Revisiting extraction and subextraction patterns from arguments

Luis Miguel Toquero-Pérez
University of Southern California

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. 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. This version of antilocality can derive *Freezing* effects, Huang's (1982) *CED*, and their exceptions; and *Comp-trace* effects and their neutralization, extending them to subextraction. 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. The version adopted here will allow extraction of the whole XP, but disallow extraction of its specifier due to the lack of an 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 (1a) and (1b) 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)   a.   *Which city$_1$ did [pictures of t$_1$] cause a scandal?
       b.   Which city$_1$ did you see [pictures of t$_1$]?

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 subextraction or that are (internally or externally) merged in a non-subject position are suitable to such an 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 specifiers. The asymmetry in (1a) and (1b) is captured by these grammatical constraints which, as they stand, make the following predictions: (i) movement out of any XP that is generated in or moves up to a specifier position is banned; and (ii) movement out of a shifted or scrambled object should not be possible either. Neither (i) nor (ii) are borne out, though, as shown by (2) and (3):

(2)   Who$_2$ did Kima call [friends of t$_2$]$_1$ up t$_1$?

(3)   German
      [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?'

(2) violates both the *CED* and *Freezing* if, following Johnson (1991) and Boeckx and Hornstein (2005), the direct object [friends of *who*] 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. (3) 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*, e.g. (4); but

both operations are possible when the XP is the object regardless of the presence/absence of the complementizer (5):

(4) a. What$_1$ did they say [(*that) t$_1$ caused a scandal]?
     b. Which city$_1$ did they say [(*that) [a picture of t$_1$] caused a scandal]?

(5) a. What$_1$ did they say [(that) the picture caused t$_1$]?
     b. Which city$_1$ did they say [(that) the picture caused [a scandal about t$_1$]]?

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 2003, Abels 2003, Bošković 2005, Erlewine 2016, 2020 a.o.) that is at play in cases like (1a) and (4) where the launching and landing sites of movement do not meet the necessary distance requirements for the operation to take place: the *wh*-element moves from Spec,TP to Spec,CP in the matrix clause in (1a), and it undergoes the same movement path in example (4) in the embedded clause, assuming CP is a phase, before moving onto the matrix. This movement is "too short" in that it only crosses the maximal projection of TP. In the grammatical cases, the distance between landing and launching sites is larger: from the complement of V to Spec,*v*P in (1b); a similar movement will have occurred first for the shifted object and scrambling cases, (2) and (3), and then from that position subextraction to Spec,CP will have happened. As opposed to (1a) and (4), the movement operations involve crossing more than the containing maximal projection, which makes them "long enough".

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 (6a) and (6b):

(6) a. Who$_1$ did McNulty show t$_1$ a picture of Baltimore?
     b. *Which person$_1$ did McNulty show [a friend of t$_1$] 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.[1] This results in the desired con-

---

1. See fn. 37 for details on why the PMC should hold in the grammar.

trast 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, i.e. the Applicative Phrase (ApplP) must be lower than the VP, I will argue that this analysis is problematic for several reasons and conclude that ApplP should in fact be higher than VP.

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, I present data from English, Spanish, German and Norwegian.[2] Based on the data gathered from these languages, I will demonstrate 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 (2020) antilocality; and Bošković's (2018) labelling approach. 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: Subsection 4.1 motivates and spells out

---

2. There were a total of 15 native speakers of English. All of them were American English speakers. 5/15 were linguists, the other 10 had no knowledge or training in linguistics. For the Spanish judgments, I consulted a total of 15 Iberian Spanish speakers, none of whom were linguists. The speakers all spoke the northern-central variety of Spanish. Three native speakers of German provided with grammaticality judgments, one of whom was a linguist. Two of them spoke München German while the other spoke the variety from the Berlin-Leipzig area. Four Norwegian speakers provided judgments. All of them were non-linguists and came from the Oslo area. Judgments were collected over email in most cases providing minimal pairs to check gradience in acceptability. The informants were also asked directly "on a scale of 1–7, (1 being terrible, 7 being great) how good they found sentence X?". It is also worth noting that in languages like Spanish and also in Norwegian where the preposition is pied-piped with the *wh*-element after movement, a context was provided when necessary to avoid attachment ambiguities. The judgments were also elicited in person when possible. The same procedure was followed here, but speakers would give feedback on the test sentences; for instance, if the test sentence was slightly unacceptable, they would minimally alter the test sentence so that the grammaticality was improved.

the particular definition of antilocality and shows how far we can get with only this constraint. Subsection 4.2 introduces the definition of the PMC adopted here to bridge the gap created by antilocality. Subsection 4.3 concentrates on indirect objects, where the interaction of antilocality and the PMC is illustrated in detail. Section 5 connects the proposal in Section 4 to the generalizations in Section 3 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, surveying the reasons why they are empirically insufficient or theoretically problematic. Section 7 is the conclusion.

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.[3] This claim is illustrated by the following examples from English in (7):

(7)   English
     a.   What$_1$ did you see t$_1$?
     b.   Which city$_1$ did you see [pictures of t$_1$]?
     c.   What$_1$ did you bet that they had seen t$_1$?
     d.   Which city$_1$ did you bet that they had seen [pictures of t$_1$]?

The examples show that *wh*-movement of the whole object is possible regardless of whether the extraction takes place from a root (7a) or an embedded clause (7c). In addition, subextraction from the object in both root and embedded contexts is also possible as illustrated by (7b) and (7d) respectively. The same is found in many other languages including the ones below:[4]

---

3. 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.

4. I would like to thank an anonymous reviewer for providing with the idiomatic translations of (10c) and (10d).

(8) Spanish
 a. [Qué]$_1$ has   visto t$_1$?
      what   have.2SG seen
      'What have you seen?'
 b. [De qué   ciudad]$_1$ has   visto [fotos   t$_1$]?
      of which city   have.2SG seen   pictures
      'Which city have you seen pictures of?'
 c. [Qué]$_1$ quieres   que tus   primos vean t$_1$?
      what   want.2SG that your cousins see
      'What do you want your cousins to see?'
 d. [De qué   ciudad]$_1$ quieres   que tus   primos vean [fotos   t$_1$]?
      of which city   want.2SG 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   t$_1$ gelesen?
      what for a   book   have you   read
      'What kind of book did you read?'
      [Was]$_1$ hast du   [t$_1$ für ein Buch] gelesen?
      what   have you   for a   book   read
      'What kind of book did you read?'
 b. [Was]$_1$ glaubst du, dass Franz t$_1$ gelesen hat?
      what   think   you that Franz   read   has
      'What do you think Franz has read?'
 c. [Was]$_1$ glaubst du, dass Franz [t$_1$ 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   t$_1$ til bursdagen?
      what   gave you children   for birthday
      'What did you give the children for their birthday?'
 b. [Hva]$_1$ gav du barna   [t$_1$ 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   Peter lager   t$_1$?
      what   want you that Peter cause?
      'What do you want Peter to create?'
 d. [Hva]$_1$ vil   du at   Peter lager [t$_1$ for noe]?
      what   want you that Peter cause   for something
      'What kind of thing do you want Peter to create?'

If we compare these data to counterpart in-situ subjects,[5] 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, work by Deal (2009) has argued that *there*-existential constructions, in fact, involve merger of *there* in *v*P 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.[6] Thus, examples like (11) should be treated as illustrating movement from the object position.

(11)   English
    a.   What$_1$ is there t$_1$ on the wall?
    b.   Which city$_1$ is there [a picture of t$_1$] on the wall?
    c.   What$_1$ did you bet that there is t$_1$ on the wall?
    d.   Which city$_1$ did you bet that there is [a picture of t$_1$] 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. In (12), when the complement of the verb has been *wh*-moved, the subject remains low in (12a) following the verb, while in (12b) it appears higher preceding the verb.

(12)   Spanish in-situ & derived subjects in non-subject *wh*-questions
    a.   [Qué]$_1$ dices   que causó el año pasado una foto   de Baltimore t$_1$?
          what   say.2SG that caused the year past   a   picture of Baltimore
          'What did you say a picture of Baltimore caused last year?'
    b.   [Qué]$_1$ dices   que una foto   de Baltimore causó el año pasado t$_1$?
          what   say.2SG that a   picture of Baltimore caused the year past

This optionality is only found in embedded clauses. In matrix clauses, the subject must always be low (Torrego 1984, Suñer 1994, Gallego 2007, 2010) (13).

---

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

6.   As a reviewer points out, Deal's (2009) argument about expletive *there* and its only occurrence with unaccusatives is not cross-linguistically robust: some Scandinavian languauges allow the equivalent of expletive *there* with unergative verbs. On top of this, Icelandic allows this same *there* with transitives. For a stronger cross-linguistic analysis that discusses Scandinavian and other Germanic languages, I recommend Richards and Biberauer (2006). For the purpose of the present paper, the only thing that matters is that the expletive is merged in the specifier of *v*P and the DP associate remains lower. I want to thank this reviewer for this suggestion.

(13) Obligatory inversion in matrix clauses

    a. [Qué]$_1$ causó el año pasado una foto de Baltimore t$_1$?

       what causó el año pasado una foto de Baltimore t$_1$?

       'What did a picture of Baltimore cause last year?'

    b. *[Qué]$_1$ una foto   de Baltimore causó el   año pasado t$_1$?

       what   a   picture of Baltimore caused the year past

If the subject undergoes *wh*-movement, landing on TP first is not allowed as the ungrammaticality of (14) with the stranded quantifier preceding the verb indicates. Thus, we can conclude that movement must happen in one fell swoop from the in-situ position, i.e. Spec,*v*P to Spec,CP, presumably banning movement from Spec,TP to Spec,CP. This can be seen in both matrix and embedded contexts. I have changed the *wh*-element to the plural to make the stranding of the quantifier more natural.

(14) impossibility of quantifier stranding in TP

    a. [Qué fotos]$_1$ (* todas t$_1$) causaron   el   año pasado (todas t$_1$) un

       what pictures   all   caused.3PL the year past   all   a

      escándalo?

      scandal

      'What pictures caused all a scandal last year?'

    b. [Qué fotos]$_1$ dices   que (* todas t$_1$) causaron   el   año pasado

       what pictures say.2SG that   all   caused.3PL the year past

      (todas t$_1$) un escándalo?

       all   a scandal

      'What pictures did you say that all caused a scandal last year?'

The examples in (15) show the same for subextraction: subextraction can only take place from the lower subject position. We can conclude, based on all these data from Spanish, that extraction and subextraction go hand in hand.

(15) Impossibility of extraction from subject in TP

    a. [De qué ciudad]$_1$ (*[unas fotos   t$_1$]) causaron   el año pasado

       of what city   some pictures   caused.3PL the year past

      ([unas fotos   t$_1$]) un escándalo?

      some   pictures   a   scandal

      'Of which city did some pictures of cause a scandal last year?'

    b. [De qué ciudad]$_1$ dices   que (*[unas fotos   t$_1$]) causaron   el año

       of what city   said.2SG that   some pictures   caused.3PL the year

      pasado ([unas fotos   t$_1$]) un escándalo?

      past   some pictures   a   scandal

      'Of which city did you say some pictures of cause a scandal last year?'

The German data in (16) show the same patterns as the English and Spanish counterparts with respect to the position the subject occupies before extraction and during 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 to Spec,CP. One argument in favor of the in-situ (or at least lower than Spec,TP) position is the position of the subject with respect to the particle *denn*. Assuming that the particle *denn* marks the left edge of the middle-field, i.e. the left edge of TP, (Bayer 2012, Bayer and Salzmann 2013, Haider 2017), one can use its position as a delimiter for raised subjects. If the subject is raised, it will appear above *denn*; but if it has not, it will appear below it. The floating quantifier in (16a) is below the particle indicating that the extraction happened from a lower position. If the floating quantifier is above the particle, the extraction is not allowed (16c). In the case of subextraction, given that the stranded DP *für Ameisen* 'for ants' follows the particle in (16b), the extraction must have taken place from the in-situ position and the operation is permitted. Subextraction from the position above the particle *denn* is disallowed (16d).

(16) German   ((b) is from Diesing 1990; (a), (c) & (d)) are adaptations)
  a. [Wer]$_1$ hat   denn (alles t$_1$) einen Postbeamten gebissen?
        who   have PRT all   a   postman   bit
        'Who all bit the postman?'
  b. [Was]$_1$ haben denn [t$_1$ für Ameisen] einen Postbeamten gebissen?
        what have PRT   for ants   a   postman   bit
        'What kind of ants bit the postman?'
  c. [Wer]$_1$ hat   (?* alles t$_1$) denn einen Postbeamten gebissen?
        who   have   all   PRT a   postman   bit
        'Who all bit the postman?'
  d. *[Was]$_1$ haben [t$_1$ für Ameisen] denn einen Postbeamten gebissen?
        what   have   for ants   PRT a   postman   bit
        'What kind of ants bit the postman?'

A stronger argument that subject raising is not obligatory comes from Haider (1990) who notes that it is possible to topicalize a *v*P with the subject in it as shown in (17):

(17) German *v*P topicalization
  a. [Ein Außenseiter gewonnen]$_1$ hat hier noch nie   t$_1$.
        an outsider.NOM won   has here yet never
        'An outsider has never won here'
  b. [Ein Außenseiter gewonnen]$_1$ hat das Derby   noch nie   t$_1$
        an outsider.NOM won   has the Derby.ACC yet never
        An outsider has never won the Derby yet'

The displaced elements in (17) 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.[7]

Norwegian behaves similar to German, 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 (18a); if the subject is low, negation should precede the subject (18b). The extraction and subextraction data are in (18c) and (18d).

(18)   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.   ??Her har forelesningen **tydeligvis ikke** utløst   en skandale.
        here has lecture   evidently not caused a   scandal
        'The lecture has evidently not caused a scandal here'
   c.   [Hva]$_1$ har **tydeligvis ikke** t$_1$ utløst   en skandale?
          what   has evidently not   caused a   scandal
        'What has evidently not caused a scandal?'
   d.   [om   hvilket tema]$_1$ har **tydeligvis ikke** [forelesningen t$_1$] utløst   en
          about which topic   has evidently not   lecture   caused a
        skandale?
        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 (19a) and the variants in (19b) and (19c):[8]

---

7. Thanks to Stefan Keine (p.c.) for pointing this out to me.
8. See Merchant (2013) for the same argument using ellipsis in English.

(19)   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 vP, we now have evidence to
argue that it is the VP that is fronted and not the whole vP as in German. There-
fore, this result crucially does not undermine the hypothesis that the subject can
stay low.

     As the reader might have noticed, only examples of extraction and subextrac-
tion from root clauses have been provided for German and Norwegian. The rea-
son is that 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,
addressing the effects deserves its own section, rather than a addressing them at
the bottom of this subpart. 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 under-

gone 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 (20b) and (21b) below:

(20)   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** [t_1 to have won the election].

(21)   V-Particle constructions
    a.   Mary **made out** [[*friends of McNulty*] to be fools].
    b.   Mary **made** [*friends of McNulty*]$_1$ **out** [t_1 to be fools].

As the examples illustrate, the DP can intervene between the matrix verb and the matrix adverb as in (20b), and between the matrix verb and the particle following as in (21b). Thus, this can be taken as evidence for movement of the DP. If we look at the A'-counterpart of (20b) and (21b), this is what we find:[9]

(22)   ECM + intervening adverbs: A'-movement
    a.   Them$_1$, I believe t_1 **with all my heart** t_1 to have won the election
    b.   %Who$_2$ do you believe **with all your heart** [a friend of t_2] to have won the election?
    c.   %Who$_2$ do you believe [a friend of t_2]$_1$ **with all your heart** t_1 to have won the election?

(23)   V-Particle constructions: A'-movement
    a.   Them$_1$, Mary **made** t_1 **out** t_1 to be fools?
    b.   %Who$_1$ did Mary **make out** [friends of t_1] to be fools?
    c.   %Who$_2$ did Mary **make** [friends of t_2] **out** to be fools?

<div align="right">adapted from Lasnik (2001)</div>

---

9. 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.

Lasnik (2001, 112) marks the sentences in (23b) and (23c) as ungrammatical. He relies on *Freezing* and claims that extraction out of an element that has itself undergone movement is inhibited. My informants report opposite judgments to his, thus the percent symbol (%). All the informants I have consulted agree that both pairs (22c) and (23c) are awkward but still acceptable. 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 (22b) and (23b): 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.

If one accepts the sentences provided, then these examples suggests 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 (2) of Section 1. More instances of these examples are given in (24):

(24)   English: Object Shift
   a.   Kima called up [friends of McNulty].
   b.   Kima called [friends of McNulty]$_1$ up t$_1$.
   c.   Them$_1$, Kima called t$_1$ up t$_1$?
   d.   %Which person$_1$ did Kima call up [friends of t$_1$]?
   e.   %Which person$_2$ did Mary call [friends of t$_2$]$_1$ up t$_1$?

If, as argued by Johnson (1991) and Boeckx and Hornstein (2005), the object [friends of McNulty] undergoes object shift to a higher position (24b) (presumably the specifier of vP), what the example in (24e) shows is that subextraction from that position after movement is allowed, at least for some speakers.

Object shift is also attested 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: (25b) vs. (25c).

(25)   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 t$_1$
        Not saw   pictures of Baltimore   McNulty

   c. *No estaba [fotos   de Baltimore]$_1$ McNulty viendo t$_1$
        not was   pictures of Baltimore   McNulty seeing
        'McNulty wasn't looking at pictures of Baltimore'

Example (25a) is the baseline with the object in-situ, and (25b) is the object-shifted counterpart: the DP has undergone movement to the specifier of *v*P (see Gallego 2013 for more details). The (sub)extraction can be tested with the following paradigm in (26) which the main verb has moved to T:

(26) a. [Qué cosas]$_1$ no vio (todas t$_1$) McNulty (t$_1$)?
          what things not saw all   McNulty
         'What didn't McNulty see?'

     b. [De qué ciudad]$_2$ no vio [fotos   t$_2$]$_1$ McNulty t$_1$?
          Of what city   not see pictures   McNulty
         'Which city didn't McNulty see pictures of?'

The quantifier *todas* 'all' has been stranded in the shifted object position indicating that extraction of *qué* took place from there. We already know, from Subsection 2.1, that subextraction is possible if the object remains in its base position, but (26b) indicates that such an operation is licit if the object has abandoned its base position as well.

    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), following the work of Svenonius (2001), 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. Vikner (2005) does not provide with the relevant examples but (27a) below supports his claim. In addition, as (27c) demonstrates, it is possible to subextract from the shifted object.[10]

---

10. All of my Norwegian informants (4/4) accept sentences like (27b) in which a full DP has object shifted between the main verb and its particle, but they show a preference for the non-shifted alternative. This same intuition is transferred to the subextraction context: while having *vekk* 'away' after the verb is the preferred option, they also allow it at the end of the sentence.

(27) Norwegian: V-Particle constructions
    a. Peter har ikke kastet   vekk en bok   om   sult
        Peter has not thrown away a   book about hunger
        'Peter hasn't thrown away a book about hunger'
    b. $^{?}$Peter har ikke kastet   [en bok   om   sult]$_1$ vekk t$_1$
        Peter has not thrown   a   book about hunger away
        'Peter has not thrown a book about hunger away'
    c. $^{?}$[Om   hvilket tema]$_2$ har ikke Peter kastet   [en bok   t$_2$] vekk t$_1$
         about which topic   has not Peter thrown   a   book   away
        'Which topic hasn't Peter thrown a book about away?'

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 (28):

(28) German: Scrambling   (Stefan Keine p.c.)
    a. Keiner hat sorgfältig ein Büch 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 t$_1$ gelesen
        No-one has a   book about-that carefully   read
        'Nobody has carefully read a book about that'

The baseline is in (28a), and the declarative counterpart with scrambling is in (28b) where the direct object *ein Buch darüber* 'a book about that' has moved above the adverb *sorgfältig* 'carefully'. What is more, as Abels (2007) notes, it is possible that scrambling feeds *wh*-movement or other A'-operations, but crucially not the opposite. 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 in (29):[11]

(29) 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?'

---

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

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 these examples is the following: the *wh*-element in (29a) and
(29c) 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, (29b) is ill-formed because either scrambling has
crossed a CP boundary; or, because *wh*-movement has occurred first giving rise
to cross-over effects.

    Considering that scrambling can feed *wh*-movement, (sub)extraction should
in principle be possible. Some examples of extraction and subextraction from
scrambled constituents are given in (30):

(30)   a.   [Was]$_2$ hat keiner   t$_1$ sorgfältig t$_1$ gelesen?
             what has no-one   carefully   read
            'What has nobody carefully read?'
       b.   [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 carefully read a book about?'
                                                        Adapted from Abels (2007)

As pointed out by Abels, *wh*-movement of the scrambled constituent can happen
as in (30a). On top of that, movement out of a scrambled element is also possible
as evidenced by (30b).

    Therefore, this subsection has provided substantial cross-linguistic evidence
for the possibility of both extraction of and subextraction from a moved con-
stituent, specifically a direct object. These data suggest that there are few restric-
tions on extraction and subextraction in the lower clausal domain. In the follow-
ing section, we move onto the TP domain.

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
from the lower clausal domain. I am now 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 the 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*:[12] the latter renders the sentence ungrammatical. An example of this asymmetry is provided in (31):

(31)   English: That-trace effects (extraction)
    a.   *Who$_1$ did they say that t$_1$ considered running for mayor?
    b.   What$_1$ did they say that a friend of McNulty considered running for t$_1$?

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 (32a). This restriction is not applicable to subextraction from objects: (32b)

(32)   English: That-trace effects (subextraction)
    a.   *Which person$_2$ did they say that [a friend of t$_2$]$_1$ t$_1$ considered running for mayor?
    b.   Which city$_1$ did they say that McNulty considered taking [a picture of t$_1$]?

(32a) shows that when extraction is not allowed, neither is subextraction. 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 <which person>] 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 (33):

(33)   *That-trace* effects in *tough*-constructions
                                   ((33a) taken from Brillman and Hirsch 2016)
    a.   *Who$_1$ is it possible t$_1$ for t$_1$ to see Mary?
    b.   *Which person$_1$ is it possible t$_1$ for [a friend of t$_1$] to see Mary?

The effects in (31a) and (32a) can be circumvented, as shown by Bresnan (1977), if an adverb intervenes between the complementizer and the trace. Another clear

---

12. 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 (2020) for a recent approach.

    c. *[Was]$_2$ glaubst du, [dass [allen   t$_2$]$_1$ denn Peter t$_1$ kaufen würde?]
        what think you that all.ACC   PRT Peter buy   would
      'What do you think that Peter would buy all?'
                                        adapted from Bayer and Salzmann (2013)
    d. *[Was]$_2$ glaubst du, [dass [t$_2$ für Bücher]$_1$ denn Peter t$_1$ kaufen würde?]
        what think you that   for books PRT Peter buy   would
      'What kind of books do you think that Peter would buy?'
                                        adapted from Bayer and Salzmann (2013)

Assuming that the subject has raised to TP and is therefore the highest specifier, a sentence like (37a) is unacceptable. The same occurs with subextraction in (37b). Some important data points to support their observation are (37c) and (37d): here the direct object has scrambled to a higher position (i.e. TP) above the particle *denn* and thus both extraction and subextraction are banned.[14] 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 (37), if the relevant phrase occupies a lower position or there are some elements that intervene. This prediction is borne out: if the launching site of movement is separated from the complementizer by an intervening phrase, the movement is licit. However, if it is the case that extraction takes place from immediately below *dass*, the operation is illicit. Bayer and Salzmann's (2013, 319: fn.10) show that stranded quantifiers are a useful test to restrict the choice between different trace positions. Thus, if a quantifier is stranded, the lower the trace the better, (38).[15]

(38) *That-trace* neutralization: adverb intervention   (Bayer and Salzmann 2013)
    a.   [Wem]$_1$ glaubst du, [dass (?* allem t$_1$) beim Busfahren (✓ allem t$_1$)
        who.DAT think   you that   all   at-the bus-riding   all
        schlecht wird?]
        nauseous becomes
        'Who do you think will become nauseous during the bus ride?'
    b.   [Wer]$_1$ glaubst du, dass (?* alles t$_1$) 1933 (✓ alles t$_1$) in Hamburg
        who.NOM believe you that   all   1933   all   in Hamburg
        (✓ alles t$_1$) in der Regierung   war?
           all   in the government was
        'Who all do you believe was 1933 in Hamburg in the government?'

---

14. It is possible that the sentence in (37b) is ruled out on independent grounds. Haider (2017) shows that scrambling of a *wh*-element over a discourse particle like *denn* is degraded. But Haider does not provide an explanation for why this is the case.

15. We can assume that in terms of derivational ambiguity, the grammar only computes derivations that converge.

The first Example (38a) illustrates intervention of an adverbial between the trace and the complementizer obviating any *that-trace* effects. Similarly in (38b), 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.

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. Below, I restrict myself to only the varieties that show the effects.[16] 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 (43):

(39) Norwegian: *that-trace* effects
   a. *[Hva]$_1$ vil   du [at   t$_1$ 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 t$_1$] ikke lage   en
          about which topic want you   that   lecture   not cause a
       skandale?]
       scandal
       'About which topic do you want the lecture to not cause a scandal?'

The sentences in (39) are marked with a star which suggests that the subject has left a trace adjacent to the complementizer. In the subextraction case in (39b), 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 (39a), Lohndal (2007, 64) reports, citing Holmberg (2000), that such sentences are improved when negation is inserted (44).[17]

(40) Hvem sa   du at   *(ikke) hadde kommet ?
      who said you that   not   had   come
      'Who did you say that not had come?'   Lohndal (2007, 64: ex. 26)

---

16. As a reviewer points out, one should not deny the fact that many varieties of Norwegian do not show instances of the effects. I direct the reader to Lohndal (2007, 2009) for details on these varieties. Later in the paper, I address some possible approaches on how to account for *non-that-trace* effect grammars.

17. My informants agree.

Holmberg (2000) analyzes the contrast between (39a) and (41a) 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. The subject following the adverbs and negation occupies a lower position in the extended projection of V, i.e. AspP (Bentzen 2009). Thus, once again, we might expect that, if the subject remains lower instead of raising to TP, such effects are not encountered (41):

(41)   *That-trace* neutralization: intervention
    a. [Hva]$_1$ vil du [at tydeligvis ikke t$_1$ skal   lage en skandale?]
        what want you that evidently not   should cause a scandal
        'What do you want to evidently not cause a scandal?'
    b. [Om hvilket tema]$_1$ vil du [at tydeligvis ikke [forelesningen t$_1$]
        about which topic want you that evidently not   lecture
        skal   lage en skandale?]
        should cause a scandal
        'About which topic do you want a lecture 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.[18] Failure to do so would result in automatic ungrammaticality. Just for the purpose of illustration, (46) provides some sentences in which the subject has undergone movement to a higher position, as indicated by the stranded quantifier *todos* "all":

(42)   Spanish: raising of subject and lack of VS inversion
    a. *[Quiénes]$_2$ [todos t$_2$]$_1$ causaron el año pasado t$_1$ un escándalo?
        who   all   caused   the year past   a   scandal
        'Who caused a scandal last year?'

---

18. A reviewer wonders whether one needs to assume that preverbal subjects are actually in A'-positions. This has been argued for by Barbosa (2001) building on work by Bonet (1990), Zubizarreta (1997) and Alexiadou and Anagnostopoulou (1998). However, this assumption is controversial and has been contested by Suñer (1994), Gallego (2007, 2010) and Etxepare and Gallego (2020). Thus, I will follow the latter and assume that the preverbal position of the subject is still an A-position. If this is in fact an A'-position associated with Topichood, then it would follow that subsequent *wh*-movement is banned according to Abels's (2007) generalization: *wh*-movement > Topicalization but *Topicalization > *wh*-movement.

b. *[Quiénes] querías   que [todos t_2]$_1$ causaran el   año pasado t_1 un
       who   wanted.2SG that   all   cause   the year past   a
    escándalo?
    scandal
    'Who did you want to cause a scandal last year?'

c. *[De qué equipo]$_2$ [todos entrenadores t_2]$_1$ causaron el   año pasado t_1
        of what team   all   coaches   caused   the year past
    un escándalo?
    a   scandal
    'Of which team did all coaches cause a scandal last year?'

d. *[De qué equipo]$_2$ querías   que [todos entrenadores t_2]$_1$ causaran el
        of what team   wanted.2SG that   all   coaches   caused   the
    año pasado t_1 un escándalo?
    year last   a   scandal
    'Of which team did you want all coaches to cause a scandal last year?'

The examples in (42) 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 (14), a clear difference in acceptability is obtained.[19]

   This section has provided more evidence for a parallel behavior of extraction
and subextraction. This time I have concentrated on the TP domain, and espe-
cially on *that-trace* effects and how they can be neutralized. In the case of English,
I also discussed the absence and presence of *do*-support. Spanish, as illustrated
in the paragraphs above, does not show these effects. However, it shares the

---

19. 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:

(i) 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 t_1] ha recompensado a Kima?
            of McNulty   with how-much money a   friend   has rewarded   to Kima
         Int.:'McNulty, With how much money has a friend of rewarded Kima?'

The sentence in (ia) 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 (ib).
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.

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.[20] The actual position that indirect objects occupy is subject to a lot of debate and this paper contributes to that debate in Section 4.3 where some arguments are given to support a particular structural representation. However, let's assume for now that they occupy a specifier position (Huang 1982; Larson 1988).

There is a lot of variation with respect to A'-movement of the indirect object. Here I will report judgments of some American English speakers distributed from different areas of the country including but not limited to Southern California, Northern Virginia and North East Pensylvannia.[21] Many speakers of these varieties allow *wh*-extraction of the indirect object in Double Object Constructions (DOC) but it does not allow to subextract from them.[22] This contrast is given in (43):

(43)   English: DOC
    a.   Who$_1$ did McNulty show t$_1$ a picture of Baltimore?
    b.   *Which person$_1$ did McNulty show [a friend of t$_1$] 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

---

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

21.   There is a lot of work that addresses cross-linguistic variation in the status of A'-extraction of the indirect object in English. While it is important to acknowledge that such variation does exist, it is beyond the scope of this paper to provide a detailed overview and subsequent analysis. This remains an open question to be addressed in future work. I direct the interested reader to the work of Siewierska and Hollman (2007), Haddican (2010), Haddican (2012) and references cited therein.

22.   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:

(ii)   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.

object is. The relevant datapoints are in (44), where again subextraction in (44b) is very degraded compared to the extraction counterpart in (44a):

(44)   Spanish: DOC
   a.   [A quién]$_1$ le ha mostrado McNulty t$_1$ 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 t$_1$] una foto   de
         of who   CL has shown   McNulty to a friend   a   picture of
        Baltimore?
        Baltimore
        'Which person has McNulty shown a friend of a picture of Baltimore?'

The asymmetry here presented can also be found in German (45) (Müller 1998, 2010; Stefan Keine p.c.) and Norwegian (46) (Lohndal 2011). 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).

(45)   German: DOC
   a.   [Welchen Buch]$_1$ hat man   t$_1$ einen Preis   gegeben?
        which   book has one.NOM a   prize.ACC given
        Int: 'What has someone given a prize?'
   b.   *[Worüber]$_1$ hat man   [einem Buch   t$_1$] einen Preis   gegeben?
        about-what has one.NOM a   book.DAT a   prize.ACC given
        Int.: 'What has someone given a book about a prize?'
                                                                    (Müller 1998, 10: ex 25)

(46)   Norwegian: DOC   (Lohndal 2011, 164: ex 2)
   a.   [Hvem]$_1$ gav du t$_1$ pakker til bursdagen?
        who   gave you gifts   for birthday
        'Who did you give gifts for their birthday?'
   b.   *[Hva]$_1$ gav du [t$_1$ 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 compelling 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

Table 1 summarizes the data presented so far:[23]

Table 1. Extraction-subextraction data summary

		Extraction	
		Yes	**No**
Subextraction	**Yes**	In situ objects In situ subjects ECM Scrambling & object shift Comp-trace neutralization	Unattested
	No	Indirect objects	Comp-trace effects Derived matrix subjects

The top left cell of Table 1 gathers the syntactic domains 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 ones that allow neither: namely, specifier of TP and complementizer-trace effects (including *tough* constructions). It also shows that there is no attested case where subextraction is possible but extraction is not. 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 (43) to (46), subextraction is not possible even though the extraction condition is met. That said, I propose a generalization concerning Table 1 in (47):

(47)   **Extraction-Subextraction generalization**
        For subextraction out of a DP to be allowed
        a.   extraction of this DP must be possible; and
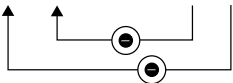   b.   the DP must not be an Indirect Object.

23. Table 1 concentrates on the varieties of the languages that show the extraction-subextraction asymmetry and that are subject to *comp-trace* effects. However, as mentioned in fn.20 there is a lot of variation regarding these two contexts.

We have seen that extraction out of objects and subjects, broadly speaking, is premissible. Besides, given the data, extraction out of indirect objects is never allowed. Thus, I propose the following novel generalization in (48):[24]

(48) **Argument Islandhood Hierarchy**
Indirect objects are stronger islands than vP subjects; which are stronger islands than direct objects:
IO > Subjects > DO

The generalization follows the spirit of Rizzi (1990, 1997) weak-strong island dichotomy and accounts for the well-known facts that direct objects are typically never islands for extraction, while extraction from subjects is degraded and emphasizes that indirect objects involve an even higher degree of opacity. Although no reference to derived subjects is made in (48), extraction from Spec,TP subjects should be as bad as from IOs, due to structural identity. If we assume that v and C are phase heads (Chomsky 2000, 2001a; Legate 2003; Rackowski and Richards 2005; Van Urk and Richards 2015 a.o.), as I will explicitly assume in the next section, when the *wh*-subject of the embedded clause is below an overt C (e.g. *that*), *wh*-subject is one projection below the edge of the phase. In the case of indirect objects, I provide arguments in 4.3 for why indirect objects should be considered high applicatives. That is, a projection above the VP and below the vP. Thus, just like TP subjects, extraction out of them is out because they are one projection below the edge of the vP phase.

Looking at the different contexts surveyed, there seems to be a common factor in the ungrammatical cases: the opaque DP that also cannot move is located in a specifier immediately below the position that would host the moving element. The structural description is provided in (49):

(49) $[_{XP}$ ZP/WP $[_{X'}$ X $[_{YP}$ $[_{ZP}$ t$_{WP}$ $[_{Z'}$ Z $]]$ $[_{Y'}$ Y … $]]]]$

The structural desctiption in (49) indicates that launching and landing sites for the moving elements are too close: a moved subject in Spec,TP targeting the specifier of probing CP for example. Therefore, we should have a generalization like (50) in mind that will help us develop a more accurate analysis:

(50) **The Spec-to-Spec Confinement Generalization (to be revised)**
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; and
b.   movement out of ZP to the specifier of XP is not allowed.

The generalization in (50) is stated as a condition on movement operations and provides a structural description of the ungrammatical data: an element in a specifier is restrained from leaving its position if its potential landing site is too close, i.e. the specifier is confined. However, in light of indirect objects in 4.3 I will show that the generalization in (50) needs to be stronger.

That said, my analysis will try to overcome these problems, capture the data in Table 1 and provide an explanation for the generalizations in (47) and (50). In order to do so, I propose that a revised version of Erlewine's (2016, 2020) 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 will be shown, antilocality is very strong and makes some undesirable predictions for contexts including 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 first outline 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:

(51)   **Antilocality**
       Movement of an XP must cross a projection line.

(52)   **Principle of Minimal Compliance**
       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 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, 2020) and has been adopted by others including Brillman and Hirsch (2016). The exact definition is provided in (53):

(53)   **Erlewine's (2020) most recent Spec-to-Spec Anti-Locality**[25]
       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 (54) is ungrammatical because, as the tree indicates, no maximal projection other than TP is crossed.

(54)   *Who$_1$ did they say that t$_1$ considered running for mayor?
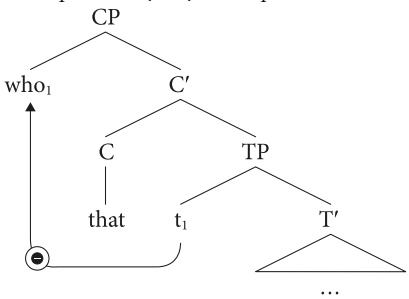
However, as defined, this makes the prediction that A′-movement out the DP in the specifier of TP should be allowed. Assuming that DPs are phases

---

25. In his previous work (Erlewine 2014, 2016, 2017) and on the pre-published version of his 2020 paper, Erlewine considers antilocality to constrain *only* A′-movement. Of course this is problematic for many reasons. First of all, if antilocality is a constraint that operates on movement in general, why should it be restricted to A′-movement? This would rule out Spec,*v*P to Spec,TP subject movement assuming that there are not hidden intermediate projections. Other proponents of the A′-exclusiveness of Spec-to-Spec antilocality (Brillman and Hirsch 2016) suffer from the same problem.

(Matushansky 2005; Heck and Zimmermann 2004; Bošković 2005; Davis 2020a) 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 (55):

(55) *Which person$_1$ did they say that [a friend of t$_1$] considered running for mayor?
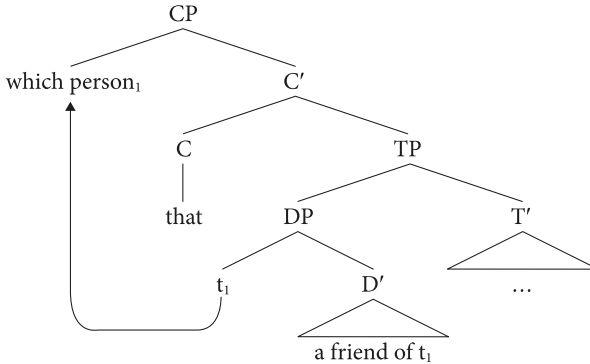
Therefore, this definition of antilocality is not strict enough because it is not able to account for the asymmetry. One of the challenges to face is, then, the extension of Spec-to-Spec antilocality to Spec-of-Spec-to-Spec Antilocality following the assumption that an element that is extracted out of 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": a head X is a full copy of a lexical item L that is inserted in the structure from the lexicon; a phrase XP, and also non-minimal terminal nodes X', are a copy of syntactic features of the element that they immediately dominate. Thus, if X' is projected by X, and XP by X', it then follows that all non-terminal X nodes are copies of the head of the construction, itself a copy of a lexical item L.[26] Therefore, a head X that has projected a phrase XP is (immediately or

---

26. No distinction is made between XP and X' from this point on 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 rep-

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 (56), and what we mean by "crossing a projection line" in (58):

(56) **Antilocality**
      Movement of an XP must cross a *Projection Line*.

(57) **Projection Line**
      A *PL* of α = {x | x = a projection of α}

(58) **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 (58) 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 possible if every single member that has been projected by a head, what we call a *PL* in (57), 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 (59) where *PLs* are marked with a square.[27]

(59) Structural representation of *PLs*
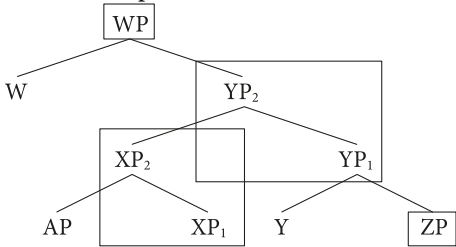
According to the structure in (59), WP and ZP constitute *PLs* because they are the only projections of the head W and Z respectively: *PL* of W = {WP} and the *PL* of

---

resentation. Thus, the fact that specifiers are sisters to XPs does not entail that they are treated as adjuncts.
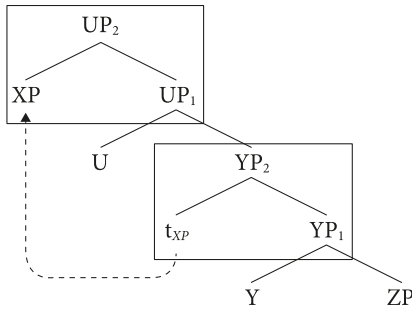
27. I want to emphasize that antilocality constraints are radically different from standard generative constraints on movement. Typically, movement operations cannot cross too many bounding nodes as indicated by the Subjacency Condition (Chomsky 1973) or the barriers framework (Chomsky 1986). Antilocality has to be understood as an economy condition (Abels 2003), i.e. subject to the Last Resort (Chomsky 2001b; Collins 2002; Abels 2003): "an element may only be merged or moved if that leads to the immediate satisfaction of a previously unsatisfiable feature". If nothing is gained from re-merge, then the movement operation is ruled out.

$Z = \{ZP\}$. 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_1 and YP_2, and XP_1 and XP_2 participate in the *PL* of Y and X respectively as members of the set, they do not count as *PL*s on their own.
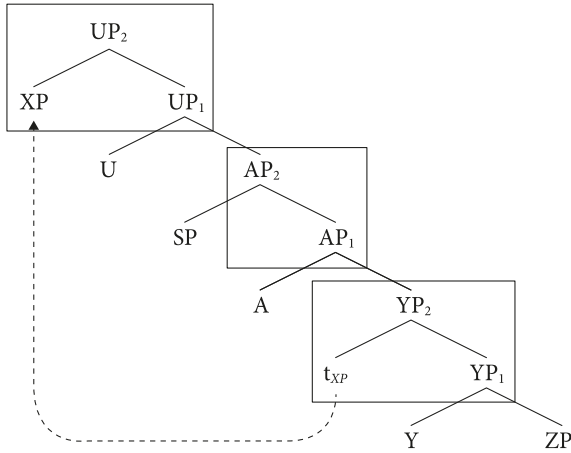
Now, given these formal definitions and our version of antilocality, how does the system make the right predictions for extraction and subextraction? A schematic representation is provided in (60):

(60) Licit vs. illicit patterns
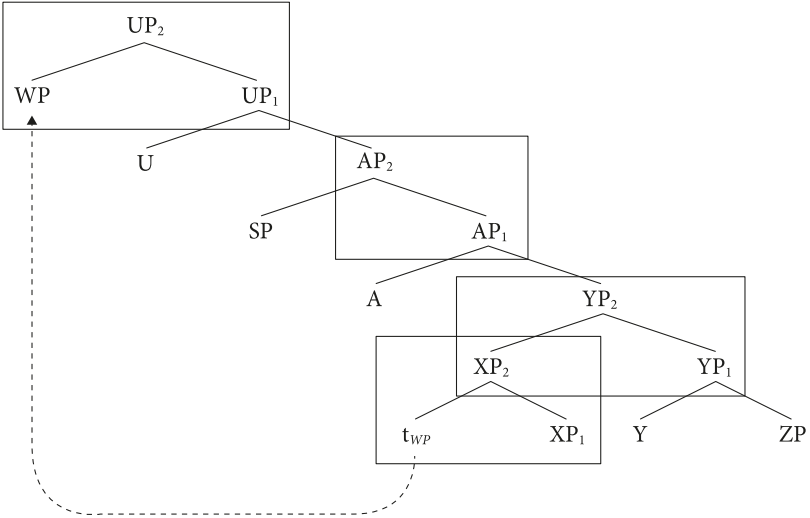
    a. *Extraction

$$
\begin{array}{c}
UP_2 \\
XP \quad\quad UP_1 \\
U \quad\quad YP_2 \\
t_{XP} \quad\quad YP_1 \\
Y \quad\quad ZP
\end{array}
$$

    b. Licit Extraction

$$
\begin{array}{c}
UP_2 \\
XP \quad\quad UP_1 \\
U \quad\quad AP_2 \\
SP \quad\quad AP_1 \\
A \quad\quad YP_2 \\
t_{XP} \quad\quad YP_1 \\
Y \quad\quad ZP
\end{array}
$$

c. *Subextraction

UP$_2$

WP   UP$_1$

U   YP$_2$

XP$_2$ YP$_1$

t$_{WP}$ XP$_1$ Y   ZP

d. Licit subextraction

UP$_2$

WP   UP$_1$

U   AP$_2$

SP   AP$_1$

A   YP$_2$

XP$_2$ YP$_1$

t$_{WP}$ XP$_1$ Y   ZP
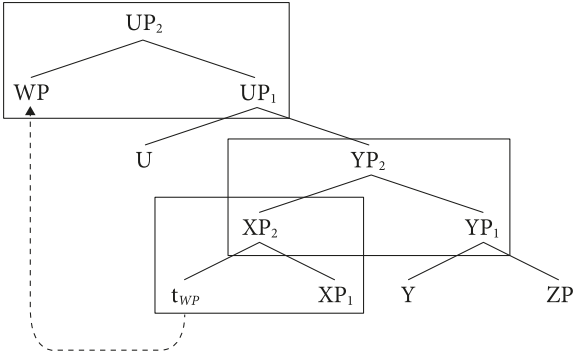
The structure in (60a) represents an illicit extraction pattern: XP is dominated by a projection of Y (i.e. YP$_2$) which is dominated by a projection of U (namely, UP$_1$). The movement of XP only crosses two individual members of independent *PL*s, not a complete set (i.e. YP={YP$_1$, YP$_2$} and UP={UP$_1$, UP$_2$}) which would make them a *PL*. The movement is thus ruled out by antilocality in (56). In other words, as the boxes marking *PL*s show, XP moves from inside one box into another box; the movement does not jump over either of the boxes. On the contrary, (60b) shows an extraction path that crosses over the *PL* of A: {AP$_1$, AP$_2$}. Illicit subextraction in (60c) is precluded by antilocality because only a single member of the *PL*s of X, Y and U is crossed, namely XP$_2$, YP$_2$ and UP$_1$. (60d) contrasts with (60c) since WP crosses the set of projections of A on its way to UP.

Given these structures, it now follows that both (54) and (55) are ruled out by antilocality. I will provide 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 (14) for subjects. The crossed *PLs* are indicated in the tree with a square:

(61)   What$_1$ did you see t$_1$?
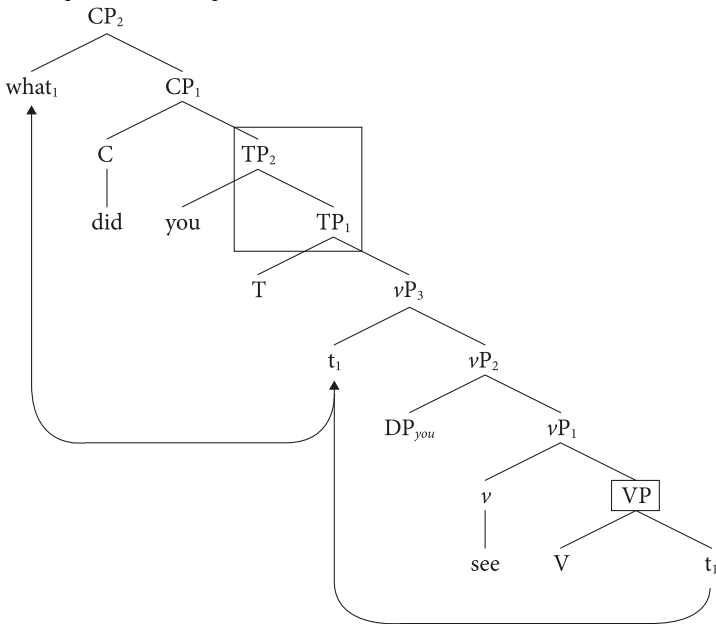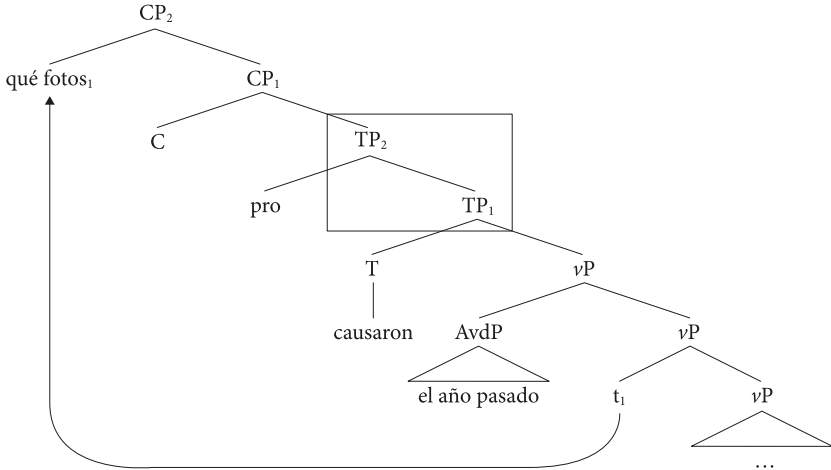
(62)   Which city₁ did you see [a picture of t₁]?
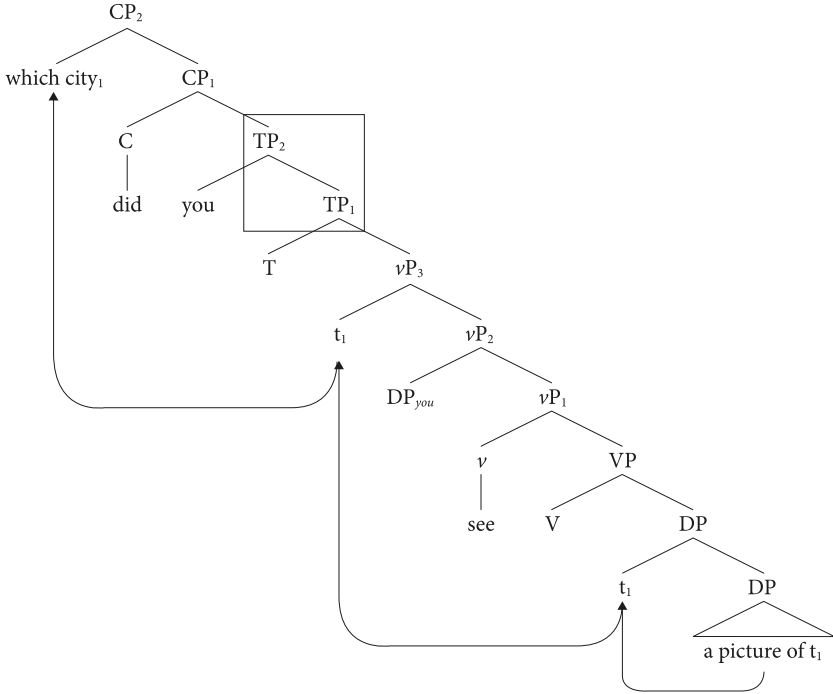
The *wh*-element in (62) first undergoes movement to the specifier of the DP, under the aforementioned assumption that DP is a phase; and from there it moves to the specifier of *v*P, following the assumption that accusative *v* is a phase head and movement has to proceed via its edge (Chomsky 2001a, Legate 2003 a.o.).[28] This movement does not violate antilocality in (56) 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 TPs nodes). For (61) everything remains the same except for the fact that there is no movement through Spec,DP given that this is only necessary for subextraction.
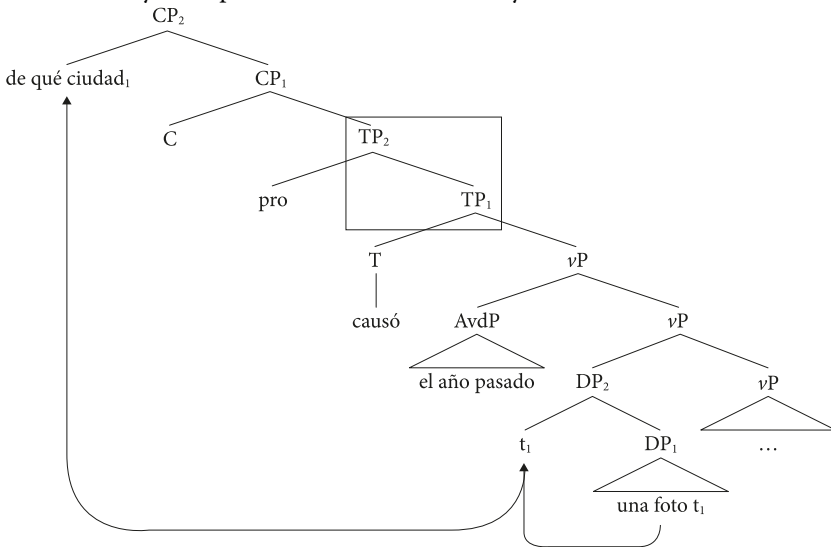
In (63) and (64) the subject remains in-situ in the specifier of *v*P. When C probes for a potential goal, it finds the *wh*-element and brings it to its specifier. This movement operation does not violate antilocality given that the *wh*-element crosses T's *PL*.

---

28.   The claim that *v*P is a phase has been contested. See Keine (2016, 2019).

(63) [Que fotos]$_1$ causaron el año pasado t$_1$ un escándalo?
      what pictures caused   the year past   a scandal
   'Which pictures caused a scandal last year?'

```
                    CP₂
             ┌───────┴───────┐
       qué fotos₁          CP₁
                       ┌────┴─────┐
                       C         TP₂
                              ┌────┴────┐
                             pro       TP₁
                                   ┌────┴────┐
                                   T        vP
                                   │    ┌────┴────┐
                                causaron AvdP     vP
                                       ┌──┴──┐ ┌───┴───┐
                                    el año pasado t₁   vP
                                                      ┌┴┐
                                                      …
```

(64) [De qué ciudad]$_1$ causó el año pasado [una foto   t$_1$] un escándalo?
       of what city   caused the year past   a   picture   a scandal
      'Of which city did a picture cause a scandal last year?'

```
                   CP₂
            ┌────────┴────────┐
      de qué ciudad₁        CP₁
                        ┌─────┴─────┐
                        C          TP₂
                               ┌────┴────┐
                              pro       TP₁
                                    ┌────┴────┐
                                    T        vP
                                    │    ┌────┴────┐
                                  causó AvdP       vP
                                      ┌──┴──┐  ┌───┴───┐
                                   el año pasado DP₂    vP
                                             ┌───┴───┐ ┌┴┐
                                             t₁     DP₁ …
                                                 ┌───┴───┐
                                                una foto t₁
```

These two derivations concern English and Spanish respectively; but they can be applied to the other languages as well. In the case of German and Norwegian, arguments such as topicalization and the position of negation (and adverbs)

respectively were given to support the claim that subjects remain lower than TP and thus (sub)extraction is allowed. This is schematically represented in (65):

(65) $[_{CP} [_{CP} C [_{TP} T [_{vP} [_{DP1} [_{DP2} DP2] [_{DP1} \dots]] [_{vP} v]]]]]$
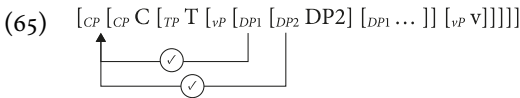
4.1.2 *Movement of and out of moved XPs: Object shift and scrambling*

Antilocality as defined in (56) also makes the right predictions for movement out of moved constituents such as ECM and verb-particle constructions. It was shown in (20)–(21) 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 (66) have a structure as in (67). I follow Johnson's (1991) analysis in which the verb in English raises to a higher position, Johnson calls it μP, above what is in current terms vP. The DP object undergoes shift to the specifier of vP and then the main verb escapes vP to derive the correct word order V>O> Particle.[29]

(66)   ?Which person$_2$ do you believe [a friend of t$_2$]$_1$ with all your heart t$_1$ to have won the election?

(67)   a.   Embedded TP

---
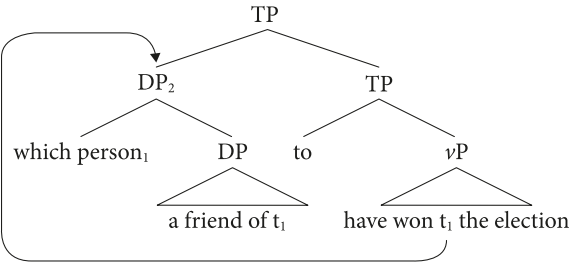
29. In Johnson's (1991) case the verb moves to T. There is not substantial evidence for head movement onto T in English which is why I do not want to commit to this claim here (Emonds 1976; Pollock 1989; Roberts 2010 among others).
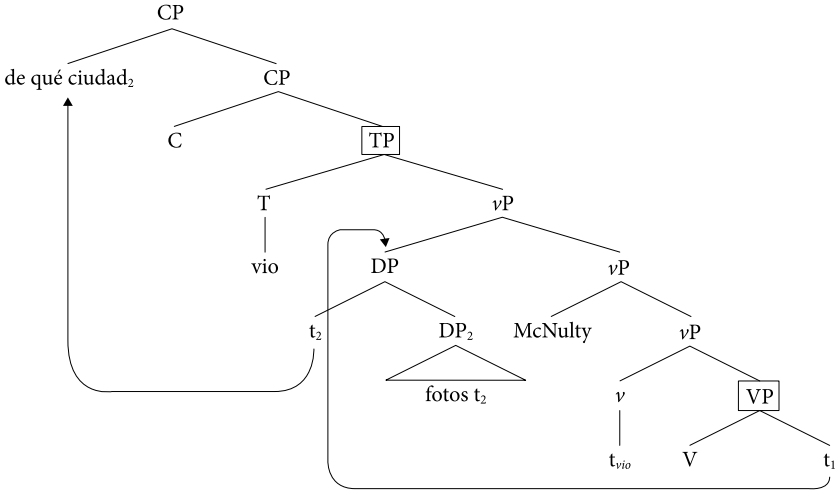
b.   Matrix clause: *PL* of V is crossed
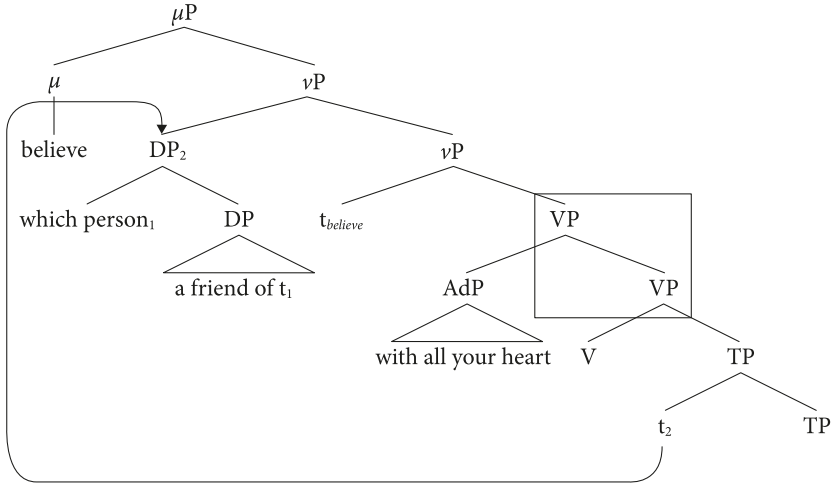
c.   Matrix clause: *PL* of μ and T are crossed
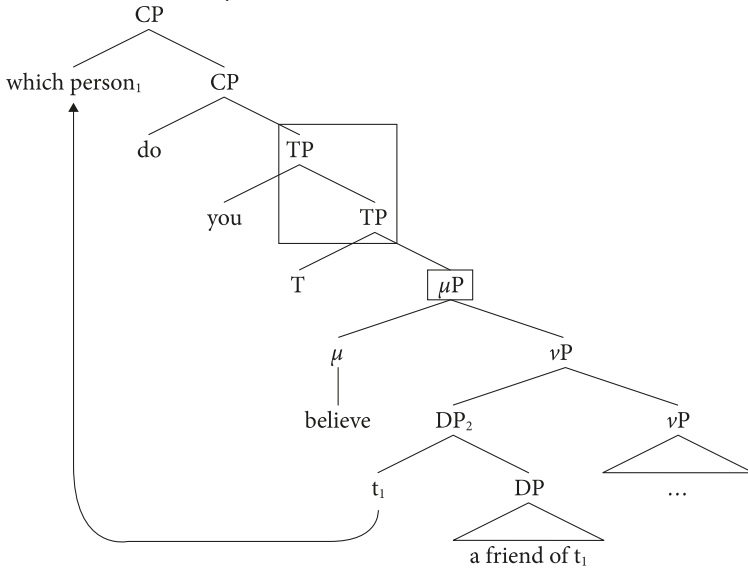
The trees in (67) illustrate that the DP [a friend of <which person>] starts out low in the embedded clause, and raises to the specifier of TP for case that receives from the matrix *v* (67a). The next operation in (67b), which seems to be optional, is movement to the specifier of matrix *v*P. It is in this position that 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 (67c).[30] 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. A Spanish example is used to illustrate the point in (68):

(68) [De qué ciudad]$_2$ vio [fotos   t$_2$]$_1$ McNulty t$_1$?
        Of which city   see   pictures   McNulty
      'Which city did McNulty see pictures of?'

CP
de qué ciudad$_2$ CP
                  C   TP
                        T   vP
                        vio   DP   vP
                              t$_2$ DP$_2$ McNulty   vP
                                    fotos t$_2$ v   VP
                                                    t$_{vio}$ V   t$_1$

In this example, the object has shifted to a higher position which, I assume following Gallego (2013), is the edge of the *v*P. Upon doing so, the DP has crossed the *PL* of V. Once there, subextraction takes place: the *wh*-element lands in the specifier of CP after moving past T's *PL*. The same would occur for extraction except that the whole shifted DP would undergo movement from Spec,*v*P to Spec,CP.

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.[31] Low scrambling is illustrated by (69) whose structural representation is in (70):

---

30. 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* AdvP sequence. What the exact position that may be is not the main question of this paper.
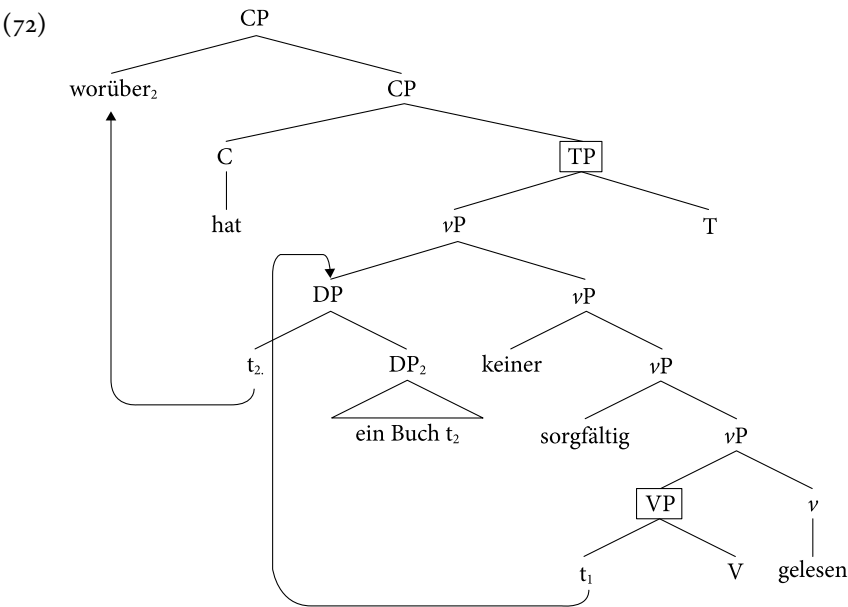
31. On the contrary, there are certain restrictions regarding *mittelfeld* scrambling: targeting the edge of TP is ruled out on independent grounds such as scope (Bayer and Salzmann

(69)   [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?'

(70)
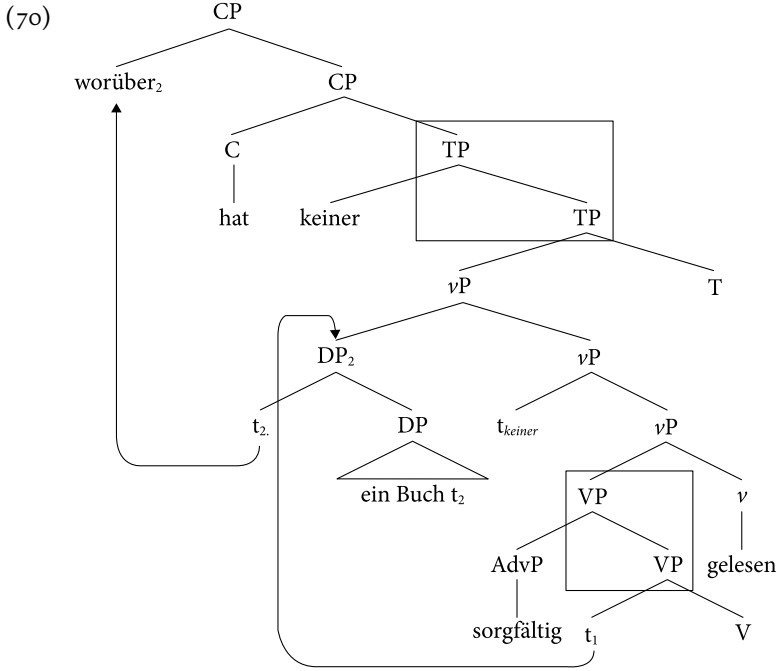
Assuming that *sorgfältig* "carefully" is a low manner adverb adjoined to VP (Cinque 1999), in (70) the subject has moved to TP and the object DP has scrambled to the edge of the *v*P satisfying antilocality in that first movement. The second movement operation is subextraction which also satisfies antilocality since the movement crosses the set of all TP nodes. Stefan Keine (p.c.) points out that sentences like (71) 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 (72):

(71)   [Worüber]$_2$ hat [ein Buch t$_2$]$_1$ keiner sorgfältig t$_1$ gelesen ?
       about-what has a   book   no-one carefully   read
       'What has nobody read a book about carefully?'

2013, 281–282; Haider 2017, 52–54), so *wh*-subextraction from the scrambled constituent is not testable without the possibility of getting into a confound.

(72)

```
                          CP
           ┌───────────────┴───────────┐
       worüber₂                        CP
                          ┌─────────────┴──────────┐
                          C                      ┌─TP─┐
                          │               ┌───────┴───────┐
                         hat             vP               T
                                  ┌───────┴───────┐
                                 DP              vP
                           ┌──────┴──────┐    ┌───┴────┐
                          t₂.          DP₂  keiner     vP
                                  ┌─────┴─────┐    ┌────┴────┐
                               ein Buch t₂  sorgfältig      vP
                                                       ┌─────┴─────┐
                                                    ┌─VP─┐         v
                                                 ┌──┴──┐           │
                                                t₁    V         gelesen
```
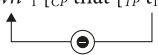
In (71) represented in (72), the subject can be in-situ (Haider 1990) and so scrambling of the object DP can have taken place to the edge of the *v*. Thus, the DP has undergone low scrambling which enables its inside material to be accessible to subsequent syntactic operations. In fact, subextraction 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 and launching sites for *wh*-movement are too close together: antilocality in (56) 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 (73a) for extraction and (73b) for subextraction:

(73)   a.   [$_{CP}$ *wh*-₁ [$_{CP}$ that [$_{TP}$ t₁ [$_{TP}$ T …]]]]

       b.   [$_{CP}$ *wh*-₁ [$_{CP}$ that [$_{TP}$ [$_{DP}$ t₁ [$_{DP}$ a friend of t₁]] [$_{TP}$ T…]]]]

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 (73a); and the higher DP, the topmost TP and the lower CP in (73b). 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 (37) and (39) 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 (33) repeated here as (74). 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 (74) is ungrammatical because the DP/*who* receives case from the complementizer:

(74)   a.   *Who₁ is it possible t₁ for t₁ to see Mary?
       b.   *Which person₁ 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 representation of the illicit movement operations is in (75):

(75)   a.   [$_{CP}$ *wh*-₁ [$_{CP}$ for [$_{TP}$ t₁ [$_{TP}$ T …]]]]

       b.   [$_{CP}$ *wh*-₁ [$_{CP}$ for [$_{TP}$ [$_{DP}$ t₁ [$_{DP}$ a friend of t₁]] [$_{TP}$ T…]]]]

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:

(76)   a.   Who$_1$ did they say that **at no time / for all intents and purposes** t$_1$considered running for mayor?
       b.   Which person$_2$ did they say that **at no time / for all intents and purposes** [a friend of t$_2$] t$_1$ 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, i.e. Culicover's (1993) PolP. These adverbs cannot be adjuncts in the traditional sense, but have to be introduced by functional heads. 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 (76) can be represented as (77):

(77)   a.

b.
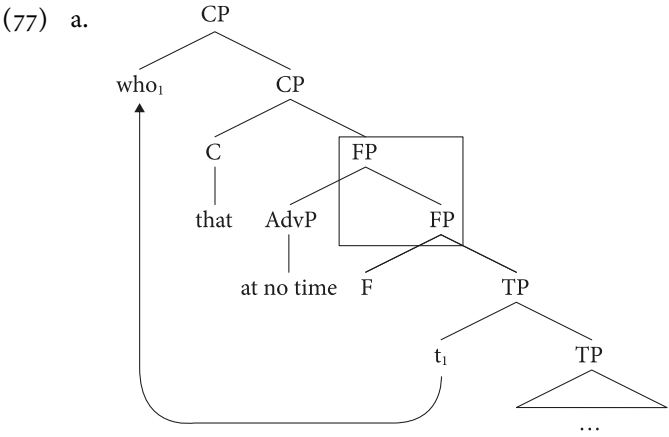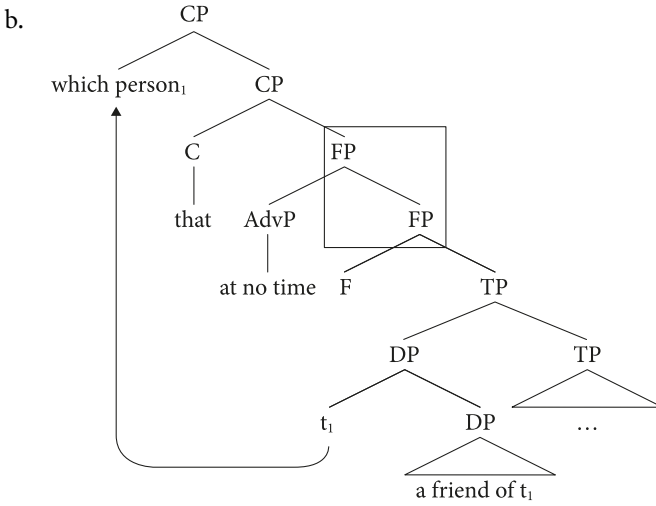
Given the representation we have assumed for the sentences in (77), the movement operation marked by the arrow does not violate antilocality. The intervening functional head projecting the FP that hosts the adverb in its specifier serves as the required *PL* to be crossed, allowing extraction and subextraction 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 (38) and (41) 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 (76). The (sub)extracted element does not land in the specifier of the embedded CP and therefore Antilocality is not violated. The analysis of *that*-less clauses that I adopt is one in which a *that*-less embedded clause is not a CP but a TP, (Hegarty 1991; Webelhuth 1992; Doherty 1993; Svenonius 1994; Bošković 1997b; Wurmbrand 2014).[32] 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. One 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 (78):

---

32. I thank an anonymous reviewer for pointing out many of these references on this topic.

(78)   a.   John didn't believe that [Mary]$_1$ Bill kissed t$_1$.
       b.   *John didn't believe [Mary]$_1$ Bill kissed t$_1$.

In addition to this, Wurmbrand (2014, 155) notes that *that-less* clauses cannot move to subject position, (79a), cannot topicalize, (79b), and cannot undergo syntactic extraposition (79c):

(79)   Adopted from Wurmbrand (2014, 155: Example (41))
       a.   [*(That) John like linguistics] was widely believed.
       b.   [*(That) John likes Mary], Jane didn't believe.
       c.   It seemed at that time [*(that) he had left].

Thus, if there is no CP layer, then the movement chain can be represented as in (80):

(80)   Movement of the subject over a null complementizer: no antilocality violation
       $[_{CP}$ who$_1$ did they say $[_{TP}$ t$_1$ considered running for mayor]]

The counterpart with the overt complementizer is in (81).[33]

(81)   Movement of the subject over an overt complementizer: antilocality or PIC violation
       $[_{CP}$ who$_1$ did they say $[_{CP}$ t$_1$ $[_{CP}$ that $[_{TP}$ t$_1$ $[_{TP}$ …]]]]]

Two potential derivations are possible. The first derivation labeled ●$_1$ yields an antilocality violation because no *PL* is crossed in the first movement operation. The second derivation labeled ●$_2$ 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

---

33. Davis (2020b) shows that English possessor extraction can strand the possessum at the edge of a *that*-less CP:

(iii)   Who$_2$ do they think $[_{CP}$ [t$_2$'s cat]$_1$ Sue found t$_1$ today]?

While this might be taken as evidence that there is a CP layer that hosts successive cyclic movement, it can also be taken as an argument for the fact that if there is no overt CP layer, TP is now the phase because it is the highest projection in V's extended projection (Bošković 2014). Alternatively, as a reviewer points out, it could be the case that C and T are conflated in these cases. Similar ideas have been proposed by Pesetsky (1995), Pesetsky and Torrego (2001) and Erlewine (2020).

be trapped in the complement of the CP phase once the next higher phase head is merged. Thus, the PIC (Chomsky 2001a) is violated.[34]

A reviewer wonders how this generalizes to German and Norwegian. I assume that German V2 *dass*-less clauses are CPs (Wurmbrand 2014). This entails that complementizerless clauses should give rise to *that-trace* effects in these languages. Although in light of examples like (82a), the prediction is a priori not borne out, Reis (1995) has analyzed these as not involving successive cyclic movement. Instead the construction should be analyzed as a bridge verb parenthetical construction. For example, *glaubst du* "do you think" appears in a position lower than what is expected if there is long distance extraction (82b); it can follow the entire question in *slifting* (Ross 1967) (82c). These data and arguments have been taken from Bayer and Salzmann (2013, 298–299) (see their paper for more details).

(82)   German from Bayer and Salzmann   (2013, 298: Example (47)–(49))
    a.   Wer   glaubst du raucht?
       who.NOM believe you smokes
       'Who do you think smokes?'
    b.   Wer   hat, glaubst du, dem   Opa   nochmal einem Schnaps
       who.NOM has believe you the.DAT grandpa again   a   schnaps
       eingeschenkt?
       poured
       'Who, do you think, poured grandpa another schnaps?'
    c.   Wer   hat dem   Opa   nochmal einem Schnaps eingeschenkt,
       who.NOM has the.DAT grandpa again   a   schnaps poured
       glaubst du?
       believe you
       'Who poured grandpa another schnaps, do you think?'

---

34. Another possibility would be to adopt Fox and Pesetsky's (2005) Cyclic Linearization as Brillman and Hirsch (2016) and Erlewine (2017) 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.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 (2017).

The claim that these constructions actually involve parentheticals instead of extraction has also been supported experimentally by Kiziak (2007, 2010): long distance extraction from V2 *dass*-less clauses is significantly degraded unless the matrix verb plays the role of a bridge verb plus a parenthetical. Given these pieces of evidence, we can conclude that V2 *dass*-less embedded clauses are CPs as argued by Wurmbrand (2014) and therefore induce *that-trace* effects.

The situation of Norwegian is more complicated considering the amount of variation with respect to complementizer-trace effects. In fact, it is the case that some speakers are reluctant to remove the complementizer altogether (Lohndal 2007, 63; Franco 2012, 2; cf. Holmberg 2000, 451). However, the data suggests that for those speakers who show sensitivity to the effects and allow complementizer drop, the situation is identical to that of German. This is illustrated in (83): the matrix clause contains a bridge verb and the embedded clause is V2 (Wiklund et al. 2007) (83a)–(83c), while non-bridge verbs lead to degraded acceptability (83d):

(83)   Norwegian
    a.   Hvem sa   du hadde utløst en skandale?
         who   said you had   caused a scandal
         'Who did you say had caused a scandal?'
    b.   Hvem hadde, sa   du, utløst en skandale?
         who   had   said you caused a scandal
         'Who, did you say, had caused a scandal?'
    c.   Hvem hadde utløst en skandale, sa   du ?
         who   had   caused a scandal   said you
         'Who had caused a scandal, did you think?'
    d. ??Hvem skrev du hadde utløst en skandale?
         who   wrote you had   caused a   scandal
         'Who did you write had caused a scandal?'

Given the similarities that Norwegian shows with German, I conclude that these *at*-less embedded clauses that are V2 should be considered CPs in the spirit of Wurmbrand (2014) and as such, they also give rise to *that-trace* effects.

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 (56) 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 (84):

(84) Who watched the movie?
    [$_{CP}$ who$_1$ [$_{CP}$ C [$_{TP}$ t$_1$ [$_{TP}$ T [$_{vP}$ t$_1$ [$_{vP}$ watched the movie?]]]]]]
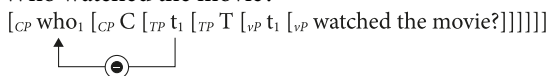
In order to avoid an antilocality violation, like the one in (84), I follow Brillman and Hirsch (2016) who propose that *wh*-subjects in matrix clauses raise as far as TP and never to CP. Since English is a residual V2 language, movement to the CP domain is usually accompanied by T-to-C movement. It 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 (85):

(85) *which person did a friend of watch the movie?
    [$_{CP}$ which person$_1$ [$_{CP}$ did$_C$ [$_{TP}$ [$_{DP}$ t$_1$[$_{DP}$ a friend of t$_1$]] [$_{TP}$ …]]]]]
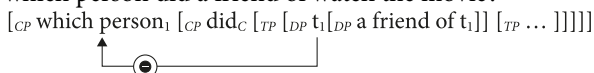
Here in (85) [which person] has undergone movement to CP and thus *do*-support is present. The sentence is, nevertheless, ungrammatical because antilocality is violated. Given the asymmetry between (84) and (85) with respect to *do*-support, I conclude that matrix subject questions target the edge of TP.[35] 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) illustrated in (86), and extraction from subject relative clauses (Chung and McCloskey 1983) in (87):

(86) 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*?

(87) 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

---

35. Even if subject-initial clauses have their subjects in Spec,TP, it may still be the case that a C is present, among others, for clause typing (as standardly assumed these days). I am grateful to a reviewer for making this observation.

adjunct that contains the parasitic gap. However, if in (86a) *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 (87), Chung and McCloskey (1983) claim that subject relatives are weaker islands than object islands. In fact, the *wh*-complement *which* in (87a) 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 (87a) allowing *wh*-elements to pass over it, but *who* has moved to the specifier of CP in (87b) creating an island.

Some recent analyses of matrix subject questions and absence of *do*-support question the Spec,TP analysis. Messick (2020) gives a series of arguments for why this should be the case. However, I would like to note two challenges to such an analysis. The first one is related to parasitic gaps and the latter to his sluicing argument. As already mentioned in this section, matrix subject extaction does not license parasitic gaps as object extraction does. If there is A'-movement of the subject from Spec,*v*P to Spec,CP as Messick argues, we would expect sentences like (86a) and (88) to be grammatical. However, this prediction is not borne out:

(88) a. *Who will ignore you [if you try to talk to *pg*]
      b. *Who met you [before you recognized *pg*]   Chomsky (1986, 54)

With respect to sluicing, Messick argues that since sluicing involves *wh*-movement to Spec,CP to escape deletion, the same must happen in matrix clauses. This is not a convincing argument because the movement path of *wh*-subjects from matrix and embedded clauses is not symmetrical as Messick himself acknowledges in his analysis. For example, in the case of the latter there is evidence that the embedded *wh*-subject has moved over the complementizer from Spec,TP due to the presence of *that-trace* effects; such effects can be bypassed if the derivation undergoes salvation by deletion (Merchant 2001) (89):

(89) Arizaphale said that some demon could save the world from Armaggeddon but I can't remember who ~~Arizaphale said t$_{who}$ that t$_{who}$ could save the world from Armaggeddon~~.

However, there is no such overt evidence that the matrix *wh*-subject has actually moved over the C head overtly either in one fell swoop, as Messick argues for, or in two separate steps. I want to suggest the possibility that the lack of *do*-support and movement to Spec,CP in subject matrix questions follow from acquisitional strategies, such as the Principle of Structural Economy (Westergaard 2009; Westergaard, Lohndal, and Alexiadou 2019) defined in (90):
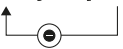
(90)   Principle of Structural Economy
    a.   only build as much structure as there is evidence for in the input.
    b.   only move elements as far as there is evidence for in the input.

According to this economy principle, the subject in matrix subject sentences should only move to T, as there is no evidence in the input available to the child for the subject to move further; just like there is no evidence for A-movement of the subject to target a position higher than T.

That said, let's go back to the claim that matrix *wh*-subjects move as far as TP.[36] If the movement happens between Spec,*v*P and Spec,TP, then this movement should be ruled out by antilocality since no *PL* is crossed. 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 in (91):

(91) $[_{TP} \text{DP}_1 \, [_{TP} \text{T} \, [_{vP} \text{t}_1 \, [_{vP} \text{watched the movie?}]]]]]]$

The DP is on the edge of the *v*P, and it targets the specifier 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 solution to this problem lies in the Principle of Minimal Compliance (PMC):

(92)   **Principle of Minimal Compliance**
    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 the same head H.

This definition of the PMC differs from previous versions proposed by Richards (1998) and Rackowski and Richards (2005).[37] It entails that no matter how close a

---

36. The fact that "*wh*- the hell" subject questions are grammatical remains unexplained under this approach if it is the case that they can only be licensed in A′-positions (see Pesetsky and Torrego 2001; Messick 2020 and references cited therein). I would like to leave this issue for future research.
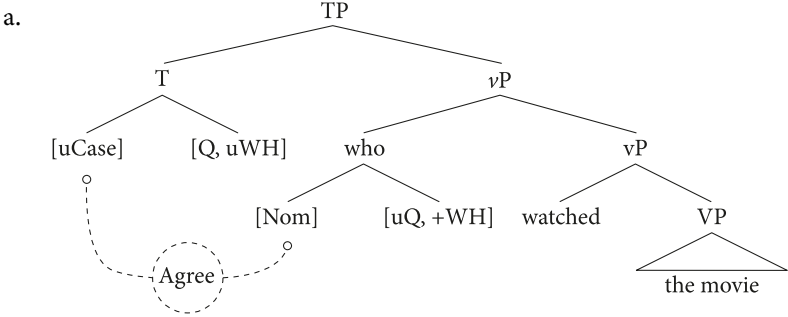
37. The Principle of Minimal Compliance is also different from standard generative constraints. While typically constraints scan the entire chunk of built structure for violations and then reject the structure if this contains any violations, the PMC determines that if one depen-

probe and a goal are, if they enter an Agree relation, the locality restrictions that may exist between them are suspended. In the following subsections, I show how the PMC allows certain movement operations but disallows others.
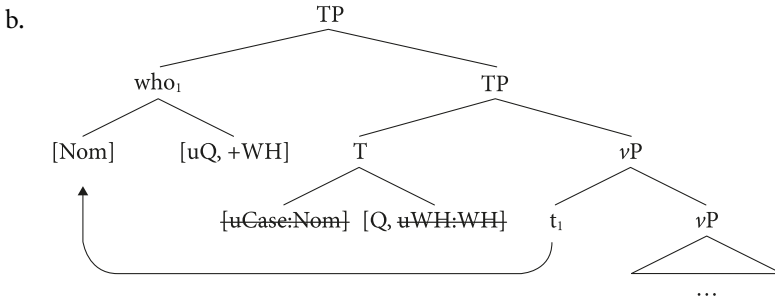
4.2.1 *The PMC and Spec,vP to Spec,TP movement*

Given the definition of the PMC, 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, uWH] features to trigger *wh*-movement of the subject and subsequent *do*-support. Thus, T must be the locus of the [Q, uWH] features. T also bears [uCase:_] that needs to get valued by the [Case: Nom] in the *wh*-subject. As a result, T can get the [uCase:_] valued and the [uWH] feature by entering an Agree relation with the *wh*-element. This is represented in the derivation in (93):

(93)   Spec,*v*P to Spec,TP movement
a.

---

b.

```
                          TP
                 ┌─────────┴─────────┐
               who₁                  TP
           ┌────┴────┐          ┌─────┴─────┐
        [Nom]   [uQ, +WH]       T           vP
                          ┌─────┴─────┐  ┌───┴───┐
                    [uCase:Nom] [Q, uWH:WH]  t₁    vP
                                                  ╱  ╲
                                                    …
```

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 (92), allowing the goal to move to its specifier. It is important to note that a sentence like (85) 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.[38]

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 predic-

---

38. Andrew Simpson (p.c.) wonders how the [Q, uWH] features 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, uWH] features only in matrix subject questions; (ii) T inherits all the features from C via "donate" (Citko 2014, 52) as a result of Feature Inheritance (Richards 2007, Chomsky 2008). This is related to a concern raised by an anonymous reviewer who wonders whether an independent Agree relation between C and the *wh*-element is possible. The assumption that I am making here is that C in matrix subject questions is present only for clause typing purposes. All the features are either originated in or inherited by T. Thus no dependecy should be established between C and the *wh*-element. Alternatively, given that T bears both A and A′ features, it is possible that this is a composite probe that looks for the closest available goal (Van Urk 2015): *who* with [Nom], [uQ, +WH]. As a result, there is only one Agree relation that values both A and A′ features simultaneously.
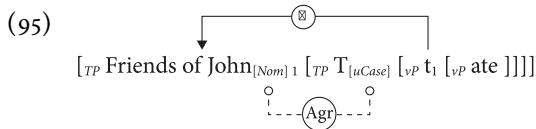
tion is borne out and we take it as evidence that the movement takes place in one fell swoop (94):

(94)   a.   *[Quiénes]$_2$ [todos t$_2$]$_1$ causaron el   año pasado t$_1$ un escándalo?
              who   all   caused   the year past   a   scandal
              'Who caused a scandal last year?'
         b.   [Quienes]$_1$ causaron el   año pasado [todos t$_1$] un escándalo?
              who   all   caused the year   past   a   scandal
              'Who caused a scandal last year?'

4.2.2   *A desirable consequence of this system*

At this point, one desirable consequence of the system that deserves to be empha-sized is how this account handles Spec,*v*P-to-Spec,TP A-movement. This is prob-lematic for previous proposals of antilocality including Erlewine (2016, 2017), Brillman and Hirsch (2016) and Deal (2017) that formulate antilocality in a way that it only applies to A′-movement as illustrated in (53). 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 (95) for A-movement of the subejct:

(95)

$$[_{TP} \text{ Friends of John}_{[Nom]\,1} [_{TP} \text{ T}_{[uCase]} [_{vP} \text{ t}_1 [_{vP} \text{ ate }]]]]$$

(Agr)

As illustrated in (95), T Agrees with the DP *Friends of John* in 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, as other antilocality approaches have done including those mentioned in the previous paragraphs, that subject raising in declarative clauses does not violate antilocality because there are invis-ible functional projections between *v*P and TP that are crossed with this move-ment. 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 (96):

(96) *John friends of watched the movie

$$[_{TP} \text{John}_1 [_{TP} \text{T}_{[uCase]} [_{vP} [_{DP[Nom]} \text{t}_1 [_{DP} \text{friends of t}_1]] [_{vP} \text{ate }]]]]$$
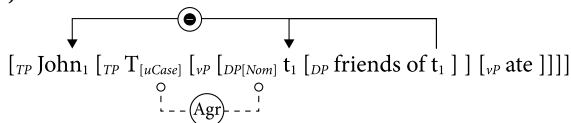
The example in (96) portrays a case in which Agree has been successfully established between the probe in T and the DP goal. This would suspend antilocality between them. However, the DP *John* which has undergone successive cyclic movement through the edge of the containing DP is still subject to antilocality. This is the exact same picture that we will see for indirect objects.

4.2.3   *What about non-that-trace grammars?*

The data reported in the paper are limited to *that-trace* grammars. It is certainly true that there is a great amount of variation when it comes to the effects both in these languages and others (Chomsky and Lasnik 1977; Sobin 1987, 2002; Boeckx 2008 and references cited therein). I would like to note, though, that despite the amount of variation typically reported in the theoretical literature just mentioned, these differences are not really supported by experimental data from English. Experimental research shows that examples containing *that-trace* effects are judged, if not completely ungrammatical, significantly worse than those examples without the effects (Cowart 1997; Chacón et al. 2015; Polinsky et al. 2015; Ritchart, Goodall, and Garellek 2016). If the presence or absence of *that-trace* effects is due to a parametric difference, it is not completely clear what differences exist between particular varieties of, let's say English, that would lead the child to appropriately set the correct parameter. Nevertheless, here I discuss three possibilities that can provide a solution to how *non-that-trace* grammars could be analyzed: the presence of additional structure between CP and TP; (c)overt agreement between C and the subject in Spec,TP; and the variability in the location of the launching site of movement.

The first possible source of variation stems from the amount of structure present in embedded clauses. In other words, assuming that the derived subject occupies the Spec,TP, the CP layer headed by the overt complementizer (which is also a phase head) does not immediately dominate the TP; but instead there is some additional projection sandwiched between CP and TP:

(97) $[_{CP} \textit{wh}_1 \quad \text{that } [_{XP} \ldots [_{TP} \text{t}_1 \text{ T} \ldots]]]$

This configuration will satisfy anti-locality and provide a treatment of the neutralization facts. I suggest Douglas (2017) and Erlewine (2020) for more details on

this. Structurally, this is similar to what Lohndal (2007, 2009) proposes but without the commitment of adopting a cartographic analysis of the left periphery.

Another possibility consists of agreement between the subject and the complementizer. Given the Agree-based approach to unlocking adopted in the paper, it is possible that some languages allow for (c)overt complementizer agreement with the subject in Spec,TP. In these languages, intermediate phase heads will bear A′-probes (Van Urk 2015) that are able to make subject extraction licit via Spec,CP as a result of feature valuation. On the contrary, other languages will have intermediate phase heads that do not participate in Agree relations with moving goals (Bošković 2007). In other words, there is a difference with respect to whether a language allows intermediate phase heads to consist of the same feature material as phase heads for final landing sites (98):

(98)   Intermediate Phase Head Divide
    a. *non-that-trace* effect grammars
        Intermediate phase heads participate in feature checking,
        e.g. $C^{o}_{[uWH]} \rightarrow_{Agree}$ Subject$_{[+WH]}$
    b. *that-trace* effect grammars
        Intermediate phase heads do not participate in feature checking,
        e.g. $C^{o} \rightarrow {}_{*Agree}$ Subject$_{[+WH]}$

Given the definition of the PMC adopted in the paper, the divide in (98a) predicts that, on the one hand, the Agree relation between the probe on C and the goal on the subject is enough to bypass antilocality. On the other hand, even though these languages lack *that-trace* effects for extraction, they should exhibit an asymmetry with subextraction, identical to what was said about indirect objects. This prediction should be checked in future work. However, the divide in (98b) predicts that the lack of Agree makes subject extraction impossible. Thus, the claim that differences in movement operations eventually boil down to the presence vs. absence of feature checking via intermediate successive cyclic movement would not be unprecedented (see Van Urk 2015 and Van Urk and Richards 2015; Bošković 2007).

The last alternative that I want to discuss is the possibility of subject extraction from a position lower than TP. This is precisely the case of Spanish and other Romance languages that can leave the subject in-situ, or in a position lower than Spec,TP. Evidence for the existence of this anti-that-trace strategy comes from the fact that stranding a quantifier in Spec,TP as a result of *wh*-movement of the subject DP is ungrammatical, e.g. (13). Similarly, languages like German make use of this strategy as well, as shown by Bayer and Salzmann (2013) and illustrated in (37) and (38): as long as (sub)extraction takes place from a position lower than the highest specifier of TP, the movement should be licit.

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. Some diagnostics to differentiate high and low applicatives are given in Table 2:[39]

Table 2. Pylkkänen's (2002) test for applicatives

Test	High Appl (e.g. Albanian)	Low Appl (e.g. English)
1. The applied argument is the "beneficiary"	✓	⊖
1′. The applied argument is the "recipient"	⊖	✓
2. Static verbs can be applicativized	✓	⊖
3. Depictives can modify the applicative	✓	⊖

 According to Pylkkänen (2002), in double object constructions in languages like English, the applicative can only bear a transfer of possession relation to the direct object, which translates in the applicative being merged lower than the voice head or *v* head and lower than the 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 (99)–(101):

---

39. For a more detailed overview I refer the reader to McGinnis (2017); Georgala, Paul, and Whitman (2008); Georgala (2010); and Bruening's extensive work on Applicatives (2010a, 2010b, 2018b a.o.).

(99)   English
        Jane did Bill a favor.

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

(101)   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 (99), "Mary" is not the recipient but the beneficiary of John's doing the favor. In
(100), "her brother" is the beneficiary of the event of Carmen's cooking: he ben-
efits 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 respec-
tively. This observation is in line with Larson's (2010) criticism of Pylkkänen's
(2002) semantics for low applicatives. For her, low applicatives must combine ear-
lier 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, illus-
trated by the following pair in (102) with two conjoined structures in which the
same entity "letter" is involved:

(102)   Larson's (2010, 702) semantics for (102a) and (102b).
        a. John wrote that letter and Bill gave Mary that letter.
        b. John wrote Mary that letter.
        c. $\exists e$[writing(e) & Agent(e,John) & Theme(e,that-letter)] & $\exists e'$ [giving(e')
            & Agent(e',Bill) & Theme(e',that-letter) & Goal(e', Mary)]
        d. $\exists e$[writing(e) & Agent(e,John) & Theme(e,that-letter) & Goal (e,Mary)]

According to Larson (2010), "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 (103). $[\![Appl_{Low}]\!]$ takes three arguments: the direct object,
the indirect object and the verb.

(103)   Low Applicative denotation   (Pylkkänen 2002)
        ⟦*Appl_Low*⟧ = λx.λy.λf.λe. f(e,x) & theme(e,x) & to-the-possession (x,y)

(104)   Pylkkänen's (2002) semantics   (taken from Larson 2010, 702)
        a.   ∃*e*[writing(e) & Agent(e,John) & Theme(e,that-letter)] & ∃*e'* [giving(e')
             & Agent(e',Bill) & Theme(e',that-letter) & to-the-possesion-of(that-
             letter,Mary)]
        b.   ∃*e*[writing(e) & Agent(e,John) & Theme(e,that-letter) & to-the-
             possesion-of(that-letter, Mary)]

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

    In addition to this, low applicatives are incompatible with static verbs: given
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 (105): "in an event of *holding*, a bag does not plausibly result in
the bag ending up in somebody's possession".

(105)   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
which according to Pylkkänen (2002) are low applicative languages (106)–(108):[40]

(106)   English
        a.   I will hold you the offer.
        b.   I will hold you a spot.

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

(108)   German   (Georgala 2010, 99)
        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.'

The final argument concerns the availability of depictive adjectives modifying the
applicative argument. Depictives are secondary predicates that refer to a nomi-
nal throughout the duration of the verbal event. The nominal modified by the

---

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

depictive must be a participant in the main event. This entails that, semantically, a depictive must take an individual and an event as arguments (Bruening 2018a). Given the semantics for low applicatives in Pylkkänen (2002) (103), this predicts that languages with low applicatives should never be able to compose with a depictive (109):

(109)   John$_i$ read Mary$_j$ the news naked$_{i/*j}$
    ✓ John was naked;
    ☉ Mary was naked

However, the ban on depictive modification does not hold if the applied argument becomes the subject under passivization (Koizumi 1994; Bruening 2018a) (110):

(110)   Mary$_i$ was read the news naked$_i$
    ✓ Mary was naked

As Bruening (2018a) observes, considering that actives and passives are truth-conditionally identical, there is no semantic restriction that restrains a depictive from modifying an indirect object. Besides, we can conclude that the indirect object must refer to a participant in the main event (e.g. the subject of the passive), contrary to what is expected under Pylkkänen's (2002) syntax-semantics for low applicatives. In fact, there are certain uses of the verb *give* in the active that enable depictive modification of the indirect object. The example in (111) is taken from Bruening (2018a, 548: ex. 29) who credits Maling (2001, 424: ex. 14). A Spanish conuterpart is in (112).[41]

(111)   a.   The nurse$_i$ gave the patient$_j$ the medication still-groggy/half-asleep$_{i/j}$.
    b.   Victorian doctors$_i$ preferred to give their female patients$_j$ a physical exam fully- dressed$_{i/j}$.

---

41. German does not allow depictives to modify the indirect object (McFadden 2006; Georgala 2010; Stefan Keine p.c.).

(iv)   Die   Krankenschwester$_i$ gab Maria$_j$ die   Medizin krank$_{i/*j}$.
    the.NOM nurse   gave Maria the.ACC medicine ill
    'The nurse$_i$ gave Maria$_j$ the medicine ill$_{i/*j}$'   Stefan Keine (p.c.)
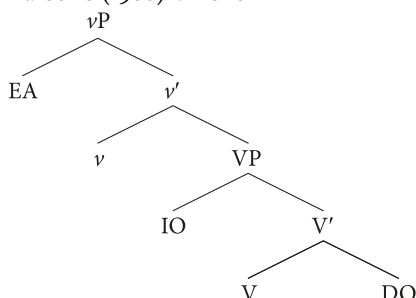
However, even if this test does not hold, the other arguments provided in this section are strong enough to consider that the indirect object in German is generated higher than the VP, as assumed by Pylkkänen (2002)
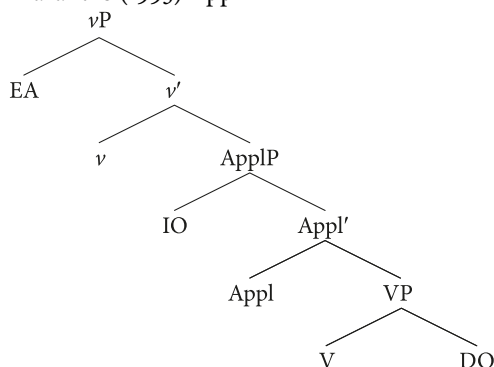
(112) Spanish

La enfermera$_i$ le dio al   abuelo$_j$ su medicación medio
the nurse.FEM CL gave to-the grandfather.MAS his medication half
dormido$_{*i/j}$/ grogui$_{i/j}$
asleep.MAS   grogui
'The nurse$_i$ gave the granfather$_j$ the medication half-asleep$_{*i/j}$ / stillgroggy$_{i/j}$'

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 (113); 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 (114). 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.

(113)   Larson's (1988) VP shell

```
                vP
           ┌─────┴─────┐
          EA           v′
                   ┌────┴────┐
                   v        VP
                        ┌────┴────┐
                       IO         V′
                               ┌──┴──┐
                               V     DO
```

(114)   Marantz's (1993) ApplP

```
                vP
           ┌─────┴─────┐
          EA           v′
                   ┌────┴────┐
                   v       ApplP
                        ┌────┴────┐
                       IO        Appl′
                               ┌──┴──┐
                             Appl    VP
                                  ┌──┴──┐
                                  V     DO
```

4.3.2 *Indirect objects: The hows and whys of extraction but not subextraction*

The second part of the Extraction-Subextraction generalization in (47) 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,ApplP, and the potential landing site in Spec,*v*P. This has an immediate consequence for the The Spec-to-Spec Confinement Generalization in (50). Considering that indirect objects occupy a specifier position that structurally corresponds to the illicit structural description in (49), they should be predicted to be immobile. However, they can move as a whole. Thus, the generalization is too weak as stated in (50) and needs to be revised to render the stronger version in (115) to include the bolded information:

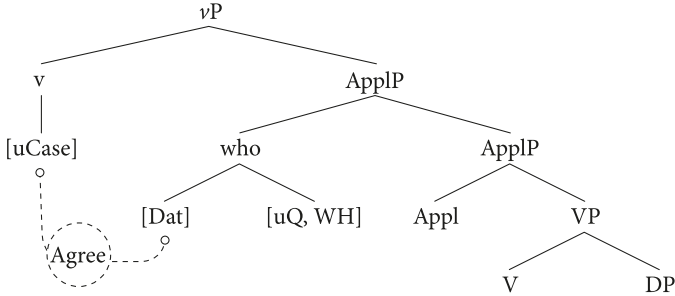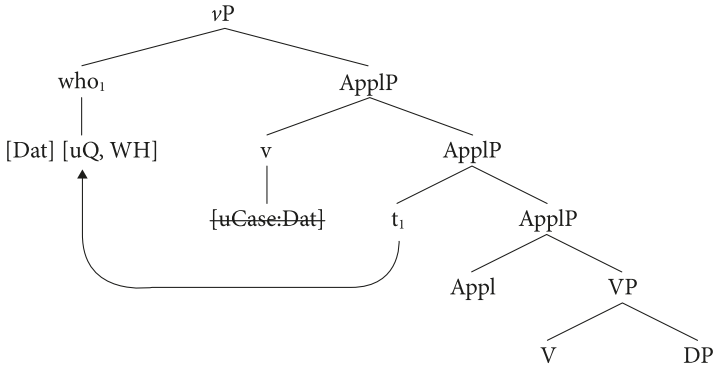(115) **The Spec-to-Spec Confinement Generalization (final version)**
       If X takes YP as a complement and YP has an element ZP in its specifier, two conditions must hold:
       a.   movement of ZP to the specifier of XP is not allowed **unless ZP is an indirect object**; and
       b.   movement out of ZP to the specifier of XP is not allowed.

That is why I have proposed that the asymmetry illustrated by indirect objects is obtained via the PMC: 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.[42] This is illustrated in (116) for extraction and (117) for subextraction with data from English.

---

42. My definition of the PMC in (92) resembles Preminger's (2019) formulation. However, the purpose and the domain of application is different. Preminger's formulation of the PMC would enable movement of the DP after Agree has occurred via merge "unlocking" the ApplP and extracting the DP. If there is Agree between *v* and the DP, the DP would be the one that is "unlocked" making movement of the D head possible. This is not what we want here.

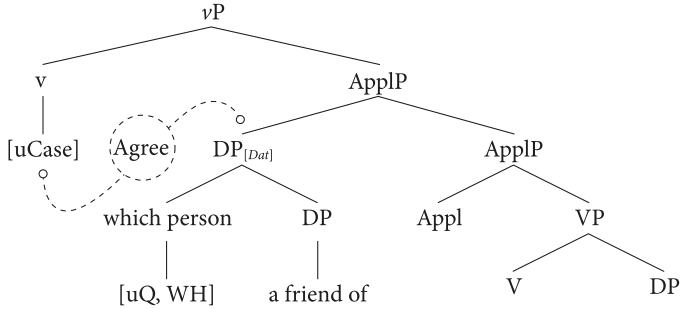(116)   Who did you show the picture?
   a.   Agree(*v,wh*)

vP
 ├─ v
 │   [uCase]
 │   (Agree)
 └─ ApplP
      ├─ who
      │   ├─ [Dat]
      │   └─ [uQ, WH]
      └─ ApplP
           ├─ Appl
           └─ VP
                ├─ V
                └─ DP

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

vP
 ├─ who₁
 │   [Dat] [uQ, WH]
 └─ ApplP
      ├─ v
      │   [uCase:Dat]
      └─ ApplP
           ├─ t₁
           └─ ApplP
                ├─ Appl
                └─ VP
                     ├─ V
                     └─ DP

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 [uQ] feature on the indirect object functions as an EPP feature à la Bošković (2007) triggering successive cyclic movement through phase edges.[43] Thus, antilocality is bypassed and the movement to the phase edge is allowed. This is how we get the extraction facts.

Regarding subextraction, blocking movement out of the indirect object proceeds as illustrated with the derivation in (117):
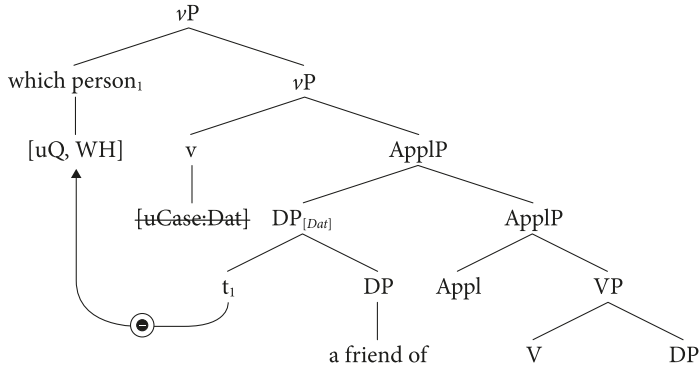
---

43. 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.

(117)   *Which person did you show [a friend of] the picture?

  a.   Agree(*v*, DP)

```
                                vP
              ┌─────────────────┴─────────────────┐
              v                                  ApplP
              │                         ┌──────────┴──────────┐
          [uCase]  (Agree)  DP[Dat]                         ApplP
                   ┌────────┴────────┐              ┌─────────┴─────────┐
              which person          DP            Appl                 VP
                   │                 │                         ┌────────┴────────┐
              [uQ, WH]          a friend of                    V                DP
```

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

```
                              vP
                ┌─────────────┴─────────────┐
          which person₁                    vP
                │                ┌──────────┴──────────┐
            [uQ, WH]             v                    ApplP
                          ┌──────┴──────┐     ┌─────────┴─────────┐
                     [uCase:Dat]  DP[Dat]                       ApplP
                               ┌────┴────┐         ┌──────────────┴──────────┐
                              t₁        DP        Appl                       VP
                                         │                          ┌────────┴────────┐
                                    a friend of                     V                DP
```

In (117), the Agree relation between *v* and the DP-indirect object still takes place because *v* needs to get its [uCase:_] valued with the DP. Nevertheless, what (117b) 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.[44]

Moving onto Spanish and German, it has been argued that active *v* is not the locus of case for the indirect object (Cuervo 2003 for Spanish, and Georgala 2010 for German) given the unavailabilty of recipient passives.[45] Such proposals entail

---

44. The same can be applied to Norwegian.

45. 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 is 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.

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 (118) together with how the PMC and antilocality work:

(118)   Multiple Probes: Agree (v,IO) and Agree (v,DO)
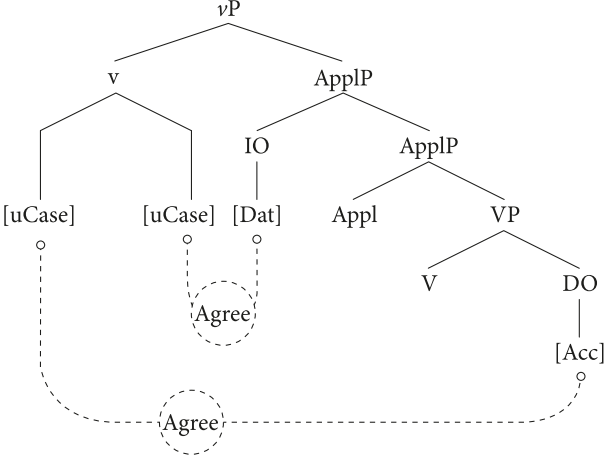
The tree in (118) shows that v undergoes two Agree operations triggered by each of its [uCase:_] probes: one for dative case with the indirect object and one for accusative case with the direct object. For the system to ensure that the correct case is assigned to the correct argument, the features must be ordered (Müller 2010). This means that the highest case feature on the stack has to be the one corresponding to the indirect object. Once that feature is discharged, the case feature responsible for agreeing with the direct object comes to be on top of the stack and allows for the dependecy to be well established. The indirect object has already participated in a feature valuation relation, and thus is not an intervener (Rizzi 1990; Chomsky 1995).

(119)   Correct feature stacking on v

v	IO	DO	v	IO	DO
[uCase:_]$_{dat}$ →	[Dat]	[Acc]	~~[uCase: dat]~~$_{dat}$	[Dat]	[Acc]
[uCase:_]$_{acc}$			[uCase: acc]$_{acc}$		___↑

The opposite ordering would result in an incorrect case assignment due to intervention: the indirect object will always be a closest potential goal because it is higher than the direct object. The [uCase:_]$_{acc}$ feature comes first on the stack and
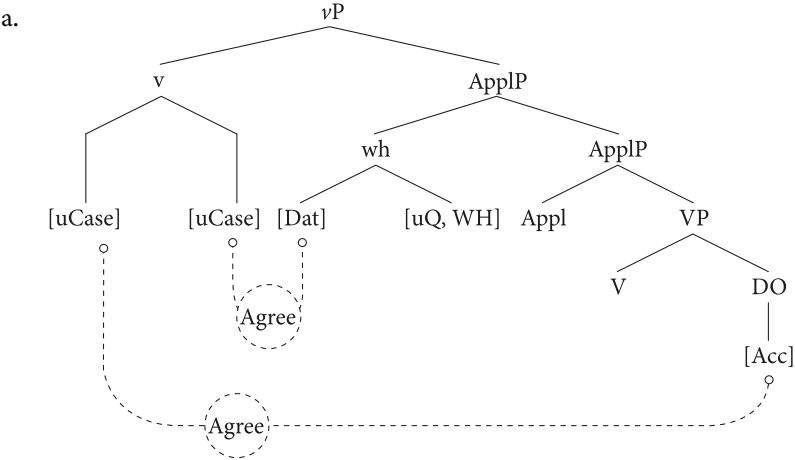
the closest available goal is the indirect object with [Dat] case. The [uCase:_]$_{dat}$ comes next and agrees with the only goal available, i.e. the direct object with [Acc] case. This derivation should crash upon transfer because the features have not been correctly valued. This is illustrated (120):

(120)   Incorrect feature stacking on v

v	IO	DO	v	IO	DO
[uCase:_]$_{acc}$ →	[Dat]	[Acc]	~~[uCase: dat]~~$_{acc}$	[Dat]	[Acc]
[uCase:_]$_{dat}$			[uCase:_]$_{dat}$	_____ ↑	

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 and accusative case; in the case of Spanish, the fact that there is agreement is visible with clitic doubling.[46]

That said, 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. First, there is agreement between the probes on v and the indirect and direct object DPs (121a). As a result, antilocality between v and the indirect object is bypassed and the indirect object can move to the Spec,vP (121b).
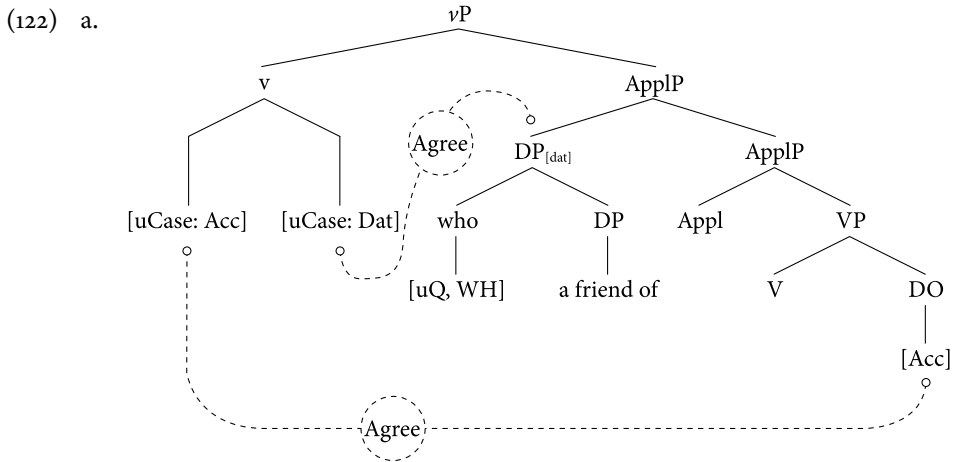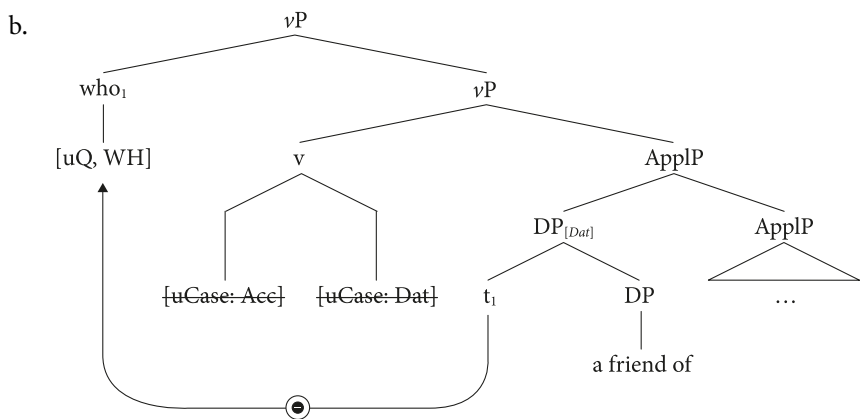
(121)   Step-by-step extraction in IO:
a.

---

46.   See Preminger (2019) for further details.

b.

vP

who₁

[Dat] [uQ, WH]

vP

v

[uCase: Acc]   [uCase: Dat]

ApplP

t₁

ApplP

Appl   VP

V   DO

[Acc]
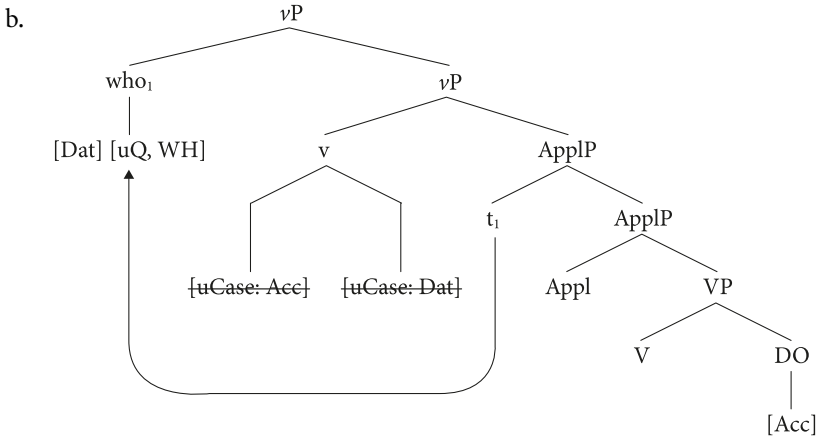
With respect to subextraction, the Agree relations between the probes on *v* and the goals have to be established (122a). However, the *wh*-element in the specifier of the indirect object cannot move to the higher phase edge: it violates antilocality since no independent Agree has been established between the probing element and the *wh*-element (122b).

(122)   a.

vP

v

[uCase: Acc]   [uCase: Dat]

ApplP

Agree

DP[dat]

who

[uQ, WH]

DP

a friend of

ApplP

Appl   VP

V   DO

[Acc]

Agree

b.

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 (47) and (50), revised as (115), repeated below:

(47)   **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.

(117)   **The Spec-to-Spec Confinement Generalization (final version)**
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 last 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 (56) which made a crucial distinction between the individual members of a projection of X and the set of all members of that pro-

jection, 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 "fettering" effects. However, the PMC can intervene and *unfetter* the relevant DP in the form of the Agree. By *unfetter*, 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 bypassed via PMC.

I also extracted the novel observation that there is a hierarchy of opacity with respect to arguments which I labeled the *Argument Islandhood Hierarchy* in (48) repeated below:

(48) **Argument Islandhood Hierarchy**
Indirect objects are stronger islands than *v*P subjects; which are stronger islands than direct objects:
IO > Subjects > DO

While it is well known that extraction out of objects is typically more permissible than extraction out of subjects, a fact that this paper has contributed to support, it is also the case that extraction out of indirect objects is even more degraded than extraction out of subjects. If the subject is in-situ or within the right periphery of the clause (*v*P or the extended *v*P), extraction tends to be acceptable. Nevertheless, extraction out of indirect objects is never possible, at least in the light of data presented here. I propose that the opacity effects follow from the structural position that both subjects and indirect objects occupy with respect to phase heads. In-situ subjects are located at the phase edge and do not have to move through an intermediate landing site satisfying antilocality. In the case of direct objects, even though they are not at a phase edge (unless they undergo object shift or some sort of scrambling), their movement path is long enough to satisfy antilocality.

On the other hand, indirect objects are not located at the phase edge, but they are one projection below the edge of the phase. What is more, there is a parallelism between indirect objects and derived subjects in Spec,TP: they are also located one projection below the edge of the phase and are strong islands for extraction. This is illustrated in (123) and (124). When there is no CP layer, this structural condition is not met and thus movement out of the derived subject is possible.

(123)   Derived subjects

```
            CP
           /  \
              C′
             /  \
            C    TP
                /  \
         Subj_derived   T′
                      /____\
                        …
```

(124)   Indirect objects

```
            vP
           /  \
    Subj_in-situ   v′
              /  \
             v    ApplP
                 /   \
                IO    Appl′
                     /____\
                       …
```
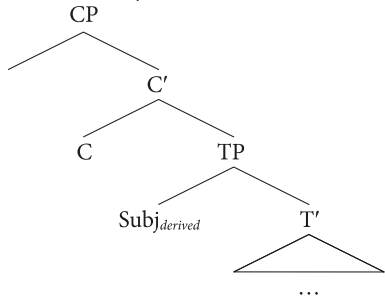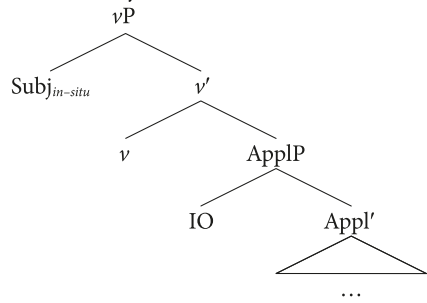
The paper has also looked at applied arguments. One of the challenges for the proposal was the low applicative hypothesis (Pylkkänen 2002). If the hypothesis is correct and the applicative argument is generated lower in the verbal domain, this would be fatal for the proposal put forth here. Nevertheless, I have provided some arguments, building on work by Larson (2010), Georgala (2010) and Bruening (2018a) that have demonstrated that applicative arguments actually occupy a higher position than what was originally argued 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:

(125)   **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.[47]

Therefore, the data and the analysis presented here have served to provide a formal explanation for three novel generalizations: the "Extraction-Subextraction Generalization", the "The Spec-to-Spec Confinement Generalization" and the "Argument Islandhood Hierarchy". In the next section, before I conclude the paper, I provide some alternative analyses 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, 2020) *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

---

47. Carol-Rose Little (p.c.) points out that Ch'ol, a Mayan language of Southern Mexico, disallows extraction out of applied arguments, which are considered to be high applicatives (Little 2020b, 2020a). For example, they allow depictive modification of the applied argument (Vázquez Álvarez 2011, 375–377). Thus, we can take this as evidence that the generalization regarding the opacity of the indirect object also holds outside the range of languages probed in this paper.

(v) *Majki$_i$ ta' a-choñ-be   karu [i-chich   t$_i$] ?
     who   PFV A2-sell-APPL car   A3-sister
     'whose sister did you sell a car to?'   Little (2020a, 13: ex. 31a)

I am incredibly grateful to Carol-Rose Little for this example and these references, and for discussion of the extraction and subextraction data.

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 member of a chain that is made of more than one link is banned. In other words, in order to not disrupt the uniform status of chains, extraction out of an already created chain is precluded. Based on these conditions, he proposes that a derivation of (126) should proceed as follows:

(126)   *$[_{CP}$ Whose books$_2$ do $[_{TP}$ you think $[_{CP}$ t$_2$ that $[_{TP}$ [t$_2$ [reviews of t$_2$]]$_1$ [$_{vP}$ t$_1$ 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 (126a). The extracted element has to step in the outer specifier of the DP given Takahashi's (1994) Shortest Movement Condition (126b). However, the Uniformity Corollary on Adjunction is violated [whose books] is adjoined to a non-trivial chain, i.e. a chain that is composed of at least more than one link: {reviews of whose books$_{vP}$, reviews of whose books$_{TP}$}.

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 (124): 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.

A reviwer wonders how the proposal fits with Rizzi's (2006) work of 'Freezing'. Here are some thoughts. Rizzi (2006), and subsequent work, proposes that an element is frozen if it lands on a particular position in the structure where the moving element checks or is assinged a scope-discourse interpretive feature, i.e. what he calls "a criterion". Rizzi refers to this as Criterial Freezing:

(127)   Strong Criterial Freezing: A phrase meeting a criterion is frozen in place.

Typically this criterial position is the specifier of a functional projection in the left periphery. Rizzi mostly concentrates on *wh*-island constructions, whereby movement out of a *whether* CP is disallowed because the *wh*-element has already received its scope-discourse properties by moving through the Spec,CP in the embedded clause. If Rizzi's approach to freezing is on the right track, we might want to conclude that there are many more 'Criterial Positions' available in the structure than what originally proposed. And while many are in the left periphery, some are also located in the eventive core, e.g. indirect objects.

One might wonder, though, what is the criterion that the indirect object has met that has caused it to be frozen in place; if, as Rizzi argues, freezing criteria are typically related to scope-discourse properties. On the other hand, one could probably assume that the freezing effect caused by meeting a criterion is parallel to the fettering-immobility effect caused by antilocality proposed here: as an economy condition on movement operations, antilocality bans any movement from position X to position Y if nothing is gained from this operation; similarly, one could say the same about Rizzi's notion of criterion: what would be gained with the movement, if the features/properties that the moving element required have already been satisfied? Thus, in terms of the proposal, both accounts of freezing resort to the economy of the derivation.

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 (128):

(128)   **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 (128) 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 (128) 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 *v*P, 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), (128) 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, 2020) antilocality

Many versions of antilocality have been proposed since Bošković (1997a), including but not limited to, Abels (2003), Bošković (2005), Erlewine's (2016, 2020) and Branan's (2019). Erlewine's, which has already been mentioned in Subsection 4.1, and also Branan's (2019), are 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.

One advantage over Erlewine's earlier versions of antilocality (e.g. 2014, 2016, 2017) 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 "hidden" functional projections between *v*P and TP. Besides, very short movement operations that should, in principle, violate antilocality might not if Agree is involved via the PMC. 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 bypassed 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, 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 bypass 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 *v*P. These movements are driven by the needs of the particular head (i.e. *v*, T,C) to value their [uF] feature on a DP. In short, this version of antilocality is thought of as an Economy Condition that emphasizes that movement must happen only if it is strictly necessary that it does as mentioned in fn.26.

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):

(129)   **Antilocality**
Movement of A targeting B must cross a projection distinct from B where unlabeled projections do not count.

(130)   **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, their extraction is predicted to be impossible.

(131)   a. $[_? [_{DP} \text{IO}_{who}] [_{ApplP} \text{Appl}^o ... [_{VP} ...]]]$
b. $[_{vP} v^o [_? [_{DP} \text{IO}_{who}] [_{ApplP} \text{Appl}^o ... [_{VP} ...]]]]$
c.   *$[_? \text{IO}_{who} [_{vP} v^o [_? t_{who} [_{ApplP} \text{Appl}^o ... [_{VP} ...]]]]]$

When the *wh*-DP indirect object merges with the ApplP, this results in an unlabeled structure since there is no feature agreement between Appl^o and its specifier (129a). When *v* is merged, the structure is as in (131b). If labeling via (130b) has failed, the only way the *?* constituents can receive a label is via movement (130a). Movement of *who* to the specifier of *v*P violates antilocality in (129) 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 (130)

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 (132).

(132) $[_{DP}$ a friend of which person $] \Rightarrow [_{?}$ which person$_1$ $[_{DP}$ 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 (132) 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. 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, one 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.

Acknowledgements

References

Abels, Klaus. 2003. "Successive Cyclicity, Anti-locality, and Adposition Stranding." PhD diss., University of Connecticut.

Abels, Klaus. 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. https://doi.org/10.1075/livy.7.04abe

Alexiadou, Artemis, and Elena Anagnostopoulou. 1998. "Parametrizing Agr: word order, V-movement and EPP-checking." *Natural Language and Linguistic Theory 16* (3): 491–539. https://doi.org/10.1023/A:1006090432389

Baltin, Mark. 2001. "A-movement." In *The handbook of contemporary syntactic theory*, edited by Mark Baltin and Chris Collins, 226–254. Oxford: Blackwell. https://doi.org/10.1002/9780470756416.ch8

Barbosa, Pilar. 2001. "On Inversion in Wh-questions in Romance." In *Subject Inversion in Romance and the Theory of Universal Grammar*, edited by Aafke C. Hulk and Jean-Yves Pollock, 20–59. Oxford University Press.

Bayer, Josef. 2012. "From modal particle to interrogative marker: a study of German *denn*." In *Functional Heads. The Cartography of Syntactic Structures*, edited by Laura Brugé, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro, and Cecilia Poletto, 13–28. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199746736.003.0001

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. https://doi.org/10.1515/9781614510796.275

Beck, Sigrid. 1996. "Wh-Constructions and Transparent Logical Form." PhD diss., Universität Tübingen.

Bentzen, Kristine. 2009. "Subject positions and their interactions with verb movement." *Studia Linguistica 63* (3): 261–291. https://doi.org/10.1111/j.1467-9582.2009.01162.x

Boeckx, Cedric. 2008. *Bare syntax*. Oxford: Oxford University Press.

Boeckx, Cedric, and Norbert Hornstein. 2005. "A Gap in the ECM Paradigm." *Linguistic Inquiry 36* (3): 437–441. https://doi.org/10.1162/0024389054396926

Bonet, Eulália. 1990. "Subjects in Catalan." *MIT Working Papers in Linguistics 13*:1–26.

Bošković, Željko. 1997a. "Coordination, Object Shift, and V-Movement." *Linguistic Inquiry 28* (2): 357–365.

Bošković, Željko. 1997b. *The Syntax of Nonfinite Complementation: An Economy Approach*. 247. Cambridge, Massachusetts: MIT Press.

Bošković, Željko. 2005. "On the locality of left branch extraction and the structure of NP." *Studia Linguistica 59* (1): 1–45. https://doi.org/10.1111/j.1467-9582.2005.00118.x

Bošković, Željko. 2007. "On the Locality and Motivation of Move and Agree: An Even More Minimal Theory." *Linguistic Inquiry 38* (4): 589–644. https://doi.org/10.1162/ling.2007.38.4.589

Bošković, Željko. 2014. "Now I'm a Phase, Now I'm Not a Phase: On the Variability of Phases with Extraction and Ellipsis." *Linguistic Inquiry 45* (1): 27–89. https://doi.org/10.1162/LING_a_00148

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

Bošković, Željko. 2018. "On Movement out of Moved Elements, Labels, and Phases." *Linguistic Inquiry 49* (2): 247–282. https://doi.org/10.1162/ling

Branan, Kenyon. 2019. "Locality and anti-locality: the logic of conflicting requirements." Ms., National University of Singapore.

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. https://doi.org/10.1162/002438998553798

Browning, Marguerite. 1996. "CP Recursion and that-t Effects." *Linguistic Inquiry 27* (2): 237–256.

Bruening, Benjamin. 2010a. "Ditransitive Asymmetries and a Theory of Idiom Formation." *Linguistic Inquiry 41* (3): 519–562. https://doi.org/10.1162/LING_a_00012

Bruening, Benjamin. 2010b. "Double Object Constructions Disguised as Prepositional Datives." *Linguistic Inquiry 41* (1): 281–305.

Bruening, Benjamin. 2018a. "Depictive Secondary Predicates and Small Clause Approaches to Argument Structure." *Linguistic Inquiry 49* (3): 537–559. https://doi.org/10.1162/ling_a_00281

Bruening, Benjamin. 2018b. "Double Object Constructions and Prepositional Dative Constructions are Distinct: A Reply to Ormazabal and Romero 2012." *Linguistic Inquiry 49*:123–150. https://doi.org/10.1162/LING_a_00268

Burzio, Luigi. 1986. *Italian Syntax*. Reidel Publishers. https://doi.org/10.1007/978-94-009-4522-7

Chacón, Dustin Alfonso, Michael Fetters, Margaret Kandel, Eric Pelzl, and Colin Phillips. 2015. "Indirect learning and language variation: Reassessing the *that*-trace effect." Ms., University of Maryland & Yale University.

Chomsky, Noam. 1973. "Conditions on Transformations." In *A Festschrift for Morris Halle*, edited by Stephen Anderson and Paul Kiparsky, 232–286. New York: Holt Rinehart / Winston.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht, The Netherlands: Foris Publications.

Chomsky, Noam. 1986. *Barriers*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 1995. *The minimalist program. 420*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 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.

Chomsky, Noam. 2001a. "Derivation by Phase." In *Ken Hale: A Life in Linguistics*, edited by Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001b. "Derivation by phase." In *Ken Hale: a life in language*, edited by Michael Kenstowicz, 1–52. Cambridge: MIT Press.

Chomsky, Noam. 2008. "On Phases." In *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, edited by Robert Freidin, Carlos Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge: MIT Press. https://doi.org/10.7551/mitpress/9780262062787.003.0007

Chomsky, Noam. 2013. "Problems on Projection." *Lingua 130*:33–49. https://doi.org/10.1016/j.lingua.2012.12.003

Chomsky, Noam, and Howard Lasnik. 1977. "Filters and control." *Linguistic Inquiry 8* (3): 425–504.

Chung, Sandra, and James McCloskey. 1983. "On the interpretation of certain island facts in GPSG." *Linguistic Inquiry 14* (4): 704–713.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. New York: Oxford University Press.

Citko, Barbara. 2014. *Phase Theory: An Introduction*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139644037

Collins, Chris. 2002. "Multiple verb movement in Hoan." *Linguistic Inquiry 33* (1): 1–29. https://doi.org/10.1162/002438902317382161

Cowart, Wayne. 1997. *Experimental syntax: Applying objective methods to sentence judgments*. Thousand Oaks/London/New Delhi: SAGE Publications.

Cuervo, Cristina. 2003. "Datives at Large." PhD diss., Massachusets Institue of Technology.

Culicover, Peter. 1993. "Evidence against ECP accounts of the that-t effect." *Linguistic Inquiry 24* (3): 557–561.

Davis, Colin. 2020a. "Crossing and stranding at edges: On intermediate stranding and phase theory." *Glossa: a journal of general linguistics 5* (1): 1–32. ISSN: 2397-1835. https://doi.org/10.5334/gjgl.854

Davis, Colin. 2020b. "Further insights and inquiries into possessor extraction in English." Colloquim Handout presented at the University of Southern California, Los Angeles. 10-12-20.

Deal, Amy Rose. 2009. "The Origin and Content of Expletives: Evidence from "Selection"." *Syntax 12* (4): 285–323. https://doi.org/10.1111/j.1467-9612.2009.00127.x

Deal, Amy Rose. 2017. "Syntactic ergativity as case discrimination." In *WCCFL 34*, edited by Aaron Kaplan, Abby Kaplan, Miranda McCarvel, and Edward Rubin, 141–150. Cascadilla Proceedings Project.

Diesing, Molly. 1990. "The Syntactic Roots of Semantic Partition." PhD diss., UMASS-Amherst.

Doherty, Cathal. 1993. "Clauses without *that*: The case for bare sentential complementation in English." PhD diss., University of California Santa Cruz.

Douglas, Jamie. 2017. "Unifying the *that*-trace and anti-*that*-trace effects." *Glossa: a journal of general linguistics 2*:1–28. https://doi.org/10.5334/gjgl.312

Emonds, Joseph E. 1976. *A Transformational Approach to English Syntax*. New York, New York: Academic Press.

Engdahl, Elisabet. 1983. "Parasitic Gaps." *Linguistics and Philosophy 6* (1): 5–34. https://doi.org/10.1007/BF00868088

Erlewine, Michael Yoshitaka. 2014. "Anti-locality and Kaqchikel Agent Focus." In *31st West Coast Conference on Formal Linguistics*, edited by Robert Santana-LaBarge, 150–159. Arizona State University, Phoenix: Cascadilla Proceedings Project.

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

Erlewine, Michael Yoshitaka. 2017. "Why the null complementizer is special in complementizertrace 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.

Erlewine, Michael Yoshitaka. 2020. "Anti-locality and subject extraction." *Glossa 5* (1): 1–38. https://doi.org/10.5334/gjgl.1079

Etxepare, Ricardo, and Ángel Gallego. 2020. "No nominative case in Spanish." Ms., IKER-CNRS & Universidad Autónoma de Barcelona.

Fanselow, Gisbert. 2001. "Features, theta-roles, and free constituent order." *Linguistic Inquiry 32* (3): 405–437. https://doi.org/10.1162/002438901750372513

Fox, Danny, and David Pesetsky. 2005. "Cyclic Linearization of Syntactic Structure." *Theoretical Linguistics 31* (1–2): 1–46. https://doi.org/10.1515/thli.2005.31.1-2.1

Franco, Irene. 2012. "Subject requirement, complementizers and optionality." Ms., University of Leiden.

Gallego, Ángel. 2007. "Phase Theory and Parametric Variation." PhD diss., Universidad Autónoma de Barcelona.

Gallego, Ángel. 2010. *Phase Theory*. Amsterdam: John Benjamins. https://doi.org/10.1075/la.152

Gallego, Ángel. 2013. "Object shift in Romance." *Natural Language and Linguistic Theory 2* (2): 409–451. https://doi.org/10.1007/s11049-013-9188-6

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. 2003. "Successive Cyclicity Under (Anti-)Local Considerations." *Syntax 6* (3): 260–312. https://doi.org/10.1111/j.1467-9612.2003.00063.x

Haddican, William. 2010. "Theme-goal ditransitives and Theme passives in British English dialects." *Lingua 120*:2424–2443. https://doi.org/10.1016/j.lingua.2009.11.003

Haddican, William. 2012. "Object movement symmetries in British English dialects: Experimental evidence for a mixed case/locality approach." *Journal of Comparative Germanic Linguistics1 15*:189–212. https://doi.org/10.1007/s10828-012-9051-x

Haddican, William, and Anders Holmberg. 2018. "Object symmetry effects in Germanic: Evidence for the role of case." *Natural Language and Linguistic Theory 37* (1): 91–122. https://doi.org/10.1007/s11049-018-9404-5

Haider, Hubert. 1990. "Topicalization and Other Puzzles of German Syntax." In *Scrambling and Barriers*, edited by Günther Grewendorf and Wolfgang Sternefeld, 93–112. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/la.5.06hai

Haider, Hubert. 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. https://doi.org/10.1002/9781118358733.wbsyncom048

Heck, Fabian, and Malte Zimmermann. 2004. "DPs and Phases." Ms., University of Leipzig and HU Berlin.

Hegarty, Michael. 1991. "Adjunct extraction and chain configuration." PhD diss., MIT.

Hein, Johannes. 2019. "Verb movement and the lack of verb-doubling VPtopicalization in Germanic." Ms. University of Postdam.

Holmberg, Anders. 1986. "Word Order and Syntactic Features." PhD diss., Stockholm.

Holmberg, Anders. 2000. "Scandinavian Stylistic Fronting: how any category can become an expletive." *Linguistic Inquiry 31* (3): 445–483. https://doi.org/10.1162/002438900554406

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. https://doi.org/10.1162/ling_a_00322

Huang, C.-T. James. 1982. "Logical relations in Chinese 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. https://doi.org/10.1007/BF00134751

Johnson, Kyle. 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.

Keine, Stefan. 2016. "Probes and their horizons." PhD diss., Ph.D. dissertation, University of Massachusetts Amherst.

Keine, Stefan. 2019. "Selective opacity." *Linguistic Inquiry 50* (1): 13–62. https://doi.org/10.1162/ling_a_00299

Kiziak, Tanja. 2007. "Long extraction or parenthetical insertion? Evidence from judgement studies." In *Parentheticals*, edited by Nicole Dehé and Yordanka Kavalova, 121–144. Amsterdam: John Benjamins. https://doi.org/10.1075/la.106.08kiz

Kiziak, Tanja. 2010. *Extraction Asymmetries: Experimental Evidence from German.* Amsterdam: John Benjamins. https://doi.org/10.1075/la.163

Koizumi, Masatoshi. 1994. "Secondary Predicates." *Journal of East Asian Linguistics 3* (1): 25–79. https://doi.org/10.1007/BF01733149

Larson, Richard. 1988. "On the double object construction." *Linguistic Inquiry 19* (3): 335–392.

Larson, Richard. 2010. "On Pykkänen's Semantics for Low Applicatives." *Linguistic Inquiry 41* (4): 701–704. https://doi.org/10.1162/LING_a_00020

Lasnik, Howard. 1995. "Case and Expletives Revisited: On Greed and Other Human Failings." *Linguistic Inquiry 26* (4): 615–634.

Lasnik, Howard. 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. https://doi.org/10.1007/978-94-010-0991-1_5

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. https://doi.org/10.1162/ling.2003.34.3.506

Little, Carol-Rose. 2020a. "Left branch extraction, object shift, and freezing effects in Tumbalá Ch'ol." *Glossa: a journal of general linguistics 5*:1–29. https://doi.org/10.5334/gjgl.988

Little, Carol-Rose. 2020b. "Mutual dependencies of nominal and clausal syntax in Ch'ol." PhD diss., Cornell University.

Lohndal, Terje. 2007. "That-t in Scandinavian and Elsewhere: Variation in the Position of C." *Working Papers in Scandinavian Syntax 79*:47–73.

Lohndal, Terje. 2009. "Comp-t effects: Variation in the position and features of C." *Studia Linguistica 63*:204–232. https://doi.org/10.1111/j.1467-9582.2009.01159.x

Lohndal, Terje. 2011. "Freezing effects and objects." *Journal of Linguistics 47*:163–199. https://doi.org/10.1017/S0022226710000010

Maling, Joan. 2001. "Dative: The Heterogeneity of the Mapping among Morphological Case, Grammatical Functions, and Thematic Roles." *Lingua 111*:419–464. https://doi.org/10.1016/S0024-3841(00)00039-5

Marantz, Alec. 1993. "Implications of Asymmetries in Double Object Constructions." In *Theoretical Aspects of Bantu Grammar 1*, edited by Sam A. Mchombo, 113–151. Stanford University: CSLI Publications.

Matushansky, Ora. 2005. "Going through a phase." In *Perspectives on phases*, edited by Martha McGinnis and Norvin Richards, 151–81. Cambridge, MA: MIT Working Papers in Linguisitcs.

McFadden, Thomas. 2006. "German inherent datives and argument structure." In *Datives and Other Cases: Between argument structure and event structure*, edited by Daniel Hole, André Meinunger, and Werner Abraham, 49–77. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.75.05mcf

McGinnis, Martha. 2017. "Applicatives." In *The Wiley Blackwell Companion to Syntax*, 2nd, edited by Martin Everaert and Henk van Riemsdijk, 1–32. Amsterdam: John Wiley & Sons. https://doi.org/10.1002/9781118358733.wbsyncom124

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. https://doi.org/10.1075/slcs.75.09mci

Merchant, Jason. 2001. *The syntax of silence: sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.

Merchant, Jason. 2013. "Voice and Ellipsis." *Linguistic Inquiry 44* (1): 77–108. https://doi.org/10.1162/LING_a_00120

Messick, Troy. 2020. "The derivation of highest subject questions and the nature of the EPP." *Glossa: a journal of general linguistics 5*:1–12. https://doi.org/10.5334/gjgl.1029

Müller, Gereon. 1998. *Incomplete Category Fronting: A Derivational Approach to Remnant Movement in German*. Dordrecht ; Boston ; London: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-017-1864-6

Müller, Gereon. 2010. "On Deriving CED Effects from the PIC." *Linguistic Inquiry 41* (1): 35–82. https://doi.org/10.1162/ling.2010.41.1.35

Nissenbaum, Jon. 2000. "Covert movement and parasitic gaps." In *Proceedings of the North East Linguistic Society*, edited by Masako Hirotani, 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. 1968. "Deep and surface constraints in syntax." PhD diss., Cambridge, MA: Massachussetts Institute of Technology.

Pesetsky, David. 1995. *Zero syntax. 351*. Cambridge, Massachusetts: MIT Press.

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. https://doi.org/10.1002/9781118358733.wbsyncom108

Pesetsky, David, and Esther Torrego. 2001. "T-to-C Movement: Causes and Consequences." In *Ken Hale: A life in language*, edited by Michael Kenstowicz, 355–426. MIT Press.

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 Wyngærd, 279–310. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/la.191.10pla

Polinsky, Maria, Carlos Gómez Gallo, Peter Graff, Ekaterina Kravtchenko, Adam Milton Morgan, and Sturgeon Anne. 2015. "Subject islands are different." In *Experimental syntax and island effects*, edited by Jon Sprouse, 286–309. Cambridge: Cambridge University Press.

Pollock, Jean-Yves. 1989. "Verb Movement, UG and the Structure of IP." *Linguistic Inquiry 20* (3): 365–424.

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. https://doi.org/10.5334/gjgl.315

Pylkkänen, Liina. 2002. "Introducing Arguments." PhD diss., Massachusetts Institute of Technology.

Rackowski, Andrea, and Norvin Richards. 2005. "Phase Edge and Extraction: A Tagalog Case Study." *Linguistic Inquiry 36* (4): 565–599. https://doi.org/10.1162/002438905774464368

Reis, Marga. 1995. "Extractions from Verb-Second Clauses in German?" In *On Extraction and Extraposition in German*, edited by Uli Lutz and Jürgen Pafel, 45–88. Amsterdam: John Benjamins. https://doi.org/10.1075/la.11.03rei

Richards, Marc. 2007. "On feature inheritance." *Linguistic Inquiry 38*:563–572. https://doi.org/10.1162/ling.2007.38.3.563

Richards, Marc, and Theresa Biberauer. 2006. "True Optionality: When the grammar doesn't mind." In *Minimalist Essays*, edited by Cedric Boeckx, 35–67. Amsterdam: John Benjamins. https://doi.org/10.1075/la.91.08bib

Richards, Norvin. 1997. "Competition and disjoint reference." *Linguistic Inquiry 28* (1): 178–187.

Richards, Norvin. 1998. "The principle of minimal compliance." *Linguistic Inquiry 29* (4): 599–629. https://doi.org/10.1162/002438998553897

Ritchart, Amanda, Grant Goodall, and Mark Garellek. 2016. "Prosody and the That-Trace Effect: An Experimental Study." In *33rd West Coast Conference on Formal Linguistics*, edited by Kim Kyeong-min, Pocholo Umbal, Trevor Block, Queenie Chan, Tanie Cheng, Kelli Finney, Mara Katz, Shopie Nickel-Thompson, and Lisa Shorten, 320–328. Somerville, MA: Cascadilla.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht, Holland: Foris Publications. https://doi.org/10.1515/9783110883718

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

Rizzi, Luigi. 1997. "The fine structure of the left periphery." In *Elements of grammar*, edited by Liliane Haegeman, 281–337. Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-011-5420-8_7

Rizzi, Luigi. 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.

Roberts, Ian. 2010. *Agreement and head movement: clitics, incorporation, and defective goals*. Cambridge: MIT Press. https://doi.org/10.7551/mitpress/9780262014304.001.0001

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.

Siewierska, Anna, and Willem Hollman. 2007. "Ditransitive clauses in English with special reference to Lancashire dialect." In *Structural-functional studies in English grammar*, edited by Mike Hannay and Gerard Steen, 83–102. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.83.06sie

Sobin, Nicholas. 1987. "The variable status of Comp-trace phenomena." *Natural Language and Linguistic Theory 5* (1): 33–60. https://doi.org/10.1007/BF00161867

Sobin, Nicholas. 2002. "The Comp-trace effect, the adverb effect, and minimal CP." *Journal of Linguistics 38*:527–560. https://doi.org/10.1017/S0022226702001652

Suñer, Margarita. 1994. "V-Movement and the licensing of argumental whphrases in Spanish." *Natural Language and Linguistic Theory* 12:335–372. https://doi.org/10.1007/BF00993148

Svenonius, Peter. 1994. "Dependent Nexus: Subordinate Predication Structures in English and the Scandinavian Languages." PhD diss., University of California at Santa Cruz.

Svenonius, Peter. 2001. "On Object Shift, Scrambling, and the PIC." *MIT Working Papers in Linguistics* 39:267–289.

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 Spanish 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. https://doi.org/10.1016/j.lingua.2006.06.003

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.

Van Urk, Coppe, and Norvin Richards. 2015. "Two Components of Long-Distance Extraction: Successive Cyclicity in Dinka." *Linguistic Inquiry* 46:113–155. https://doi.org/10.1162/LING_a_00177

Vázquez Álvarez, Juan Jesús. 2011. "A grammar of Chol, a Mayan language." PhD diss., Austin, TX: University of Texas Austin.

Vikner, Sten. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages. 294.* Oxford: Oxford University Press.

Vikner, Sten. 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.

Webelhuth, Gert. 1992. *Principles and Parameters of Syntactic Saturation.* Oxford: Oxford University Press.

Westergaard, Marit. 2009. "Microvariation as diachrony: A view from acquisition." *Journal of Comparative Germanic Linguistics* 12 (1): 49–79. https://doi.org/10.1007/s10828-009-9025-9

Westergaard, Marit, Terje Lohndal, and Artemis Alexiadou. 2019. "The asymmetric nature of V2: Evidence from learner languages." In *The Sign of the V – Papers in Honour of Sten Vikner*, edited by Ken Ramshøj Christensen, Henrik Jørensen, and Johanna Wood, 709–733. Aarhus: AU-TRYK: Aarhus University. https://doi.org/10.7146/aul.348.121

Wexler, Ken, and Peter Culicover. 1980. *Formal Principles of Language Acquisition.* Cambridge, MA: MIT Press.

Wiklund, Anna-Lena, Gunnar Hrafn Hrafnbjargarson, Kristine Bentzen, and Thorbjörg Hróarsdóttir. 2007. "Rethinking Scandinavian verb movement." *Journal of Comparative Germanic Linguistics* 10 (3): 203–233. https://doi.org/10.1007/s10828-007-9014-9

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

Wurmbrand, Susi. 2014. "The Merge Condition: A syntactic approach to selection." In *Minimalism and Beyond. Radicalizing the interfaces*, edited by Peter Kosta, Steven Franks, Teodora Radeva-Bork, and Lilia Schürcks, 130–166. Amsterdam: John Benjamins. https://doi.org/10.1075/lfab.11.06wur

Zubizarreta, Maria Luisa. 1997. *Word order, Prosody and Focus*. Cambridge, Massachusetts: MIT Press.

Address for correspondence

Luis Miguel Toquero-Pérez
Department of Linguistics
Univeristy of Southern California
3601 Watt Way, Grace Ford Salvatori 301
Los Angeles, CA 90089
United States

toquerop@usc.edu
https://orcid.org/0000-0002-9049-8150

Publication history