In (*do-*)support of phrasal auxiliary movement

July 4, 2018

It has been assumed that auxiliaries move to T in English, based on facts like in (1), where the auxiliary verb precedes sentential negation.

(1)   Gromit has not eaten any cheese.

Generally, it has been assumed that auxiliary movement is uniform, and there is string-vacuous aux-movement even in the absence of negation (Emonds 1970). In this paper, I argue for a view in which the affirmative counterpart to (1), as in (2), allows the auxiliary to remain low in its base position (Baker 1991).

(2)   Gromit has eaten some cheese.

On this alternative view, head movement to T is optional unless there is an intervening maximal projection between the two heads, in which case movement becomes obligatory.

If we assume that auxiliary movement is feature driven, this view on head movement also challenges standard assumptions about movement, on which EPP-features on a given head uniformly trigger movement. I argue for a more general Merge-based perspective on the EPP – the Extended EPP, or E2P2 – that permits this new description and explains other facts about the English auxiliary system and *do*-support as a reflex of this mechanism.

Specifically, I will challenge the stipulation that the EPP is exclusively satisfied by specifier-creation and argue that complements may satisfy the EPP as well. On this view EPP properties can be satisfied if a probe merges in any way with its goal. If its goal happens to be the head that it selected for, the EPP property is satisfied vacuously without movement. Otherwise, the goal must move to the probe.

This account of auxiliary movement builds on Matushansky (2006) and Harizanov and Gribanova (2018) by allowing this type of head movement to occur in the syntax. Syntactic head movement here is structure building and non-local; it does not obey the head movement constraint, allowing an auxiliary to move over a NegP to satisfy a higher EPP property on T. This analysis also offers a principled explanation of *do*-support if we adopt the possibility of pure feature movement (Chomsky 1995, Yuan 2015) from v to T.

On this proposal, the fact that negation triggers auxiliary movement and *do*-support is a direct consequence of the theory; any non-verbal maximal projection in the domain of T that dominates the nearest verb will trigger movement of that verb to T because such a projection fails to satisfy the relevant EPP property[1]. If no auxiliary verb is present, v can satisfy this EPP property by moving to T, where it becomes pronounced as *do*. V is stranded in situ because its adjunction to v is a post-syntactic word-building process, which occurs *after* auxiliary movement, causing v to move on its own.

This analysis derives additional support from the distribution of sentential *not* in finite and non-finite clauses. The optionality in *not*-placement in non-finite clauses contrasts with obligatory post-auxiliary placement in finite clauses.

(3)   a. For Gromit to not finish Wallace's cheese would be a shame.
      b. For Gromit not to finish Wallace's cheese would be a shame.
      c. Gromit has not finished Wallace's cheese.
      d. *Gromit not has finished Wallace's cheese.

I argue that in (3a,b), negation may either remain in situ or undergo head movement to T in response to an extended EPP property on T for sentential negation. In finite contexts, these two properties, namely the EPP for verbal elements and the EPP for sentential negation, interact via normal locality conditions on probing to yield the correct word order.

1   Auxiliaries and E2P2

1.1   Facts about the English auxiliary system

We know from Chomsky (1957) and subsequent work the following facts about the English auxiliary system and negation: (I) auxiliary verbs move to T,

(4)   a. Gromit has not eaten cheese recently.
      b. *Gromit not has eaten cheese recently.
      c. Gromit is not eating cheese at the moment.
      d. *Gromit not is eating cheese at the moment.

(II) multiple auxiliaries follow a fixed hierarchy: *modal - have - bePROG - bePASS - main verb*, and (III) when the vP is separated from T by negation or emphasis, and there is no auxiliary that can move to T, the otherwise blocked auxiliary *do* is used.

(5)   a. Gromit did not eat cheese at the scene of the crime.

---

[1]Adverbs are claimed not to trigger auxiliary movement or *do*-support because they are invisible to selection. They merge within a phrase and do not affect labeling along the clausal spine.

    b. No, Gromit *did* eat cheese at the scene of the crime!

    c. *Gromit not ate cheese at the scene of the crime.

The examples in (4) constitute evidence of auxiliary movement based on the assumption that sentential negation is generated between T and auxiliaries. We know this because of data like (6). When auxiliaries are forced to remain in situ, negation surfaces above them (unless it is constituent negation, as in (6b), as evidenced by the polarity of the tag question).

(6)   a. Gromit will not have liked the cheese, will he/*won't he?

       b. Gromit will have not liked the cheese, won't he/*will he?

       c. For Gromit to not have stolen Wallace's cheese...

While the above examples are evidence that auxiliaries move to T across negation, the claim that auxiliaries move string vacuously in the absence of such an intermediate projection is not a logical necessity (without further assumptions). If we allow the possibility that auxiliaries remain in situ in the absence of negation, the paradigm we need to account for is represented in (7)[2].

(7)   a. $[_{TP}$ Gromit **T** $[_{vP}$ v $[_{VP}$ walked $]]]$

       b. $[_{TP}$ Gromit **T** $[_{AuxP}$ has $[_{vP}$ v $[_{VP}$ walked $]]]]$

       c. $[_{TP}$ Gromit **did** $[_{NegP}$ not $[_{vP}$ v $[_{VP}$ walk $]]]]$

       d. $[_{TP}$ Gromit **has** $[_{NegP}$ not $[_{AuxP}$ t $[_{vP}$ v $[_{VP}$ walked $]]]]]$

Allowing auxiliaries to remain in situ in (7b) shows that (7a,c) parallel (7b,d). In the absence of negation, auxiliaries and main verbs may remain in their base positions. When negation intervenes, something needs to happen. Either the auxiliary moves to T or we get *do*-support.

1.2 Proposal

To formalize this idea, I propose that there is a property of T that can be satisfied just in case an auxiliary or a verb-like element heads T's sister, or moves to T. This property is a probe with a generalized EPP property, whose goal is a verb-like constituent. This generalized EPP property cares only about whether T merges with an element that bears the relevant feature. If that feature is a property of the head of T's complement, this EPP property is satisfied by virtue of the fact that T selected for and merged with its complement. If that feature is not on the head of T's complement, it must move to T as a specifier.

---

[2]The trees in this paper are drawn to focus on auxiliary movement, not subject movement. All of the trees have subjects in Spec TP that I assume have moved there from Spec vP. This may make some of the derivations look counter cyclic, because it looks like the subject has moved to Spec TP before the auxiliary has moved to T. This countercyclicity is not part of my proposal, and is not a necessary component of the derivation.

I furthermore adopt a view of head movement like that proposed by Matushansky (2006), in which a moving head first forms a specifier by normal cyclic merge to a higher head, followed by an obligatory process of *m-merge* that forms an indivisible unit from the two heads. With these ideas as a background, we can imagine that head movement could also satisfy an EPP property via specifier creation.

Step 1:

```
          TP
         /  \
     Gromit   T′
             /  \
          has    T′
                /  \
               T    NegP
                   /  \
                 not    vP
                       /  \
                   t_has  walked
```

Step 2:

```
          TP
         /  \
     Gromit   T′
             /  \
          has    NegP
                /  \
              not    vP
                    /  \
                t_has  walked
```
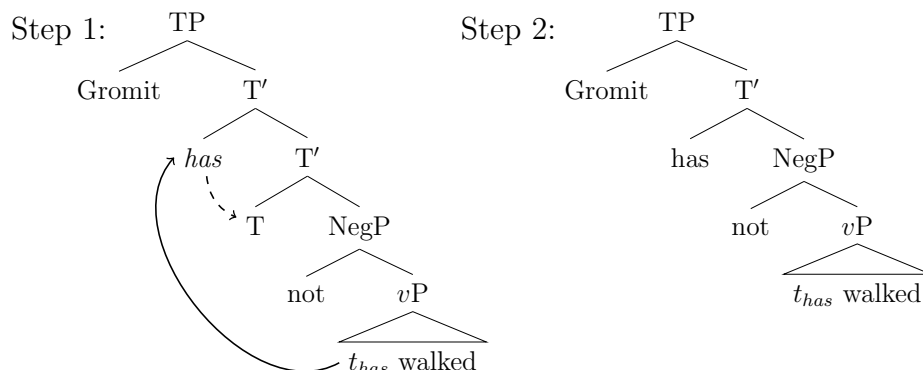
Figure 1: Schematic of head movement for Matushansky (2006).

If we accept that auxiliary movement can be syntactic (Matushansky (2006), Harizanov and Gribanova (2018)), this extension of the EPP follows from familiar notions about MERGE and projection labeling. As formalized by Chomsky (1995), MERGE takes two elements and forms an unordered set from the pair. Standard definitions of the EPP stipulate a preference for checking features via specifier creation, but in principle there is no reason to prefer feature-checking via merging as a specifier to merging as a complement. Whatever features are projected onto a complement phrase should be equally visible to a head as those on a specifier since the merge statements should be identical in either case. If a phrase projects the features of its head, we then predict that complement heads should satisfy EPP properties as well as specifiers[3].

I will call this generalized view of the EPP the *Extended EPP* (henceforth E2P2), which is intended to replace previous versions of the EPP. I propose that in addition to a φ-probe, T also has a probe that searches for verb-like elements (perhaps to host

---

[3]This proposal is reminiscent of discussion about Comp-to-Spec Anti-locality (Abels 2003). Given a head X, Comp-to-Spec Anti-locality rules out movement from the complement of X to the specifier of X on the theoretical grounds that such a movement step creates no new relationship between X and its complement, and is therefore redundant. By extension, given that heads select for the category of their complement's *head*, we would not expect feature-driven movement of a head Y to its selecting head X to ever be obligatory because such a derivational move would be syntactically vacuous. While I leave aside the discussion of whether to rule out such movement altogether, there shouldn't in principle be any syntactic motivation for the EPP to *require* movement of a constituent with which the probe has already merged.

T's φ and tense features). These two probes have an E2P2 property, that can be satisfied in the following ways (where x^v is a verbal element, and x$^\varphi$ is a φ-goal):

1. via movement, i.e. x^v/x$^\varphi$ moves to T

2. if x^v/x$^\varphi$ heads T's sister

With Chomsky (1995), Alexiadou and Anagnostopoulou (1998), Pesetsky and Torrego (2004), and others, I will propose that agreement is a precursor to satisfaction of the EPP.
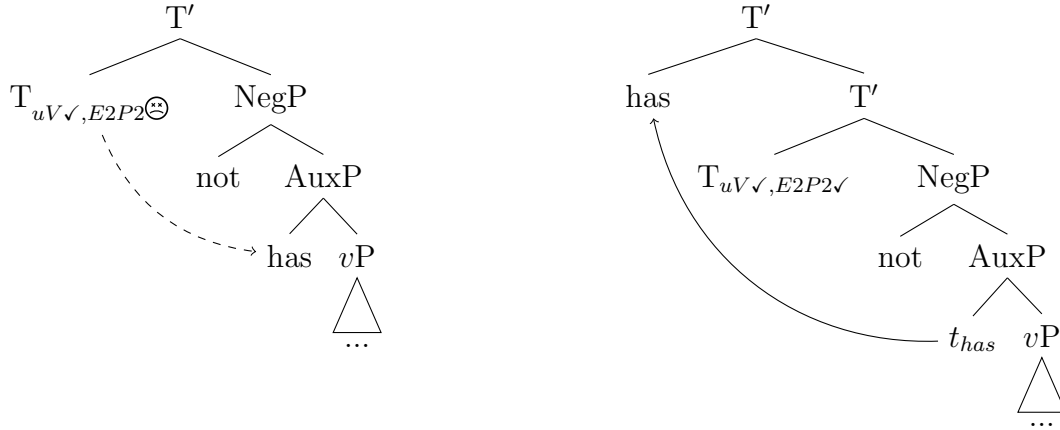
Since a goal for the φ-probe on T will generally be found inside vP and will not be the head of T's complement, this formalization of the EPP predicts that φ-goals in English must *always* move to Spec TP, just like the original EPP predicts. On the other hand, goals for T's other probe include auxiliary verbs and v, which head T's complement phrase in affirmative clauses. They are therefore able to satisfy T's E2P2 property for verbal elements via complementation if no non-verbal head comes between them and T. The presence of a NegP above the verbal domain will force auxiliary movement because negation is assumed here not to be a viable goal for this probe.

We now apply E2P2 to the structures in (7). In (7a) (repeated in the tree below), v heads T's sister. Assuming that v is a verbal element, T finds and agrees with v. Since v heads T's sister, it can also satisfy E2P2 in situ. At PF, I assume the root combines with v, yielding an inflected verb.
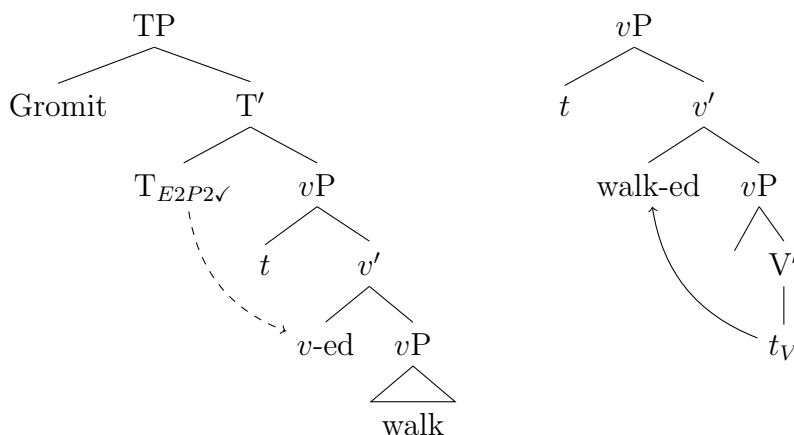
Similarly if T's sister is an auxiliary verb, T will agree with the auxiliary, which can satisfy T's E2P2 property in situ. However in (7c,d) (repeated below), T's sister is a NegP rather than a verbal projection[4]. Therefore, T's E2P2 property for verbal elements is unsatisfied unless a verbal element moves to T.
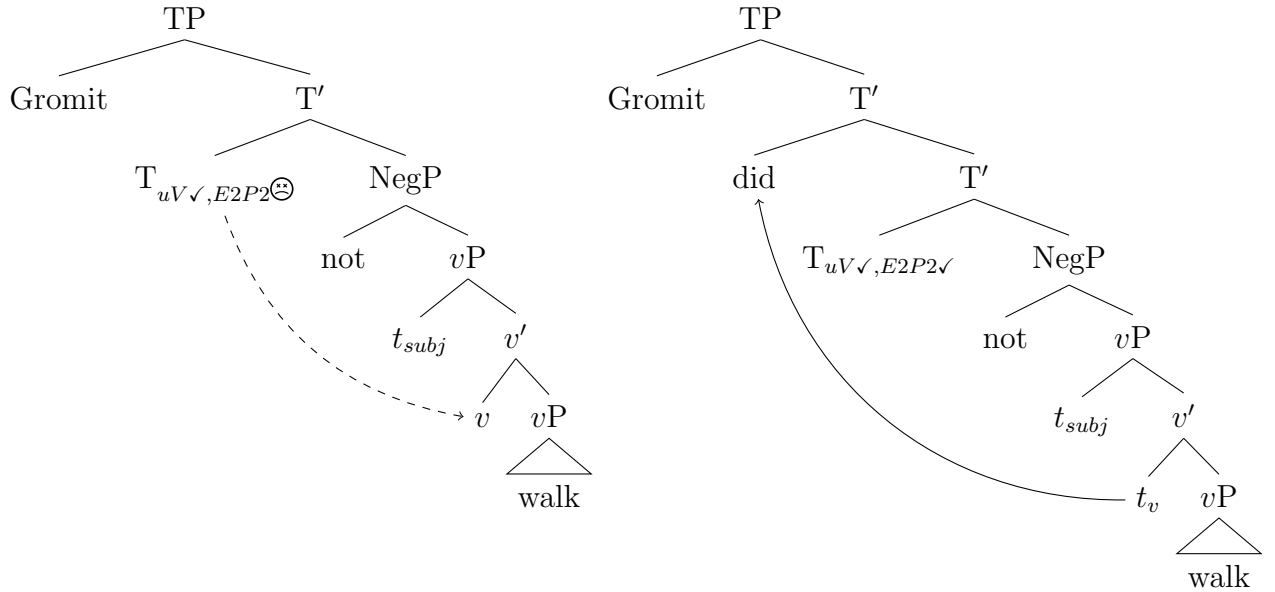
---

[4]I assume that negation is a maximal projection on the clausal spine. In contrast I assume that other adverbs merge as specifiers within a phrase. This assumption is necessary to explain why adverbs don't trigger *do*-support in sentences like *Gromit quickly ate the cheese.*

T′
  T$_{uV\checkmark,E2P2}$☹   NegP
                      not   AuxP
                          has vP
                                 ...

T′
  has   T′
      T$_{uV\checkmark,E2P2\checkmark}$ NegP
                            not   AuxP
                                t_{has} vP
                                          ...

If an auxiliary verb is available, it moves to T. If there is no auxiliary verb available, I propose that v moves to T (featural movement (Chomsky 1995, Yuan 2015)), and the following rule applies.

Do-support (inspired by Bobaljik 1995): v is pronounced as *do* when separated from vP by a maximal projection

Harizanov and Gribanova have argued for a distinction between phrase-building syntactic head movement and word-building postsyntactic amalgamation. I take V to v movement to be a local word-building operation involving a category-less root and a verbal head, contrary to auxiliary movement, which appears to be non-local (e.g. crosses negation, can move to C without morphological consequences, etc.). If this is the case, then V-to-v raising must be post-syntactic. If V-to-v raising is post-syntactic, then when no auxiliary is present, T will attract v on its own. It is under these circumstances, I propose, that it is realized as auxiliary *do*. On this proposal, the difference between modern English and French (and older variants of English) in whether main verbs are able to raise to T could be traced to whether V to v movement is regular syntactic phrase-building movement, or post-syntactic amalgamation.

Tree 1:

TP
- Gromit
- T′
  - $T_{uV\checkmark,E2P2}$☹
  - NegP
    - not
    - vP
      - t_{subj}
      - v'
        - v
        - vP
          - walk

Tree 2:

TP
- Gromit
- T′
  - did
  - T′
    - $T_{uV\checkmark,E2P2\checkmark}$
    - NegP
      - not
      - vP
        - t_{subj}
        - v'
          - t_v
          - vP
            - walk

In summary, the modification of the EPP that I propose can be distilled to the following statement: **An EPP (henceforth E2P2) property on a head H for a feature x can be satisfied if H merges with a bearer of x.** Assuming that a phrase projects the features of its head, E2P2 can be satisfied by either its complement's head or its specifier[5].

The predictions of this theory will depend in part on whether we adopt economy conditions or an Anti-locality constraint on movement (Abels 2003, Erlewine 2016). The central claim of this theory is that a complement head has two options for satisfying an E2P2 property, one which involves movement and another which does not. If we imagine that vacuous movement is disallowed for reasons of economy or Anti-locality, we predict only one option for satisfying E2P2 in any given derivation: 1) complement heads *must* remain in situ, 2) non-complement heads must move.

However if we remove these constraints on movement, we would predict optionality for complement heads, namely they can either remain in situ or move. Evidence from the distribution of negation may argue in favor of this view, as will be discussed in section 2.

---

[5]vP ellipsis poses an apparent puzzle to this theory. Assuming that vP ellipsis is the deletion of all structure below a T head, and adopting the proposal that auxiliaries can remain in situ in the absence of negation, we predict that auxiliaries should be deletable as well, contrary to fact. If this theory is correct, it may suggest that ellipsis sites are introduced by a null head bearing the e-feature proposed by Merchant (2001). In this case, the null head would be introduced between T and the elided material (namely the rest of the verb phrase), which makes satisfaction of E2P2 for verbal elements contingent on movement to T.

2   Negation in non-finite clauses

With two options to satisfy E2P2, we might expect that if we had an overt T head, we would see optionality in word order between the complement head and the T head. The complement head can either remain in situ and surface to the right of the T head, or can move cyclically to a specifier of T and m-merge with T, thus surfacing to the left of T.

TP
  subj.   T′
      X_{+E2P2} vP
            aux ...

TP
  subj.   T′
      aux-X_{+E2P2} vP
            t_{aux} ...
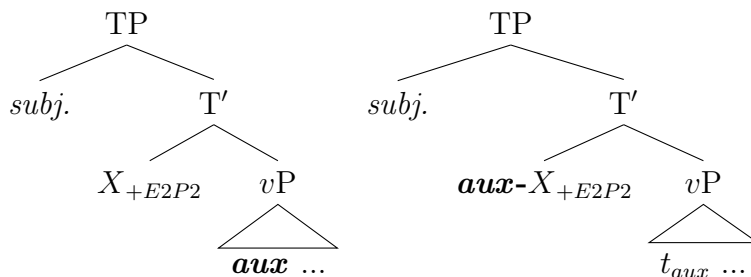
In English, non-finite T is overt (*to*). However, non-finite T does not appear to have an E2P2 property attracting verbal elements because it is not an inflectional affix. Therefore auxiliaries will not move to T, and we won't be able to use non-finite T to check for optional auxiliary movement. However, we do see word order optionality with sentential negation.

(8)   a. For Gromit to not finish Wallace's cheese would be a shame.
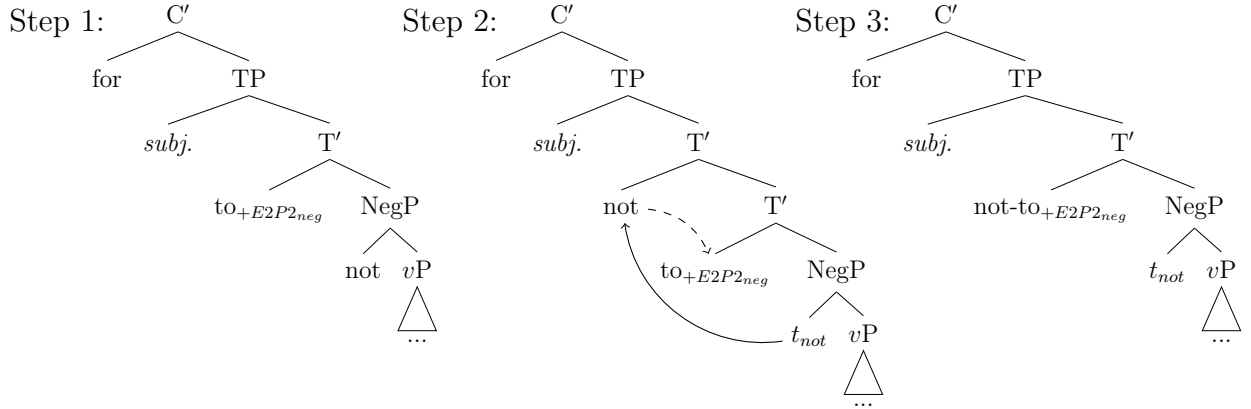   b. For Gromit not to finish Wallace's cheese would be a shame.

The *not* in both (8a,b) can take sentential scope (more on this in the appendix), and speakers cannot discern a difference in meaning between them. Some speakers prefer the order *not to* to *to not*, but produce both freely (see appendix for corpus data). I argue that we can model these facts with an E2P2 property as well based on novel adverb data.

When *not* surfaces above *to*, we see in (9f) that they must be linearly adjacent. No adverb may come between *not* and *to* in this configuration.

(9)   a. For Gromit **to not** occasionally finish the cheese would be a shame.
      b. For Gromit occasionally **to not** finish the cheese would be a shame.
      c. For Gromit **to** occasionally **not** finish the cheese would be a shame.
      d. For Gromit **not to** occasionally finish the cheese would be a shame.
      e. For Gromit occasionally **not to** finish the cheese would be a shame.
      f. *For Gromit **not** occasionally **to** finish the cheese would be a shame.
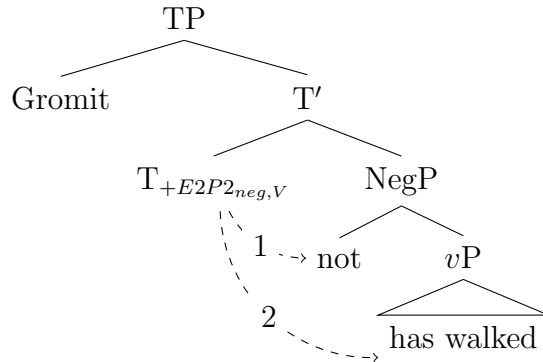
An E2P2 property on T for sentential negation could explain both the word order optionality and the indivisibility of *not to*. If non-finite T merges with a NegP,

not is the head of T's sister and could satisfy an E2P2 property for negation on T. Alternatively it can move to T, first forming a specifier and then m-merging to form an indivisible unit *not-to*.

Step 1:

```
        C′
       /  \
     for   TP
          /  \
       subj.  T′
             /  \
    to_{+E2P2_{neg}}  NegP
                      /  \
                   not    vP
                          △
                          ...
```

Step 2:

```
        C′
       /  \
     for   TP
          /  \
       subj.  T′
             /  \
          not    T′
                /  \
   to_{+E2P2_{neg}}  NegP
                     /  \
                 t_{not}  vP
                          △
                          ...
```

Step 3:

```
        C′
       /  \
     for   TP
          /  \
       subj.  T′
             /  \
  not-to_{+E2P2_{neg}}  NegP
                        /  \
                    t_{not}  vP
                             △
                             ...
```

Given this view of negation and auxiliary movement, we might wonder: if *not* can move to T, why do we see the order *not to* but not *not have/be/modal*? To account for this contrast, I propose that the drivers of head movement behave like the drivers of other types of syntactic movement in that they obey locality conditions. Combining a locality based view of probing with syntactic m-merge, we will see that in finite clauses, no matter how E2P2 is satisfied, the resulting order is always predicted to be *have/be/modal not*.

When T has an E2P2 property for negation as well as an E2P2 property for verbal elements, and there is negation in the structure, there are two ways for both E2P2 properties to be satisfied: 1) *not* remains in NegP and *aux* moves to T, 2) both *not* and *aux* move to T. The first option readily predicts the order *aux not*. To get the order *aux not* from option two, we can adopt a view of probing that is based on locality.

```
                    TP
                   /  \
             Gromit    T′
                      /  \
     T_{+E2P2_{neg,V}}    NegP
                          /  \
                  1 ⇢ not     vP
                             /  \
                  2 ⇢   has walked
```

The other crucial assumption we need is that m-merge applies cyclically in the syntax as soon as it can. This assumption is compatible with Matushansky's original proposal, in which m-merge applies early enough in the derivation to cause future

9

movement operations to target the newly formed complex head rather than a sub-component of it. With these two assumptions we see in Figure 4 how they work together to yield the correct word order.
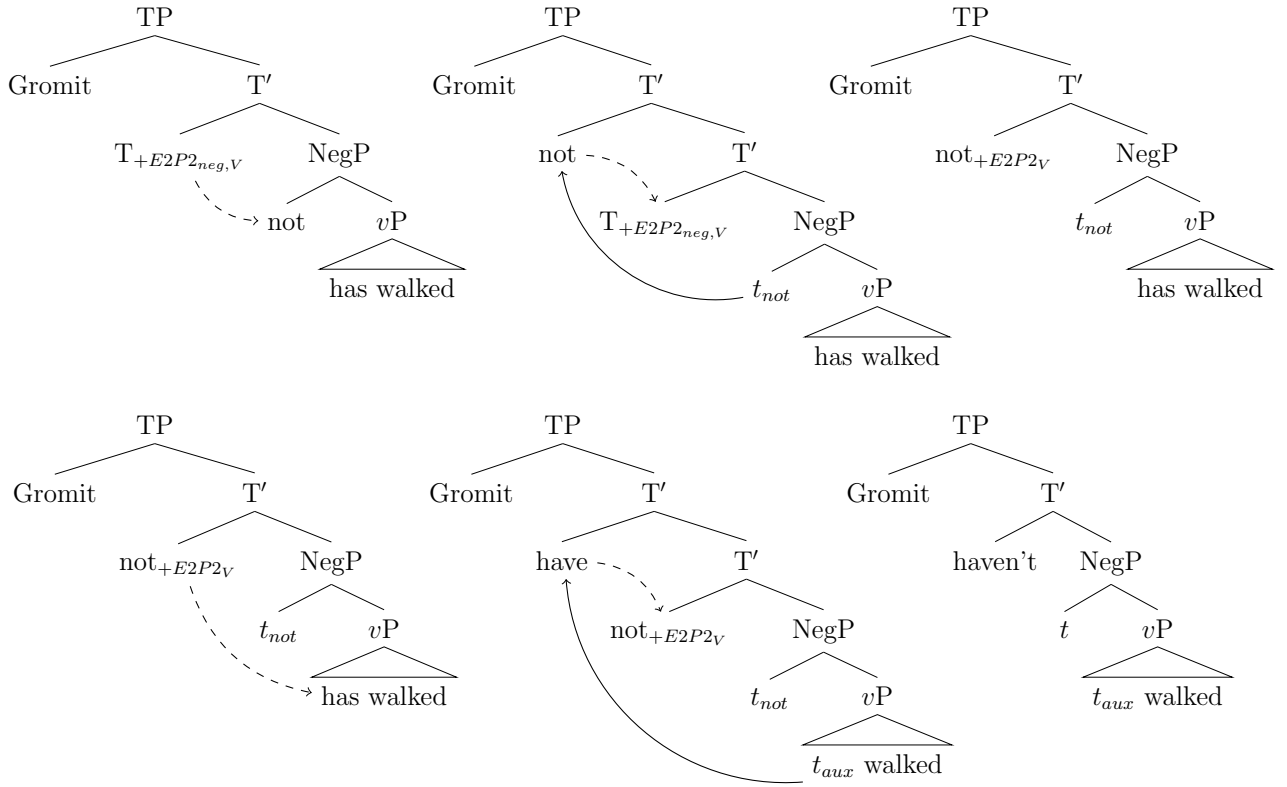

Figure 2: T first probes for and finds *not*, which may move to T (illustrated in the first row). Then T probes for and finds the auxiliary (illustrated in the second row), by which point *not* has formed a complex head with T. The auxiliary then moves cyclically to T, surfacing above *not*.

As seen in the above figure, *not* first moves from NegP up to a specifier of T and then m-merges with T. Afterward, *has* moves up to form a specifier above T (which is now *not*) and then m-merges to it. I assume that T sees and moves *not* before it sees and moves *has*. There is no tucking in (Richards 2005) because the moving constituent m-merges and forms a complex head before the next constituent can move up to the T domain. Therefore, the second constituent only has the option of forming a specifier above the complex head and we always get cyclic movement. If there were no m-merge, or if m-merge occurred later in the derivation, we might expect tucking in. However with these assumptions, the only possible order is *has not*.

In non-finite clauses, since *to* is an instance of T from the start of derivation, and does not reach T via movement, if *not* moves, the order is always *not to*. Modals behave similarly to auxiliaries in that the only possible order is *modal not*. If this

analysis is correct, it is further evidence that modals are not instances of T, but rather move to T from below negation.

These results suggest an explanation for the contraction *n't* as well. I propose that the contraction *n't* is the result of an auxiliary or modal m-merging with a *not* in T. Evidence for this proposal is that *not* can only adjoin to verbs in T to form *n't*. *Not* cannot adjoin to anything else to form *n't*.

(10)   a. I can't eat the cheese.
        b. I won't...
        c. I haven't...
        d. I don't...
        e. *I often tryn't to pay attention. (meaning *I often try not to pay attention.*)
        f. *For hern't to understand would be a shame. (meaning *For her not to understand...*)
        g. *It would be a shame to haven't seen the movie.

I propose that every case of *n't* is a case where *not* has moved to T. The fact that *not* only optionally becomes *n't* parallels the fact that *not* optionally moves[6].

In summary, I have shown that we can capture facts about negation in non-finite clauses with a neg probe that has an E2P2 property, if we allow movement to apply without economy or Anti-locality constraints. This sort of optionality is a bit puzzling. We might wonder why negation would ever choose to move to T if an equally viable, less involved derivation exists in which negation remains in situ. At this point we can either accept that parts of the grammar allow for true optionality, and do not have economy or Anti-locality constraints, or we can posit the existence of a null intervening head between T and negation in the *not to* examples to get the same effect.

One might balk at the idea that T has a probe for negation given that negation doesn't appear in every derivation. A potential perspective on this problem is that while T does indeed always have a probe for negation, this probe is allowed to fail without crashing the derivation (Preminger 2011).

3   T-to-C Movement

At this point we have seen how E2P2 can explain auxiliary movement to T as well as its interaction with negation in both finite and non-finite contexts. I will now show that we can extend this analysis to T to C movement and that such an extension in fact explains previously ill-understood facts about *do*-support.

---

[6]One question this raises is why we never see *to-n't* (or rather *n't-to* since *not* is merging from above *to*). I'm not sure why *not* doesn't form a contraction with *to* when it m-merges to T. This could be the result of a phonological/morphological constraint.

It is often assumed that movement to C is a two step process: 1) V to T movement, which feeds 2) T to C movement. However this assumption proves problematic if we take movement to C to also be the result of an E2P2 property on C. If a viable goal for C's E2P2 property heads T (after V to T movement), C's E2P2 property should be automatically satisfied by virtue of T being C's sister. Movement to C appears to always be obligatory when applicable, contrary to this prediction.

(11) a. What has Gromit eaten?   (12) a. *What Gromit has eaten?
      b. What is Gromit eating?   b. *What Gromit is eating?
      c. What does Gromit like to eat?   c. *What Gromit likes to eat?

If we take movement to C to be a different head movement process, i.e. PF head movement or head movement not driven by an E2P2 property, etc., we can maintain the assumption that movement to C goes through T and explain the contrast between (11a/b) and (11d/e) well enough. However, we might have problems understanding the contrast between (11c) and (11f). In sentences like (11c), T will not have triggered *do*-support by the time C is merged, because there are no intervening heads between T and the verb. And yet, there is *do*-support in C. If this movement is truly T to C movement, we need to have a second version of *do*-support specific to C in which empty T nodes get converted to *do*, or C has a default value of *do* if nothing moves to it.

I propose that we can have a unified theory of *do*-support that avoids these problems if we abandon the notion that movement to C is a two step process. In other words if we allow T to C movement to actually be *v to C* movement[7], we can account for the facts with just one E2P2 property on C that can be inherited by T under the right conditions (Chomsky 2005).
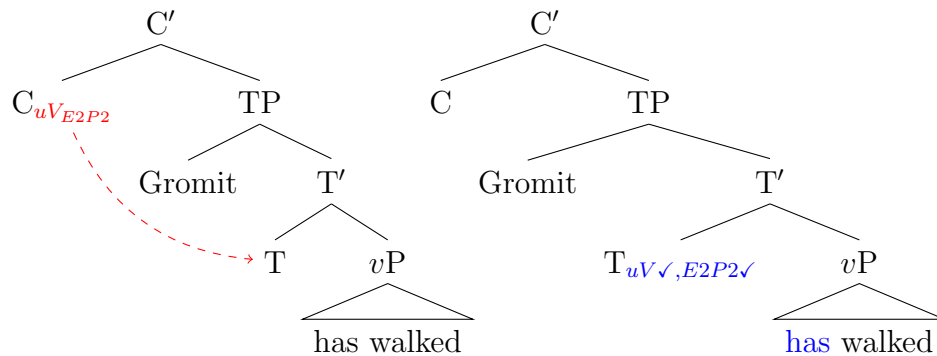

Figure 3: E2P2 is born on C but can be inherited by T if C has no other probes.

---

[7]Harizanov and Gribanova (2018) allow for this possibility in V2 environments cross-linguistically.
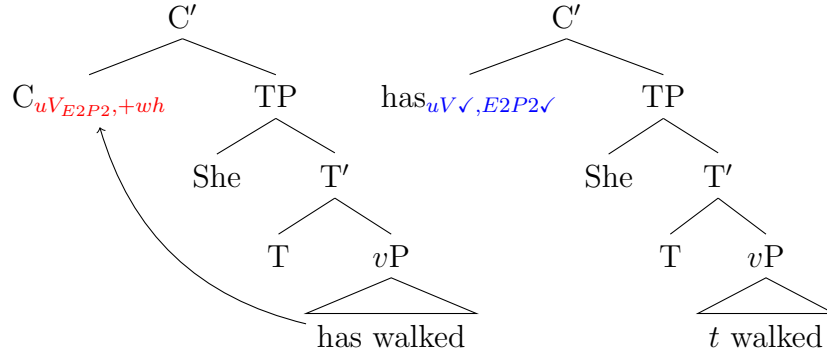
Figure 4: If C has other probes on it (like +wh), C keeps its E2P2 property, triggering obligatory v to C movement

In this proposal, there is only one E2P2 property for verbal elements that is shared between T and C depending on whether there are other active probes on C. If the E2P2 feature is on T, movement to T by the verbal element is optional in the absence of negation or emphasis because T is adjacent to the rest of the verb phrase. If the E2P2 feature is on C, T is always an intervening head between C and the verbal element. Therefore head movement to C is always obligatory to satisfy E2P2, and *do*-support is the same featural movement that we saw for V to T movement[8].

3.1 Countercyclicity and Free merge

One might worry about the countercyclic nature of this process. If T inherits its E2P2 feature from C, all movement to T must tuck in under C instead of building new structure. To avoid this problem, we can adopt a free-merge view of movement (Chomsky 2016) in which movement applies freely and randomly throughout a derivation. At the end of each derivation, if all of the features are in the right configuration to get checked, the derivation is successful. If at the end of the derivation some features go unchecked, the derivation crashes. In this way, derivations with V to T movement allow for movement to T to happen before C is merged. Once C is merged, only the derivations where an E2P2 property on T/C is satisfied will succeed[9].

---

[8]Counterfactual inversion is an apparent outlier in this paradigm in that it doesn't trigger *do*-support even though it allows other auxiliaries to invert (e.g. *Had/*did I go(ne) to the store...*). However, counterfactual inversion's unique aspectual requirements may indicate that it is a separate phenomenon: perhaps triggered by a different probe with specific semantics. Such a probe could coincidentally trigger auxiliary movement by searching for features that are only hosted on auxiliary verbs (or on *if*, which could satisfy the probe by heading C), without that probe necessarily looking for verbal elements in general.

[9]Of course this raises the possibility of random auxiliary movement to T before C in T to C derivations. However as long as we assume that the auxiliary doesn't project on T here because it doesn't check its features at T, movement of the auxiliary to C is still predicted to be obligatory.

4 Conclusion

In this paper I illustrated that we can capture classic facts about the English auxiliary system with an extension of the EPP, called E2P2. This proposal is a departure from previous analyses because it allows auxiliary verbs to remain in situ in the absence of other maximal projections between them and T. This proposal argues for feature-driven auxiliary movement to T in English, which builds on work by Matushansky, Harizanov and Gribanova suggesting that such head movement happens in the syntax. On this analysis, *do*-support is featural movement from v to T, which parallels auxiliary movement. With evidence from the distribution of negation in non-finite contexts, I discussed the potential relationship this proposal could have with economy and Anti-locality constraints on movement. More evidence is needed to decisively say how and whether such constraints should be active in this type of head movement.

Appendix A

	CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK	ZA	NG	GH	KE	TZ	JM
1	NOT TO	379614	82423	25313	83569	20784	29794	15281	16516	9126	11037	6543	8824	8899	8527	7270	8843	9467	7776	7243	5721	6658
																				2.320 seconds		
1	TO NOT	58632	18347	4765	10214	1990	5429	2496	2542	737	1345	845	1384	1288	1082	899	1184	753	637	1045	640	1010
																				2.234 seconds		
1	TO OCCASIONALLY NOT	2					1	1														
																				1.125 seconds		
1	OCCASIONALLY NOT TO	1				1																
																				1.141 seconds		
1	NOT TO OCCASIONALLY	1	1																			
																				0.805 seconds		

Figure 5: Corpus data about the *not to*/*to not* alternation mentioned in section 3. It looks like *not to* is used much more widely than *to not* (though both are very high frequency), but this number is inflated due to matches from prepositional *to* or expressions like *not to worry*. Any of the missing permutations brought up zero matches on the corpus site.

Scope judgements reveal that negation in both *not to/to not* can take sentential scope.

(13)   a. All the arrows have not been hitting the target. (all > not, not > all)

           b. All the arrows have been not hitting the target. (all > not, *not > all)

(14)   a. For all the arrows to have been not hitting the target recently, you must be tired. (all > not, *not > all)

           b. For all the arrows to not have been hitting the target recently, you must be tired (all > not, not > all)

           c. For all the arrows not to have been hitting the target recently, you must be tired (all > not, not > all)

References

[1] Abels, Klaus, *Successive Cyclicity, Anti-locality, and Adposition Stranding.* Doctoral dissertation, UConn, 2003.

[2] Alexiadou, Artemis, and Anagnostopoulou, Elena, "Parametrizing AGR: word order, v-movement and EPP-checking." *Natural Language and Linguistic Theory.* 16:491-539. 1998.

[3] Baker, C. L., "The Syntax of English Not: The limits of Core Grammar". Linguistic Inquiry, Volume 22, No. 3, pg. 387-429, 1991.

[4] Bobaljik, Johathan. *Morphosyntax: the syntax of verbal inflection.* MIT Dissertation, 1995.

[5] Chomsky, Noam, *Syntactic Structures.* Mouton and Co, The Hague, 1957.

[6] Chomsky, Noam, *The Minimalist Program.* MIT Press, 1995.

[7] Chomsky, Noam, *On Phases.* In Freidin, R., Otero, C., & Zubizarreta, M. (eds.) Foundational Issues in Linguistic Theory Essays in Honor of Jean-Roger Vergnaud, 2005.

[8] Chomsky, Noam, *Puzzles about phases.* 2016.

[9] Cinque, Guglielmo, *Adverbs and Functional Heads.* Oxford University Press, 1999.

[10] Emonds, Joseph, *Root and structure-preserving transformations.* Doctoral Dissertation, MIT, 1970.

[11] Erlewine, Michael Yoshitaka, "Anti-locality and optimality in Kaqchikel Agent Focus". *Nat Lang Linguist Theory*, 2016.

[12] Gribanova, Vera, and Harizanov, Boris, *Whither head movement?.* NLLT, 2018.

[13] Grohmann, Kleanthes K., *Prolific domains: ON the anti-locality of movement dependencies.* John Benjamins. 2003.

[14] Hankamer, Jorge, *Auxiliaries and Negation in English.* UC Santa Cruz, 2011.

[15] Lobeck, Anne, *Ellipsis: Functional heads, licensing, and identification.* Oxford University Press, 1995.

[16] Matushansky, Ora, *Head Movement in Linguistic Theory.* Linguistic Inquiry, 2006.

[17] Merchant, Jason, *The Syntax of Silence.* Oxford Studies in Theoretical Linguistics, 2001.

[18] Merchant, Jason, *Fragments and ellipsis.* Linguistics and Philosophy 27: 661-738, 2005.

[19] Pesetsky, David, and Torrego, Esther, "The Syntax of Valuation and the Interpretability of Features". In *Phrasal and clausal architecture: Syntactic derivation and interpretation*, p. 262-294. 2004.

[20] Potsdam, Eric, *NegP and subjunctive complements in English*. Linguistic Inquiry 28, 533-541, 1997.

[21] Preminger, Omer, *Agreement and Its Failures* Linguistic Inquiry Monographs, 2014.

[22] Richards, Norvin, *Featural Cyclicity and the Ordering of Multiple Specifiers*. MIT Linguistics, 2005.

[23] Richards, Norvin, *Uttering Trees*. MIT Press, 2010.

[24] Yuan, Michelle, *Person Restrictions in South Baffin Inuktitut: An Argument for Feature Movement*. Proceedings of WSCLA 19, 2015.

[25] Zwicky, Arnold M., and Pullum, Geoffrey K., "Cliticization vs. Inflection: English N'T". *Language*, Vol. 59, No. 3 (Sep., 1983), pp. 502-513. 2009.