Revisiting the COMP-trace effect:
Syntax after all?

Grant Goodall[*]

Abstract. The COMP-trace phenomenon has long resisted easy explanation or even consensus as to whether it is ultimately due to the syntax or something else. Two analyses are examined here, one based on a principle of syntax (Anti-locality) and the other based on a principle of sentence planning (Principle of End Weight; PEW). Three cases are presented in which the Anti-locality analysis predicts that a COMP-trace effect will arise, while the PEW analysis does not. In all three cases, involving inversion in matrix clauses, inversion in embedded clauses, and clauses headed by prepositional complementizers, the COMP-trace effect does seem to occur, suggesting that the Anti-locality analysis is correct. This result is compared to earlier evidence that suggested that the PEW analysis was empirically superior, and a new way of understanding these results is proposed that is compatible with the evidence presented here in favor of Anti-locality.

Keywords. COMP-trace effect; Anti-locality; sentence planning; L2 syntax; *wh-*movement

1. Introduction. Long-distance filler-gap dependencies are known to be constrained in many ways, and a number of these constraints make some intuitive sense. Extraction out of structurally complex environments, for example, such as a clause embedded within a DP (a "complex NP"), is heavily degraded relative to extraction out of simpler environments, such as a complement clause. Similar degradation occurs when there are overlapping dependencies, such as occurs with extraction out of a *wh*-clause. Intuitively, it is perhaps not surprising that structural complexity or the presence of competing dependencies would have such an effect on acceptability. The fact that these effects "make sense" at some level does not resolve the question of what accounts for them, but it might at least give us some hints of avenues to explore as we search for the correct account.

Not all constraints on filler-gap dependencies are of this type, however, in that some make very little intuitive sense. A prime example of this is the COMP-trace effect, illustrated in (1).

(1)   a. Who do you think [__ saw Mary] ?
         b. Who do you think [Mary saw __] ?
         c. *Who do you think [that __ saw Mary] ?
         d. Who do you think [that Mary saw __] ?

Extraction out of finite complement clauses is normally allowed freely, as seen in (1a-b), but when the clause is introduced by the complementizer *that*, extraction of the subject is disallowed, as in (1c), while extraction of non-subjects continues to be possible, as in (1d). This pattern, first noted in Perlmutter (1968, 1971) and much discussed in the decades since (see Pesetsky 2017; Cowart & McDaniel 2021 for overviews) is surprising in that it is not at all obvious why *that* should have such an effect on extraction or why the effect should apparently be limited to subjects. If anything, one might expect non-subject extraction to be the type that exhibits substantial degradation, given that it results in a longer dependency and the dependency must cross an intervening DP (e.g., *Mary* in (1d)).

Perhaps because of the counterintuitive nature of the pattern in (1), there has been an unusually wide variety of analyses and there is currently no consensus view of what lies behind it. One point of controversy, in fact, is whether this phenomenon is syntactic in nature or has its roots in some other domain, such as prosody or sentence processing. In recent work, Boyoung Kim and I found evidence that seemed to favor an analysis in terms of sentence planning over one based purely on syntax (Kim & Goodall 2024). Here I return to these two types of analyses armed with new sets of data and I make two claims.[1] First, the new evidence seems to point in the opposite direction of what we found in Kim & Goodall (2024). That is, it supports a syntactic analysis over one based on sentence planning. Second, despite the apparent conflict between our older results and the new evidence, they can be reconciled in a way that seems promising if we adopt a syntactic analysis along with plausible assumptions about the acquisition process.

2. Two analyses. Although the unexpected and counterintuitive nature of the COMP-trace phenomenon has led to a variety of analyses, most have assumed that the effect arises because of something that goes wrong in the syntax in cases like (1c). Other approaches have been taken, however, spurred on partly by the realization that only preverbal subject gaps trigger the effect (as in Rizzi's 1982 evidence from Italian and much subsequent work), leading to the conjecture that it is not the subject status of the gap *per se* that leads to ill-formedness, but rather the fact that the gap is in the leftmost or highest position within the clause. Syntactic analyses are still possible, of course, even if it is not the subject status of the gap in (1c) that prohibits extraction, but the apparent lack of a need to refer to the subject in the analysis makes it possible to also imagine analyses based on other components of the grammar or on extra-grammatical factors.

Here we examine two analyses of the COMP-trace effect that take these considerations into account. Both claim that the effect arises when the gap is leftmost/highest within the clause, but one attributes this to syntactic factors and the other to factors relating to sentence planning. I refer to each by the name of the principle that is claimed to be crucially involved: Anti-locality in the case of the syntactic account and Principle of End Weight (PEW) in the case of the sentence planning account.

---

[1] This study is part of a larger research project on the COMP-trace effect. The project includes formal sentence acceptability experiments, but here I mainly present traditional judgment data, based on my own judgments and those of others with whom I have consulted. It is worth noting that there is always some variability in this type of data, and perhaps especially so with regard to COMP-trace phenomena. In some early work, this led researchers to think that the existence of the COMP-trace effect in English is subject to idiosyncratic or dialect variation (e.g., Sobin 1987), but careful experimental work in recent years suggests very strongly that this is not the case. That is, the traditional judgment that (1c) is significantly less acceptable than the other sentences in (1) appears to be a very robust and essentially universal effect among English-speakers (e.g., Chacón 2021; Cowart & McDaniel 2021). For the particular phenomena explored in this paper, acceptability experiments will provide new and valuable types of information, but I do not expect them to refute the more informally collected judgments reported here.

2.1. ANTI-LOCALITY. Among current syntactic analyses of the COMP-trace effect, the most influential idea is that the first step of movement in (1c) is "too short" (e.g., Bošković 2016; Erlewine 2016, 2020; Pesetsky 2021). That is, the *wh*-phrase first moves to the specifier of C before undergoing further movement into the higher clause, as sketched in (2).

(2)   *Who do you [think [that [___ saw Mary]]]?

This first step is claimed to be impossible because of a general principle called Anti-locality, shown in (3).

(3)   Anti-locality: Movement of a phrase from the Specifier of XP must cross a maximal projection other than XP.   (Erlewine 2020)

Anti-locality rules out the configuration in (2), because movement of the specifier of TP crosses only TP itself as it moves into the specifier of CP (see Bošković 2016; Pesetsky 2021 for different formulations of this idea).

In sentences where movement is out of a lower position in the clause, as in movement of the object in (1d), TP counts as "a maximal projection other than XP", so moving into the specifier of CP does not violate Anti-locality. Similarly, in languages where subjects may originate in a lower, postverbal position, movement into the specifier of CP will be in accord with Anti-locality.

In cases where *that* is not present, there is no intermediate CP structure (e.g., Bošković 1997, 2016; Rizzi & Shlonsky 2007; Pesetsky 2021), so the first step of movement is longer, going directly into the higher clause. TP will be crossed, but so will other maximal projections in the higher clause, so Anti-locality is respected. For this reason, there is no asymmetry between subject and object extraction from embedded clauses without *that*, as seen above in (1a-b).

2.2. PRINCIPLE OF END WEIGHT. In McDaniel et al. (2015), the COMP-trace effect is derived from principles of sentence processing. The analysis takes as its point of departure the idea that the clause is the default major planning unit in sentence production, but that speakers may plan the matrix and embedded clauses as a single unit under some circumstances. McDaniel et al. assume that this joint planning of the two clauses is associated with the absence of *that* in finite embedded clauses, an assumption that is supported by findings that *that* is used less often when advance planning of the embedded clause is possible (Ferreira & Dell 2000), and that gaps are especially burdensome for the sentence production system when they are at the beginning of a planning unit. They attribute this second assumption to the PEW from Wasow (2002).

(4)   Principle of End Weight (PEW): Phrases are presented in order of increasing weight.
                                                                              (Wasow 2002: 3)

McDaniel et al. take "weight" in this case to refer to syntactic and semantic complexity, which means that gaps, which presumably represent significant syntactic complexity, are problematic for PEW when they are at the beginning of a planning unit. The complexity of gaps derives at least in part from the fact that there is no overt material to process, but syntactic structure (including the full DP, if one adopts the copy theory of movement) must nonetheless be posited.

Putting all of this together allows McDaniel et al. to derive the COMP-trace effect. If clauses with *that* are a separate planning unit, then having a gap at the beginning of that clause would violate PEW, while having a gap further to the right would not. This accounts for the subject/non-subject asymmetry seen in (1c-d). If clauses without *that* are jointly planned together with the

matrix clause, then both subject and non-subject gaps will be compatible with PEW, since both types of gaps are far to the right within the planning unit. This then accounts for the lack of an asymmetry in clauses without *that*, as in (1a-b).

2.3. DISTINGUISHING BETWEEN THE TWO ANALYSES. The Anti-locality analysis and the PEW analysis are similar in that neither one rules out a subject gap *per se* in an embedded *that*-clause. In rough descriptive terms, however, both generally prohibit a gap in the initial position within the clause. In the Anti-locality analysis, this is because only a gap in the highest specifier position will violate Anti-locality; any lower specifier will "cross a maximal projection other than XP" as it moves into the CP layer. In the PEW analysis, this is because the clause is taken to be a planning unit and a gap in the initial position is a blatant violation of PEW; any later gap will be intermingled among heavier elements in the planning unit. A schematization of what causes a violation in each of the two analyses is presented in Figure 1.

$[_{CP}\rule{1cm}{0.15mm}\ C\ [_{TP}\ \rule{1cm}{0.15mm}$ $[_{\text{planning unit}}\ \rule{1cm}{0.15mm}$

Anti-locality analysis   PEW analysis

Figure 1. The configuration that violates Anti-locality (left panel) and the configuration that violates PEW (right panel)

Though the two analyses make similar predictions, they are clearly not identical, and it is possible to tease out differences that can then be tested empirically. Specifically, the Anti-locality analysis claims that sentences like (1c) are bad because there is an instance of movement as in the first step of (2) (i.e., Anti-locality-violating movement), while the PEW analysis claims that it is because there is a gap at the beginning of the planning unit. To distinguish between the two analyses, then, we should look for cases such as (5a) or (b).

(5) a. The left panel of Figure 1, but not the right: Movement violates Anti-locality but the resulting gap is not at the beginning of a planning unit.
   b. The right panel of Figure 1, but not the left: The gap is at the beginning of a planningunit, but the movement does not violate Anti-locality.

The two analyses make opposite predictions in these cases. For (5a), Anti-locality predicts a sharp decline in acceptability, while PEW does not, and in (5b), PEW predicts low acceptability while Anti-locality does not. Finding and testing cases like these should thus, in principle, allow us to decide between the two analyses.[2]

---

[2] Brooke Larson points out to me that another way to distinguish between the two analyses is to look for cases where advance planning of the embedded clause is facilitated, but the syntactic structure remains the same. Larson notes that this might be done by inserting filler words like *um* or *like* (Boomer 1965), which arguably allow for more planning time, as in (i).

(i) Who do you think that … um … like saw Mary?

Under the PEW analysis, one might expect amelioration here, relative to (1c), but under the Anti-locality analysis, one presumably would not. This intriguing suggestion is well worth exploring, though there would be challenges in designing an effective experiment. Audio stimuli would probably need to be used (Sedarous & Namboodiripad 2020), and any amelioration seen in (i) would need to be compared to possible amelioration whenever filler expressions are inserted before ill-formed gaps (e.g., *Who do you wonder when Mary will … um … like talk to?*). It is at least conceivable that filler expressions might interrupt syntactic processing and/or distract listeners to the extent

Kim & Goodall (2024) investigate a case along the lines of (5b). In that paper, we explore the COMP-trace phenomenon among L2 English speakers, a population that arguably has a very reduced ability for advance planning in sentence production (see Kim & Goodall 2024 for discussion of findings from the literature that support this conclusion). In particular, we expect these speakers to be much less able to do joint planning of the matrix and embedded clauses, so all embedded clauses, with or without *that*, will thus be separate planning units. This would predict that a gap at the beginning of any embedded clause, including one without *that*, would result in a PEW violation. This prediction is borne out, as seen by the results of an acceptability experiment in Figure 2.
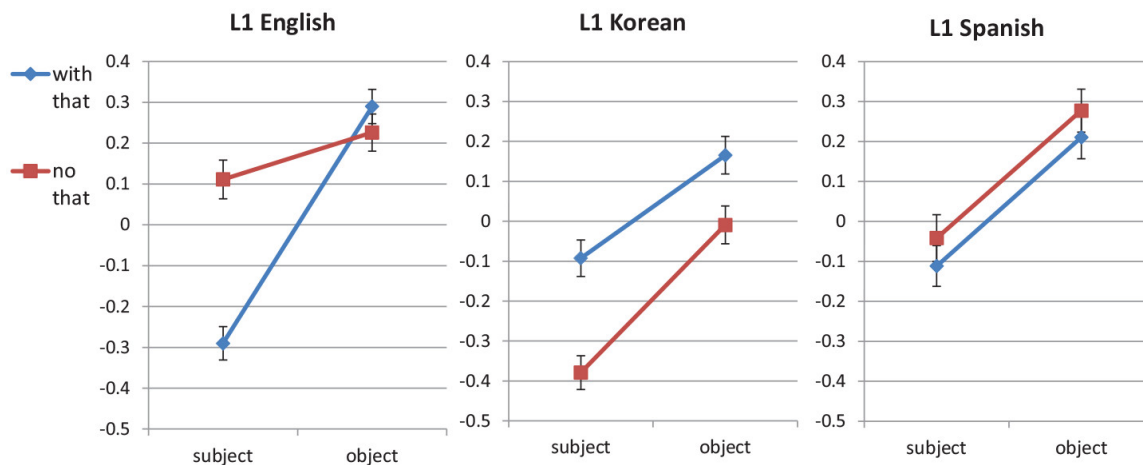

Figure 2. Results from Kim & Goodall (2024) (in z-scores) showing acceptability of extraction from embedded clauses in English among three groups of speakers

As seen in Figure 2, L1 English speakers show a very large decline in acceptability when a subject is extracted from a *that*-clause, as expected. L2 English speakers (L1 Korean or L1 Spanish), on the other hand, show a similar decline in both clauses with *that* and clauses without *that*. This difference between L1 and L2 makes sense if the L2 speakers have difficulty doing advance planning of embedded clauses. All types of embedded clauses will thus tend to be separate planning units, so a gap at the beginning of the clause will violate PEW, whether or not *that* is present.

Subject extraction from an embedded clause without *that* among these L2 speakers thus seems to be a case of (5b): Anti-locality is not violated, but PEW is. Given that this type of extraction shows the same degradation for these speakers as the standard *that*-trace case, the PEW analysis thus seems to make the correct prediction, and in Kim & Goodall (2024) we argue that this fact provides evidence in favor of the PEW analysis and against the Anti-locality analysis.

3. Evidence. Now that we have seen a case of (5b), let us now turn to cases of (5a): instances where Anti-locality is violated, but where the gap does not seem to be at the beginning of a sentence planning unit. The Anti-locality analysis of course predicts that such cases will be ungrammatical, but the PEW analysis does not. We examine three of these cases here, and we will see that in all three, the Anti-locality analysis seems to make the correct prediction. These

---

that the ill-formedness of the sentence is not perceived as much as it would be otherwise. In order to be relevant to evaluating the PEW analysis, the amelioration seen for (i) would need to be greater than the type of amelioration just described.

cases thus seem to push us to the opposite conclusion of that reached by Kim & Goodall (2024), a fact to which we return in section 4.

3.1. INVERSION IN MATRIX CLAUSES. Under the Anti-locality analysis, the locus of ungrammaticality in COMP-trace sentences as in (1c) is found in the first step of movement, from the specifier of TP to the specifier of CP. Such movement crosses TP and no other maximal projection, so it is too local to satisfy the constraint. When C is overt, as it is in this case (given the presence of *that*), then CP is obligatorily projected and there is no way to evade the Anti-locality violation: moving higher than CP would violate locality constraints and moving to a position within CP violates Anti-locality, as we have seen. Under this analysis, we thus expect any movement from the specifier of TP to the specifier of CP to be disallowed when C is overt.

This prediction can be tested in instances where C is overt by virtue of T-to-C movement, as with subject-auxiliary inversion in *wh*-questions in English. It will be particularly helpful to use auxiliaries like *do* and *ought* (as opposed to ones like *can* or *will*), since these are known to be possible only when T-to-C movement has occurred, as seen in (6) and (7).

(6)   a. * They **did** hire Mary.   (* with non-emphatic reading)
        b.   Who **did** they hire ___ ?   (OK with non-emphatic reading)

(7)   a. * They **ought** hire Mary.   (OK with *to*: They ought to hire Mary.)
        b.   Who **ought** they hire ___ ? (* with *to*: *Who ought they to hire?)

When the subject is extracted, auxiliary *do* and *ought* are not possible, as seen in (8) and (9).

(8)   a. *Who **did** ___ hire Mary? (* with non-emphatic reading)
        b. Who hired Mary?

(9)   a. *Who **ought** ___ hire Mary?
        b. Who ought to hire Mary?

In the (a) examples, C is overt, as evidenced by *did* in (8a) and *ought* (without *to*) in (9a), so the position of *who* to the left of this suggests that movement has been to the specifier of CP. Anti-locality predicts that this movement will be prohibited, as it in fact appears to be. (8a) and (9a) are thus excluded for the same reason that standard *that*-trace cases (as in (1c)) are, an idea first proposed by Koopman (1983). The only way to avoid the Anti-locality violation here is to not have an overt C (thus avoiding the need for a CP layer, presumably), as in (8b) and (9b).

As we have seen, then, (8a) and (9a) are only derivable through Anti-locality-violating movement, so the Anti-locality analysis correctly predicts that they are unacceptable. What is notable for our purposes, though, is that the gaps in these cases do not appear to be at the beginning of a sentence planning unit, given that the sentence planning unit is presumably the entire clause. (8)–(9) thus constitute a case of (5a), where Anti-locality predicts unacceptability, but PEW does not, thus offering us a way to compare the competing hypotheses. The Anti-locality analysis seems to clearly come out the winner here.

3.2. INVERSION IN EMBEDDED CLAUSES. One might object to the above conclusion on the grounds that (8)–(9) do not involve embedded clauses and therefore might not represent the same phenomenon as the standard COMP-trace effect, contrary to Koopman's (1983) claim. Under this view, we would not expect the PEW analysis to have anything to say about (8)–(9), and they thus become irrelevant to deciding between it and the Anti-locality analysis.

To address this potential concern, we can look for cases where we find the same type of movement (from the specifier of TP to the specifier of CP when C has an overt auxiliary) but in an embedded clause. Such cases will not be easy to find, since inversion in English is generally disallowed in embedded clauses, but they can be created by taking advantage of the fact that fronted negative expressions require inversion, even in embedded clauses, as seen in (10).[3]

(10)   a. They said [that [at no time] **did** Mary hire that person].
        b. They said [that [under no circumstances] **ought** Mary hire that person].

We can then extract a *wh*-phrase out of the embedded subject position to see whether we can re-create the type of movement we examined in the previous section. Movement out of the embedded clause always incurs some amount of degradation, and we would expect even more in sentences like (10) where there is already a fronted negative expression within the embedded clause. Nonetheless, we can compare extraction out of subject position with extraction out of object position, as in (11) and (12).[4]

(11)   a. ?Who did they say [that [at no time] **did** Mary hire __] ?
        b. *Who did they say [that [at no time] **did** __ hire Mary] ?
        c. ?Who did they say [that [at no time] __ hired Mary] ?

(12)   a. ?Who did they say [that [under no circumstances] **ought** Mary hire __] ?
        b. *Who did they say [that [under no circumstances] **ought** __ hire Mary] ?
        c. ?Who did they say [that [under no circumstances] __ **ought** to hire Mary] ?

In the (a) examples, we see that *wh*-extraction out of the embedded clause is possible in principle, albeit with some decline in acceptability, presumably for the reasons mentioned above. In the (b) examples, we see that extraction out of the subject position is not possible when there is T-to-C movement (i.e., when there is an overt auxiliary in C), but the (c) examples show that this is possible when there has not been T-to-C movement.

The facts in (11)–(12) are essentially what the Anti-locality analysis predicts. In the (a) examples, the first step of movement does not originate in the specifier position of TP, but it does cross TP (and presumably other maximal projections). In this way, it satisfies Anti-locality. The (b) examples are the same except that here the first step of movement originates in the specifier of TP, so Anti-locality is violated. In the (c) examples, movement also originates in the specifier of TP, but since T-to-C movement has not occurred, there is no "lower" CP here and movement can target a higher layer (perhaps the "upper" CP headed by *that*), thus satisfying Anti-locality. These three scenarios are illustrated in (13), which shows the initial step of movement in the embedded clauses of (11a-c).[5]

[3] Another possible place to look for such cases would be varieties of English that allow inversion in embedded clauses, as is sometimes claimed for some varieties in the British Isles (Filppula 2000) and for Indian English (Bhatt 2004).
[4] The relatively acceptable sentences are marked with "?" in recognition of the fact that extraction out of an embedded clause with a fronted negative is likely to cause more than the usual amount of degradation, but consultants suggest that the contrast between the ? and * sentences here is clear and sharp.
[5] It is conceivable that the first step of movement in (c) is to a position within the layer associated with the fronted negative expression, rather than to the "upper" CP layer. Whatever analysis one adopts here would presumably be the same as what is needed for the well-known "adverb effect" cases (Bresnan 1977; Culicover 1983), discussed below.

(13) a. [$_{CP}$ that [[at no time] [$_{CP}$ **did** [$_{TP}$ Mary hire __]]]]

  b. [$_{CP}$ that [[at no time] [$_{CP}$ **did** [$_{TP}$ __ hire Mary]]]]

  c. [$_{CP}$ that [[at no time] [$_{TP}$ __ hired Mary]]]

It is not easy to imagine how the PEW analysis could account for these facts. We would have to say that the gap in the (b) examples is at the beginning of a sentence planning unit, but that the gap in the (c) examples is not. It is not at all clear on what basis we could make that distinction. The (b) examples thus seem to be another case where we have Anti-locality-violating movement, but where the gap is not plausibly analyzed as being at the beginning of a sentence planning unit. This is then another case of (5a), and since Anti-locality makes the correct prediction here, we again have evidence in favor of the Anti-locality analysis and against the PEW analysis.

It is worth pointing out that one would expect that the examples in (11a) and (12a) might be problematic, since they appear to require two CP projections, one to host T-to-C movement (the "lower" CP) and one headed by *that* (the "upper" CP). Whether this situation contributes to degradation or not, the important point is that the same situation obtains in (11b) and (12b), yet there is a noticeable decline in acceptability for (11b)–(12b) relative to the baseline judgments in (11a)–(12a). Under the Anti-locality analysis, this degradation relative to the baseline results from the violation of Anti-locality in the first step of movement. It is also worth noting that (11c) and (12c) exemplify the "adverb effect" (Bresnan 1977; Culicover 1983), in which COMP-trace effects seem to disappear when adverbial expressions intervene between *that* and the subject position. Both the Anti-locality and PEW analyses are able to account for this effect, though, so it is less relevant to our concerns here.

3.3. PREPOSITIONAL COMPLEMENTIZERS. Another potentially relevant case comes from ACC-*ing* gerund complement clauses, as in (14).

(14) I can imagine [Mary hiring the right person].

These are a kind of ECM structure, in that case for the embedded subject is provided by the matrix verb. Standard tests for subjecthood suggest that *Mary* in (14) truly is the subject of the embedded clause, as seen in the relatively high acceptability of the sentences in (15) and the truth-value equivalence of the sentences in (16).

(15) a. I can easily imagine **there** being a flaw in the analysis.
  b. I can't easily imagine **it** snowing on campus.
  c. I can easily imagine **all hell** breaking loose.
  d. I can easily imagine **advantage** being taken of their inexperience.

(16) a. I can easily imagine **the doctor** examining **the patient**.
  b. I can easily imagine **the patient** being examined by **the doctor**.

Case for the subject in these gerund complements comes from the main verb (*imagine*) in these examples, but it may also come from a preposition, as in (17).

(17) I am interested in [Mary hiring the right person].

The fact that a preposition is providing case in (17) does not seem to change the subject status of *Mary*, as seen by the tests in (18) and (19).

(18)   a.   I am interested in **there** being a large audience for the presentation.
           b.   I am interested in **it** snowing on Tuesday, if that can be arranged.
           c.   I am interested in **all hell** breaking loose.
           d.   I am interested in **advantage** being taken of their inexperience.

(19)   a.   I am interested in **the doctor** examining **the patient**.
           b.   I am interested in **the patient** being examined by **the doctor**.

Perhaps not surprisingly, it is also possible to perform extraction out of this type of gerund, as seen with the matrix verb *imagine* in (20).

(20)   a.   Who can you imagine [__ hiring Mary] ?
           b.   Who can you imagine [Mary hiring __] ?

In the prepositional case, however, a familiar asymmetry emerges, as seen in (21) (see Kayne 1979).

(21)   a.   *Who are you interested [in [__ hiring Mary]] ?
           b.   Who are you interested [in [Mary hiring __]] ?

That is, extraction of the subject seems to be degraded in (21a), even though this effect does not obtain with extraction of the object, as in (21b). This asymmetry only occurs in the prepositional case, however, as seen in the lack of contrast between (20a) and (20b) with matrix *imagine*.

These facts receive a ready explanation under the Anti-locality account. In (20), there is no overt C, so no CP, and movement of either the subject or the object is able to occur without violating Anti-locality. In (21), however, the presence of the preposition (or more precisely, the prepositional complementizer) means that there is additional structure above the TP. Regardless of the categorial status of this element (i.e., whether we take it to instantiate P or C), it will project structure, and locality will plausibly force the first step of movement to be to the specifier of this projection.[6] This first step of movement will thus violate Anti-locality in (21a) in that it will leave the specifier of TP and then cross TP without crossing any other maximal projection. The initial step of movement for (20a) and for (21a) is shown here in (22a) and (22b), respectively.

(22)   a. imagine [$_{TP}$ __ hiring Mary]

           b. interested [$_{CP}$ in [$_{TP}$ __ hiring Mary]]

It is much more difficult to see how the PEW account could account for these facts. If (21a) is out because the gap is at the beginning of a planning unit, it is not clear why (20a) would be different. Both *imagine* and *interested in* may take either a DP or a gerund clause as a complement. The preposition (*in* in the case of *interested in*) is lexically specified by the matrix predicate and it occurs regardless of whether the complement is a DP or a clause. Unlike *that* in finite

---

[6] As Brooke Larson points out to me, saying that *in* in (20a) instantiates P and that movement is thus to the specifier of P might be problematic, since one might then expect that any case of preposition-stranding would result in an Anti-locality violation. I do not address this question further here, however. See Emonds (1985) for discussion of the relation between P and C.

complement clauses, the preposition here is not optional and its presence does not plausibly signal a new planning unit or anything else.

The examples in (20) and (21) thus seem parallel in every way except that (21) has some additional structure (the preposition and its projection) that (20) does not. For the Anti-locality analysis, the presence or absence of this structure is crucial for determining whether extraction of the subject is possible. For the PEW analysis, on the other hand, this structure in and of itself is essentially irrelevant, unless it affects the planning unit status of the clause (and as we have seen, there is no reason to think that it does). We are thus confronting once more an example of (5a), a case where the Anti-locality analysis predicts we should see an effect and the PEW analysis does not, and once more, the Anti-locality analysis appears to make the correct prediction.

4. Conclusion. We have now seen three sets of data where the Anti-locality account and the PEW account make diverging predictions, and in every case, the predictions of the Anti-locality account seem to be correct. In all three cases, there is plausibly an instance of movement from the specifier of TP to the specifier of CP, which should clearly violate Anti-locality, but at the same time, there is no reason to think that the gap left behind by this movement is at the beginning of a separate planning unit, which would be prohibited under the PEW analysis. The three cases examined thus all seem to be straightforward examples of configurations that are disallowed by Anti-locality but not by PEW, so the fact that we do observe a substantial decline in acceptability constitutes an argument in favor of the Anti-locality account.

One initially puzzling aspect of this conclusion is that it is at odds with the conclusion reached by Kim & Goodall (2024) in a study comparing L1 and L2 speakers of English. In that article, we argue that L2 speakers have a reduced ability to perform advance sentence planning, with the result that embedded clauses are more likely to be separate planning units for them than for L1 speakers, even for embedded clauses without *that* (where the absence of *that* has been claimed to signal joint planning of the matrix and embedded clauses). We show that L2 speakers display a kind of COMP-trace effect in finite embedded clauses generally, both with and without *that*, a finding that is just what the PEW analysis predicts. If L2 speakers treat all embedded clauses as separate planning units (unlike L1 speakers, who only do this with *that*-clauses), then we expect to find a severe drop in acceptability whenever there is a gap at the beginning of the embedded clause, and this is just what the results show.

Looking at the Kim & Goodall results this way, they constitute a case where we have a gap at the beginning of a planning unit but without the kind of movement that would result in an Anti-locality violation (as in (5b) above). That is, subject extraction out of an embedded finite clause without *that* should be able to proceed without stopping in the specifier of CP, since without an overt complementizer, CP does not need to be projected. Movement of the subject will thus cross more than just TP as it enters the matrix clause, and Anti-locality will be satisfied. In this sense, then, the results suggest that the COMP-trace is best accounted for by the PEW analysis and not the Anti-locality analysis.

If this is correct, however, how can it be that the evidence we examined above points in exactly the opposite direction? The answer, I believe, lies in the fact that it is plausible to believe that L1 and L2 speakers differ in how they do extraction out of clauses without *that*. For L1 speakers, we have been assuming that when *that* is not present, CP is not projected. We have not been explicit about how this occurs, but it is sometimes claimed that it is made possible by a special operation, such as "truncation" in Rizzi & Shlonsky (2007). If so, it is reasonable to assume that L2 learners need some time to acquire this operation (especially given that the operation

does not seem to exist in many speakers' L1). Until that happens, L2 speakers will project CP and do *wh*-movement through the specifier of CP whether *that* is present or not, and this means that subject-extraction out of an embedded clause will always result in an Anti-locality violation. As learners begin to acquire the truncation operation (or whatever the exact mechanism is), they will begin to differentiate between clauses with *that* and clauses without *that*, and subject-extraction from clauses without *that* will become increasingly possible as speakers begin to do the longer-distance first step of movement that we have been assuming for L1 speakers. In the meantime, however, we expect to see the kinds of effects observed in Kim & Goodall (2024) where clauses with and without *that* behave uniformly with regard to extraction. In Kim & Goodall (2024), we assumed that L1 and L2 speakers differed in how they are able to plan embedded clauses, whereas here, we are assuming that they differ in the structure they assign to embedded clauses without *that* (with consequences for the ability to do movement out of such clauses).[7] Both sets of assumptions are plausible and in fact, both are likely correct, but the second set allows us to reconcile the L1 results that we have seen here with the L2 results from Kim & Goodall (2024) under a single Anti-locality analysis. Both L1 and L2 speakers are constrained by Anti-locality, under this view, but some (all L1 speakers, presumably, and perhaps some L2 speakers) are able to perform an operation that removes the CP layer in clauses without *that*, which in turn allows subject-extraction to occur without violating Anti-locality. In this way, then, the Anti-locality analysis can account for all the data that we have been considering.

Before concluding, I will make three further observations. First, the data reported here are based on the previous literature and on informal judgments performed by myself and others. This traditional style of data collection is of course valuable, but there are clear ways in which formal experiments could give us added value. This is particularly true in the case of extraction out of clauses headed by prepositional complementizers, as in section 3.3 above. It is often said that the COMP-trace effect here is weaker than it is with *that*-clauses, but in fact, this can't really be determined on the basis of traditional judgments. It may be that subject-extraction out of a prepositionally-headed clause is less unacceptable than the equivalent extraction out of a *that*-clause, but from an experimental perspective, such a difference would not be surprising. The prepositionally-headed clause is a gerund clause, and it is known that extraction out of nonfinite clauses like this generally seems to cause less degradation than extraction out of finite clauses such as *that*-clauses (e.g., Michel & Goodall 2013; Goodall 2022). Even if subject-extraction leads to an Anti-locality violation in the two cases in exactly the same way and the violation results in the same amount of degradation in the two cases, one would still expect the COMP-trace effect in the prepositional case to be more acceptable than the effect in the *that*-clause case, just because the baselines would be different. To tell if the size of the effect is different, one would need to do an experiment with a 2x2 factorial design, comparing subject- and object-extraction out of gerund clauses that are not headed by prepositions with those that are (as in the sentences in (20) and (21)), and then compare the size of that interaction with analogous sentences with and without *that* (as in the sentences in (1)).[8] It may turn out to be the case, though of course it is too early to tell, that the effect size is the same even though the prepositional COMP-trace

---

[7] Under either account, one would expect that as learners' exposure to the language increases, their *that*-trace effect would become more L1-like, either because of an increased ability to plan embedded clauses (in the Kim & Goodall account) or because of an increased ability to remove the CP layer in embedded clauses (in the account proposed here). See Kim & Goodall (2024) for discussion of what is known regarding the learning trajectory in this area.

[8] I currently have experiments of this type in progress. See Goodall (2021) for further details on this general type of analysis.

sentence (as in (21a)) is more acceptable than the *that*-clause COMP-trace sentence (as in (1c)). In fact, if the effect size turns out to be the same, this would be additional evidence that both sentences violate the same principle.

Second, I have adopted a particular formulation of Anti-locality, that of Erlewine (2020) (given in (3)), but there are others, and I have not attempted to determine which are compatible with the facts discussed here and which are not (see Richards forthcoming for an overview of different versions of Anti-locality). With regard to the reconciliation between the sets of data presented here and those of Kim & Goodall (2024) in particular, some formulations of the principle may turn out to be better suited than others.

Third, despite the fact that the Anti-locality analysis seems to be empirically superior, the PEW analysis appears to have a conceptual advantage in the sense that its basic principles fit naturally within a broader view of sentence planning. The idea that gaps are difficult, for instance, in a way similar to traditionally "heavy" constituents, and that such elements might be strongly avoided at the beginning of a sentence planning unit seems at least conceivable, and it opens up the possibility that the COMP-trace effect could be derived without stipulation. The Anti-locality analysis, on the other hand, at least as it has been presented here, appears more stipulative. If it is true that overly local movement is prohibited, one would want to know why this is the case, and specifically, why the grammar is unable to move an element from the specifier of TP to the specifier of CP. One would hope that this constraint follows from more basic properties of the mechanism of movement, but I have not reviewed attempts to explore that here (see Bošković 2016; Richards forthcoming for discussion) and will simply leave the issue unresolved for now.

Whatever the source of Anti-locality, however, and in fact regardless of whether Anti-locality as formulated is ultimately the correct principle, the data examined here strongly suggest that there is something about movement from the specifier of TP to the specifier of CP that leads to severe degradation in acceptability, and at least at first blush, this generalization is most naturally captured within the syntax.

References

Bhatt, Rakesh M. 2004. Indian English: Syntax. In Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton (eds.), *A handbook of varieties of English: Morphology and syntax*, 1016–1030. Berlin / New York: Mouton de Gruyter.

Boomer, Donald. S. 1965. Hesitation and grammatical encoding. *Language and Speech* 8(3). 148–158. https://doi.org/10.1177/002383096500800302.

Bošković, Željko. 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge, MA: MIT Press.

Bošković, Željko. 2016. On the timing of labeling: Deducing comp-trace effects, the subject condition, the adjunct condition, and tucking in from labeling. *The Linguistic Review* 33(1). 17–66. https://doi.org/10.1515/tlr-2015-0013.

Bresnan, Joan. 1977. Variables in the theory of transformations. In Peter W. Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax,* 157–196. New York: Academic Press.

Chacón, Dustin. 2021. Acceptability (and other) experiments for studying comparative syntax. In Grant Goodall (ed.), *The Cambridge handbook of experimental syntax*, 180–208. Cambridge: Cambridge University Press.

Cowart, Wayne & Dana McDaniel. 2021. The *that*-trace effect. In Grant Goodall (ed.), *The Cambridge handbook of experimental syntax*, 258–277. Cambridge: Cambridge University Press.

Culicover, Peter W. 1993. Evidence against ECP accounts of the *that*-t effect. *Linguistic Inquiry.* 24. 557–561. http://www.jstor.org/stable/4178827.

Emonds, Joseph E. 1985. *A unified theory of syntactic categories*. Dordrecht: Foris.

Erlewine, Michael Yoshitaka. 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language & Linguistic Theory* 34(2). 429–479. http://doi.org/10.1007/s11049-015-9310-z.

Erlewine, Michael Yoshitaka. 2020. Anti-locality and subject extraction. *Glossa: a journal of general linguistics 5*(1). 84. https://doi.org/10.5334/gjgl.1079.

Ferreira, Victor S. & Gary S. Dell. 2000. Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology* 40. 296–340. https://doi.org/10.1006/cogp.1999.0730.

Filppula, Markku. 2000. Inversion in embedded questions in some regional varieties of English. In Ricardo Bermúdez-Otero et al. (eds.), *Generative theory and corpus studies: A dialogue from 10 ICEHL*, 439–453. Berlin: Mouton de Gruyter.

Goodall, Grant. 2021. Sentence acceptability experiments: What, how, and why. In Grant Goodall (ed.), *The Cambridge handbook of experimental syntax*, 7–38. Cambridge: Cambridge University Press.

Goodall, Grant. 2022. D-linking, non-finiteness, and cross-linguistic variation in island phenomena. In Grant Goodall (ed.), *Theory and experiment in syntax*, 309–340. New York: Routledge.

Kayne, Richard S. 1979. Two notes on the NIC. In Adriana Belletti, Luciana Brandi & Luigi Rizzi (eds.), *Theory of markedness in generative grammar,* 317–346. Pisa: Scuola Normale Superiore di Pisa.

Kim, Boyoung & Grant Goodall. 2024. The source of the *that*-trace effect: New evidence from L2 English. *Second Language Research* 40, 79–102. https://doi.org/10.1177/02676583221104604.

Koopman, Hilda. 1983. ECP effects in main clauses. *Linguistic Inquiry* 14(2). 346–350. https://www.jstor.org/stable/4178332.

McDaniel, Dana, Wayne Cowart, Cecile McKee, & Merrill F. Garrett. 2015. The role of the language production system in shaping grammars. *Language* 91, 415-441. https://doi.org/10.1353/lan.2015.0021

Michel, Dan & Grant Goodall. 2013. Finiteness and the nature of island constraints. In Nobu Goto, Koichi Otaki, Atsushi Sato & Kensuke Takita (eds.), *Proceedings of GLOW in Asia IX 2012: The main session*, 187–197.

Perlmutter, David M. 1968. *Deep and surface structure constraints in syntax*. Cambridge, MA: MIT dissertation.

Perlmutter, David M. 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.

Pesetsky, David. 2017. Complementizer-trace effects. In Martin Everaert & Henk van Riemsdijk (eds.), *The Wiley Blackwell companion to syntax* (2nd edition). Hoboken, NJ: Blackwell. https://doi.org/10.1002/9781118358733.wbsyncom108.

Pesetsky, David. 2021. Exfoliation: Towards a derivational theory of clause size. Manuscript. https://lingbuzz.net/lingbuzz/004440.

Richards, Norvin. Forthcoming. Anti-locality. In Kleanthes Grohman & Evelina Leivada (eds.), *The Cambridge handbook of minimalism*. Cambridge: Cambridge University Press.

Rizzi, Luigi. 1982. Negation, wh-movement and the null subject parameter. In Luigi Rizzi (ed.), *Issues in Italian syntax,* 117–185. Dordrecht: Foris.

Rizzi, Luigi & Ur Shlonsky. 2007. Strategies of subject extraction. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics,* 115–160. Berlin: Mouton de Gruyter.

Sedarous, Yourdanis & Savithry Namboodiripad. 2020. Using audio stimuli in acceptability judgment experiments. *Language and Linguistics Compass* 14. e12377. https://doi.org/10.1111/lnc3.12377.

Sobin, Nicholas. 1987. The variable status of COMP-trace phenomena. *Natural Language & Linguistic Theory* 5(1). 33–60. https://www.jstor.org/stable/4047574.

Wasow, Thomas. 2002. *Postverbal behavior*. Stanford, CA: CSLI Publications.