Abstract: Problems of Projection: Extensions

I assume here the general minimalist research program, which I think is well-motivated on grounds of learnability, explanatory success, and the very limited information on origin of the human language faculty. In particular, I assume the conclusions of Chomsky (2013), including the abandonment of the endocentricity stipulation of X-bar theory and its descendants, and the separation of projection (labeling) from the principles of construction of expressions. A labeling algorithm, keeping to minimal search, assigns labels to expressions {X,Y} constructed by iterated Merge (external EM or internal IM); labeling yields no new category If one of X, Y is a head, labeling is trivial: minimal search yields the head as a label. If neither is a head, labeling is possible only if search of X and Y yields agreeing heads, meaning that if one or the other was raised it is now in its *criterial position* in Luigi Rizzi's sense. Assume further that at the CI interface, and for the rules of externalization, syntactic objects must be labeled. It follows that IM is successive-cyclic leading to a criterial position, and is forced to ensure labeling.

A further question is what Rizzi calls "the halting problem": why is there no further movement from a criterial position? A simple solution is outlined here that keeps to the minimalist assumptions just sketched. Further questions arise about special properties of subjects of CP: the Extended Projection Principle (EPP) and the Empty Category Principle (ECP). These can be unified under the labeling theory assumed, and the analysis can be extended to the second phase v*P, where the object that is raised – in accordance with the raising-to-object analysis of Lasnik and Saito (1991), tracing back to work of Paul Postal's – is in a structural position analogous to subjects of CP. ECP is violated for v*P, and sometimes for CP (escape from the "that-trace filter"). The reason for these apparent violations is the same, under the analysis presented, which also entails a revision of standard approaches to head-raising and sharpening of notions of phase-based memory. Work tracing back to Rizzi (1982) has shown that null-subject languages apparently differ in these properties, the parametric difference relating to "rich agreement." The basic distinctions also fall into place under the presented analysis. Several other anomalies of earlier proposals are also discussed and overcome.

I will assume here the general framework of Chomsky (2013), henceforth POP, which carried forward a reanalysis of the basic properties of language that have been a focus of inquiry since the early days of generative grammar, of course with roots tracing far back in history.[1]

To summarize briefly, work in generative grammar, since its origins, has focused on I-language (previously called "grammar," in one of the senses of this ambiguous term), considered to be an internal property of an individual, grounded in a basic principle BP: each I-language can be regarded as a computational procedure that generates an unbounded array of hierarchically structured expressions, each with an interpretation at the Sensorymotor (SM) and Conceptual-intentional (CI) interface – in the former case, usually sound, but other modalities are possible as

[1] I am particularly indebted to Riny Huijbregts and Hisa Kitahara for comments on earlier versions of this paper, and many very valuable discussions.

well. One goal of linguistic theory (UG) is to determine the fixed properties of BP and the options of variation. Naturally, one seeks the simplest account of UG. One reason is just normal science: it has long been understood that simplicity of theory is essentially the same as depth of explanation. But for language there is an extra reason: UG is the theory of the biological endowment of the language faculty, and each complication of UG therefore poses a barrier to some eventual account of the origin of language, to the extent that this can be attained.[2]

In the best case, phenomena would be explained by interaction of the simplest computational operation – Merge, with its two logically possible subcases, Internal Merge IM (automatically yielding "the copy theory of movement") and External Merge EM -- interacting with general principles of minimal computation MC. The *Strong Minimalist Thesis SMT* articulates this goal.

There are a number of reasons to suspect that the thesis is a reasonable one. Almost all of the properties of language, beyond the most superficial, are acquired with little or no evidence; and language seems to be largely dissociated from other cognitive functions, and without significant animal analogues. Furthermore, though evidence about origin of language is slight, it does seem to indicate that the language faculty emerged recently in evolutionary time, and there is overwhelming evidence that it has not evolved in any significant way (if at all) since humans left Africa, probably not long after it emerged. If so, then it is likely that a version of SMT holds in a fundamental way.[3]

SMT would gain significant support if it can be shown that pursuit of the thesis solves problems that have regularly come to light as efforts to construct generative grammars have been undertaken; even more so, if it yields conclusions that bear on more far-reaching questions about the nature of language. That seems to be the case. It may be useful to clarify some of the major results, which tend to be lost in the intriguing details of execution.

One result is an explanation for the strange and ubiquitous phenomenon of displacement. Since the early 1950s, it has always seemed to me – and still does – that the ubiquity of the phenomenon and its central role in syntax and semantics are a sufficient reason to rule out all versions of Phrase Structure Grammar (PSG) as candidates for theories of human language, even unrestricted rewriting systems with universal weak generative capacity. Until recently, displacement had always been assumed to be a curious imperfection of language, an odd complication that is never introduced into invented symbolic systems. But under SMT, it follows that displacement under IM is the simplest case, and that it would be a curiosity if it were not found. Correspondingly, any stipulation that blocks it carries a heavy empirical burden, and

[2] There is a burgeoning literature on "evolution of language," of doubtful quality. For discussion, see Hauser et al. (2014).

[3] For discussion, see Chomsky (2010), POP, Berwick and Chomsky (2011). It is commonly argued that this approach cannot be correct because of the complexity of language (e.g., Corballis, 2011), but the criticism is irrelevant to the proposals to which critics refer, which crucially distinguish core systems from externalization: the former arguably close to (or completely) invariant, and apparently quite simple, the latter complex for principled reasons discussed in these papers, and perhaps involving little evolution. Corballis's own proposal in terms of episodic memory falls under the general critique of Hauser, *op. cit.*; specifically, there is no connection between episodic memory and the properties of language to be explained, with the added problem that episodic memory, as he notes, is acquired by children after language competence is well-established. Similar comments hold of other proposals, including popular ideas about theory of mind and language origin.

of course that is even more true for other mechanisms devised to yield what comes free under SMT.

A second conclusion, and one that is far-reaching in its import, is that IM provides the right forms for semantic interpretation, including quite intricate cases, but the wrong forms for SM interpretation. These are derived by another operation of MC, deleting copies. Observed expressions therefore pose serious perceptual/parsing problems, the familiar filler-gap problems. We therefore find a deep asymmetry between the two interfaces: the core principles of language design are well adapted to the conceptual-intentional interface, but yield problems for perception and parsing, conclusions supported by many other properties of language.

A third conclusion is that we have an immediate explanation for a long-standing puzzle, going back to the 1950s: why do languages universally adhere to the computationally complex UG principle of structure-dependence of rules, always ignoring much simpler processing principles relying on linear order? The asymmetry of BP is, in fact, the only known explanation for this curious phenomenon.

A fourth conclusion is a suggestion about another curious phenomenon: the fact that languages have unvalued features, assigned values in certain structural positions. These features mark phases, a particular execution of strict cyclicity, well-motivated on grounds of computational efficiency; and it may be that their only motivation is to do so. I'll assume here that the phases are CP and v*P, for reasons discussed elsewhere, leaving open the controversial status of DP.

The apparent asymmetry of BP provides additional reasons for returning to a traditional conception of language as essentially an instrument for construction and interpretation of thought – in effect providing a "language of thought" (LOT). Externalization appears to be a secondary phenomenon, in large part a reflex of properties of the SM systems that are employed. And various uses of externalized language, such as communication, are then even more peripheral to the fundamental design and nature of language, contrary to widely held dogmas that have little basis.

Such conclusions illustrate the "ideal case" of linguistic inquiry, a goal towards which it should strive, in my opinion. Let us now consider how the ideal case can be extended farther.

The earliest work in generative grammar focused on several core properties of language: compositionality, displacement, order, and projection. In the earliest versions, order and projection were assigned to PSG, displacement to transformational grammar (TG), and compositionality to interaction of the two. By the late 1960s, PSG was replaced by X'-theory, eliminating many redundancies and arbitrary stipulations. Over the years, further simplifications have improved both X'-theory and TG.

X'-theory radically simplified PSG, eliminating many stipulations, but also introduced a new one: that constructions are necessarily endocentric. That was an error I believe, for reasons discussed in POP. In fact, exocentricity is common, including all cases of IM and many others. Considerable artificiality and complexity have been introduced in the attempt to impose endocentricity. Accordingly, there is no notion of SPEC, though I will continue to use it for expository purposes, along with some other familiar notations, like trace.

Under SMT assumptions, pursued in POP, the division of labor is different from earlier frameworks. Order is relegated to externalization. Compositionality and displacement are provided by the simplest computational operation, binary Merge. What remains is projection, which is a qualitatively different property: theory-internal, and not, like the others, easily detectable in presented utterances. Optimally, projection should be reducible to a labeling algorithm LA, a special case of minimal search (like Agree), which in turn falls under MC. LA seeks heads H within its search domain (observing the Phase Impenetrability Condition PIC – that is not searching too far). Since the same labeling is required at CI and for the processes of externalization (though not at SM, which has no relevant structure), it must take place at the phase level, as part of the Transfer operation.

Crucially, LA does not yield a new category as has been assumed in PSG and its various descendants, including X' theory. Under LA, there is no structure $[_{\alpha}\, X]$, where α is the label of X. LA simply determines a property of X for externalization and CI. It is therefore advisable to abandon the familiar tree notations, which are now misleading. Thus in the description of an [XP, [YP, ZP]] structure, there is no node above either of the two merged constituents. There is no label for the root of the branching nodes.

One should also view with caution the resort to multimensionality, late merge, sidewards movement, etc., along with the complex tree notations used standardly with lines connecting various nodes. Tree notations may misleadingly suggest that these operations are special cases of Merge. They are not. They are new operations, and therefore carry a burden of proof.

There is also a general methodological point that should be kept in mind concerning "exotic constructions" such as ACD, parasitic gaps, or others for which the learner as little or no evidence (as is typically the case when the evidence is semantic). It is highly unlikely that they involve mechanisms other than those that account for simple and familiar constructions. There would be no way to learn such mechanisms, and it is implausible to think that they are properties of UG. These considerations impose significant constraints on investigation of these topics. Such investigation has often been highly revealing, but remains descriptive – posing problems to be solved -- until this methodological condition is met.

LA is trivial for {H, XP} structures, H a head. In this case, LA selects H and the usual operations apply. The interesting cases are {XP, YP}, neither a head, in which case LA finds {X,Y}, the respective heads of XP, YP, and there is no label unless they agree. In that case, the label is the pair of the agreeing elements. An element raised by IM to create this structure is in what Rizzi calls a "criterial position." It follows that IM is successive-cyclic, driven by labeling failures, continuing until a criterial position is reached.

One consequence is an explanation for the puzzling phenomenon of raising-to-object. Work tracing back to Paul Postal 40 years ago shows convincingly that in ECM constructions, such as "John expects Bill to leave," the object *Bill* has raised from embedded subject to surface object. Lasnik and Saito (1991 and Lasnik (2002) added further evidence, and showed how the result can be obtained by rule systems that apply elsewhere. The embedded subject raises to SPEC of *expect*, which then raises to the phase head v*.

The double raising restores the original word order, but with a different structure, yielding interesting semantic effects. The puzzle is why such rules should apply, restoring the original

word order and with no empirical evidence available to the language learner, a striking illustration of the unlearnability of syntactic and semantic structures. A solution is provided by labeling theory, as shown in POP, relying on no new principles, hence adhering to the methodological condition for exotic constructions – super-exotic in this case, since the learner has no evidence at all. We return to the case.

Another puzzling phenomenon may also receive a natural explanation in these terms. It was observed by Guasti and Rizzi (2012) that in SPEC-head structures, where SPEC is the clausal subject, agreement tends to be richer than in long-distance agreement with subject remaining in situ. This is the opposite of what one would expect on perception-parsing grounds. In SPEC-head constructions agreement is adjacent, hence need not be signaled by visible morphology, which is more needed in the long-distance case. The reason may be that in SPEC-head constructions rich agreement is necessary for labeling. In long-distance agreement there is no labeling issue. Once again optimal language design appears to conflict with efficiency of perception and use.

Another long-standing puzzle has to do with the status of head-head constructions: Which is the label? There is a plausible answer in terms of an approach to categories developed in detail by Hagit Borer (2005a,b, 2013), also discussed by Alec Marantz (1997; Embick and Marantz, 2008). Assume that the substantive elements of the lexicon are roots, unspecified as to category, and that their category as nominal, verbal, etc., derives from merger with a functional element n, v, etc. It is as if languages are a kind of idealized proto-Semitic. Suppose further that roots alone are too "weak" to serve as labels. Then structures of the form {K, R}, where R is a root and K a category marker, will be labeled K, as required. Other {H, H} structures are unlabellable, and the problem dissolves – largely, perhaps completely. We will see directly that there is independent support for this conclusion.

Let's turn now to a range of problems beyond what is discussed in POP. One is what Rizzi (2013) calls "criterial freezing," "the halting problem": when XP reaches its criterial position, it cannot raise any further. Consider (1), *t* the lower copy of "which dog":

(1) *[$_{\beta}$ which dog do you wonder [$_{\alpha}$ *t* [$_{\gamma}$ C_Q John likes *t*']]]

If *which dog* remains in its criterial position *t*, then labeling of α proceeds as in POP, with the agreeing Q features resolving the ambiguity of LA. What then bars (1), in which *which dog* has raised from its criterial position?

Rizzi presents an elegant solution, but it requires that we assume that projection yields a new category, as in PSG and its descendants. But under SMT assumptions, this should be dispensable. Suppose, then, that we take (1) as it stands, without any new category assigned by projection.

At β, t is invisible for the reasons discussed in POP. Therefore α is labeled Q, the head of ɣ, and is interpreted as a yes-no question, with Aux-raising and rising intonation. But this is gibberish, crashing at CI and solving the problem.[4]

[4] Hisatsugu Kitahara points out (pc) that the analysis appears in Chomsky (1995), though without recognition of this problem.

The other relevant case is subject-predicate constructions. In these, the subject transfers within the CP phase, so is not accessible for raising from the criterial position, with some interesting exceptions to which we return.

Kitahara notes that in Japanese, raising from the SPEC-Q criterial position is permissible, unlike English. The reason, he argues, is that the Japanese question particle *ka* can double as the counterpart of "whether," so raising from the SPEC-Q position yields a grammatical sentence. There are further complications that he explores in a forthcoming paper.

Let us move on to some additional problems. One long-standing puzzle has been several apparently exceptional properties of subjects, specifically EPP and ECP, and their cross-linguistic variation. One might suspect that these properties should be unified and given a principled explanation, which should also extend to the second phase, v*P. This would be a substantial step towards SMT, if it can be achieved.

Let us approach the questions in terms of labeling theory. In these terms, there is indeed a similarity: in both cases, at the labeling stage, the subject must be visible, not a copy, in SPEC position in [C [SPEC-TP]].

This observation suggests that T is similar to roots: T is too "weak" to serve as a label. With overt subject, the SPEC-TP construction is labeled $<\phi, \phi>$ by the agreeing features. Therefore, English satisfies EPP. Suppose that the subject raises to SPEC-CP, and is therefore invisible to LA for the usual reasons. What remains visible is T alone, which cannot label.[5]

Support for this approach comes from null-subject languages like Italian. As discussed by Rizzi (1982), ECP appears to be violated freely in these languages.[6] Following Rizzi, the parametric difference is commonly attributed to richness of agreement. In terms of labeling theory, Italian T, with rich agreement, can label TP and also {SPEC, TP}; for English, with weak agreement, it cannot, so that SPEC must be visible when LA applies. Therefore English-type languages observe ECP, but null subject languages do not.

What about EPP for Italian (and null subject languages generally)? It has usually been assumed that EPP holds for these as well, with a null pro expletive subject. But there is little reason for this assumption, which may be a historical accident: English was investigated first, and it was natural to extend the principles discovered for English to other languages. Dropping the assumption, we can conclude that Italian lacks EPP just as it does not conform to ECP, thus unifying EPP and ECP in terms of labeling, and keeping the parametric difference in terms of richness of agreement.

Note that raising of subject to SPEC-T is not blocked for Italian. It is just not necessary. That leaves open a range of interesting questions, which I will put aside here, about just where the subject and VP are when the subject does not raise to SPEC-T.[7]

[5] We return to some apparent exceptions.
[6] Rizzi (1982), chap. IV. This was a major contribution to the then-emerging Principles & Parameters framework, showing how properties cluster around a single parameter, and going well beyond what is discussed here.
[7] See Belletti (2001, 2004).

As discussed in POP, there are independent labeling reasons why SPEC of v*P must raise to SPEC-TP. The crucial cases for EPP, then, are constructions lacking external arguments: in English-type languages an expletive is required. Such languages appear to be rather unusual, and furthermore, what fills the subject position seems typically to have strange properties. In English, it is an adverb (and not all cases have a sensible adverbial interpretation); in Icelandic, a demonstrative; in Hebrew, a meaningless particle alternating with negation. In German, French and other languages there are special constructions, which furthermore raise problems. Thus in the counterpart to sentences with in-situ subjects, such as "there were destroyed three buildings," the expletive subject has contradictory agreement properties: with the copula and the in-situ subject.[8] And there are other oddities.

This carries us part way towards unifying ECP and EPP, with their parametric differences. Let's now move on to the second phase, v*P.

Compare the relevant structures of the two phases CP (2) and v*P (3):

(2) $[_{\delta}$ C $[_{\alpha}$ Tom T $[_{\beta}$ t v*…]]] ("Tom read a book")

(3) $[_{\delta}$ v* $[_{\alpha}$ John R $[_{\beta}$ t …]]] ("they expected John to win")

In (2), *Tom* has raised from the trace position as SPEC of v*P. In (3), *John* has raised from the trace position as well, and is now in the same structural position as the subject *Tom* in (2). Since R is universally too weak to label, it follows that the analogue of EPP for v*P holds for both English and Italian, but here EPP refers to raising-to-object. Just as English T can label TP after strengthening by SPEC-T, so R can label RP after object-raising. That takes care of the very strange raising-to-object phenomenon already mentioned, with no new rules involved, a methodological necessity, as noted.

Suppose now that we replace the subjects *Tom* and *John* by *who*, and look into when it can be raised further. The relevant constructions after subsequent raising are the following:

(2') $[_{\gamma}$ who do you v* $[_{\varepsilon}$ think $[_{\delta}$ C $[_{\alpha}$ t T β]]] ("who do you think that read the book")

(3') $[_{\gamma}$ who do you $[_{\delta}$ v* $[_{\alpha}$ t R β]]] ("who do you expect to win")

In (2'), ECP is observed for English, but not for Italian, as we have seen. But in (3'), ECP is universally violated. The raised object *who* can raise further to SPEC-CP. Why should this be so?

To see why, let's return to CP. As well-known, ECP can be violated in English if the complementizer *that* deletes, as in (4):

(4) $[_{\gamma}$ who do you v* think Ø $[_{\alpha}$ t T β]] ("who do you think read the book").

[8] These examples are marginal in English but are normal in English-type languages. See Chomsky (2001).

The phenomena are described by the *that-trace filter*, but the filter is actually stated backwards: it is the violation of ECP by *that*-deletion that is the idiosyncratic case.[9]

Keeping to the major cases, the order of rules for (4) is the following:

(5)(1) Inheritance

(2) IM of *who* in α (EPP)

(3) Labeling of α as <ϕ, ϕ>

(4) C → Ø, so that *who* can remain in situ and still be accessible to IM in the next phase

(5) Transfer.

But transfer of what? The phase head C has disappeared, along with the syntactic object {C, α}. The natural assumption is that phasehood is inherited by T in step (1) along with all other inflectional/functional properties of C (ϕ-features, tense, Q), and is activated on T when C is deleted. Therefore, what is transferred in (4) is β. We therefore preserve the computational simplifications of PIC, and it follows that movement must be successive strict cyclic, universally, with visible effects in some languages at v*P (Indonesian, Dinka) just as at CP (in many languages).[10]

More specifically, all inflectional/functional properties of C are inherited and the unvalued ones are deleted from C. All are activated at T when C is deleted.[11]

Note that ɤ in (4) is now a single phase including [$_{\alpha}$ who T….]. α has already been labeled, in the embedded CP phase, which has now been de-phased by disappearance of C. We must assume, then, that IM of *who* to the matrix clause doesn't de-label α. The basic principle is that memory is phase-level -- as, e.g., in distinguishing copies from repetitions. Hence at the phase level ɤ in (4), if some element (here α) has already been labeled, the information is available and α stays labeled.

Merely for expository purposes, we can think of labeling as assigning the feature "label" to heads that are labeled, hence to T in step (5(3)). And nothing removes it.

---

[9] A variety of similar phenomena in other languages are reviewed and analyzed in Rizzi and Shlonsky (2007). On English dialects that permit subject extraction with *that* present, see Sobin (1987, 2002). A simple assumption is that in these dialects the complementizer in fact deletes in the syntax, as in other dialects, but leaves a phonetic residue, invisible to the syntactic operations; perhaps like the appearance of agreement elements in the complementizer after inheritance by T in some languages.

[10] But it remains to show that previous analyses can be properly restated in this simplified framework.

[11] There is more to say about this matter. Thus the Q feature of C is substantive, so C_Q is not deleted under inheritance. But after inheritance, Q functions at T, as we see in cases where subject wh-phrases remain in situ in embedded interrogatives. The phase-head feature of C is not substantive, so the phase-head C can delete, phase-hood becoming active at T, where the feature is inherited. Is phase-hood activated at T when C does not delete? That raises questions related to timing of operations within CP, to which we return at the end. Note also that spell-out of inherited features is a different issue, related to externalization, not activation for the purposes of syntax-semantics.

The analysis works, and involves no new principles. But the familiar methodological question, already mentioned, hovers in the background here. Complementizer-deletion (and similar phenomena elsewhere) is an "exotic" property, so that the analysis should be motivated by non-exotic constructions.

With that in mind, let us return to v*P in (3).

Recall that the raised object in v*P is structurally parallel to the subject in CP. Therefore we can duplicate (5) step-by-step for v*P. In step (4), we want v* to become invisible, like C (in the idiosyncratic case of *that* deletion), so that the object within α may stay in situ, now at edge, just like the subject of CP after C-deletion,[12] and is therefore accessible to extraction at the next phase.

Why then should v* become invisible, universally? Recall that R raises to v*, forming an amalgam [R, v*]. What exactly is it?

Conventional treatments of head-raising assume that the raised element is adjoined to its host – pair-merged to it, in our terms. But that is not correct. For example, when V raises to T, a collection of inflectional features, the result is not V affixed to T, but T affixed to V. More generally, the conventional theory of head-raising seems to have the story backwards: the host should be affixed to the raised element. There is no conceptual argument against this: head-raising is a unique operation, with special properties, and there is no conceptual preference for one or the other form of amalgamation, hence no reason to reject what seems to be accurate empirically.

Accordingly, raising of R to v* yields an amalgam with v* adjoined to R, and the affix is invisible to the labeling algorithm. Note that although R cannot label, the amalgam [R-v*] can.

We therefore have principled reasons for the analysis in (5) for the non-exotic construction v*P, and it simply carries over to the case of deletion of complementizer.

That largely takes care of ECP, EPP, and the parametric variation, unifying these in accord with SMT. These are the special cases involving subjects: surface subjects in TP and their structural analogues in v*P after raising-to-object. The analysis consists of a number of small parts, each plausible in itself, each in accord with SMT. They fit together to yield a range of interesting conclusions, which suggests that there is something right about the analysis.

Let us turn to a few of the questions that arise nearby.

Suppose that in (5) the operation (5(2)) does not take place: the object does not raise. Then when R raises to v*, its copy will be invisible in α, which will be labeled by β without head-raising. Therefore, to ensure that head raising takes place, it is necessary for step (2) in (5) to precede step (3): raising of the object must be obligatory and precede labeling. Suppose the order of (2), (3) is optional. In that case, raising of object will be optional – as in fact argued by Lasnik and Saito. That is problematic, for reasons discussed in POP, leaving the question open.

[12] And the external argument in v*P, not indicated here.

If the order is optional, that may provide a way to approach a question not mentioned above. Consider again the structure (2'):

(2') [$_{\gamma}$ who do you v* [$_{\varepsilon}$ think [$_{\delta}$ C [$_{\alpha}$ t T β]]]]

Prior to raising of *who* to matrix SPEC, the object of *think*, δ, should raise to SPEC of ε, with the root *think* then raising to v*, the ordinary raising to object analysis. But questions arise about how ε should be labeled, since the raised object in this case lacks the relevant features; there is no agreement. There are also other questions about the status of this construction, but none of these would arise if δ remains in situ, as it can if the order is optional.

Note the difference from the halting problem. In that case the relevant construction is (6), with *which book* having raised illegitimately from its criterial position as SPEC-Q in β:

(6) *which book [$_{\alpha}$ they wonder [$_{\beta}$ which book Q [he read]]]

At phase β, the complement of Q is transferred, and the edge *which book Q* remains to be labeled at phase α. But at phase α, the lower copy of *which book* is invisible for the usual reasons, so β is never labeled by <Q, Q>.[13] Therefore, Rizzi's *criterial freezing* holds, and the halting problem is overcome along the lines sketched earlier.

There are related issues about the order of operations. The operations (1), (2) of (5) are counter-cyclic: the phase head C is introduced, permitting inheritance to T under (5(1)), prior to (5(2)), IM of the subject to SPEC-T in α (EPP). This analysis of EPP is carried over from POP and earlier papers, but it is problematic, as pointed out by Epstein, Kitahara, and Seely (2012), because it involves a substitution operation that is ternary, even though only narrowly so. They present an ingenious alternative keeping strictly to Merge and cyclicity, but there may be an even simpler approach.

POP takes over the earlier assumption that operations all apply at the phase level, though as Samuel Epstein observed some years ago, that cannot be true for EM, which has to apply freely for the phase even to be constructed – and in fact, every approach to language that even permits syntactic objects presupposes EM or some more complex alternative.[14]

The simplest conclusion, then, would be that Merge applies freely, including IM. Hence IM can apply prior to merge of C and inheritance, and the problems are resolved. There are several reasons why this simple suggestion was not adopted in earlier work. The first is that Transfer must know whether IM has applied in order to distinguish copies from repetitions, but assuming phase-level memory, there is no need for IM to apply along with Transfer at the phase level to achieve this result. Another reason has been the lingering idea, carried over from earlier work, that each operation has to be motivated by satisfying some demand. But there is no reason to

[13] Suppose that the wh-phrase remains in situ, as subject of the embedded phrase. In that case, it is transferred in the lower phase β and is no longer accessible, just like subjects generally. Note that labeling requires not just matching but agreement of the paired heads; see POP. Agreement holds for a pair of features <valued, unvalued>. The Q feature of C is valued, so the corresponding feature of a wh-phrase must be unvalued, its interpretation as relative, interrogative, exclamative determined by structural position.

[14] A usual choice is concatenation, which presupposes Merge, adding order and an operation to remove the structure that is strongly generated.

retain this condition. Operations can be free, with the outcome evaluated at the phase level for transfer and interpretation at the interfaces.

Accordingly, we can invert the order in (5), revising it to (8), illustrating with the v*P analogue to CP in (7) and *t* the lower copy of DP:

(7) [v* [$_{\alpha}$ DP [R [$_{\beta}$ t ….]]]

(8) (1) form R-β by EM
(2) IM of DP in α (EPP)
(3) Merge v*, reaching the phase level
(4) Inheritance
(5) Labelling; α is labelled <ϕ, ϕ>
(5) R raises to v* forming R with v* affixed, hence invisible, so phasehood is activated on the copy of R, and DP (which can be a wh-phrase) remains in situ, at the edge.
(6) transfer of β

The same analysis carries over to CP, whether or not C deletes.

That revision overcomes the Epstein et al. objection to the former treatment of EPP, no longer requiring a (weakly ternary) substitution operation in addition to Merge.

Daniel Seely points out that under the cyclic analysis in (8), we lose an earlier explanation for the fact that in subject-predicate ({C, {DP, TP}}) constructions, it is T rather than D that raises to C. We can no longer appeal to the fact that raising takes place before DP is raised to SPEC-TP. The likely answer is suggested by the v*P analogue. Here R, not the head D of its DP SPEC, universally raises to v* for principled reasons of root-categorization, just as inheritance relates v* and R, not v* and D. Probably the same is true of the CP counterpart.

There are unsettled questions about head-raising, in particular, as to whether it falls within (narrow) syntax or externalization. A strong argument for the latter is that it typically does not feed semantic interpretation, unlike A- and A'-movement. Thus in the now-classic analysis of V-raising by Jean-Yves Pollock, interpretation of the verb is the same whether it raises or not.

An argument that head-raising should be syntactic is that it appears to be successive-cyclic, as in R → v* → T → C. These steps are quite different in character, however. R-raising is universal in the framework assumed here, and might be syntactic without affecting the status of the other cases. Furthermore, T-raising falls under the different category of V-second phenomena, which could be assimilated to externalization if it were restricted to the root; but it sometimes is not, as in Dutch. The issue therefore remains unclear, along with numerous other questions that arise when we look further.

Returning to our main concern here -- the special properties of CP subjects (EPP and ECP) and their v*P analogues -- we have an explanation and unification for both phases in terms of labeling, and a simple parameter for the CP phase. There are, as always, innumerable complications to be dealt with when we seek to move beyond into new territory. But the account seems to extend the ideal case discussed in POP to new domains. It adheres to SMT throughout:

nothing is invoked beyond the simplest computational operation Merge and reasonable interpretations of general principles of MC.

Bibliography:

Belletti, A. (2001). "Inversion" as Focalization. In A. Hulk and J-Y Pollock (eds.). Inversion in Romance and the Theory of Universal Grammar. Oxford: Oxford University Press.

Belletti, A. (2004). Aspects of the Low IP Area. In L. Rizzi (ed.). The Structure of IP and CP. New York: Oxford University Press.

Borer, H. (2005a). In Name Only. Structuring Sense, Volume I. Oxford: Oxford University Press.

Borer, H. (2005b). The Normal Course of Events. Structuring Sense, Volume II. Oxford: Oxford University Press.

Borer, H. (2013). Taking Form: Structuring Sense Volume III. Oxford: Oxford University Press.

Chomsky, N. (2001). Derivation by Phase. In M. Kenstowicz (ed.), Ken Hale: a Life in Language. Cambridge: MIT press.

Chomsky, N. (2013). Problems of Projection. Lingua 130: 33-49.

Corballis, M. (2011). Language and Episodic Sharing. In R. Menzel and J. Fischer (eds.), Animal Thinking and Comparative Cognition. Cambridge: MIT press.

Embick, D., and A. Marantz. (2008). Architecture and Blocking. Linguistic Inquiry 39.1: 1-53.

Epstein, S., H. Kitahara, and D. Seely. (2012). Structure Building that Can't be. In M. Uribe-Etxebarria and V. Valmala (eds.). Ways of Structure Building. Oxford: Oxford University Press.

Guasti, M-T, and L. Rizzi (2012). Agreement and Tense as Distinct Syntactic Positions: Evidence from Acquisition. In G. Cinque (ed.), Functional structure in DP and IP: The Cartography of Syntactic Structures. New York : Oxford University Press.

Hauser, M., C. Yang, I. Tattersall, R. Berwick, M. Ryan, J. Watumull, N. Chomsky, and R. Lewontin (2014). The Mystery of Language Evolution. Frontiers in Psychology, 7 May.

Lasnik, H. (2007). Clause-mate Conditions Revisited. Glot International, 6.4, April 2002. 94–96.

Lasnik, H., and M. Saito (1991). On the Subject of Infinitives. In L. M. Dobrin, L. Nichols, and R. M. Rodriguez (eds.), Papers from the 27th Regional Meeting of the Chicago Linguistic Society, Part I: The General Session, 324–343. Chicago: Chicago Linguistic Society, University of Chicago.

Marantz, A. (1997). No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of your own Lexicon. University of Pennsylvania Working Papers in Linguistics 4.2: 201-225.

Rizzi, L. (1982). Issues in Italian Syntax. Dordrecht: Foris.

Rizzi, L. (2013). Cartography, Criteria and Labeling. Ms, Universities of Siena and Geneva.

Rizzi, L., and U. Shlonsky (2007). Strategies of Subject Extraction. In H-M Gärtner and U. Sauerland (eds.). Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics. Berlin: Mouton de Gruyter.

Sobin, N. (1987). The Variable Status of Comp-Trace Phenomena. Natural Language & Linguistic Theory 5.1: 33-60.

Sobin, N. (2002). The Comp-trace Effect, the Adverb Effect and Minimal CP. Journal of Linguistics 38: 527-560.