Head Movement in *Problems of Projection*[*]

Vicki Carstens, Norbert Hornstein, and T. Daniel Seely
University of Missouri, University of Maryland, and Eastern Michigan University

1. The Problem

A central property of grammatical processes is that they are structure dependent. Chomsky 2013 (henceforth POP) illustrates this well-established fact with Yes/No question formation in English. In (1a) the fronted *can* must relate to the matrix verb *swim* and cannot link to the linearly closer *fly*.[1] Thus (1)a has the paraphrase (1)b, not (1)c:

(1) a. Can eagles that fly swim?
b. Is it the case that eagles that fly can swim?
c. #Is it the case that eagles that swim can fly?

This follows if grammatical operations like T-to-C must be structure dependent, based on hierarchical relations rather than linear proximity.

Taking this assumption as a starting point, POP asks two related questions: (a) why must syntactic operations exploit hierarchical conceptions of proximity and never linear conceptions? [2] and (b) assuming that "inversion depends on locality independent of category" (POP: 43), why does T-to-C move T rather than (a subpart of) the expression in Spec T?

POP's answer for (a) is that the objects of grammatical manipulation are only hierarchically specified. They have no linear order, the latter arising from the mapping to the sensory and motor systems (S&M) at spell out (SO). Thus, until an object has been transferred to S&M, phrase markers are unordered and so operations that transform them (including T-to-C I-merge) cannot refer to such order. The idea, both simple and elegant, is that grammars cannot use absent information.[3]

---

[*] Thanks to Samuel Epstein and Hisa Kitahara for teaching us much about POP; to the University of Michigan for organizing the 2013 Linguistics Society of America's summer institute where the ideas for this paper originated; to the University of Missouri's South African Education Program for funding Xhosa research reported in §2.4; and to Brandon Fry for eloquently articulating in EKS's LSA class that things other than locality should be considered in the analysis of T-to-C.

[1] As Chomsky notes, there are parallels that don't involve fronted T, among them that the adverb *instinctively* must modify *fly*, and cannot modify *swim*:

(i) Instinctively, eagles that swim fly

[2] POP points out that linear conceptions are in some sense "simpler" in that they can be defined with visible properties like word order rather than more abstract conceptions of phrasal structure.

[3] Though the solution described has obvious virtues, there are alternative hypotheses that would work equally well. For example, it is well-known that subjects are islands to extraction. If this were so, then the relative clause would not be a potential launch site for T-to-C movement. Note that if islandhood is explained in terms of Transfer (viz. the

POP gives a variant of this answer to (b) as well. Here's the proposal: Were (2) the structural input for T-to-C, then we would expect either (i) that D or T could both move to C as they are equidistant from C, or (ii) that nothing could move to C as there is no unambiguously closest candidate, given the equidistance of D and T from C. [4] While explicit discussion of T-to-C in POP assumes (i), POP's approach to labeling suggests that rule application is deterministic and ambiguity gums up the works, hence (ii). [5] Neither approach is free of the problems we will raise in this paper.

(2) [C [$_{TP}$ [D NP] [$_{T'}$ T vP]]]]

POP's proposed solution to the locality puzzle in (2) is the following: The reason that only T moves to C is that D (and the phrase that contains it) is in its vP-internal base position when C is Merged. The structural input to T-to-C movement is not (2) but (3), since the external argument only raises after C-T Feature Inheritance (Chomsky 2007, 2008, Richards 2007). Thus at the crucial point there is nothing as close to C as T, and that's why T alone can move.

(3) *Chomsky's claim: only T can raise to C because the C-T relationship is established when EA is still in situ*

[C [$_{TP}$ T...[$_{vP}$ [D NP] v...]]]

Summing up, T and not D raises to C because D is not there to raise. As in the solution to problem (a), the key assumption is that the derivation cannot exploit absent information.

In what follows we concentrate on POP's approach to (b). We argue that when a fuller range of head movement operations are considered, POP's conclusion, viz. that D does not count because it is not *there* to be moved (I-merged), turns out to be inadequate in a very important way: the crucial ambiguity regarding movement should arise in any configuration of the form in (4), where α and β correspond to XPs or intermediate

---

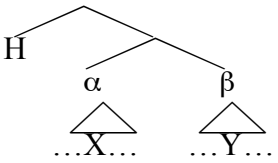
reason it is not a source of movement is that it is not there), then an explanation similar in kind to the one POP offers would be available; only the matrix T is visible at the point where the operation would apply. In what follows we abstain from discussing this alternative and stick to the assumptions in POP.

[4] Labels such as *TP*, *T'*, *vP*, are employed for convenience. POP suggests that labels are not an available part of syntactic objects, but the argument relies only on the assumption that D and T in (2) are equidistant from C (and see (4) for a label-free schematic).

[5] POP proposes that both labeling and identification of candidates for internal Merge are based on minimal search. The labeling algorithm cannot freely choose between α and β in a configuration like (4). Compare two passages from POP: 43, the first about I-Merge: "…inversion could just as well yield 'eagles [young are flying]' rather than 'are [young eagles flying]' as the interrogative counterpart to 'young eagles are flying'." The second is about labeling: "The interesting case is SO = {XP, YP}. Here minimal search is ambiguous…There are then two ways in which SO can be labeled: (A) modify SO…or (B) X and Y are identical in some relevant respect, providing the same label…"

projections in the traditional X' schema and H, X, and Y are heads.[6] But POP's solution is not applicable in most such cases. One consequence of this state of affairs is that, when combined with the VP Internal Subject Hypothesis (VPISH), POP assumptions falsely predict that D-raising should always compete with V-to-T movement. This either blocks verb raising altogether or allows D-raising as a licit alternative, depending on the choice of (i) or (ii).

(4) *X and Y are equidistant from H*



We present several arguments from V-to-T movement in favor of the more traditional view that head-movement is sensitive to categorial and other features of the target and the moving item, not just to locality. If this is true of V-to-T, it stands to reason it could be true of T-to-C as well.[7]

The second half of our paper addresses T-to-C movement in Wh-questions. It is well-known that local subject Wh-questions disallow T-to-C while all other direct Wh-questions require it.[8]

(5) a. Which boys (*did) eat the pizza
b. Which pizza *(did) the boys eat
c. Which boys$_1$ *(did) you say t_1 ate the pizza

We argue in §3 that, like the facts of V-to-T movement, this asymmetry shows that factors other than locality are involved in motivating and constraining head-movement. See Pesetsky & Torrego 2001 for a proposal that T and a subject Wh-phrase are indeed equidistant from the local C. Both have features relevant to C, but the subject has more of them; its movement hence blocks T-to-C. In §3 we compare this approach to a potential alternative based on a POP proposal that subject Wh surface in Spec, TP. We show that such an account also relies crucially on features of T and C to explain the distribution of T-to-C movement, a state of affairs that seems at odds with the spirit of POP's purely locality-based approach to (1).

---

[6] While POP does not present a theory of head-movement, the argument it presents relies crucially on analysis of T-to-C as syntactic. We adopt this view and generalize it. See Roberts 2010 for arguments that head-movement is syntactic, including the licensing effect of T-to-C on subject NPIs: **Anybody didn't leave* is ill-formed, but *Didn't anybody leave?* is fine.

[7] Chomsky 2007:21, comparing C and v*, notes that "T may or may not raise to C, but V must raise to v* which therefore is an affix." Consider however English T, some but not all choices of which are affixes (i.e. past versus future). Analogously interrogative C might be affixal and declarative C not. See §3.3.2 for a slightly modified interpretation.

[8] The *do* in (5) is unstressed. Stressed *do* is permitted in (5)a with an emphatic interpretation, but this is irrelevant to the T-to-C movement question explored here.

This paper is structured as follows. §2 explores V-to-T movement. §3 addresses the distribution of T-to-C in Wh-questions and how to account for it. §4 concludes.

2. V-to-T: the equidistance problem

2.1 Introduction

As noted above, POP's proposed solution to (b), i.e. to why there is a T-to-C relation but not a D-to-C relation, exploits the fact that in this case, the configuration [H [XP, YP]] arises when XP raises from a lower position after H is Merged. The solution therefore cannot be extended to many familiar cases of head-movement. We begin in §2.2 by illustrating the problem with respect to V-to-T movement across an external argument. §2.3 briefly considers whether it is relevant that in SVO languages, the subject raises beyond the landing site of V-movement, yielding a potential remedy for this problem in terms of the simultaneity of phase-level processes. §2.4 shows that subject-in-situ VSO languages resist such an account. §2.5 adds that V-raising across negation is also ruled out by Chomsky's solution to (b).

2.2 V-raising across EA

Consider the structural relations between T and material in vP. By the logic of POP, V+v and D, the immediate daughters to vP and DP respectively, are equally close to T. Therefore the puzzle POP raised for T-to-C movement arises here as well. Why is it that in (6) there can be a v+V-to-T relation but not a D-to-T relation? Minimal search should obtain an ambiguous result. Yet V+v can raise to T in familiar languages, and D cannot. [9]

(6) *Minimal search should find D and v equidistant from T*

...[$_T$ T [$_\alpha$ ...DP... vP]]

Chomsky's argument that C-to-T is evidence of a C-T relation prior to EA-raising seems to be nullified by the recurrence of the same phenomenon at a lower point in the tree. A parallel account of v-to-T movement would have to claim that a relation existed between T and v prior to Merge of the EA. However, in this case, there is no apparent lower position to shunt the external argument to in order to finesse the problem. Thus, if VPISH is correct, POP's proposed solution to (b) above is too narrow to account for the regularities of v-to-T movement.[10, 11] If the external argument raised to Spec, TP before

[9] For expository ease we briefly delay discussion of C's potential role in this via FI.
[10] Of course, one might reconsider VPISH, locating the Merge position of the subject elsewhere. Exploring alternatives to the VPISH lies beyond the scope of the present paper. We think that the problems of head-movement for POP will recur in any case (consider N-to-D movement, and VR across Negation to be discussed in §2.5 below).
[11] It is also worth noting that whereas C-T and v-V share features through Feature Inheritance in the current framework, there is no analogous relationship between T and v. Were the EA to originate lower than Spec, vP, this fact and the operation of phasal Transfer would seem to necessitate EA raising to Spec, vP if it is ever to interact with

v-to-T movement, its unpronounced copy in vP would be invisible to T under POP assumptions, hence solving this problem. But this possibility seems to conflict with Chomsky's account of C-to-T, which as we have seen assumes that the subject does not raise before C is Merged.

2.3 Simultaneity of phase-level processes

Recall that POP adopts both the VPISH and the Feature Inheritance hypothesis of Chomsky 2007, 2008 and Richards 2008. Under FI, only phase heads have edge and Agree features, and hence after operations in the vP phase are completed, there are no further operations until C is merged. Chomsky 2008:151 writes, "It follows that the edge and Agree properties of P[hase] H[ead] apply in parallel: EF raises XP…to Spec, PH, while Agree values all uninterpretable features and may or may not raise XP to form an A-chain….the edge and Agree features of the probe apply in either order, or simultaneously…" Among the consequences of this is that raising of EA to Spec, TP after C is merged does not count as counter-cyclic, because all phase-level processes are simultaneous.

This idea might provide a way around the ordering problems we have discussed that is consistent with POP assumptions, if we extend the approach to head-movement. Suppose EA raises to Spec, TP, v raises to T (see (7)a) and, language-particular factors permitting, in a question T may also raise to C at the same time (see (7)b). If all processes in a phase are simultaneous, it seems possible in principle to arrive at a licit instance of (V-to-)v-to-T by this route (though see Epstein, Kitahara, and Seely 2012 for a different view on the status of raising EA after C is merged, with potential implications for v-to-T here).[12]

(7) a. C [T [EA v...]] → C [EA V+v+T [$_{\alpha}$ <EA> <v> ...]] *Statements with VR*
b. C_Q [T [EA v...]] → V+v+T+C_Q [EA <V+v+T> [$_{\alpha}$ <EA> <v> ...]] *Y/N Q with VR*

But the next section argues that the phenomenon of V-to-v-to-T movement is not restricted to circumstances in which EA raises. The approach outlined here cannot extend to such cases.

C/T. See Kandybowicz 2008 on problems for the combination of PIC and Feature Inheritance; further discussion would take us too far afield.

[12] Note that, in relation to interpretations (i) and (ii) regarding the results of equidistance (see discussion leading to (2) in §1), this hypothesis is by no means fool-proof. It might provide a way to view the subject as irrelevant to the licitness of v-raising, thus working around the stalling effect of equidistance sketched out in (ii). But if so, v could be similarly viewed as irrelevant to potential D-raising so that raising of either candidate is again expected to be possible as in (i). The EPP might be considered to provide a basis for electing v-raising over D-raising in this circumstance. Given further problems to be discussed in §2.4, we leave this aside.

2.4 Subject-in-situ VSO languages

VSO order in languages such as Irish and Xhosa constitutes a major class of problems for connecting the facts of V-to-T movement to EA raising.[13]

McCloskey 1996, 2005 provides persuasive empirical evidence that a sentence like (8) is the product of V-to-v-to-T across the in situ subject. By the logic of POP's approach to (1), D of the subject should be equidistant from T and the crucial ambiguity should arise.

(8) a. Sciob an cat an t-eireaball de-n luch
snatched the cat the tail from-the mouse
'The cat cut the tail off the mouse'

b. *Subparts of EA and V+v are equally close to T, but only V+v raises*

[$_{TP}$ ...T [$_{vP}$ [the cat] [$_{v'}$ snatched+v [$_{VP}$ <snatched> from the mouse]]]]
Minimal search fails to single out V+v alone

Carstens & Mletshe 2013 provide examples of VSO in Xhosa embedded clauses with overt complementizers. They argue that in a clause with default (= Class 17) subject agreement like (9)a, the subject remains very low in the structure. The verb raises across it and adjoins to a middle-field inflection identified in the Bantu linguistics literature as Mood (see (9)b); (9) adapted from Carstens & Mletshe 2013).[14] The subject cannot surface higher, as (9)c,d illustrate. The locations of C and T morphemes make it particularly unlikely that EA has raised to Spec, TP in the licit example (as it would have to for the proposal in §2.2 to be applicable). Had it so raised, the subject would be expected to surface between *okokuba* – 'that' and the future auxiliary *be*, contrary to fact.

(9) a. …okokuba ku-be ku-fund-a wena i-si-Xhosa
that 17SA-FUT 17SA-study-MOOD you 7-7-Xhosa
'…that you will study Xhosa'
[Lit: that will study you Xhosa]

b. [$_{CP}$ that [$_{TP}$ FUT [$_{MoodP}$ study+MOOD … [$_{\alpha}$ you …<study+v> Xhosa…]]]]

c. *…okokuba **wena** ku-be ku-fund-a i-si-Xhosa

d. *…okokuba ku-be **wena** ku-fund-a i-si-Xhosa

We conclude that VR is possible in Xhosa across an EA that remains low. On the interpretation of POP assumptions in (ii) verb-movement should be blocked here (like labeling in ambiguous cases; see note 5), but this is clearly not what happens. On the other interpretation, sketched out in (i), D-movement should be a licit alternative since

[13] It is not our intent to suggest that VSO has this derivation universally; only that there are languages where it works this way.

[14] We disregard a low FocusP between vP and T proposed in Carstens & Mletshe's analysis. Its inclusion reproduces exactly the same problem of head movement across the low subject again, one notch higher in the structure.

subparts of EA are as close to Mood as V+v is (see (10)). For the sake of completeness (11) shows that D cannot raise instead of V.[15]

(10) *EA and V+v are equally close to Mood.*

[$_{MoodP}$ … Mood … [$_{\alpha}$ [EA] [study+v [$_{VP}$…]]]]

(11) *D cannot move instead of V:*

a. …okokuba ku-be ku-fund-a lo mntwana i-si-Xhosa
that 17SA-FUT 17SA-study-MOOD 1this 1child 7-7-Xhosa
'that this child will study Xhosa'

b. *…okokuba ku-be **lo**(-a) m-ntwana ku-fund(-a) i-si-Xhosa
that 17SA-FUT 1this(-MOOD) 1-1child 17SA-study(-MOOD) 7-7-Xhosa
'that this child will study Xhosa'

2.5 V-raising across intervening negation

A second class of problems for POP (and the strategy sketched out in §2.2) lies in one of the standard diagnostics for V(to-v)-to-T movement. Following Pollock 1989, we take the presence of negation between the surface position of V and its object to indicate that V has raised out of VP.

(12) a. Je n' aime *pas* les fraises
I ne like not DET strawberries
'I don't like strawberries'

b. [$_{TP}$ SU…V+v+T [$_{NegP}$ Neg [$_{vP}$ <SU> <v> ...]]]

VR to T across negation is unexpected under POP assumptions because only hierarchical relations are involved in the calculus of closeness. If category does not matter, then it is irrelevant whether or not the raised item is a verb, and Neg itself ought to be a candidate for movement to T, contrary to fact. If we assume that the subject does not pass through a Spec, NegP en route to Spec TP, then Neg should be the only candidate (see (13)a). If we assume instead that the subject moves as far as a Spec, NegP, then minimal search should yield an ambiguous result; Neg and D in the subject should compete for raising to T (13)b. In neither case is raising of the verb expected to cross Negation.

(13) a. *Minimal search should find and raise Neg to T, not v*

T [$_{NegP}$ Neg [$_{vP}$ EA v...]]

[15] (11)b illustrates that the derivation fails regardless of whether or to what the Mood suffix *–a* attaches. In more traditional approaches, such an affix (overt or null) can select for the category of what raises for it to attach to. Recall however POP's proposal that T-to-C should be explained by "locality independent of category" (see citation in §1). We see no principled reason why the expectation should not hold equally here.

b. *If EA raises to Spec, Neg, minimal search should yield ambiguous results*

T [$_{NegP}$ EA Neg [$_{vP}$ <EA> v…]]

3. Wh-questions: the Subject/non-subject asymmetry

3.1 The problem

Consider now a second case of English T-to-C movement not addressed in POP. T-to-C occurs in Wh-questions (WHQ) as well as Yes/No questions. Subject and non-subject Wh-questions display a well-known asymmetry in that T-to-C is *required* in matrix questions unless the moving WH is coming from the local Spec T:[16]

(14) a. Which boys (*did) eat the pizza
b. Which pizza *(did) the boys eat
c. Which boys$_1$ *(did) you say t_1 ate the pizza

As (14) demonstrates, T-to-C is obligatory in all direct English WHQ except local subject questions like (14)a, where it is forbidden. (15) sketches out the relevant structure before WH movement for the three examples prior to C-to-T Feature Inheritance (Chomsky 2007, 2008). Shading indicates areas to which Transfer has applied; following Chomsky 2000 we assume the VP complement of v* Transfers before WH moves to Spec C. This forces all but local subject WHs to move to the edge of the local vP phase edge, as is commonly assumed. Local subjects are externally Merged to the vP phase edge.

(15) a. C [$_{TP}$ T [$_{vP}$ which boys [v [$_{VP}$ eat the pizza]]
b. C [$_{TP}$ T [$_{vP}$ which pizza [$_{vP}$ you [v [$_{VP}$ eat <which pizza>]]]]
c. C [$_{TP}$ T [$_{vP}$ which boys [$_{vP}$ you [v [$_{VP}$ say …]]]]

The problem for deriving (14)a-c should be evident: there is no difference between subject and non-subject WHQ in terms of the relation between T and C or in terms of distance between C and WH, and hence no clear basis on which to predict when T-to-C movement applies.

3.2 A solution

There is a simple way around this problem. Suppose subject WH movement proceeds via Spec T, rather than directly from Spec v. Thus the two cases involve different configurations, sketched in (16):

(16) a. C [$_{TP}$ which boys [T [$_{vP}$ <which boys> [v [$_{VP}$ eat the pizza]]]]
b. C [$_{TP}$ …T …[$_{vP}$ which pizza [you v [$_{VP}$ eat <which pizza>]]]]

In (16)b, T is clearly closer to C than the Wh-phrase is, in marked contrast to (16)a. This assumption allows a potential account of the suppression of T-to-C movement in subject questions like (14)a on the basis of locality, much in the spirit of the POP attempt to explain why C rather than D raises in (3). But in a departure from POP assumptions,

[16] As noted in footnote 8, stressed *do* is irrelevant here.

reference to features is crucial to ensure that a WH subject prevents T-to-C while a non-WH subject like that in (3) or (14)c does not. See Pesetsky & Torrego 2001 for an argument that T and a subject in Spec, TP are equally distant from C, as in POP. In Pesetsky & Torrego's account, both have features relevant to C, but a WH-subject has more of them, making a derivation that raises the WH-subject more economical than a derivation that raises both.[17]

This analysis is incompatible with POP assumptions in several ways. First, as we saw in §2, equidistant candidates for raising should either yield completely free alternatives or eliminate the possibility of raising altogether depending on one's interpretation (see note 5 and (i) and (ii) of §1). Second, POP proposes that WH-subjects never raise to Spec,CP. Hence either some POP assumptions must be abandoned or some alternative account of this asymmetry must be found.

3.3 A Possible Alternative

3.3.1 Subject questions in POP

POP proposes that subject and non-subject questions differ in a way that is potentially relevant to the asymmetry of concern. While non-subject WH move to Spec, CP, WH-subjects surface in Spec TP.

As noted in §1 and §2.3, POP follows Chomsky 2007, 2008 and Richards 2007 in assuming that T obtains its unvalued phi-features from C. POP adds to this a proposal that the Q feature of an interrogative C is among those that T inherits, on the basis that "features of an LI cannot move independently of the feature bundle to which they belong" (POP:47). Copies of all of C's features including Q are therefore inherited by T in a bundle.

(17) *Feature Inheritance in POP*: $C_{[Q,\ uPhi\ldots]}\ T \rightarrow C_{[Q1,\ uPhi\ldots]}\ T_{[Q2,\ uPhi\ldots]}$

Phrases obtain labels in POP by means of a feature-seeking algorithm that applies at the phase level. In the configuration $[_{\alpha}$ XP, YP], α can obtain a label if XP and YP share a "prominent feature." In subject questions, the copy of Q on T agrees with the Wh-subject and the constituent formerly known as TP is labeled QP (see (18)a).[18] In contrast, it is the Q feature on C that agrees with a non-subject WH, labeling CP as QP (see (18)b). Sharing of prominent features between the raised EA and T labels their containing phrase

[17] We gloss over some details of Pesetsky & Torrego's analysis, which proposes that two options do exist if the subject is not interrogative. Where the object is WH, either T or the equidistant subject can raise to the CP edge, but an exclamative interpretation results in the latter case (*What a silly book Mary bought!*). We think that the necessity of T-to-C in Yes/No questions raises some problems for this approach similar to those we discuss in relation to POP, but they lie outside this paper's scope. See §3.3 for discussion of T-to-C in Yes/No questions and a suggestion on how it might work.

[18] T also agrees with the WH-subject in phi-features. It isn't clear from POP's discussion how this factors into the labeling in (18)a.

PhiP. Crucially, POP argues that Feature Inheritance must be construed as copying "leaving Q in its original position for selection and labeling" (POP:47, note 47).

(18) a. $C_{[Q1,\ uPhi\ldots]}$ [$_{QP}$ [How many mechanics] [$T_{[Q2,\ \sout{uPhi}\ldots]}$ fixed the cars?]]
b. [$_{QP}$ [How many cars] $C_{[Q1,\ uPhi\ldots]}$ [$_{PhiP}$ [the mechanics] $T_{[Q2,\ \sout{uPhi}\ldots]}$ ...]]

POP does not explain why T should ever raise to C or why it cannot do so here. In prior treatments including that of Pesetsky & Torrego mentioned above, C has a feature that requires valuation in a local relation with a matching feature of T or the subject. But in the POP system, there is no comparable motivation for T-to-C movement in non-subject questions. Feature Inheritance ensures that C and T's features are the same ones, whether valued or unvalued. T therefore has nothing that C does not also have. C needs only its own Q-feature in order to interact with a non-subject WH-phrase for labeling or to participate in selection. It is therefore puzzling that T should have to raise. If we suppose that T-to-C simply happens freely in a move-alpha sort of way, with locality the only relevant issue, it is not clear why it should be barred in (18)a. Crucially, C is present in the derivation to supply T's features. The impossibility of T-to-C movement in subject questions therefore cannot be attributed to C being absent.

3.3.2 A speculation about multiple Qs

It seems to us in principle possible that POP's duplication of the Q-feature might yield a novel account of the subject/non-subject asymmetry for T-to-C in WH-questions. Our proposal relies on one crucial assumption: *both copies of Q must end up in a Spec, head relation with an interrogative operator*. In a subject question like (15)a and (18)a, C would lower to T (see (19)). [19] In a non-subject question like (15)b,c or (18)b. where the interrogative operator moves to Spec, CP, T must raise to place its copy of Q in a Spec, head relation with WH (see (20)):[20]

(19) <$C_{[Q1,\ uPhi\ldots]}$> [$_{QP}$ [How many mechanics] [$C_{[Q1,\ uPhi\ldots]}$+$T_{[Q2,\ \sout{uPhi}\ldots]}$ fixed the cars?]]

(20) [$_{QP}$ [How many cars] $T_{[Q2,\ \sout{uPhi}\ldots]}$+$C_{[Q1,\ uPhi\ldots]}$ [$_{PhiP}$ [the mechanics] <$T_{[Q2,\ \sout{uPhi}\ldots]}$> ...]]

To extend this analysis to Yes/No questions requires the further assumption that such questions involve a null Q-operator in Spec C. Roberts 1993, Grimshaw 1993, and Radford 2004 propose that such a null operator is indeed base-generated in Spec, CP of a

[19] It is important that this hypothetical lowering operation be distinguished from the FI operation that copies the features of C onto T. The latter process does not remove the copied features of C from it. In contrast, the former leaves no copy in CP.

[20] Given that labels are determined by a late algorithm, a possible interpretation is that T can't raise to C in a subject question because this would interfere with labeling TP as QP. By extension, though, we might expect labeling of TP as PhiP to be compromised by raising T to C in Yes/No and non-subject WH-questions. We leave this aside.

Yes/No question. Assuming this, the approach we have sketched above might be extended to them.[21]

(21) a. Did the students leave?
b. [$_{QP}$ OP $T_{[Q2, \sout{uPhi}...]}$+$C_{[Q1, uPhi...]}$ [$_{PhiP}$ [the students] <$T_{[Q2, \sout{uPhi}...]}$> [$_{vP}$ leave...]]]

This is the only interpretation of the POP system that we have been able to cook up which might capture the distribution of T-to-C in questions.

Like Pesetsky & Torrego's approach to T-to-C movement, the possibility we have sketched out here ties head-movement to features of the moving element and its target.

In sum, there is a straightforward way of accounting for the distribution of T-to-C in matrix questions if we assume that WH moves to Spec C in all cases. A different account might be possible under POP assumptions, including copying C's features to T and moving subject WHs only to Spec, TP.[22] But if something along these lines can be made to work, it will be further evidence that head-movement is driven and constrained by factors other than pure locality, as we have argued to be the case for V-to-v-to-T. A natural move is to extend our conclusions to the question of why T and not D raises to C in questions, contrary to POP's approach.

3.3.3 Unsolved mysteries

POP's proposals about Feature Inheritance raise some important conceptual questions. Under the interpretation of Feature Inheritance as copying, it is not clear why the unvalued Phi left on C do not cause the derivation to crash. POP's footnote 47 suggests that these phi-features may be "deleted or given a phonetic form (as in West Flemish) hence [are] invisible at the next phase."

Chomksy 2001, 2007, 2008 argues that transferring features before valuation is "too early" because unvalued features cause crash; and that transferring features after valuation is "too late" because they cannot be distinguished from intrinsically valued interpretable features and will therefore also cause a crash (see also Epstein, Kitahara & Seely 2010 for discussion). The upshot is that unvalued features must obtain values and be immediately transferred. They cannot be retained on a phase head, either valued or unvalued. Chomsky's two suggestions – that C's uPhi can be deleted unvalued, or given phonetic

[21] How to explain the lack of scope ambiguities in Yes/No questions is an open question for such an account, since OP might be merged in a clause lower than the matrix. This would yield impossible interpretations consistent with answering (i) with (iii) as well as (ii):

(i) Did Bill hear that Mary left?
(ii) Yes, he did.
(iii) *Yes she did.

[22] To repeat, there are many technical details left unaddressed. However, we believe that any analysis exploiting POP's core idea will face the problems outlined here.

form – are not consistent with these prior positions on the status of uPhi, but POP contains no explanation as to how these inconsistencies are to be resolved.[23]

4. Concluding Remarks

POP's explanatory ambitions fall short of its goals, as revealed when one compares its outcomes with those of traditional accounts. In more conventional analyses of head raising, selection for category and other features work in concert with hierarchical locality considerations to dictate what moves. Thus, for example, V moves to T because of some specific requirements of T that V meets; D fails to have the relevant features or properties and hence is not a potential mover in these cases (the same is true for Neg). Similarly, in cases of WHQs, features of T and C (and under Pesestsky and Torrego's account, the subject) determine when movement does and does not apply. The features of relevance are *ad hoc*, however. POP eschews such devices and tries to provide a more principled account. Sadly, the account appears to be incompatible with broad classes of phenomena and with standard assumptions about clausal architecture and the Merge location of subjects. Moreover, it is at least questionable *whether* one wants too principled an account for these cases. T-to-C movement and verb-raising are parametric options. It seems to us that features (including reference to category) are a reasonable way of distinguishing these grammar-specific options. If so, then POP's ambitions are misdirected. Sometimes *ad hoc* is just what we need.[24]

References

Carstens, Vicki & Loyiso Mletshe. 2013. Implications of Xhosa expletive constructions. LingBuzz 001750.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130: 33-49.

Chomsky, Noam. 2007. Approaching UG from below. In *Interfaces +Recursion = Language?*, ed. by U. Sauerland and H.-M. Gartner, 1-29. Berlin: Mouton de Gruyter.

Chomsky Noam. 2008. On Phases. In Foundational Issues in Linguistic Theory, ed. by Robert Freidin, Carlos P. Otero and Maria Luisa Zubizarreta, 133-166. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. In *Ken Hale: a life in language.* Ed. By M. Kenstowicz, 1-52. Cambridge: MIT Press.

Epstein, Sam, Hisa Kitahara, and Daniel Seely. 2010. Uninterpretable features: what are they and what do they do? in *Exploring Crash-Proof Grammars*, ed. by Michael Putnam, 125-142. Amsterdam: John Benjamins.

[23] POP refers the reader to Ouali 2008 for discussion of the options and their consequences. Ouali argues that in subject questions, C does not give phi-features to T; and also assumes (i) that unvalued uPhi will cause the derivation to crash, and (ii) that valued uF on phase-heads are licit. It would take us too far afield to do a full review and comparison here. We leave these points for readers to explore.

[24] *Ad hoc* does mean "to the point."

Epstein, S. D., H. Kitahara, and T. D. Seely. 2012. "Structure Building That Can't Be!" In M. Uribe-etxebarria and V. Valmala eds., Ways of structure building, in Oxford Studies in Theoretical Linguistics D. Adger and H. Borer series eds, Oxford, 253-270.

Grimshaw, Jane. 1993. Minimal projection, heads, and optimality. Draft ms., Rutgers University.

Kandybowicz, Jason. 2008. On the incompatibility of Feature Inheritance and the PIC. Ms., Swarthmore College.

McCloskey, James. 2005. Predicates and Heads in Irish Clausal Syntax. In *Verb First: On the Syntax of Verb-Initial Languages*, Andrew Carnie, Heidi Harley, and Sheila Ann Dooley, editors, John Benjamins: 155–174.

McCloskey, James. 1996. 'On the Scope of Verb Movement in Irish,' *Natural Language and Linguistic Theory* 14, 47–104.

Ouali, Hamid. 2008. On C-to-T f-feature transfer: the nature of agreement and anti-agreement in Berber. In: D'Alessandro, Rl, Hrafnbjargarson, G.H., Fischer, S. (Eds.), *Agreement Restrictions*. Berlin: Mouto de Gruyter, pp. 115-180.

Pesetsky, David and Esther Torrego. 2001. T-to-C movement: causes and consequences. in M. Kenstowicz (ed.) *Ken Hale: A Life in Language*. Cambridge: MIT Press, pp. 355-426.

Pollock 1989. Verb movement, Universal Grammar and the structure of IP. *Linguistic Inquiry* 20.3: 365-424.

Radford, Andrew. 2004. *Minimalist Syntax: exploring the structure of English.* Cambridge: Cambridge UP.

Richards, M. 2007. On feature inheritance: an argument from the phase impenetrability condition. Linguistic Inquiry 38: 563-572.

Roberts, Ian. 1993. *Verbs and Diachronic Syntax.* Dordrecht: Kluwer Publications.

Roberts, Ian. 2010. *Agreement and head-movement: clitics, incorporation, and defective goals.* Cambridge: MIT Press.