# Anti-locality*

Norvin Richards, MIT

norvin@mit.edu

## 1. Locality and anti-locality

Locality is classically concerned with the proper relations between the participants in the movement operation. Much current work focuses on a requirement that a Probe find the closest possible Goal (see Branan and Erlewine, this volume, for discussion of requirements of this kind).

Once a Probe has found the closest possible Goal, where should the new position in the movement chain be created? One possibility is that the answer to this question has nothing to do with locality. In an approach that posits the Extension Condition of Chomsky (1993), for example, the position of newly Merged material is invariably fixed; the new copy of the movement chain will have to be in a position that extends the tree.

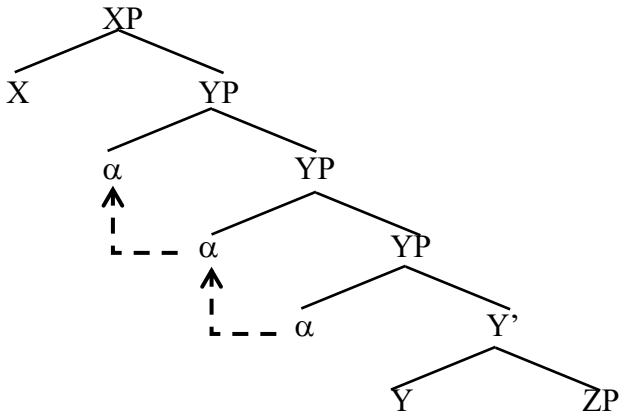
Another strand of work on locality posits conditions on possible landing sites, in addition to the conditions on the relations between Probes and Goals; movement operations are to land in positions that are as close as possible to the original position of the moved item (see, for example, the Minimize Chain Links principle of Chomsky and Lasnik (1993), and subsequent work along the same lines). We might imagine that the condition requiring the Probe to find the closest possible Goal is part of a more general condition on relations between points in the syntactic tree; such relations must always be as short as possible.

If we posit conditions along these lines, an important question arises: how short can movement operations be? Saito and Murasugi (1999), for example, working in a theory in which

* I am very grateful to the editors, to an anonymous reviewer, and to an audience at MIT for feedback on the ideas expressed here. The responsibility for any remaining mistakes is entirely mine.

movement can in principle land anywhere, point out that in such a theory, if movements are to be as short as possible, there must be some minimal length for movement operations (and see Bošković 1994, 1997 for related arguments). If there were not, then movement of a phrase sitting in the specifier of some YP, for example, might be forced to start by adjoining the phrase to YP itself, and in fact then to adjoin it again to a higher position in YP, ad infinitum, with the consequence that movement out of YP would be impossible.

(1)




We might take this difficulty as an argument that some of Saito and Murasugi's assumptions about movement are wrong: we might, for example, decline to join them in assuming that movement can in principle land anywhere, or we might decide that there is in fact no condition constraining the grammar to make movement steps as short as possible. But the case depicted in (1) demonstrates that if there is to be anything like Minimize Chain Links, there must also be some account of why movement operations cannot always be maximally short.

A substantial literature explores the possibility that movement dependencies have a minimum possible length. In what follows I will sketch definitions of anti-locality that have been offered in the literature (section 2), review some of the kinds of work that anti-locality has been taken to be able to do for us (section 3), and finally turn to the question of why anti-locality conditions should exist at all (section 4).

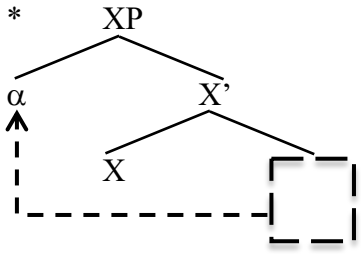
## 2. What is anti-locality?

Much of the current literature on anti-locality focuses on three proposals, namely those of Grohmann (2003), Abels (2003), and Erlewine (2016)[1].

For Grohmann (2003 and subsequent work), the clause is to be divided into three domains: the Ω-domain, corresponding to the CP layer; the Φ-domain, corresponding to the TP/IP layer; and the Θ-domain, corresponding to the VP layer. Grohmann's version of anti-locality bans movement within one of these domains; thus, for example, movement may cross from the Θ-domain into the Φ-domain (as it does when the subject moves from the internal to the external position), but movement within the Θ-domain is not possible (so phrases may not move from one theta-position to another).

Abels (2003) (and see Pesetsky and Torrego 2001 for a similar proposal) develops a version of anti-locality that rules out movement of the complement of XP to a specifier of XP:

(2) *



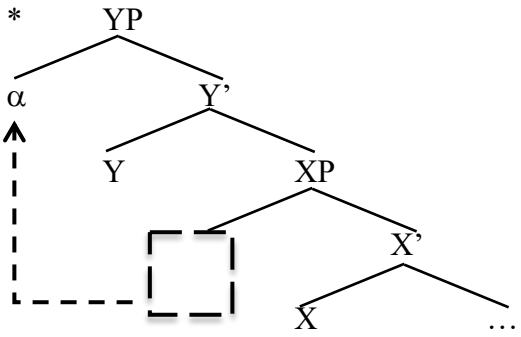

Abels makes the interesting point (to which we will return in section 4) that his version of anti-locality has the virtue of being plausibly relatable to a principle requiring operations to be motivated (which he refers to as Last Resort). In a tree like the one in (2), he claims, any relation between α and X which might be created by the movement operation ought to have already existed before the movement operation takes place. Prior to movement, X and α are already in a

[1] These are not the only proposals in the literature. Saito and Murasugi (1999), for example, require movement operations to cross one full node; by this metric, the movements in (1) are excessively short, since they cross only a segment of the YP node.

mutual c-command relation, and if we think that movement is driven by a need to check features, it is hard to see how the movement operation in (2) could improve the structure.

Erlewine (2016) proposes that movement of a phrase from the specifier of XP must cross a maximal projection other than XP:

(3)




A number of researchers have posited variations on Erlewine's (2016) proposal. For Brillman (2017), movement must cross a specifier of a phrase other than XP; for Zyman (2021), the condition in (3) holds just if XP is a phase.

**3. What does anti-locality do?**

The literature on anti-locality has found more possible uses for anti-locality than I can adequately cover[2]. Much of the literature centers on derivations which are blocked, putatively because they require movement operations which would be too short. I will concentrate on a few of these, but see Saito and Murasugi (1999), Ticio (2005), Bošković (1994, 1997, 2016), and Zyman (2021) for more discussion.

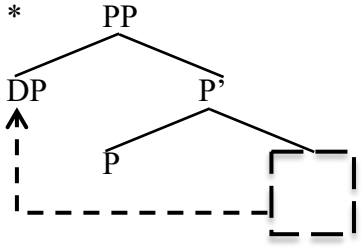
**3.1. Preposition stranding**

Abels (2003, 2012) offers an account of the widespread ban on preposition stranding which makes use of his version of anti-locality, together with the Phase Impenetrability Condition (PIC) of Chomsky (2000, 2001). The PIC states that at some point after a phase has been constructed,

[2] See Grohmann (2011) for a summary of the literature up to that point.

parts of it become inaccessible to further computation; in one formulation, everything c-commanded by the head of a phase becomes unavailable to syntactic operations. As Abels points out, if we pair this condition with his version of anti-locality, we arrive at the result that extraction of the complement of a phase head should be impossible once the phase has become impenetrable. For example, if PP is (at least in most languages[3]) a phase, then stranding of P should require movement of the complement of P to first become a specifier of P, so that it can escape phase impenetrability:

(4) *




The movement in (4), however, is blocked by Abels' version of anti-locality, and preposition stranding is therefore correctly blocked. More generally, Abels points out, we expect phase heads to be impossible to strand; the fact that movement of TP cannot strand C is another example of the same kind.

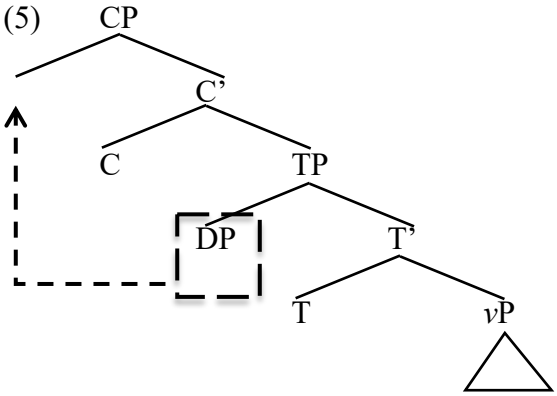
### 3.2. Subject extraction

Since subjects are closer to C than objects, ordinary extraction of the subject is in danger of being excessively local. On a theory positing anti-locality, we might therefore expect extraction of subjects to require extraordinary measures.

[3] Abels (2012) has an extensive discussion of how to account for the minority of the world's languages in which preposition stranding is possible. One possibility, which he rejects, is that they are languages in which PP is not in fact a phase. He endorses another approach, in which P takes as its complement, not the DP object itself, but an additional functional projection; DP may then move through the specifier of PP in a way which obeys anti-locality, leaving the additional functional projection behind.

Bošković (1994, 2016), Ishii (1999), Brillman and Hirsh (2016), Douglas (2017), Amaechi and Giorgi (2019), and Erlewine (2020) are among the works which apply this reasoning to complementizer-trace effects (including the English *that*-trace effect). These approaches assume a structure something like the one in (5):

(5)




Erlewine's (2016, 2020) Spec-to-Spec anti-locality would ban movement of the subject from the Specifier of TP to the Specifier of CP, in a tree like the one in (5), as excessively short. Movement of anything lower than the subject, by contrast, would be acceptable. Anti-locality therefore combines with existing assumptions about the structure of the clause to allow us to capture contrasts like the ones in (6):

(6) a. *Who do you think **that** __ left?

b. Who do you think **that** you saw __ ?

c. Where do you think **that** we should go __ ?

As Erlewine points out, an anti-locality account offers an explanation for the classic observation (Bresnan 1977, Culicover 1993, Rizzi 1997) that the effect in (6a) is obviated in the presence of high adverbs (Culicover 1993, 557):

(7) the man that Leslie said [that *(*for all intents and purposes*) __ was the mayor of the city]

On an anti-locality account, we can regard the italicized phrase in (7) as introducing enough structure between the subject position and the specifier of CP for movement between those two positions to be legitimate.

Of course, we must also capture the contrast between (6a) and (8):

(8) Who do you think __ left?

Here the literature entertains various possibilities. One is that the embedded clause in (8) literally lacks a CP layer; the subject moves directly into the matrix clause, perhaps skipping matrix VP and landing at the edge of matrix *v*P, obeying anti-locality. Another possibility, considered by Erlewine (2020), is that (8) represents an example in which C and T are Merged together as a single complex head, with the subject as its specifier; this option is not available for the examples in (8), since C and T can be seen to be pronounced in different heads[4].

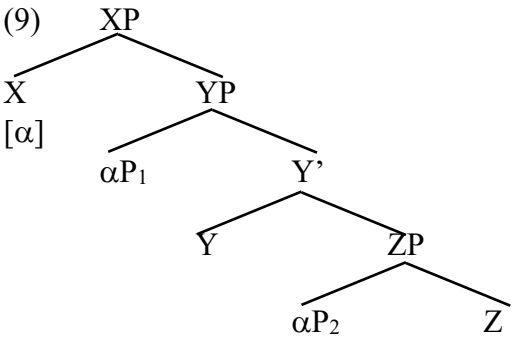
Another potential solution to the anti-locality problem in (5) would be to extract the subject from a lower position, rather than from the external subject position. Several researchers have proposed approaches to anti-agreement along these lines; wh-movement takes place from a position below the ordinary position of subjects, and wh-extracted subjects therefore do not trigger the kind of agreement on the verb that they ordinarily would (Cheng 2006, Schneider-Zioga 2007, Erlewine 2016). See also Baier (2017) for an argument that not every instance of anti-agreement can be handled in this way.

[4] The first of these approaches is tailored to languages like English, in which the version of C which allows subject extraction is not pronounced. There are also languages with complementizer-trace effects in which both versions of the complementizer are pronounced (Perlmutter 1971, Kandybowicz 2009, Sulemana 2018); such languages are potentially problematic for approaches of this kind.

### 3.3. Unexpectedly long movement

Some comparatively recent proposals argue that considerations of anti-locality may sometimes force the choice of the longer of two possible movements (Deal 2019, Ko 2020, Branan 2021). This work invites us to consider trees of the form in (9):

(9)




Here we have a head X bearing a probe searching for a feature α, with two potential Goals for the probe which are both c-commanded by X ($\alpha P_1$ and $\alpha P_2$). Ordinary conditions on locality would lead us to expect the probe to prefer the closer Goal, namely $\alpha P_1$. However, if anti-locality bans movement of $\alpha P_1$, the idea goes, the more distant Goal $\alpha P_2$ may undergo movement instead. For Deal (2019), this idea is part of an account of raising to ergative in Nez Perce; a phi-probe is able to raise the theme past a higher applicative argument, because movement of the applicative argument would violate anti-locality.

The cases of this kind that have so far been discussed all involve apparent A-movement, and a natural next question for this approach is whether A-bar movement ever behaves in this way. Given the anti-locality-based accounts of the conditions on subject extraction, we might expect to find languages in which Superiority effects behave as expected, except when a subject is involved, when a lower wh-phrase must always move past a subject wh-phrase if one is available.

## 4. What drives anti-locality?

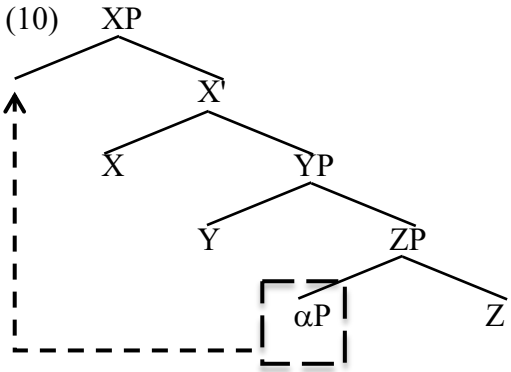
A point often made in the literature on anti-locality (see, for example, Abels 2003, Grohmann 2003, Baier 2017, Erlewine 2016, 2020, Brillman 2017, and Deal 2019) is that the predictions of a theory of anti-locality depend, not only on the formulation of anti-locality under consideration, but also on the architecture of the tree. For Abels (2003), for example, anti-locality is responsible for the fact that phase heads cannot be stranded by movement of their complements; this generalization can account for the ban on movement of TP that strands C, on the crucial assumption that C is a phase head that takes TP as its complement. If, on the other hand, CP is decomposed into a number of different projections in the "CP domain", and if the phase head in this domain is not the head that takes TP as its complement, then Abels' result no longer holds, as he points out. Similarly, Erlewine's (2020) Spec-to-Spec anti-locality would ban movement of the subject from the specifier of *v*P to the specifier of TP[5], if T takes *v*P as its complement[6]. Approaches to anti-locality like Abels' and Erlewine's, then, have the potential to constrain the possible kinds of trees we can posit, though I think it is fair to say that at this point in our understanding, the constraints are often straightforward ones to comply with, since we are comparatively poor in proposals which rule out the postulation of syntactic structure wherever it is theoretically convenient to do so.

One of Brillman's (2017) motivations for modifying Erlewine's (2016, 2020) version of anti-locality is to make it less sensitive to decisions we might make about clausal architecture. The movement in (10) obeys Erlewine's version of anti-locality, but not Brillman's:

[5] Erlewine (2016) deals with this problem by restricting his version of anti-locality so that it does not apply to A-movement.

[6] See also Newman (2020) for another proposal about how A-movement might interact with anti-locality; for her, the requirement that middles contain an adverb (*This bread cuts *(easily)*) follows from anti-locality, as the adverb provides enough structure for A-movement past it to be acceptably anti-local.

(10)




For Erlewine, movement of αP from the specifier of ZP to the specifier of XP is well-formed, because the movement escapes YP, and Erlewine's version of anti-locality requires only that αP move out of a phrase other than the ZP in which it began. Brillman's version of anti-locality is stricter; for movement of αP to be legitimate, the intervening YP must itself have a specifier. As Brillman points out, this version of Erlewine's anti-locality covers all of the ground that Erlewine intended it to; his well-formed examples invariably involve movement past intervening specifiers. Moreover, she says, her version of anti-locality should be impervious to future discoveries about clausal architecture, as long as these do not involve the discovery of new specifiers; no matter how many clausal heads are introduced to the clausal spine, the predictions of her version of anti-locality will remain the same[7].

As mentioned above, Abels' version of anti-locality has the virtue of being comparatively easy to think of as a consequence of some general condition that movement operations be driven by the need to change the structure in certain specified ways. In the remainder of this vignette I will outline a motivation for the particular kind of anti-locality developed by Brillman (2017). We will see that a version of Brillman's particular formulation of anti-locality, along with an account of the exceptions to it, will follow if we assume the motivations for movement posited in

[7] Deal (2019) and Erlewine (2020) point out that the comparative 'fragility' of Erlewine's version of anti-locality can be seen as a virtue; his version rules out certain kinds of possible claims about clausal architecture, and thus imposes stronger constraints on the possible theories we can entertain.

Contiguity Theory (Richards 2010, 2016, Branan 2018, Yamada 2019). In fact, Brillman (2017) herself suggested that this might be the case, and I will try to show that she was right.

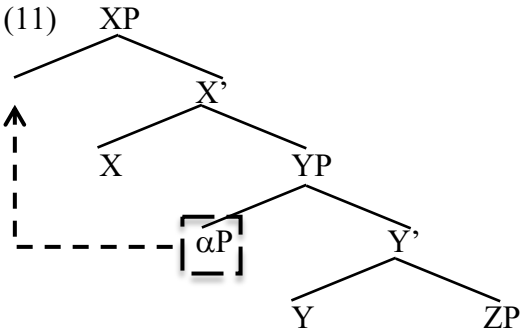
Contiguity Theory posits two motivations for phrasal movement, called *Affix Support* and *Contiguity*. Affix Support is a condition on affixes with certain morphophonological properties: it requires suffixes to be linearly preceded by material to which the rules for stress assignment in the language make reference. Affix Support is the Contiguity-theoretic version of the EPP; if a language has suffixal Tense, and the rules for stress in the language do not make reference either to the Tense suffix or to any morpheme before it, then a phrase (which will invariably contain stress) must move to a position preceding Tense (that is, to a specifier of TP). The account predicts the distribution of classic EPP effects in a range of languages: for example, English and French have EPP effects, and Italian, Spanish, and Catalan do not, because of differences in how stress is assigned in the verb in these languages (Richards 2016).

The second driving force for movement, Contiguity, can be thought of as a condition on the distribution of prosodic prominence: it says, in effect, that between a Probe and a Goal, there must be no phrases which are more prosodically prominent than the Goal. The effects of Contiguity depend on the conditions on the placement of prosodic prominence in a given language, and in some cases Contiguity may be satisfied without any movement at all. When Contiguity does trigger movement, it is because there is some phrase linearly between the Probe and the Goal which is more prosodically prominent than the Goal; the Goal must move past the offending phrase.

Suppose these are the correct conditions on movement, and suppose further that movement cannot take place unless it satisfies one of these conditions. What would this mean for anti-locality?

We would expect movement driven by *Affix Support* not to be subject to considerations of anti-locality; this type of movement is driven by a need to place a phrase in a position linearly preceding a head, and any movement that accomplishes this, no matter how short, ought to be licit. On the other hand, when movement is driven by *Contiguity*, the requirement is simply that the affected phrase be closer to the Probe than another phrase. Such a requirement ought not to be able to drive movement which does not cross an offending phrase. In the tree in (11), in other words, the movement shown could be driven by Affix Support, but not by Contiguity:

(11)



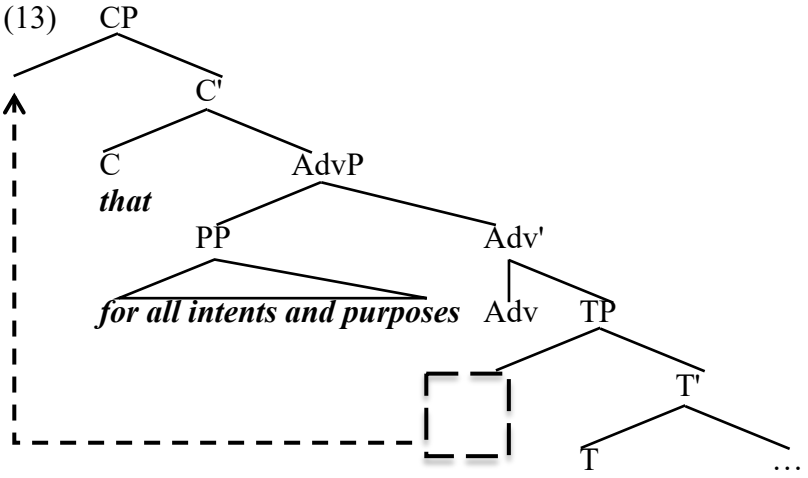

The movement in (11) does place a phrase in a position preceding X, and if X is in need of Affix Support, the movement in (11) could be licensed. But if the relation between X and αP is one of Contiguity, there is no need of movement to create this relation in the tree in (11); there are no phrases linearly intervening between X and αP. The movement in (11) would therefore achieve no purpose, and if movement is only allowed if it is motivated, movement in this case should be blocked. In the tree in (11), αP is the specifier of the complement of YP, but this result will hold for more complex trees as well; in general, we expect this kind of movement to be possible just if it crosses a linearly intervening phrase. This is, very nearly, Brillman's (2017) version of anti-locality.

A reviewer correctly points out a difference between Brillman's (2017) anti-locality and the one derived here, which can be illustrated using the effects of high adverbs on the *that*-trace effect, an example of which is repeated here in (12):

(12) the man that Leslie said [that *(*for all intents and purposes*) __ was the mayor of the city]

Recall that Erlewine's (2016, 2020) version of anti-locality requires movement out of XP to cross a maximal projection other than XP, and Brillman's (2017) version adds the requirement that the maximal projection other than XP which is crossed must itself have a specifier. For this version of anti-locality to account for the contrast in (12), we must imagine that the italicized adverb is the specifier of a dedicated projection:

(13)

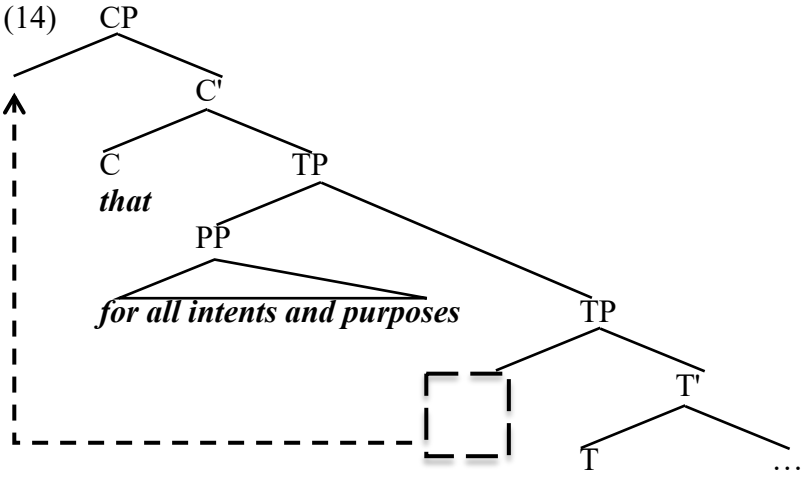



In the Erlewine and Brillman approaches to anti-locality, movement from the specifier of TP to the specifier of CP is legitimate in (13) because of the AdvP intervening between them; movement crosses a maximal projection other than the TP in which the subject began, which is what these versions of anti-locality require[8]. In the Contiguity-theoretic approach to anti-locality

[8] Brillman points out a difference in this regard between her approach and Erlewine's. For her, if the specifier of AdvP were not present, the *that*-trace effect would reappear, since AdvP cannot make movement of the subject sufficiently non-local unless it has a specifier. For Erlewine, she points out, movement of the subject would be possible in (12) even if AdvP had no specifier. In other words, it is important for Erlewine, though not for Brillman, that AdvP be present only when there is an actual adverb in the structure; adverbs must force the introduction of new material to the clausal spine.

just described, movement is anti-local if it crosses a specifier, and there is no need to posit an AdvP for adverbs to Merge into; an adverb adjoined to TP would also make movement of the subject possible, as long as the adverb was in a position linearly intervening between C and the subject:

(14)




Whether there is a strong reason to prefer one of the trees in (13-14) is a question I will have to leave for future research.

I have tried to show here that accepting Contiguity Theory as a theory of when movement is motivated allows us to explain the existence of a certain type of anti-locality effect, and also predicts that certain types of movement will be exempt from anti-locality. Movement of the subject to the specifier of TP in English, for example, is one of the kinds of movement that Richards (2016) attributes to the effects of Affix Support, and we therefore expect this particular kind of A-movement not to be subject to anti-locality. On the other hand, wh-movement is generally driven by Contiguity, and should therefore exhibit anti-locality effects, which has indeed been claimed to be true.

**References**


Abels, Klaus. 2003. *Successive cyclicity, anti-locality, and adposition stranding*. Doctoral dissertation, University of Connecticut.

Abels, Klaus. 2012. *Phases: an essay on cyclicity in syntax*. Walter de Gruyter, Berlin.

Amaechi, Mary, and Doreen Georgi. 2019. Quirks of subject (non-)extraction in Igbo. *Glossa* 4(1):69.1-36.

Baier, Nico. 2017. Antilocality and antiagreement. *Linguistic Inquiry* 48.367-377.

Bošković, Željko. 1994. D-structure, theta-criterion, and movement into theta-positions. *Linguistic Analysis* 24.247-286.

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

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* 33.17-66.

Branan, Kenyon. 2018. *Relationship preservation*. Doctoral dissertation, MIT.

Branan, Kenyon. 2021. Locality and anti-locality: the logic of conflicting requirements. *Linguistic Inquiry*.

Branan, Kenyon, and Michael Yoshitaka Erlewine. This volume. Locality and (minimal) search.

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

Brillman, Ruth. 2017. *Subject/non-subject extraction asymmetries: the view from* tough-*constructions.* Doctoral dissertation, MIT.

Brillman, Ruth, and Aron Hirsh. 2016. An anti-locality account of English subject/non-subject asymmetries. *Proceedings of CLS 50*.

Cheng, Lisa Lai-Shen. 2006. Decomposing Bantu relatives. In Christopher Davis, Amy Rose Deal, and Youri Zabbal (eds.), *Proceedings of NELS 36*, pp. 197-216. Graduate University of Massachusetts Graduate Linguistic Student Association, Amherst, Mass.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale and Samuel J. Keyser (eds.), *The view from building 20*, pp. 1-52. MIT Press, Cambridge, Mass.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In Roger Martin, David Michaels, and Juan Uriagereka (eds.), *Step by Step: essays on Minimalist syntax in honor of Howard Lasnik*, pp. 89-155. MIT Press, Cambridge, Mass.

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

Chomsky, Noam, and Howard Lasnik. 1993. The theory of principles and parameters. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann (eds.), *Syntax: an international handbook of contemporary research*, pp. 506-569. Walter de Gruyter, Berlin.

Culicover, Peter. 1993. Evidence against ECP accounts of the *that*-t effect. *Linguistic Inquiry* 24.557-561.

Deal, Amy Rose. 2019. Raising to Ergative: remarks on applicatives of unaccusatives. *Linguistic Inquiry* 50.388-415.

Douglas, Jamie. 2017. Unifying the *that*-trace and anti-*that*-trace effects. *Glossa* 2(1):60.1-28.

Erlewine, Michael Yoshitaka. 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language and Linguistic Theory* 34:429-479.

Erlewine, Michael Yoshitaka. 2020. Anti-locality and subject extraction. *Glossa* 5(1):84.1-38.

Fox, Danny, and David Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31.1-45.

Grohmann, Kleanthes. 2003. *Prolific domains: on the anti-locality of movement dependencies*. John Benjamins.

Grohmann, Kleanthes. 2011. Anti-locality: too-close relations in grammar. In Cedric Boeckx (ed.), *Oxford handbook of linguistic minimalism*, pp. 260-290. Oxford University Press.

Ishii, Toru. 1999. Cyclic spell-out and the *that-t* effects. In Sonya Bird, Andrew Carnie, Jason Haugen, and Peter Norquest (eds.), *Proceedings of WCCFL 18*, pp. 220-231. Cascadilla Press, Somerville, Mass.

Kandybowicz, Jason. 2009. Embracing edges: syntactic and phono-syntactic edge sensitivity in Nupe. *Natural Language and Linguistic Theory* 27.305-344.

Ko, Edwin. 2020. Feeding agreement: anti-locality in Crow applicatives of unaccusatives. Handout from talk given at NELS 51, UQAM.

Newman, Elise. 2020. Facilitator effects in middles and more. *Glossa* 5(1): 62.1-40.

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

Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: causes and consequences. In Michael Kenstowicz (ed.), *Ken Hale: a life in language*, pp. 355-425. MIT Press, Cambridge, Mass.

Richards, Norvin. 2010. *Uttering trees*. MIT Press, Cambridge, Mass.

Richards, Norvin. 2016. *Contiguity Theory*. MIT Press, Cambridge, Mass.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegman (ed.), *Elements of grammar*, pp. 281-337. Kluwer, Dordrecht.

Saito, Mamoru, and Keiko Murasugi. 1999. Subject predication within IP and DP. In Kyle Johnson and Ian Roberts (eds.), *Beyond principles and parameters*, pp. 167-188. Springer, Dordrecht.

Schneider-Zioga, Patricia. 2007. Anti-agreement, anti-locality, and minimality: the syntax of dislocated subjects. *Natural Language and Linguistic Theory* 25.403-446.

Sulemana, Abdul-Razak. 2018. Q-particles and the nature of covert movement: evidence from Bùlì. ms., MIT.

Ticio, M. Emma. 2005. Locality and anti-locality in Spanish DPs. *Syntax* 8.229-286.

Yamada, Akitaka. 2019. Contiguity Theory and Crystallization: wh-phrases and concord adverbs in Japanese. In Hisao Tokizaki (ed.), *Phonological externalization, vol. 4*, pp. 1-22. Sapporo University, Sapporo.

Zyman, Erik. 2021. Antilocality at the phase edge. *Syntax* 24.510-556.