

Revisiting extraction and subextraction patterns from arguments

Luis Miguel Toquero-Pérez

toquero@usc.edu

Abstract

Extraction and subextraction tend to receive separate attention in syntax, which leads to the assumption that they should be analyzed independently, even though they both illustrate an asymmetry between subjects and objects. By looking at various phenomena in English, German, Spanish and Norwegian I propose that this parallel behavior is not accidental, but that there is a previously unnoticed generalization: subextraction is allowed iff extraction is possible and the target of subextraction is not an indirect object. In order to account for the facts, I propose that a revised version of Spec-to-Spec antilocality (Erlewine 2016) is necessary: movement of and out of an XP must cross a *Projection Line (PL)* (Brody 1998), i.e. the set of all projections of a head. I show that this version of antilocality is appropriate to derive *Freezing* effects, Huang's (1982) *CED*, and their exceptions; and *Comp-trace* effects and their neutralization, including subextraction cases. However, antilocality on its own cannot derive the extraction-subextraction asymmetry in indirect objects. I propose that the Principle of Minimal Compliance (PMC) (Richards 1998) can suspend antilocality if agree between a probe and a goal has happened first. The version adopted here will allow extraction of the whole XP, but disallow extraction of its specifier due to the lack of a prior agree relation. Antilocality and the PMC combined also make the right predictions in other domains such as the lack of *do*-support in matrix subject questions and A-movement of the subject in declarative clauses, providing evidence that antilocality is a constraint that should apply to (at least) both A and A'-movement.

Keywords : Extraction, subextraction, freezing, CED, antilocality, Principle of Minimal Compliance

1 Introduction

The relevance of subject islands to syntactic theory has been clear since Ross's (1967) dissertation. In the course of the generative enterprise, several proposals have attempted to explain the asymmetry between (1) and (2) where the only difference is the position from which subextraction of the *wh*-element takes place (the subject position or the object position respectively):

- (1) **[Which city]₁ did [pictures of t₁] cause a scandal?*
- (2) *[Which city]₁ did you see [pictures of t₁]?*

There are several approaches that have tried to explain this asymmetry between objects and subjects. Some of them have been concerned with the fact that only those elements that do not move prior to the subextraction or that are (internally or externally) merged in a subject position are suitable to such operation. To mention two classic analyses, there is Wexler and Culicover's (1980) *Freezing* that prevents extraction out of moved constituents and Huang's (1982) *Condition on Extraction Domains* that bans extraction from XPs on subject positions. The asymmetry in (1) and (2) is captured by these grammatical constraints which, as they stand, make the following predictions: movement out of subjects should never be possible assuming that they occupy the specifier position of *v* and movement out of a shifted or scrambled object should not be possible either. These predictions are not borne out, though:

- (3) *[Who]₂ did Kima call [friends of t₂]₁ up t₁?*
- (4) **German**
[Worüber]₂ hat keiner [ein Buch t₂]₁ sorgfältig t₁ gelesen?
 about-what has no-one a book carefully read
 'What has nobody read a book about carefully?'

(3) violates both the *CED* and *Freezing* if, following Johnson (1991) and Boeckx and Hornstein (2005), the direct object [friends of t₁] has undergone object shift from the complement position to a higher projection: the former is violated because the DP lands in a specifier (i.e. a subject position), the latter is violated under the assumption that object shift should yield the DP frozen. (4) is an instance of movement out of a scrambled constituent in German which again should be ruled out by both constraints.

Although much attention has been devoted to subextraction, most analyses (with some exceptions, e.g. Bošković 2016, 2018) seem to only focus on extraction out of XPs as an independent phenomenon; but they are not concerned with extraction of the XPs themselves. That is, is there any relation between extraction and subextraction at all? For instance, if we look at *that-trace* effects we can find the following minimal pair where both extraction of and subextraction from a

subject are not possible on the presence of the complementizer *that*:

- (5) a. [What]₁ did they say [(**that*) t₁ caused a scandal]?
 b. [Which city]₁ did they say [(**that*) [a picture of t₁] caused a scandal]?

Therefore, the goal of this paper is to revisit extraction of and subextraction domains from arguments and investigate whether there is a common pattern in the syntactic behavior of the two phenomena. In fact, I argue for a generalization according to which subextraction is possible if extraction is allowed. The hypothesis that I put forth in this paper is that only those movement operations that are not too local will render extraction and subextraction cases licit. In other words, there is an antilocality violation (Bošković 1997b,2016, Grohmann and Etxepare 2003, Abels 2003, Bošković 2005, Erlewine 2016, 2019 a.o.) that is at play in cases like (1) through (5) where the movement launching and landing sites do not meet the necessary distance requirements for the operation to take place. In addition, we see cases where antilocality seems to be violated and yet extraction is possible. Such is the case of indirect objects as in (6-a) and (6-b):

- (6) a. [Who]₁ did McNulty show t₁ a picture of Baltimore?
 b. *[Which person]₁ did McNulty show [a friend of t₁] a picture of Baltimore?

In order to capture this asymmetry, I propose that a revised form of the Principle of Minimal Compliance (PMC) first proposed by Richards (1998) (see Rackowski and Richards 2005 and Preminger 2019 for other applications of the PMC) can salvage the derivation: an Agree relation between a probe and a goal will enable the goal to be accessible to subsequent syntactic operations regardless of locality conditions with respect to that probe. This results in the desired contrast in (6): the Agree relation is between a head (i.e. *v*) and the indirect object, not between the head and the specifier of the indirect object, which entails that the *wh*-element in the specifier is still subject to antilocality. Although such an analysis assumes a structure of indirect objects that goes against the standard Low Applicative hypothesis (Pylkkänen 2002) for the languages that will be here illustrated, arguments against the transfer-of-possession relation and the observation that static verbs cannot be applicativized are provided which support the structural representation that is being followed in the paper.

Besides, if extraction and subextraction are governed by a general principle of the grammar we may want to look at a number of languages and observe whether there is a common pattern across them that can help us gain a better understanding of this phenomenon. For the purpose of this paper, data from English, Spanish, German and Norwegian have been collected. Based on the data gathered from these languages, it will be demonstrated that previous proposals in the literature face empirical challenges when trying to provide a uniform treatment of the phenomena under study. Some of these proposals include, but are not limited to, Wexler and Culicover's (1980) *Freezing*,

and Huang's (1982) *CED*, which have already been mentioned in this introduction; Erlewine's (2019) antilocality, which states that movement of a phrase must cross a maximal projection but makes inaccurate predictions with respect to subextraction from subjects in embedded TPs and indirect objects; and Bošković's (2018) *Labels*, which predicts that if an element first moves to the specifier of the containing phrase, that phrase should not be able to move. These are assessed in their own section where they are compared to the empirical coverage made in the current proposal.

The paper is organized as follows. Section 2 presents the data from the four languages sampled and provides arguments when necessary for the position of certain phrases in the clause. Section 3 takes stock and provides the generalizations that the paper is set to explain. Section 4 is concerned with the analysis and the proposal and it has three main parts: the first part in subsection 4.1 motivates and spells out the particular definition of antilocality and shows how far we can get with only this constraint. The second part in 4.2 introduces the definition of the PMC adopted here to bridge the gap created by antilocality. The third part in 4.3 concentrates on indirect objects, where the interaction of antilocality and the PMC is more salient. Section 5 connects the proposal in section 4 to the generalizations and provides a brief summary of how the generalizations are explained. Section 6 provides an overview of some of the alternative analyses that have been proposed in the literature, outlining their empirically insufficient or theoretically problematic. Section 7 is the conclusion where the main findings and predictions are brought together and discussed.

2 Extraction and Subextraction data

In this section, the data corresponding to the extraction and subextraction patterns under study will be presented. Each syntactic domain that is being looked at has its own subsection. Data from English, Spanish, German and Norwegian is provided to illustrate the phenomena.

2.1 In-situ Direct Objects and in-situ Subjects

As noted in the literature on movement (Ross 1967, Huang 1982, Rizzi 1997 to cite a few), A'-extraction of and subextraction from an object that has not moved are usually widely acceptable.¹ This claim is illustrated by the following examples from English:

- (7) English
- a. [What]₁ did you see t₁?
 - b. [Which city]₁ did you see [pictures of t₁]?
 - c. [What]₁ did you bet that they had seen t₁?
 - d. [Which city]₁ did you bet that they had seen [pictures of t₁]?

The examples show that *wh*-movement of the whole object is possible regardless of whether the extraction takes place from a root (7-a) or an embedded clause (7-d). In addition, subextraction from the object in both root and embedded contexts is also possible as illustrated by (7-b) and (7-c) respectively. The same is found in many other languages including the ones below:

(8) Spanish

- a. [Qué]₁ has visto t₁?
what have.you seen
'What have you seen?'
- b. [De qué ciudad]₁ has visto [fotos t₁]?
of which city have.you seen pictures
'Which city have you seen pictures of?'
- c. [De qué ciudad]₁ quieres que tus primos vean [fotos t₁]?
of which city want.you that your cousins see pictures
'Which city do you want your cousins to see pictures of?'

(9) German

- a. [Was für ein Buch]₁ hast du t₁ gelesen?
what for a book have you read
'What kind of book did you read?'
- b. [Was]₁ hast du [t₁ für ein Buch] gelesen?
what have you for a book read
'What kind of book did you read?'
- c. [Was]₁ glaubst du, dass Franz t₁ gelesen hat?
what think you that Franz read has
'What do you think Franz has read?'
- d. [Was]₁ glaubst du, dass Franz [t₁ für ein Buch] gelesen hat?
What think you that Franz for a book read has
'What kind of book do you think Franz has read?'

(10) Norwegian ((a-b) from Lohndal 2011: 164-5)

- a. [Hva]₁ gav du barna t₁ til bursdagen?
what gave you children for birthday
'What did you give the children for their birthday?'
- b. [Hva]₁ gav du barna [t₁ for noe] til bursdagen?
what gave you children for something for birthday
'What kind of things did you give the children for their birthday?'
- c. [Hva]₁ vil du at tydeligvis Peter lage t₁?
what want you that evidently Peter cause?
'What do you want Peter to evidently cause?'
- d. [Hva]₁ vil du at tydeligvis Peter lage [t₁ for noe]?
what want you that evidently Peter cause for something

'What kind of thing do you want Peter to evidently cause?'

If we compare these data to counterpart in-situ subjects,² we notice that there is no difference regarding extraction and subextraction patterns. English does not have in-situ subjects *per se*. *There*-existential constructions have been analyzed by Burzio (1986), Lasnik (1995) and Chomsky (1995, 2000) as involving merger of the expletive in TP directly. However, recent work by Deal (2009) has argued that *there*-existential constructions, in fact, involve merger of *there* in vP and subsequent movement to TP. Thus, the DP associate of the expletive stays in situ inside the VP. An argument provided by Deal (2009) is that *there* can only occur with unaccusative verbs; this indicates that *there* never occurs in structures that have a *v* introducing an external argument. Thus, examples like (11) should be treated as illustrating movement from the object position.

(11) English

- a. [What]₁ is there t₁ on the wall?
- b. [Which city]₁ is there [a picture of t₁] on the wall?
- c. [What]₁ did you bet that there is t₁ on the wall?
- d. [Which city]₁ did you bet that there is [a picture of t₁] on the wall?

However, it is the case that for other languages it is easier to determine the lower position of the subject. For instance, in-situ subjects might be an optional or obligatory strategy for question formation. That is the case of Spanish:

(12) Spanish

- a. [Qué]₁ causó el año pasado t₁ un escándalo?
what.PL caused the year past a scandal
'What caused a scandal last year?'
- b. [De qué ciudad]₁ causó el año pasado [una foto t₁] un escándalo?
of what city caused the year past a picture a scandal
'Of which city did a picture cause a scandal last year?'
- c. [Qué]₁ querías que causara el año pasado t₁ un escándalo?
what want.you that caused the year past a scandal
'What did you want to cause scandal last year?'
- d. [De qué ciudad]₁ querías que causara el año pasado [una foto t₁] un
of what city wanted.you that caused the year past a picture a
escándalo?
scandal
'Of which city did you want a picture to cause a scandal last year?'

The Spanish data in (12) show that the *wh*-subject is moved in one fell swoop from the in-situ position, i.e. the specifier of vP. It has been argued in the literature that Spanish requires the subject

to remain low because of the obligatory inversion rule in questions (see Rizzi 1982, Torrego 1984, Suñer 1994): this is achieved via head movement of the verb to T. In addition, one could take the position of the stranded DP subject after subextraction has taken place in (12-b) and (12-d) as evidence for this hypothesis.

The German data in (13) pattern identically to English and Spanish with respect to the position the subject occupies before extraction and during the subextraction. Given that German is V2 it is difficult to determine whether the subject has to obligatorily move to TP first or move directly from Spec-vP. One argument in favor of the in-situ status is the same that was said about Spanish: assuming that the particle *denn* is sitting in TP (Haider 2017), given that the stranded DP *für Ameisen* for ants' follows the particle, the extraction must have taken place from the in-situ position. A stronger argument that subject raising is not obligatory comes from Haider (1990) who notes that it is possible to topicalize a vP with the subject in it as shown in (14):

(13) German (Diesing 1990)

- a. [Was für Ameisen]₁ haben denn t₁ einen Postbeamten gebissen?
 what for ants have PRT a postman bit
 'What kind of ants bit the postman?'
- b. [Was]₁ haben denn [t₁ für Ameisen] einen Postbeamten gebissen?
 what have PRT for ants a postman bit
 'What kind of ants bit the postman?'

(14) German vP topicalization

- a. [Ein Aussenseiter gewonnen]₁ hat hier noch nie t₁.
 an outsider.NOM won has here yet never
 'An outsider has never won here'
- b. [Ein Aussenseiter gewonnen]₁ hat das Derby noch nie t₁.
 an outsider.NOM won has the Derby.ACC yet never
 'An outsider has never won the Derby yet'

The displaced elements in (14) must include the subject. If this constituent is smaller than TP, then this is an argument that the subject does not have to move to Spec-TP in German. Nevertheless, movement to TP could still be optional.³

Norwegian behaves similar to the German data above, given its V2 status. In fact, what the exact position of the subject is might be hard to determine. A good argument comes from the position of negation and adverbs. Negation is located between TP and vP (Vikner 1995, Svenonius 2002): thus, if the subject is high, negation should follow the subject; if the subject is low, negation should precede the subject.

(15) Norwegian

- a. Her har **tydeligvis ikke** forelesningen utløst en skandale.
 here has evidently not lecture caused a scandal
 ‘The lecture has evidently not caused a scandal here’
- b. [Hva]₁ har **tydeligvis ikke** t₁ utløst en skandale?
 what has evidently not caused a scandal
 ‘What has evidently not caused a scandal?’
- c. [om hvilket tema]₁ har **tydeligvis ikke** [forelesningen t₁] utløst en skandale?
 about which topic has evidently not lecture caused a scandal
 ‘about which topic has the lecture evidently not caused a scandal?’

One must note at this point that the VP topicalization test that we used to determine the non-obligatory movement of the subject to TP in German is not allowed in Norwegian. Nevertheless, we should not take this as an argument against the possibility that the subject can actually appear low in Norwegian. As Hein (2019) notes, contrary to German, it is the VP that is fronted in Norwegian and also other Scandinavian languages (see Platzack 2012 for Swedish). The argument that Hein provides is originally proposed by Johnson (2004) and revolves around the ambiguity of the adverb *igjen* ‘again’. According to Johnson (2004), ‘again’ (and thus its Norwegian counterpart) is ambiguous between a repetitive and a restitutive reading depending on whether it adjoins to vP or VP. In Norwegian Hein (2019) notes that the ambiguity can be resolved if topicalization of the verb phrase is involved. The baseline sentence is in (16-a) and the variants in (16-b) and (16-c):⁴

(16) Norwegian VP topicalization (Hein 2019: 45-46)

- a. Terje åpnet døra igjen.
 Terje opened door again
 ‘Terje opened the door again’
 ✓Repetitive: Terje has opened the door before
 ✓Restitutive: The door had been open before and was closed in the meantime.
- b. [(Å) åpne døra igjen] gjorde faktisk Terje.
 to open door again did actually Terje
 ‘As for opening the door again, Terje did in fact do it’
 ⊖Repetitive: Terje has opened the door before
 ✓Restitutive: The door had been open before and was closed in the meantime.
- c. [(Å) åpne døra] gjorde Terje igjen
 to open door did Terje again
 ‘As for opening the door, Terje did it again’
 ✓Repetitive: Terje has opened the door before
 ⊖Restitutive: The door had been open before and was closed in the meantime.

Given that the restitutive reading is associated with adjunction to the VP while the repetitive reading is associated with adjunction to the ν P, we now have evidence to argue that it is the VP that is fronted and not the whole ν P as in German. Nevertheless, this result crucially does not undermine the hypothesis that the subject can stay low.⁵

As the reader might have noticed, only extraction and subextraction from root clauses has been provided for German and Norwegian. I would like to briefly explain why. German and Norwegian show, as will be illustrated in subsection 2.3, *that-trace* effects and those effects are neutralized if the relevant XP remains in a lower position in the clause (i.e. sometimes the XP happens to be in-situ). Thus, this organization leaves room for a closer examination of the embedded clause in the relevant languages and enables to keep the contrast handier. In the next section, examples of movement out of moved elements are provided.

2.2 Movement out of moved elements: ECM in English, Object Shift and Scrambling

In the preceding section of the paper, it was shown that extraction out of unmoved elements (i.e. "subjects" and "direct objects") is allowed. Now, I turn my attention to cases of extraction and subextraction after the constituent has undergone movement. These include cases like ECM verbs in English, object shift and scrambling.

Although the analysis of ECM in English is not settled ground, one possible approach is a subject-to-object raising analysis (Lasnik and Saito 1991) according to which, as an option, the subject DP from the embedded clause can end up in the matrix clause, typically for case reasons. One argument for this type of analysis is provided by the fact that matrix adverbs can intervene between the ECM subject and the rest of the embedded clause; and, if the verb involves a particle such as *make out*, the raised DP can interrupt the V-Particle sequence. This is shown in (17-b) and (18-b) below:

- (17) ECM + intervening adverbs
- a. I believe **with all my heart** [[*a friend of McNulty*] to have won the election].
 - b. I believe [*a friend of McNulty*]₁ **with all my heart** [t₁ to have won the election].
- (18) V-Particle constructions
- a. Mary **made out** [[*friends of McNulty*] to be fools].
 - b. Mary **made** [*friends of McNulty*]₁ **out** [t₁ to be fools].

As the examples illustrate the DP can intervene between the matrix verb and the matrix adverb as in (17-b), and between the matrix verb and the particle following as in (18-b). Thus, this can be

taken as evidence for movement of the DP. If we look at the A'-counterpart of (17-b) and (18-b), this is what we find:⁶

- (19) ECM + intervening adverbs: A'-movement
- a. [Them]₁, I believe t₁ **with all my heart** t₁ to have won the election
 - b. (?) [Who]₂ do you believe **with all your heart** [a friend of t₂] to have won the election?
 - c. (?) [Who]₂ do you believe [a friend of t₂]₁ **with all your heart** t₁ to have won the election?
- (20) V-Particle constructions: A'-movement (Adapted from Lasnik 2001)⁷
- a. [Them]₁, Mary **made** t₁ **out** t₁ to be fools?
 - b. (?) [Who]₁ did Mary **make out** [friends of t₁] to be fools?
 - c. (?) [Who]₂ did Mary **make** [friends of t₂] **out** to be fools?

These examples suggest that not only the subject of an ECM clause can raise to the matrix clause, but also that even after movement, it can still participate in extraction and allows movement out of itself.

Moreover, if it is the case that such operations can take place, we should be able to see the same pattern in what Johnson (1991) and Boeckx and Hornstein (2005) consider "object shift", which is exhibited by V-Particle constructions that do not select for an ECM clause. This prediction is in fact borne out as already illustrated in example (3) of section 1 and repeated here:

- (21) English: Object Shift
- a. Kima called up [friends of McNulty].
 - b. Kima called [friends of McNulty]₁ up t₁.
 - c. [Them]₁, Kima called t₁ up t₁?
 - d. (?) [Which person]₁ did Kima call up [friends of t₁]?
 - e. (?) [Which person]₂ did Mary call [friends of t₂]₁ up t₁?

If, as argued by Johnson (1991) and Boeckx and Hornstein (2005), the object [friends of McNulty] undergoes object shift to a higher position (21-b) (presumably the specifier of vP), what the example in (21-e) shows is that subextraction from that position after movement is allowed at least for some speakers.

Object shift is also found in Spanish as argued by Gallego (2013). He notes that if the verb escapes the vP and undergoes movement to a higher head, the shift is allowed. This can be tested with the following paradigm in which the main verb moves to T:⁸

- (22) Spanish: Object Shift
- a. McNulty no vio fotos de Baltimore
McNulty not saw pictures of Baltimore
'McNulty did not see pictures of Baltimore'
 - b. No vio [fotos de Baltimore]₁ McNulty t₁
Not saw pictures of Baltimore McNulty
 - c. [Qué]₁ no vio (t₁) McNulty (t₁)?
what not saw McNulty
'What didn't McNulty see?'
 - d. [De qué ciudad]₂ no vio [fotos t₂]₁ McNulty t₁?
Of which city not see pictures McNulty
'Which city didn't McNulty see pictures of?'

Example (22-a) is the baseline with the object in-situ, and (22-b) is the object-shifted counterpart: the DP has undergone movement to the specifier of vP (see Gallego 2013 for more details). We know, from subsection 2.1, that subextraction is possible if the object remains in its base position, but these data indicate that such an operation is licit if the object has abandoned its base position as well (22-d).

Object shift in Scandinavian languages is only restricted to pronouns with the exception of Icelandic (Holmberg 1986, Vikner 2005). This makes it impossible to construct equivalent examples to the ones provided for Spanish. However, Vikner (2005: 397-399) indicates that V-Particle constructions in Norwegian behave the same way as V-Particle constructions behave in English: the particle can either precede a full DP complement or follow it. If full-DP object shift in Norwegian patterns like English or Spanish, we predict that subextraction should be possible. Unfortunately Vikner (2005) does not provide with the relevant examples and, at the moment, my informants have not provided me with the relevant Norwegian data.

Up until this point, the data have served to support the claim that extraction out of a moved object is allowed. All the examples that have been presented focus on object shift. One could wonder whether these extraction facts are simply particular to this phenomenon. Nevertheless, data from German refute this hypothesis given that German lacks object shift but displays scrambling. What is more, as Abels (2007) notes, it is possible that scrambling feeds *wh*-movement or other A'-operations. Abels (2007) credits Beck (1996: 6-7), Sauerland (1996), Wiltschko (1997) and Fanselow (2001) for the original observations. An argument for this is the lack of cross-over effects:⁹

- (23) Lack of cross-over effects in German (taken from Abels 2007: 18-19)
- a. Welchen Studenten_i unterstützen seine_i Eltern?
which.ACC student.ACC support his parents?
'which student_i do his_i parents support?'

- b. *Welchen Studenten_i glauben seine_i Eltern, dass Maria unterstützt?
 which.ACC student.ACC believe his parents that Maria supports
 ‘Which student_i believe his_i parents that Maria supports?’
- c. Welchen Studenten_i glaubt Maria, dass seine_i Eltern unterstützen?
 which.ACC student.ACC believes Maria that his parents support
 ‘Which student_i does Maria believe that his_i parents support?’

The logic behind this examples is the following: the *wh*-element in (23-a) and (23-c) must have first undergone scrambling above the possessive in the embedded clause to obviate the cross-over effects and then it must have A'-moved to the matrix clause; on the contrary, (23-b) is ill-formed because either scrambling has crossed a CP boundary; or, because *wh*-movement has occurred first giving rise to cross-over effects. Some examples of extraction and subextraction from scrambled constituents are given below:

(24) German: Scrambling (Stefan Keine p.c.)

- a. Keiner hat sorgfältig ein Buch darüber gelesen.
 No-one has carefully a book about-that read
 ‘Nobody has carefully read a book about that’
- b. Keiner hat [ein Buch darüber]₁ sorgfältig t₁ gelesen
 No-one has a book about-that carefully read
 ‘Nobody has carefully read a book about that’
- c. [Was]₂ hat keiner t₁ sorgfältig t₁ gelesen?
 what has no-one carefully read
 ‘What has nobody carefully read?’
- d. [Worüber]₂ hat keiner [ein Buch t₂]₁ sorgfältig t₁ gelesen?
 about-what has no-one a book carefully read
 ‘What has nobody carefully read a book about?’

The baseline is in (24-a), and the declarative counterpart with scrambling is in (24-b) where the direct object *ein Buch darüber* ‘a book about that’ has moved above the adverb *sorgfältig* ‘carefully’. As pointed out by Abels, *wh*-movement of the scrambled constituent can be assumed to have happened as in (24-c). On top of that, movement out of a scrambled element is also possible as evidenced by (24-d).

Therefore, this subsection has provided substantial cross-linguistic evidence for the possibility of both extraction of and subextraction from a moved constituent, specifically a direct object. The data here suggests that there are few restrictions on extraction and subextraction in the lower clausal domain. In the following section, we move onto the TP domain and discuss what happens when the extraction takes place from higher specifiers.

2.3 The TP domain: *that-trace* effects, lack of *do*-support and their neutralization

The previous sections have been concerned with extraction and subextraction in the lower clausal domain. I am moving into the higher clausal domain, i.e. TP, where we start to see certain restrictions on what can be extracted and subextracted. The focus of this subsection is on *that-trace* effects in the Germanic languages that have been examined so far and lack of *do*-support in English. These facts are interspersed with their neutralized counterparts.

Data from which extraction of objects from embedded clauses has already been presented in subsection 2.1, but as first noted by Perlmutter (1968) there is an asymmetry between these and extraction of the subject out of an embedded clause headed by the complementizer *that*:¹⁰ the latter renders the sentence ungrammatical. An example of this asymmetry is provided below:

- (25) English: That-trace effects (extraction)
- a. *[Who]₁ did they say that t₁ considered running for mayor?
 - b. [What]₁ did they say that a friend of McNulty considered running for t₁?

Assuming that subjects in English have to raise to TP, the original observation is that the complementizer cannot immediately precede the trace or the copy left after the extraction. Many analyses have looked at these effects in isolation, but they have not tried to compare them to what happens if a DP raises to TP, and then a *wh*-element moves from inside the raised DP. Subextraction from a subject in TP is ungrammatical as the following example indicate:

- (26) English: That-trace effects (subextraction)
- *[Which person]₂ did they say that [a friend of t₂]₁ t₁ considered running for mayor?

(26) shows that where extraction is not allowed, neither is subextraction. Besides, this may hint at the idea that it is not the notion of immediate precedence in the sense of linearity that matters: the DP [a friend of] immediately precedes the trace and yet the sentence is unacceptable. Another instance of complementizer-trace effects in English is seen in *tough*-constructions. The relevant pair of examples is in (27):

- (27) That-trace effects in tough-constructions ((27-a) taken from Brillman and Hirsch 2016)
- a. *[Who]₁ is it possible t₁ for t₁ to see Mary?
 - b. *[Which person]₁ is it possible t₁ for [a friend of t₁] to see Mary?

The effects in (25-a) and (26) can be circumvented, as shown by Bresnan (1977), if an adverb intervenes between the complementizer and the trace. Another clear solution to neutralize these effects is omitting the complementizer. A pair of examples is provided below for the relevant

extraction and subextraction patterns:

- (28) That-trace effects neutralization: adverb intervention
- a. [Who]₁ did they say that **at no time / for all intents and purposes** t₁ considered running for mayor?
 - b. [Which person]₂ did they say that **at no time / for all intents and purposes** [a friend of t₂] t₁ considered running for mayor?
- (29) That-trace effects neutralization: complementizer omission
- a. [Who]₁ did they say t₁ considered running for mayor?
 - b. [Which person]₂ did they say [a friend of t₂] t₁ considered running for mayor?

Following on the peculiar properties for extraction and subextraction that are found in the TP domain in English, the lack of *do*-support in subject matrix questions is worth mentioning too. *Do*-support is prohibited in matrix questions in English as (30-a) shows. For subextraction, the presence or absence of *do*-support does not make a difference (30-c) with respect to the acceptability of sentence:

- (30) Lack of *do*-support in matrix questions
- a. *Who did consider running for mayor?
 - b. Who considered running for mayor?
 - c. *[Which person]₁ (did) [a friend of t₁] consider(ed) running for mayor?

With the exception of *do*-support, which is particular to English, similar complementizer-trace effects are observed in German and Norwegian as well. Bayer and Salzmann (2013) explore this phenomenon in German and extract the following generalization: extraction and subextraction from the highest clausal structural position is impossible. Some examples taken from Bayer and Salzmann (2013) include the following:

- (31) German: That-trace effects (Bayer and Salzmann 2013)
- a. ??*[Wem]₁ glaubst du, [dass t₁ schlecht wird?]
 who.DAT think you that nauseous becomes
 ‘Who do you think will become nauseous?’¹¹
 - b. *[Was]₂ glaubst du, [dass [t₂ für Bücher]₁ denn Peter t₁ kaufen würde?]
 what think you that for books PRT Peter buy would
 ‘What kind of books do you think that Peter would buy?’

Assuming that the subject has raised to TP and is therefore the highest specifier, a sentence like (31-a) is unacceptable. An important data point to support their observation is (31-b): here the

direct object has scrambled to a higher position (i.e. TP) above the particle *denn* and thus subextraction is banned. It is possible that the sentence in (31-b) is ruled out on independent grounds. Haider (2017) shows that scrambling of a *wh*-element over a discourse particle like *denn* is degraded. But, for the sake of the argument, if all that matters for a successful extraction and subextraction operation is height, we would then expect not to find *that-trace* effects or restrictions as in (31), if the relevant phrase occupies a lower position or there are some elements that intervene. This prediction is borne out:

- (32) *That-trace* neutralization: intervention (Bayer and Salzmann 2013)
- a. [Wem]₁ glaubst du, [dass beim Busfahren t₁ schlecht wird?]
 who.DAT think you that at.the bus.riding nauseous becomes
 ‘Who do you think will become nauseous during the bus ride?’
 - b. [Wer]₁ glaubst du, [dass (*t₁) 1933 (t₁) in Hamburg (t₁) Bürgermeister war?]
 who.NOM think you that 1933 in Hamburg mayor was
 ‘Who do you think was the mayor in Hamburg in 1933?’

The first example (32-a) illustrates intervention of an adverbial between the trace and the complementizer obviating any *that-trace* effects. Similarly in (32-b), there are at least two positions indicated with traces from which the subject could have been extracted satisfactorily creating the necessary distance from the complementizer.¹²

Turning now to Norwegian, given that there might be multiple subject positions optionally available for the subject to move into (Nylsen 1997, Svenonius 2002, Bentzen 2009), Norwegian is more liberal than other Scandinavian languages when it comes to *that-trace* effects. In addition, Lohndal (2007) notes that there is a lot of inter and intra speaker variation regarding these effects. However, the variety spoken by the informants analyzed here shows instances of the effects. Again, just like it was done in subsection 2.2 adverbs and negation can be used to control for the position of the subject in the clause:

- (33) Norwegian: *that-trace* effects
- a. *[Hva]₁ vil du [at t₁ lage en skandale?]
 what want you that cause a scandal
 ‘What do you want to not cause a scandal?’
 - b. *[Om hvilket tema]₁ vil du [at [forelesningen t₁] ikke lage en skandale?]
 about which topic want you that lecture not cause a scandal
 ‘About which topic do you want the lecture to not cause a scandal?’

The sentences in (33) are marked with a star which suggests that the subject has left a trace adjacent to the complementizer. In the subextraction case in (33-b), this is clear considering that the stranded subject appears above negation. Although we may have our doubts as to what the exact position of

the subject is in (33-a), Lohndal (2007: 64) reports, citing Holmberg (2000), that such sentences are improved when negation is inserted.¹³ Holmberg (2000) analyzes the contrast between (33-a) and (34-a) as an instance of the *adverb effect* shown for English: the only difference is that in Norwegian, the intervening adverb *tydeligvis* 'evidently' moves to Spec-TP and behaves as an expletive. Since the adverb occupies the highest position in the TP there is no need for the subject to undergo A-movement to that position which makes extraction possible. Thus, once again we might expect that if the subject remains lower instead of raising to TP such effects are not encountered:

(34) *That-trace* neutralization: intervention

- a. [Hva]₁ vil du [at tydeligvis ikke t₁ lage en skandale?]
 what want you that evidently not cause a scandal
 'What do you want to evidently not cause a scandal?'
- b. [Om hvilket tema]₁ vil du [at tydeligvis ikke [forelesningen t₁] lage en
 about which topic want you that evidently not lecture cause a
 skandale?]
 scandal
 'About which topic do you want to evidently not cause a scandal?'

Nothing in this section has been said about Spanish. The main reason is because in Spanish, as in other null subject languages like Italian, it is obligatory to leave the subject in-situ (Rizzi 1982, Torrego 1984 a.o.) which entails that no *that-trace* effects are evidenced. Failure to do so would result in automatic ungrammaticality. Just for the purpose of illustration, here are some sentences in which the subject has undergone movement to a higher position, as indicated by the stranded quantifier *todos* 'all':

(35) Spanish: raising of subject and lack of VS inversion

- a. *[Quiénes]₂ [todos t₂]₁ causaron el año pasado t₁ un escándalo?
 who.PL all.PL caused the year past a scandal
 'Who caused a scandal last year?'
- b. *[Quiénes] querías que [todos t₂]₁ causaran el año pasado t₁ un escándalo?
 who.PL wanted.you that all.PL cause the year past a scandal
 'Who did you want to cause a scandal last year?'
- c. *[De qué equipo]₂ [todos entrenadores t₂]₁ causaron el año pasado t₁ un
 of what team all coaches caused the year past a
 escándalo?
 scandal
 'Of which team did all coaches cause a scandal last year?'
- d. *[De qué equipo]₂ querías que [todos entrenadores t₂]₁ causaran el año
 of what team wanted.you that all coaches caused the year
 pasado t₁ un escándalo?
 last a scandal

‘Of which team did you want all coaches to cause a scandal last year?’

The examples in (35) demonstrate that regardless of whether it is a matrix or an embedded clause extraction and subextraction from the derived subject position is illicit. If these are compared to the counterparts in which the subject remains low in (12), a clear difference in acceptability is obtained.¹⁴

This section has provided more evidence for a parallel behavior of extraction and subextraction. This time in the TP domain concerning *that-trace* and how they can be neutralized. In the case of English, the absence and presence of *do*-support was overviewed as well. Spanish, as illustrated in the paragraphs above, does not show these effects. However, it shares the commonality with the other languages that extraction and subextraction are not allowed when the subject is sitting high in the clause. The next subsection looks at the status of Indirect Objects and concludes the presentation of the data.

2.4 Indirect Objects

In this final section of the data presentation, we turn back to the lower clausal domain to look at indirect objects.¹⁵ The actual position indirect objects occupy is subject to a lot of debate and this paper contributes to that debate in later sections where some arguments are given to support a particular structural representation. However, let’s assume now that they occupy a specifier position as originally captured by the *CED* (Huang 1982).

English (at least some speakers) allows *wh*-extraction of the indirect object in Double Object Constructions but it does not allow to subextract from them.¹⁶ This contrast is given in (36):

- (36) English: DOC
- a. [Who]₁ did McNulty show t₁ a picture of Baltimore?
 - b. *[Which person]₁ did McNulty show [a friend of t₁] a picture of Baltimore?

This is not an isolated fact about English. Spanish also exhibits this same contrast in which subextraction is not allowed even if extraction of the whole indirect object is. The relevant paradigm is in (37), where again subextraction in (37-b) is very degraded compared to the extraction counterpart in (37-a):

- (37) Spanish: DOC
- a. [A quién]₁ le ha mostrado McNulty t₁ una foto de Baltimore?
to who CL has shown McNulty a picture of Baltimore
‘Who has McNulty shown a picture of Baltimore?’
 - b. *[De quién]₁ le ha mostrado McNulty [a un amigo t₁] una foto de Baltimore?
of who CL has shown McNulty to a friend a picture of Baltimore

‘Which person has McNulty shown a friend of a picture of Baltimore?’

The asymmetry here presented can be found in German (38) and Norwegian (39) as well. In both sets of examples, the same format as above is followed: first the the *wh*-extraction of the DP is provided in (a), followed by subextraction in (b):

(38) German: DOC

- a. [Was]₁ hat man t₁ einen Preis gegeben?
 what has one.NOM a prize.ACC given
 Intended: ‘What has someone given a prize?’
- b. *[Worüber]₁ hat man [einem Buch t₁] einen Preis gegeben?¹⁷
 about-what has one.NOM a book.DAT a prize.ACC given
 Intended: ‘What has someone given a book about a prize?’

(39) Norwegian: DOC (Lohndal 2011: 164)

- a. [Hvem]₁ gav du t₁ pakker til bursdagen?
 who gave you gifts for birthday
 ‘Who did you give gifts for their birthday?’
- b. *[Hva]₁ gav du [t₁ for en fyr] pakker til bursdagen?
 what gave you for a guy gifts for birthday
 ‘What kind of guy did you give gifts for their birthday?’

These data show again that there seems to be a strong correlation between extraction and subextraction and presents a difference with respect to the previous indirect object data: extraction was allowed but subextraction was not. Whatever it is, this difference is to be explained.

The contrast that has been mentioned with respect to extraction and subextraction in double object constructions is interesting because it establishes an asymmetry that had not been encountered in the data so far. That is, it is "special" given that extraction is possible but sub-extraction is not. These facts might pose a challenge to the desire to formulate a non-disjunctive generalization and, more importantly, to a uniform, parsimonious treatment of extraction and subextraction patterns. Nevertheless, in the following sections a generalization is stated based on the data that has been obtained and presented in this section and a subsequent analysis that fits all these patterns is also articulated.

3 Taking Stock: Extraction-Subextraction Generalization

The data regarding extraction and subextraction has been presented, and now it is time that we look at it to extract a generalization that can group all the facts together. A table that summarizes the data is given in Table 1:

Table 1: Extraction and Subextraction domains

		Extraction	
		Yes	No
Subextraction	Yes	In-situ objects In-situ subjects ECM Scrambling & object shift <i>That-trace</i> neutralization	?
	No	Indirect objects	Spec TP <i>That-trace</i> effects

The top left cell of Table 1 gathers the contexts in which extraction and subextraction are possible (i.e. in-situ objects and subjects, ECM subjects, scrambled and shifted objects and neutralized *that-trace* effects) and the bottom right includes the three contexts that allow neither: specifier of TP and complementizer-trace effects (including *tough*-constructions). It also shows that no case where subextraction is possible but extraction is not has been found. I know of no example that would fill this part of the table. However, given that the cross-linguistic data here is limited to 4 languages a question mark has been assigned in case future research is able to find one. At this point, there is good evidence to assume that extraction is a necessary condition for subextraction to take place. However, this cannot be all there is to it given the existence of the bottom left cell: as illustrated by examples (36) to (39), subextraction is not possible even though the extraction condition is met. That said, I propose a generalization concerning Table 1 in (40):

(40) **Extraction-Subextraction generalization**

For subextraction out of a DP to be allowed

- a. extraction of this DP must be possible; and
- b. the DP must not be an Indirect Object.

In addition to this, even though it is not stated in the generalization in (40), I want to emphasize a common factor shared by all the contexts displayed in Table 1 and that will be important for the analysis. Given the description and organization of the data in the previous section, one distributional feature shared by all contexts in which both extraction and subextraction is disallowed is that launching and landing sites for the moving elements are too close: a moved subject in Spec-TP targeting the specifier of probing CP, or an indirect object in the specifier of an Applicative projection targeting the specifier of the immediately higher projection vP. Therefore, we should have a generalization like (41) in mind that will help us develop a more accurate analysis:

(41) **The Freezing Generalization**

If X takes YP as a complement and YP has an element ZP in its specifier, two conditions

must hold:

- a. movement of ZP to the specifier of XP is not allowed unless ZP is an indirect object; and
- b. movement out of ZP to the specifier of XP is not allowed.

The generalization in (41) is stated as a condition on movement operations. It should not be confused with Wexler and Culicover's (1980) or Rizzi's (2006) *Freezing* principle. In fact, (41) should be understood as a descriptive observation of the facts that can be drawn from Table 1 instead of a universal principle that holds in the grammar.

That said, my analysis will try to overcome these problems, capture the data in Table 1 and provide an explanation for the generalizations in (40) and (41). In order to do so, I propose that a revised version of Erlewine's (2016, 2019) antilocality is necessary to account for the very local movement restriction between launching and landing sites. In addition, I will argue that the Principle of Minimal Compliance (PMC Richards 1998) via Agree can debar antilocal relations between a probe and a DP, but not the specifier of that DP. This makes the DP accessible to subsequent syntactic operations with the probe triggering the Agree relation. The details of this proposal are spelled out in the following section.

4 Antilocality + the PMC

The goal of this section is to demonstrate that antilocality should not only restrict movement *of* a DP but also movement *out of* a DP. The working hypothesis is that too short a movement, whatever the movement type is, should be ruled out. However, as it will be shown antilocality is very strong and makes some undesirable predictions for certain contexts like indirect objects. That is why, I propose that if Agree takes place, then antilocality can be suspended giving rise to licit extraction patterns. Before moving to the specific mechanics of the proposal, I am simply anticipating the assumptions that will be motivated through this paper. Each of them will be carefully spelled-out and argued for as this section moves forward:

(42) **Antilocality**

Movement of an XP must cross a projection line.

(43) **Principle of Minimal Compliance**

If a head H with probes $P_{[X],[Y]}\dots$ Agrees with a Goal G in a feature X, G is accessible to subsequent syntactic operations regardless of locality conditions with respect to probes on the same head H.

I will first concentrate on spelling out how we get to the exact definition for antilocality. I will then show how antilocality is able to account for almost every cell of Table 1 but that it has its limits; I will then use this as an opportunity to introduce the PMC. Together, these two conditions will also be argued to make the right predictions in A-movement contexts, such as subject raising in declarative clauses.

4.1 Antilocality

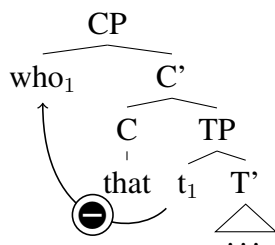
One of the most recent definitions of antilocality has been proposed by Erlewine (2016, 2019) and has been adopted by others including Brillman and Hirsch (2016). The exact definition is provided in (44):

(44) **Erlewine's (2016, 2019) Spec-to-Spec Anti-Locality**

A'-movement of a phrase from the specifier of XP must cross a maximal projection other than XP. Movement from position α to β crosses γ if γ dominates α but does not dominate β .

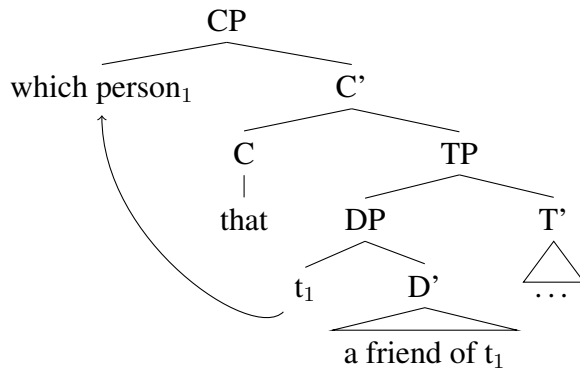
The formulation was originally proposed to account for anti-agreement effects in Kaqchikel, but it has also been motivated by the presence of complementizer-trace effects in English. The idea is that for A'-movement of the subject in an embedded clause to take place, the moving DP must cross at least one maximal projection other than the one immediately dominating the DP. Thus, a sentence like (45) is ungrammatical because, as the tree indicates, no maximal projection other than TP is crossed.

(45) *[Who]₁ did they say that t₁ considered running for mayor?



However, as defined, this makes the prediction that A'-movement out the DP in the specifier of TP should be allowed. Assuming that DPs are phases and that movement has to target their edge for successive cyclicity, the *wh*-element in the specifier of DP would cross a maximal projection other than the dominating DP: TP. And yet the sentence is ungrammatical. This is shown in (46):

(46) *[Which person]₂ did they say that [a friend of t₂]₁ t₁ considered running for mayor?



Therefore, this definition of antilocality is not strict enough because it is not able to account for the asymmetry. One of the challenges to face is, then, the extension of Spec-to-Spec antilocality to Spec-of-Spec-to-Spec Antilocality following the assumption that an element that is extracted out of a DP first moves through the specifier of that DP. Assuming, as we are, that movement of a specifier of a specifier is precluded whenever movement of the specifier is, the standard formulation of antilocality in terms of crossing a "maximal projection" is not enough. The desired result can be achieved if the phrasal node that dominates the specifier "does not count".

A possible way to do this is the following. Let's assume that phrases projected by a particular head share the features inherited by this head via projection. Let's assume further, following Brody (1998), that *projection* should be understood as "XP is a partial copy of X".¹⁸ Therefore, a head X that has projected a phrase XP is (immediately or nonimmediately) dominated by a partial copy. Call the set of all projected occurrences of a head X (assuming bare phrase structure) a *Projection Line (PL)* of X. We can then formulate antilocality as in (47), and what we mean by "crossing a projection line" in (49):

(47) **Antilocality**

Movement of an XP must cross a *Projection Line*.

(48) **Projection Line**

A *PL* of $\alpha = \{x \mid x = \text{a projection of } \alpha\}$

(49) **Crossing a Projection Line**

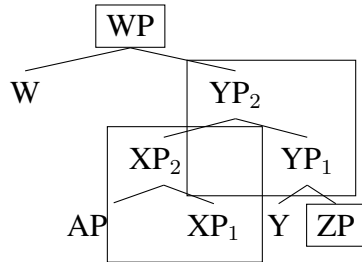
An element α crosses a *Projection Line (PL)* on its way to a position β iff

- a. all members of that *PL* dominate α 's launching site; and
- b. no member of that *PL* dominates β .

The definition in (49) ensures that for movement to satisfy antilocality it is not enough to only cross a maximal projection. In fact, movement of an element is considered to be satisfactory if every single member that has been projected by a head, what we call a *PL* in (48), is left behind

after extraction; and the landing site is outside the domain of the *PL* that has been crossed. This can be schematically represented in the tree in (50) where *PLs* are marked with a square:

(50) Structural Representation of *PLs*

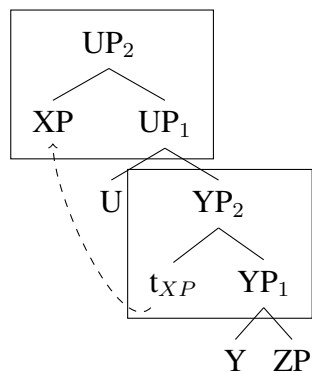


According to the structure in (50), *WP* and *ZP* constitute *PLs* because they are the only projections of the head *W* and *Z* respectively: *PL* of *W* = {*WP*} and the *PL* of *Z* = {*ZP*}. The set of *YPs* ({*YP*₁, *YP*₂}) is the *PL* of *Y* because it encompasses all the members with the label *YP* that have been projected by *Y*. Likewise, the set of *XPs* is also a *PL*. However, one must note that, though the individual *YP* and *XP* participate in the *PL* as members of the set, they do not count as *PLs* on their own.

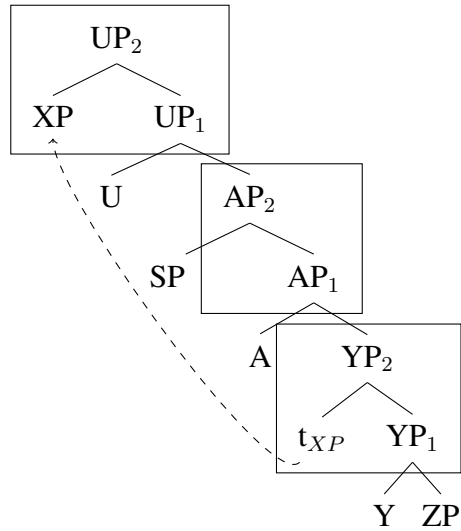
Now, given these formal definitions and our version of antilocality, how does the system make the right predictions for extraction and subextraction? A schematic representation is provided in (51):

(51) Licit vs. Illicit patterns

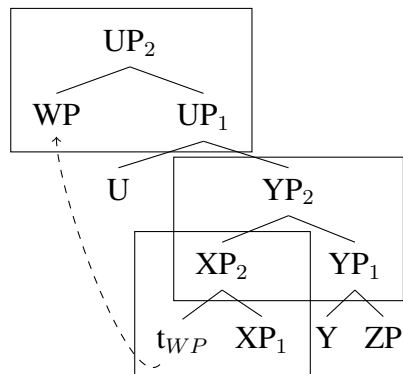
a. *Extraction



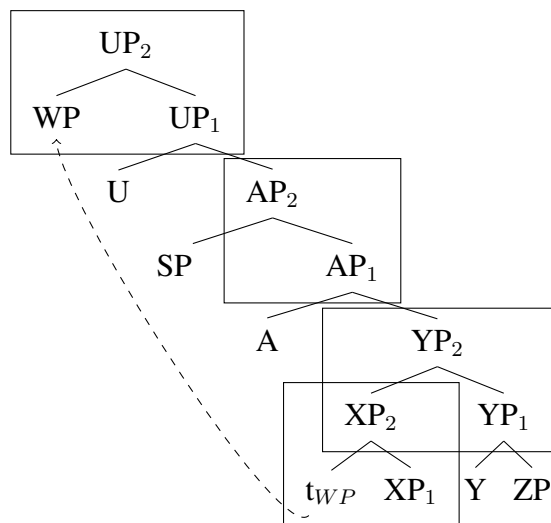
b. Licit Extraction



c. *Subextraction



d. Licit Subextraction



The structure in (51-a) represents an illicit extraction pattern: XP is dominated by a projection of Y (i.e. YP₂) which is dominated by a projection of U (namely, UP₁). The movement of XP only crosses two individual members of independent *PLs*, not a complete set (i.e. YP={YP₁, YP₂} and

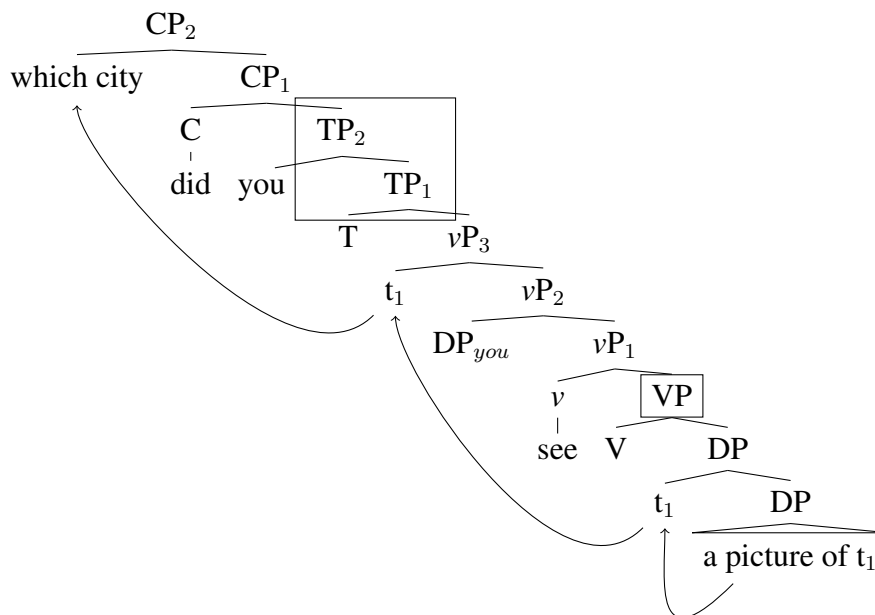
UP={UP₁, UP₂}) which would make them a *PL*. The movement is thus ruled out by antilocality in (47). In other words, as the boxes marking *PLs* show, XP moves from inside one box into another box; the movement does not jump over either of the boxes. On the contrary, (51-b) shows an extraction path that crosses over the *PL* of A: {AP₁, AP₂}. Illicit subextraction in (51-c) is precluded by antilocality because XP₂ is not a *PL* of X. Only the set of XPs constitutes a *PL* of X: {XP₁, XP₂}. The status of the YP and UP that are crossed remains as in (51-a). (51-d) contrasts with (51-c) since WP crosses the set of projections of A on its way to UP.

Given these structures, it now follows that both (45) and (46) are ruled out by antilocality. I will provide with the details behind *that-trace* effects later in this section after spelling out the nitty gritty of in-situ subjects and objects and movement out of moved elements.

4.1.1 In-situ subjects and objects

With these assumptions and structural representations in mind, we can now go on to derive the data from section 2. In the case of in-situ direct objects and subjects, the distance between the launching site of the (sub)extraction and the landing site (intermediate or final) has to be antilocal enough for the movement operation to be possible. We can illustrate this with the examples in (7) for object (sub)extraction and (12) for subjects. Unless stated otherwise, the trees will only show the structure for the subextraction counterparts. The reason behind this is that subextraction and extraction pattern alike with respect to the *PLs* they cross on their way to the landing site as illustrated by the diagrams in (51). The crossed *PLs* are indicated in the tree with a square:

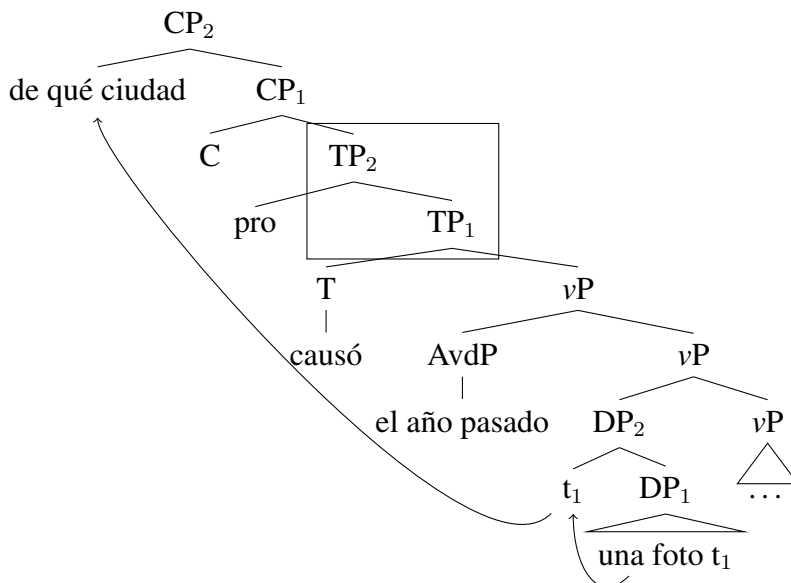
(52) [Which city]₁ did you see [pictures of t₁]?



The *wh*-element in (52) first undergoes movement to the specifier of the DP, under the aforementioned assumption that DP is a phase; and from there it moves to the specifier of *v*P, following the assumption that accusative *v* is a phase and movement has to proceed via its edge (Chomsky 2001, Legate 2003 a.o.). This movement does not violate antilocality in (47) because it crosses the *PL* of *V*. When the probing phase head *C* is merged it attracts the *wh*-element to its specifier. This movement is also compliant with our definition of antilocality because, here, the *wh*-element crosses the *PL* of *T* (i.e. the set of the two *TP*s nodes).

In (53) the subject remains in-situ in the specifier of *v*P. When *C* probes for a potential goal, it finds the *wh*-element *de qué ciudad* 'of which city' and brings it to its specifier. This movement operation does not violate antilocality given that the *wh*-element crosses *T*'s *PL*.

- (53) [De qué ciudad]₁ causó el año pasado [una foto t₁] un escándalo?
 of what city caused the year past a picture a scandal
 'Of which city did a picture cause a scandal last year?'



These two derivations use English and Spanish respectively; but they can be applied to the other languages as well, as long as subjects and objects occupy structurally low positions with respect to where the landing site is located. In the case of German and Norwegian, arguments such as topicalization and the position of negation (and adverbs) respectively were given to support this claim. This is schematically represented in (54):

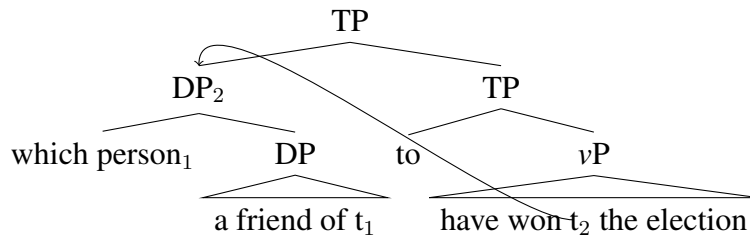
- (54) [_{CP} [_{CP} C [_{TP} T [_{vP} [_{DP1} [_{DP2} DP₂] [_{DP1}...]]] [_{vP} v]]]]]
-

4.1.2 Movement of and out of moved XPs: Object shift and scrambling

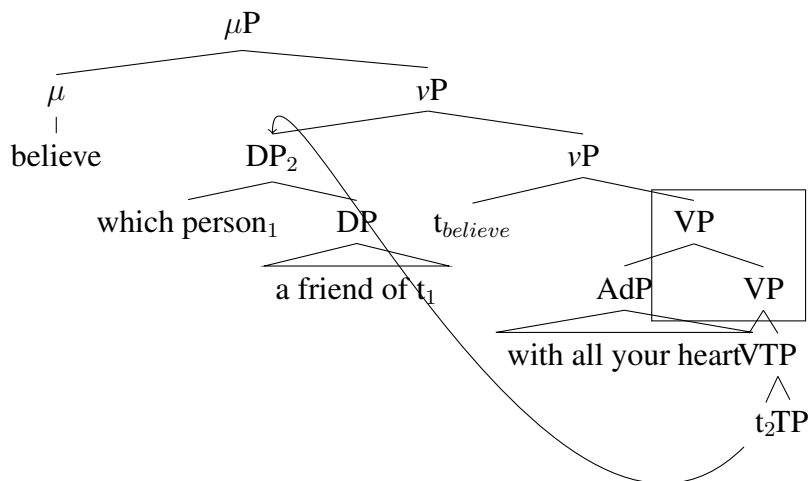
Antilocality as defined in (47) also makes the right predictions for movement out of moved constituents such as ECM and verb-particle constructions. It was shown in (17)-(18) that an XP can surface between the matrix verb and matrix adverbials, and between the matrix verb and its particle. And the examples showed that extraction and subextraction are possible. Thus, sentences like (55) have a structures as in (56):

(55) (?) [Which person]₂ do you believe [a friend of t₂]₁ with all your heart t₁ to have won the election?

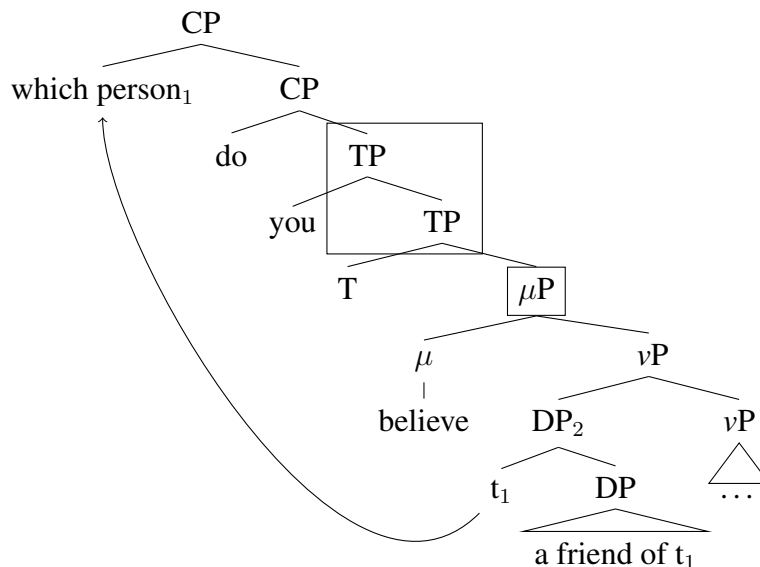
(56) a. Embedded TP



b. Matrix clause: *PL* of V is crossed



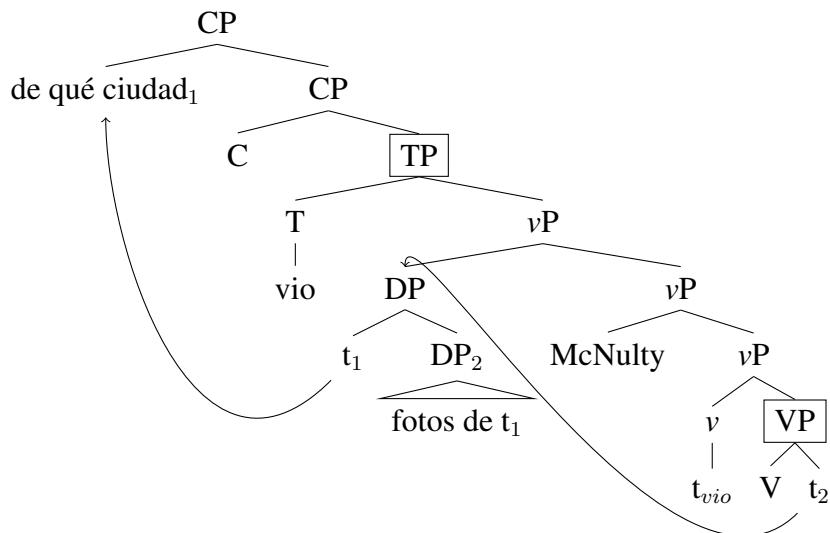
c. Matrix clause: *PL* of μ and T are crossed



The trees in (56) follow Johnson's (1991) analysis in which the verb in English raises to a higher position to accommodate the intervening DP.¹⁹ In this case, the DP [a friend of which person] starts out low in the embedded clause, and raises to the specifier of TP for case that receives from the matrix v (56-a). The next operation in (56-b), which seems to be optional, is movement to the specifier of matrix vP . It is in this position that the subextraction occurs. None of these steps violate antilocality: (i) the DP crosses the PL of embedded v , (ii) it then crosses the PL of matrix V , and (iii) the subextracted wh -element crosses the PL s of μP and TP in (56-c).²⁰ If the optional movement operation had not occurred to the matrix clause, the subextraction would have taken place from the specifier of embedded TP, once the DP receives case. In that scenario, a PL would also be crossed: matrix V 's.

The same logic can be applied to object shift. Here is a Spanish example to illustrate the point with a different language.

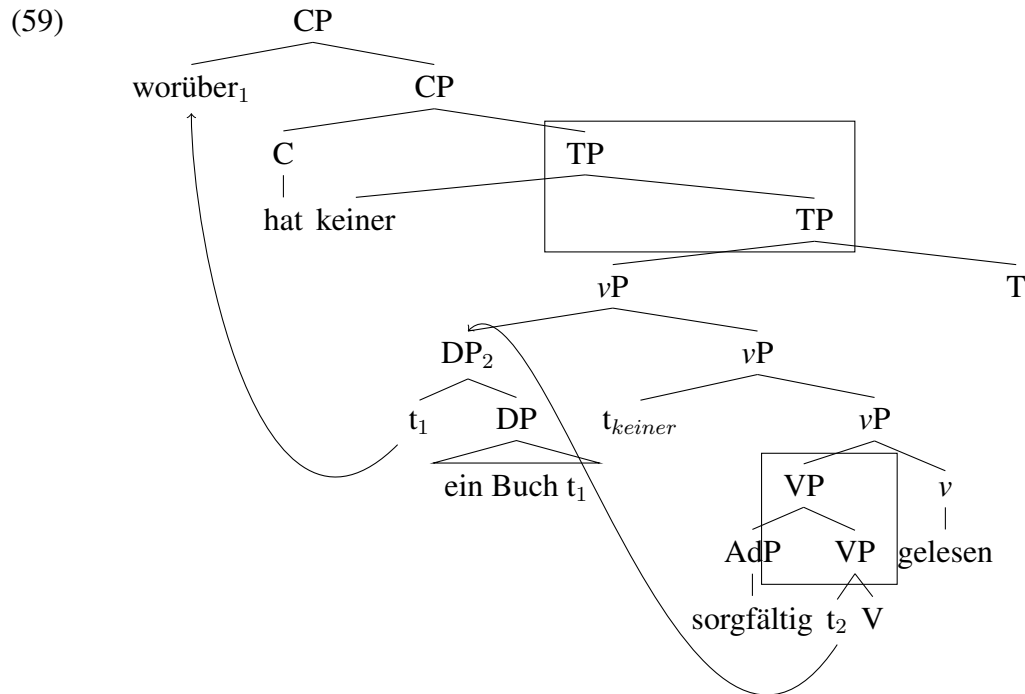
- (57) [De qué ciudad]₂ vio [fotos t₂]₁ McNulty t₁?
 Of which city see pictures McNulty
 'Which city did McNulty see pictures of?'



In this example, the object has shifted to a higher position which, I assume following Gallego (2013), is the edge of the vP . Upon doing so, the DP has crossed the the PL of V . Once there, subextraction takes place: the wh -element lands in the specifier of CP after moving past T's PL .

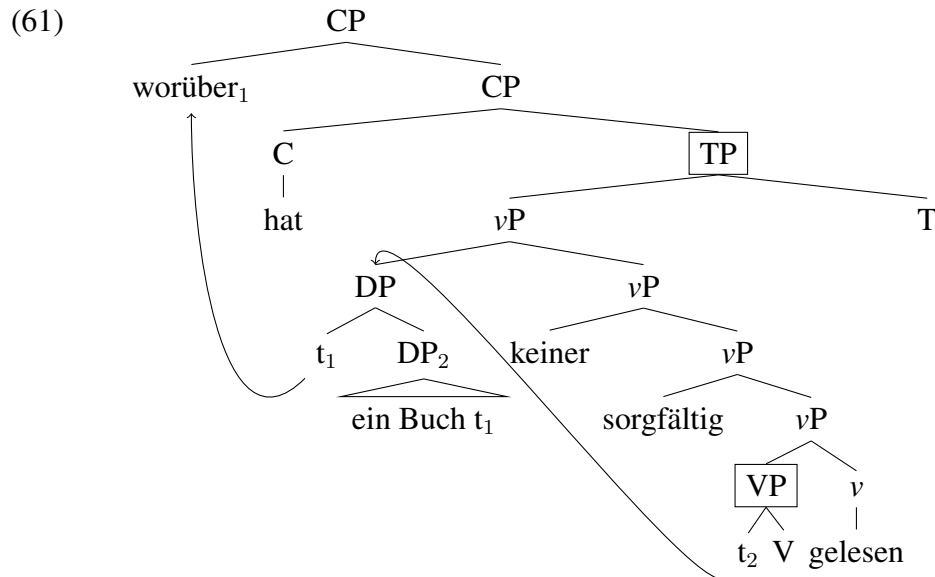
Movement out of moved constituents also encompasses German scrambling. Crucially, as pointed out by Müller (1998: 11) if the scrambling takes places low, then subextraction is allowed. On the contrary, *mittelfeld* scrambling targeting the TP is ruled out on independent grounds (see Haider 2017: 52-54), so wh -subextraction from the scrambled constituent is not testable without the possibility of getting into a confound. Low scrambling is illustrated by (58) whose structural representation is in (59):

- (58) [Worüber]₂ hat keiner [ein Buch t_2]₁ sorgfältig t_1 gelesen ?
 about-what has no-one a book carefully read
 'What has nobody read a book about carefully?'



Assuming that the adverb lies low, in (59) the subject has moved to TP and the object DP has scrambled to the edge of the vP satisfying antilocality in that first movement. The second movement operation is subextraction which also satisfies antilocality since the movement crosses the set of all TP nodes. Stefan Keine (p.c.) points out that sentences like (60) are also grammatical in German where the DP object has scrambled over the external argument of the verb. A potential representation for such a sentence is given in (61):

- (60) [Worüber]₂ hat [ein Buch t₂]₁ keiner sorgfältig t₁ gelesen ?
 about-what has a book no-one carefully read
 'What has nobody read a book about carefully?'




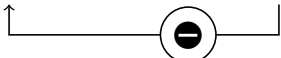
In (60) represented in (61), the subject can be in-situ (Haider 1990) and so scrambling of the object DP can have taken place to the edge of the v . Thus, the DP has undergone low scrambling which enables its inside material to be accessible to subsequent syntactic operations. In fact, subextraction of *worüber* ‘about that’ does not violate antilocality given that the set of all TP nodes is crossed with this movement.

4.1.3 XPs in TP: *That-trace* effects, their neutralization and *do*-support

The scrambling cases have served as a good transition to indicate what happens when the landing and launching sites for *wh*-movement are too close together: antilocality in (47) is violated given that the movement does not meet the distance requirements. This brings us into the extraction and subextraction restrictions illustrated for the TP domain. For English, German and Norwegian these restrictions included *that-trace* effects in embedded clauses. In the case of Spanish, these effects are not present and the extraction and subextraction facts from matrix and embedded TPs can be due to other factors such as the lack of V-S inversion.

The classic problem with *that-trace* effects is that the complementizer cannot immediately precede a copy or a trace of the "subject" or, as German shows (Bayer and Salzmann 2013), the highest XP in the TP domain. If we think about the distance that the highest XP in TP has to cover to get to CP, we note that this is very short. In fact, if we apply antilocality as defined in this paper we can remove the notion of "immediately precedes" and provide an adequate explanation for the phenomenon: the movement is simply too short and does not cross a *PL*. Moreover, the notion of "immediately precedes" is not very appropriate given that subextraction is also subject to these effects; and, yet the complementizer does not immediately precede the launching site. In fact, this

was the main motivation to modify previous versions of antilocality. We can illustrate this again with the schema in (62-a) for extraction and (62-b) for subextraction:

- (62) a. $[_{CP} wh_{-1} [_{CP} \text{that} [_{TP} t_1 [_{TP} T \dots]]]]$

- b. $[_{CP} wh_{-1} [_{CP} \text{that} [_{TP} [_{DP} t_1 [_{DP} \text{a friend of } t_1]]] [_{TP} T \dots]]]]$


As the schematic representations show, the movement of the *wh*-element from spec-TP to spec-CP only crosses individual members of two different *PLs*: the topmost TP and the lower CP in (62-a); and the higher DP, the topmost TP and the lower CP in (62-b). Therefore, there is an antilocality violation that precludes both extraction and subextraction in this syntactic domain. This same schema can be used for German and Norwegian as well (see examples (31) and (33) respectively) for those cases in which the launching site of the movement is preceded by the complementizer.

The same logic can be applied to the ban on extraction and subextraction in *tough*-constructions in English such as (27) repeated here as (63). The traditional analysis, since Chomsky (1981), is one such that *who* never receives a case feature and so the derivation crashes. However, as brought into attention by Brillman and Hirsch (2016), this cannot be the reason why (63) is ungrammatical because the DP/*who* receives case from the complementizer:

- (63) a. *[Who]₁ is it possible t₁ for t₁ to see Mary?
 b. *[Which person]₁ is it possible t₁ for [a friend of t₁] to see Mary?

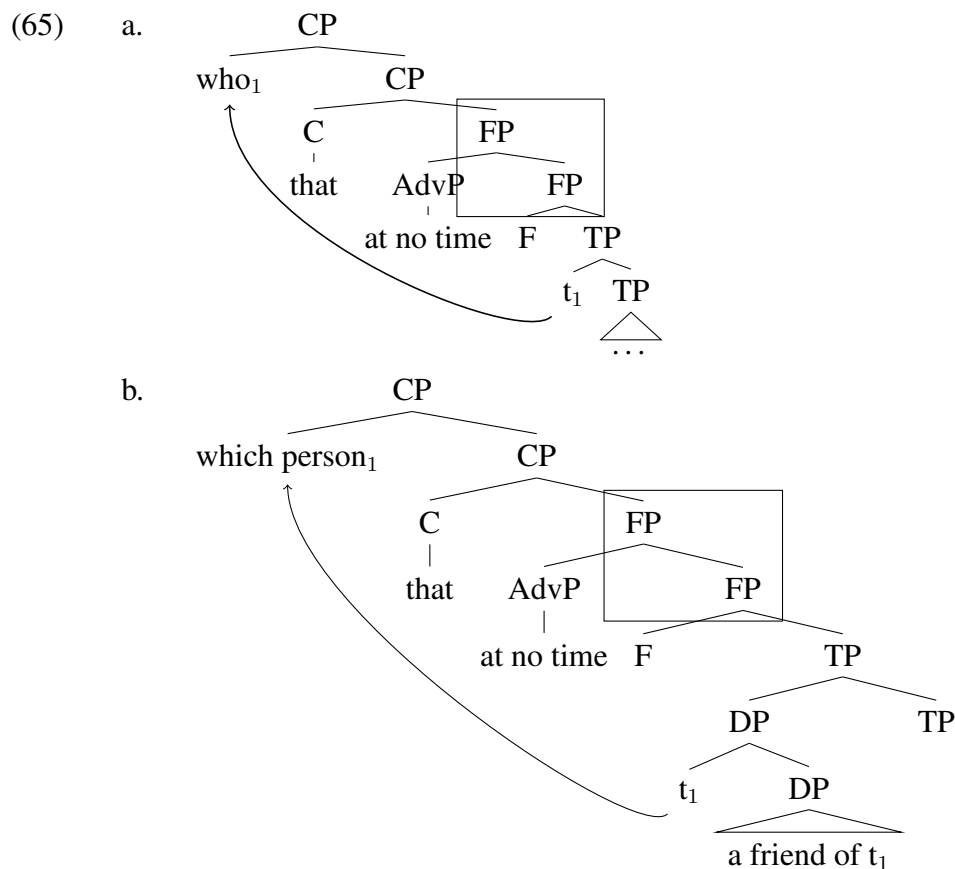
Thus, following Brillman and Hirsch (2016), I propose that these sentences are precluded for the same reason *that-trace* effects are: the derivation involves a first movement step from the specifier of embedded TP to the closest CP. And this violates antilocality because no *PL* is crossed. The schematic representations in (62) can be used to illustrate the illicit movement step in both extraction and subextraction cases if we substitute *that* with *for*.

If it is the case that very short movement is what is causing *that-trace* effects to arise here, our account should also predict why these effects are neutralized whenever high adverbs intervene. In fact, this prediction is borne out:

- (64) a. [Who]₁ did they say that **at no time / for all intents and purposes** t₁ considered running for mayor?
 b. [Which person]₂ did they say that **at no time / for all intents and purposes** [a friend of t₂] t₁ considered running for mayor?

For an antilocality account like the one that is being proposed here, it is crucial that the position of

these adverbs is not in the TP domain. Otherwise, (i) either antilocality is not the right approach or (ii) if it is, we should never expect to find these neutralizations. It is not the case that (ii) holds, and considering the arguments that have been given in favor of this antilocality approach, let's reject (i). In fact, it has been argued in the literature (Culicover 1993, Watanabe 1993, Browning 1996 a.o.) that these cases of neutralization by intervention involve an extended CP or an extra layer between CP and TP (Culicover's 1992 PolP for instance). What the actual label of the projection might be is not as relevant for our purposes as the fact that there is an intervening projection between CP and TP which is not part of TP. That said, the sentences in (64) can be represented as below:



Given the representation we have assumed for the sentences in (65), the movement operation marked by the arrow does not violate antilocality. The intervening functional head projecting the FP that hosts the adverb in its specifier serves as the required *PL* to be crossed, allowing extraction and subextraction to occur. Thus, the neutralization follows. This should not be taken to be only applicable to *that* but to other complementizers. It is not the case, however, that *tough*-constructions can be rescued this way because no maximal projection can be placed between *for* and TP. The same prediction is borne out for German and Norwegian (see examples (32) and (34) in which the launching site is distal enough from the embedded spec-CP).

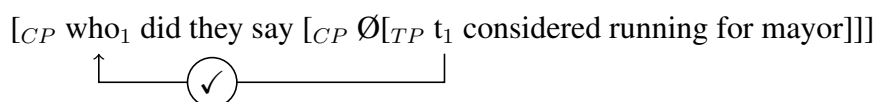
It was also noted in subsection 2.3 that complementizer-trace effects could be neutralized if the

complementizer is omitted: I propose that subject movement in *that*-less clauses follows a different (sub)extraction path than movement of the subject in (64). The (sub)extracted element does not land in the specifier of the embedded CP and therefore Antilocality is not violated. A possible analysis of *that*-less clauses that I adopt is one in which a *that*-less embedded clause is not a CP but a TP, as proposed by Bošković (1997b). If there is no CP layer to begin with, then there should not be movement to its edge. In this case, the *wh*-element moves directly from Spec-TP to the matrix clause. An argument for the lack of a CP layer in such contexts is provided by Bošković (2016: 16, fn. 15): if *that*-less clauses are in fact TPs, topicalization inside the clause should not be allowed given that such an operation targets the left periphery of CP. This prediction is borne out:

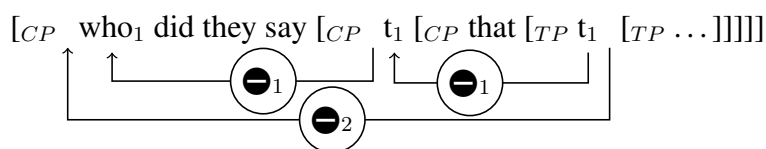
- (66) a. John didn't believe that [Mary]₁ Bill kissed t₁.
 b. *John didn't believe [Mary]₁ Bill kissed t₁.

Thus, if there is no CP layer, then the movement chain can be represented as in (67). The counterpart with the overt complementizer is in (68). In the case of the latter, two potential derivations are possible. The first derivation labeled \ominus_1 yields an antilocality violation because no *PL* is crossed in the first movement operation. The second derivation labeled \ominus_2 incurs in a violation of the PIC: if C is a phase head and requires elements to move through its edge before escaping the embedded clause, movement in one fell swoop would entail that the *wh*-element should be trapped in the complement of the CP phase once the next higher phase head is merged. Thus, the PIC (Chomsky 2001) is violated.

- (67) Movement of the subject over a null complementizer: no antilocality violation



- (68) Movement of the subject over an overt complementizer: antilocality or PIC violation

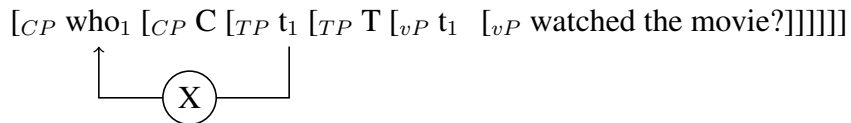


Another possibility would be to adopt Fox and Pesetsky's (2005) Cyclic Linearization as Brillman and Hirsch (2016) and Erlewine (2017, 2019) do. However, Andrew Simpson (p.c.) notes that the combination of Cyclic Linearization and the PIC would result in a very powerful model, whose compatibility is not clear; so it may not be easy to have both in a single analysis.²¹

The last piece of data that needs to be talked about is the absence of *do*-support in matrix questions. This case does not look straightforward if antilocality in (47) is correct: if the *wh*-element

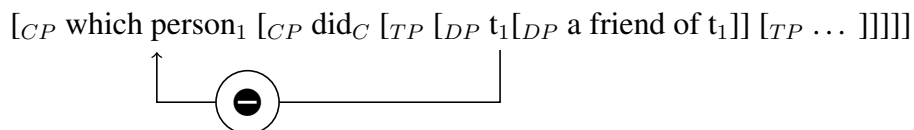
subject undergoes *wh*-movement from spec-TP to spec-CP, no *PL* would be crossed and antilocality would be violated:

(69) Who watched the movie?



In order to avoid an antilocality violation, like the one in (69), I follow Brillman and Hirsch (2016) who propose that *wh*-subjects in matrix clauses raise as far as TP and never to CP. Since English is a residual V2 language, movement to the CP domain is usually accompanied by T-to-C movement. It is not surprising then that if *wh*-subjects do not move all the way to CP, *do*-support does not take place either. The correlation between *do*-support and movement to CP is evidenced by subextraction from the specifier of the subject in TP. If the subextraction from the edge of TP occurs, we expect it to target the CP domain and therefore require *do*-support. This is borne out and illustrated in the schema in (70):

(70) *which person did a friend of watch the movie?



Here in (70) [which person] has undergone movement to CP and thus *do*-support is present. The sentence is, nevertheless, ungrammatical because antilocality is violated. Given the asymmetry between (69) and (70) with respect to *do*-support, it can be concluded that matrix subject questions target the edge of TP. Some other arguments to support the hypothesis that *wh*-subjects remain in TP come from the lack of parasitic gaps in subject questions (Engdahl 1983), and extraction from subject relative clauses (Chung and McCloskey 1983):

(71) Lack of parasitic gaps in subject questions (Brillman and Hirsch 2016, from Engdahl 1983)

- a. *who hired Mary without talking to *pg*?
- b. who did Mary hire without talking to *pg*?

(72) Extraction from subject relative clauses (Chung and McCloskey 1983)

- a. Isn't that the song which Paul and Stevie were the only ones [who wanted to record <which>]?
- b. *Isn't that the song which Paul and Stevie were the only ones [who George would let <who> record <which>]?

The parasitic gap data can be explained if only A'-movement is able to license parasitic gaps. According to Nissenbaum (2000), A'-movement creates a host for the adjunct that contains the parasitic gap. However, if in (71-a) *who* does not undergo A'-movement, no host is created so the parasitic gap is not licensed. It also follows that if there is no A'-movement in subject relatives either, the subject relative is not an island and therefore movement out of it can happen. Regarding the relative clause data in (72), Chung and McCloskey (1983) claim that subject relatives are weaker islands than object islands. In fact, the *wh*-complement *which* in (72-a) has been extracted without giving rise to an island-effect. On the contrary, extraction of *which* from a non-subject relative clause results in ungrammaticality. This asymmetry can be explained if *who* is the specifier of TP in (72-a) allowing *wh*-elements to pass over it, but *who* has moved to the specifier of CP in (72-b) creating an island.

However, if the movement happens between Spec-*v*P and Spec-TP, then it should be ruled out by antilocality since no *PL* is crossed. We would have the same configuration we are trying to preclude in (69). I, thus, propose that the Principle of Minimal Compliance can come to the rescue and debar the locality condition between the probes in a head and its goal.

4.2 The Principle of Minimal Compliance

As just noted, such a definition of antilocality precludes other movement operations such as Spec-*v*P to Spec-TP DP movement. This would be captured by the tree in (51): the DP is on the edge of the *v*P, and it targets the edge of the TP by only moving over a node of *v*P and a node of TP, instead of crossing their complete set. If this were the case, this account would be predicting that this movement should never be allowed. The key to this problem lies in the Principle of Minimal Compliance (PMC):

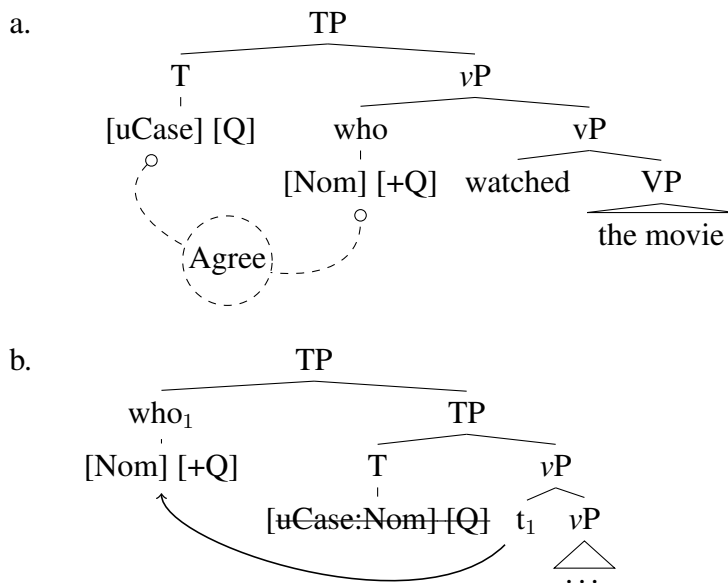
(73) The Principle of Minimal Compliance (PMC)

If a head H with probes $P_{[X],[Y]...}$ Agrees with a Goal G in a feature X, G is accessible to subsequent syntactic operations regardless of locality conditions with respect to probes on H.

This definition of the PMC differs from previous versions proposed by Richards (1998) and Rackowski and Richards (2005). It entails that no matter how close a probe and a goal are, if they enter an Agree relation, the locality restrictions that may exist between them are suspended. That said, we can now explain why Spec-*v*P to Spec-TP movement is allowed for subject matrix questions. For instance, in subject matrix questions, given the hypothesis followed here that *wh*-subjects raise as far as TP, C must not bear the [Q] feature to trigger *wh*-movement of the subject and subsequent *do*-support. Thus, T must be the locus of the [Q] feature. T also bears [uCase:

] that needs to get valued by the [iCase: Nom] in the *wh*-subject. As a result, T can get the the [uCase:] valued and the [Q] feature by entering an Agree relation with the *wh*-element. This is represented in the derivation in (74):

(74) Spec-*vP* to Spec-TP movement



The head T has multiple probes and when looking for a goal it finds *who*. T and *who* match in terms of [Case] features; and, upon establishing this Agree relation, the antilocality restriction is debarred by the PMC (73), allowing the goal to move to its specifier. It is important to note that a sentence like (70) is also ruled out by this PMC+antilocality account. In this example, subextraction of [which person] from the DP in Spec-TP violates antilocality: no *PL* is crossed. And since there is no independent Agree dependency between C and [which person], the movement is banned.²²

The issue of matrix subject questions in the other languages is not as complex. As it has been mentioned several times, German and Norwegian do not require their subjects to move to TP. Thus, in matrix subject questions the *wh*-element is extracted directly from its base-generation position. In the case of Spanish, movement of the subject must take place from the base generated position. The subject does not move to the specifier of TP. A test for this is the formation of a matrix subject question with a stranded quantifier. If a quantifier can be stranded in any position a DP has passed through, we should not be able to find an acceptable sentence in which the stranded quantifier appears in Spec-TP and the remnant has *wh*-moved to CP. But we should be able to find a case in which the quantifier is stranded in-situ and the *wh*-element has targeted the CP periphery. This prediction is borne out and we take it as evidence that the movement takes place in one fell swoop:

(75) a. *[Quiénes]₂ [todos t₂]₁ causaron el año pasado t₁ un escándalo?
 who.PL all.PL caused the year past a scandal

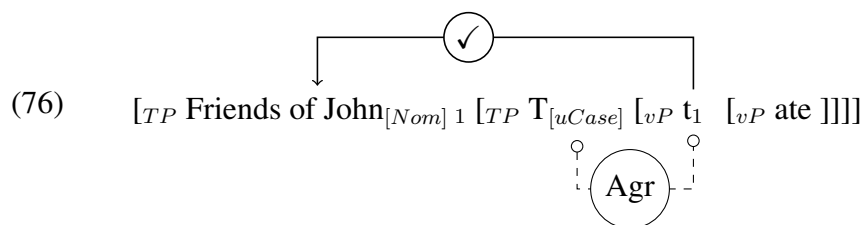
‘Who caused a scandal last year?’

- b. [Quienes]₁ causaron el año pasado [todos t₁] un escándalo?
 who.PL caused the year past all.PL a scandal
 ‘Who caused a scandal last year?’

4.2.1 A desirable consequence of this system

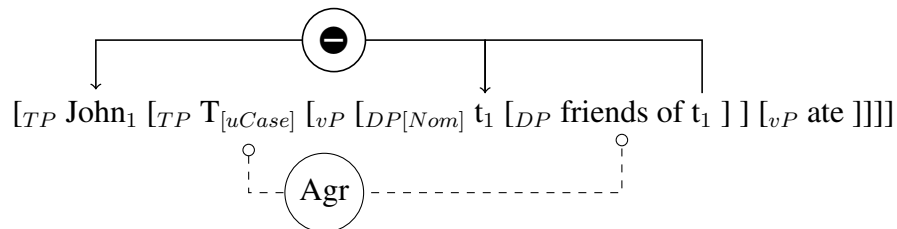
At this point, one desirable consequence of the system that deserves to be emphasized is how this account handles Spec-*v*P-to-Spec-TP A-movement. This is problematic for previous proposals of antilocality including Erlewine (2016, 2019) (and Brillman and Hirsch 2016 following him) that formulate antilocality in a way that it only applies to A'-movement as illustrated in (44). Nevertheless, this is not completely satisfactory because if antilocality is a universal principle of UG that restricts movement, why should A-movement be exempt from it?

One consequence of the account proposed here is that this distinction is no longer necessary because it allows antilocality to be more general: both A and A'-movement are subject to antilocality. This is illustrated in (76) for A-movement of the subject:



As illustrated in (76), T Agrees with the DP *Friends of John* in a case. As a result of this Agree relation, the antilocality restrictions between the probe on T and the DP are eliminated, given the PMC, and the movement is able to occur. This has the advantage that there is no need to stipulate that subject raising in declarative clauses does not violate antilocality because there are invisible functional projections between *v*P and TP that are crossed with this movement. In addition, this antilocality+PMC account predicts that a DP in Spec-*v*P should pattern like indirect objects with respect to extraction and subextraction. Movement of the whole DP to Spec-TP is allowed but subextraction out of it is not. And this is also correct:

- (77) *John friends of watched the movie



The example in (77) portrays a case in which Agree has been successfully established between the

probe in T and the DP goal. This would suspend antilocality between them. However, the DP *John* which has undergone successive cyclic movement through the edge of the containing DP is still subject to antilocality. This is the exact same picture that we will see for indirect objects.

4.3 Indirect Objects: why the asymmetry?

Up until this point, I have provided an explanation for the correlation between extraction and subextraction that exists (at least) in the four languages that are being examined here. It has been argued that antilocality makes the right predictions when launching and landing sites are too close to each other. However, this restriction can be suspended if Agree takes place first. This was specially relevant for the last piece of data regarding the lack of *do*-support. And the interaction of both is of crucial importance to argue why one can extract an indirect object but not subextract from it. But before I get into the "hows and whys", it is crucial that I spell out the structural representation of indirect objects that is followed here.

4.3.1 Against the low applicative hypothesis

I argue following Georgala, Paul, and Whitman (2008), Georgala (2010) and Larson (2010) that English, German, Spanish and Norwegian do not have a low applicative (contra Pylkkänen 2002). In fact, the applicative argument in these languages is higher than the VP.

Pylkkänen (2002) hypothesizes that there are 2 types of Applicatives: a high applicative and a low applicative. She identifies three(-four) main diagnostics to differentiate high and low applicatives but I will only concentrate on (78-a) and (78-b):²³

- (78) Diagnostics for the position of applicative arguments
- a. The theta role of the applied argument.
 - b. Compatibility with static verbs.
 - c. Compatibility with depictive modification.
 - d. Transitivity restrictions on unergatives.

According to Pylkkänen (2002), in double object constructions in languages like English, the applicative only bears a transfer of possession relation to the direct object, which translates in the applicative being merged lower than the voice head or *v* head. The theta role of the applied argument is that of "recipient", never "beneficiary". However, one can easily find examples in which the applied argument is the beneficiary of the event:

- (79) English
Jane did Bill a favor.

Table 2: Pylkkänen's (2002) tests for Applicatives

<i>Test</i>	<i>High Appl</i> (Albanian)	<i>Low Appl</i> (English)
1. The applied argument is the "beneficiary" of the event	YES	NO
1'. The applied argument is the "recipient"	NO	YES
2. Static verbs can be applicativized	YES	NO

(80) Spanish

- a. Carmen le cocinó el pastel de cumpleaños a su hermano.
 Carmen CL baked the cake of birthday (to) her brother
 Intended: 'Carmen baked her brother the birthday cake so that he did not have to it'

(81) German (Georgala 2010: 99, originally from McIntyre 2006)

- a. Er klopfte und sie machte ihm (die Tür) auf
 he.NOM knocked and she.NOM made him.DAT the door.ACC open
 'He knocked and she opened the door for him.'
- b. Sie haben mir das Leben kaputtgemacht
 they.NOM have me.DAT the life.ACC ruined
 'I had them ruin my life.'

In (79), "Mary" is not the recipient but the beneficiary of John's doing the favor. In (80-a), "her brother" is the beneficiary of the event of Carmen's cooking: he benefits from Carmen's baking so that he did not need to do it himself. The German cases are examples of applied arguments as beneficiary and maleficiary respectively. This observation is in line with Larson's (2010) criticism of Pylkkänen's (2002) semantics for low applicatives. For her, low applicatives must combine earlier in the derivation to establish a local relation with the direct object, while only high applicatives can combine higher in the structure because they require access to the event variable of the verb. Larson notes a problem with this semantics, illustrated by the following pair with two conjoined structures in which the same entity "letter" is involved:

(82) Larson's (2010: 702) semantics for (82-a) and (82-b).

- a. John wrote that letter and Bill gave Mary that letter.
 b. John wrote Mary that letter.
 c. $\exists e[\text{writing}(e) \ \& \ \text{Agent}(e, \text{John}) \ \& \ \text{Theme}(e, \text{that-letter})] \ \& \ \exists e' [\text{giving}(e') \ \& \ \text{Agent}(e', \text{Bill}) \ \& \ \text{Theme}(e', \text{that-letter}) \ \& \ \text{Goal}(e', \text{Mary})]$

- d. $\exists e[\text{writing}(e) \ \& \ \text{Agent}(e,\text{John}) \ \& \ \text{Theme}(e,\text{that-letter}) \ \& \ \text{Goal}(e,\text{Mary})]$

According to this semantics, "the writing of the letter by John" and "that letter coming into the possession of Mary" does not entail that "John wrote (addressed) the letter to Mary". This is captured by associating Mary to the giving event e' , instead of the writing event e . However, Pylkkänen's semantics would contain the undesired entailment. The denotation for the low applicative in Pylkkänen's (2002) dissertation is in (83). $\llbracket \text{Appl}_{Low} \rrbracket$ takes three arguments: the direct object, the indirect object and the verb.

- (83) Low Applicative denotation (Pylkkänen 2002)

$\llbracket \text{Appl}_{Low} \rrbracket = \lambda x.\lambda y.\lambda f.\lambda e. f(e,x) \ \& \ \text{theme}(e,x) \ \& \ \text{to-the-possession}(x,y)$

- (84) Pylkkänen's (2002) semantics (taken from Larson 2010: 702)

- a. $\exists e[\text{writing}(e) \ \& \ \text{Agent}(e,\text{John}) \ \& \ \text{Theme}(e,\text{that-letter})] \ \& \ \exists e' [\text{giving}(e') \ \& \ \text{Agent}(e',\text{Bill}) \ \& \ \text{Theme}(e',\text{that-letter}) \ \& \ \text{to-the-possession-of}(\text{that-letter},\text{Mary})]$
 b. $\exists e[\text{writing}(e) \ \& \ \text{Agent}(e,\text{John}) \ \& \ \text{Theme}(e,\text{that-letter}) \ \& \ \text{to-the-possession-of}(\text{that-letter}, \text{Mary})]$

Therefore, this semantic argument for the distinction between high and low applicatives does not seem appropriate.

The second argument for the low applicative hypothesis is based on the observation that static verbs cannot be applicativized. In other words, based on the claim that low applicatives involve a transfer of possession, Pylkkänen (2002) argues that they would make no sense with verbs that are completely static. One example is the verb *hold*: "in an event of *holding*, a bag does not plausibly result in the bag ending up in somebody's possession".

- (85) a. *I hold him the bag.
 b. *I held you the pot.

High applicatives, on the other hand, should have no problem combining with such verbs. But this prediction is also borne out for English, Spanish and German:²⁴

- (86) English

- a. I will hold you the offer.
 b. I will hold you a spot.

- (87) Spanish

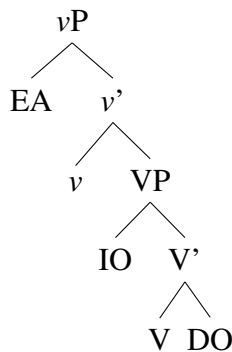
- a. McNulty le mantiene la oferta a Bunk.
 McNulty CL maintains the offer (to) Bunk
 'McNulty holds Bunk the offer'

(88) German (Georgala 2010: 99)

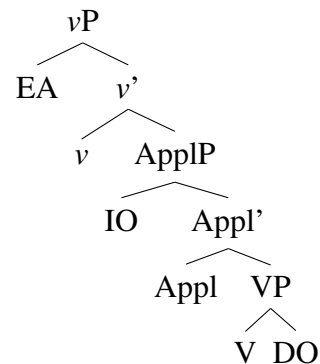
- a. Eva hat Jan den Rucksack zwei Stunden gehalten
 Eva.NOM has Jan.DAT the backpack.ACC two hours.ACC held
 ‘Eva held the backpack for Jan for two hours.’

Based on these data that have been presented here, I propose that the most adequate structural representation is one such that the applicative argument is higher than the VP. This entails that we should adopt a structure like the one proposed by Larson (1988), where the applied argument is introduced in the specifier of V; or a structure à la Marantz (1993), in which there is an ApplP that takes the VP as its complement, and introduces the applied argument in its specifier. For ease of exposition, the two are illustrated in (89) and (90) respectively. Structurally the choice between the two does not have any negative consequences for the extraction and subextraction facts discussed in the remainder of the section. Although data from Norwegian has not been presented, I follow Holmberg, Sheehan, and Van der Wal (2019) who adopt the structure à la Marantz (1993) for Norwegian.

(89) Larson’s (1988) VP shell



(90) Marantz’s (1993) ApplP



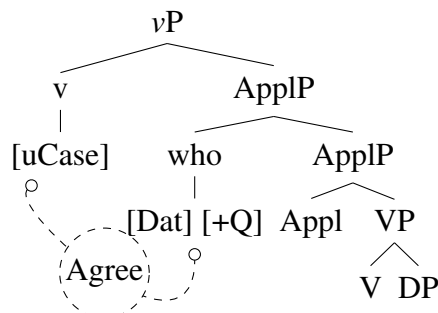
4.3.2 Indirect Objects: the hows and whys of extraction but not subextraction

The second part of the Extraction-Subextraction generalization in (40) makes specific reference to the observation that even if extraction of the indirect object is allowed, subextraction is not. Antilocality, on its own, cannot explain why this asymmetry exists: it would disallow both movement operations (Spec-to-Spec and Spec-of-Spec to Spec) given the proximity of indirect objects, base-generated in Spec-AppIP, and the potential landing site in Spec-*v*P. That is why I have proposed that this asymmetry is obtained via Agree: the probing head *v* Agrees with the DP-indirect object; and, as a result, antilocality between *v* and the DP-indirect object is suspended. This Agree relation has only been established with the DP, which means the following: the DP can move as a whole, but its specifier alone cannot since no Agree dependency exists between *v* and the specifier of that DP. That is, subextraction would still be subject to antilocality.²⁵ This is illustrated in (91)

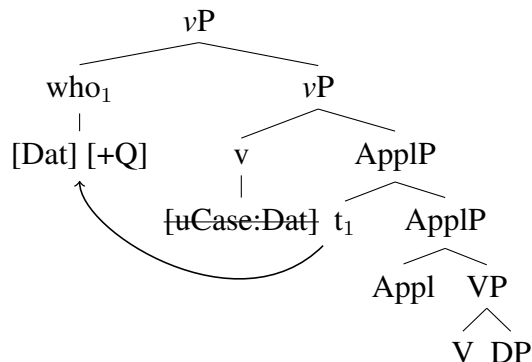
for extraction and (92) for subextraction with data from English.

(91) Who did you show the picture?

a. Agree(v, wh)



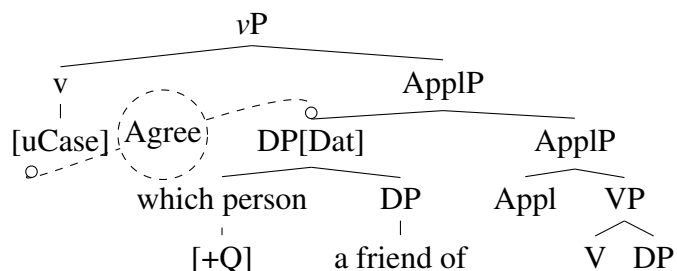
b. Move(wh) to Spec- vP

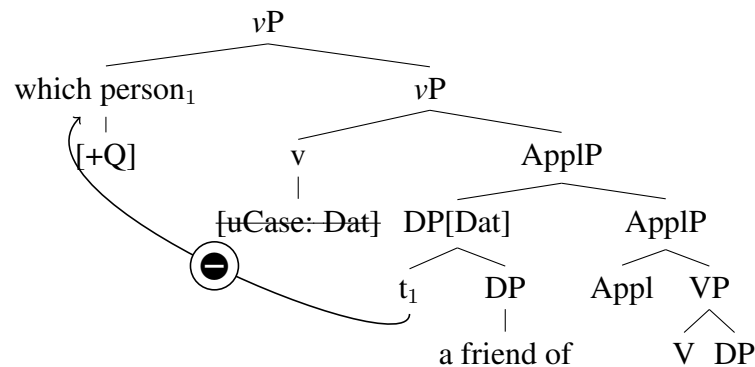


The active v phase head carries a probing feature $[uCase]$. This probe matches the case feature of the wh -element generated in the specifier of the $Appl$ head. The $[Dat]$ feature for case in the indirect object gets valued by the $[uCase:]$ on the v head. Once $Agree(v, wh)$ has been established, the goal is free from locality conditions with respect to the probes on v . The $[+Q]$ feature on the indirect object functions as an EPP feature à la Bošković (2007) triggering successive cyclic movement through phase edges.²⁶ Thus, antilocality is suspended and the movement to the phase edge is allowed. This is how we get the extraction facts.

(92) *Which person did you show a friend of the picture?

a. Agree(v, DP)

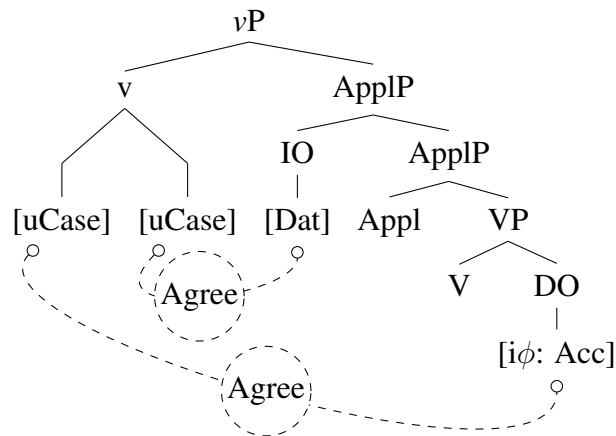


b. Move(*wh*) to Spec-*v*P: Antilocality violation

In (92), the Agree relation between *v* and the DP-indirect object still takes place because *v* needs to get its [uCase] valued with the DP. This entails that if the DP were to move as a whole, antilocality would not be violated. Nevertheless, what (92-b) shows is movement of the specifier of the DP. This movement operation violates antilocality on the basis that no *PL* is crossed on the *wh*-element's way to the phase edge. The PMC does not come to the rescue here because no independent Agree relation has been established between the probes on the *v* head and the *wh*-element.

The same can be applied to Norwegian. Moving onto Spanish and German, it has been argued that as active *v* is not the locus of case for the indirect object (Cuervo 2003 for Spanish, and Georgala 2010 for German) given the unavailability of recipient passives.²⁷ Such proposals entail that the directionality of Agree in these languages is dual: probes can look up (i.e. Agree(Appl, IO)) or down (i.e. Agree(*v*, DO)). Instead, I want to propose that active *v* has "multiple probes" in these languages. With this, I mean that *v* has at least two probes and each of them enters an Agree relation with a different potential goal: in this case, the indirect and direct objects. The feature specification of active *v* for these languages is shown in (93) together with how Agree and antilocality work:

(93) Multiple Probes: Agree (*v*,DO) and Agree (*v*,IO)



The tree in (93) shows that v undergoes two Agree operations triggered by each of its ϕ -probes: one for accusative case with the direct object and one for dative case with the indirect object. The claim that there is Agreement between the phase head and the indirect object can be supported by Preminger's (2019) observation that overt morphological exponents in a DP are the result of a prior Agree relation: in the case of German, this agreement is manifested with overt dative case; in the case of Spanish, the fact that there is agreement is visible with clitic doubling.²⁸

With these assumptions and representations in mind we can now explain the extraction-subextraction asymmetry with respect to indirect objects in Spanish and German in a manner that is parallel to English and Norwegian active sentences:

(94) Step-by-step extraction in IO:

- a. Agree(v , DO) and Agree(v , IO).
- b. Antilocality between v and indirect object is suspended as a result of Agree.
- c. The indirect object can *wh*-move to the v P domain.
- d. The *wh*-element in the specifier of the indirect object cannot move to the higher phase edge: it violates antilocality.

The crucial step here is (94-b). Subextraction is still subject to antilocality on the basis that the Agree relation is established between v and the indirect object as a whole. Thus, only moving the edge of the indirect object violates antilocality.

With this in our toolkit, we can provide an explanation for why A'-movement of the specifier of the indirect object is not allowed: the specifier is still subject to antilocality, and the PMC does not rescue it because it has not entered a prior agree relationship with the local probe. Moreover, the proposal here provides empirical arguments for the preference of a ϕ -Agree relation that involves a c-commanding probe.

5 The generalizations explained

This paper set out to provide an explanation for the following generalizations in (40) and (41) repeated below:

(95) **Extraction-Subextraction generalization**

For subextraction out of a DP to be allowed

- a. extraction of this DP must be possible; and
- b. the DP must not be an Indirect Object.

(96) **The Freezing Generalization**

If X takes YP as a complement and YP has an element ZP in its specifier, two conditions must hold:

- a. movement of ZP to the specifier of XP is not allowed unless ZP is an indirect object; and
- b. movement out of ZP to the specifier of XP is not allowed.

The latter generalization observes that if the launching and landing sites of a movement chain are too close without a lot of intervening material in between, extraction and subextraction should not be possible. That said, I proposed an antilocality constraint as in (47) which made a crucial distinction between the individual members of a projection of X and the set of all members of that projection, i.e. a *Projection Line*: for (any type of) movement to be allowed it must cross a *Projection Line*. Illicit cases of extraction and subextraction can, thus, be understood as violating antilocality. In other words, antilocality has "freezing" effects. However, the PMC can intervene and *thaw* the relevant DP in the form of the Agree. By *thaw*, I mean that the relevant DP is able to move as a whole but extraction out of it is not allowed. As a result of this, what on the surface is an apparent violation of antilocality is, in fact, not because antilocality has been suspended via PMC.

The paper has also looked at applied arguments. One of the challenges for the proposal was the low applicative hypothesis (Pylkkänen 2002). If the hypothesis is correct and the applicative argument is generated lower in the verbal domain, this would be fatal for the proposal put forth here. Nevertheless, I have provided some arguments, building on work by Larson (2010) and Georgala (2010), that have demonstrated that applicative arguments actually occupy a higher position than what originally assumed by Pylkkänen (2002). The fact that indirect objects are extractable but opaque in active sentences has been derived as a consequence of Agree between the [uCase] probes on v and the matching features on the indirect object DP. This Agree operation enabled the indirect object to A'-move. As it has been formulated, the proposal makes some very strong

typological predictions regarding applicative arguments. I am outlining these below:

(97) **Typological Predictions**

- a. If a language has a high applicative, subextraction should, in principle, not be allowed unless independently ruled in.
- b. If a language has a low applicative, subextraction should, in principle, be possible unless independently ruled out.

Answering these questions is desired to test the empirical and conceptual adequacy of the proposal. However, providing an answer to these questions is beyond the scope of the paper. Thus, I leave them for future research.

Therefore, the data and the analysis presented here have served to provide a formal explanation for the "Extraction-Subextraction Generalization" and the "Freezing Generalization". In the next section, before I conclude the paper, I provide some alternative analysis that have been proposed and underline some of their empirical and/or conceptual inadequacies.

6 Some previous alternatives

It was mentioned in the introduction that there have been several approaches concerned with extraction and subextraction facts. Most of these have concentrated in either the extraction part or the subextraction part. However, as I have argued for here, there seems to be a strong correlation between the two. In this section, I intend to assess how some of these approaches make undesirable predictions to account for the data presented in this paper and, more specifically, how they fall short in enhancing our understanding of these phenomena. I will only concentrate on four of these and some subsequent proposals inspired by them: Wexler and Culicover's (1980) *Freezing*, Huang's (1982) *CED*, Erlewine's (2016, 2019) *antilocality*, and Bošković's (2018) *Labels*.

6.1 Wexler and Culicover's (1980) *Freezing* and subsequent formulations

Wexler and Culicover (1980) observe that extraction out of a constituent is not allowed if the constituent has undergone movement or if its base generation order has been modified by a transformational rule. They call this *The Generalized Freezing Principle*. I will only concentrate on the first part here.

Thus formulated, the Generalized Freezing Principle predicts that if a DP moves from its base position to a derived position in a specifier, movement out of that DP should not be possible. However, as it has been shown by the data concerning object shift, scrambling and subextraction

from neutralized *comp-trace* effect environments, this principle is not very appropriate on the basis of empirical evidence. In fact, subextraction is possible in these domains.

In a similar spirit, Takahashi (1994) develops an account of extraction out of a syntactic domain based on two conditions: the Shortest Movement Condition, which states that a moving element must reach its landing site by a series of short successive movement steps along the structural spine; and, the Chain Uniformity Condition, which states that chains should not be modified, but kept uniform. On top of these, he proposes the Uniformity Corollary on Adjunction according to which adjunction to a part of a non-trivial chain is not allowed. Based on these conditions, he proposes that a derivation of (98) should proceed as follows:

- (98) * $[_{CP}$ Whose books₂ do $[_{TP}$ you think $[_{CP}$ t₂ that $[_{TP}$ [t₂ [reviews of t₂]]₁ $[_{VP}$ t₁ caused a scandal]]]]
- a. $[_{CP}$ that $[_{TP}$ [reviews of whose books] $[_{VP}$ [reviews of whose books] caused a scandal]]
- b. $[_{CP}$ that $[_{TP}$ [**whose books** [reviews of **whose books**]] $[_{VP}$ [reviews of whose books] caused a scandal]]

The original constituent [reviews of whose books] and its copy form a uniform chain after the movement step to TP. Extraction of [whose books] takes place from the higher link in the chain (98-a). The extracted element has to step in the outer specifier of the DP given Takahashi's (1994) Shortest Movement Condition (98-b). However, the Uniformity Corollary on Adjunction is violated [whose books] is adjoined to a non-trivial chain: {reviews of whose books, reviews of whose books}.

Although it is true that extraction out of in-situ elements is predicted under this account, it also faces some challenges. First of all, assuming that the element to be extracted undergoes movement to the specifier of the containing DP, movement out of shifted or scrambled objects should be precluded just like it was shown for (98): the shifted/scrambled DP creates a uniform movement chain that is disrupted by the internal movement of the subextracted element in its specifier. Another challenge is concerned with indirect objects. If it is the case, as shown here, that indirect objects do not move, then Takahashi (1994) would predict that extraction out of them would be allowed. Though, it is not.

This is not meant to be an exhaustive list of *Freezing* analyses, but a small sample.²⁹ The point that I would like to make here is that, even though we should not eliminate the notion of "freezing" from syntax, it should not be used as a strong principle in the UG. Instead, we should use it as a descriptive label to formulate generalizations on movement that will help us elaborate a subsequent analysis.

6.2 Huang's (1982) *CED* and its descendants

Huang (1982) made a very influential proposal according to which extraction out of subjects and adjuncts should be ruled out uniformly assuming that they both occupy specifier positions. Huang (1982) called this *Condition on Extraction Domains* which he defined as in (99):

(99) **Condition on Extraction Domains**

A phrase A may be extracted out of a domain B only if B is properly governed.

The notion of government refers to a local licensing of the domain, from which the extraction will take place, by some lexical head. That is, DP complements are properly governed by V, whereas subjects are not licensed this way. The *CED* in (99) makes very strong predictions with respect to extraction from non-complements: it should never be possible or, in other words, extraction out of subject is out whenever extraction out of adjuncts is out (and viceversa).

Although (99) seems to make the right predictions for adjuncts (see Truswell 2007 for licit extraction out of adjuncts, though), the expectation is not always fulfilled by subjects or (non-complements). English seems not to have in-situ subjects in vP , but Spanish, German and Norwegian do; and it has been shown in this paper that extracting out of the in-situ position is allowed. Besides, if after movement, an XP must land in a specifier position (e.g. a non-complement position), (99) will preclude subextraction regardless of where that specifier position is in the structure. This is also very strict as indicated by the availability of moving out of moved elements including derived subject positions in the specifier of TP (i.e. neutralization of *comp-trace* effects).

Uriagereka's (1999) Multiple Spell-Out attempts to maintain the difference in status between complements and specifiers, by developing a theory in which the specifier must be linearized before it is merged with its XP sister. In other words, the subject forms a syntactic compound when it enters the derivation which makes it an island for extraction. The linearization is forced by a requirement at PF motivated by Kayne's (1994) Linear Correspondence Axiom. On the contrary, complements do not need to be linearized before being merged with the relevant head. Again, this has the advantage of treating adjuncts and subject arguments as being governed by the same grammatical principles, and establishes a well known asymmetry between subjects and objects. Nevertheless, just like the *CED*, it is too strict: according to this linearization algorithm, specifiers should always be islands for extraction. This prediction is not borne out.

Therefore, the *CED* and subsequent analysis relying on the subject-adjunct vs. object distinction need to be modified. The account proposed here has looked at this asymmetry from a different perspective, given that adjuncts have been left out of the paper; but it has been able to introduce certain flexibility with respect to extraction out of subjects. Instead of placing the emphasis on the structural fact that a DP is moved to or generated in the specifier of a projection of X, the proposal

has focused on the distance between the launching and landing sites of movement. This has enabled the necessary flexibility to allow extraction out of certain subject DPs, while, at the same time, has maintained the subject-object asymmetry: objects are generated lower than subjects and so, generally, the movement path is longer than from specifier to specifier.

6.3 Erlewine's (2016, 2019) *antilocality*

Many versions of antilocality have been proposed since Bošković (1997a), including but not limited to, Abels (2003), Bošković (2005), and Erlewine's (2016, 2019). Erlewine's, which has already been mentioned in subsection 4.1, is formulated in terms of crossing a maximal projection other than the one immediately dominating the moving element. One of the weaknesses was the inability of extending the Spec-to-Spec movement violation to the Spec-of-Spec-to-Spec movement as illustrated with *that-trace* effects. Moreover, subextraction from other contexts, such as indirect objects would be predicted to be possible: the subextracted element would cross the maximal projection of the containing DP, and the maximal projection of the ApplP that has the DP in its specifier.

Moreover, one advantage over Erlewine's is that the way he formulates antilocality is specifically targeted to A'-movement. However, if antilocality is a constraint that operates on movement in general, why should it be restricted to A'-movement? One advantage of antilocality, as I formulated it here, is that it is expected to hold for both A and A'-movements which others had found challenging to encompass without the need of further stipulations, such as a covert array of functional projections between *vP* and *TP*. Very short movement operations that should, in principle, violate antilocality might not if Agree is involved. This is not new of this paper, as it has already been argued to hold in longer distance extraction contexts to void locality restrictions by Van Urk (2015). The difference with Van Urk's (2015) is that here we are dealing with shorter distance extraction; thus, it makes sense that the antilocality restrictions are also subject to be suspended by the PMC. In fact, if we adopt the definition of antilocality that has been put forth here we can see that the PMC is operative in A-movement domains such as subject raising to *TP* which would otherwise violate antilocality.

The current proposal has been able to bridge this gap by assuming that it is not a maximal projection that needs to be crossed, but a *Projection Line*: the set of all projections of a head. Moreover, I believe that the current proposal is conceptually stronger than those versions of antilocality that rely on "counting nodes". Why should the computational system be counting the number of maximal projections that are crossed to determine if the movement path is licit? If one thinks about Abels's (2003) version, movement of the complement of *X* to the specifier of *X* is precluded because there is no previously unsatisfiable feature that will be valued as a result of this

movement. The proposal argued for here relies on the idea that movement must also be motivated. If we think of a projection as a partial copy of a head X , then movement within the same PL is not driven by the necessity of satisfying a feature, since this should have already been obtained via merge. Also movement from PL to the immediately higher PL does not seem to be feature motivated either. And if it is, then the PMC-Agree will take place to void antilocality and require the XP to become part of the PL of X (e.g. A-movement of the subject). In the case of movement crossing a PL , we observe cases like object shift, ECM raising or *wh*-movement from vP . These movements are driven by the needs of the particular head (i.e. v , T , C) to value their $[uF]$ feature on a DP.

6.4 Bošković's (2016, 2018) labelling approaches

Bošković (2016) develops an analysis to capture a series of *a priori* unrelated phenomena by using antilocality and the labeling algorithm (Chomsky 2013):

(100) **Antilocality**

Movement of A targeting B must cross a projection distinct from B where unlabeled projections do not count.

(101) **Labeling Algorithm** (Chomsky 2013)

When two non-terminal nodes, e.g. XP and YP , are merged, labeling can take place in two different ways:

- a. either one of the non-terminal nodes moves out and so there is no problem for minimal search given that copies/traces do not count for labeling; or
- b. XP and YP agree in terms of features and so they are identical for labeling.

By means of these two mechanism, Bošković is able to derive the *CED* and the Subject Condition. However, his proposal has the weakness of indirect objects. If it is the case, as I have argued here, that indirect objects occupy a position higher than VP , extraction out of them is predicted to be possible.

- (102)
- a. $[_? [_? \text{who} [_{DP} \text{IO}]] [_{AppIP} \text{AppI}^0 \dots [_{VP} \dots]]]$
 - b. $[_{vP} v^0 [_? [_? \text{who} [_{DP} \text{IO}]] [_{AppIP} \text{AppI}^0 \dots [_{VP} \dots]]]]$
 - c. $*[_? \text{who} [_{vP} v^0 [_? [_? <\text{who}> [_{DP} \text{IO}]] [_{AppIP} \text{AppI}^0 \dots [_{VP} \dots]]]]]$

Movement of *who* to the specifier of the DP in (102-a) creates an unlabeled constituent. When the ? indirect object merges with the $AppIP$, this also results in an unlabeled structure since there is no feature agreement between $AppI^0$ and its specifier. When v is merged, the structure is as in (102-b). If labeling via (101-b) has failed, the only way the ? constituents can receive a label is

via movement (101-a). Movement of *who* to the specifier of *vP* violates antilocality in (100) since it does not cross a maximal projection that is not unlabeled.

In a similar fashion, Bošković (2018) provides an attempt to derive the subject condition and other phenomena only from the labelling algorithm proposed in (101)

Bošković (2018) notes that most of the literature concerned with the subject condition has placed the ungrammaticality in the subextraction step of the derivation. However, he proposes to shift the perspective on the issue. Assuming that DPs are phases, and that movement out of a phase has to target its edge first, Bošković puts forth the following assumption: when an element moves to the edge of a phase as a result of successive cyclic movement, the result is an unlabeled projection as in (103).

(103) $[_{DP} \text{ a friend of which person }] \Rightarrow [? \text{ which person}_1 [_{DP} \text{ a friend of } t_1]]$

That said, Bošković (2018) hypothesizes that only phases can move. But a necessary condition for an XP to be a phase is that the phrase be labeled. Therefore, the resulting ?P in (103) is not a phase, *ergo* it cannot undergo any movement. In other words, he reverses the concept of *Freezing* to the prohibition of the first movement step.

This approach has to face the challenges of extraction out of a shifted or scrambled objects, but also of a derived subject when *comp-trace* effects are neutralized. It is also unclear what predictions this labelling approach would make for indirect objects, and nothing about them is mentioned in the proposal. Therefore, though conceptually appealing because it reduces the ban on subextraction to very primitive syntactic principles, it lacks some of the empirical coverage provided in this paper. As other approaches, it does not predict a distinction between the V domain and the T domain. As Deniz Rudin (p.c.) points out, it seems like whatever the real story is, it has to be sensitive to the distinction between those domains, not just to domain-blind properties of the constituent being extracted from.

7 Conclusion

The main goal of the paper has been to establish a correlation between extraction and subextraction effects that is captured by the generalizations in section 3. The purpose was to come up with a uniform analysis that could enable us to cover those gaps which previous proposals for subextraction such as *Freezing*, *CED* or even *Labels* cannot encompass or that are language specific. The proposal that I have followed here is an antilocality approach, according to which movement operations are banned if the distance covered by the moving element is too short. The formulation of antilocality in terms of Brody's (1998) concept of *Projection Line* as the set of XPs projected by X has been proved to be stricter and have a wider scope than other versions of antilocality, by

extending the Spec-to-Spec restriction to also hold for Spec-of-Spec to Spec movement. However, antilocality at the same time is too strong and, on its own, it would predict that indirect objects should, by structural definition, be "frozen". This is not a desirable prediction, which is why Agree can sometimes void locality restrictions between probes and their goals. This is not only operative in A'-contexts but also in A-movement operations.

That said, the proposal is aimed to have a wider scope and be able to account for the cross-linguistic variation regarding extraction and subextraction patterns. Most of the literature on this issue, specially regarding subextraction, has tended to only focus on particular languages. By presenting data from various languages, I have broadened the horizons, connected the facts, and demonstrated that the proposal makes the right predictions in, at least, a small subset of languages. Ultimately, we would like to know how far the generalizations can be extended and whether there might be exceptions to provide a better characterization of the phenomena.

Notes

1. I am using the label "object" here to refer to the complement of the V head. That is, the internal argument of the lexical verb.

2. The label "subject" in this paper is used to refer to the external argument of the verb.

3. Thanks to Stefan Keine (p.c.) for pointing this out to me.

4. See Merchant (2013) for the same argument using ellipsis in English.

5. A problem for the hypothesis would be one such that the vP is fronted in Norwegian but the subject remains stranded. This would be evidence for the hypothesis that the subject has to raise to TP. Thus, the best test to determine the position of the subject is the use of sentential negation and adverb placement as indicated above.

6. In order to show that there has been A'-movement from the derived position, I have chosen topicalization of a pronoun in the extraction cases. This makes it clear that the DP must have first moved to be in a local relation with the matrix verb in ECM cases. For V-Particle constructions, a pronoun DP must obligatorily break the V-Particle sequence illustrating that there must have been movement to a lower position first before A'-extraction.

7. Lasnik(2001: 112) marks the sentences in (20-b) and (20-c) as ungrammatical. He relies on *Freezing* and claims that extraction out of an element that has itself undergone movement is inhibited. My informants report contrary judgments to his. All the informants I have consulted agree that both pairs (18-c) and (20-c) are awkward but still acceptable, thus the question mark (?). In fact, all of my informants were not able to find a contrast between these and the counterparts in which prior movement of the DP to the matrix clause does not occur as in (18-b) and (20-b): both sentences are equally acceptable. It is true though that the grammaticality could be contested since we are aware that the acceptability of these sentences is not universal to every speaker of English.

8. Gallego (2013) notes that examples with the progressive *estar* disallow object shift to which I agree:

- (104) a. [_{TP} Ayer estaba [_{vP} Juan leyendo un libro]]
 yesterday was Juan reading a book
 ‘Yesterday Juan was reading a book’
 b. *[[_{TP} Ayer estaba [_{vP} [un libro]₁ Juan leyendo t₁]]]
 yesterday was a book Juan reading

He mentions i (Gallego 2013: 425, fn. 14) that other auxiliaries yield the same outcome. This is not the case for perfect auxiliary *haber* ‘have’ in my dialect (at least) as the examples below illustrate:

- (105) a. Ayer ha leído Juan un libro
 yesterday has read Juan a book
 ‘Yesterday Juan (has) read a book’
 b. Ayer ha leído un libro Juan
 yesterday has read a book Juan

This is not evidence against his generalization à la Holmberg (1986) but rather evidence for an asymmetry in the head movement of past participles and progressive forms: the former undergo movement whereas the latter do not (i.e. affix hopping).

9. Wiltschko (1997) attributes the lack of superiority effects in German to the fact that scrambling has occurred before A’-movement of the XP.

10. The literature on *that-trace* effects is extensive and this section is not meant to be a literature review of the phenomenon. For those readers interested in knowing what some of the previous analyses are, I suggest that they read Pesetsky (2017) for an overview of the effect, and Erlewine (2019) for a recent approach.

11. As Bayer and Salzmann (2013) indicate, when parsed as if the *wh*-element has moved to the highest specifier of TP, the sentence is degraded and not accepted by everyone. If the parsing does not place the gap in TP, then the sentence is supposed to be acceptable.

12. In their footnote 10, Bayer and Salzmann (2013:319) show that stranded quantifiers are useful to test to restrict the choice between different trace positions. Thus, if a quantifier were stranded in (32-b), the lower the trace the better. This is shown below:

- (106) [Wer]₁ glaubst du, dass (?* alles t₁) 1933 (✓ alles t₁) in Hamburg (✓ alles t₁) in der
 who believe you that all 1933 all in Hamburg all in the
 Regierung war?
 government was
 ‘Who all do you believe was 1933 in Hamburg in the government?’

13. My informants agree.

14. It is possible to have the subject raise to a higher position in Spanish A’-operations when the displaced element is a non-argument (Torrego 1984, Suñer 1994, Gallego 2010). However, even if this is the case subextraction is not allowed:

- (107) a. ¿Con cuánto dinero un amigo de McNulty ha recompensado a Kima?
 with how-much money a friend of McNulty has rewarded to Kima

‘With how much money has a friend of McNulty rewarded Kima?’

- b. *[**de McNulty**]₁, ¿*Con cuánto dinero* [un amigo t₁] ha recompensado a Kima?
 of McNulty with how-much money a friend has rewarded to Kima
 Intd.: ‘McNulty, With how much money has a friend of rewarded Kima?’

The sentence in (i-a) shows that when a non-argument is A'-moved the subject does not have to stay in-situ. However, extraction out of the subject in this higher position is also degraded (i-b). In this case, topicalization has been used since as noted by Abels (2007) *wh*-movement can feed topicalization in certain circumstances whereas it is not that clear that *wh*-movement can feed other *wh*-movement operations. This is not one of them, though. Cases like this do not posit a challenge for the analysis that is proposed in this paper.

15. I am using the label "indirect object" here to refer to the recipient or applied argument in a double object construction.

16. Lohndal (2011: 185, fn. 28) cites an example from Baltin (2001:251-252, fn. 2) in which Baltin shows *wh*-movement of an indirect object to be possible. The example is the following:

- (108) a. John taught Sally French.
 b. Who did John teach French?

These examples are not as peripheral as the literature notes. In fact, the informants consulted did not have any problems with their acceptability.

17. Taken from Müller (1998: 10, ex.[25-a]).

18. No distinction is made between XP and X' from this point in the paper. They all are treated as partial copies of a projecting head. Intermediate XPs are differentiated from maximal XPs in that the former is dominated by its copy while the latter is not. This is just for ease of representation. Thus, the fact that specifiers are sisters to XPs does not entail that they are treated as adjuncts.

19. In Johnson's (1991) case the verb moves to T. There is not substantial evidence for head movement onto T in English which is why I do not want to commit to this claim here.

20. This structure may not be the conventional representation of ECM verbs assuming that the main verb moves further than *v*. Nothing crucial relies on the labels used here. What is being illustrated is that when the optional subject-to-object raising takes place it must break the matrix *v* AdP sequence. What the exact position that may be is not the main question of this paper.

21. Under Cyclic Linearization, movement to the edge of the phase has to occur in order to yield an appropriate ordering statement: when the C head is overt, movement to its left is obligatory because the *wh*-element has to be linearized to the left of *that*. However, if C is null, the moved subject in the specifier of TP is already the leftmost pronounced element in the embedded clause. This means that even if the movement to the edge of the CP does not occur, no contradictory ordering statement will be given. In other words, movement of the subject from the embedded TP to the matrix clause will not yield as outcome an ordering paradox. This is fully compatible with the definition of antilocality adopted in the paper. However, I will follow Andrew Simpson's suggestion and not adopt this model here. For further details of how this would

work, I refer the reader to Erlewine (2019).

22. Andrew Simpson (p.c.) wonders how the [Q] feature comes to be part of T's feature stack. There are two possibilities: (i) T is inserted into the structure from the work space with the [Q] feature only in matrix subject questions given that it is the highest functional head to be merged in its subarray; (ii) T inherits all the features from C via "donate" (Citko 2014: 52) as a result of Feature Inheritance (Richards 2007, Chomsky 2008).

23. For arguments against the other diagnostics, I refer the reader to Harley and Jung (2015). For the opposite view, more in line with Pytkänen's (2002) hypothesis, I suggest that the reader takes a look at Bruening's extensive work on Applicatives (2010, 2018a, 2018b a.o.).

24. I would like to thank Brad Foley for the English examples.

25. My definition of Agree in (73) resembles Preminger's (2019) formulation. However, the purpose and the domain of application is different. His formulation of the PMC would enable movement of the D head after the Agree has occurred. This is not what we want here.

26. For Bošković (2007), successive cyclic movement is motivated by the need of the XP to move. Thus, the movement triggering feature is not part of the probing head, but is part of the moving goal. This greed-based approach to movement ensures that the goal is accessible to higher probes, which eventually will match the goal's feature. The fact that a goal has to move does not mean that it is free from locality as illustrated by subextraction data or *that-trace* effects.

27. Norwegian allows symmetric passives. It has been argued (Haddican and Holmberg 2018, and Holmberg, Sheehan, and Van der Wal 2019) that the language has optionality on how to form a passive: (i) it can form a passive like English in which case, v is the locus of case features but not a phase in the passive; or (ii) it can form a passive like German and Spanish and thus v is the locus of case features for the recipient and the theme in the active but only for the recipient in the passive.

28. See Preminger (2019) for further details.

29. See Corver's (2017) for a recent overview on *Freezing*.

References

- Abels, Klaus. 2003. "Successive Cyclicity, Anti-locality, and Adposition Stranding." PhD diss., University of Connecticut.
- . 2007. "Towards a restrictive theory of (remnant) movement!" In *Linguistic Variation Yearbook*, edited by Jeroen van Craenenbroeck and Johan Rooryck, 7:53–120. 1. Amsterdam: John Benjamins.
- Baltin, Mark. 2001. "A-movement." In *The handbook of contemporary syntactic theory*, edited by Mark Baltin and Chris Collins, 226–254. Oxford: Blackwell.

- Bayer, Josef, and Martin Salzmann. 2013. “{That}-trace effects and resumption: How improper movement can be repaired.” In *Repairs: The Added Value of Being Wrong*, edited by Patrick Brandt and Eric Fuß, 275–334. De Gruyter Mouton.
- Beck, Sigrid. 1996. “Wh-Constructions and Transparent {L}ogical {F}orm.” PhD diss., Universität Tübingen.
- Bentzen, Kristine. 2009. “Subject positions and their interactions with verb movement.” *Studia Linguistica* 63 (3): 261–291.
- Boeckx, Cedric, and Norbert Hornstein. 2005. “A Gap in the {ECM} Paradigm.” *Linguistic Inquiry* 36 (3): 437–441.
- Bošković, Željko. 1997a. “Coordination, Object Shift, and {V}-Movement.” *Linguistic Inquiry* 28 (2): 357–365.
- . 1997b. *The Syntax of Nonfinite Complementation: An Economy Approach*. 247. Cambridge, Massachusetts: MIT Press.
- . 2005. “On the locality of left branch extraction and the structure of {NP}.” *Studia Linguistica* 59 (1): 1–45.
- . 2007. “On the Locality and Motivation of {M}ove and {A}gree: An Even More Minimal Theory.” *Linguistic Inquiry* 38 (4): 589–644.
- . 2016. “On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling.” *The Linguistic Review* 33 (1): 17–66.
- . 2018. “On Movement out of Moved Elements, Labels, and Phases.” *Linguistic Inquiry* 49 (2): 247–282. doi:[10.1162/ling](https://doi.org/10.1162/ling).
- Bresnan, Joan. 1977. “Variables in the theory of transformations.” In *Formal Syntax*, edited by Peter Culicover, Thomas Wasow, and Adrian Akmajian, 157–196. Academic Press.
- Brillman, Ruth, and Aron Hirsch. 2016. “An anti-locality account of English subject/non-subject asymmetries.” In *Proceedings of CLS 50*.
- Brody, Michael. 1998. “Projection and Phrase Structure.” *Linguistic Inquiry* 29 (3): 367–398.
- Browning, Marguerite. 1996. “CP Recursion and that-t Effects.” *Linguistic Inquiry* 27 (2): 237–256.
- Bruening, Benjamin. 2010. “Ditransitive Asymmetries and a Theory of Idiom Formation.” *Linguistic Inquiry* 41 (3): 519–562.

- Bruening, Benjamin. 2018a. "Depictive Secondary Predicates and Small Clause Approaches to Argument Structure." *Linguistic Inquiry* 49 (3): 537–559.
- . 2018b. "Double Object Constructions and Prepositional Dative Constructions are Distinct: A Reply to Ormazabal and Romero 2012." *Linguistic Inquiry* 49:123–150.
- Burzio, Luigi. 1986. *Italian Syntax*. Reidel Publishers.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht, The Netherlands: Foris Publications.
- . 1995. *The minimalist program*. 420. Cambridge, Massachusetts: MIT Press.
- . 2000. "Minimalist inquiries: The framework." In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, edited by Roger Martin, David Michaels, and Juan Uriagereka, 89–156. Cambridge, MA: MIT Press.
- . 2001. "Derivation by Phase." In *Ken Hale: A Life in Linguistics*, edited by Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.
- . 2008. "On Phases." In *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, edited by Robert Freidin, Carlos Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge: MIT Press.
- . 2013. "Problems on Projection." *Lingua* 130:33–49.
- Chung, Sandra, and James McCloskey. 1983. "On the interpretation of certain island facts in {GPSG}." *Linguistic Inquiry* 14 (4): 704–713.
- Citko, Barbara. 2014. *Phase Theory: An Introduction*. Cambridge, UK: Cambridge University Press.
- Corver, Norbert. 2017. "Freezing Effects." In *The Wiley Blackwell Companion to Syntax*, Second edi, edited by Martin Everaert and Henk van Riemsdijk. John Wiley & Sons.
- Cuervo, Cristina. 2003. "Datives at Large." PhD diss., Massachusetts Institute of Technology.
- Culicover, Peter W. 1993. "Evidence against {ECP} accounts of the that-t effect." *Linguistic Inquiry* 24 (3): 557–561.
- Deal, Amy Rose. 2009. "The Origin and Content of Expletives: Evidence from "Selection"." *Syntax* 12 (4): 285–323.
- Diesing, Molly. 1990. "The Syntactic Roots of Semantic Partition." PhD diss., UMASS-Amherst.
- Engdahl, Elisabet. 1983. "Parasitic Gaps." *Linguistics and Philosophy* 6 (1): 5–34.

- Erlewine, Michael Yoshitaka. 2016. "Anti-locality and optimality in Kaqchikel Agent Focus." *Natural Language and Linguistic Theory* 34 (2): 429–479.
- . 2017. "Why the null complementizer is special in complementizer- trace effects." In *A pesky set: Papers for David Pesetsky*, edited by Claire Halpert, Hadas Kotek, and Coppe van Urk, 371–380. MIT Working Papers in Linguistics.
- . 2019. "Anti-locality and subject extraction." Ms. National University of Singapore.
- Fanselow, Gisbert. 2001. "Features, theta-roles, and free constituent order." *Linguistic Inquiry* 32 (3): 405–437.
- Fox, Danny, and David Pesetsky. 2005. "Cyclic Linearization of Syntactic Structure." *Theoretical Linguistics* 31 (1-2): 1–46.
- Gallego, Ángel. 2010. *Phase Theory*. Amsterdam: John Benjamins.
- . 2013. "Object shift in Romance." *Natural Language and Linguistic Theory* 2 (2): 409–451.
- Georgala, Effi. 2010. "Why German is not an exception to the universal <IO, DO> base order of double object constructions." In *West Coast Conference on Formal Linguistics* 28. University of Southern California.
- Georgala, Effi, Waltraud Paul, and John Whitman. 2008. "Expletive and Thematic Applicatives." In *West Coast Conference on Formal Linguistics* 26, 181–189. University of California, Berkley.
- Grohmann, Kleanthes K, and Ricardo Etxepare. 2003. "Root Infinitives: A comparative view." *Probus* 15 (2): 201–236.
- Haddican, Bill, and Anders Holmberg. 2018. "Object symmetry effects in Germanic: Evidence for the role of case." *Natural Language and Linguistic Theory* 37 (1): 91–122.
- Haider, Hubert. 1990. "Topicalization and Other Puzzles of {G}erman Syntax." In *Scrambling and Barriers*, edited by Günther Grewendorf and Wolfgang Sternefeld, 93–112. Amsterdam: John Benjamins Publishing Company.
- . 2017. "Mittelfeld Phenomena. Scrambling in Germanic." In *The Wiley Blackwell Companion to Syntax*, Second Edi, edited by Martin Everaert and Henk van Riemsdijk. John Wiley & Sons.
- Harley, Heidi, and Hyun Kyoung Jung. 2015. "In Support of the PHAVE Analysis of the Double Object Construction." *Linguistic Inquiry* 46 (4): 703–730.

- Hein, Johannes. 2019. "Verb movement and the lack of verb-doubling VP-topicalization in Germanic." Ms. University of Postdam.
- Holmberg, Anders. 1986. "Word Order and Syntactic Features." PhD diss., Stockholm.
- . 2000. "Scandinavian Stylistic Fronting: how any category can become an expletive." *Linguistic Inquiry* 31 (3): 445–483.
- Holmberg, Anders, Michelle Sheehan, and Jenneke Van der Wal. 2019. "Movement from the double object construction is not fully symmetrical." *Linguistic Inquiry* 50 (4): 677–722.
- Huang, C.-T. James. 1982. "Logical relations in {C}hinese and the theory of grammar." PhD diss., Massachusetts Institute of Technology.
- Johnson, Kyle. 1991. "Object positions." *Natural Language and Linguistic Theory* 9 (4): 577–636.
- . 2004. "How to be Quiet." In *Proceedings from the 40th Annual Meeting of the Chicago Linguistics Society*, edited by Nikki Adams, Adam Cooper, Fey Parrill, and Thomas Wier, 1–20. Chicago, Illinois: Chicago Linguistics Society.
- Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, Massachusetts: MIT Press.
- Larson, Richard. 1988. "On the double object construction." *Linguistic Inquiry* 19 (3): 335–392.
- . 2010. "On Pykkänen's Semantics for Low Applicatives." *Linguistic Inquiry* 41 (4): 701–704.
- Lasnik, Howard. 1995. "Case and Expletives Revisited: On Greed and Other Human Failings." *Linguistic Inquiry* 26 (4): 615–634.
- . 2001. "Subjects, Objects and the EPP." In *Objects and other subjects: Grammatical functions, functional categories, and configurationality*, edited by William D Davies and Stanley Dubinsky, 103–121. Dordrecht: Kluwer Academic Press.
- Lasnik, Howard, and Mamoru Saito. 1991. "On the Subject of Infinitives." In *Chicago Linguistics Society 27*, edited by Lise M Dobrin, Lynn Nichols, and Rosa M Rodriguez, 27:324–343. Chicago University.
- Legate, Julie Anne. 2003. "Some Interface Properties of the Phase." *Linguistic Inquiry* 34 (3): 506–516.
- Lohndal, Terje. 2007. "{That}-t in Scandinavian and Elsewhere: Variation in the Position of C." *Working Papers in Scandinavian Syntax* 79:47–73.
- . 2011. "Freezing effects and objects." *Journal of Linguistics* 47:163–199.

- Marantz, Alec. 1993. "Implications of Asymmetries in Double Object Constructions." In *Theoretical Aspects of {B}antu Grammar 1*, edited by Sam A Mchombo, 113–151. Stanford University: CSLI Publications.
- McIntyre, Andrew. 2006. "The interpretation of German datives and English "have"." In *Datives and Other Cases*, edited by Daniel Hole, André Meinunger, and Werner Abraham, 185–211. Philadelphia: John Benjamins.
- Merchant, Jason. 2013. "Voice and Ellipsis." *Linguistic Inquiry* 44 (1): 77–108.
- Müller, Gereon. 1998. *Incomplete Category Fronting: A Derivational Approach to Remnant Movement in German*. Dordrecht ; Boston ; London: Kluwer Academic Publishers.
- Nissenbaum, Jon. 2000. "Covert movement and parasitic gaps." In *Proceedings of the North East Linguistic Society*, edited by Masako Hirotoni, Andries Coetzee, Nancy Hall, and Ji-yung Kim, 541–556. Rutgers University: Graduate Linguistic Student Association.
- Nylsen, Øystein. 1997. "Adverbs and A-shift." *Working Papers in Scandinavian Syntax* 59:1–39.
- Perlmutter, David M. 1968. "Deep and Surface Structure Constraints in Syntax." PhD diss., Ph.D. dissertation, Massachusetts Institute of Technology.
- Pesetsky, David. 2017. "Complementizer-Trace effects." In *The Wiley Blackwell Companion to Syntax*, Second Edi, edited by Martin Everaert and Henk van Riemsdijk. John Wiley & Sons.
- Platzack, Christer. 2012. "Cross Germanic variation in the realm of support verbs." In *Comparative Germanic Syntax: The state of the art*, edited by Peter Ackema, Rhona Alcorn, Caroline Heycock, Dany Jaspers, Jeroen van Craenenbroeck, and Guido Vanden Wyngaerd, 279–310. Amsterdam/Philadelphia: John Benjamins.
- Preminger, Omer. 2019. "What the PCC tells us about "abstract" agreement, head movement, and locality." *Glossa: a journal of general linguistics* 4 (1): 1–42. ISSN: 2397-1835.
- Pylkkänen, Liina. 2002. "Introducing Arguments." PhD diss., Massachusetts Institute of Technology.
- Rackowski, Andrea, and Norvin Richards. 2005. "Phase Edge and Extraction: A {T}agalog Case Study." *Linguistic Inquiry* 36 (4): 565–599.
- Richards, Marc. 2007. "On feature inheritance." *Linguistic Inquiry* 38:563–572.
- Richards, Norvin. 1998. "The principle of minimal compliance." *Linguistic Inquiry* 29 (4): 599–629.

- Rizzi, Luigi. 1982. *Issues in {I}talian syntax*. Dordrecht, Holland: Foris Publications.
- . 1997. “The fine structure of the left periphery.” In *Elements of grammar*, edited by Liliane Haegeman, 281–337. Dordrecht: Kluwer Academic Publishers.
- . 2006. “On the form of chains: Criterial positions and ECP effects.” In *On Wh movement*, edited by Norbert Corver and Lisa Cheng, 97–133. Cambridge, MA: MIT Press.
- Ross, John. 1967. “Constraints on variables in syntax.” PhD diss., Massachusetts Institute of Technology.
- Sauerland, Uli. 1996. “Guess how?” In *Proceedings of ConSoLE*, 297–311.
- Suñer, Margarita. 1994. “V-Movement and the licensing of argumental wh-phrases in Spanish.” *Natural Language and Linguistic Theory* 12:335–372.
- Svenonius, Peter. 2002. “Subject Positions and the Placement of Adverbials.” In *Subjects, Expletives, and the EPP*, edited by Peter Svenonius, 199–240. New York: Oxford University Press.
- Takahashi, Daiko. 1994. “Minimality in movement.” PhD diss., University of Connecticut.
- Torrego, Esther. 1984. “On Inversion in {S}panish and some of its effects.” *Linguistic Inquiry* 15 (1): 103–129.
- Truswell, Robert. 2007. “Extraction from adjuncts and the structure of events.” *Lingua* 117:1355–1377.
- Uriagereka, Juan. 1999. “Multiple spell-out.” In *Working minimalism*, edited by Samuel Epstein and Norbert Hornstein, 251–282. Cambridge, Massachusetts: MIT Press.
- Van Urk, Coppe. 2015. “A uniform syntax for phrasal movement: A case study of Dinka Bor.” PhD diss., MIT.
- Vikner, Sten. 1995. *Verb Movement and Expletive Subjects in the {G}ermanic Languages*. 294. Oxford: Oxford University Press.
- . 2005. “Object Shift.” In *The Blackwell Companion to Syntax*, edited by Martin Everaert and Henk van Riemsdijk, 392–436. Oxford: Blackwell Publishing.
- Watanabe, Akira. 1993. “Larsonian {CP} Recursion, Factive Complements, and Selection.” In *Proceedings of the North East Linguistic Society 23*, edited by Amy J Schafer, 523–537. University of Ottawa: Graduate Linguistic Student Association.
- Wexler, Ken, and Peter Culicover. 1980. *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.

Wiltschko, Martina. 1997. "Superiority in {G}erman." In *West Coast Conference on Formal Linguistics*, edited by Emily Curtis, James Lyle, and Gabriel Webster, 431–445. University of Washington: Stanford Linguistics Association.