March 2015

On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling*

Željko Bošković
zeljko.boskovic@uconn.edu

Abstract
The paper provides a uniform account of a number of locality effects, in particular, the Subject Condition, the Adjunct Condition, Richards's (2001) tucking in effect, and the full Comp-trace paradigm, including (in addition to the basic cases) relative and extraposed clauses, the impossibility of short-subject topicalization, French *que-qui* alternation, and the effect of wh-movement on agreement in languages like Kinande, which is discussed in some detail. The account is based on a proposal that there is a difference in the timing of labeling between the basic case where a head and a phrase merge and the case where two phrases merge, as well as a particular labeling-based approach to antilocality.

0. Introduction

The goal of this paper is to provide a uniform account of a number of locality effects, in particular: (a) the Subject Condition (1); (b) the Adjunct Condition (2); (c) the full Comp-trace paradigm, including the basic case in (3) with the improvement with intervening adverbs (4) and relative clauses (6), as well as the null C case with extrapositions like (5), French *que-qui* alternation and the effect of wh-movement on agreement in languages like Kinande, which will be discussed in some detail (with a comparison of subject wh-movement in Kinande and Kaqchikel as well as subject wh-movement and object wh-movement in Kinande); (d) the impossibility of short-subject topicalization (8) and zero-subject relatives (7); and (e) Richards's (2001) tucking in effect.

(1) ?*Who did [friends of t] see Mary?
(2) ?*What did you fall asleep [after Peter had bought t]
(3) *Who do you think that t will leave Mary?
(4) Who do you think that under no circumstances t would leave Mary?
(5) *Who is it likely t likes Mary?
(6) the stone that t broke the window
(7) *John picked up the stone t broke window.
(8) *John, t likes Mary.

It will be shown that all these cases can be accounted for in a unified manner in the labeling framework, given the proposal made here that there is a difference in the timing of labeling between the basic case where a head and a phrase merge and the case where two phrases merge, and a particular labeling-based approach to antilocality. I will first discuss the ingredients of the analysis,

*This work is based upon research supported by the NSF under Grant BCS-0920888. For helpful comments and suggestions, I thank the participants of my 2014 UConn seminar and the International Workshop in Linguistics at Dokkyo University, Tokyo. The paper is dedicated to the memory of Yoshiyuki Shibata.

namely antilocality and the labeling framework (Chomsky 2013, 2014, see also Epstein, Kitahara, and Seely 2014, Carstens, Hornstein and Seely 2013, Rizzi 2013, Saito 2013, Collins 2014, Takita, Goto, and Shibata 2014, among others), and then turn to the account of the cases noted above.

1. Ingredients

1.1. Antilocality

It is standardly assumed that there is an upper bound on movement—movement cannot be too long. A number of authors have argued that movement also cannot be too short. (The ban on movement that is too short is referred to as antilocality in Grohmann 2003). There is a battery of arguments for antilocality in the literature, some of the relevant works being Bošković (1994, 1997, 2005, 2014a), Saito and Murasugi (1999), Ishii (1999), Abels (2003), Grohmann (2003), Ticio (2005), Boeckx (2005), Jeong (2006). These works do not all adopt the same definition of antilocality (see Grohmann 2011 for an overview). Thus, Bošković (2005, 2014a) argues that Move must cross at least one full phrase (not just a segment). One of the effects of this definition of antilocality is that it blocks complement-to-Spec movement, which deduces Abels's (2003) generalization that complements of phasal heads cannot move, one of the arguments for antilocality offered in the literature. Another argument concerns the unacceptability of extraction of NP-adjuncts, noted by a number of authors (e.g. Huang 1982, Chomsky 1986, Culicover and Rochement 1992) and illustrated by (9). Assuming that these adjuncts are NP-adjoined and that DP is a phase, the derivation in (9) is ruled out by antilocality (movement to SpecDP, required by the PIC, crosses only a segment.)

(9) *From where$_i$ did John meet $[_{DP}\, t_i\, [_{NP}\, [_{NP}$ girls] t_i]]

In this paper I will argue for a particular view of antilocality that is adjusted to Chomsky's (2013) system which allows unlabelled projections.

1.2. Labelling

Chomsky (2013) proposes a theory of labeling where in the case where a head and a phrase merge, the head projects (more precisely, provides the label for the resulting object).[1] Chomsky suggests two ways of implementing labeling in the case where non-minimal projections (i.e. phrases) are merged: through prominent feature sharing or traces, where traces are basically ignored for the purpose of labeling. (10) illustrates the former: when *what* is merged with interrogative C (actually CP) both the wh-phrase and the CP have the Q-feature, what is projected (i.e. determines the label of the resulting object) then is the Q-feature.[2]

(10) I wonder $[_{CP}$ what$_i$ $[_{C'}$ C [John bought t_i]]]

The latter case is illustrated by (11).

[1]For much relevant discussion of labeling pre-dating Chomsky (2013), see Chametzky (2000), Collins (2002), Seely (2006), Hornstein (2009), and Hornstein and Nunes (2008).

[2]Like Chomsky (2013), I will continue using CP and SpecCP for such cases for ease of exposition.

(11) What$_i$ do you think [$_{CP}$ t'$_i$ [$_{C'}$ that [John bought t$_i$]]]
(12) v [$_{VP}$ think [$_?$ what [$_{CP}$ that [John bought t$_i$]]]

The timing of labeling here has rather interesting consequences for antilocality, a derivational ban on movement that is too short, which I continue to assume: Move must cross at least one phrase. Chomsky assumes that there is no feature sharing between the intermediate, declarative complementizer *that* and the wh-phrase that passes through its edge in (11) (which essentially follows Bošković 2002, 2007, 2008). Consequently, labeling through feature sharing is not an option here. The embedded clause then cannot be labeled at the point of movement of *what* to its edge, as indicated in (12) by using ?-notation. When v is merged, *what* moves away. The element merged with the CP now being a trace, it is ignored for the purpose of labeling, hence ? is labeled as CP after movement of *what*. Only at this point the status of t'$_i$ in (11) can be determined as the Spec of CP. At the point of movement (12), ? is not a CP, in fact it is not a phrasal projection at all, it is simply undetermined regarding that issue. Since there is no labeling before movement, at the point of movement there is no crossing of a phrase even if *that* projects after wh-movement, with t'$_i$ in SpecCP. To make the issue clearer, we can adopt the following definition of antilocality (cf. Bošković 2014b), adjusted to the framework that allows unlabelled objects, the intuitive idea here being that movement does not cross B if it involves merger with B. (In effect, (13) requires crossing of a labeled projection.)

(13) Antilocality: Movement of A targeting B must cross a projection distinct from B (where unlabelled projections are not distinct from labeled projections).

Antilocality is still satisfied in (12) because the movement that targets vP crosses VP. Were VP to be missing in (12), movement of *what* to vP would violate antilocality.[3] Note that I assume that labeling can take place as soon as it can be accomplished (this will be refined below), otherwise it would not be possible to label structures where both relevant elements move.

Bošković (2014b) shows that this approach to labeling/antilocality deduces the Complex NP Constraint, i.e. the ban on extraction out of nouns modified by clauses. Bošković takes Chomsky's proposal that vP functions as a phase as indicating that the highest projection in the thematic domain functions as a phase. He argues that there is no theta-marking nP in the traditional Noun Phrase (TNP) of the object nominal in (14) (and more generally complex NPs), which makes NP the highest thematic projection here, hence a phase. As a result, movement in (14) must target the edge of CP and the edge of NP. Since there is no feature-sharing between *that* and the wh-phrase that merges with it, the object that is created by their merger is not labeled at the point it is created. In the next step *rumors* merges with this object, the resulting object being labeled as NP via the base-

[3]Antilocality/crossing can also be defined by using the notion *terms of*, e.g. as follows: Movement must cross a labeled category where movement from X to Y crosses Z if X but not Y is a term of Z and Y is not merged with Z.

step of the labeling algorithm. The wh-phrase then merges with the NP. This movement, however, violates antilocality.[4]

(14) ??What$_i$ did you hear t_i [$_{NP}$ rumors [$_?$ t_i that [$_{IP}$ John bought t_i]]]?

At any rate, what is important for our purposes is that under the conception of antilocality argued for here movement must cross a labeled projection.[5]

I will now show that the labeling system and antilocality enable us to deduce a number of locality effects, providing a unified treatment for all of them. An important ingredient of the account will be the proposal that labeling via the base step of the algorithm (when a head and a phrase are merged) can be done immediately, while labeling in the case of a merger of two non-minimal projections takes place when the structure is sent to the interfaces.

In Chomsky (2013), labeling in the cases where a head and a phrase are merged (*the base step*) is done rather differently from the cases where two phrases undergo merger: labeling of the base step occurs via minimal search (MS), the same operation as Agree Closest, a syntactic mechanism falling under minimal computation. MS does not determine the label when two phrases merge. Given the difference, I argue for a timing difference in labeling. I will refer to the proposal below as TOL (timing of labeling). Since the labeling of the base step is done through essentially a syntactic mechanism, it takes place when the relevant configuration is created. Labeling in the case of merger of two phrases occurs when the relevant structure is sent to the interfaces, given Chomsky's assumption that unlabeled objects are uninterpretable (see section 5 for a more extensive discussion, where it is noted that without TOL, it is not even possible to determine the points of spell out (basically, phase determination requires TOL); see also section 5 for another way of deducing TOL, where labeling in the case of a head-complement merger takes place for a strictly syntactic reason, namely, subcategorization, the underlying assumption being that satisfying subcategorization requires that the element with the requirement to take a complement projects (otherwise, there would be no head-complement relation here), see here Chomsky 2000;[6] pending this discussion the reader can simply take TOL at face value).

2. CED effects

2.1. The Subject Condition

[4]See Bošković (2014b) for discussion of the full paradigm regarding extraction out of traditional Noun Phrases as well as its consequences for extraction out of other domains and the theory of phases more generally.

[5]Erlewine (2014) defines antilocality as simply stating that A'-movement from SpecXP must cross a phrase other than XP. This is basically the result of the above discussion, though only for successive-cyclic movement (this will be revised below), and not confined to A'-movement. The labeling framework also makes possible a more natural statement of the condition; in fact, we are still basically capturing here Bošković's (2005, 2014a) intuition that movement must cross a phrase, unlabelled projections not being counted as phrases due to their underdetermined status.

[6] The more specific requirement regarding the kind of a complement can actually be satisfied even if the complement itself is unlabelled, see Collins (2002).

The system provides a rather natural account of the traditional Subject Condition, i.e. the ban on extraction out of subjects located in SpecIP.[7] The ban is illustrated by (15).

(15) *I wonder who$_i$ [friends of t_i] left

Since subjects are phases (being DPs, and DP are phases), whatever moves out of a subject must first move to its edge. Both the merger of *who* with the subject DP and the merger of the subject DP with the IP involve merger of two phrases. Given TOL, the result of the mergers is labeled only when the structure is sent to the interfaces, not at the point of merger (the subject is given in italics).[8]

(16) … [$_{?2}$ *[$_{?1}$ who [$_{DP}$ subject]]* [$_{IP}$ I...[$_{vP}$

The next step involves merger with C, with C, a head, projecting. The wh-phrase then targets CP. The movement violates antilocality. (After the movement, ?1 is labeled as DP and ?2 as IP (through feature sharing), but that is too late to save the derivation.)

(17) …[$_{CP}$ C [$_{?2}$ *[$_{?1}$ who [$_{DP}$ subject]]* [$_{IP}$ I...[$_{vP}$

The ban on extraction out of subjects is thus deduced.[9]

---

[7] For recent perspectives on the CED, see also Boeckx (2008) and Müller (2010).

[8] Klaus Abels (p.c.) notes an alternative derivation where the subject DP from (16) moves alone from the edge of vP to TP, followed by movement of *who* to SpecCP. (Note that I assume split IP, as a result of which subject A-movement does not violate antilocality. There are many arguments in the literature that there is additional structure between vP and TP, e.g. concerning languages which have intermediate verb movement where the verb is lower than the finite verb in Romance, which is located in T, but higher than in English, where it is located in v (Belletti 1990, Stjepanović 1999, Cinque 1999, Bošković 2001), languages like Icelandic which quite clearly have two distinct subject positions above the subject theta-position (see Bobaljik and Jonas 1996) and the distribution of floating quantifiers, which also requires richer clausal structure (Bošković 2004)).The derivation represents a more general issue for the labeling-framework (not simply the analysis proposed here), since in the labeling framework the italicized subject DP is a maximal projection even after *who* merges with it (but see Rizzi 2013). There are several ways of ruling out this derivation (while in the case under consideration such a derivation needs to be excluded, in work in preparation I argue that such derivations should not be excluded in principle (for relevant discussion see also Ott in press), i.e. I argue that there are acceptable instances of such derivations that involve traditional X'-movement). One possibility is that we are dealing here with a semantic issue since the base-subject position ends up being occupied by different elements at different points of the derivation, first by *friends of who* and then by *who* (note that *who* would label the element in the edge of vP in (16) after *friends of who* (*subject* from (16)) moves away). Intuitively, we would then be dealing here with the subject theta-role being assigned to two different elements at different points of the derivation. (Note also that, given split Infl, I will interchangeably use the terms I(P) and T(P), with the understanding that TP is not the first/only projection in the inflectional domain of the clause.)

[9]Note that the problem noted above with respect to extraction from subjects located in SpecIP does not arise with extraction from subjects that remain at the vP edge. Stepanov (2007) claims that such extraction is indeed allowed. There is, however, some controversy regarding whether extraction out of subjects in SpecvP is possible; thus, Uriagereka (2012) claims that it isn't. Such extraction can be either allowed or disallowed in the current system depending on how several additional issues that are involved in such extraction are resolved. It should also be noted that Chomsky (2008) discusses some examples where extraction from

2.2. The Adjunct Condition

The analysis may be extendable to the ban on extraction out of adjuncts, illustrated by (2), if adjuncts are adjoined to complements of phasal heads, i.e. VP and IP. (18) gives the structure for the former case. Given TOL, since adjunct merger involves merger of two maximal projections, its result is not labeled upon merger.[10] (For ease of exposition, all labels in this section will be given at the left edge of the brackets.)

(18) $[[[\ \ldots\ _{VP}\ ]\ K\ _{?}]\ _{vP}]$

Assuming adjuncts are phases (CPs, DPs, or PPs, all of which have been argued to be phases, see Bošković 2014a,b and references therein), movement out of an adjunct, given in italics in (19), has to target the adjunct, resulting in an unlabelled object for reasons discussed above. Further movement has to target vP, which violates antilocality.

(19) $[[[\ \ldots\ _{VP}]\ \textit{[wh-phrase [K(adjunct)]}\ _{?}\textit{]}\ _{?}]\ _{vP}]$

A question, however, arises regarding wh-adjuncts, as in (20).

(20) How did John leave?

A number of authors have argued that wh-adjuncts that are located in SpecCP are actually base-generated in that position (see for example Law 1994, Uriagereka 1988), and Stepanov (2001) argues that wh-adjuncts are merged differently from their non-wh-counterparts due to the presence of the Q-morpheme, which can be implemented as merging them with the Q. In either case, the problem noted above regarding (17) would not arise in (20) (in the latter case, the adjunct would always cross the QP which is created by Q-adjunct merger). Both of these approaches assume that wh-adjuncts and their non-wh-counterparts are not base-generated in the same position. On a more speculatory note, this (i.e. assuming such a difference) in itself opens up another possibility, which

---

subjects is allowed in English. While the grammaticality status of those cases is somewhat controversial (see for example the references in Gallego and Uriagereka 2006), their defining property is that they involve passive/ergative subjects and that the moved element must be a PP, P-stranding being disallowed, as discussed in Broekhuis (2005), Gallego and Uriagereka (2006), and Lohndal (2007), who quite convincingly argue based on these properties (and additional evidence) that the examples in question do not involve extraction from either the surface or the base subject position, giving them a derivation that makes them irrelevant to our current concerns (see also Boeckx 2008).

[10] It should be noted that, following Hornstein and Nunes (2008) and Hunter (2010), Bošković (2014b) argues that adjunction does not require labeling for interpretation, which under Chomsky (2013) means that the result of adjunction is not labeled at all. (In fact, Bošković suggests that not labeling should be taken as the defining property of what is referred to as adjunction, segmentation being dispensable). In other words, under this treatment of adjunction TOL is actually irrelevant here; labeling is not simply delayed with adjuncts, it does not take place at all.

is that wh-adjuncts are generated as adjuncts, but not as adjuncts to complements of phasal heads, which would allow them to extract. [11]

3. Tucking in

Based on a variety of cases, Richards (2001) shows that in multiple-specifier constructions, after one specifier is created the second specifier has to be created below the exisiting specifier, tucking in under it, not on top of the existing specifier.

Consider (21), a multiple wh-fronting construction from Bulgarian, a language which places all fronted wh-phrases in SpecCP (see Rudin 1988) (for ease of exposition I will assume that the embedded CP is the only intermediate phase in this example, which means that movement to the matrix-clause must proceed via the embedded-clause SpecCP).

(21) a. Koj_i kogo_j misliš $[_{CP}$ t_i t_j če $[_{IP}$ t_i e udaril $t_j]$?
who where think-2s that has hit
'Who do you think hit whom?'
b. cf. *Kogo koj misliš [če e udaril]?

The subject wh-phrase is higher than the object wh-phrase prior to undergoing movement. As a result, the subject wh-phrase must move first to the embedded clause SpecCP (this is the standard account which basically treats ordering of fronted wh-phrases in Bulgarian as a Superiority effect). An additional CP-Spec is then created by movement of *kogo* 'what'. Richards (2001) argues that the additional Spec is created below the original Spec, i.e. by tucking in under the original Spec. As a result, *koj* is still higher than *kogo* prior to movement to the matrix clause. *Koj* then must move first to the matrix SpecCP, with *kogo* tucking in under it. The ordering of fronted wh-phrases in Bulgarian thus illustrates Richards's tucking-in effect.

YongSuk Yoo (p.c.) observes that Richards's tucking in can also be derived from the labeling system if only labeled categories can be targets of movement.[12] Consider again (21) in light of this proposal. As discussed above, *koj* moves first, merging with the CP headed by *that*. Since, as discussed above, the relevant elements are not involved in feature-sharing, the result of the merger is not labeled (recall that labeling is possible only after the wh-phrase moves away). We then have (22) prior to movement of *kogo*. Notice now that given that only labeled categories can be targeted

[11] In fact, something along these lines could be used for Truswell's (2011) observation that some adjuncts allow extraction. This can be captured in the current system if these adjuncts are not adjoined to complements of phasal heads. E.g., if an adjunct is adjoined to vP instead of VP, extraction out of it will not be banned. A broader question, however, remains: Why are most adjuncts adjoined to complements of phasal heads? Why do most but not all adjuncts disallow extraction is really a question for everyone, the current account merely states it in a particular way.

[12] One issue needs to be addressed here. We cannot require that external merge involves only labeled categories, otherwise phrases targeted for successive-cyclic movement, like the embedded CP in (11), could not undergo further merger; i.e. the CP in question could not be merged with the V. While I leave working out the details of the issue for another occasion, the intuition here seems to be clear. What is targeted by external merge is the element whose requirement external merge satisfies. In the case in question it is V, since V has to take a complement. No issue regarding the (now more general) requirement that merge targets only labeled categories then arises here.

by movement, *kogo* can move to the edge of the CP in question only be merging below *koj*, i.e. by merging with CP; merging with the *koj*+CP object is not an option, since this object is not labeled. This yields Richards's tucking in effect.

(22) $[_{?}$ koj $[_{CP}$ če $[_{IP}$ …kogo..]?

However, this in itself does not suffice to account for the tucking-in effect in the matrix clause, or more generally for simple questions like (23). Here, since the relevant C is interrogative, the wh-phrase does undergo feature-sharing with the CP that it merges with. While in (22) the CP cannot be labeled until *koj* moves away, this is not the case in (23). Recall now the proposal made above: labeling via the base step of the algorithm (when a head and a phrase are merged) can be done immediately while labeling in the case of a merger of two non-minimal projections takes place when the structure is sent to the interfaces.

(23) a. Koj$_i$ kogo$_j$ t$_i$ e udaril t$_j$?
who where has hit
'Who hit whom?'

b. cf. *Kogo koj e udaril?

This means that the result of merger of C and the IP can be labeled immediately (at the point of merger), while the result of merger of the CP and *koj* cannot be labeled immediately. As a result, (23) has the following structure prior to movement of *kogo*.

(24) $[_{?}$ koj $[_{CP}$ C-Q $[_{IP}$ …kogo…]]]?

The only way *kogo* can move to a labeled category is if it targets CP, yielding Richards's tucking-in effect for this case too. As far as I can tell, other cases of tucking-in that Richards discusses can also be handled in this manner. I conclude therefore that under the particular view of labeling adopted here, Richards's tucking in effect follows from the labeling mechanism.[13]

4. Local subject movements

I will now turn to various types of local subject movements, which exhibit rather interesting behavior crosslinguistically. What I refer to as local subject movement here is movement from SpecTP to the CP that immediately dominates the TP in question. In many cases such movement is clearly blocked. However, there are cases where it seems to be allowed. I am not aware of any attempts at a uniform account of all relevant cases (see for example (3)-(8)). In this section I will show that given TOL, the labeling system can provide a uniform account of the rather complex paradigm pertaining to the domain of local subject movement. However, given the complexity of the paradigm in question, some issues will need to be left unresolved or without proper independent support, which I hope to return to in future work.

[13] See Yoo (in preparation) for a slightly different deduction of tucking in.

A number of works have shown that subject movement to SpecCP cannot proceed via SpecIP (see Bošković 2008, Erlewine 2014, Holmberg and Hróarsdóttir 2003, Rizzi 1990, Rizzi and Shlonsky 2007 among many others for various languages). Thus, in many languages, like Kinande, verbal morphology that arises as a result of agreement between T and the subject located in its Spec cannot be present when the subject undergoes wh-movement, which is standardly taken to indicate that wh-subjects do not move via SpecTP, hence they do not license the usual agreement morphology that occurs with subjects located in SpecTP. The effect is illustrated by the following Kinande paradigm, where the usual subject-agreement morphology from (25a) cannot be present under wh-movement (25b) (it is not possible to simply drop the agreement morphology; rather, a different morphological marker appears on the verb, as in (25c). A labeling-based account of this will be provided in section 4.6.)

(25) a. Kambale **a.**langira Marya
Kambale agr.saw Mary
b. *Iyondi yo **a.**langira Marya
who C agr.saw Mary
c. Iyondi yo **u.**langira Marya
who C anti-agr.saw Mary (Schneider-Zioga 1995)

Another argument to this effect is provided by several varieties of Italian. Thus, Brandi and Cordin (1989) show that Trentino and Florentino have different agreement with postverbal and preverbal subjects. Wh-movement of subjects is necessarily accompanied by postverbal agreement. The relevant pattern is illustrated below with Florentino.

(26) a. Le ragazze l' hanno telefonato.
the girls CL_{3pl} has_{3pl} phoned (Campos 1997)
b. Gl'- ha telefonato delle ragazze
CL_{3sm} has_{3sm} telephoned some girls
`Some girls have telephoned.'
c. Quante ragazze gli ha parlato con te?
how-many girls CL_{3sm} has_{3sm} spoken with you
`How many girls talked to you?'
d. *Quante ragazze le hanno parlato con te?
how-many girls CL_{3pf} has_{3pf} spoken with you (Brandi and Cordin 1989)

Consider also Icelandic (27). (27a) shows that an intervening experiencer blocks agreement with a nominative object, hence the obligatory singular on the matrix verb in (27a). An NP trace does not exhibit this blocking effect, as shown by (27b). However, the blocking effect is still present in (27c,d). If the experiencer in in (27c-d) could undergo the same kind of movement to SpecTP it undergoes in (27a) before undergoing wh-movement, the experiencer blocking effect should be voided in (27c-d) since the intervening element would be an NP-trace, just as in (27b). The Icelandic paradigm in question thus also indicates that wh-movement via SpecTP is not possible.

(27) a. Það virðist/*virðast einhverjum manni [hestarnir vera seinir]
EXPL seems/seem some man.DAT the-horses.NOM be slow
'It seems to some man that the horses are slow.'
b. Mér virðast t_{NP} [hestarnir vera seinir]
me.DAT seem.PL the-horses.NOM be slow
c. Hvaða manni veist þú að virðist/*virðast t_{wh} [hestarnir vera seinir]
which man.DAT know you that seems/seem the-horses be slow
'To which man do you know that the horses seem to be slow?' (Holmberg & Hróarsdóttir 2003)
d. Hverjum mundi/??mundu hafa virst t_{wh} [hestarnir vera seinir]
who.DAT would.3SG/would.3PLhave seemed the-horses.NOM be slow
'To whom would it have seemed that the horses are slow?' (Nomura 2005)

What the above discussion indicates is that the configuration in (28), stated in the traditional obligatory labeling system, should be disallowed.

(28) $[_{CP}\ wh_i\ [_{IP}\ t_i$

This in fact follows straightforwardly from the labeling system argued for here, given TOL and the labeling-based approach to antilocality. Given TOL, merger of the wh-phrase and TP does not result in immediate labeling. The object in question is then merged with C (interrogative or non-interrogative, recall that the effect in question is found with both interrogative and non-interrogative Cs). The resulting object is labeled as CP via the base labeling algorithm. The wh-phrase then merges with the CP. The movement, however, does not cross a labeled projection, violating antilocality, i.e. (13).

(29) $[_{?}\ wh_i\ [_{CP}\ [_{?}\ t_i\ [_{IP}$

4.1. The *that*-trace effect

The account can be extended to a number of other local subject movements, including the traditional *that*-trace effect, illustrated by (30).[14] Movement from the IP-edge to the CP-edge in (30) also involves the configuration in (29), hence it violates antilocality (see (31)).[15]

(30) *Who_i do you think that t_i left Mary ?

[14] Starting with Perlmutter (1971), the *that*-trace effect has generated a great deal of interest, particularly within the Government and Binding framework (see e.g. Chomsky 1981, 1986, Kayne 1984, Lasnik and Saito 1984, 1992, Pesetsky1981, Rizzi 1990). While the phenomenon was generally ignored in early Minimalism, probably because it was considered too hard to explain, minimalist accounts of the *that*-trace effect have begun to emerge more recently (see e.g. Boeckx 2008, Branigan 2005, Buesa García 2011, Hoge 2001, Ishii 2004, Kim 2008, Lohndal 2009, Mizuguchi 2008, Pesetsky and Torrego 2001, Richards 2001, Rizzi and Shlonsky 2007, Roussou 2002, Szczegielniak 1999).

[15] See also Erlewine (2014) and Brillman and Hirsch (2014) for an antilocality account of the *that*-trace effect and the improvement in (32) (the antilocality account of the *that*-trace effect goes back to Bošković 1997 and Ishii 1999).

(31) Who$_i$ do you think [$_?$ t$_i$ [$_{CP}$ that [$_?$ t$_i$ [$_{IP}$ left Mary]]]]

The antilocality account also captures the well-known improvement in (32). Browning (1996) and Watanabe (1993) argue that such cases involve CP-recursion while Culicover (1992) argues that they involve a PolP; what is important for our purposes is that there is a phrase between IP and CP, hence movement of the subject from IP to the highest CP, which functions as a phase (assuming that the highest clausal projection functions as a phase, as in Bošković 2014a (see also Bošković 2014b, Wurmbrand 2014a), does not violate antilocality. (All this extends to the (improvement of) Comp-*t* effects with other complementizers, see Culicover 1992; note also that the rescuing effect here does not depend on inversion, see Browning 1996).[16]

(32) Leslie is the person who I said [$_?$ t$_i$ that [$_{CP/PolP}$ at no time [$_?$ t$_i$ [$_{IP}$ considered running for public office]]]] (Browning 1996)

What about the cases like (34), which do not exhibit a Comp-trace effect?

(34) Who do you think t left Mary?

There are several possibilities for analyzing such cases given the proposals that were independently made in the literature for embedded non-interrogative clauses which are not introduced by *that* in English. Thus, such cases can be treated as in Rizzi (2006), namely as involving truncation of the CP+IP structure (confined to clauses with non-overt subjects (IP-internally) in the V-complement position). The issue that arose in (30) then would not arise in (34) since the relevant structure is missing.

A number of authors have argued that *that*-less embedded clauses in English are quite generally IPs, even in simple examples like *John believes Mary left* (see Bošković 1997 and references therein). The IP analysis (with a similar restriction regarding the distribution of the option in question) would also easily capture (34).[17] [18]

[16]There is a controversy regarding whether non-adverbial topics can rescue *that*-trace effects, see Culicover (1992) and Browning (1996). Notice that not all subject extractions discussed in this paper can be improved with adverb insertion. I leave open why this is the case (presumably in some cases we could be dealing with an intervention effect (topic movement across a topic; multiple topics are in fact not possible); it is also possible that prosodic considerations are relevant in some cases (PF has in fact been implicated even in the *that*-trace effect, see de Chene 1995 and Kandybowicz 2006, 2008).

[17]One potentially relevant phenomenon from the perspective of the truncation and the bare IP analysis is the impossibility of embedded topicalization in the absence of *that*. If topicalization targets the CP field, the lack of a CP field would lead to the impossibility of embedded topicalization. These analyses thus capture the ungrammaticality of (ib). (Bošković 1997 in fact gives the unacceptability of (ib) as one of the arguments for the bare IP analysis, and Rizzi 2006 also gives this as supporting evidence for his truncation analysis).

(i) a. John didn't believe that Mary, Bill kissed.
b. *John didn't believe Mary, Bill kissed.

[18] Extraction out of subjects needs to be reconsidered under this analysis.

(i)?*Who$_i$ do you think [$_{CP}$ that [$_{IP}$ [friends of t$_i$] left]]?

(ii) ??Who$_i$ do you think [$_{IP}$ [friends of t$_i$] left]?

(i) can be handled as discussed in section 2.1. Regarding (ii), as noted in Bošković (1992), extraction from subjects of *that*-less clauses is better than extraction from subjects of clauses headed by *that*. It is, however,

Pesetsky's (1992) account on which *that*-less embedded clauses are CPs headed by a null C that undergoes movement to V opens up another avenue for analyzing (34). There are a number of well-documented cases where head-movement voids locality violations, in particular, by voiding the phasehood of the phrase whose head undergoes movement (see den Dikken 2007, Gallego and Uriagereka 2007, Bošković 2013, 2014b for various approaches to the issue). For example, Galician has a rather interesting phenomenon of D-to-V incorporation which voids islandhood effects (see Uriagereka 1988, Bošković 2013). To illustrate, Galician disallows movement from definite NPs. However, the violation is voided when D incorporates into the verb. Bošković (2014b) argues that the source of the definiteness effect is that movement cannot proceed through the Spec of definite DPs. The problem with (35a) is then that v cannot attract the wh-phrase without violating the PIC. Turning to (35b), Bošković (2014b) argues that movement of the phasal head voids the phasehood of DP, as a result of which v can attract the wh-phrase without violating the PIC.[19]

(35) a. *e de $quén_i$ viche $[_{DP}$ o retrato $t_i]$?
and of who saw(you) the portrait
b. e de $quén_j$ viche-lo_i $[_{DP}$ t_i retrato $t_j]$?
and of whom saw(you)-the portrait
'so, who have you seen the portrait of?' (Uriagereka 1988)

---

still degraded. Putting aside the contrast and assuming (ii) still needs to be ruled out, the question is why it is degraded if the embedded clause is an IP. Its unacceptability actually follows straightforwardly follows from the current system given Hiraiwa's (2005) claim that the edge of the edge of phase XP is not at the edge of XP for the purpose of the PIC, i.e. it is not accessible from outside of the XP, which Bošković (2014b) formalizes by requiring Y to merge with a projection of phasal head X to count as the edge of XP. Since the edge of the edge of phase XP does not merge with a projection of X, it is not at the edge of XP. As discussed above, Bošković (2014a) argues the highest clausal projection counts as the phase (see also Bošković 2014b, Wurmbrand 2014a). If CP is missing in (ii), this makes IP a phase. Movement from IP is then possible only from the edge of IP, given the PIC. However, the element moving out of the subject, which moves to the edge of the subject as discussed above, is at the edge-of-the-edge configuration with respect to IP (since it does not merge with a projection of I), which means that it is not accessible from outside of the IP. To be accessible from outside of the IP, it first needs to move to the edge of the IP, either by adjoining to IP or by moving to a higher IP-Spec. This movement, however, violates antilocality, as stated in (13). (For ease of exposition, the parenthesis in (iii) indicate what the relevant labels would be when the structure is sent to the interfaces. Note also that under this analysis, (i) and (ii) are not ruled out in exactly the same way, which may be desirable given that they do not display exactly the same degree of deviance.)

(iii) …$[_{?2(IP)}$ who $[_{?1(DP)}$ t $[_{DP}$ subject]] $[_{IP}$ I...

[19]It is standardly assumed that there are a number of projections below DP in (35), which I ignore here. (Bošković 2014b actually assumes that NP is a phase, see that work for an approach to the PIC where the PIC is not violated here in spite of the phasehood of the NP.) See den Dikken (2007), Gallego and Uriagereka (2007), Bošković (2013, 2014b), and Wurmbrand (2014b) for a number of additional cases of this type and Bošković (2014b) for discussion of restrictions on where the effect in question occurs. (Bošković argues that the effect occurs only when a phasal head moves to a phasal head (in the cases under consideration, C and the article move to the V+v complex, which contains the phasal head v). Bošković also shows that the Complex NP Constraint effect from (14) is voided in languages that have N-to-D movement like Setswana, another instantiation of the effect in question. (Note that for Bošković (2014b), the-phasehood-voiding movement of a phasal head cannot turn a higher non-phase into a phase; for den Dikken (2007) and Gallego and Uriagereka (2007) it can, but not in this case due to V-to-v movement.)

The phase-voiding effect of head movement provides a straightforward account of (34) under Pesetsky's (1992) null C incorporation analysis. The incorporation voids the phasehood of the CP hence movement need not proceed through the edge of CP in (34). As a result, the problem that arose in (30) does not arise in (34). (Pesetsky's analysis thus enables us to provide a uniform treatment of the contrast in (30)/(34) and the contrast in (35), where head movement voids the phasehood of CP in (34) and phasehood of DP in (35b)).[20]

Another alternative is provided by Bošković's (2011) account of the contrast between (30) and (34) which is based on Chomsky and Lasnik's (1977) proposal that (37) is derived from (36) via deletion of *that*.

(36) Mary thinks that John left.
(37) a. Mary thinks John left.
b. Mary thinks ~~that~~ John left.

There is a long-standing line of research going back to Chomsky (1972) where movement out of an island leads to *-marking of the island, with the locality violation being repaired if the *-marked element is deleted in PF, as in Ross's (1969) cases where movement out of an island is repaired if the island is elided, illustrated by (38) (see Merchant 2001, Lasnik 2001, Hornstein, Lasnik, and Uriagereka 2003, Fox and Lasnik 2003, among many others, for this line of research).

(38) a. *Ben will be happy if Mary fires one of the students, but she didn't know which student$_i$ Ben will be happy [if she fires t_i].
b. Ben will be happy if Mary fires one of the students, but she didn't know which student$_i$ ~~Ben will be happy[if she fires t_i].~~

Working within this line of research, Bošković (2013) argues that with PIC/antilocality violations with phase XP, what is *-marked is the head of the phase. (Bošković 2013 treats (35) in this manner too; see Bošković 2013, Riqueros 2013, and Talić 2014 for evidence for this position). Consider in this light (30). As discussed above, movement of the subject to SpecCP induces an antilocality violation, which under the proposal made in Bošković (2013) leads to the *-marking of the head of the relevant phase, namely *that*.

(39) *Who$_i$ do you think [$_?$ t_i [$_{CP}$ that* [$_?$ t_i [$_{IP}$ left Mary]]]]

In (30), the *-marked element remains in the final PF representation, inducing a violation. However, in (34), the *-marked element is deleted in PF (under Chomsky and Lasnik's 1977 analysis of *that*-less clauses), which removes the locality violation. Under this analysis, which follows Bošković (2011, 2013), the contrast between (30) and (34) receives the same treatment as the contrast in (38).[21]

[20]Under this analysis, *??who do you think friends of t left* can be accounted for in the same way as under the IP analysis (see footnote 18): If the embedded CP cannot be a phase, IP is the highest clausal projection that can be a phase, hence should count as the phase under this analysis.
[21]Notice that in contrast to the *that*-trace effect, with movement out of a subject in SpecIP two phases are implicated in the violation, since the initial position of movement has the DP phase as its sister and the final

There are thus several ways of treating (34), I will leave teasing them apart for future research.

4.2. Extraposition

The current analysis also provides a new perspective on a curious property of extraposed clauses like (40) which has resisted a satisfactory account.

(40) It is likely/appears (that) John bought a house.

While these extraposed clauses allow object extraction they disallow subject extraction, as noted in Kayne (1984), Stowell (1981), Bošković and Lasnik (2003): compare in this respect the object extraction cases in (41)/(43), and subject extraction cases in (42)/(44) (the examples are taken from Bošković and Lasnik 2003:538).[22] Adjuncts pattern with objects, as in (45)-(46) (see Bošković and Lasnik 2003).

(41) What is it likely (that) John will read?
(42) *Who is it likely will read the book?
(43) Who does it appear that Mary likes?
(44) ?*Who does it appear likes Mary?
(45) How is it likely [(that) John fixed the car t]?
(46) How does it appear [(that) John fixed the car t]?

A number of authors have argued extraposed clauses are not complements but VP-Specs/adjuncts (e.g. Reinhart 1980, Stowell 1981, Bošković 2002, 2014b). Recall now that the options appealed to above regarding (34) were restricted to clauses in the V-complement position (truncation/bare IPs being possible only in this position).[23] This means that the clause in (47) must be a CP. As a result,

---

position the CP phase. Riqueros (2013) argues on independent grounds that in such situations it is the lower phase that is *-marked (basically what is *-marked is the first phase that c-commands the launching site of movement). This means that in the case of extraction out of subjects, what is *-marked is the DP phase, hence deletion of *that* does not improve the construction.

[22]While Kayne (1984), Stowell (1981), and Bošković and Lasnik (2003) give unacceptable examples of subject extraction with both raising adjectives and raising verbs, Kayne (1984:18) also observes that there is some speaker variation with respect to the raising verb case. He suggests that this is an instance of interference from parentheticals or that the extraposed clause is actually a complement (see the discussion below) for the speakers who find such examples acceptable.

[23] The situation is a bit more complicated under the C-to-V analysis; if C-to-V were to be forced here even the null C examples without extraction would be incorrectly ruled out. We thus may need to assume that C-to-V is in principle optional or that it simply does not occur in the context currently under consideration (i.e. that the null C here is not an affix (see below) or that the movement takes place only in the presence of v). There is, however, another, rather straightforward option that has the desired effect. Pesetsky (1992) argues that C-to-V takes place because of the affix nature of the null C. Bobaljik (1995), however, suggests that there are two ways of satisfying an affix requirement: through head movement or through PF merger under adjacency. Assume that they are both in principle available in the case under consideration (i.e. null C), which was in fact proposed (though implemented slightly differently) in Kim (2008). The former option is ruled out in the extraposition case since it would involve C-lowering. PF-merger under adjacency is,

the problem that arose in (30) also arises here: the construction is ruled out by antilocality because movement from the IP edge to the CP edge does not cross a labeled projection.[24]

(47) *Who does it appear t likes Mary?

Traditionally, overt C and null C were assumed to differ with respect to their locality-licensing properties: thus, in the Government and Binding framework it was often assumed that null C counts as a proper governor while complementizer *that* does not. Capturing the obvious phonological difference between the two in a principled way in syntactic terms, which is needed under this analysis, has proven rather tricky, given that syntax should not even know about phonological realization of particular lexical items. The issue does not arise under the current analysis. In fact, under the current analysis in examples like (47) we are dealing with a Comp-trace effect with a null C. There is then no need to posit any difference between *that* and the null C with respect to syntactic locality under the current analysis; they both raise a locality problem for subject wh-movement.

4.3. The ban on short subject topicalization

The above account also captures the ban on short-subject topicalization in English, i.e. the impossibility of the local topicalization option for the subject in (48)).

(48) *I think that John$_i$, [$_{IP}$ t$_i$ likes Mary]].

Lasnik and Saito (1992) provide a number of arguments that local/vacuous subject topicalization is not possible. To cite only one argument here, they note that if short subject topicalization were allowed we would expect that, as in (49), *John* and *himself* can be coindexed in (50), which is not the case.

---

however, still an option and can take place in (i). Notice also that the contrast between (i) and (ii), noted in Bošković and Lasnik (2003), can still be captured: neither head movement nor PF merger under adjacency is an option in (ii), while the latter is possible in (i). (Bošković and Lasnik 2003 in fact restate Pesetsky's 1992 analysis in terms of PF merger. As far as I can tell, combining the two in a way discussed here does not have any obvious undesirable consequences (see also Kim 2008); the facts regarding the distribution of null-C clauses discussed in Pesetsky 1992 and Bošković and Lasnik 2003 can still be captured.)

(i) a. It appeared C John had left.

b. *It appeared at that time C John had left.

[24] The account cannot be maintained as is under the *that*-deletion analysis of the *that*-trace effect. Here's a potential alternative that is consistent with that analysis: Following the line of research in Moro (1997), Hornstein and Witkos (2003), and Sabel (2000), Bošković (2014b) suggests that in this kind of construction the expletive and its clausal associate are generated as a constituent VP/AP internally. In particular, they are generated as the Spec and the complement of a linker-like projection FP (this could be den Dikken's 2006 Relator Projection). Suppose now that what is special about the *that*-less clauses in question is that they are actually IPs, i.e. they do not involve *that*-deletion. Under Bošković's (2014b) approach to phases, where the highest projection in the functional domain of a clause functions as a phase, the FP (rather than IP) should then function as a phase. The subject will then have to move from the IP-edge subject position to the edge of FP. Since the IP is not labeled at the point of movement, the movement in question does not cross a labeled projection, violating antilocality. The problem does not arise with object and adjunct movement.

(49) John$_i$ thinks that himself$_i$ Peter likes.
(50) *John$_i$ thinks that himself$_i$ likes Peter.

The account of the impossibility of subject extraction in the context in (29) can be straightforwardly extended to the ban on short subject topicalization. Given TOL, when *John* merges with IP the resulting object cannot be labeled immediately. The head that hosts topicalization then enters the structure. When the subject moves to merge with the head in question the movement violates antilocality for the same reason it does in (29).

(51) *I think that John$_i$ [$_?$ t$_i$ [$_{IP}$ likes Mary].

4.4.The ban on short subject zero relatives

The above account provides a new perspective on the distribution of *that* in relative clauses. A well-known puzzle with relative clauses is that they do not display the *that*-trace effect.[25]

(52) the stone Op$_i$ that t$_i$ broke the window

From the current perspective, the key to the lack of the *that*-trace effect in relative clauses lies in the optionality of *that* in (53)-(54).

(53) the stone that Mary threw
(54) the stone Mary threw

Kayne (1984) (see also Bošković 1997) observes the contrast in (55) regarding the possibility of a resumptive pronoun, which would be surprising if the relative operator is located in the same position in both constructions.

(55) a. *The book Op I was wondering whether I would get it in the mail
b. The book Op that I was wondering whether I would get it in the mail

I will therefore assume that relative Op is not in the same position in relatives with and without *that*. In (54), the relative operator is located in the Spec of a relative-clause dedicated projection, which I will refer to as RelP, which is obligatory in all relative clauses. The CP headed by *that*, on the other hand, is present optionally; it is present when *that* is present (with Op in its Spec under the assumption that it moves to the edge of the relative clause). (53)-(54) then have the structures in (56)-(57) respectively.[26]

[25] For ease of exposition I adopt here an approach where the nominal head originates external to the relative clause.
[26] The precise projection labels and the details of the structure do not really matter here, see Rizzi (1997) for a different perspective. (Note that in many languages a *that*-CP can occur even above indirect questions). See also Deal (2014).

(56) the stone [$_{CP}$ Op$_i$ that [$_{RelP}$ [Mary threw t$_i$]]]
(57) the stone [$_{RelP}$ Op$_i$ [Mary threw t$_i$]]

This immediately explains the lack of the *that*-trace effect in (52). As before, subject merger with IP does not result in immediate labeling. However, in contrast to (30), movement of the subject to the Spec of *that* does not violate antilocality in (58). In fact, the RelP rescues the derivation from the *that*-trace/antilocality effect here in the same way that CP/the intervening adverb does in (32).

(58) the stone [$_{CP}$ Op$_i$ that [$_{RelP}$ [$_?$ t$_i$ [$_{IP}$ broke the window]]]]

Evidence for this analysis is provided by the impossibility of short zero-subject relatives, noted by Bresnan (1972). While *that* is optional with object relatives under consideration, it is obligatory with subject relatives.

(59) John picked up the stone that broke the window.
(60) *John picked up the stone broke the window.

This follows straightforwardly under the current analysis. In contrast to (58), where the operator moves to CP (which is a phase here), in (61), the operator can only move to SpecRelP (which is a phase here as the highest clausal projection). This movement, however, violates antilocality.

(61) *the stone [$_{RelP}$ Op$_i$ [$_?$ t$_i$ [$_{IP}$ broke the window]]]

Zero subject relatives are expected to be possible in *pro*-drop languages (where *that* is not obligatory in the first place), since in such languages subject movement to IP is not obligatory. Pesetsky (1982) and Bošković (1997) show that this is indeed the case.

(62) Chi'e faccenda Op$_i$ tocca a noi t$_i$
'This is a matter (that) does not concern us' (15th century Italian, Rizzi 1990)

4.5. Agreement under wh-movement
4.5.1. Kaqchikel

The account of the English paradigm discussed above gets an interesting confirmation from a Kaqchikel paradigm discussed in Erlewine (2014). Erlewine observes that there is different morphology in Kaqchikel depending on whether or not subject moves to SpecIP. As in a number of other languages, a subject moving to SpecCP is not allowed to pass through SpecIP (i.e. the usual morphology that accompanies subjects that move to SpecIP cannot be present in that case), which can be captured as discussed above. Significantly, Erlewine shows that when there is a phrase between IP and CP, a wh-subject can pass through SpecIP on its way to SpecCP (i.e. the morphology that accompanies subject movement to SpecIP is then present). This is exactly what is

expected. Due to the presence of this phrase, as in (32) and in contrast to (29), subject movement from IP to CP does not violate antilocality in (63).[27]

(63) $[_{?/CP}\ wh_i\ [_{CP}\ [_{FP}\ [_{?/IP}\ t_i\ [_{IP}$

Consider the relevant Kaqchikel paradigm (all the data below are taken from Erlewine 2014). (64)a illustrates the usual morphology that accompanies subject movement to SpecIP (following Aissen 1999, Erlewine argues that SpecIP is linearized to the right in Kaqchikel).[28] This morphology is not possible under wh-movement of the subject, as in (64)b. Instead, the so-called agent-focus affix (AF) appears. AF is not possible with object wh-movement (64)c or long-distance subject-movement; it occurs only with the local step of subject movement to SpecCP (even in long-distance questions, see (64)d). Interestingly, just like addition of an adverb above IP rescues *that*-trace violations in English, addition of an adverb in Kaqchikel makes available the appearance of usual verbal morphology under wh-movement, the AF morphology not being available in this case (64)e. These facts can be accounted for just like the local-subject movement data from English discussed above. A subject moved to SpecIP undergoes feature-sharing with I, which results in the labeling of the object created by subject-IP merger. However, this happens only when the structure is sent to the interfaces. As a result, if the subject that merges with IP moves to merge with CP right above it, antilocality is violated (no labeled projection is crossed). The problem does not arise in (64)e since the presence of the adverb introduces additional structure as a result of which subject that merges with IP can move to merge with CP and still cross a labeled projection.[29]

(64) a. Iwir x-Ø-u-tëj ri wäy ri a Juan.
yesterday COM-B_{3sg}-A_{3sg}-eat the tortilla Juan
`Yesterday, Juan ate the tortilla.'

b. Achike *x-Ø-u tëj /✓x-Ø-tj-**ö** re wäy
who COM-B_{3sg}-A_{3sg}-eat /COM-B_{3sg}-eat-**AF** the tortilla
`Who ate the tortilla?'

c. Achike ✓x-Ø-u tëj /*x-Ø-tj-**ö** ri a Juan.
what COM-B_{3sg}-A_{3sg}-eat / COM-B_{3sg}-eat-**AF** Juan
`What did Juan eat?'

d. Achike n- Ø-a-b'ij rat [chin *x-oj-r-tz'ët/ ✓x-oj-tz'et-ö roj]?
who INC-B_{3sg}- A_{2sg}-think 2sg that COM-B_{1pl}-A_{3sg}-see COM-B_{1pl}-see-AF 1pl

[27] Erlewine also argues for an antilocality account. The analysis of Kaqchikel adopted here follows the gist of Erlewine's account, with some modifications and an adaptation to the labeling framework.

[28] Kaqchikel verbal morphology is rather complex (see Erlewine 2014 and references cited therein, especially Preminger 2011). I focus here on constructions involving what Erlewine refers to as Set A morphology (still glossing over some morphological complexities; note also that I assume that in this context the presence of Set B morphology is conditioned on the presence of Set A morphology), since this morphology requires the presence of the subject in SpecIP.

[29] I have argued in section 4.4. that relatives are not structurally uniform in English: English has relatives with only one projection above IP and relatives with two projections above IP, which is the case with *that*-relatives. At least some Mayan languages use AF with local subject relatives (see Coon, Pedro, and Preminger in press), which means that these Mayan languages have only the first option from English.

'Who do you think saw us?'

e. Achike kanqtzij x-Ø-u tëj /*x-Ø-tj-**ö** re wäy
who actually COM-B_{3sg}-A_{3sg}-eat/ COM-B_{3sg}-eat-**AF** the tortilla
`Who actually ate the tortilla?'

The data in (65) provide further confirmation. Wh-phrases and indefinites have the same form in Kaqchikel. Erlewine shows that they are both licensed by movement to projections above IP, the indefinite-licensing projection, referred to below as QP, being lower than CP. Consider then (65)b. If object is the indefinite, it will move to QP, which is located right above IP. As a result, subject can move to IP and then to CP without violating antilocality. Usual verbal morphology is then possible on this reading in (65)b. On other hand, on the unavailable who-did-someone-see reading, the object moves to CP and the subject moves to QP. Since QP is right above IP, subject movement to IP and then to QP violates antilocality, hence the ungrammaticality of (65)b on the reading in question. As before, the AF can save the derivation in question, hence this reading is available in (65)a.

(65) a. Achike k'o x- Ø-tz'et- **ö**?
Who ∃ COM-B_{3sg}-see-**AF**
✓`who did someone see?'
*`Who saw someone?'

b. Achike k'o x- Ø-utz' ët?
Who ∃ COM-B_{3sg}-A_{3sg}-see
*`Who did someone see?'
✓`Who saw someone?'

Now, there was an implicit assumption above that, except in the case of *that*-relatives and constructions with pre-IP adverbs, English has only a single CP projection above IP; in other words, CP is not always split. Kaqchikel provides a confirmation of this; CP cannot always be split (more precisely, uniformly split) in Kaqchikel either, otherwise subject movement through IP would not be only selectively available. The conclusion from the above discussion is that there is no uniform split CP field that is present either crosslinguistically or in all constructions of a single language.[30] This is in fact not a surprising conclusion, given that even superficially, the left-periphery shows quite a bit of variation crosslinguistically.

[30] Erlewine (2014) in fact also explicitly argues (primarily based on Mayan languages) that a clause includes only those functional projections that are independently motivated—finely articulated sequences of functional projections proposed in works like Rizzi (1997) and Cinque (1999) are not always present (see in fact Rizzi 1997:314-315 on this position). Note also that Abels's (2003) account of the immobility of IPs that are dominated by CP, illustrated by (i), in terms of his generalization that phasal complements cannot move also requires that the CP field is not split in this case. (Since CP is a phase, the PIC requires IP movement through SpecCP, but antilocality blocks it because it is too short; (iia) is ruled out by the PIC and (iib) by anti-locality).
(i)*[His mother likes Mary]$_i$ everyone believes that t_i
(ii) a. *[$_{CP}$ IP$_i$ [$_{C'}$ C t_i
b. *IP$_i$ [$_{CP}$ [$_{C'}$ C t_i

There is, however, one point that still needs to be captured regarding Kaqchikel. Usual subject morphology requires movement to SpecIP. However, when the subject fails to move to SpecIP the usual subject morphology is not simply dropped: it is replaced by AF. Furthermore, the AF morphology appears only as last resort, to save the derivation where subject movement to SpecIP fails to occur for independent reasons. Why is that the case? A proposal made in Chomsky (2014) can be productively applied here. Chomsky (2014) restates the traditional EPP effect as a labeling effect. He suggests that languages differ regarding whether or not T is strong enough to label on its own. English T is not strong enough to label on its own, hence the subject needs to move to TP to strengthen it, making labeling possible. On the other hand, in Italian TP can label on its own, hence subject movement to TP is not necessary.[31]

Kaqchikel should then be like English, requiring movement to TP to strengthen T so that it can label. I suggest therefore that AF is inserted as last resort to enable T to label when subject movement to SpecTP fails to take place.[32] This is the reason why AF occurs only in these contexts. However, given the obvious difference between English and Italian in the richness of subject agreement morphology, Chomsky ties the difference in the strength of T between English and Italian to the richness of agreement morphology. The analysis just suggested treats Kaqchikel and English in the same way although, in contrast to English, Kaqchikel has rich verbal morphology. There is, however another way of looking at the English/Italian/ Kaqchikel paradigms in question where the issue in question does not arise. Let us assume that, as in Italian, the usual morphology in Kaqchikel is enough to strengthen T, enabling it to label. However, this morphology can only be licensed if the subject is located in SpecTP.[33] As a result, if the subject does not move to SpecTP, the morphology cannot be present, hence T is then like English: it is weak and cannot label on its own. This is what rules out constructions where the agreement morphology is simply dropped in the absence of movement to SpecTP. As already proposed above, the AF insertion can still be considered a last resort strategy to strengthen T so that it can label. This explains the complementary distribution between the usual subject morphology and AF. When the usual subject morphology is present, which is possible any time subject moves to SpecTP, T is strong enough to label, hence there is no need for AF-insertion. When such morphology is not present, T cannot label hence AF is inserted to strengthen it. AF-insertion basically turns English-style, morphologically poor T into Italian-style, morphologically rich T. This analysis thus captures both the fact that verbal morphology cannot be simply dropped when the subject fails to move to SpecTP in Kaqchikel, as well as the last-resort nature of AF. It should, however, be noted that Kaqchikel is not

[31]Chomsky's (2014) proposal regarding English requires a slight modification of the position taken in section 3 regarding targets of movement: movement can target an unlabelled projection iff the movement takes place in order to label that projection, i.e. to make its labeling possible. (The driving force of movement in this case is to provide a label for the projection in question, which was not the case in the examples discussed in section 3. As a result, the difference between the case currently under consideration and the examples from section 3 is that the sister of the moved element in the former is a labeled projection (the labeling is done right after the merger), while this is not the case in the latter, which can also be appealed to to state the relevant difference).

[32] In more traditional terms, AF takes care of the EPP property in these contexts. (Erlewine in fact argues that T here has the EPP property in Kaqchikel).

[33] This is simply a different perspective on Erlewine's (2014) claim that the morphology in question is associated with T having the EPP property.

quite like either English or Italian in Chomsky's analysis of these two languages: Kaqchikel T is strong enough to label on its own, as expected given its morphological richness. In this respect, it is like Italian, not like English T. However, the relevant morphology can only be licensed in Kaqchikel when subject moves to SpecTP.[34] The motivation for the presence of the subject in SpecTP is thus different in Kaqchikel and English. In English, the subject moves there to make labeling possible, while in Kaqchikel the subject moves there for morphology-licensing. In this respect, AF is more like subject movement in English. The AF insertion takes place strictly for labeling reasons, just like subject movement in English. However, while this is always required in English (except in one context, where English uses a very similar mechanism to AF, see section 4.7.), it is required in Kaqchikel only in one, well-defined context.

4.5.2. Kinande

[34] Kaqchikel is actually quite similar to Trentino and Fiorentino, as discussed in Brandi and Cordin (1989) and Rizzi (1990). In Trentino and Florentino, subjects agree only if they are located in SpecIP (i-ii); postverbal subjects do not agree. (As illustrated in (26) for Florentino, wh-moved subjects cannot agree, which confirms that wh-movement from SpecIP is imposible ((26) involves the configuration in (29)). This can be interpreted as indicating that "usual" agreement can only be licensed if the subject is located in SpecIP in Trentino and Florentino, just like in Kaqchikel.

(i) Le	ragazze	l'	hanno	telefonato		Florentino
the	girls	CL_{3pl}	has_{3pl}	phoned		
`The girls have phoned.'						(Campos 1997)
(ii) a. Gl'-		ha	telefonato	delle	regazze	Florentino
b. Ø		Ha	telefoná	qualche	putela	Trentino
CL_{3sm}		has_{3sm}	telephoned	some	girls	
`Some girls have telephoned.'						(Brandi and Cordin 1989)

However, when the subject does not move to SpecIP, as in (ii), Trentino and Florentino can still use default verbal morphology, 3.p.sg (this is not an option in Kaqchikel due to the rather complex rules for verbal agreement morphology that Kaqchikel has, see Erlewine 2014). This is the same morphology as in regular preverbal subject constructions where the subject is 3.p.sg.m., hence Infl with this morphology is strong enough to label.

Like Kaqchikel, Trentino and Florentino thus require the subject to be located in SpecIP to license regular subject-verb agreement morphology. When this does not happen, Trentino and Florentino use default agreement morphology, which plays the same role as AF in Kaqchikel. However, it does not have the same last resort flavor as Kaqchikel AF; in contrast to Kaqchikel AF, it is not present only when the subject for independent reasons cannot move to SpecIP—i.e. it is not a last-resort mechanism. (One could try to argue that the impossibility of AF in (64)a is not due to the last-resort nature of AF but Case. The difference between Trentino and Florentino and Kaqchikel would then be that the subject must move to a position that c-commands I in Kaqchikel (in constructions under consideration), either to SpecIP or to a higher position like SpecCP, for Case-licensing, which is not the case with Trentino and Florentino. This can be captured if, as argued in Bošković (2007) and discussed in footnote 42, Case-licensing normally requires the DP to c-command its Case source; one of the special mechanisms to get around this requirement discussed in Bošković (2007) would then be available in Trentino and Florentino, but not in Kaqchikel (in the relevant context). Morphological ergativity of Kaqchikel interferes with investigating this option. However, this option seems implausible for Kinande, which will be analyzed in the same way as Kaqchikel below. However, in Kinande it is clear that the subject need not move for Case reasons; it can stay in its base-generated position with exactly the same agreement morphology that occurs in normal subject-in-SpecIP transitive constructions, as long as another element (the object) moves to SpecIP to license the morphology in question (see footnote 35).)

The above account of Kaqchikel can also be extended to Kinande: the usual morphology in Kinande is enough to strengthen T, enabling it to label (cf. (25a)). However, this morphology can only be licensed if the subject is located in SpecTP.[35] As a result, if the subject does not move to SpecTP, the morphology cannot be present, hence T is then like English: it is weak and cannot label on its own. This is what rules out constructions where the agreement morphology is simply dropped in the absence of movement to SpecTP (compare **Iyondi yo langara Marya* with (25c)). What happens in such contexts is that what is traditionally referred to as anti-agreement (AA) morphology is inserted (cf. (25c)). I suggest that, like AF in Kaqchikel, AA is inserted as last resort to enable T to label when subject movement to SpecTP fails to take place, which means that the usual subject agreement morphology is not licensed in such cases. This is then the reason why AA occurs only in these contexts. As in the case of Kaqchikel AF, this account explains the complementary distribution between the usual subject morphology and AA in Kinande. When usual subject morphology is present, which is possible any time subject moves to SpecTP, T is strong enough to label, hence there is no need for AA-insertion.

There is, however, a difference between Kaqchikel and Kinande. In contrast to Kaqchikel (cf. (64)c), long-distance contexts do not require AA-morphology in Kinande (I omit the agreement between the wh-phrase and the complementizer where this information is not relevant).

(66) [$_{CP}$ iyondi [$_{C'}$ yo [Kambale a-alengekanaya [$_{CP}$ ng' a-kahuka ebiken]]]]
who that Kambale agr.thought comp agr.cook yams
`Who did Kambale think is cooking yams?' (Schneider-Zioga 2007)

Following Schneider-Zioga (2007), I suggest that the reason for this is that what occurs in the embedded SpecTP in (25a), which also licenses the usual morphology, is *pro*.[36] In other words, *pro* here functions as a resumptive, with the wh-phrase base-generated in the matrix SpecCP.[37]

[35]The relevant element actually does not have to be the subject, it can be a locative in the locative inversion construction or the object in the subject object inversion construction (which is not a passive); in both of these cases the element in SpecTP (not the subject, which remains in its base position) undergoes traditional subject agreement, which confirms that such agreement requires an element in SpecTP (the element can also be a *pro*, see below). It should be noted that there are some interpretational requirements on subjects in Kinande which are similar to Chinese. It is well-known that there is a specificity/topicality requirement on subjects located in SpecTP in Chinese. Kinande has a similar requirement on clause-initial subjects located in SpecTP (I reanalyze from this perspective the relevant data noted in Schneider-Zioga 2007 in work in preparation).

[36] It is not out of question that we are dealing here with a pronoun that undergoes deletion in PF, see e.g. Holmberg (2005). ((i), where SM is the usual subject agreement morphology, shows that null subject constructions are possible in Kinande.)

(i) a-ká-ly-a o-mu-tsérê
SM-TM-eat-FV AUG-C3-rice
'(someone) eats rice'

[37] This kind of a base-generated wh-dependency is apparently not possible in Kaqchikel, hence the grammaticality of (64)d with regular morphology and ungrammaticality with AF. It should be noted here that Kinande is extremely permissive with respect to the possibility of base-generation of wh-phrases in SpecCP; in fact, Kinande quite generally disallows any movement out of CPs, long-distance dependencies being created through *pro* or a series of Operators that are base-generated in SpecCPs (see Boeckx 2003, Bošković 2008, and especially Schneider-Zioga 2009, who shows that the ban on movement out of clauses extends even to A-raising—even A-raising is clause-bounded in Kinande). We are thus dealing with a much broader

(67) [$_{CP}$ iyondi [$_{C'}$ yo [Kambale a-alengekanaya [$_{CP}$ ng' *pro* a-kahuka ebiken]]]]

That this is an option in Kinande is confirmed by the fact that such constructions behave in all relevant respects just like constructions in which an object wh-phrase is present in SpecCP and associated with a resumptive clitic: like the resumptive clitic constructions, the resumptive *pro* constructions are island-insensitive (see (68) for the object+resumptive case and (69) for the subject+*pro* case) and do not induce WCO effects (see (70).

(68) a. **ekihi** kyo Yosefu a-kabula [iyondi nga y' u-ka-**ki**-gula]
what that Jospeh agr-wonder who if that anti.agr-tense-cl-buy
`What does Joseph wonder who is buying it?
b. ebaruha **yahi** yo w-asiga [isi-wu-li w-asoma-**yo**]
letter which that you-left neg-you-be you-read-cl
`Which article did you leave before you read it? (Schneider-Zioga 2007)

(69) a. [**yondi**$_k$ yo Yosefu a-kabula [ekihi$_j$ nga ky' ***pro***$_k$ a-kalangira t$_j$]
who that Joseph agr-wonder what if that agr-sees
`Who does Joseph wonder what (he) sees?
b. omukali **ndi**$_k$ yo w-asiga [***pro***$_k$ isy-a-lyagua]
woman who that you-leave neg-agr-spoke
`Which woman did you leave before (she) spoke?' (Schneider-Zioga 2007)

(70) a. **iyondi**$_i$ yo mama wiwe$_i$ a-kalengekanaya ati omugalimu a-a-**mu**-nzire kutsibu
who that mother his agr-thinks that teacher agr-past-cl-loves best
`Who does his mother think that the teacher loves him the best?'
b. **iyondi**$_j$ yo mama wiwe$_j$ a-kalengekanaya ati ***pro***$_j$ a-anzire ebitabu
who that mother his agr-thinks that agr.loves books
`Who$_j$ does his$_j$ mother think loves books?' (Schneider-Zioga 2007)

I suggest that the reason why the *pro*-as-a-resumptive strategy is not an option in (25b) (if it were, regular morphology rather than AA would occur here) is the well-known ban on local subject resumptives, i.e. the ban on a resumptive pronoun in SpecTP that is associated with a wh-phrase in SpecCP of the same clause (see Aoun and Li 1990, Boeckx 2003:83-91, Borer 1984, McCloskey 1991, Bošković 2009).[38]

There is also an alternative explanation. More generally, as noted in footnote 37, all movement in Kinande is clause-bounded; Kinande uses a variety of resumptive strategies (*pro* in an A-position, or a null Operator in an agreeing SpecCP which is co-indexed with a higher clause Operator) to establish long-distance dependencies. We are thus simply dealing here with an independent property of Kinande and the often-noted last-resort nature of resumptivization, which is used when movement is not an option (movement is an option in Kinande in a single-clause environment, but not in long-distance contexts).

The null operator strategy is illustrated by (71).

---

difference between the two languages which is abundantly independently motivated (see the references cited above).

[38] Note that we are not dealing here with an Avoid Pronoun Principle kind of effect, see Boeckx (2003).

(71) Ekihi kyo wasiga [$_{island}$ embere Marya aminye [nga kyo wasoma ___]]
what wh-agr you:leave before M knew comp wh-agr you:read
`What did you leave before Mary knew you had read?' (Schneider-Zioga 2009)

Here a null Operator is moved to the embedded SpecCP (notice that we are dealing here with an agreeing C), and coindexed with the wh-phrase in the matrix SpecCP, voiding the adjunct island effect. Schneider-Zioga (2009) shows that even in non-island examples like (72) the fronted element cannot be reconstructed into the embedded clause to license the bound variable reading, which confirms that there is no long-distance movement taking place here, in accordance with the clause-bounded nature of movement in Kinande.[39]

(72) e**kitabu kiwe**$_{k/*j}$ kyo ngalengekanaya [$_{CP}$ nga.kyo [**obuli mukolo**]$_j$ akasoma _ kangikangi
book his wh-agr I.think C.wh-agr every student read regularly
'(It is) His$_{k/*j}$ book that I think [every student]$_j$ reads regularly.' (Schneider-Zioga 2009)

In this respect, consider also the AA-examples in (73), involving an adjunct island. The examples in question do not involve standard agreement, hence no *pro* in SpecTP of the most embedded clause (in contrast to (69)b), which has AA-morphology.

(73) a. *omukali ndi yo wasiga [$_{island}$ embere___ wabuga]
woman who wh-agr you:left before spoke
`Which woman did you leave before (she) spoke?'
b. omukali ndi yo wasiga [$_{island}$ embere Kambale anasi [$_{CP}$ ko yo ___ wabuga]]
woman who wh-agr you:left before K knew that wh-agr spoke
`Which woman did you leave before Kambale knew that (she) spoke?' (Schneider-Zioga 2009)

(73)a) is ruled out due to a locality effect. However, (73)b) is acceptable. The only difference between (73)a) and (73)b) is that the theta-position of the wh-phrase is separated from the adjunct boundary by an agreeing CP in (73)b. (73)b) can then be derived as follows: a null Operator is base-generated in the most embedded CP, and undergoes movement to the local SpecCP (without moving to SpecTP), with the AA inserted to strengthen T. (The null element in question can be the same null element as the one located in SpecTP in e.g. (66)/(67), i.e. this may simply be *pro* moving to SpecCP. I will refer to such *pro* as null operator for ease of exposition). The operator is coindexed with the wh-phrase that is base-generated in the matrix SpecCP. Superficially, the AA-morphology is not just last resort in Kinande, since nothing here prevents the option of having *pro* in SpecTP and regular agreement in the most embedded clause, i.e. this option is in principle available (cf. for example (69)b). But not for the numeration in question: (73)b) contains an

[39] As Schneider-Zioga (2009) notes, the lack of superiority effects even in cross-clausal superiority contexts confirms the lack of long-distance wh-movement.
(i) ekihi kyo ndi anasi nga kyo Josefu abula
what wh-agr who know if wh-agr Joseph bought
Lit. `What does who know if Joseph bought?' Schneider-Zioga (2009)

agreeing C that requires an operator (a wh-phrase or a null element co-indexed with a wh-phrase) in its SpecCP. Since the only element that can undergo movement to the embedded SpecCP is the subject, local subject movement to SpecCP is then forced here, which disallows regular subject agreement morphology. The AA-morphology then only appears not to be last resort in Kinande: while there are cases where regular subject-agreement morphology and the AA-morphology alternate, such cases have different Cs: the AA-morphology is possible only in the context of an agreeing C, which requires movement to SpecCP; furthermore, it is the only option in such cases.[40]

Kinande is interesting in another respect. Chomsky (2014) suggests more or less uniform treatment of T and V when it comes to labeling. Interestingly, Kinande objects (in ditransitive and ECM environments, see Bošković 2008 and Schneider-Zioga 1995) also obligatorily trigger agreement, which is not possible under wh-movement. However, with wh-movement this agreement is simply dropped, which suggests that T and V should not be treated in the same way when it comes to labeling. Bošković (2008) suggests an analysis of Kinande where the object in the environments in question in Kinande must undergo object shift in order to be Case-marked, agreement being a reflex of this movement. Under Bošković's (2008) analysis, these objects can actually be case-marked as long as they c-command the verb (see footnote 42); movement to SpecCP also suffices for this purpose. If object shift targets SpecVP, as suggested in Chomsky (2014), an object cannot undergo object shift and then proceed with wh-movement since the next step of movement would need to target vP, violating antilocality.[41] Wh-movement via the object shift position is then not an option, hence the impossibility of agreement in wh-contexts (the agreement morphology being licensed by the element in the object-shift position, on a par with TP). Under this analysis, T and V are not treated uniformly with respect to labeling. While T needs to be phi-strengthened for labeling, this is not the case with V (when object agreement occurs, it's a result of a case-licensing movement, see footnote 42); this is why some subject agreement morphology is always obligatory with T, even in wh-contexts where the wh-subject is prevented from moving to SpecTP for independent reasons, which is not the case with object agreement morphology. At any rate, Kinande indicates that T and V should not be treated in the same way when it comes to labeling.[42]

[40] Is there any evidence that the AA-morphology is indeed last resort? If it were not, then in any construction, even non-wh constructions, it should be possible to introduce AA-morphology and leave the subject in SpecvP, with nothing filling the SpecTP position. This, however, is not possible. The AA-morphology is restricted to the contexts requiring local subject wh-movement.

[41] Note that if object shift targets SpecVP extraction from shifted objects is expected to be impossible (it involves the same kind of configuration in the relevant respect as extraction from subjects located in SpecIP). Stepanov (2007) shows that extraction from shifted objects is indeed impossible (see also Bošković 1997, 20002, Lasnik 1999).

[42] Interestingly, objects in simple transitives do not trigger object agreement morphology, only objects in ECM configurations do (note that Bošković 2008 treats double object constructions as in the line of work originating with Kayne (1984) (see especially den Dikken 1995), where ditransitive constructions involve an ECM-style configuration: the goal and the theme are generated in a small clause that excludes the higher verb). What is important here is that objects in simple transitive constructions c-command the verb even without movement. In other words, the descriptive generalization is that accusative-marked elements must c-command the case-marking verb: they trigger object agreement (hence move to SpecVP) only if other independently required operations do not put them in a position where they c-command the verb (wh-movement or lexical insertion itself). This can be straightforwardly accommodated under the approach to case-licensing argued for in Bošković (2007, 2008), where the DP needs to c-command the verb. This also

4.5. French *que-qui* alternation

I now turn to the French *que-qui* alternation, where *qui* occurs under subject wh-movement. A well-known aspect of this alternation is that *qui* occurs only with local subject movement.

(74) a. *Quelle étudiante crois-tu [t' que [t va partir]]?
`Which student do you believe that is going to leave?'
b. Quelle étudiante crois-tu [t' qui [t va partir]]?
`Which student do you believe QUI is going to leave?'
c. Quelle étudiante crois-tu [t' que [Marie va aider t]]?
`Which student do you believe that Marie is going to help?'
d. * Quelle étudiante crois-tu [t' qui [Marie va aider t]]?
`Which student do you believe QUI Marie is going to help?' (Rizzi and Shlonsky 2007)

The distribution of *qui* is in this respect very similar to AF in Kaqchikel, which also only occurs with local subject questions. In light of this similarity I suggest treating the two in the same way.

Taraldsen (2001) makes a very interesting proposal to treat *qui* as *que+il*, which is further expanded on in Rizzi and Shlonsky (2007), who also argue that this *il* (referred to below as *–il*) is not quite the same as regular 3p.sg expletive *il*.[43] The analysis raises a number of interesting questions, I simply refer the reader to Taraldsen (2001) and Rizzi and Shlonsky (2007) for relevant discussion. What is important for our purposes is that we are dealing here with insertion of a pronominal element, which takes place if something goes wrong with local subject wh-movement (i.e. in the configuration in (29), which, as discussed above, is disallowed). Insertion of this *-il* thus has a last resort flavor; it can be considered to be a rescue strategy for local subject wh-movement constructions. In this respect, this *-il*-insertion in fact works in exactly the same way as AF insertion in Kaqchikel. The above account of Kaqchikel can then be extended to French, with a small difference in that French T can be treated exactly like English T when it comes to labeling (given the relative poverty of French verbal morphology). As in Chomsky's (2014) proposal regarding English, French T is too weak to label. Normally, the subject moves to merge with T to strengthen it, so that labeling is possible. In a local subject wh-movement context (see (29)), such movement is not possible: As discussed above, the object formed as a result of such merger cannot be labeled

---

confirms that object-agreement morphology should not be tied to labeling and should be treated differently from subject-agreement morphology.

[43] For example, they show that while *-il* has phi-features it does not have the same phi-feature specification as regular expletive *il*, being valued during the syntactic derivation. (Basically, Rizzi and Shlonsky 2007 show that in contrast to *il*, which is 3p.sg, *-il* agrees with the wh-subject in phi-features (thus, Rizzi and Shlonsky 2007 show that it is plural when the subject is plural, as revealed through liaison). Given that *–il* occurs only in the context where the subject moves to SpecCP, from where it c-commands Infl, in light of Bošković's (2007, 2008) proposal regarding case discussed in footnote 42, we can assume that the subject in SpecCP probes the IP domain, getting case this way and also valuing the phi-features of *–il*. When the phi-features of the pronominal in question are not valued this way, they get default 3.p.sg specification, which is what happens with "regular" *il*.

immediately. As a result, subject movement to SpecCP violates antilocality.[44] In other words, wh-subjects cannot pass through SpecTP on their way to SpecCP. The latter movement cannot be avoided; without it we would get a PIC violation. The former movement can be avoided if the language has another way of strengthening T for labeling (in more traditional terms, satisfying its EPP requirement). The suggestion is then that this is precisely what this *-il* insertion accomplishes. It is then treated in exactly the same way as Kaqchikel AF-insertion. It is a last-resort strategy which takes place to strengthen T in the contexts where the subject is independently prevented from moving to SpecTP.

4.6. Some alternatives

In this section I note several recent alternatives to the proposed account of local subject wh-movement constructions. Consider Chomsky's (2014) account of the *that*-trace effect in English. As discussed above, Chomsky argues that T is too "weak" to label in English when it merges with vP (or whatever its complement is, given split Infl). The subject then must move to SpecTP for labeling purposes (under (phi) feature sharing between the subject and T). Since T is too weak to label, if the subject were to move away there could be no labeling (given that traces are ignored for the purpose of labeling). This is the source of the *that*-trace effect: In constructions like (3), the subject must move to SpecCP given that CP is a phase; since traces are ignored for the purpose of labeling and T itself is too weak to label, TP cannot be labeled in such constructions. What about *Who do you think left*? Chomsky argues that in such examples C transfers all of its features to T, including the phasehood property. In clauses with a null complementizer, *that* is actually deleted: when this occurs, T becomes a phase. Subject can now move away from SpecTP because the label that was created when the subject and T merged can be now preserved, the crucial assumption here being that once a phase is completed all labels are set hence movement away will not "delabel" the relevant object, i.e. in the case at hand, movement of the subject will not "de-label" TP.[45]

The analysis appears to leave a number of constructions discussed above unaccounted for. Under this analysis subject wh-movement via SpecTP should be possible only when there is no CP. This, however, cannot be right given the data discussed above: we have seen that if there is another projection between TP and CP, subject movement via SpecTP is possible (i.e. we have seen a number of cases of this type).[46]

Rizzi (2006, 2013) also proposes rather interesting analyses where subject cannot move to SpecCP from SpecIP. In fact, he proposes an even stricter account where subject that moves to SpecTP can no longer move at all since he considers SpecTP a Criterial Freezing Position, from

---

[44] As in English and Kaqchikel, there is an improvement with an intervening adverb (thus, (i) is better than (74)a), which can be captured in the manner discussed above.

(i) ?Quelle étudiante$_i$ crois-tu que dans deux jours t_i va partir?
which student believe-you that in two days is-going to-leave

[45] Under the phase-labeling approach, we actually might expect labeling to be set only for what is sent to spell-out, hence not for the phasal-edge area. However, Chomsky's analysis of *Who do you think left* then would not work.

[46] I point out another, independent issue with Chomsky's analysis. If all labeling is done only at the phasal level, many constructions where both the head and the element in its Specifier move would be left unlabelled.

which no movement is possible (see Rizzi 2013 for an account of this). This analysis has the same problem as Chomsky's analysis regarding constructions where subject wh-movement apparently does proceed via IP.

4.7. Short subject questions

The grammaticality of simple English wh-questions like (75) now becomes puzzling. Given the above discussion, such examples cannot be derived by having the wh-phrase move to SpecIP and then to SpecCP since this derivation would violate antilocality.

(75) Who left?

There are in fact a number of proposals in the literature to derive (75) without involving wh-movement via SpecTP. Thus, a number of authors have argued that the wh-subject in such constructions in facts stays in SpecTP, an account that is fully compatible with the system argued for here.

There are, however, some arguments against this analysis (for relevant discussion, see also Agbayani 2000, An 2007, Boeckx 2003, Pesetsky and Torrego 2001, among others). Consider for example the distribution of the *hell*-phrases (cf. Ginzburg and Sag 2000, Pesetsky and Torrego 2001). If the data in (76)-(78) are taken to indicate that *the hell* can only occur with wh-phrases in SpecCP, it then follows that the subject wh-phrase in (79) must be located in SpecCP, not SpecTP.

(76) Who bought what?
(77) What the hell did John buy?
(78) *Who bought what the hell?
(79) Who the hell arrested Mary?

While the argument is rather tempting, it might not be completely conclusive, since one could try to argue that for some reason *hell*-phrases are not possible with wh-phrases located within vP. There are, however, other arguments that the wh-phrase does not remain in SpecIP in subject questions. Thus, Mizuguchi (2014) observes that if that were the case, we would expect (80) to be ambiguous, on a par with (81), which is not the case.

(80) Who loves everyone? (who>everyone; *everyone>who)
(81) Someone loves everyone. (someone >everyone; everyone>someone)

Furthermore, the West Ulster English (WUE) data in (82)-(83), noted by McCloskey (2000), provide a rather conclusive argument that subject wh-phrases do not stay in SpecIP.

(82) Who_i was arrested all t_i in Duke Street?
(83) *$They_i$ were arrested all t_i last night. (McCloskey 2000)

While, in contrast to Standard English, WUE allows Q-float under wh-movement, just like standard English, WUE disallows (83). The ungrammaticality of (83) indicates that a subject located in SpecIP cannot float a quantifier in the postverbal position in passives. It must then be the case that *who* in (82) is not located in SpecIP; if that were the case (82) should be ungrammatical on a par with (83) since subjects located in SpecIP cannot float a quantifier in this context. This example in fact also rules out the derivation where *who* in (82) moves to SpecCP via SpecTP. If that were the case, the quantifier in (82) would still be floated under movement to SpecIP, which (83) indicates is not possible. In fact, we have seen above that this derivation would anyway be ruled out by antilocality, on a par with a host of similar cases. Based on the above considerations, McCloskey (2000) in fact argues that *who* in (82) moves directly to SpecCP, without moving to SpecIP. The reasoning seems sound; notice also that the derivation in question conforms with antilocality. But what happens with the usual requirement for the subject to move to SpecIP in English in such constructions?

I will take the above data to provide clear motivation that local subject wh-questions in English are derived through direct movement to SpecCP and give two suggestions why the subject-movement-to-SpecIP requirement is apparently voided in this context (for alternative accounts, see McCloskey 2000 and Bošković 2004).

4.7.1. Q in T

Chomsky (2008, 2013) argues that there is C-T association through which the heads in question share features. This means that when there is a Q-feature in C, there is also a Q-feature in T. Assume that this is indeed the case. I also suggest that, like AF in Kaqchikel and *–il* in French, the Q-feature strengthens T so that T can label. Given that labeling is the reason for the movement of the subject to SpecTP there is then no need for the subject to move to SpecTP in this context; the subject can move directly to SpecCP.[47] Notice that under this account, the voiding of the traditional EPP requirement in English is confined to a +wh-CP context, i.e. to constructions where TP is immediately dominated by interrogative CP. As a result, the account does not extend to other short subject A'-movement contexts discussed above, namely short subject topicalization, short-subject relativization, and the *that*-trace effect.

4.7.2. T-to-C

[47] Why can't the subject then stay vP internally in non-subject wh-questions, as in **What has bought John* (the issue noted in this footnote also arises under the alternative analysis discussed in section 4.7.2 and can be handled in the same way)? There are two reasons: As discussed above, Bošković (2007, 2008) argues that the subject needs to c-command its case assigner, T, for case-licensing reasons: this in itself would force subject movement to SpecTP (as discussed in Bošković 2007, some languages have a way of case-marking the subject that does not require the subject to c-command T, such constructions could then be acceptable in such languages under this view). Alternatively, Chomsky (2013) suggests that the subject that stays in SpecvP makes labeling of vP impossible (i.e. the subject needs to move away so that vP can be labeled; Chomsky also suggests that this requirement may not hold in all languages, referring to German).

There is an alternative analysis where the *wh*-subject in (75) also moves directly to SpecCP. English matrix wh-questions involve I-to-C movement.[48] This is the case in (75) too. (Under the analysis about to be proposed, the T-to-C movement from usual questions would also take place in embedded wh-subject questions as last resort since that would be the only way to ensure proper labeling). T-to-C movement eliminates the problem that arises due to the inability of T to label on its own. Once T moves away, given that traces are ignored for the purpose of labeling, the result of T+vP merger is labeled by vP (this would give us two vacuous vP projections; as far as I can tell, no obvious problems arise because of this; notice also that if, as several works have suggested, the subject must be located in the Spec of a projection that has T in its head position in order to get case, this does happen under the analysis currently under consideration).[49] Notice that this analysis is again specific to +Wh-CP environments, i.e. contexts where TP is immediately dominated by a +wh-CP hence it does not extend to other constructions involving short subject movement discussed above, namely, short-subject topicalization, short-subject relativization, and the *that*-trace effect, where subject movement through SpecTP still must take place.

5. Conclusion

The paper has provided a uniform account of a number of locality effects, in particular, the Subject Condition, the Adjunct Condition, Richards's (2001) tucking in effect, and the full Comp-trace paradigm, including (in addition to the basic case and its improvement with intervening adverbs) relative and extraposed clauses, the impossibility of short-subject topicalization, French *que-qui* alternation, and the effect of subject wh-movement on agreement in languages like Kinande. It should, however, be noted that in spite of the wide coverage, the system argued for here does not purport to offer a comprehensive theory of all of locality of movement—there are certainly cases that are not covered by the current analysis (though the reader is referred to Bošković 2014b for an account of a generalized version of the Complex NP Constraint which extends to all complements of all lexical heads which is also crucially based on the approach to antilocality argued for here).[50] The account proposed here is based on the (slightly modified) labeling framework of Chomsky (2013, 2014) and a labeling-based approach to antilocality, which treats it as a derivational constraint on movement steps that essentially requires movement to cross a labeled category (another way to think of this is that antilocality is defined on phrases, as in Bošković (2005, 2014a),

[48] It is sometimes assumed that this is not the case with subject questions due to the lack of *do*-support. However, such questions do not show *do*-support for independent reasons discussed already in Chomsky (1957). Under the PF merger account of English verbal morphology (Halle and Marantz 1993, Bobaljik 1995, Lasnik 1995), which goes back to Chomsky (1957), *do* is inserted as last resort when the tense affix is not PF-adjacent to the verb hence cannot hop onto it. This is what happens in matrix questions, where, due to I-to-C movement, the subject intervenes between the Tense affix and the verb. However, this is not the case in subject questions even if I moves to C. Since the subject is located in SpecCP, the tense affix in C is PF-adjacent to the verb (no phonologically realized material intervenes), hence there is no need for *do*-support, which is then blocked.

[49] In fact, as discussed in Bošković (2008), due to such case-licensing, SpecCP in the construction under consideration would count as an ambiguous A'/A position hence the grammaticality of examples like *Who seems to himself to be smart*.

[50] See also Bošković (in preparation) for a labeling-based account of the freezing effect, i.e. the ban on movement out of moved elements.

but that unlabelled elements do not count as phrases). The guiding intuition of previous works on islands (e.g. Chomsky's 1986 *Barriers*), and more generally locality trouble makers, is that they are in a sense strong, with something about their strength blocking movement. The current approach is different in this respect: what makes something impenetrable is that it is underdetermined (due to the lack of a label, which makes movement from (some) islands too short). In other words, the problem with the relevant contexts is not that movement out of them is too long, as has been standardly assumed, but that it is too short.[51]

While a number of other proposals were made in the course of the discussion, one proposal needs to be emphasized: the account of the phenomena discussed in the paper is crucially based on the proposal that there is a difference in the timing of labeling between the basic case where a head and a phrase merge (the base step) and the case where two phrases merge, where the former takes place in the course of the derivation and the latter takes place as part of the spell-out operation. As discussed above, in Chomsky (2013), the label in the former case is determined rather differently from the latter case: in the former case the label is determined via minimal search (MS), the same operation as Agree Closest, a syntactic mechanism falling under minimal computation. Since the labeling of the base step is done through essentially a syntactic mechanism, I have argued that it takes place when the relevant configuration is created. On the other hand, labeling in the case of merger of two phrases occurs when the relevant structure is sent to the interfaces, given Chomsky's assumption that unlabeled objects are uninterpretable. It should, however, be noted that the latter assumption is not obvious, i.e. it is not obvious that labels are indeed needed for interpretation. At least with respect to the phenomena discussed here, nothing would go wrong in the syntax itself if the result of merger of two phrases is never labeled. If labels were not to be needed for interpretation (contra Chomsky 2013; for relevant discussion, see Chametzky 2000, Collins 2002, Seely 2006, Hornstein 2009, Hornstein and Nunes 2008), such cases then would need not be labeled at all (notice that adjunction, which in footnote 10 was suggested not to require labeling for interpretation, also involves merger of two phrases). If the current discussion is on the right track, the result of head-complement merger still needs to be labeled. However, if labels are not needed for interpretation, the labeling here would then need to take place for a strictly syntactic reason. It seems plausible that in this case labeling may be required by subcategorization, i.e. that satisfying subcategorization requires that the element with the requirement to take a complement projects (i.e. determines the label of the resulting object), otherwise, there would be no head-complement relation here; see in this respect Chomsky (2000).[52] There may even be a more pressing reason for

[51] The actual situation is slightly more complicated. Movement needs to struck the right balance: it cannot be too short or too long. In the cases in question it is not possible to accomplish that.

[52] Another way to look at this is in terms of Chomsky's (2000, 2001) assumption that all merge must have a reason. In the case of head-complement merger, the merger is driven by subcategorization requirements (which, as discussed above, also determine labeling); non-complement merger would have more than one possible motivation: e.g. because of the PIC (for the relevant element to avoid being sent to spell out, see Bošković 2007), for feature licensing (which would hold with movement to criterial positions—see Bošković 2007 for an implementation of this), or to enable labeling (where the relevant head is unable to label on its own—this case might ultimately boil down to satisfying subcategorization requirements under the above assumption about such requirements and labeling). Obviously, the discussion in this footnote, and the concluding section more generally, raises a number of questions, addressing which would go way beyond the scope of this paper, which merely offers some speculations regarding potential avenues for research that the issues discussed in this paper may open up.

immediate labeling for the head-complement case. Phases are taken to determine the point of spell out, i.e. when the structure is sent to the interfaces. If labeling occurs strictly for interpretive reasons we would expect it to occur at this point. But a serious chicken-or-the-egg style question arises then: as discussed in Bošković (in preparation), phasehood determination requires labeling, i.e. phases do not really exist prior to labeling: to know whether something is a phase we need to know its label (see Bošković in preparation for evidence that unlabelled elements cannot be phases; it is shown in that work that the well-known ban on movement out of moved elements can be deduced given that unlabelled objects cannot be phases). Since phases determine the points of spell out, without any labeling structure cannot be sent to the interfaces, which in turn is necessary for labeling under a purely interpretative approach to labeling. The problem does not arise if head-complement merger is labeled immediately since this is actually all that is needed to determine points of spell out.

References

Abels, Klaus. 2003. Successive cyclicity, antilocality, and adposition stranding. Doctoral dissertation, University of Connecticut.

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

Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory* 17:673-711.

An, Duk-Ho. 2007. Syntax at the PF interface. Doctoral dissertation, University of Connecticut, Storrs.

Aoun, Joseph, and Yen -Hui Audrey Li. 1991. The interaction of operators. In *Principles and parameters in comparative grammar*, ed. by Robert Freidin, 163-181. Cambridge, Mass.: MIT Press.

Belletti, Adriana. 1990. *Generalized verb movement*. Turin: Rosenberg and Sellier.

Bobaljik, Jonathan. 1995. Morphosyntax: The syntax of verbal inflection. Doctoral dissertation, MIT, Cambridge, Mass.

Bobaljik, Jonathan, and Dianne Jonas. 1996. Subject positions and the roles of TP. *Linguistic Inquiry* 27:195-236.

Boeckx, Cedric. 2003. *Islands and Chains*. Amsterdam: John Benjamins.

Boeckx, Cedric. 2005. Some notes on bounding. Ms., Harvard University, Cambridge, Mass.

Boeckx, Cedric. 2008. *Bare Syntax*. Oxford: Blackwell.

Borer, Hagit. 1984. *Parametric syntax*. Dordrecht: Foris.

Bošković, Željko. 1992. Clausal selection, subjacency, and minimality. Ms., University of Connecticut, Storrs.

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*. Cambridge: MIT Press.

Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena.* Amsterdam: Elsevier Science.

Bošković, Željko. 2002. A-movement and the EPP. *Syntax* 5:167-218.

Bošković, Željko. 2004. Be careful where you float your quantifiers. *Natural Language and Linguistic Theory* 22:681-742.

Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59:1-45.

Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38: 589-644.

Bošković, Željko. 2008. On successive cyclic movement and the freezing effect of feature checking. In *Sounds of silence: Empty elements in syntax and phonology*, ed. by Jutta M. Hartmann, Veronika Hegedüs, and Henk van Riemsdijk, 195-233. Amsterdam: Elsevier.

Bošković, Željko. 2009. On relativization strategies and resumptive pronouns. In *Studies in Formal Slavic Phonology, Morphology, Syntax, Semantics and Information Structure. Proceedings of FDSL 7, Leipzig 2007*, ed. by Gerhild Zybatow, Uwe Junghanns, Denisa Lenertová, Petr Biskup, 79-93. Frankfurt am Main: Peter Lang.

Bošković, Željko. 2011. Rescue by PF deletion, traces as (non)interveners, and the *that*-trace effect. *Linguistic Inquiry* 42:1-44.

Bošković, Željko. 2013. Traces do not head islands: What can PF deletion rescue? In *Deep insights, broad perspectives: Essays in honor of Mamoru Saito*, ed. by Yoichi Miyamoto, Daiko Takahashi, Hideki Maki, Masao Ochi, Koji Sugisaki, Asako Uchibori, 56-93.Tokyo:Kaitakusha.

Bošković, Željko. 2014a. Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45: 27-89.

Bošković, Željko. 2014b. From the Complex NP Constraint to Everything: On deep extractions across categories. Ms., University of Connecticut, Storrs. Downloadable at http://ling.auf.net/lingbuzz/002278

Bošković, Željko. in preparation. The ban on movement out of moved elements in the phasal/labeling system. Ms., University of Connecticut.

Bošković, Željko, and Howard Lasnik. 2003. On the distribution of null complementizers. *Linguistic Inquiry* 34:527-546.

Brandi, Luciana and Patricia Cordin. 1989. Two Italian dialects and the null subject parameter. In *The null subject parameter*. Springer.

Branigan, Philip. 2005. The Phase Theoretic Basis for Subject-Aux Inversion. Ms., Memorial University of Newfoundland.

Bresnan, Joan W. 1972. Theory of complementation in English syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Brillman, Ruth and Aron Hirsch. 2014. An anti-locality account of English subject/non-subject asymmetries. Paper presented at the 50th meeting of the Chicago Linguistic Society.

Broekhuis, Hans. 2005. Extraction from subjects: some remarks on Chomsky's *On phases*. In *Organizing the grammar. Studies in honor of Henk van Riemsdjik*, ed. Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz and Jan Koster, 59-68. Berlin: Mouton de Gruyter.

Browning, Maggie. 1996. CP recursion and *that*-t effects. *Linguistic Inquiry* 27:237-255.

Buesa García, Carlos. 2011. The interaction between locality and the subject-gap restriction in Spanish questions. In *Proceedings of the North East Linguistic* Society 39, ed. by Suzi Lima, Kevin Mullin and Brian Smith, 163-176. Amherst: GLSA, University of Massachusetts.

Campos, Héctor. 1997. On subject extraction and the anti-agreement effect in Romance. *Linguistic Inquiry*.

Carstens, Vicki, Norbert Hornstein and T. Daniel Seely. 2013. Head movement in problems of projection. Ms., University of Missouri/University of Maryland/Eastern Michigan University.

Chametzky, Robert. 2000. *Phrase structure: From GB to Minimalism*. Blackwell.
de Chene, Brent. 1995. Complementizer-trace effects and the ECP. *Geneva Generative Papers* 3.1: 1–4. Département de Linguistique Générale, Université de Genève.
Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
Chomsky, Noam. 1986. *Barriers*. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 2000. Minimalist inquires. *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels and Juan Uriagereka, 89-155. Cambridge MA: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 1-52. Cambridge: MIT Press.
Chomsky, Noam. 2008. On phases. In *Foundational issues in Linguistic theory. Essays in honor of Jean-Roger Vergnaud* ed by Robert Freidin, Carlos Peregrin and Maria Luisa Zubizarreta. 133-166. Cambridge: MIT Press.
Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33-49.
Chomsky, Noam 2014. Problems of projection: extensions. Ms., MIT, Cambridge.
Chomsky, Noam, and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 11:1-46.
Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective.* Oxford University Press.
Collins, Chris. 2002. Eliminating labels. In *Derivation and explanation in Minimalist syntax* ed. By Samuel David Epstein and T. Daniel Seely, 42-64. Wiley.
Collins, Chris. 2014. Merge(X.Y) = {X,Y}. Ms., New York University.
Coon, Jessica, Pedro Mateo Pedro, and Omer Preminger. In press. The role of case in A-bar extraction asymmetries: Evidence from Mayan. *Linguistic Variation*.
Culicover, Peter. 1992. The adverb effect: Evidence against ECP accounts of the *that-t* effect. *NELS 23:* 97-11.
Culicover, Peter and Michael Rochemont. 1992. Adjunct extraction from NP and the ECP. *Linguistic Inquiry* 23:496-501.
Deal, Amy Rose. 2014. Cyclicity and connectivity in Nez Perce relative clauses. Ms., University of California Santa Cruz.
Dikken, Marcel den. 1995. *Particles*. Oxford: Oxford University Press.
Dikken, Marcel den. 2006. *Relators and Linkers: The syntax of predication, predicate inversion and Copulas*. Cambridge, MA: MIT Press.
Dikken, Marcel den. 2007. Phase extension: Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33:1–41.
Epstein, Samuel David, Hisatsugu Kitahara and T. Daniel Seely. 2014. Labeling by minimal search: Implications for successive-cyclic A-movement and the conception of the postulate “Phase”. *Linguistic Inquiry* 45:463-481.
Erlewine, Michael. 2014. Anti-locality and optimality in Kaqchikel Agent Focus. Ms., MIT.
Fox, Danny, and Howard Lasnik. 2003. Successive-cyclic movement and island repair: The difference between sluicing and VP-ellipsis. *Linguistic Inquiry* 34:143-154.
Gallego, Ángel J. and Juan Uriagereka. 2006. Conditions on sub-extraction. Ms., Universitat Autònoma de Barcelona and University of Maryland, College Park.

Gallego, Ángel J. and Juan Uriagereka. 2007. Sub-extraction from subjects: A phase theory account. In *Romance Linguistics 2006*, ed. by José Camacho, Nydia Flores-Ferrán, Liliana Sánchez, Viviane Déprez and María José Cabrera, 149-162. Amsterdam: John Benjamins.

Ginzburg, Jonathan and Ivan A. Sag. 2000. *Interrogative Investigations*. CSLI Publications.

Grohmann, Kleanthes. 2003. *Prolific domains: On the antilocality of movement dependencies*. Amsterdam: John Benjamins.

Grohmann, Kleanthes. 2011. Antilocality: Too close relations in grammar. In *The Oxford handbook of linguistic minimalism*, ed. by Cedric Boeckx, 260-290. Oxford: Oxford University Press.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 111-176. Cambridge, Mass.: MIT Press.

Hiraiwa, Ken. 2005. Dimensions of symmetry in syntax: Agreement and clausal architecture. Doctoral dissertation, MIT, Cambridge, Mass.

Hoge, Kerstin. 2001. *That*-t effects in English and Yiddish. In *The Minimalist Parameter*, ed. by Galina Alexandrova and Olga Arnaudova. 233-48. Amsterdam: John Benjamins.

Holmberg, Anders, and Thorbjörg Hróarsdóttir. 2003. Agreement and movement in Icelandic raising constructions. *Lingua* 113: 997-1019.

Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and universal grammar*. Cambridge: Cambridge University Press.

Hornstein, Norbert and Jairo Nunes. 2008. Adjunction, labeling and bare phrase structure. *Biolinguistics* 2:57-86.

Hornstein, Norbert, and Witkoś, Jacek. 2003. Yet another approach to existential constructions. In *Grammar in Focus: Festschrift for Christer Platzack*. eds. Lars-Olof Delsing, Cecilia Falk, Gunlög Josefsson and Halldór Ármann Sigurðsson. Lund University, Department of Scandinavian.

Hornstein, Norbert, Howard Lasnik and Juan Uriagereka. 2003. The dynamics of islands: Speculations on the locality of movement. *Linguistic Analysis* 33:149-175.

Huang, C. -T. James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT, Cambridge, Mass.

Hunter, Tim. 2010. Relating movement and adjunction in syntax and semantics. Doctoral Dissertation, University of Maryland, College Park.

Ishii, Toru. 1999. Cyclic spell-out and the *that*-trace effect. *Proceedings of WCCFL* 18: 220-231.

Ishii, Yasao. 2004. The phase impenetrability condition, the vacuous movement hypothesis, and *that*-effects. *Lingua* 114:183-215.

Jeong, Youngmi. 2006. The landscape of applicatives. Doctoral dissertation, University of Maryland, College Park.

Kandybowicz, Jason. 2006. Comp-trace effects explained away. In *Proceedings of the 25th West Coast Conference on Formal Linguistics*, ed. by Donald Baumer, David Montero, and Michael Scanlon, 220-228. Somerville: Cascadilla Press.

Kandybowicz, Jason. 2008. *The grammar of repetition: Nupe grammar at the syntax phonology interface*. Amsterdam: John Benjamins.

Kayne, Richard. 1984. *Connectedness and binary branching*. Dordrecht: Foris.

Kim, Sun-Woong. 2008. English C moves downward as well as upward: An extension of Bošković and Lasnik's (2003) approach. *Linguistic Inquiry* 39:295-307.

Lasnik, Howard. 1995. Verbal morphology: *Syntactic structures* meets the Minimalist Program. In *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero*, ed. by Héctor Campos and Paula Kempchinsky, 251-275. Washington, D. C.: Georgetown University Press.

Lasnik, Howard. 1999. *Minimalist analysis*. Oxford: Blackwell.

Lasnik, Howard. 2001. When can you save a structure by destroying it? In *Proceedings of the North East Linguistic Society* 31, ed. by Min-Joo Kim and Uri Strauss, 301-320. Amherst: GLSA, University of Massachusetts.

Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15:235-289.

Lasnik, Howard, and Mamoru Saito. 1992. *Move α: Conditions on its application and output*. Cambridge: MIT Press.

Law, Paul. 1993. On the base position of wh-adjuncts and extraction. Paper presented at the 67th Annual Meeting of the Linguistic Society of America, Los Angeles, California, January 1993.

Lohndal, Terje. 2007. Subextraction and the freezing effect: A Scandinavian case study. Ms., University of Oslo.

Lohndal, Terje. 2009. Comp-T effects: Variation in the position and features of C. *Studia Linguistica* 63:204-232.

McCloskey, James. 1991. Resumptive pronouns, A'-binding, and levels of representation in Irish. In *Syntax and Semantics 23: The Syntax of the Modern Celtic Languages*, ed. by Randall Hendrick, 199-248. New York: Academic Press.

McCloskey, James. 2000. Quantifier Float and wh-movement in an Irish English. *Linguistic Inquiry* 31:57-84.

Merchant, Jason. 2001. *The syntax of silence.* Oxford: Oxford University Press.

Mizuguchi, Manabu. 2008. Derivation, minimalism and *that*-trace effects. *English Linguistics* 25:56-92.

Mizuguchi, Manabu. 2014. Phases, labeling, and wh-movement of the subject. Presented at the 32nd Conference of the English Linguistic Society of Japan, Gakushuin University, Tokyo, November 8.

Moro, Andrea. 1997. *The raising of predicates. Predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.

Müller, Gereon. 2010. On Deriving CED Effects from the PIC. *Linguistic Inquiry* 41:35-82.

Nomura, Masashi. 2005. Nominative case and AGREE(ment). Doctoral dissertation, University of Connecticut, Storrs.

Ott, Dennis. In press. Symmetric merge and local instability: Evidence from split topics. *Syntax*.

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

Pesetsky, David. 1981. Complementizer-trace phenomena and the nominative island condition. *The Linguistic Review* 1:297-343.

Pesetsky, David. 1992. Zero Syntax Vol. *2*. Ms., MIT, Cambridge.

Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 355-426. Cambridge, Mass.: MIT Press.

Preminger, Omer. 2011. Agreement as a fallible operation. Doctoral Dissertation, Massachusetts Institute of Technology.

Reinhart, Tanya. 1980. On the position of extraposed clauses. *Linguistic Inquiry* 11:621-624.

Richards, Norvin. 1999. Complementizer cliticization in Tagalog and English. *Proceedings of AFLA* 16:297-312.

Richards, Norvin. 2001. *Movement in language: Interactions and architectures*. Oxford: Oxford University Press.

Riqueros, José. 2013. Spanish Nominal(ization) Patterns. Doctoral dissertation, University of Connecticut, Storrs.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge: MIT Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*. ed. by Liliane Haegeman, 281-337. Dordrecht: Kluwer.

Rizzi, Luigi. 2006. On the form of chains: Criterial positions and ECP effects. In *Wh-movement: Moving on*, ed. by Lisa Lai-Shen Cheng and Norbert Corver, 97-133. Cambridge: MIT Press.

Rizzi, Luigi. 2013. Cartography, criteria and labeling. Ms., Université de Genève.

Rizzi, Luigi and Ur Shlonsky. 2007. Strategies of Subject Extraction. In *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics,* ed. by H.-M.Gärtner and U. Sauerland, 115-16. Berlin: Mouton de Gruyter.

Ross, John Robert. 1969. Guess who? In *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society,* ed. by Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan, 252-286. Chicago: University of Chicago, Chicago Linguistic Society

Roussou, Anna. 2002. C, T, and the subject: That-t phenomena revisited. *Lingua* 112:13-52.

Rudin, Catherine. 1988. On multiple questions and multiple *wh*-fronting. *Natural Language and Linguistic Theory* 6:445-501.

Sabel, Joachim. 2000. Expletives as features. In *The Proceedings of the Nineteenth West Coast Conference on Formal Linguistics,* ed. by Roger Billerey and Brook Danielle Lillehaugen, 101-114. Somerville, Mass.: Cascadilla Press.

Saito, Mamoru. 2013. Case and labeling in a language without φ-feature agreement. Ms., Nanzan University.

Saito, Mamoru, and Keiko Murasugi. 1999. Subject predication within IP and DP. In *Beyond principles and parameters,* ed. Kyle Johnson and Ian Roberts, 167-188. Dordrecht: Kluwer.

Schneider-Zioga, Patricia. 1995. Specifier/head agreement in Kinande. *Cahiers Linguistiques d'Ottawa* 23:67-93.

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

Schneider-Zioga, Patricia. 2009. Wh-agreement and bounded and unbounded movement. In *Merging Features* ed. By Jose Brucart, Anna Gavorro and Juame Sola. Oxford University Press.

Seely, T. Daniel. 2006. Merge, derivational c-command, and subcategorization in a label-free syntax. In *Minimalist Essays,* ed. Cedric Boeckx. Amsterdam: John Benjamins.

Stepanov, Arthur. 2001. Cyclic domains in syntactic theory. Doctoral dissertation, University of Connecticut.

Stepanov, Arthur. 2007. The End of CED? Minimalism and Extraction Domains. *Syntax* 10: 80-126.

Stjepanović, Sandra. 1999. What do second position cliticization, scrambling, and multiple wh-fronting have in common. Doctoral dissertation, University of Connecticut, Storrs.

Stowell, Timothy. 1981. Origins of phrase structure. Doctoral dissertation, MIT.

Szczegielniak, Adam. 1999. *That*-t effects crosslinguistically and successive cyclic movement. *MIT Working Papers in Linguistics* 33: 369-393. Cambridge, Mass.: Department of Linguistics, MITWPL.

Takita, Kensuke, Nobu Goto, and Yoshiyuki Shibata. 2014. Labeling through Spell-Out. Presented at GLOW in Asia X, National Tsing Hua University, Taiwan.

Talić, Aida. 2014. Extraordinary complement extraction: PP-complements and inherently case-marked nominal complements. *Studies in Polish Linguistics* 8: 127-150.

Taraldsen, Knut T. 2001. Subject extraction, the distribution of expletives and stylistic inversion. In *Subject Inversion in Romance and the theory of universal grammar* 6. 16-182. New York: Oxford University Press.

Ticio, Emma. 2005. Locality and anti-locality in Spanish DPs. *Syntax* 8: 229-286.

Truswell, Robert. 2011. *Events, phrases, and questions*. Oxford: Oxford University Press.

Uriagereka, Juan. 1988. On government. Doctoral dissertation, University of Connecticut, Storrs.

Uriagereka, Juan. 2012. *Spell-out and the minimalist program*. Oxford: Oxford University Press.

Watanabe, Akira. 1993. Larsonian CP Recursion, factive complements, and selection. *NELS* 23:523-537.

Wurmbrand, Susi. 2014a. Tense and aspect in English infinitives. *Linguistic Inquiry* 45:403-447.

Wurmbrand, Susi. 2014b. Tagalog infinitives: Consequences for the theory of phases, voice marking and extraction. Ms., University of Connecticut, Storrs.

Yoo, YongSuk. In preparation. Labeling algorithm and its consequences for multiple movement. Ms., University of Connecticut, Storrs.