating locality seems suspiciously redundant from a Minimalist perspective. We have seen that, unlike phases, the theory of RM can provide a unified account of both successive cyclic A'-movement and A-movement, and can potentially be extended to accommodate all forms of construal under Hornstein's (2001, 2009) theory. Chomsky's (2001) arguments for the elimination of Equidistance from the theory, meanwhile, are fleeting, and directed only against a weakly relativized theory in any case. RM also reduces computational complexity considerably, and can be grounded in processing and 'third factors'; by contrast, phases, like the barriers and rigid bounding nodes of which they are clearly to some extent reincarnations, are stipulations, and appear to be computationally non-optimal in a number of fatal respects. It would therefore seem Minimalistically expedient to eliminate phases altogether, and to readopt instead the richer Equidistance-based and strongly relativized account of RM as a fully unified locality component.

References

Aboh, E. 2010. "Information Structure Begins with the Numeration." *Iberia: An International Journal of Theoretical Linguistics*, 2(**1**): 12-42

Adams, Jack. A. 1987. "Historical review and appraisal of research on the learning, retention and transfer of human motor skills." *Psychological Bulletin*, 101: 41-74.

Biberauer, T., Richards, M. 2006. "True Optionality: when the grammar doesn't mind." In Boeckx, C. (ed.) *Minimalist Essays*. Amsterdam: John Benjamins, 35-67

Biberauer, T., Roberts, I. 2009. "Cascading Parameter Changes: internally driven change in Middle and Early Modern English." In Eythórsson, Th. & J.T. Faarlund (eds.). *Grammatical Change and Linguistic Theory: the Rosendal Papers.* Amsterdam: John Benjamins.

Boeckx, C., Grohmann, K. 2007. "Putting Phases in Perspective." *Syntax*, 10(**2**): 204-222

Boeckx, C, Hornstein N, Nunes, J. 2010. *Control as Movement.* Cambridge: CUP.

Campbell, Jamie I. D. 1991. "Conditions of error priming in number-fact retrieval." *Memory and Cognition*, 19: 197-209.

Carnie, A. 2010. *Constituent Structure.* Oxford: OUP.

Carstens, V. 2003. "Rethinking Complementizer Agreement." *Linguistic Inquiry.* 34(**3**)

Chandler, Christopher C. 1991. "How memory for an event is influenced by related events: Interference in modified recognition tests." *Journal of Eperimental Psychology: Learning, Memory, and Cognition*, 17: 115-125.

Chomsky, N. 1957. *Syntactic Structures.* The Hague: Mouton.

Chomsky, N. 1965. *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, N. 1973. "Conditions on Transformations." In S.R. Anderson and P. Kiparsky (eds.), *A Festschrift for Morrris Halle*, New York: Holt, Rinehart, and Winston, 232-86

Chomsky, N. 1981. *Lectures on Government and Binding.* Berlin: Mouton de Gruyter

Chomsky, N. 1986. *Barriers.* Cambridge, MA: MIT Press.

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

Chomsky, N. 1995b. "Bare Phrase Structure," in G. Webelhuth (ed.), *Government and binding theory and the Minimalist Program.* Oxford: Blackwell, 383-440.

Chomsky, N. 1998/2000. "Minimalist Inquiries: The Framework," in R. Martin, D. Michaels and J. Uriagereka (eds.), *Step by Step.* Cambridge, MA: MIT Press, 91-155.

Chomsky, N. 2001. "Derivation by Phase" In M. Kenstowicz (ed.), *Ken Hale: A life in language*, Cambridge, MA: MIT Press, 1-52.

Chomsky, N. 2004. "Beyond Explanatory Adequacy." In A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures.* Oxford: OUP, 104-131.

Chomsky, N. 2005a. "Three Factors in Language Design." *Linguistic Inquiry*, 36: 1-22.

Chomsky, N. 2005b. "On Phases" In R. Freidin, C.P. Otero & M.-L. Zubizaretta (eds.), *Foundational issues in linguistic theory.* Cambridge, MA: MIT Press.

Chomsky, N. 2007. "Approaching UG from below" In Uli Sauerland and Hans Martin Gartner (eds.) *Interfaces + Recursion = Language?* New York: Mouton de Gruyter, 1-29.

Chomsky, N. 2009. "On Language and Cognition." In Ozsoy, S., Nakipoğlu, M. (eds.), *Linguistics Edition 73*, Lincom Academic Publishers

Chung, Sandra and James McCloskey. 1987. "Government, barriers, and small clauses in modern Irish." *Linguistics Inquiry*, 18: 173-237.

Dailey-McCartney, Anna, Victor Eskenazi, Chia-Hui Huang. 2002. "Response to Ura (1994), Varieties of Raising and the Feature-Based Bare Phrase Structure Theory." *Working Papers of the Linguistics Circle*, 15.

Dayal, V. 2003. "Multiple Wh-Questions" in M. Everaert and H. van Riemsdijk (eds.), *The Blackwell Companion to Syntax.* Oxford: Blackwell.

Epstein, S. D., Seely, T. D. 2002. "Rule Applications as Cycles in a Level-Free Syntax" in Epstein, S. D., Seely, T. D. (eds.), *Derivation and Explanation in the Minimalist Program.* Malden, MA.: Blackwell.

Frampton, J., Gutmann, S. 2002. "Crash Proof Syntax." In Epstein, S. D., Seely, T. D. (eds), *Derivation and Explanation in the Minimalist Program.* Malden, MA: Blackwell.

Haegeman, L. M.V. 1994. *Introduction to Government and Binding Theory.* Oxford: Blackwell.

Halle, M. 1995. "Feature Geometry and Feature Spreading." *Linguistic Inquiry*, 26: 1-46.

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

Hornstein, N., Nunes, J., and Grohmann, K,. 2005. *Understanding Minimalism.* Cambridge: CUP.

Hornstein, N. 2009. *A Theory of Syntax.* New York: CUP.

Hornstein, N., Boeckx, C., Nunes, J. 2010. *Control as Movement.* Cambridge: CUP.

Huang, J. 1982. *Logical Relations in Chinese and the Theory of Grammar.* PhD dissertation, MIT.

Kayne, R. S. 1989. "Facets of Romance past participle agreement." In P. Benincà (ed.), *Dialect Variation and the Theory of Grammar*, Dordrecht: Foris, 85-103.

Kayne, R. S. 1994. *The Antisymmetry of Syntax.* Cambridge, MA: MIT Press.

Kluender, Robert. 2004. "Are subject islands subject to a processing account?" In Vineeta Chand (ed.), *Proceedings of 23rd West Coast Conference on Formal Linguistics.* Somerville, MA: Cascadilla Press, 475-499

Koopman, H, Sportiche, D. 1991. "The Position of Subjects." *Lingua*, 85: 211-258.

Larson, R. 1988. "On the Double Object Construction." *Linguistic Inquiry*, 19, 335-391.

Legate, J.A. 2003 "Some Interface Properties of The Phase." *Linguistic Inquiry*, 34(**3**): 506-516.

Li, C, N., Thompson, S. 1981. *Mandarin Chinese: a Functional Reference Grammar.* Berkeley and Los Angeles, California: University of California Press.

Li, Y-H Audrey. 1990. *Order and Constituency in Mandarin Chinese.* Dordrecht: Kluwer Academic Publishers.

Maki, H., Baoill, D. 2005 "Two Notes on Wh-Movement in Modern Irish: Subject/Object Asymmetries and Superiority." *Gengo Kenkyu*, 128: 1-31

Miyagawa, S. 1993. "LF Case-checking and minimal link condition." *Case and Agreement II, MITWPL*, 19: 213-254.

Mobbs, I. 2011. "Minimalism and the parser. Part I." Syntax and Biolinguistics Cluster, University of Cabridge. Available at: http://www.academia.edu/681927/Minimalism_and_the_Parser_Part_I

Ortega-Santos, I. 2011. "On Relativized Minimality, memory and cue-based parsing." *Iberia*, 3(*1*): 35-64.

Pesetsky, D. and Torrego, E. 2001 "T-to-C movement: causes and consequences.' In M. Kenstowicz (ed.), *Ken Hale: a Life in Language.* Cambridge: MA: MIT Press, 355-426

Phillips, C. 1995. *Order and Structure.* Ph.D. thesis, MIT.

Pritchett, Bradley. 1991. "Subjacency in a principle-based parser." In Robert Berwick, Steve Abney & Carol Tenny (eds.), *Principle-based parsing: computation and psycholinguistics.* Dordrecht: Kluwer, 301-345.

Radford, A. 1994. *Minimalist Syntax.* Cambridge: CUP.

Richards, N. 1997. *What moves where when in which language?* PhD thesis, MIT.

Rizzi, L. 1990. *Relativized Minimality.* Cambridge, MA: MIT Press.

Rizzi, L. 2002. "Locality and Left Periphery." In A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures, Vol.2.* Oxford: OUP.

Rizzi, L. 2004. "On the Form of Chains." In Adriana Belletti (ed.), *Structures and beyond: The cartography of syntactic structures, Vol. 3.* New York: OUP, 223-251.

Rudin, C. 1988. "On Multiple Questions and Multiple Wh-Fronting." *Natural Language and Linguistic Theory*, 6: 445-501.

Sabel, J. 2002. "Intermediate traces, reconstruction, and locality effects." In A. Alexiadou (ed.), *Theoretical Approaches to Universals.* Amsterdam: John Benjamins, 259-313.

Sauerland, U. 2003. "Adjunction with A-movement." *Linguistic Inquiry*, 34(**2**): 308-314.

Shi, Dingxu. 1990. "Is there Object-to-Subject Raising in Chinese?" *Proceedings of BLS 16*, 305-314.

Sportiche, D. 1998. "Movement, agreement and case." In *Partitions and Atoms of Clause Structure.* London: Routledge, 88-243.

Stepanov, A. 2001. "Late Adjunction and Minimalist Phrase Structure." *Syntax*, 4(**2**): 94-125.

Takae, T. 2002. *The syntax of possession in Japanese.* Outstanding Dissertations in Linguistics. London: Routledge.

Teng, Shou-shin. 1974. "Double Nominative in Chinese." *Language.* 50: 455-473.

Tsujioka, T. 2002. *The Syntax of Possession in Japanese.* New York: Routledge.

Ura, Hiroyuki. 1994. "Varieties of Raising and the Feature Based Bare Phrase Structure Theory." *MIT Occasional Papers in Linguistics*, 7.

Van Dyke, Julie A. & Richard L. Lewis. 2003. "Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities." *Journal of Memory and Language*, 55: 157-166.

Zwart, J.-W. 1997 "Transitive Expletive Constructions and the Evidence Supporting the Multiple Specifier Hypothesis." in W. Abraham, E. van Gelderen (eds.), *German: syntactic problems – problematic syntax.* Tubingen: Niemeyer, 105-134

John Torr
Department of Theoretical & Applied Linguistics
Faculty of Modern and Medieval Languages
Sidgwick Avenue
Cambridge, CB3 9DA

jpt40@cam.ac.uk