On the restricted interaction of subjects and parasitic gaps*

Colin P. B. Davis ⋄ University of Konstanz ⋄ colin.davis@uni-konstanz.de

Draft of 14 November 2022

1 Introduction

The goal of this paper is to clarify the interaction of subjects and parasitic gaps (PGs)—a multifaceted topic that has mostly been discussed in scattered fashion in previous literature. Here I integrate a variety of facts about PGs in English, which I argue present a coherent picture—that subjects and PGs interact productively and straightforwardly, except when independent factors intervene. This data gathering process was accompanied by seeking confirmation of these patterns from native speaker judgments. The contrasts I discuss here are corroborated by 10 native speakers so far (including the author).

The nature and limitations of A-bar movement from subject position will be central to this paper. This topic is itself an active area of debate. In many languages, English among them, it is clear that a *wh*-phrase that originates in a non-subject position must move (ignoring multiple-*wh* questions), as we see below:

(1) *Obligatory non-subject wh-movement*
   a. **What**$_1$ will you eat t_1?
   b. * Will you eat **what**?

However, when the *wh*-phrase is the subject, there would be no change in word order whether it moves or not (setting the related issue of T to C movement aside for now):

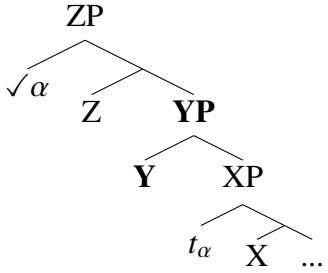
(2) *Two potential analyses of subject wh-phrases*
   a. [$_{CP}$ **Who**$_1$ [$_{TP}$ t_1 will eat the cake]]?
   b. [$_{CP}$ [$_{TP}$ **Who** will eat the cake]]?

For reasons like this, subject A-bar movement is often difficult to diagnose. Since *wh*-movement is obviously required from non-subject positions (1), it is often assumed that *wh*-movement also occurs from subject positions (2a). However, some work argues that there is typically no clause-internal subject A-bar movement, as in (2b) above (George 1980; Chung and McCloskey 1983; Agbayani 2000; Brillman and Hirsch 2016; Carstens et al. 2017; Gallego 2017; Erlewine 2017, 2020). Using data about parasitic gaps, I will argue that clause-internal subject A-bar movement is indeed (usually) impossible, though we will also see a principled exception to this generalization.

*Thanks to feedback from Michael Yoshitaka Erlewine, Alex Grosu, Claire Halpert, Romi Hill, Julia Horvath, Bozhil Hristov, Elango Kumaran, Idan Landau, Jean-Philippe Marcotte, Luis Miguel Toquero-Perez, Molly Rolf, Tal Siloni, Hooi Ling Soh, George Walkden, and Adam Woodnut, as well as audiences at the University of Minnesota, Tel Aviv University, Ben Gurion University, and the University of Konstanz.

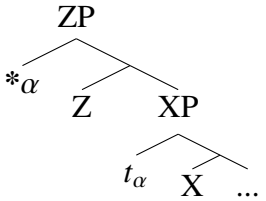
I argue that the correct predictions about the interaction of subjects and PGs emerge from a hypothesis about the limitations of movement termed *anti-locality* (Bošković 1997; Ishii 1999; Grohmann 2003; Abels 2003; Erlewine 2015, 2017, 2020, and more). This hypothesis states that movements that are too short fail, though several different versions of this constraint have been proposed. I will focus on a version of anti-locality stating that movement from one specifier to another must cross over at least one phrase (Bošković 2005; Brillman and Hirsch 2016; Erlewine 2015, 2017, 2020; Davis 2020b, a.o.). Given this constraint, it is possible for a phrase α to move from spec-XP to spec-ZP in the following schema, since YP sits between XP and ZP:

(3) *A schema for movement that is long enough*



But if YP were absent, this movement would fail due to being too short:[1]

(4) *A schema for movement that is too short*



Several of the works cited above argue that this constraint is responsible for preventing clause-internal subject A-bar movement in many contexts: If subjects must move to spec-TP for case/EPP reasons before A-bar moving to spec-CP, that A-bar movement will fail since movement from spec-TP to spec-CP is too short:

(5) *Prediction of anti-locality: Movement from spec-TP to spec-CP cannot occur*

a. * $[_{CP}$ **Who** $[_{TP}$ *t* will eat the cake]]?

b. ✓ $[_{CP}$ $[_{TP}$ **Who** will eat the cake]]?

As mentioned, my exploration of anti-locality and the nature of subject movement uses facts about PGs (Engdahl 1983; Nissenbaum 2000; Culicover and Postal 2001, a.o.). PGs are, roughly speaking, "extra" gaps that can occur in constituents crossed-over by an A-bar movement. PGs are productive in object positions:

[1]To be more precise, if α inhabits spec-XP and XP is the sister of Z, movement of α to spec-ZP would violate anti-locality. If a phrase YP intervenes between XP and ZP such that YP dominates XP but not ZP, this movement of α succeeds. See Erlewine (2020) for additional discussion of and recent citations for this proposal.

(6) *Object PGs in clausal adjuncts*

a. [What movies]$_1$ did Mary [claim she liked t_1 [in order to get you to see **PG**$_1$]]?

b. John's the guy $\varnothing_1$ that they said they'll [hire t_1 [if I criticize **PG**$_1$ publicly]]. (Nissenbaum 2000, p. 30)

Importantly, PGs in subject positions are often unacceptable (Kayne 1983; Munn 1992):

(7) *Unacceptable PGs in subject position*

a. Who$_1$ did you slap t_1 [because **they/*PG**$_1$ ate your lunch?]

b. That's the guy who$_1$ I fired t_1 [after **he/*PG**$_1$ insulted me.]

c. What$_1$ will you eat t_1 [if **it/*PG**$_1$ is confirmed to be healthy]?

However, it will also be important that subject PGs sometimes succeed, as we'll see. I will argue that the facts about PGs and subjects support the proposal that a principle such as anti-locality bans clause-internal subject A-bar movement.[2]

1.1 Contents of the paper

Next, section 2 provides background on anti-locality through an explanation of its relation to the *that*-trace effect, which will be relevant later on this paper. Section 3 overviews the basic properties of PGs, and explains why they are relevant to the topic of anti-locality. Section 4 provides an analysis of the contrast between subject and object PGs, and also discusses a variety of related facts and predictions that clarify the nature of these phenomena. Section 5 extends these considerations to PGs in PPs, which I argue are also constrained by anti-locality. Section 6 concludes.

2 Background on anti-locality

While the PG evidence will suggest that clause-bounded A-bar movement of subjects usually doesn't occur, it is clear that cross-clausal subject A-bar movement does:

(8) *Subject wh-movement from a lower clause*
Who$_1$ did you say [t_1 is silly]?

However, when the subject of an embedded clause moves away, that clause cannot have an (overt) complementizer. This is known as the *that-trace effect* (Perlmutter 1968; Pesetsky 2017).

(9) *The that-trace effect*

a. **Who**$_1$ did you say [**(*that)** t_1 is silly]?

b. That's the person **who**$_1$ I think [**(*that)** t_1 should leave]

[2]Anti-locality is not the only way to account for the basic patterns I focus on, but I argue that it helps predict certain details that we will see later. In particular, the hypothesis that string-vacuous movement is banned (George 1980; Chomsky 1986, a.o.) would also predict at least some of the facts I will discuss today. Previous literature also argues that some of the facts I discuss here emerge from the *Empty Category Principle* (ECP). I set aside further discussion and comparison of these alternative analyses for the meantime.

In contrast, cross-clausal movement of a non-subject is compatible with the presence of a complementizer:

(10) *Complementizer allowed with non-subject movement*

a. **What**$_1$ did you say [$_{CP}$ **(that)** you want t_1]?

b. **Where**$_1$ do you think [$_{CP}$ **(that)** we should go t_1]?

Therefore it is clear that this restriction is specifically about subject movement. Furthermore, Bresnan (1977) observed that there is a way around the *that*-trace effect—adding an adverb after the complementizer:

(11) *Additional adverb repairs the that-trace effect*

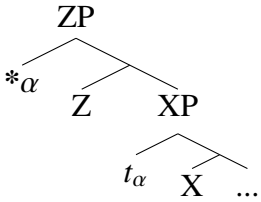
a. **Who**$_1$ did you say [$_{CP}$ **(that)** unfortunately t_1 is not very smart at all]?

b. That's the person **who**$_1$ I heard [$_{CP}$ **(that)** just yesterday t_1 bought a duck]

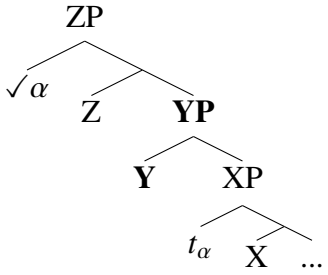
Several recent works have argued that the *that*-trace effect, and its repair by the addition of an adverb, is attributable to the anti-locality constraint I introduced above, illustrated again below:

(12) *Anti-locality*

a. *A movement that is too short*



b. *Movement made possible by crossing more structure*



The anti-locality account of the *that*-trace effect depends on the interaction of anti-locality and *phase theory* (Chomsky 2000, 2001; Citko 2014, a.o.). The essence of phase theory is that syntactic structures are built in a "chunk-by-chunk" manner, due to the way that the syntactic derivation is related to the other components of the grammar (phonology, semantics, etc). Such chunks are termed "phases", widely regarded to include CP, vP, and often DP.[3] One of the characteristic properties attributed to phases is that, when something moves from a phase, it must reach the phase edge before moving further. If CP is a phase, it is thus necessary for movement to reach spec-CP before exiting CP:

[3]Though there are many unresolved issues about which phrases count as phases. See Davis (2020a,b) for discussion.

(13) *Movement to CP edge feeds further movement*
✓ What did you say [$_{CP_{[Phase]}}$ t that [$_{TP}$ you ate t]]?

Importantly, when we attempt to extract the subject of an embedded CP, anti-locality and phase theory predict a conflict. If movement of a *wh*-subject through spec-CP is required, but anti-locality prevents movement from spec-TP to spec-CP, then we expect the derivation to fail:

(14) *Embedded subject movement causes a phase/anti-locality conflict*
* **Who**$_1$ did you say [$_{CP_{[Phase]}}$ t_1 **that** [$_{TP}$ t_1 ate the beans]]?

This prediction fits the description of the *that*-trace effect. If embedded clauses without *that* are bare TPs (Doherty 1997; Brillman and Hirsch 2016), then for such clauses both the phase problem and the anti-locality problem are irrelevant.[4] In this case, we correctly predict that the embedded subject can be extracted:

(15) *Subject extraction from CP-less clause succeeds*
Who$_1$ did you say [$_{TP}$ t_1 ate the beans]?

What about the fact that the addition of an adverb circumvents the *that*-trace effect? If the addition of an adverb below C introduces more structure between TP and CP, then we predict that anti-locality will not prevent movement from spec-TP to spec-CP in this situation (Brillman and Hirsch 2016; Erlewine 2017, 2020):

(16) *Adverb repairs that-trace effect by introducing more structure*
Who$_1$ did you say [$_{CP_{[Phase]}}$ t_1 **that** [$_{XP}$ unfortunately [$_{TP}$ t_1 ate the beans]]]?

We now have a theory for the *that*-trace effect and its avoidance. The concepts mentioned here will be relevant at many points in this paper's analysis. Before proceeding to the account, it is first necessary to make clear why PGs are relevant to anti-locality in the first place. I do this next.

3 Why parasitic gaps are relevant

Typical phrasal movement leaves behind an obvious corresponding gap, which in current syntactic theory is usually marked *t* for "trace":

(17) *Typical movement leaves behind a gap*
What$_1$ did you eat t_1?

A characteristic property of such movement is that it cannot exit certain constituents, which are termed "islands":

[4]Erlewine (2017) offers an alternative version of this account in which CP is not necessarily absent, but must be silent in order to prevent a linearization problem, building on Fox and Pesetsky (2005).

(18) *Some islands*

a. *Adjunct island*
   * [Whose birthday]$_1$ did you cry [because I forgot t_1]?

b. *Subject island*
   * Who$_1$ do [pictures of t_1] scare you?

c. *Complex NP island*
   * [How many hotdogs]$_1$ did you hear a rumor [that I managed to eat t_1]?

This means that we typically do not expect to see a moved phrase and its corresponding gap separated by an island. For my purposes it is convenient to focus on clausal adjuncts, which are often islands. Some of these are stronger islands than others, but nevertheless, many of them indeed clearly block or degrade movement:

(19) *Clausal adjunct islands*

a. *?? Tell me [**which paper**]$_1$ you ate fried chicken for lunch [after giving them comments on $\boldsymbol{t}_1$].

b. *?? [**What assignment**]$_2$ did you go home [because you need to finish $\boldsymbol{t}_2$ tonight]?

c. *?? I think I know [**what kind of pet**]$_3$ you'd move out of town [if your roommate bought $\boldsymbol{t}_2$].

However, if there is a well-formed A-bar movement elsewhere in the structure, it is often possible for an island in that structure to have a gap co-referent with the moved phrase. This is exactly what a PG is.[5] PGs are very productive in clausal adjuncts:

(20) *PGs in clausal adjuncts*

a. Who$_1$ did you forget about t_1 [after talking to **PG**$_1$]?

b. [What kind of cake]$_3$ would you eat a piece of t_3 [if I decided to bring **PG**$_3$ to the party]?

c. Who$_1$ did you tell t_1 about our idea [in order to impress **PG**$_1$]?

d. Tell me [which paper]$_1$ I should read t_1 [before giving you comments on **PG**$_1$]

e. This is a dish [$\varnothing_2$ that I know a lot about t_2 [because I make **PG**$_2$ every week]].

Why can a PG, and the moved phrase that it is associated with, be separated by an island? Much previous literature has argued that this is because PGs do not involve movement from an island, but rather A-bar movement of a null operator within the island (Contreras 1984; Stowell 1985; Chomsky 1986; Browning 1987; Nissenbaum 2000, a.o.).[6] This means that what we call a PG is

[5] PGs do not occur only in islands, but using an island makes it clear that a given gap is indeed parasitic.

[6] The null operator approach to PGs is in contrast to "shared antecedent" theories, for which PGs involve genuine extraction of a variety resembling the Across-The-Board (ATB) movement from coordinate structures. As Nissenbaum (2000) and Nissenbaum and Schwarz (2011) discuss, asymmetries in reconstruction for principle A, principle C, and variable binding all show that PGs involve a separate operator, and are thus not reducible to ATB extraction configurations. Additionally, as Culicover and Postal (2001) discuss, there is a consensus in the literature that at least in English PGs are nominals, though ATB movement is not category-specific in this way, further supporting the distinctness of PGs and ATB gaps. Munn (2001) argues for a unification of PGs and ATB contexts that makes a different distinction: Munn proposes that PGs involve null pronominals (equivalent to the null operators mentioned above), and that some instances of ATB movement are in fact PG-like null pronoun configurations.

just the trace of a silent operator's movement:

(21) *Operator movement within containing island forms PG*
Who$_1$ did you forget about t_1 [OP after talking to t_{OP}(=PG$_1$)]?

How do we know that this operator actually moves inside the island? If it does need to move, we predict that a PG will fail if we place another island inside of the first, in such a way that it would block the operator's movement. In other words, while we have seen that an island can separate a PG from the moved phrase that it matches, we expect that a PG cannot be separated by more than one island. Many previous works have shown that this is indeed the case (Kayne 1983; Chomsky 1986; Cinque 1990; Postal 1994), as the following example shows by combining an adjunct and relative clause:

(22) *PG-forming operator cannot move from a second island inside the first*
* Who$_1$ did you insult t_1 [OP after meeting a guy [who likes t_{OP}(=PG)]]?

Here are a few more relevant examples:

(23) *PG licensing across multiple islands fails*

a. *Relative clause island plus adjunct island*
* Who$_1$ did you talk to t_1 [after meeting someone [who knows **PG**$_1$]].

b. *Subject island plus adjunct island*
* Durian is a fruit [which$_1$ I tried t_1 for the first time [after [every variety of **PG**$_1$] was sent to me by someone who really likes them]].

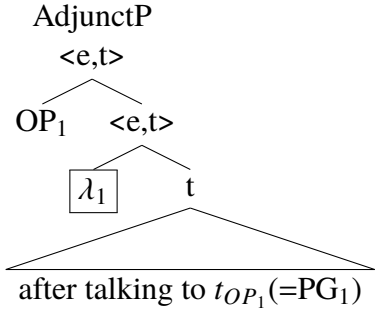
c. *Adjunct island in adjunct island*
* Guess who$_1$ I ironically ran into t_1 [after taking the other hallway [because I wanted to avoid **PG**$_1$]].

I will thus assume that PGs require movement of a null operator within the island. By exploring the constraints on PGs, we can find out whether this silent phrase's movement verifies the predictions of anti-locality or not. Before we do that, though, it will be useful to say a little more about the motivation for the operator movement that facilitates PGs.

3.1 The operator must move for semantic reasons

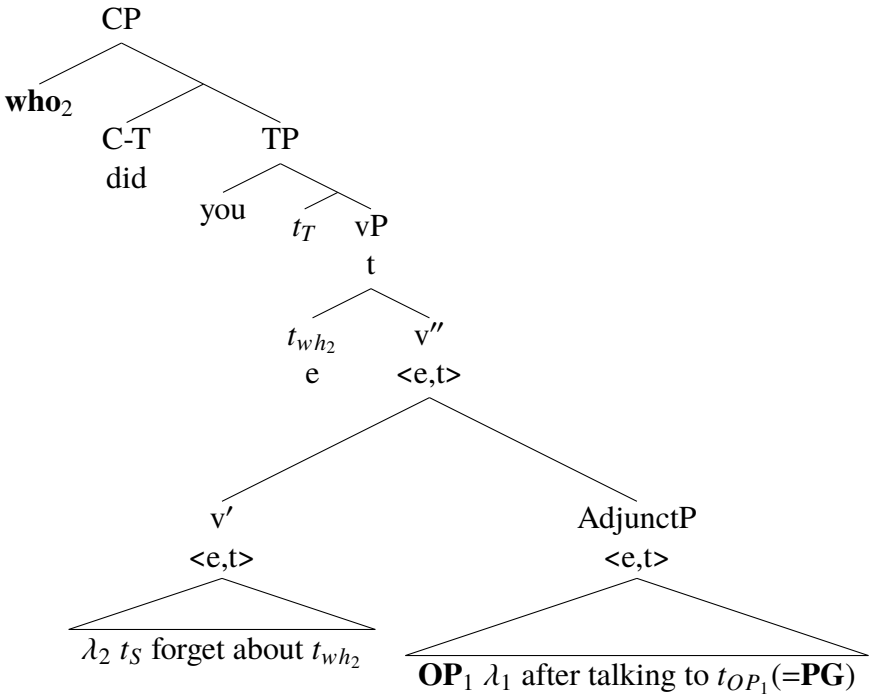
Nissenbaum (2000) argues that PG-formation requires the operator to move to the edge of the island for semantic reasons. Specifically, Nissenbaum argues that this movement must occur in order to trigger the semantic rule of Predicate Abstraction (Heim and Kratzer 1998). Though the operator is itself semantically content-less, when it moves and triggers this rule, it makes the island into an unsaturated predicate—a function with an empty individual argument position:

(24) *Semantic effect of operator movement for an adjunct island*



When this constituent merges in a structure containing an independently well-formed movement chain, that moved phrase can saturate this predicate, filling in the missing semantic argument of the function. This results in the trace of the operator becoming co-referent with the "true" gap (both of which co-refer with the moved phrase), creating what we call a PG. Nissenbaum (2000) argues that this process successfully occurs when a PG-containing clausal adjunct merges in the edge of the vP phase, through which the PG-licensing phrase moves, leaving a semantic reflex via Predicate Abstraction which allows the vP and the PG-containing adjunct to combine via Predicate Modification. This derivation relies on the same principles as the analysis of relative clauses in Heim and Kratzer (1998): operator movement, Predicate Abstraction, Predicate Modification, in addition to the successive-cyclic movement predicted by phase theory.

(25) *Syntactic/semantic derivation for a PG in a clausal adjunct*



I set aside the details of the semantic derivation here. All that matters for this paper is that there is a semantic reason why PG formation requires the null operator to reach the edge of its island. If such movement did not occur, Predicate Abstraction would not apply to the PG-containing adjunct, which would thus be of the wrong semantic type to merge in a tree like (25) above.

Importantly, if the operator must move to the edge of the PG-containing adjunct, then if that

movement would conflict with anti-locality, we expect a corresponding PG to be unacceptable. I argue that the facts about how PGs and subjects interact verify this prediction, in such a way that indicates that clause-internal subject A-bar movement is usually banned.

4 Analyzing the interaction of subjects and parasitic gaps

Culicover and Postal (2001) note that there is a tendency in the literature to conclude that subjects and PGs do not interact, or at least do so in a restricted way. Though the discussion of this topic is scattered, important observations about it were made in the very first article on PGs—Engdahl (1983). Engdahl pointed out that, assuming that *wh*-subjects do undergo some clause-internal A-bar movement, it does not appear that such movement can license PGs:

(26) *If clause bounded subject A-bar movement exists, it doesn't license PGs*

a. [Which articles]$_1$ t_1 got filed by John [without him reading them/*PG$_1$]? (Engdahl 1983, ex. 53)

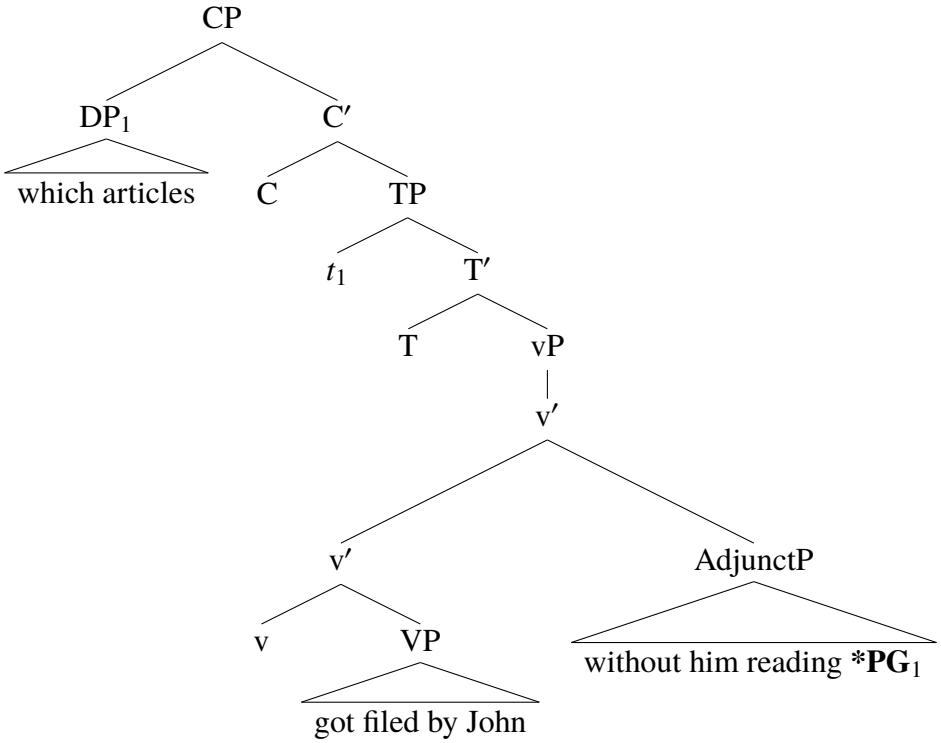
b. * That's the person [$_{CP}$ who$_1$ t_1 fired me [because I insulted PG$_1$]]

c. * Tell me [$_{CP}$ what$_1$ t_1 scared you [when you found PG$_1$ under the bed]]

If anti-locality bans such movement, then we correctly make the prediction that PGs here should fail. However, Engdahl identifies another reason why PG licensing should not work here. To paraphrase, A-bar movement of the subject from spec-TP to spec-CP would not actually structurally cross over the adjuncts in (26), assuming that they attach to the VP (in the updated theory in Nissenbaum (2000), the vP). As Nissenbaum discusses in detail, the PG-containing island needs to be attached within the movement path of the licensing phrase, otherwise semantic composition will fail.[7]

[7]Engdahl argued that it is important that the "true" gap does not c-command the PG. This constraint has come to be known in the literature as the *anti-c-command condition*. This condition is subject to a number of interesting qualifications, as Nissenbaum discusses. In my opinion it is more straightforward to make the generalization that the PG-container must be structurally crossed by A-bar movement of the licenser, since all interpretable PG structures I know of fit this description.

(27) *A-bar movement doesn't cross island → No PG*



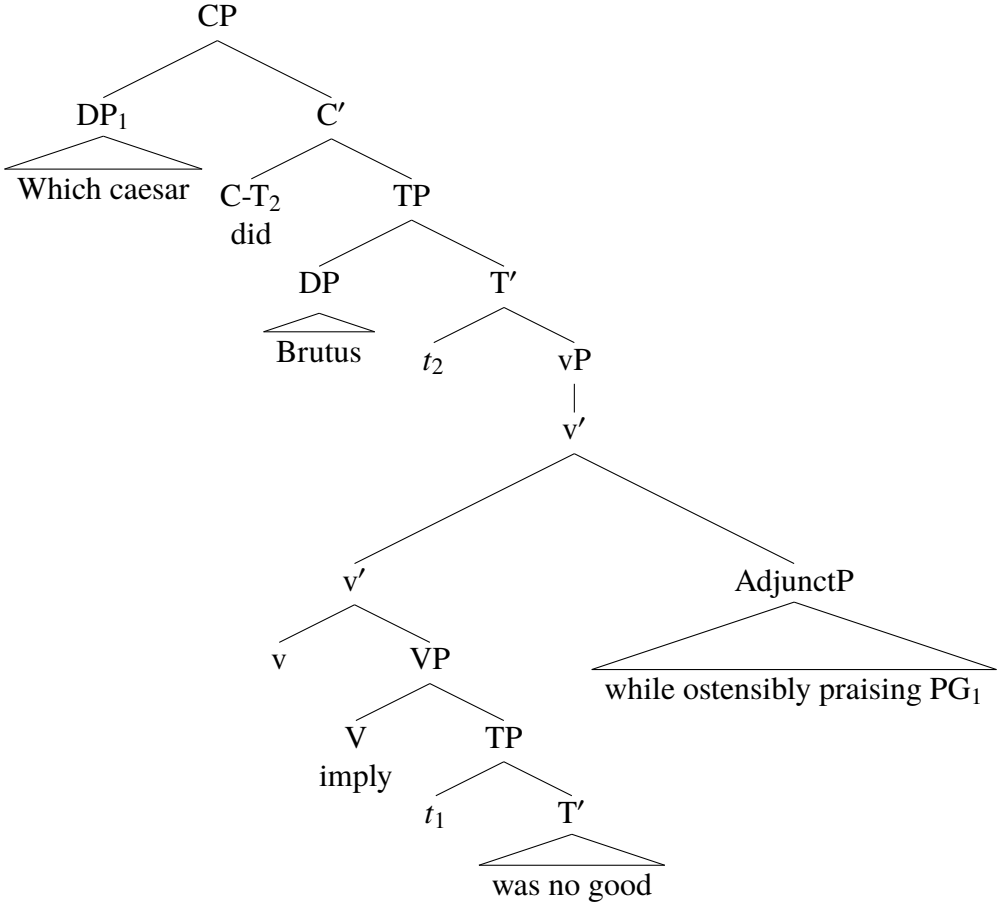
If PG-containing adjuncts merge in the vP, we predict that we should be able to get subject A-bar movement to license a PG by doing the following: Build a bi-clausal structure, where the PG-containing adjunct attaches to the higher vP. Extract the lower subject into the edge of the main clause, thus crossing that adjunct. Engdahl reports an example that verifies this prediction, and based on my research so far, such configurations do generally seem acceptable:

(28) *Cross-clausal subject extraction licenses a PG in the main clause's adjunct*

a. [Which caesar]$_1$ did Brutus [imply [t_1 was no good] [while ostensibly praising PG$_1$]]? (Engdahl, ex. 60)
b. Remind me who$_1$ you [found out [$_{TP}$ t_1 likes cats] [after talking to PG$_1$ about animals]]
c. This is the guy who$_1$ I [said [$_{TP}$ t_1 is stupid] [because I wanted to insult PG$_1$]]

Here's a tree to illustrate:

(29) *Successful PG licensing in main clause by cross-clausal subject extraction*



If clause-bounded subject A-bar movement is banned, then it is expected that cross-clausal movement as in the above tree will be the only way for a subject to license a PG.

So far in this paper, all PG examples have involved non-subject PGs. We've seen that (when the structure is right) such PGs can be licensed either by non-subject movement (20) or subject movement (28). Next let's examine PGs in subject positions, which are more significant.

While PG-licensing by subject movement is possible in principle, we've seen that it is more restricted. Therefore in order to achieve licensing of a subject PG, the safest strategy will be to first attempt licensing by movement of a non-subject. It turns out that non-subject A-bar movement cannot license a PG in the subject position of a mono-clausal adjunct:

(30) *Non-subject movement fails to license PG in subject of mono-clausal adjunct*

a. Who$_1$ did you slap t_1 [because **they/*PG**$_1$ ate your lunch?]

b. What$_1$ will you eat t_1 [if **it/*PG**$_1$ is discovered to be healthy]?

c. That's the guy who$_1$ I fired t_1 [after **he/*PG**$_1$ insulted me]

While I will argue that anti-locality predicts this fact, first I will consider a potential confound. In some languages, it has been observed that there is a requirement for a PG, and the moving phrase that licenses it, to match in case / semantic role. See for instance Kiss (1985) on Hungarian, and Franks (1992, 1993, 1995) on Russian and other Slavic languages. If this is also true for English, then perhaps the configuration in (30) above is no good due to the mismatch between subject and

non-subject. However, Engdahl shows that for English there are acceptable examples like (28a) above, repeated below, where subject movement licenses a non-subject PG.

(31) *A PG succeeding despite subject / non-subject mismatch*
[Which caesar]$_1$ did Brutus imply [t_1 was no good] while ostensibly praising PG$_1$?

If a mismatch in case or semantic roles were the issue with the examples in (30) above, we would expect the configuration in (30) to improve when we try to license the subject PG with subject A-bar movement. To give this configuration the best chance of succeeding, we should use cross-clausal subject movement, which we've seen in (28) above can license at least non-subject PGs. Even when we control for these factors, a PG in the subject position of a mono-clausal adjunct fails:

(32) *Subject movement cannot license subject PG in a mono-clausal adjunct*

a. Who$_1$ did you say [t_1 is a jerk] [because **they/*PG**$_1$ ate your lunch?]

b. That's the guy who$_1$ I will suspect [t_1 hates dogs] [if **he/*PG**$_1$ turns out to have a cat].

c. Remind me what$_1$ you told us [t_1 is a bad idea to eat] [after **it/*PG**$_1$ gave you a stomachache]

Since it is clear that a matching violation is not responsible for this unacceptability, we have good reason to instead look for a structural problem. A few previous works such as Kayne (1983) and Munn (1992) note the same fact (though do not control for subject versus object status to make sure that this is indeed a structural issue). I will argue that anti-locality can be productively invoked as the constraining structural factor.

Recall that as described in the previous section, PGs are formed by movement of an operator from the PG position, to the edge of the island:

(33) *Operator movement within containing island*
Who$_1$ did you forget about t_1 [OP after talking to t_{OP}(=PG$_1$)]?

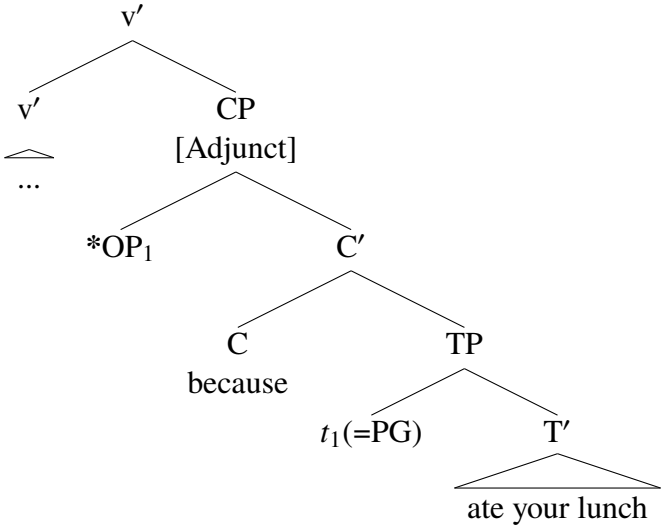
In the case of a PG in the subject position of a mono-clausal adjunct, it would be necessary for the operator to move from spec-TP to the edge of the island. I hypothesize that such clausal adjuncts are CPs, which are headed by words like *because, after, if* and so on. To form a subject PG in such adjunct CPs, it would be necessary for an operator to move from spec-TP to spec-CP. However, such movement is banned by anti-locality:[8]

(34) *Operator movement from subject position within island is impossible*
* Who$_1$ did you [$_{vP}$ say [$t_1$is a jerk] [$_{CP}$ OP because [$_{TP}$ t_{OP}(=PG$_1$) ate your lunch]]]?

Thus anti-locality accurately predicts the unacceptability of PGs in the subject position of mono-clausal adjuncts.

[8]It would not matter if words like *because* and *after* are in fact instances of P in these structures, since movement from spec-TP to spec-PP would still be banned by anti-locality as I argue.

(35) *Anti-locality blocks subject operator movement in mono-clausal adjunct*



This theory predicts that subject PGs should succeed when the PG is the subject of an embedded clause in a bi-clausal adjunct. This is because operator movement from the lower TP to the higher CP in a bi-clausal adjunct would not violate anti-locality (assuming no CP in the embedded clause). There are a few examples from previous literature which fit this description:

(36) *PGs in embedded subject position* (see also Munn (1992), ex. 49)

a. ? This is the student $\varnothing_1$ everyone thinks t_1 is clever [because John said PG_1 was clever]
(Engdahl, ex. 59)

b. ?? the person $\varnothing_1$ that you consulted t_1 [because you thought PG_1 understood the problem]
(Browning 1987)

Though complex, at least some instances of this configuration seem acceptable, clearly more so than examples with PG subjects in mono-clausal adjuncts.

(37) *More PGs in embedded subject position*

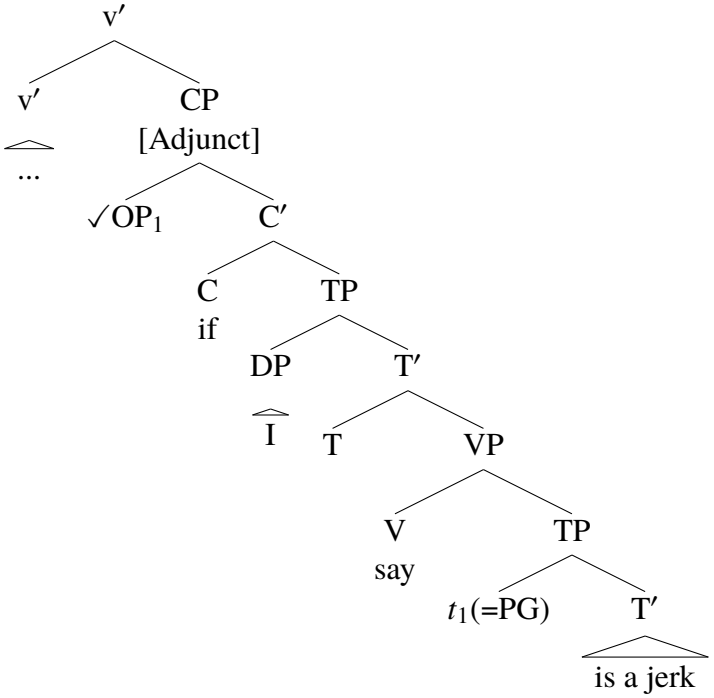
a. Who_1 did you avoid t_1 [after Mary said (*that) [$_{TP}$ PG_1 is a jerk]]?

b. This is a snack $\varnothing_1$ I eat t_1 every day [since I suspect (*that) [$_{TP}$ PG_1 improves my digestion]]

c. Let me tell you [which students]$_1$ I punished t_1 [after finding out [PG_1 have been stealing my cookies]].

That such examples should be better is exactly what we expect:

(38) *Operator movement from embedded subject position respects anti-locality*



In summary: PGs fail in the subject position of mono-clausal adjuncts. Anti-locality accounts for this fact, since it predicts the impossibility of operator movement from spec-TP to spec-CP within the adjunct clause. This theory also predicts that PGs in embedded subject positions should improve, since operator movement is long enough to be legal in this situation.[9]

4.1 An accurate prediction about anti-locality avoidance

In section 2 above, I summarized a theory in which the *that*-trace effect stems from anti-locality, which can be avoided by the inclusion of an adverb between TP and CP:[10]

[9]The analysis presented here is also compatible with an ATB extraction analysis of PGs. Under such an analysis, the normal gap and PG are both formed by genuine movement paths, which unite at a higher point in the structure, resulting in one moved phrase visible on the surface which corresponds to two gaps. Assuming that CP is a phase, the movement path within the adjunct clause would need to reach spec-CP before moving on out of the adjunct. However, if that movement is initiated from spec-TP, anti-locality will prevent such a derivation from succeeding.

[10]We predict that in examples like (36-37), the *that*-trace effect should apply to the operator movement from embedded subject position, and thus prevent the embedded clause in the adjunct from having a complementizer. Munn (1992) provides an example verifying this prediction. Furthermore, given the discussion of the *that*-trace effect in section 2 above, we expect use of an adverb below the embedded complementizer to repair such examples. Tentatively I claim that this is correct:

(i) a. Who$_1$ will you think t_1 is a jerk [if I say (*that) PG$_1$ is a jerk]?
b. This is a snack $\varnothing_1$ that I eat t_1 every day [because I suspect (*that) PG$_1$ might be good for me]
c. Let me tell you [which students]$_1$ I punished t_1 [after sadly finding out (*that) PG$_1$ have been stealing my cookies].

(39) *The that-trace effect and its repair*

a. * **Who**$_1$ did you say [$_{CP}$ **that** t_1 is silly]?

b. **Who**$_1$ did you say [$_{CP}$ **that** unfortunately t_1 is not very smart at all]?

Specifically, recall that this effect arises due to anti-locality's ban on movement from spec-TP to spec-CP, though inclusion of an adjunct between TP and CP circumvents anti-locality as discussed:

(40) *Prediction of anti-locality: Movement from spec-TP to spec-CP cannot occur*

a. * [$_{CP}$ **Who** [$_{TP}$ t will eat the cake]]?

b. ✓ [$_{CP}$ [$_{TP}$ **Who** will eat the cake]]?

If adverbs allow circumvention of anti-locality by adding more structure, then we expect insertion of an adjunct between TP and CP to make clause-bounded subject movement possible. Furthermore, if such a configuration actually has subject movement in it, we should be able to detect that movement by placing a PG in the adjunct. Since following Engdahl (1983) and Nissenbaum (2000) a PG is only possible when the constituent that contains it is structurally crossed over by the licensing phrase, a successful PG in this situation should only be possible if subject movement from spec-TP to spec-CP really did occur. An example of precisely this sort is reported by Haegeman (1984), and the native speakers that I have consulted agree that this configuration is productive:

(41) *Adverb facilitating clause-internal subject movement (+PG)*

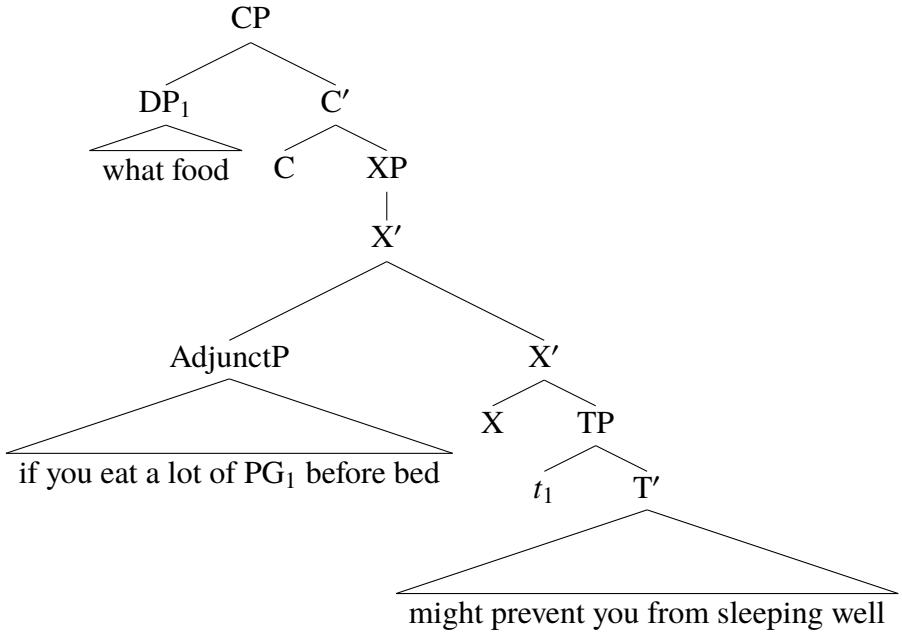
a. a note which$_1$ [unless we send back PG$_1$] t_1 will ruin our relationship
(Haegeman, ex. 9)

b. Let me tell you who$_1$, [despite nobody liking PG$_1$ at all], t_1 is probably gonna get promoted.

c. [What food]$_1$, [if you eat a lot of PG$_1$ before bed], t_1 might prevent you from sleeping well?

This is precisely what the anti-locality theory predicts. The inclusion of an adjunct between TP and CP should be able to co-occur with additional structure which permits such movement, as the following tree illustrates, and as the facts verify:

(42) *Clause-bounded subject movement permitted by intervening adjunct*



4.2 An incorrect prediction and a solution

I have argued that anti-locality prevents the formation of PGs in the subject position of mono-clausal adjuncts since the needed operator movement would be too short:

(43) *Failed operator movement from subject position within island*
* Who$_1$ did you [$_{vP}$ say [$t_1$is a jerk] [$_{CP}$ OP because [$_{TP}$ t_{OP}(=PG$_1$) ate your lunch]]]?

We predict that the addition of an adverb between TP and CP in the PG-containing adjunct should facilitate the needed operator movement. My research has shown that this prediction is in fact incorrect:

(44) *No PG in subject position, even with intervening adverb*
a. * Who$_1$ did you slap t_1 [because **unfortunately PG**$_1$ ate your lunch?]
b. * What$_1$ will you eat t_1 [if **eventually PG**$_1$ is confirmed to be healthy]?
c. * That's the guy who$_1$ I fired t_1 [after **surprisingly PG**$_1$ insulted me]

Adverbs are possible in the needed position, as we can see by replacing the PGs with pronouns:

(45) *High adverbs allowed in clausal adjuncts*
a. Who$_1$ did you slap t_1 [because **unfortunately they**$_1$ ate your lunch?]
b. What$_1$ will you eat t_1 [if **eventually it**$_1$ is confirmed to be healthy]?
c. That's the guy who$_1$ I fired t_1 [after **surprisingly he**$_1$ insulted me.]

Since such adverbs are independently legal, this fact is indeed a puzzle for the anti-locality approach I've adopted here.

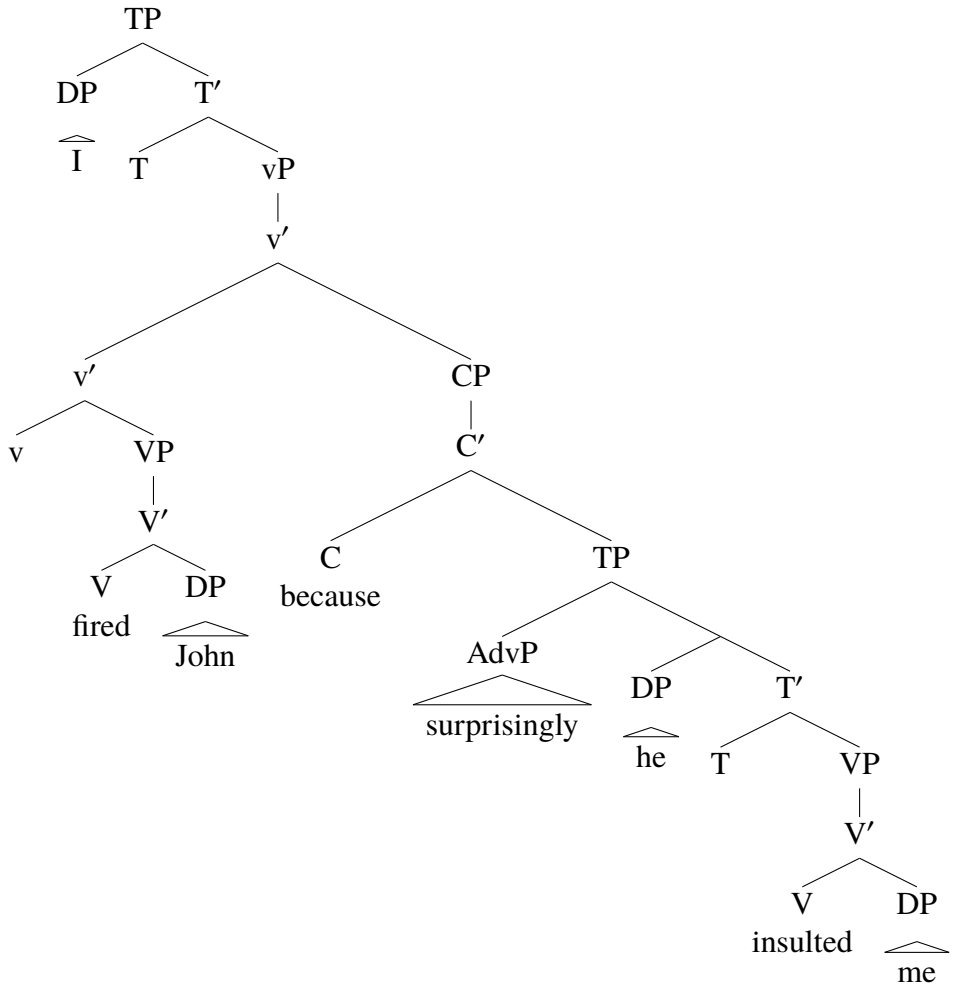
I suggest that this fact stems from a difference in the internal structures possible for typical CPs

headed by *that*, versus the sorts of adjunct CPs that can host PGs. As mentioned above, several relevant works argue that adverbs ameliorate the *that*-trace effect due to introducing additional structure between TP and CP:

(46) *Adverb resolves that-trace effect by introducing more structure*
Who$_1$ did you say [$_{CP}$ t_1 **that** [$_{XP}$ unfortunately [$_{TP}$ t_1 ate all the beans]]]?

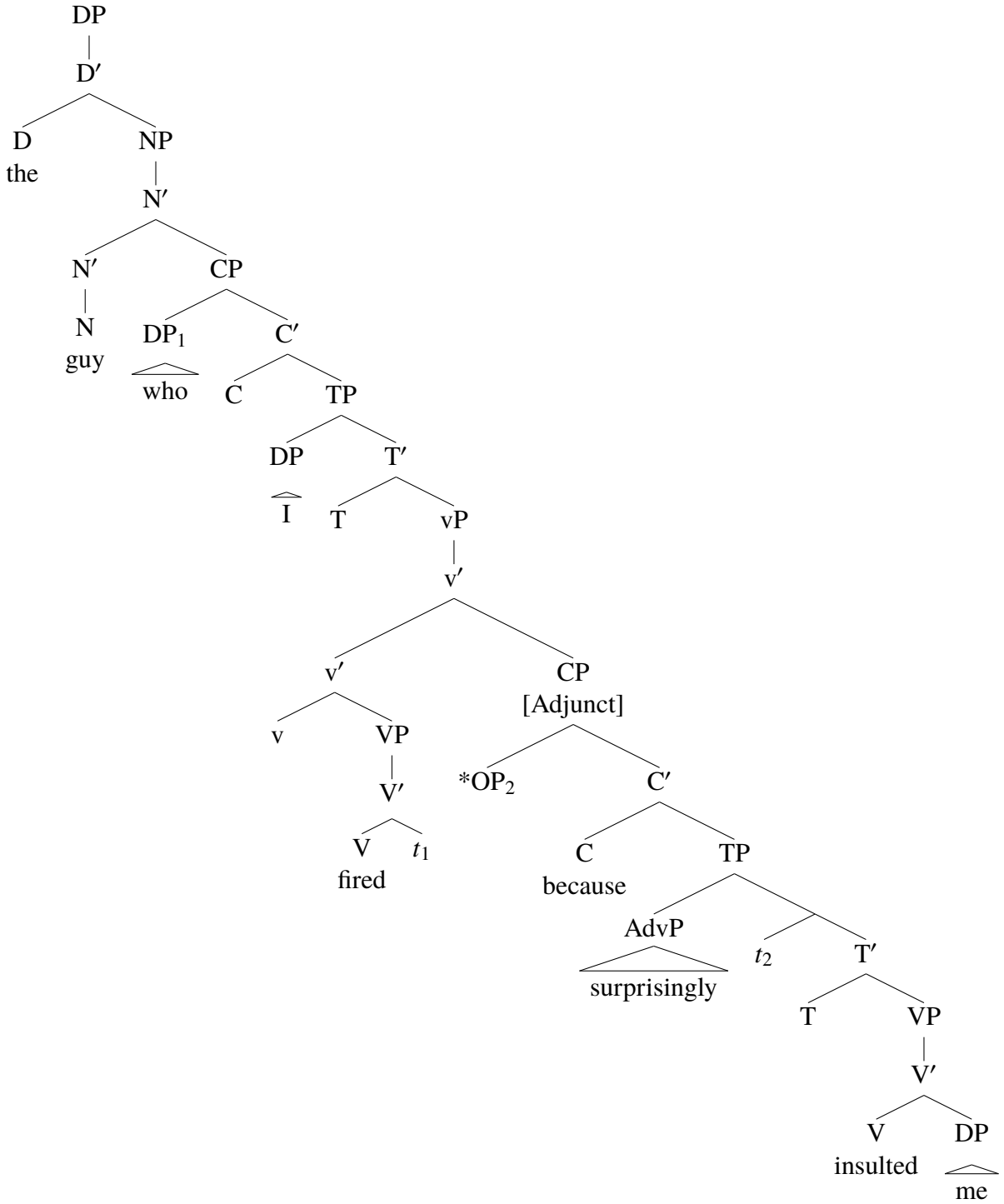
In the above structure, the presence of the XP containing the adverb is what is vital. I suggest that this XP cannot be merged in adjunct CPs. This would entail that when we do see a high adjunct in such CPs as in (45), it sits in the edge of the TP rather than being hosted by an additional projection:

(47) *High adjunct in adjunct CP attached in TP*



In this situation, the adjunct does not co-occur with structure that dominates TP. Instead, the adjunct is simply inside of the TP, but the TP is still immediately dominated by CP. This in this situation movement from the specifier of TP to CP will remain banned by anti-locality, as the following tree illustrates:

(48) *High adjunct in adjunct CP attached in TP doesn't circumvent anti-locality*



= * The guy who I fired because surprisingly PG insulted me

This analysis entails that the left periphery of the relevant clausal adjuncts is structurally impoverished, compared to *that*-CPs where adverb amelioration does succeed. Assuming that *that*-CPs are essentially matrix-like, this result is analogous to previous findings that generally, embedded clauses are relatively syntactically reduced (see for instance the *Penthouse Principle* of Ross (1973)).

5 Extension: Parasitic gaps in PPs

In this section, I will show how the concepts discussed above make the right predictions about another configuration, involving PGs in PPs. First, note that it is possible to have PGs in DPs:

(49) *PGs in DPs*

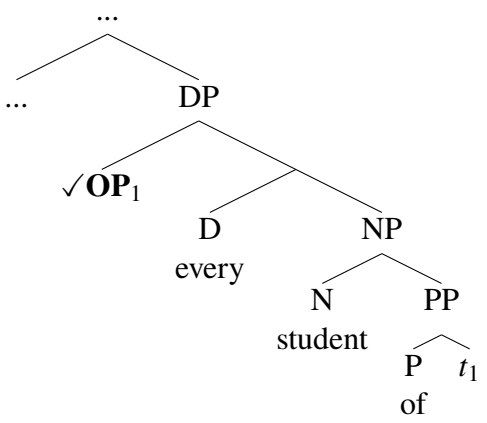
a. Who$_1$ would [every student of PG$_1$] love to throw a pie at t_1?

b. Tell me who$_1$ [a statue of PG$_1$] would surprise t_1

c. John's the guy who$_1$ I showed [the best friend of PG$_1$] a silly picture of t_1.

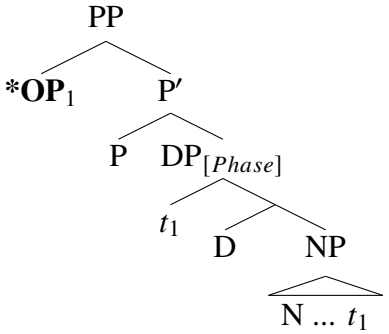
Under the operator theory of PGs, the examples above would need to involve movement to spec-DP from the complement of NP, which certainly obeys anti-locality:

(50) *Successful movement of OP within DP*



Let's consider what we predict for a configuration with a PG inside of a DP that is contained by a PP. It is common to assume that DP is a phase (Bošković 2005, 2016; Newell 2008; Newell and Piggott 2014; Syed and Simpson 2017; Simpson and Park 2019, a.o.). If so, a PG-forming operator would need to move through spec-DP on its way to the edge of PP in order to derive a PG in a DP in a PP. However, notice that this movement from spec-DP to spec-PP would violate anti-locality:

(51) *Operator movement from DP edge to PP edge: Predicted to be banned*



Consequently, we predict a PG inside of a DP that is in a PP to be unacceptable. This prediction appears accurate:

(52) *Attempted PGs in DPs in PPs*

a. * This is the guy who$_1$ it seems [to every student of PG$_1$] that I told a very mean joke about t_1

b. * Remind me [which student]$_1$ you told an awful rumor about t_1 [to every friend of PG$_1$]

c. * Tell me [which student]$_1$ you sent an awful picture of t_1 [to every friend of PG$_1$]

Thus in this domain as well, anti-locality leads us to the correct predictions about the distribution of PGs.

6 Conclusion

I've argued that facts about the interaction of PGs and subjects in English indicate that clause-internal subject A-bar movement is usually banned. I pursued an anti-locality approach to this ban, which I argued makes a number of correct predictions about when subject PGs will be either impossible or allowed. These results reveal that subjects and PGs interact in a principled and expected manner, with any gaps in the distribution of their interaction attributable to the independent influence of anti-locality.

6.1 Note about another analysis of subject A-bar movement

See Messick (2020) and references therein for discussion of the theory that subjects A-bar move directly from their θ-position to spec-CP, without passing through spec-TP.

(53) *Subject A-bar movement directly to spec-CP*

[$_{CP}$ **Who** [$_{TP}$ will [$_{vP}$ t eat the cake]]]?

Such a theory is not obviously compatible with the findings that I have discussed here, but there are nevertheless some interesting arguments that such subjects do indeed make it to spec-CP in at least some contexts. For instance, if A-bar movement in relative clauses occurs to trigger Predicate Abstraction which makes the relative CP the right type to combine with NP (Heim and Kratzer 1998), then for semantic reasons the *wh*-subject of a relative clause should be forced to move. While this presentation supports a theory in which subjects cannot usually move to spec-CP, it is possible that different A-bar constructions have other properties, and that such movement can be forced under certain conditions. Erlewine (2015) argues that anti-locality is not an absolute principle, but rather a violable constraint. If this is correct, then we indeed expect anti-locality to not always assert its influence. For the facts I have focused on in this paper, however, anti-locality appears to behave in a uniform way.

References

Abels, Klaus. 2003. *Successive-Cyclicity, Anti-locality, and Adposition Stranding*. Doctoral dissertation, University of Connecticuit.

Agbayani, Brian. 2000. Wh-subjects in English and the vacuous movement hypothesis. *Linguistic Inquiry* 31:703–713. Https://doi.org/10.1162/002438900554523.

Bošković, Željko. 1997. *The Syntax of Nonfinite Complementation: An Economy Approach*. Cambridge, MA: MIT Press.

Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studica Linguistica* 59.

Bošković, Željko. 2016. Getting really edgy: On the edge of the edge. *Linguistic Inquiry* 45. Https://doi.org/10.1162/LING_a_00203.

Bresnan, Joan. 1977. Variables in the theory of transformations. In *Formal syntax*, ed. Peter Culicover, Thomas Wasow, and Adrian Akmajian. Academic Press.

Brillman, Ruth, and Aaron Hirsch. 2016. An anti-locality account of English subject/non-subject asymmetries. In *Proceedings of Chicago Linguistic Society 50*, ed. Ross Burkholder, Carlos Cisneros, and Emily R. Coppess. Chicago Linguistic Society.

Browning, M. 1987. *Null Operator Constructions*. Doctoral dissertation, MIT.

Carstens, Vicki, Norbert Hornstein, and Daniel Seely. 2017. Head movement in Problems of Projection. *The Linguistic Review: Special issue on Labeling* 33:67–86. Https://doi.org/10.1515/tlr-2015-0014.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press. Linguistic Inquiry Monographs.

Chomsky, Noam. 2000. Minimalist Inquiries. In *Step by step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. Roger Martin, David Michales, Juan Urigareka, and Samuel Jay Keyser, 89–155. MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. In *Ken hale: A life in language*, ed. Michael Kenstowicz. MIT Press.

Chung, Sandra, and James McCloskey. 1983. On the interpretation of certain island facts in GPSG. *Linguistic Inquiry* 14:704–713.

Cinque, Guglielmo. 1990. *Types of a′-dependencies*. Cambridge, MA: MIT Press.

Citko, Barbara. 2014. *Phase Theory: An Introduction*. Cambridge University Press, Cambridge: UK.

Contreras, Heles. 1984. A note on parasitic gaps. *Linguistic Inquiry* 15:698–701.

Culicover, Peter, and Paul Postal, ed. 2001. *Parasitic Gaps*. Cambridge, MA: MIT Press.

Davis, Colin. 2020a. Crossing and stranding at edges: On intermediate stranding and phase theory. *Glossa* 5:1–32. Http://doi.org/10.5334/gjgl.854.

Davis, Colin. 2020b. *The Linear Limitations of Syntactic Derivations*. Doctoral Dissertation, MIT.

Doherty, Cathal. 1997. Clauses without complementizers: Finite IP complementation in English. *The Linguistic Review* 14:197–220.

Engdahl, E. 1983. Parasitic Gaps. *Linguistics and Philosophy* 6.

Erlewine, Michael Yoshitaka. 2015. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language & Linguistic Theory* 34. Https://doi.org/10.1007/s11049-015-9310-z.

Erlewine, Michael Yoshitaka. 2017. Why the Null Complementizer is Special in Complementizer-Trace Effects. In *A Pesky Set: Papers for David Pesetsky*, ed. Claire Halpert, Hadas Kotek, and Coppe van Urk, 232–288. Cambridge, MA: MITWPL.

Erlewine, Michael Yoshitaka. 2020. Anti-locality and subject extraction. *Glossa* 5:1–38.
Fox, Danny, and David Pesetsky. 2005. Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31. Https://doi.org/10.1515/thli.2005.31.1-2.1.
Franks, Steven. 1992. A prominence constraint on null operator constructions. *Lingua* 88:1–20.
Franks, Steven. 1993. On parallelism in across-the-board dependencies. *Linguistic Inquiry* 24:509–529.
Franks, Steven. 1995. *Parameters of Slavic morphosyntax*. Oxford Universty Press.
Gallego, Angel. 2017. The EPP in labeling theory: Evidence from Romance. *Syntax* 20:384–399. Https://doi.org/10.1111/synt.12139.
George, Leland. 1980. *Analogical generalization in natural langauge syntax*. Doctoral dissertation, Massachusetts Institute of Technology.
Grohmann, Kleanthes. 2003. *Prolific domains: On the anti-locality of movement dependencies*. Amsterdam: John Benjamins.
Haegeman, Liliane. 1984. Parasitic Gaps and Adverbial Clauses. *Journal of Linguistics* 20:229–232.
Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.
Ishii, Toru. 1999. Cyclic spell-out and the *that*-trace effect. In *Proceedings of WCCFL 18*, ed. A. Carnie, J. D. Haugen, and P. Norquest, 220–231. Somerville, MA: Cascadilla Press.
Kayne, Richard. 1983. Connectedness. *Linguistic Inquiry* 14:223–249.
Kiss, Katalin. 1985. Parasitic Chains. *The Linguistic Review* 5:41–74.
Messick, Troy. 2020. The derivation of highest subject questions and the nature of the EPP. *Glossa* 5:1–12.
Munn, Alan. 1992. A null operator analysis of ATB gaps. *The Linguistic Review* 9:1–26.
Munn, Alan. 2001. Explaining Parasitic Gap Restrictions. In *Parasitic Gaps*, ed. Peter Culicover and Paul Postal, 369–402. Cambridge, MA: MIT Press.
Newell, Heather. 2008. *Aspects in the morphology and phonology of phases*. Doctoral dissertation, McGill University.
Newell, Heather, and Glyne Piggott. 2014. Interactions at the syntax-phonology interface: Evidence from Ojibwe. *Lingua* 150:332–362. Http://dx.doi.org/10.1016/j.lingua.2014.07.020.
Nissenbaum, John. 2000. *Investigations of covert phrase movement*. Doctoral dissertation, MIT.
Nissenbaum, John, and Bernhard Schwarz. 2011. Parasitic degree phrases. *Natural Langauage Semantics* 19:1–38. Https://doi.org/10.1007/s11050-010-9061-7.
Perlmutter, David. 1968. *Deep and surface constraints on syntax*. Doctoral dissertation, Massachusetts Institute of Technology.
Pesetsky, David. 2017. Complementizer-Trace Effects. In *The Wiley Blackwell Companion to Syntax, Second Edition*, ed. Martin Evearet and Henk van Riemsdijk. Wiley-Blackwell.
Postal, Paul. 1994. Parasitic and pseudo-parasitic gaps. *Linguistic Inquiry* 25:63–117.
Ross, Haj. 1973. The penthouse principle and the order of constituents. In *You take the high node and i'll take the low node*, ed. Corum, Smith-Stark, and Weiser, 397–422. Chicago: Chicago Linguistic Society.
Simpson, Andrew, and Soyoung Park. 2019. Strict vs. free word order patterns in Korean and cyclic linearization. *Studica Linguistica* 73:139–174.
Stowell, Tim. 1985. Licensing conditions on null operators. In *Proceedings of the 4th West Coast Conference on Formal Linguistcs (WCCFL)*. Center for the Study of Language and Information (CSLI), Stanford University.

Syed, Saurov, and Andrew Simpson. 2017. On the DP/NP status of nominal projections in Bangla: Consequences for the theory of phases. *Glossa: a journal of general linguistics* 2:1–24. Https://doi.org/10.5334/gjgl.186.