Head-head relations in *Problems of Projection* [*]

Vicki Carstens, Southern Illinois University
Norbert Hornstein, University of Maryland
T. Daniel Seely, 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.[1] In (1)a, fronted *can* must relate to the matrix verb *swim* and cannot link to the linearly closer *fly*.[2] 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? [3] and (b) assuming that "inversion depends on locality independent of

[*] Thanks to Samuel Epstein and Hisatsugu Kitahara for sharing insights into 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; 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, and for helpful comments on an earlier draft; and, finally, to two anonymous reviewers for ideas, comments and suggestions that improved the final version.

[1] This was first discussed by Chomsky (1968), see also Piatelli-Palmarini (1980).

[2] 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

[3] 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.

category" (PoP: 43), why does T-to-C move T rather than (a subpart of) the expression in Spec, TP?

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.[4]

PoP gives a variant of this answer to (b) as well. Here's the proposal: Were (2) the structural input for T-to-C,[5] then we would expect that either of D or T could move to C as they are equidistant from C.[6]

---

[4] Though the solution described has obvious virtues, there are plausible alternatives. 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 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.

[5] Labels such as *TP*, *T'*, *vP*, are employed for convenience. PoP suggests that labels are not an available part of syntactic objects, and the argumentation we are presenting here does not rely in any way on labels; only on the assumption that D and T in (2) are equidistant from C (see (4) for a label-free schematic).

[6] PoP proposes that both labeling and identification of candidates for raising to C are based on minimal search. The labeling algorithm cannot freely choose between alternatives in a configuration like (2) (in category neutral terms, between α and β of (4) to come). 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…" Our working assumption for this paper is that the stalling effect is restricted to ambiguous

(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), on the assumption that 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.[7]

(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 projections in the traditional X' schema and H, X, and Y are heads.[8] 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

---

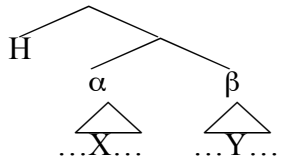
results for the labeling algorithm. We leave exploration of the basis for this difference to future research.

[7] See §2.3 on some potential problems for this "timing" approach to T-to-C.

[8] The argument that PoP 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.

predict that raising D to T from an external argument should always be a licit alternative to V-to-T movement.

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



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 (see (5)a versus (5)b,c).[9]

(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. This is true in each of several possible approaches to the asymmetry. Pesetsky & Torrego 2001 argue 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. A very different approach in terms of affix-hopping relies crucially on categorical features of potential hosts. Yet a third possibility, building on a PoP proposal that subject WH surface in Spec, TP, refers crucially to Q features in accounting for the asymmetry. We show that each potential account is at odds with the spirit of PoP's purely locality-based approach to (1). The results of our investigation of T-to-C are thus more compatible with a revised approach to head-movement sketched out in Chomsky 2014. In

[9] 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.

Chomsky 2014, the subject raises to Spec, TP before C merges. The featural relation of inheritance is involved in determining that only V raises to v*, and T to C.

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 critiques some related aspects of Feature Inheritance in PoP. §5 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 out of vP: assuming the tail of a chain is invisible, this might yield a potential remedy in terms of the relative timing of EA and V-raising. §2.4 shows that V-to-T is possible in low-subject VSO languages and D-to-T is not an available alternative; hence EA raising is not the key to this asymmetry. §2.5 adds that V-raising across negation is also erroneously 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? In fact, D and V-v should be equally available for a relation to T, since both are equally close to it. Minimal

search should obtain an ambiguous result. Yet V-v can raise to T in familiar languages, and D cannot.[10]

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

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

PoP's argument that C-to-T is evidence of a C-T relation prior to EA-raising is weakened by the recurrence of the same phenomenon at this lower point in the clause. A parallel account of V-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-v-to-T movement.[11, 12]

2.3 A timing account?

Suppose the external argument raised to Spec, TP before V-v-to-T movement. Its unpronounced copy in vP would be invisible to T under PoP assumptions, hence solving the problem we presented in §2.2.

(7) a. *EA raises making v unambiguously closest to T*

[$_{TP}$ EA T [$_{vP}$ <EA> [v VP]]

[10] For expository ease we briefly delay discussion of C's potential role in this through the Feature Inheritance hypothesis.

[11] 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 verb-raising across Negation to be discussed in §2.5 below).

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

b. *V-v subsequently raises to T*

[$_{TP}$ EA v+T [$_{vP}$ <EA> v VP]]

This timing provides a potential explanation for why it is true that only V-v, and not D, can raise to T in languages with verb raising. But it gives rise to other difficulties. For one thing, it appears to conflict with PoP's account of T-to-C. Recall that PoP explains the fact that only T (and not D of Spec, TP) raises to C in questions by assuming that the C-T relation is established BEFORE the EA raises from Spec, vP to Spec, TP. Crucially, PoP assumes that D of the subject cannot raise to C because the subject is "not there" (see (3), repeated below).

(3) *PoP'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...]]]

To maintain PoP's timing solution to the original puzzle while at the same time adopting the timing solution we just suggested for (6), we would have to assume that T-to-C and EA-raising *both* precede V-v-to-T. This is highly stipulative, suspiciously counter-cyclic, and incompatible with the fact that V-v-to-T appears to feed T-to-C movement in many languages. We illustrate with Standard Norwegian in (8)a (from Taraldsen 1986:8, who adopts a V-to-C analysis; the representation in (8)b is a slight update).[13]

(8) a. Hvilket spørtsmål skjønte Jens ikke?
which question understood Jens not
'Which question did Jens understand?'

b. [$_{CP}$ Hvilket spørtsmål skjønte+T+C [$_{TP}$ Jens [$_{T'}$ ...ikke...]]]
which question understood Jens not

[13] Underlying word order and structure below T' in Norwegian are irrelevant here. Many aspects of verb second syntax and head-movement in general have been debated in recent years, but we take (8)b to represent a fairly standard analysis of V2, consistent with the PoP approach to inversion that we are critiquing. The issue at hand is how to predict what inverts as opposed to what the precise nature of inversion is.

Another problem arises for *any* appeal to timing in relation to these issues. 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…"

This approach raises questions about whether it is even possible to claim that the C-T relation necessarily precedes EA raising, as PoP does; or to couch a potential account of V-v-to-T in terms of timing. Any timing solution is incompatible with simultaneity, since simultaneity expressly *denies* timing.

Summing up, we have explored the possibility of a timing solution to the puzzle of V-v-to-T across EA in Spec, vP. We have argued that such a solution is incompatible with PoP's approach to T-to-C. Lastly we have argued that, in fact, *any* timing solution is incompatible with the simultaneity of application of operations internal to a phase proposed in Chomsky 2007, 2008.

We turn in §2.4 to a final piece of evidence against such a timing approach, namely that it is founded on the assumption that competition between EA and V-v-to-T are avoided *only if* EA raises to Spec, TP. §2.4 shows that V-v-to-T is the only option in various cases where EA does NOT raise to Spec, TP.

2.4 Low subject VSO languages

VSO clauses in languages such as Standard Arabic, Middle Egyptian, and Xhosa provide evidence that V-v-to-T is possible across subjects that remain low.[14] In such cases no timing approach could explain why D-raising from EA does not compete with V-v-to-T raising.

The morpho-syntax of VSO varies considerably across languages, and there are accordingly a number of approaches to it. In Standard Arabic, where only SVO is accompanied by full subject agreement, Mohammad 2000 argues that a sentence like (9) is the product of V-to-v-to-T across the in situ subject (see also Fassi Fehri 1993, Melebari & Seely 2012, and see Kramer 2009 for an analysis of Middle Egyptian along the same lines). By the logic of PoP's approach to (1), in a VSO language with this derivation, D of the subject should be equidistant from T and the crucial ambiguity should arise.[15]

(9) a. tuhibbu kull-u ʔumm-in ʔibn-a-ha
loves.3SG.FEM every-NOM mother-GEN son-ACC-her
'Every mother loves her son'

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

[TP ...T [vP [every mother] [v' love+v [VP <love> her son]]]]
Minimal search fails to single out V+v alone

[14] It is not our intent to suggest that VSO involves such derivations universally; only that there are languages where it works this way.

[15] A previous version of this paper stated that McCloskey argues for such an analysis of VSO in Irish. While this was true of McCloskey 1996, in McCloskey's more recent work the subject raises out of its base position and the verb lands higher (see McCloskey 2011, 2012, and numerous other works; thanks to Jim McCloskey personal communication for pointing this out to us). Regardless of the eventual landing sites, under any scenario where the verb moves across the subject a version of the same problem arises for the PoP approach (see discussion of (8) in §2.3, and footnote 16 on the possibility of short subject movement in Xhosa VSO clauses).

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 (10)a, the subject remains very low in the structure. The verb raises across it and adjoins to a middle-field inflection below Tense, identified in the Bantu linguistics literature as Mood (see (10)b); (10) adapted from Carstens & Mletshe 2013).[16] The subject cannot surface higher as (10)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 explored 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.

(10) 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 verb-raising is possible in Xhosa across an EA that remains low. As (11) illustrates, this means that prior to V+v-raising, D of the EA is at least as close to Mood as V+v is. But the contrast in (12) shows that D cannot raise instead of V.[17]

[16] 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. Because EA of VSO constructions raises into Spec of the low FocusP in Carstens & Mletshe's account, Xhosa VSO clauses are not incompatible with the PoP proposal that the labeling algorithm LA forces EA raising out of vP. See §3.3 on a contradiction between the LA approach to EA raising and the PoP proposal that Feature Inheritance leaves a copy of the inherited features on the phase head.

[17] (12)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

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

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

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

a. …okokuba ku-be ku-fund-a lo mntwana <kufund> 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) <lo> 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 §2.3) lies in one of the standard diagnostics for V-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.

(13) 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> …]]]

Recall PoP's assertion that "inversion depends on locality independent of category," hence only hierarchical relations are involved in the calculus of closeness relevant to T-to-C movement. If we extend this idea to V-v-to-T movement, then it cannot be relevant whether 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, since it is most local to T (see (14)a). Neg's structural relation to T parallels that of T to C in (3), repeated below. If we assume instead that the subject occupies Spec, NegP at a point before verb-raising, then

---

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.

minimal search should yield an ambiguous result. Neg and D in the subject should compete for raising to T in (14)b, just as PoP argues would be true of T-to-C across a subject in Spec, TP in the hypothetical (2). In neither case is raising of the verb expected to cross Negation.

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

T [$_{NegP}$ Neg [$_{vP}$ EA v…]]

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

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

(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…]]]

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:[18]

(15) 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 (15) demonstrates, T-to-C is obligatory in all direct English WHQ except local subject questions like (15)a, where it is forbidden. (16) 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, CP. This

[18] As noted in footnote 9, stressed *do* is irrelevant here.

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.

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

The problem for deriving (15)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 Pesetsky & Torrego's 2001 solution

There is a simple way around this problem. Suppose with Pesetsky & Torrego 2001, and contra Chomsky 2007, 2008, that subject WH movement in English proceeds to Spec, CP via Spec TP, rather than directly from Spec, vP.[19] Representations of the two cases are sketched in (17) (relative locations of T and the Wh-phrase being crucial).

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

In (17)b, T is clearly closer to C than the *Wh*-phrase is, in marked contrast to (17)a. This assumption allows a potential account of the suppression of T-to-C movement in subject questions like (15)a on the basis of locality, much in the spirit of the PoP attempt to explain why C rather than D raises in (3). That is, in (17)b T raises to C since T is

[19] As two anonymous reviewers point out, there is substantial evidence that subject Wh-movement proceeds not from Spec, TP but from the subject's base position in many languages (see McCloskey 2000, Bošković 2015, Brandi & Cordin 1989, Schneider-Zioga 2007, Erlewine to appear among many others). We take English that-trace effects to be evidence that English subjects do not typically have recourse to a lower extraction site. (i) shows that when such an alternative exists, that-trace effects disappear.

(i) How many people did you say that *(there) were __ in the room?

unambiguously close(est) unlike in (17)a. But in a departure from PoP assumptions, reference to features is crucial to ensure that a WH subject prevents T-to-C in (17) while a non-*Wh*-subject like that in (3) or (15)c does not, assuming subject and T are equidistant from C. Following 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.[20]

This analysis is incompatible with PoP assumptions in several ways. First, as we saw in §2, equidistant candidates for raising should yield completely free alternatives; hence it would be expected that T and (subparts of) the *Wh*-subject would be equally able to raise to C(P) in a subject question. Second, PoP proposes that *Wh*-subjects never raise to surface in local Spec,CP, in contrast with Pesetsky & Torrego's analysis. Hence either some PoP assumptions must be abandoned or some alternative account of this asymmetry must be found.

3.3 An affix-hopping alternative

Two anonymous reviewers suggest an alternative in terms of affix-hopping that we think it is instructive to explore. Consider again (16), a PoP-style derivation under which C merges before the subject raises to Spec, TP. Assume that T raises to C in all questions including the subject question in (16)a. After syntactic movements have applied, the results will be (18).

(16) a. C [$_{TP}$ T [$_{vP}$ which boys [$_{v'}$ eat+v [$_{VP}$ <eat> the pizza]]]]

[20] 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 PoP-inspired suggestion on how it might work.

b. C [$_{TP}$ T [$_{vP}$ which pizza [$_{vP}$ the boys [v [$_{VP}$ <eat> <which pizza>]]]]]

(18) a. [$_{CP}$ Which boys T+C [$_{TP}$ <which boys> <T> [$_{vP}$ eat+v [$_{VP}$ <eat> the pizza]]]]
b. [$_{CP}$ Which pizza T+C [$_{TP}$ the boys <T> [$_{vP}$ eat+v [$_{VP}$ <eat> <which pizza>]]]]

The reviewers suggest that affix-hopping, sensitive only to pronounced material, applies successfully in (18)a (see (19)a). In (18)b, the overt subject *the boys* blocks it, yielding the necessity for *do*-support in C, as shown in (19)b (see Chomsky 1957, Lasnik 1995, Lasnik et al 2000, Halle & Marantz 1993 among many others).

(19) a. [$_{CP}$ Which boys __ [$_{TP}$ <which boys> __ [$_{vP}$ eat+v+PST [$_{VP}$ <eat> the pizza]]]]
b. [$_{CP}$ Which pizza do+T+C [$_{TP}$ the boys __ [$_{vP}$ eat+v [$_{VP}$ <eat> the pizza]]]]

While we see this as a viable alternative to Pesetsky & Torrego's approach, its compatibility with PoP assumptions seems equally questionable. PoP's claim regarding T to C is that "inversion depends on locality independent of category," and it assumes that D of a subject is in principle as likely a candidate as T is for raising to C. The question that arises in connection with (19)b then is why affix hopping should not be equally a matter of pure locality, independent of category? But if this were the case, there would seem to be no reason for D of the subject in (19)b not to be a potential target for affix-hopping.

(19)' b. [$_{CP}$ Which pizza __ [$_{TP}$ the+PST boys __ [$_{vP}$ eat+v [$_{VP}$ <eat> the pizza]]]]

If, on the other hand, the affix-hopping operation is recognized to be categorically sensitive, the solution is viable.[21]

[21] Another question that arises in connection with the affix-hopping approach arises from the assumption in Chomsky 2007, 2008 that when the subject is counter-cyclically merged to Spec, TP after C is merged, C cannot "see" the subject at all. Assuming this, and without advance knowledge of the facts, it would be difficult to predict that the overt subject in (19)b would not be invisible to affix-hopping the same as the moved subject in (19)a. A stipulated distinction between morphology and syntax in this regard might work but seems an uninsightful position, worth avoiding if possible.

Summing up, the subject/non-subject asymmetry for T-to-C movement can be captured under the assumption that a wh-subject in Spec, TP suppresses it a la Pesetsky & Torrego, or by hypthesizing that affix-hopping applies across the subject trace but refuses to attach to D. Under either approach, reference to categorical and other features is critical.

3.4 A speculation re: the treatment of subject questions in PoP

3.4.1 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 of local subject questions in PoP. As two anonymous reviewers note, there are some interesting arguments against this conclusion in the syntactic literature including McCloskey 2000, Agbayani 2000, Ginzburg & Sag 2000, and Pesetsky & Torrego 2001); but some of these are neutralized by the approach in PoP, under which TP of a subject question inherits all the properties of an interrogative CP. For the sake of argument, we assume the proposal is viable and explore its potential consequences for T-to-C movement.

As noted in §1 and §2.3, PoP follows Chomsky 2007, 2008 and Richards 2007 in assuming that T obtains features from C. In addition to the unvalued phi-features these works discuss, PoP adds 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 C's features including Q are therefore inherited by T in a bundle. 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). We illustrate in (20):

(20) *Feature Inheritance in PoP*: a. $C_{[Q, uPhi...]}$ T → b. $C_{[Q1, uPhi1...]}$ $T_{[Q2, uPhi2...]}$

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 (21)a).[22] In contrast, it is the Q feature on C that agrees with a non-subject WH, labeling CP as QP (see (21)b). In this case, sharing of prominent features between the raised EA and T labels their containing phrase PhiP. Once these labelings have taken place, no further movement is possible for the XPs involved because under PoP assumptions, *criterial freezing* accompanies successful sharing of a prominent feature (see discussion in PoP: 47, citing Rizzi personal communication; see also Rizzi 2013; Rizzi & Shlonsky 2007). Hence the subject WH cannot raise to Spec, CP in (21)a.

(21) a. $C_{[Q1, uPhi1...]}$ [$_{QP}$ [How many mechanics] [$T_{[Q2, \sout{uPhi2}...]}$ fixed the cars?]]
b. [$_{QP}$ [How many cars] $C_{[Q1, uPhi1...]}$ [$_{PhiP}$ [the mechanics] $T_{[Q2, \sout{uPhi2}...]}$ ...]]

PoP does not address the question of why T must raise to C in the circumstances where this is required. The facts are puzzling on PoP assumptions. Feature Inheritance ensures that C and T's features are largely identical. T therefore has nothing that C does not also have apart from their differing categories, and PoP assumptions rule out reference to category. Furthermore, C needs only its own Q-feature in order to interact with a non-subject WH-phrase for labeling in a case like (21)b, or to participate in selection.

[22] 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 (21)a. Also unclear is why heads are immune to the "freezing effect," thus why, after TP is labeled PhiP, a subject cannot raise from Spec, TP but T can raise to C. We leave these questions aside.

In prior treatments including that of Pesetsky & Torrego mentioned above, C has an additional 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. It is thus somewhat mysterious 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 (21)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.

3.4.2 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: retention of Qs at both CP and TP levels yields a deviant result. Where Q does not pair with a wh-phrase, it yields a Yes/No question (see Epstein, Seely, and Kitahara 2015). Therefore a structure like (22)a,b present gibberish combinations of Yes/No and wh-questions.

(22) a. *[$_{QP1}$ $C_{[Q1]}$ [$_{QP2}$ [How many mechanics] $T_{[Q2]}$ fixed the cars?]]
b. *[$_{QP1}$ How many cars $C_{[Q1]}$ [$_{QP2}$ the mechanics $T_{[Q2]}$ fix <wh...> ?]]

To rectify this problem, in a subject question like (22)a, C bearing Q1 would lower to T (see (23)). [23] In Yes/No questions, and in a non-subject question like (22)b, where the interrogative operator moves to Spec, CP, T must raise to place its copy of Q in C (see (24)).[24]

[23] 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.
[24] Given that labels are determined by a late algorithm, a possible alternative interpretation is that T can't raise to C in a subject question because this would interfere

(23) <$C_{[Q1, uPhi1...]}$> [$_{QP}$ [How many mechanics] [$C_{[Q1, uPhi1...]}$+$T_{[Q2, \sout{uPhi2}...]}$ fixed the cars?]]

(24) [$_{QP}$ [How many cars] $T_{[Q2, \sout{uPhi2}...]}$+$C_{[Q1, uPhi1...]}$ [$_{PhiP}$ [the mech…] <$T_{[Q2, \sout{uPhi2}...]}$> …]]

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

Like Pesetsky & Torrego's approach to T-to-C movement and like the affix-hopping possibility, the hypothesis we have sketched out here ties head-movement to features of the moving element and its landing site. Absent Q features, T-to-C (and its inversion, C lowering to T) do not happen at all.

In sum, there are several promising options for capturing the distribution of T-to-C in matrix questions: raising the subject via Spec, TP; affix-hopping from C down to V+v in subject questions, or relocating an extraneous duplicate Q to the projection containing interrogative material..[25] It lies outside this paper's scope to choose between them. What matters for our purposes is that all the possibilities take head-movement to be driven and constrained by factors other than pure locality, as we also argued to be the case for V-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.

4. Unsolved mysteries

PoP's proposals about Feature Inheritance raise some important conceptual questions. While FI is not head movement, it is an important sub-case of head-head relations

---

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.

[25] To repeat, there are many technical details and questions left unaddressed, among them why T does not raise to C in indirect questions (thanks to an anonymous reviewer for pointing this out). However, we believe that any analysis exploiting PoP's core idea will face the problems outlined here.

relevant to the material in this paper. We therefore offer a few observations on its implications.

Under the interpretation of Feature Inheritance as copying, it is not clear why the unvalued phi-features 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."

Chomsky 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 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.[26]

One more unsolved mystery concerns PoP's proposal that EA raises from vP because otherwise vP cannot be labeled. PoP claims that with EA in situ, vP constitutes an [XP, YP] configuration in which there is no shared prominent feature to form the label. The subject therefore must move. This proposal is articulated in PoP, 44: "Since β is of the form {XP, YP} it is not labeled by LA…if EA raises to surface subject – SPEC-T,

[26] 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.

using the term SPEC here and below just for exposition – then β will be labeled v*, as required" ((25) = PoP's (17)).

(25) T [$_{\beta}$ (EA) [v* [V IA]]]

Consider now the landing site of the subject, depicted in ((26)a (= PoP's (16)), where TP = α). As noted above, PoP proposes that when the EA raises to Spec, TP there is labeling of TP as PhiP because "NP and TP share prominent features, namely phi-features" (PoP: 45). *Criterial freezing* of the subject results (recall discussion of the subject *Wh*-phrase in (21) of §3.3.1).

(26) a. [$_{C}$ C [$_{\alpha}$ NP TP]] → b. [$_{C}$ C [$_{PhiP}$ NP TP]] *Subject criterially frozen*

But recall PoP's proposal that Feature Inheritance is feature copying, leaving a copy on the phase head as discussed in §3.3.1. PoP argues that "individual features cannot be moved independently of the bundle to which they belong" (PoP:47) and that C must keep a copy of Q for selection and labeling. If these proposals generalize to the phase head v*, then v* will necessarily retain a copy of uPhi in spite of v*-to-V Feature Inheritance (see (27)).

(27) *Feature Inheritance as copying: v* retains uPhi*

a. [v*$_{uPhi...}$ [$_{VP}$ V...]] → b. [v*$_{uPhi1...}$ [$_{VP}$ V $_{uPhi2...}$ ...]]

Since EA also has phi-features, vP is predicted to be labeled PhiP by feature-sharing between EA and v* (see (28)). As noted above, once labeling has taken place, no further movement is possible. Raising of the subject to Spec, TP is erroneously ruled out.

(28) *Since EA and v* both have Phi under this approach to FI, shared prominent feature should label vP as PhiP, incorrectly ruling out EA-raising*

a. [$_{\alpha}$ EA$_{\phi}$ v*$_{uPhi1...}$ [$_{VP}$...]] → b. [$_{PhiP}$ EA$_{\phi}$ v*$_{\sout{uPhi1,}...}$ [$_{VP}$ ...]]

The labeling account of EA raising and the copying approach to Feature Inheritance thus seem to be incompatible components of the PoP approach.

5. Concluding remarks

PoP's explanatory achievements 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 Pesetsky 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 (and Chomsky's 2014 retreat from PoP's pure locality approach is well-warranted). Sometimes *ad hoc* is just what we need.[27]

In addition to approaching T-to-C in terms of locality only, PoP presents some new proposals regarding Feature Inheritance as copying, and labeling as the driving force for XP-movement. While these have many interesting implications, we have pointed out

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

some inconsistencies and problems that must be resolved in order for them to be workable.

References

Agbayani, Brian. 2000. Wh-subjects in English and the vacuous movement hypothesis. *Linguistic Inquiry* 31. 703–713.

Brandi, Luciana & Patricia Cordin. 1989. Two Italian dialects and the null subject parameter. In Osvaldo Jaggli & Kenneth Safir (eds.), The null subject parameter, 111–142. Dordrecht: Kluwer.

Bošković, Željko. 2015. On the timing of labeling: deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling. Ms., University of Connecticut.

Carstens, Vicki & Loyiso Mletshe. 2015. Radical defectivity: implications of Xhosa expletive constructions. *Linguistic Inquiry* 46.2:187-242

Chomsky, Noam. 1957. *Syntactic Structures.* Paris: Mouton.

Chomsky, Noam. 1968. *Language and Mind,* Harcourt, Brace, Jovanovitch, New York.

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

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

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

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

Chomsky, Noam. 2014. Problems of Projection: Extensions. Ms, MIT.

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

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*, 253-270. In Oxford Studies in Theoretical Linguistics D. Adger and H. Borer series eds, Oxford,.

Epstein, Samuel, H. Kitahara, T. D. Seely. 2015. "'What do you wonder' is not syntactic," in Epstein, Kitahara, Seely. *Explorations in Maximizing Syntactic Minimization*. Routledge Leading Linguists Series, Routledge

Erlewine, Michael. to appear. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language and Linguistic Theory*.

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words.* Dordrecht: Kluwer Publications.

Ginzburg, Jonathan & Ivan A. Sag. 2000. Interrogative Investigations. Palo Alto: CSLI Publications.

Halle, Morris and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In K. Hale and S. J. Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 1-52. Cambridge, MA: MIT Press.

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

Kramer, Ruth. 2009. VSO and SVO word order in Middle Egyptian. In Charles G. Häberl (ed.), *Afroasiatic studies in memory of Robert Hetzron: proceedings of the 35th*

annual North American conference on Asiatic linguistics, 92-147. Cambridge: Cambridge Scholars Press.

Lasnik, Howard. 1995b. Verbal morphology: *Syntactic Structures* meets the Minimalist Program. In H. Campos and P. Kempchinsky(eds.), *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero*, 251-275. Washington, D.C.: Georgetown University Press.

Lasnik, Howard (with Marcela Depiante and Arthur Stepanov). 2000. *Syntactic Structures Revisited.* Cambridge: MIT Press.

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

McCloskey, James. 2000. Quantifier float and Wh-movement in an Irish English. *Linguistic Inquiry* 31(1). 57–84.

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

McCloskey, James. 2011. The Shape of Irish Clauses,' in Andrew Carnie (ed.) *Formal Approaches to Celtic Linguistics*, 143-178. Newcastle upon Tyne: Cambridge Scholars Publishing.

McCloskey, James. 2012. Irish Existentials in Context. To appear in *Syntax*.

Melebari, Ala'a and T. Daniel Seely. 2012. Agreement without 'exception' in Standard Arabic? Presentation at the Illinois Symposium on Semitic Linguistics, University of Illinois, Urbana-Champaign.

Mohammad, Mohammad A. 2000. *Word order, agreement, and pronominalization in Standard and Palestinian Arabic*. Amsterdam and Philadelphia: John Benjamins.

Ouali, Hamid. 2008. On C-to-T ϕ-feature transfer: the nature of agreement and anti-agreement in Berber. In: D'Alessandro, Rl, Hrafnbjargarson, G.H., Fischer, S. (eds.), *Agreement Restrictions*, 115-18. Berlin: Mouton de Gruyter.

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

Piatelli-Palmarini, M. (ed.), 1980. *Language and Learning: The Debate between Jean Piaget and Noam Chomsky.* Cambridge, MA: Harvard University Press.

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

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

Rizzi, Luigi. 2013. Cartography, criteria, and labeling. Ms, Università di Sienna.

Rizzi, Luigi and Ur Shlonsky. 2007. Strategies for subject extraction. In U. Sauerland and H.-M. Gartner eds, *Interfaces + recursion = language?; Chomsky's minimalism and the view from syntax-semantics*, 115-160. Studies in Generative Grammar 89. Berlin: Mouton de Gruyter.

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

Schneider-Zioga, Patricia. 2007. Anti-agreement, anti-locality, and minimality. *Natural language & Linguistic Theory* 25:403-446.

.Taraldsen, Knut Tarald. 1986. On verb-second and the functional content of syntactic categories. In Haider, Hubert and Martin Prinzhorn (eds.), *Verb second phenomena in the Germanic languages*, 7-25. Dordrecht: Foris.