What the PCC tells us about "abstract" agreement, head movement, and locality

Omer Preminger*
December 2016

1. Introduction

It is exceedingly common in contemporary linguistic theorizing to come across claims along the following lines: "It may appear that verbs in language *L* do not agree with their arguments, but that is just an arbitrary fact about the morpho-phonology of *L*. In other words, the relevant exponents in *L* just happen to lack segmental content. Syntactically, the relevant agreement relation is operative in *L* just as it would be in a morphologically richer language." I will refer to this type of analysis as *abstract agreement*, by analogy with abstract case (Chomsky 1981, Vergnaud 1977). For a representative example of an analysis resorting to abstract agreement, see Chomsky (2000:123ff.) on supposed agreement between *v* and the direct object in English.

The primary goal of this paper is to show that when it comes to agreement in φ-features (PERSON, NUMBER, GENDER/NOUN-CLASS), this type of reasoning is almost always mistaken. I will show that, generally speaking, there is no such thing as abstract agreement; φ-feature agreement is only there when you can see it.

The conclusion that there is generally no such thing as abstract agreement would have seemed incongruous not too many years ago, due to the following two widely held premises:

(1) PREVIOUSLY HELD PREMISES CONCERNING CASE & AGREEMENT
   a. PREMISE 1: structural case is assigned as a consequence of agreement in φ-features (Chomsky 2000, 2001)
   b. PREMISE 2: noun phrases that are not assigned inherent case must receive structural case, or else ungrammaticality arises (Chomsky 1981, Vergnaud 1977)

Taken together, these premises entailed that any noun phrase that could not plausibly be analyzed as a bearer of inherent case had to be the target of an agreement relation; and insofar as there was no morpho-phonological evidence of such a relation (as is the case with, e.g., direct objects in English), abstract agreement had to be at play.

By now, however, it has become quite clear that neither of these premises is correct.[1] The case-theoretic underpinnings of abstract agreement have therefore faded away, rendering its existence ripe for re-evaluation.

*A heartfelt thank you to Ümit Atlamaz, Vera Gribanova, Boris Harizanov, Norbert Hornstein, Ruth Kramer, Howard Lasnik, Theodore Levin, Jeff Lidz, David Pesetsky, Carolina Petersen, and Maria Polinsky, to audiences at UMD's S-Lab, at NELS 46 (Concordia), at ZAS (Berlin), and at the Stanford Workshop on Head Movement, and to the participants in the fall 2016 seminar on Head Movement, Noun Incorporation, and Clitic Doubling at UMD, for comments and helpful discussion. All errors are my own.

[1] It is not the purpose of the current paper to rehash the arguments against (1a) and against (1b), arguments which have already been presented in detail elsewhere. On (1a), see Preminger (2011a,b, 2014) (building on Bittner & Hale 1996 and Bobaljik 2008). On (1b), see Preminger (2011b) and Kornfilt & Preminger (2015) (building on Marantz 1991 and McFadden 2004). For the remainder of this paper, I will consider it an established fact that (1a) and (1b) are both false.

In the course of this investigation, we will encounter one notable exception to the generalization that there is no abstract agreement. This exception involves the agreement relation that prefigures clitic doubling. Following Rezac (2008a) and Roberts (2010), *a.o.*, I take clitic doubling to be an instance of syntactic movement. We will see that syntactic movement is not generally dependent on a prior agreement relation (contra Chomsky 2000, *et seq.*; see section 6.2). To explain why clitic doubling in particular does require a prior agreement relation, I will offer a novel perspective on the interaction of locality, head movement, phrasal movement, and the Principle of Minimal Compliance (Richards 1998, 2001).

The aforementioned clitic-doubling exception also means that the ban on abstract agreement is unlikely to have the status of a steadfast principle of grammar. I will show that these facts may instead arise by way of a conservative acquisition strategy with respect to the placement of unvalued φ-features on functional heads.

2. A note on terminology

In the context of this paper, the term 'agreement' refers to transmission of φ-feature values (PERSON, NUMBER, GENDER/NOUN-CLASS) from a noun phrase to a functional head.

Recent years have seen a flurry of reductions-to-agreement: attempts to reduce various other linguistic phenomena to the same formal operation hypothesized to underpin agreement. Examples include: Binding Theory and fake indexicals (Kratzer 2009, Reuland 2011, Rooryck & Vanden Wyngaerd 2011), negative concord (Zeijlstra 2004, 2008b), modal concord (Zeijlstra 2008a), noun-modifier concord (Baker 2008, Carstens 2000, Mallen 1997), and the formation of in-situ questions (Bobaljik & Wurmbrand 2014). These reductions are not the primary focus of this paper; though if the paper's conclusions are correct, it casts considerable doubt on the veracity of some of these reductions—in particular, those that avail themselves of abstract agreement in φ-features (see also section 8).

3. A short primer on the Person Case Constraint (PCC)

The argument I will present in sections 4–5 against abstract agreement is based on the distribution of Person Case Constraint (PCC) effects. Therefore, I will begin with a primer on the PCC itself.

The PCC (also known as the "**me-lui* Constraint") is a family of restrictions governing the PERSON features of different arguments in relation to one another, usually affecting combinations of multiple internal arguments of a single predicate. It is therefore most commonly attested with ditransitive verbs. The constraint comes in at least four varieties: *Strong*, *Weak*, *Me-First*, and *UltraStrong* (see Anagnostopoulou 2005, Nevins 2007 for discussion). I will focus here on the Strong PCC, as exemplified in Basque. At a first approximation, the effect can be defined as follows:

(2) STRONG PCC IN BASQUE: (first approximation)
In finite clauses, the direct object of a ditransitive verb must be 3rd person.

The consequences of (2) are illustrated in (3a–b):

(3) a. Zuk niri liburu-a saldu d-i-ϕ-da-zu (Basque)
you.ERG me.DAT book-ART$_{sg}$(ABS) sell 3.ABS-$\sqrt{}$-sg.ABS-1sg.DAT-2sg.ERG
'You have sold the book to me.'

b. * Zuk harakin-ari ni saldu n-(a)i-ϕ-o-zu
you.ERG butcher-ART$_{sg}$.DAT me(ABS) sell 1.ABS-√-sg.ABS-3sg.DAT-2sg.ERG
'You have sold me to the butcher.' [*Laka 1996*]

As (3a–b) already illustrate, the PCC is fundamentally asymmetric: it restricts the PERSON features of the direct object in the presence of an indirect object, but there is no corresponding restriction limiting the PERSON features of the indirect object in the presence of a direct object. The PCC is asymmetric in another way: it restricts only the PERSON features of the relevant argument—and not, for example, its NUMBER features. As Nevins (2011:944) puts it, there is no Number Case Constraint.

Another noteworthy property of the PCC is that it seems to only arise when there is overt morphology reflecting φ-agreement with the relevant arguments. This has led to the rather widespread view that the PCC is a morphological filter (see, e.g., Bonet 1991, 1994). I return to this point in section 5.

4. The PCC is syntactic

As noted above, the PCC is often thought of as a morphological effect. However, Albizu (1997) and Rezac (2008b) show that it is in fact syntactic in nature. They do so by examining two-place unaccusatives in Basque: verbs that take an absolutive argument and a dative argument, but no ergative argument (what Rezac refers to as *applicative unaccusatives*). It turns out that there are two classes of these verbs in Basque: those where the dative argument c-commands its absolutive co-argument (DAT ≫ ABS), and those where the structural relations are reversed (ABS ≫ DAT). This is demonstrated below using reflexive binding:

(4) DAT ≫ ABS:

a. Kepa-ri bere buru-a gusta-tzen zako
Kepa-DAT his head-ART$_{sg}$(ABS) like-HAB AUX
'Kepa likes himself.'

b. * Kepa bere buru-a-ri gusta-tzen zako
Kepa(ABS) his head-ART$_{sg}$-DAT like-HAB AUX

(5) ABS ≫ DAT:

a. * Kepa-ri bere buru-a ji-ten zako ispilu-a-n
Kepa-DAT his head-ART$_{sg}$(ABS) come-PROG AUX mirror-ART$_{sg}$(ABS)-LOC
Intended: 'Kepa is approaching himself in the mirror.'

b. Miren bere buru-a-ri mintzatu zaio
Miren(ABS) his/her head-ART$_{sg}$-DAT talk-PRT AUX
'Miren talked to herself.' [*Rezac 2008b:75*]

This state of affairs is different from what one finds with true, three-place ditransitives in Basque, which always adhere to an ERG ≫ DAT ≫ ABS structure (Elordieta 2001, Rezac 2008a).

Crucially, only those verbs that show DAT ≫ ABS behavior exhibit PCC effects. Compare (6) and (7b):

(6) ABS ≫ DAT verb:

Ni Peru-ri hurbildu na-tzai-ϕ-o
me(ABS) Peru-DAT approach 1.ABS-√-sg.ABS-3sg.DAT
'I approached Peru.'

(7) DAT ≫ ABS verb:

a. Miren-i gozoki-ak gusta-tzen ϕ-zai-zki-o
Miren-DAT sweet-ART$_{pl}$(ABS) like-IMPF 3.ABS-√-pl.ABS-3sg.DAT
'Miren likes candy.'

b. */?? Ni Miren-i gusta-tzen na-tzai-ϕ-o
me(ABS) Miren-DAT like-IMPF 1.ABS-√-sg.ABS-3sg.DAT
'Miren likes me.'

[*Albizu 1997:21, Rezac 2008b:73*]

The putative target form of the finite auxiliary in the ungrammatical (7b) is identical to the one in the grammatical (6). This is not merely phonological identity, but morphosyntactic identity: the two express the same set of features, {1sg.ABS, 3sg.DAT}.

These facts have several consequences. First, they show that at least in Basque, the effect in (2) (the Strong PCC in ditransitives) is actually a subcase of a slightly broader pattern:

(8) STRONG PCC IN BASQUE: (revised version)
In those finite clauses that have a DAT argument located higher than the ABS argument, the ABS argument must be 3rd person.

Because ditransitives in Basque always adhere to a ERG ≫ DAT ≫ ABS structural hierarchy, the effect described in (2) is derivable as a special case of (8).

Second, these data show that the effect in question cannot be a morphological filter: the putative auxiliary form in (7b) is morphologically identical to the one in (6). Moreover, insofar as there is a meaningful distinction between syntax and morphology, the finer hierarchical organization of arguments relative to one another is the purview of syntax, not morphology.[2,3]

How, then, does the PCC arise in syntax? And why is it sensitive to the structural hierarchy of internal arguments, in the manner shown above (i.e., DAT ≫ ABS versus ABS ≫ DAT)? A body of work by Anagnostopoulou (2003, 2005), Béjar & Rezac (2003), and others, has already provided the answer to these questions. Here, I adopt the system presented in Béjar & Rezac (2003); for the

[2]We could, of course, endow one of the datives in (6–7) with a diacritic that is missing on the other, and grant morphology access to this diacritic in evaluating PCC violations. But since there are no actual differences in the morphology between the two types of datives, this would amount to a restatement of the problem faced by morphological analyses of the PCC, not a solution to it. And it would make the correlation with structural asymmetries (4–5) accidental (cf. the syntactic analysis, surveyed in the text).

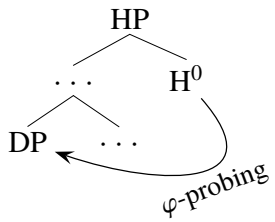
An alternative would be to grant morphology access to finer structural distinctions of the sort shown in (4–5). It seems to me, however, that this would stand in rather blatant violation of the point of modularizing the grammar in the first place. We could therefore rephrase the point being made in the text as follows: either the PCC is syntactic in nature, or else there is no meaningful distinction between syntax and morphology *qua* grammatical modules, in which case we could still say that the PCC is syntactic without any loss of generality.

[3]The same results all but rule out accounts of the PCC in terms of usage-based grammaticalization (cf. Haspelmath 2004). That is because, as evinced by (6), the target form in (7b) is in no way missing from the grammatical vocabulary of the language.

sake of completeness, the remainder of this section provides a sketch of how their system derives the PCC syntactically, and in a way that is sensitive to structural hierarchy.

Let us first consider monotransitives and unaccusative intransitives—configurations where there is only one, non-oblique internal argument.[4] In such a configuration, a probe seeking valued φ-features will reach the internal argument without impediment:

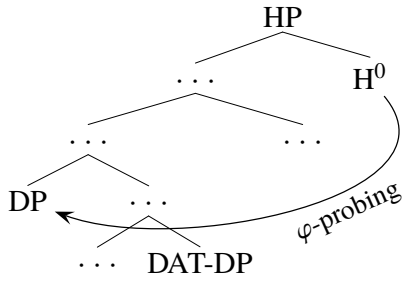
(9)



This will give rise to what one would typically call "object agreement morphology."

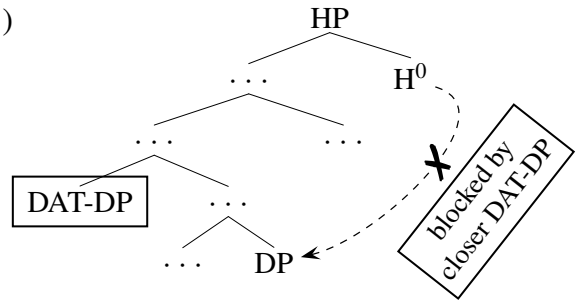
Adding a dative co-argument to (9) located *lower* than the other DP will not affect φ-probing, since minimality dictates that the closer of the two will be targeted:

(10)



However, adding a dative co-argument to (9) that is located *higher* than the other DP will result in intervention, and the disruption of φ-probing:

(11)



(On dative intervention, as well as the general inability of datives nominals to satisfy φ-probing, see Preminger 2014:129–175 and references therein.)

We now add one final ingredient, the *Person Licensing Condition* (Preminger 2011b:930–934; cf. Baker 2008, Béjar & Rezac 2003, Bruening 2005, Nichols 2001, *a.o.*):

[4]The term 'oblique' is used here as a purely descriptive means to refer to any nominal that is either enclosed in a PP, or marked with a case other than {absolutive, nominative, ergative, accusative}.

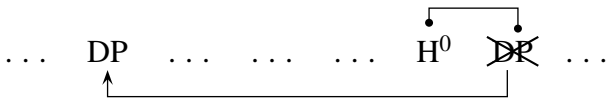
(12) PERSON LICENSING CONDITION (PLC):[5,6]
A [participant] feature on a DP that is a viable agreement target (as far as its case, etc.), and for which there is a clausemate PERSON probe, must participate in a valuation relation.

This predicts that a 1st/2nd person non-oblique internal argument in a DAT ≫ ABS structure (11) will be illicit, because its [participant] feature will fail to participate in a valuation relation with the φ-probe. In contrast, a non-oblique internal argument will be licit *regardless* of its person features in the other two structures (9–10)—because in those cases, the φ-probe can access this argument.

The result is precisely the effect exemplified using the two types of two-place unaccusatives in in (4–7) (as summarized in (8)).

Before concluding, let me address a putative alternative view of these results. Suppose that we were to concede everything stated in this section so far, but insist that the Person Licensing Condition was itself a morphological filter: the requirement would not be that [participant] on a DP participate in a (syntactic) valuation relation—but rather that, if the verb is capable of reflecting agreement morphology, 1st/2nd person morphology on a DP must be reflected on the verb. The factors determining whether agreement with the non-oblique DP did or did not occur would still be syntactic, but the constraint responsible for the PCC would operate in the morphological component.

The problem with such a view is, again, that it flies in the face of what the separation of syntax from morphology is supposed to accomplish, in the first place. Note that the PLC can be satisfied at arbitrary linear and structural distance, as far as the representation handed off from syntax to morphology is concerned; that's because the nominal targeted by the φ-probe can subsequently undergo movement to an arbitrarily distant position:

(13) 

While it has become somewhat commonplace to allow morphology to traffic in objects like 'chains', 'copies', 'traces', etc. (cf. Bobaljik 2008, Marantz 1991), the view taken in this paper is different. Insofar as there is *any* meaningful distinction between syntax and morphology, relations that can cross arbitrary linear and structural distances are strictly the purview of syntax (see also fn. 2).

Consequently, it is safe to say that there is no adequate implementation of the PLC that is truly morphological. And because the PLC is necessary to derive the kind of hierarchy-sensitive PCC effects surveyed here, there is no viable account of the PCC that is truly morphological, either.

[5]The feature [participant] is what distinguishes local (i.e., 1st/2nd person) pronouns from all other nominal expressions. See Harley & Ritter (2002:486–488), and references therein, for discussion.

[6]There is a tantalizingly simpler version of the PLC, which has been put forth in the literature:

(i) A [participant] feature on a DP must participate in a valuation relation.
[*Baker 2008:126–150, Béjar & Rezac 2003:53, a.o.; cf. Nichols 2001:525–526*]

On why this simpler formulation is in the end inadequate, see Preminger 2011b:928ff., as well as fn. 10.

5. The sensitivity of the PCC to the overtness of agreement morphology, and its consequences

As briefly noted in section 3, the PCC is famously restricted to environments in which overt φ-feature agreement with internal arguments is found.[7]

As an example, consider Hebrew ditransitives. In Hebrew, when the dative argument precedes the accusative(=non-oblique) one, the dative is the hierarchically higher of the two arguments (Landau 1994, Preminger 2006, *a.o.*).[8] This is demonstrated in (14), where gender is used to disambiguate the antecedent of the bound reflexive:

(14) DAT ≫ ACC ...

ha-mehapnet-et ta-cig la-cofe et acmo (Hebrew)
the-hypnotist-F FUT.3sg.F-introduce DAT.the-spectator.M ACC REFL.M
'The (female) hypnotist will introduce the (male) spectator to himself.'
(*lit.* 'The (female) hypnotist will introduce [to the (male) spectator] [himself].')

Thus, dative-first ditransitives in Hebrew show the same hierarchical order of internal arguments as their Basque counterparts in (3). However, Hebrew lacks overt agreement with internal arguments; accordingly, no PCC effects arise, as shown in (15). (There is, of course, overt φ-agreement with the subject in (15); but that is irrelevant to the distribution of the internal-argument PCC.)

(15) ... BUT NO PCC

ha-menahel-et ta-cig la-hem oti
the-manager-F FUT.3sg.F-introduce DAT.the-them ACC.me
'The (female) manager will introduce me to them.'
(*lit.* 'The (female) manager will introduce [to them] [me].')

This is an example of how the PCC covaries with the presence of overt agreement morphology cross-linguistically.[9] But the same is true *intra*-linguistically, as well. Consider the PCC-violating Basque sentence in (3b), repeated here:

(3) b. * Zuk harakin-ari ni saldu n-(a)i-ϕ-o-zu (Basque)
you.ERG butcher-ART$_{sg}$.DAT me(ABS) sell 1.ABS-√-sg.ABS-3sg.DAT-2sg.ERG
'You have sold me to the butcher.'

[7] This idealizes the data in one important respect: it collapses φ-agreement and clitic doubling. This distinction, and its consequences for the present discussion, are the topic of section 6.

[8] In accusative-first ditransitives, the converse is the case: the accusative argument is hierarchically higher than the dative one. See Landau (1994) and Preminger (2006) for discussion.

[9] It has been claimed that some languages that lack overt agreement with internal arguments—notably, English and Swiss German—nevertheless show PCC effects when only weak pronouns are involved (Bonet 1991, Haspelmath 2004; see also Anagnostopoulou 2008). It is worth noting, however, that the judgments in these languages that purport to differentiate sentences with 1st/2nd person weak-pronoun Themes from their 3rd person counterparts are quite subtle (as is sometimes acknowledged in the literature on this topic). This is nothing like real PCC effects, which give rise to judgments that are much more robust. Moreover, in a real PCC language, if strong-pronoun Themes can be realized without Differential Object Marking then the PCC will not be sensitive to the strong vs. weak pronoun distinction (cf. Basque). Languages like English and Swiss German lack Differential Object Marking—and so, if the effect in question were a real PCC effect, one would not expect it to be sensitive to whether the pronouns are strong or weak, in the first place. That such sensitivity nevertheless arises suggests that this effect is likely of a different nature. I therefore set it aside for the purposes of the present discussion.

In non-finite environments in Basque, including nominalizations, there is no finite agreement morphology and, in particular, overt φ-agreement with the internal arguments. If we take the very same verb as in (3b), with the very same combination of arguments, and place it in such an environment, the effects of the PCC disappear:

(16) Gaizki irudi-tzen ϕ-zai-ϕ-t [zuk ni harakin-ari
wrong look-IMPF 3.ABS-√-sg.ABS-1sg.DAT you.ERG me(ABS) butcher-ART$_{sg}$.DAT
sal-tze-a]
sold-NMZ-ART$_{sg}$(ABS)
'It seems wrong to me for you to sell me to the butcher.' [*Laka 1996*]

Note that this is a fact about overt agreement morphology, not a fact about finiteness. We can see this by comparing Basque to Spanish, for example. Spanish also has overt agreement with the internal arguments (or more precisely, it has internal argument clitics; see section 6 for further discussion). But unlike Basque, Spanish infinitives retain this agreement morphology. Accordingly, in Spanish, the PCC persists even in infinitives:

(17) a. * Juan me los recomendó (Spanish; *Rodrigo Ranero, p.c.*)
Juan CL$_{1sg}$ CL$_{3pl}$ recommend.PAST
'Juan recommended me to them.'
(okay as: 'Juan recommended them to me.')

b. * Recomendár-me-los es una sorpresa
recommend.INF-CL$_{1sg}$-CL$_{3pl}$ COP DET$_{Fsg}$ surprise
'Recommending me to them is a surprise.'
(okay as: 'Recommending them to me is a surprise.')

In sum, the PCC comes and goes together with the presence of overt φ-agreement with internal arguments; and this is true both cross- and intra-linguistically.

Let us now juxtapose these facts with the results in section 4. As noted earlier, this sensitivity of the PCC to overtness was a central motivation for the view of the PCC as a morphological filter; but the results of section 4 render such an approach untenable. What we have in the PCC, then, is a syntactic effect par excellence, which nevertheless only arises in the presence of overt agreement morphology. This begs the question: *How can something in narrow syntax be sensitive to the overtness of agreement morphology?*

As best I can tell, the only possible answer that maintains the modularity of syntax vs. morpho-phonology is that the mechanisms of agreement and dative intervention, which are implicated in the PCC, are only in place when we can see them. To put this another way, there is no such thing as "null agreement." Importantly, this refers to agreement that is null across the entire paradigm; there is of course no prohibition against particular cells being null in what is otherwise an overt paradigm, as the PCC still arises when agreement with internal arguments has some null cells but is otherwise overt.

Thus, the PCC goes away in the absence of overt agreement morphology—e.g. in Hebrew, and in non-finite clauses in Basque—not because it is a morphological filter. (We already saw in section 4 that it cannot be a morphological filter.) It goes away because in the absence of overt agreement morphology, there is simply no agreement there in the syntax. Not even "abstract" agreement.

I will refer to this as the no-null-agreement generalization:[10]

(18) THE NO-NULL-AGREEMENT GENERALIZATION
There is no such thing as morpho-phonologically undetectable φ-feature agreement.

I label this a generalization rather than a principle, for reasons that will become apparent in section 7.

6. The clitic-doubling caveat

In describing the distribution of PCC effects in section 5, I abstracted away from an important detail: the distinction between φ-agreement in the narrow sense, and clitic doubling. To accurately capture the intra- and cross-linguistic distribution of PCC effects, we need clitic doubling (and/or syntactic cliticization) of internal arguments to also count as "overt agreement morphology." To see why this is an issue at all, let us first review the central properties of the two kinds of relations.

6.1. Some background on clitic doubling

The term φ-agreement refers to a valuation relation between a functional head H^0 and DP, as the result of which the φ-feature values associated with the interpretation of the DP ([participant], [plural], etc.) come to be expressed on H^0. Agreement morphology that arises in this manner is the spellout of valued features on a functional head. There is therefore no particular reason to expect that the exponents of these features will resemble the free-standing pronouns of the language.[11] Moreover, it is possible for these exponents to exhibit allomorphy, and even suppletion, based on the (other) features of the head H^0 (see Arregi & Nevins 2008, 2012, *a.o.*). A widespread example of the latter would be the agreement exponents in one tense/aspect configuration differing from those found in another tense/aspect configuration.

Clitic doubling refers to the occurrence of a D^0-like morpheme, which is φ-feature-matched to the doubled DP, and appears alongside an appropriate host. As such, doubled clitics should not exhibit allomorphy based on the features of their host. Furthermore, we may expect that at least in some cases, doubled clitics will bear morpho-phonological resemblance to the free-standing pronouns of the language. Note: I restrict the use of the term *clitic doubling* to those languages and constructions where the full noun phrase is in argument position, and the relation between the clitic and the full noun phrase exhibits at least some properties characteristic of syntactic movement (see Anagnostopoulou 2006, to appear, for a review).

As an example of clitic doubling, consider the Basque sentence in (19):

[10] The fact that there is no null agreement is another reason (alongside those discussed in Preminger 2011b:928ff.) why we cannot say that any [participant] feature, wherever it may occur, must be licensed by agreement (cf. the formulation of the Person Licensing Condition given in (12); see also fn. 6). There are plenty of environments where 1st/2nd person pronouns can appear and not be targeted by overt φ-agreement (objects in languages without object agreement; complements of prepositions in languages where prepositions do not agree; or most any environment in languages that lack overt φ-agreement altogether). If there can be no null agreement, then we cannot say that 1st/2nd person pronouns in such environments are licensed by agreement.

[11] Diachronically, φ-agreement in the narrow sense often develops from clitics, which themselves often develop from free-standing pronouns. Consequently, it is *possible* for the forms in question to retain their resemblance. The point here is merely that once the synchronic grammar of the speakers involves φ-agreement rather than clitic doubling, there is no longer any principled reason to expect such similarity. Indeed, it is possible that sound changes associated affecting doubled clitics could serve as a catalyst for the diachronic reanalysis of clitics into φ-agreement in the narrow sense.

(19) Guraso-e-k ni-ri belarritako ederr-ak erosi
parent(s)-ART$_{pl}$-ERG me-DAT earring(s) beautiful-ART$_{pl}$(ABS) bought
d-i-zki-da-te.
3.ABS-√-pl.ABS-1sg.DAT-3pl.ERG
'(My) parents have bought me beautiful earrings.' [*Laka 1996*]

The underlined morphemes in the finite auxiliary complex are clitics, doubling the corresponding full noun phrases (Arregi & Nevins 2008, 2012, Preminger 2009).

There are several points concerning clitic doubling that merit mention, at this juncture. The first point is that clitic doubling is not, generally speaking, optional; nor is it conditioned by nominal properties like *animacy*, *definiteness*, and/or *specificity*, in the general case. Clitic doubling in (19), for example, is entirely obligatory, irrespective of the properties of the doubled nominals. In languages where clitic doubling appears to be conditioned by such nominal properties—e.g. Porteño Spanish (20a–b)—it is likely not the clitic-doubling operation itself that is sensitive to these properties. Instead, these properties regulate movement of the full noun phrase into a position from which clitic doubling is then both possible and obligatory (Diesing 1992, Merchant 2006, Sportiche 1998, *a.o.*).

(20) a. La$_{\mathbf{i}}$ oían [a Paca / a la niña / a la gata]$_{\mathbf{i}}$. (Porteño Spanish)
CL hear.PAST.3pl A Paca / A the girl / A the cat
'They listened to Paca / the girl / the cat.'

b. (*La$_{\mathbf{i}}$) buscaban [a alguien que los ayudara]$_{\mathbf{i}}$.
(*CL) search.PAST.3pl A somebody COMP CL.pl help.SBJNCT
'They were looking for somebody who could help them.' [*Suñer 1988:396*]

Importantly, these nominal properties (animacy, definiteness, specificity) are known to regulate movement of noun phrases even in languages that lack clitic doubling entirely (cf. Diesing 1997, Diesing & Jelinek 1993, *a.o.*, on Object Shift). Since the possibility already exists for phrasal movement to be sensitive to these properties, it would be redundant to build this sensitivity into the clitic-doubling operation as well (cf. also indiscriminate obligatory clitic doubling, as in (19)).

The second point concerning clitic doubling is that the doubled noun phrase is known to behave, for the purposes of locality, like traces of A-movement (Anagnostopoulou 2003, *a.o.*), which are known to be non-interveners for other φ-agreement and A-movement operations (Holmberg & Hróarsdóttir 2003, *a.o.*). As an example, consider the status of dative noun phrases in Basque. As shown by Preminger (2009), dative noun phrases that have *not* been clitic-doubled are interveners for φ-agreement in Basque—including agreement in NUMBER—when attempting to target an absolutive DP for agreement:[12]

(21) a. [[Miren-entzat]$_{PP}$ [**harri horiek**]$_{(ABS)}$ altxa-tze-n] probatu [d-**it**-u-zte]$_{aux}$
Miren-BEN **stone(s) those$_{\mathbf{pl}}$ABS** lift-NMZ-LOC attempted 3.ABS-**pl.ABS**-√-3pl.ERG
'They have attempted to lift those stones for Miren.'

[12]The data in (21a–b) are from "substandard" varieties of Basque; see Etxepare (2006:303n2) for discussion.

b. [[Lankide-e-i]$_{\text{DAT}}$ **[liburu horiek]**$_{\text{ABS}}$ irakur-tze-n] probatu
colleague(s)-ART$_{\text{pl}}$-DAT **book(s) those$_{\text{pl}}$(ABS)** read-NMZ-LOC attempted
[d-ϕ/***it**-u-(z)te]$_{\text{aux}}$
3.ABS-**sg/*pl.ABS**-√-3pl.ERG
'They have attempted to read those books to the colleagues.' [*Preminger 2009:640–641*]

In (21a) the matrix finite auxiliary can successfully target the embedded absolutive DP for agreement in NUMBER. However, when the benefactive PP in (21a) is replaced with a bona fide dative DP, as shown in (21b) (and note that, crucially, there is no clitic doubling of dative DPs across clausal boundaries), the same agreement relation is rendered impossible.

Now contrast this state of affairs with what we saw earlier with monoclausal ditransitives in Basque, e.g. (19). There, number agreement with a plural absolutive DP was possible (and, in fact, obligatory), despite the presence of a dative DP. This is not because of the relative positions of the two arguments; the dative argument in ditransitives is systematically higher than the absolutive one (see section 4). The reason number agreement with the absolutive goes through in this case is because the dative DP has been clitic-doubled, rendering it a non-intervener for subsequent probing (note the 1st person clitic *da* on the auxiliary in (19), vs. the absence of any dative clitic whatsoever on the auxiliary in (21b)).[13]

The third point regarding clitic doubling concerns its relation to syntactic *cliticization* (i.e. the occurrence of a clitic that does not seem to double a full noun phrase). As in the case of sensitivity to nominal properties, here too we can reason by appealing to that which we know is independently necessary. Since pro-drop occurs even in languages that lack clitic doubling entirely, let us assume that this is all that (syntactic) cliticization is: clitic doubling of a pro-dropped nominal. In particular, let us assume that it is simply clitic doubling of an unpronounced *pro*.

6.2. Clitic doubling and the PCC

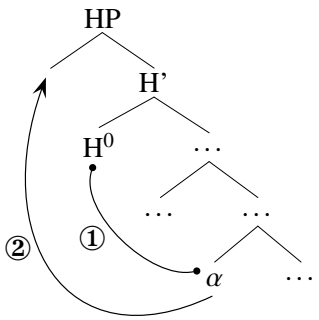
Having briefly surveyed the properties of clitic doubling and how it differs from φ-agreement in the narrow sense, we can now return to the issue that clitic doubling raises for our characterization of the distribution of the PCC. As noted at the outset of section 6, the generalization that the PCC occurs only where overt agreement morphology with internal arguments is found is only correct if the term 'overt agreement morphology' covers both φ-agreement and clitic doubling. The question, now, is why this would be so; after all, the results of section 4 demonstrate quite clearly that what is at issue when it comes to the PCC is the mechanisms of φ-agreement and intervention in syntax. Why, then, would clitic doubling also suffice to give rise to PCC effects?

One answer that we can dismiss quite easily is that clitic doubling "counts" as φ-agreement for the purposes of PCC distribution because clitic doubling *is* φ-agreement. The differences surveyed in section 6.1 show quite clearly that it is not (and see Anagnostopoulou 2006, to appear, and references therein, on further differences between the two phenomena).

[13]The relative timing of PERSON agreement, clitic doubling, and NUMBER agreement plays a crucial role here. PERSON agreement is attempted first, and is blocked by the dative intervener; this gives rise to clitic doubling of the dative; subsequent NUMBER agreement then goes through successfully (due to the clitic-doubled dative having ceased to intervene). The PERSON-specificity of the PCC is one outcome of this interplay; see Béjar & Rezac (2003) and Preminger (2009), *a.o.*, for discussion.

Let us assume, then, that clitic doubling is an instance of movement. This is an explicit part of "Big DP" analyses of clitic doubling (see Arregi & Nevins 2012, Belletti 2005, Cecchetto 2005, Craenenbroeck & Koppen 2008, Torrego 1992, Uriagereka 1995, *a.o.*); several other approaches to clitic doubling include a movement component, as well (see Sportiche 1998; Roberts 2010; Harizanov 2014, *a.o.*). But why would clitic doubling, *qua* movement, behave for the purposes of PCC distribution as though it were agreement? A seemingly promising direction involves the idea that *all* movement is prefigured by a corresponding agreement relation. This view has become very popular in recent syntactic literature, following Chomsky (2000, 2001). On this view, syntactic movement is in fact a two-step process:

(22) ① H^0 enters into an Agree relation with α
→ *subsequently*/*consequently*:
② α moves to (or undergoes Internal Merge in) the domain of H^0, *viz.* in [Spec,HP]



If this were how syntactic movement always works, it would explain why clitic doubling "counts" as agreement as far as the PCC is concerned. The explanation would go as follows: clitic doubling, being a movement operation, invariably requires a prior agreement step; and it is this agreement step that is relevant to the PCC (in the manner described in section 4).

It has become clear, however, that (22) is incorrect—at least as a general requirement. While some instances of movement certainly obey (22), there are other instances that do not. The reader is referred to Preminger (2014:157–175) for a more comprehensive discussion of the issues, but I will mention one clear counterexample to (22) here. In Icelandic, there are double-dissociations between subjecthood and nominative case. In (23), for example, the subject (*einhverjum stúdent* "some student.SG.DAT") is non-nominative, and the nominative (*tölvurnar* "computers.the.PL.NOM") is a non-subject. (See Andrews 1976, Harley 1995, Jónsson 1996, Sigurðsson 1989, Thráinsson 1979, Zaenen, Maling & Thráinsson 1985, among many others, for arguments that this is indeed the correct analysis of an example like (23).)

(23) [Einhverjum stúdent]$_1$ finnast $\boldsymbol{t_1}$ tölvurnar ljótar
some student.SG.DAT find.PL computers.the.PL.NOM ugly
'Some student finds the computers ugly.' [*Holmberg & Hróarsdóttir 2003:999*]

Of particular interest here is movement of the dative *einhverjum stúdent* ("some student.SG.DAT") to subject position. This phrase is not the target of any overt φ-agreement in (23); overt φ-agreement is controlled by the nominative.[14] The standard response to these facts is an appeal to abstract φ-agreement, in an attempt to salvage (22). The idea is that a dative subject like *einhverjum stúdent* ("some student.SG.DAT") is still agreed with, as a precursor to its ultimate movement to

subject position, but this φ-agreement just happens to lack any morpho-phonological content. This, however, is untenable: we have already seen, in section 5, that there can be no such thing as truly abstract (i.e., morpho-phonologically undetectable) φ-agreement.

It is worth noting that Icelandic is perfectly well-behaved with respect to those properties that were the topic of section 5. The PCC covaries with the presence of internal-argument φ-agreement, and Icelandic does not have internal-argument PCC effects (e.g. in ditransitives). Thus, we would not expect it to have φ-agreement with datives, overt or otherwise. There are indeed person restrictions reminiscent of the PCC in Icelandic, but they crucially affect only the nominative argument (see Sigurðsson & Holmberg 2008 for details)—exactly as one would expect if only the nominative was ever targeted for φ-agreement of any kind.

Sentences like (23) therefore provide dispositive evidence against the idea that all movement (or even just all A-movement) involves a prior φ-agreement relation. Thus, the argument that clitic doubling gives rise to PCC effects because it is a movement relation, and all movement involves φ-agreement, fails.

What we are in search of, then, is a reason why clitic doubling constructions—in contrast to movement in the general case—necessarily involve syntactic φ-agreement (which then gives rise to PCC effects).

7. Towards an account of the clitic-doubling caveat

In this section, I propose an explanation for what I have called the *clitic-doubling caveat* (section 6): the fact that clitic doubling must "count" as syntactic φ-agreement, if we are to correctly capture the distribution of the PCC. This, even though clitic doubling is *not* syntactic φ-agreement, nor can it be maintained that it is prefigured by syntactic φ-agreement simply by dint of being a movement operation. Before the explanation, I sketch in sections 7.1–7.2 an account of clitic doubling as long head movement of D^0 out of the doubled DP.

7.1. Clitic doubling as head movement

I follow Rezac (2008a), Roberts (2010), and others, in assuming that clitic doubling is an instance of head movement—specifically, head movement of D. Moreover, this instance of head movement is *non-local*, in the sense that it "skips" at least one c-commanding head in its path, thus violating Travis' (1984) Head Movement Constraint (HMC). To see this, let us consider what it would look like if clitic doubling *did* comply with the HMC. Because the clitic originates in the complement position of the lexical verb, HMC-compliant head movement cannot alter the basic constituent structure given in (24). (Notation: "$\sqrt{}$/V" = the lexical verb root; "D_{CL}" = the clitic.):

(24) {AUXILIARY/TENSE/ASPECT/FINITENESS, {TRANSITIVITY/VOICE, {D_{CL}, $\sqrt{}$/V}}}

[14] This kind of φ-agreement is optional; an alternative is to simply have "default" agreement, i.e., 3rd person singular (see Holmberg & Hróarsdóttir 2003, Sigurðsson & Holmberg 2008). Such 3rd person singular agreement would not, in any case, be agreement with the dative (see Preminger 2014:130-175 for a review of the relevant evidence).

Of course, various elements of (24) may be null in a particular construction or throughout a given language; but if they are overt, (24) is the constituent structure predicted by the HMC.[15] What we actually find, however, does not match this predicted constituent structure:

(25) [L'as]-tu fait? (French)
[CL-have]-you done
'Have you done it?'

Example (25) is an instance of (syntactic) cliticization, rather than clitic doubling per se—but recall that cliticization is assumed to simply be clitic doubling of *pro* (see section 6.1). If anything, cliticization is an even more transparent instance of movement than clitic doubling is, since it does not involve any doubling of overt material; see section 7.2 for further discussion. The actual constituent structure of an example like (25) is the following:

(26) $\{\{D_{CL}, \text{AUXILIARY/TENSE/ASPECT/FINITENESS}\}, \{\surd/V(, t_{D_{CL}})\}\}$ *[cf. (24)]*

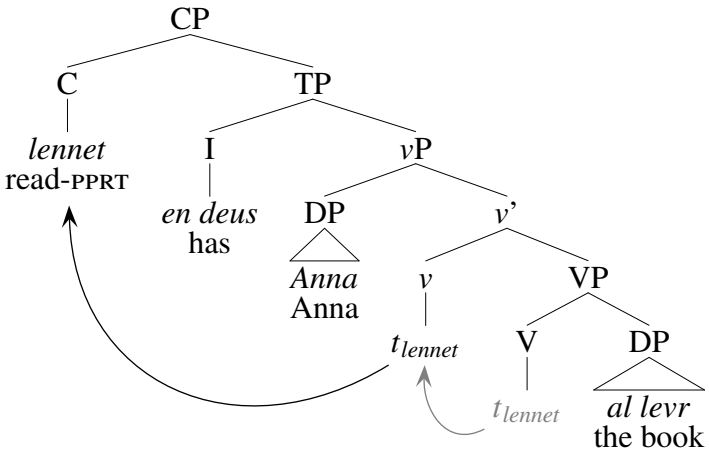
Thus, clitic doubling (or cliticization) viewed as head movement is movement of D at least as far as *v* (hence necessarily skipping over √/V), and often further still. In (25)/(26), for example, we see movement of D to T, skipping over √/V as well as *v*, and possibly other heads (e.g. Asp), too.

At this juncture, one might attempt to maintain the generality of the HMC, taking the incompatibility of (25)/(26) with the HMC as evidence against a head-movement analysis of clitic doubling and/or cliticization. Crucially, however, the HMC has proven inadequate even in empirical domains that do not involve clitics. Verb fronting in Breton (27a–b), for example, also violates the HMC; and yet every applicable diagnostic for distinguishing head movement from (remnant) phrasal movement indicates that it is indeed head movement (see Borsley & Kathol 2000, Borsley, Rivero & Stephens 1996, Jouitteau 2005, Roberts 2004, 2010, Schafer 1994, Stephens 1982):

(27) a. Lenn=a ra Anna al levr. (Breton)
read.INF=PRT does Anna the book
'Anna reads the book.'

b. Lennet en deus Anna al levr.
read-PPRT has Anna the book
'Anna has read the book.'

(28)



[*Roberts 2010:194ff.*]

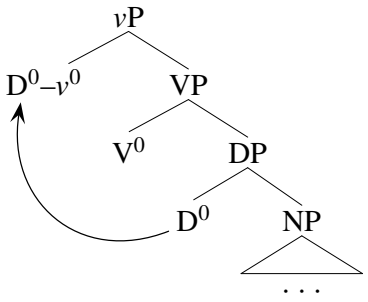
[15] According to the HMC (Travis 1984), heads can only move into the immediately c-commanding head position, and there can be no 'excorporation'—i.e., a subsequent head-movement step would have to pied-pipe the entire complex constituent formed by the previous head-movement step.

For evidence of HMC violations in additional empirical domains, see Harizanov (2016), Lambova (2004), Lema & Rivero (1990), Rivero (1991), Vicente (2007, 2009) and Wilder & Ćavar (1994).

The question is not, then, how (25)/(26) could be so given the HMC, since the HMC is not inviolable to begin with. The question is why the HMC is *often* true, and what it is about cases like (25)/(26) (as well as (27a–b)) that allows them to violate it.[16] I return to these questions in section 7.5.

In the meantime, what we can glean from (25)/(26) is that clitic doubling—including instances of syntactic cliticization—has the following general structure (though the landing site can be higher than v^0, of course):

(29)



7.2. The "double-pronunciation" problem

An immediate question raised by (29) concerns why D^0 is pronounced twice: in cases of clitic doubling, the full noun phrase need not (and often, cannot) be pronounced without its determiner also being pronounced in situ:

(30) Guraso*(-e)-k ni-ri belarritako ederr-ak erosi
parent(s)*(-ART$_{pl}$)-ERG me-DAT earring(s) beautiful-ART$_{pl}$(ABS) bought
d-i-zki-da-te. (Basque)
3.ABS-√-pl.ABS-1sg.DAT-3pl.ERG
'(My) parents have bought me beautiful earrings.' *[=(19)]*

If clitic doubling is movement of D^0, we might expect it to be pronounced only at its landing site, as is the case with many other instances of movement (including head movement).

To begin to understand the reasons for this "double pronunciation" of D^0, let us first note that the form of D^0 when it is a clitic is not always identical to its form when it is a determiner. Much of the early generative work on clitic doubling centered on varieties of Romance, where the two series of forms bear an overwhelming similarity to one another (though there are some instances where Romance determiners do differ from the corresponding clitic form). On the other side of the spectrum, Basque clitics—while similar to the free-standing *pronouns* in the language—are not particularly similar in form to Basque determiners (see Preminger 2009 and references therein). This is what we might expect if these are indeed two instances of the same element, D^0, which occur in two different morphological contexts: as the head of the extended nominal projection, and as a

[16] A related question is why there would be no clitic doubling, or syntactic cliticization, whose landing site is √/V (assuming this is indeed unattested). The view of clitic doubling as head movement allows this to be derived as a special case of Li's (1990) and Baker's (1996) *Proper Head Movement Generalization*, which forbids head movement of a functional category into a lexical category. I thank Theodore Levin for helpful discussion of this point.

clitic adjoined to a verbal head. A particular morphological context may, in a given language, give rise to allomorphy or even suppletion. This is par for the course, e.g., in the relation between finite verb forms and their infinitival counterparts. If finite T^0 (or some feature borne by it, e.g. [PAST]) is a trigger for contextual allomorphy of a given verb root, then the relation between the finite and non-finite forms of that verb will be irregular. But if, for a given verb, T^0 and its features are not an allomorphy trigger, the relation between the finite and non-finite forms of that verb will be morpho-phonologically transparent.

On this view, the hosts of Basque clitics (usually finite auxiliaries, but also a small number of verbs able to carry finite inflection) trigger a great deal of contextual allomorphy in the form of D, resulting in significant differences between the form of pronominal clitics and the corresponding determiners.[17] The hosts of Romance clitics, on the other hand, trigger very little allomorphy of this sort, resulting in a high degree of similarity between the form of clitics and the form of the corresponding determiners. This does not yet account for the double-pronunciation phenomenon; but if the analogy with finite vs. non-finite verbs is apt, we may ask whether double-pronunciation phenomena of this sort are also found in that empirical domain; and the answer is that they are. In what follows, I will consider Landau's (2006) results concerning the fronting of verbs (and verb phrases) in modern Hebrew.

Landau (2006) discusses instances of topicalization in modern Hebrew in which verbs may be fronted, with or without their arguments—henceforth, predicate clefts.[18] Representative examples are given in (31):

(31) a. li-rkod, Gil lo yi-rkod ba-xayim
INF-dance, Gil NEG FUT.3sgM-dance in.the-life
'As for dancing, Gil will never do so.'

b. li-knot et ha-prax-im, hi kant-a
INF-buy ACC the-flower-PL, she PAST.buy-3sgF
'As for buying the flowers, she has done so.' [*Landau 2006:37*]

What is crucial for our present purposes is that, in examples like these, the verb stem (*-rkod* "dance" in (31a), *-knot* "buy" in (31b)) is pronounced twice.

Landau's analysis of this instance of double pronunciation is based on the idea that the pronunciation or omission of each copy in a movement chain is negotiated at PF, and in a highly local manner. He assumes there are two different phonological requirements at play: one demanding a host for the affixes associated with T^0, and one demanding the fronted verb anchor the left edge of the intonational contour associated with predicate clefting in Hebrew. These different phonological requirements each force the pronunciation of a particular copy, resulting in the double-pronunciation effect seen in (31a–b).

I would like to propose a slightly different analysis of facts like (31a–b), one which also extends to the double pronunciation of D^0 in clitic doubling contexts. The analysis is in some sense inspired by the work of Bošković & Nunes (2007) and Nunes (2004) and, in particular, by their focus on the

[17] This is not to be confused with the "tense-invariance" generalization of Arregi & Nevins (2008, 2012) and Nevins (2011): the latter is concerned with contextual allomorphy triggered not by the *identity* of the host (e.g. T vs. N/*n*), but by features *borne by* the host (e.g. [PAST] vs. [PRESENT]). And, as these authors show, the form of Basque clitics indeed shows no allomorphy of the latter kind.

[18] To be precise, certain right-adjoined modifiers may also be fronted in this construction; see Landau 2006:38n9.

mechanics of phonological chain reduction (i.e., the suppression of pronunciation of some copies in a movement chain) as the key to understanding doubling phenomena. Their leading idea is that in order to apply phonological chain reduction, the system must first recognize that the different instances of a single syntactic object are indeed copies of one another, and that this recognition can be obscured under certain circumstances. They focus on cases where one of the two copies occurs within a larger morphosyntactic unit, whose internal structure is not accessible to the linearization algorithm. In these cases, PF cannot identify the two instances as copies of the same object, and phonological chain reduction does not apply.[19]

This cannot be the whole story, though: consider that canonical instances of verb movement to T^0 (e.g. in French) involve morphological merger of the verb with other material, as well—namely, with tense and/or agreement morphology—and yet this does not inhibit phonological reduction of the lower copy (or copies) of the verb in this case. Viewed from this perspective, the question is what sets apart predicate clefts, as well as clitic doubling (viewed as head movement), from more familiar instances of head movement like V/*v*-to-T.

Instead of this morphology-driven approach, I propose that the conditions on phonological chain reduction of head movement are as follows:

(32) CONDITIONS ON PHONOLOGICAL CHAIN REDUCTION OF HEAD MOVEMENT

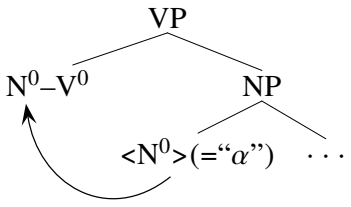
Let X^0 be a head that undergoes movement to Y^0, and let α be the lower copy of X^0. α will be phonologically deleted **iff** either of the following conditions is met:

(i) α and Y^0 are not separated by a phasal maximal projection (incl. XP)

(ii) X and Y are part of the same extended projection (Grimshaw 2000), and Y^0 c-commands α in the surface structure (i.e., no constituent containing α but not Y^0 has undergone subsequent movement to a position above Y^0)

I readily concede that, even if true, (32) is a rather unwieldy beast. However, let us concentrate first on whether or not it is a correct characterization of the facts, starting with instances of maximally local (i.e., HMC-compliant) head movement. Given condition (32.i), whenever XP is not a phase, reduction will apply. That is because, in maximally local head movement, there is no other maximal projection relevant to (32.i). This is as desired; consider, for example, classic cases of noun incorporation (Baker 1988 *et seq.*). Here, the complement of V^0 is NP, which is not a phase, and movement proceeds from N^0 to V^0. This correctly predicts that in this scenario, the lower copy of N^0 will be deleted:

(33)



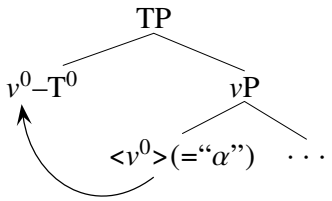
The same applies to movement from T^0 to C^0, and to movement of the verb root (V^0 or $\sqrt{}^0$) to v^0.

In the event that XP *is* a phase, reduction is still predicted to apply, so long as X and Y are part of the same extended projection (in the sense of Grimshaw 2000) and XP has not been moved from

[19] Bošković & Nunes (2007) in fact apply such an analysis to cases of predicate clefting in Vata and in Brazilian Portuguese, cases which are—at least superficially—very similar to (31a).

[Compl,Y]. This is the case, for example, with any instance of v^0-to-T^0 movement that does *not* involve predicate clefting: both v^0 and T^0 are part of the extended verbal projection, and *v*P has not been moved, meaning T^0 still c-commands α (the original position of v^0).

(34)

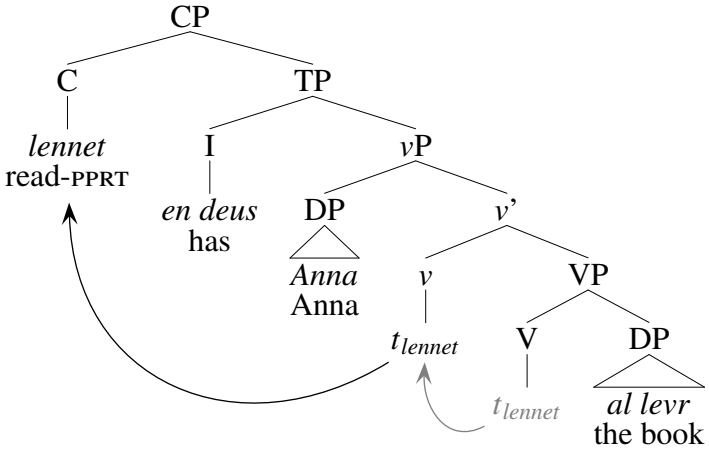


In a language with *v*P-fronting *and* v^0-to-T^0 movement, fronting of the *v*P means not only a violation of (32.i) (which never holds of v^0-to-T^0 movement), but also of (32.ii) (since the lower copy of v^0 will be fronted together with the *v*P, but T^0 will not be). Consequently, neither (32.i) nor (32.ii) holds in this scenario, and the conditions for phonological chain reduction are therefore not met. The result is the double-pronunciation effect, as is the case in modern Hebrew predicate clefting.[20] In a language that lacks v^0-to-T^0 movement, on the other hand, (32) is rendered irrelevant (as there is no v^0-to-T^0 chain to which it could apply). In such a language, *v*P-fronting will result in pronunciation of the verb only within the fronted verb phrase (as is the case, e.g., in English).

Let us now turn to the following scenario: XP is a phase, and X and Y are not part of the same extended projection. Here, there is no way either (32.i) or (32.ii) could be satisfied. This is so regardless of whether the movement in question is maximally local (i.e., HMC-compliant) or not, since adding another projection between YP and XP would not alter either of the relevant factors. This, I argue, is precisely the state of affairs when it comes to clitic doubling: it is head movement of D^0, which is part of the extended nominal projection, to a position outside of DP and within the extended projection of the verb (v^0/T^0/etc.). Given (32), we predict that clitic doubling will always be just that—*doubling*—because phonological reduction will never apply to the lower copy (the one contained in DP).

Finally, let us consider a case of non-local head movement in Breton, the structure of which is repeated below:

(35) *[=(28)]*



[20] See Landau (2006:46–50) for arguments that the fronted category in modern Hebrew predicate clefting is indeed *v*P, rather than VP.

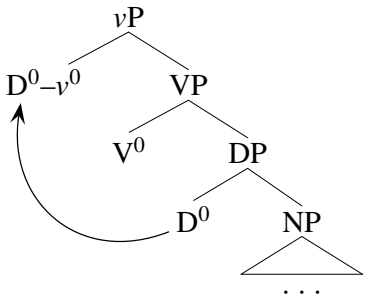
While *v*P qualifies as a phasal maximal projection situated between $Y^0(=C^0)$ and $\alpha(=v^0)$, the two heads are both part of the extended verbal projection, and the c-command relations between them have not been disrupted by subsequent movement. This means that condition (32.ii) is satisfied, and phonological reduction of v^0 is (correctly) predicted to apply.[21]

Let us now return to the theoretical status of the conditions themselves. It would be eminently fair to characterize (32.i–ii) as quite stipulative. I would nevertheless contend that they represent a (modest) step forward in understanding the *doubling* part of clitic doubling relative to existing accounts, which by and large arrive at this result by brute force. The "Big DP" analysis (Arregi & Nevins 2012, Torrego 1992, Uriagereka 1995, *a.o.*), for example, is tailored precisely to achieve this desideratum by base generating an already-doubled structure (in which a clitic and the actual to-be-doubled noun phrase form a constituent, from which the clitic subsequently sub-extracts). The same is clearly also the case for true base-generation approaches to clitic doubling (Sportiche 1998, *a.o.*). While (32) obviously begs for further explanation, it at least captures the behaviors of clitic doubling, predicate clefting, noun incorporation, and more common V/*v*-to-T-type head movement—an array which neither Landau's (2006) nor Bošković & Nunes' (2007) proposals are able to fully capture. For a proposal that could potentially derive (something like) (32) from more basic properties of syntactic movement and morphological composition, see, among others, Gribanova & Harizanov (2016).

7.3. An A-over-A-like effect blocking head movement

Combining the results of section 7.1 and section 7.2, we have in place the essential ingredients of a theory of clitic doubling as long head movement of D^0 out of its containing DP. A sample derivation—in this case, with v^0 serving as the landing site—is repeated below:

(36) *[=(29)]*

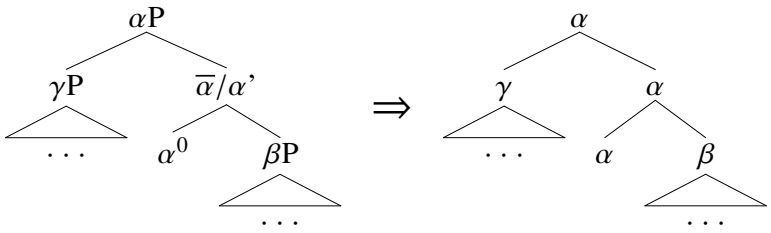


We are now finally in a position to address the central goal of section 6, namely, answering why it is that clitic doubling "counts" as syntactic φ-agreement for the purposes of the PCC. Here, I build on proposals by Hornstein (2009:72–74) and Roberts (2010:33–40). The central idea is that Bare Phrase Structure (Chomsky 1994) and Iterative Downward Search (Kitahara 1994, 1997, Koizumi 1995, Müller 1996, 1998, Takano 1994, *a.o.*) together yield an *A-over-A*-like effect, which under normal circumstances precludes head movement altogether. Crucially, however, we will see that this effect abates under particular conditions.

[21] The movement step between V^0 and v^0 satisfies both (32.i) *and* (32.ii), and so phonological reduction of V^0 is also (correctly) predicted to apply

Let us begin with the contribution of Bare Phrase Structure (BPS). The aspect of BPS relevant here is its conception of projection and, in particular, the fact that non-terminal levels of projection (previously thought of as "αP" and "$\overline{\alpha}$") are now viewed as additional instances of the very same syntactic object that constitutes the head (previously thought of as "α^0"):

(37)



Like many others, I continue to employ the pre-BPS notation for the sake of perspicuity; but the denotatum is a structure that, as far as the grammar is concerned, has the properties characterized on the righthand side of (37). Accordingly, the maximal projection ("αP"), for example, cannot be distinguished from the minimal projection ("α^0") in featural terms. The two are, by hypothesis, one and the same syntactic object, and it is logically impossible for there to be any featural distinction between an object and itself. The two can therefore only be distinguished *relationally*, by inspecting whether a given instance of the object in question dominates and/or is dominated by other instances of the same object (α).

Let us now turn to Iterative Downward Search (IDS). The idea here is that a syntactic probe seeking a viable goal will scan the structure iteratively, using a search algorithm that has at least the following properties (see Kitahara 1994, 1997, Koizumi 1995, Müller 1996, 1998, Takano 1994):[22]

(38) ADEQUACY CONDITIONS ON IDS ALGORITHM

a. If y asymmetrically c-commands x, then the algorithm for IDS will encounter y before it encounters x.

b. If y asymmetrically dominates x, then the algorithm for IDS will encounter y before it encounters x.

An example of an algorithm that meets (38a–b) is given in (39):[23,24]

[22]Definitions:

(i) y asymmetrically c-commands x **iff** y c-commands x and x does not c-command y.

(ii) y asymmetrically dominates x **iff** y dominates x and x does not dominate y.

[23]I assume that there is no actual freedom with respect to the search algorithm employed by the mental grammar. That is, the grammar employs exactly one such algorithm, and what we know about this algorithm is that it meets the conditions in (38a–b).

[24]It might appear that the example algorithm in (39) is categorically unable to return a head (an "α^0") as its output, since all the non-failing halting conditions (the ones that do not say "no goal") involve returning an "αP." But this is illusory; it is an artifact of the pre-BPS notation used in (39). For example, in step (39f), the constituent in [Compl,X] may itself be a head (i.e., non-branching). When the algorithm loops back to step (39b), all it can do is check whether the constituent in question matches the featural search criterion. It cannot determine, using featural means, whether it is an "αP" or an "α^0" (see the discussion of Bare Phrase Structure, above). Thus, if the head in question matches the featural search criterion, it will be returned as the output of the algorithm.

(39) EXAMPLE OF IDS ALGORITHM

a. Let $\mathcal{P}$ be a syntactic probe, and let XP be $\mathcal{P}$'s sister

b. QUERY: Is XP a viable goal? If so, **halt with "XP" as the search result**

c. For every specifier ZP of XP, QUERY: Is ZP a viable goal? If so, **halt with "ZP" as the search result**

d. QUERY: Is XP a phase? If so, **halt with no goal**

e. QUERY: Does X^0 have a complement? If not, **halt with no goal**

f. Return to step **(b)**, using the constituent in [Compl,X] as the new "XP"

Let us also make the following assumption considering the search criterion employed in IDS:

(40) CONDITION ON IDS SEARCH CRITERION
The criterion used to determine whether a given node counts as a viable goal for the probe must be featural.

Consider, now, the combination of (38b) and (40), as well as the consequences of BPS, discussed earlier. Condition (38b) entails that if a head has a projection other than itself, that projection will be encountered before the head. BPS entails that there is no featural basis on which different projections of the same head could be distinguished from one another. And condition (40) states that the criterion for what constitutes a viable goal must be featural. Taken together, the result is that if a head has a projection other than itself, IDS could not possibly yield the head as its search result. In particular, it will never be able to skip a maximal projection but still deem the head of that projection a viable goal, except in the trivial case where the head *is* the maximal projection. Following Hornstein (2009) and Roberts (2010), I will refer to this as an *A-over-A*-like locality condition on IDS. But note that there is no appeal here to a sui generis A-over-A principle; the effect is derived directly from the premises stated above.

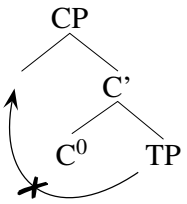
If true, this locality condition would rule out the theory of clitic doubling sketched earlier in this section, which was based on (long) movement of D^0 alone out of its containing DP. In section 7.5, I will suggest that this condition—like other locality conditions in syntax—is subject to the Principle of Minimal Compliance (Richards 1998, 2001), which means that it only holds once for any pair of relata. But before turning to that, I consider another possible response to this state of affairs.

7.4. The complementary locality conditions on head movement and phrasal movement

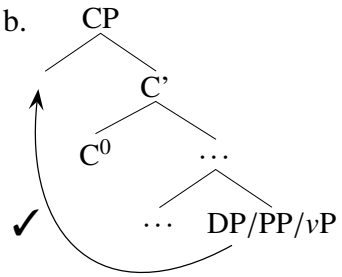
Taken at face value, the A-over-A-like condition identified in section 7.3 seems to rule out head movement altogether (except in the trivial case that the head is also the phrase). For Hornstein (2009), this suggests that head movement might be better modeled as a PF phenomenon, completely outside the purview of syntax. This is a fairly common position concerning head movement (Abels 2003, Brody 2000, Chomsky 1995, among many others), but it is often contested on the grounds that some instances of head movement appear to have interpretive effects (Hartman 2011, Lechner 2006, *a.o.*). Since it is trivially true that head movement affects pronunciation, if it turns out that it also affects semantic interpretation then it must occur in the part of the derivation that feeds both form *and* interpretation, i.e., syntax. In what follows, however, I suggest a different reason why we should be skeptical of attempts to remove head movement from the syntactic component.

As noted in section 7.1, Travis' (1984) Head Movement Constraint (HMC) is counter-exemplified by several kinds of head movement. Nevertheless, it is beyond question that the HMC *often* holds; Emonds (1970, 1976), Travis (1984), and others would not have been able to make the observations they made if this were not the case. Let us contrast this with the state of affairs when it comes to phrasal movement. Here, the literature recognizes a condition known as *anti-locality* (Abels 2003, 2012, Bošković 1994, 1997, Grohmann 2003, Ishii 1997, 1999, Kayne 2005, Murasugi & Saito 1995, Saito & Murasugi 1999). Specifically, there appears to be a ban on phrasal movement that is too local; there is a minimal amount of structural distance that phrasal movement must traverse. For the purposes of the current discussion, I will assume Abels' (2003) version of the constraint, which simply bans movement from the complement position of a given head to the specifier of the same projection (though see section 7.5 for a refinement of this proposal):

(41) a.



b.



What is less often noted, however, is that these two locality conditions—on head movement and on phrasal movement—stand in a complementary relation to one another (or near-complementary, once exceptions to the HMC are considered). The picture that emerges is that phrasal movement *cannot* be maximally local, while head movement (in most cases) *must* be maximally local. One case where this complementarity is explicitly noted is in the work of Pesetsky & Torrego (2001), who assume that it holds without exception, and build it into their Head Movement Generalization:

(42) Head Movement Generalization [*Pesetsky & Torrego 2001:363*]

Suppose a head H attracts a feature of XP as part of a movement operation.

a. If XP is the compl. of H, copy the head of XP into the local domain of H.

b. Otherwise, copy XP into the local domain of H.

In other words, (42) states that if H^0 attracts a feature on XP, then XP will move to [Spec,HP]—unless XP is the sister of H^0, in which case X^0 will head-move to H^0. This is an idealization, in that (42) entails the strict and invariant HMC, which, we have already seen, is not quite right. And even abstracting away from this issue, (42) does not *derive* the complementarity in question, it merely asserts it. In the next sub-section, I will propose a theory of the locality of head movement and its relation to phrasal movement that derives their (near-)complementarity.

What I wish to emphasize here, however, is that if this complementarity of locality conditions is real, it constitutes an argument in and of itself that head movement should remain part of syntax. Modularizing the grammar is vacuous unless different modules make use of different primitives, and access different kinds of information. Phrasal movement clearly resides within syntax proper, since it often has semantic effects as well as phonological ones. It would be quite an odd coincidence, then, if a different operation, situated in a different module (e.g. PF), ended up satisfying complementary conditions to those that phrasal movement satisfies.

Of course, the strength of this argument hinges on the precise nature of this complementarity. If the complementarity is only approximate—as the aforementioned deviations from the HMC might suggest—then the coincidence would be less pronounced, and it would perhaps be less dubious to situate the two types of movement in different modules. I will argue, however, that we can do better. In the next sub-section, I will present a theory for the locality of head movement and its interaction with phrasal movement, which, while allowing certain deviations from the HMC, has Abels-style anti-locality as its consequence. Crucially, the theory in question requires a computation that makes reference to both types of movement, and therefore requires them both to reside in the same computational module.

7.5. Head movement meets the Principle of Minimal Compliance (PMC)

7.5.1. The PMC

Richards (1998, 2001) argues for a principle that regulates the way the grammar enforces syntactic locality constraints in general. Consider, first, the following pair:

(43) a. * [Which book]$_1$ did the journalist spread the rumor that the senator wanted to ban t_1?
b. ? [Which journalist]$_1$ t_1 spread the rumor that the senator wanted to ban [which book]$_2$?

One might be tempted to explain the contrast between (43a) and (43b) in terms of whether or not the lower wh-phrase, *which book*, undergoes movement. The idea would be that (43b) is well-formed because in this example, the wh-phrase in question has not undergone movement out of the Complex NP island. What Richards shows is that such an explanation is, at best, insufficient:

(Bulgarian)

(44) a. * [Koja kniga]$_1$ razprostranjavaše žurnalistât [mâlvata če senatorât iska da
which book spread journalist rumor that senator wanted to
zabrani t_1]?
ban
'[Which book]$_1$ did the journalist spread the rumor that the senator wanted to ban t_1?'

b. ? [Koj žurnalist]$_1$ [koja kniga]$_2$ t_1 razprostranjavaše [mâlvata če senatorât iska
which journalist which book spread rumor that senator wanted
da zabrani t_2]?
to ban
'[Which journalist]$_1$ t_1 spread the rumor that the senator wanted to ban [which book]$_2$?' [*Richards 1998:607*]

In contrast to its English counterpart, the pair-list question in (44b) involves overt movement of both wh-phrases, including the one that originates within the Complex NP island, and yet it is well-formed. It is also not the case that Bulgarian simply lacks the Complex NP Constraint. As (44a) illustrates, such movement is illicit in Bulgarian, too, when not accompanied by movement of a second wh-phrase (cf. the English (43a)).

Richards shows that locality conditions such as Subjacency (or whatever subsumes Subjacency as an explanation of the Complex NP Constraint) need only be satisfied once with respect to a given landing site. Once a single Subjacency-compliant wh-chain has terminated in a given CP periphery, subsequent wh-chains landing in the same position are exempt from the similar locality conditions.

Note that this explanation generalizes to the English data in (43a–b), as well, on the assumption that apparent in situ wh-phrases in English pair-list questions do move, albeit covertly (see also Nissenbaum 2000:197–201). The converse crucially does not hold: the putative explanation of (43b) based on lack of movement of the second wh-phrase could not possibly generalize to (44b).

Importantly, such interactions are only possible among multiple wh-phrases landing at the same clausal periphery. Violations of island constraints are not ameliorated if they target a CP periphery that is not itself targeted by a separate, island-respecting movement chain. Compare (44b) with (45):

(45) * Kakvo$_1$ kazva tozi služitel na [žurnalistite, kojto$_2$ [t_2 razsledvat] t_1], če komunistite sa
what tells this official to journalists who investigate that communists AUX
zabludili redaktorite im?
deceived editors their
Intended: 'What$_1$ does this official tell journalists who$_2$ [t_2 are investigating t_1] that the communists have deceived their editors?' [*Richards 1997:256*]

Richards proposes the following principle to capture these effects:

(46) PRINCIPLE OF MINIMAL COMPLIANCE: original version
For any dependency D that obeys constraint C, any elements that are relevant for determining whether D obeys C can be ignored for the rest of the derivation for purposes of determining whether any other dependency D' obeys C.
[*Richards 1998:601; see also Richards 2001:199*]

The same principle also explains why, in pair-list questions in Bulgarian, the two wh-phrases exhibit standard superiority effects, but in tuple-list questions involving more than two elements, there are no superiority effects among the $(n-1)$ non-highest wh-phrases (see Richards 2001:282).

I will adopt a slight variation on (46) which, as far as I can tell, performs equally well with respect to the Bulgarian data discussed here, but which generalizes more readily to the head-movement scenario that is our current focus:

(47) PRINCIPLE OF MINIMAL COMPLIANCE: revised version
Once a probe P has successfully targeted a goal G, any other goal G' that meets the same featural search criterion, and is dominated or c-commanded by G (= dominated by the mother of G), is accessible to subsequent probing by P irrespective of locality conditions.

7.5.2. The A-over-A-like condition meets the PMC

Let us now reconsider head movement in light of the Principle of Minimal Compliance (PMC). Section 7.3 ended with the observation that Bare Phrase Structure (BPS), combined with Iterative Downward Search (IDS)—specifically, (38a) and (38b)—appears to ban head movement altogether. No featural search criterion could possibly be satisfied by a head without also being satisfied by the maximal projection of that same head, and the latter will be encountered by the probe first. That is, if α^0 and αP are distinct, it is αP that must be targeted. This was referred to as the *A-over-A*-like locality condition on probe-goal relations.

If this is a locality condition par excellence, though, then we predict that it would be subject to the PMC (note that (47) is formulated over *locality conditions* in general). Consequently, it is only the first relation targeting α that should be subject to this A-over-A-like condition. After

the condition has been satisfied once, subsequent relations between the same probe and (some projection of) α are, by hypothesis, exempt from it. Therefore, we predict that it should be possible for a probe H^0 to target the head of αP to the exclusion of other material in αP, as long as this is not the first relation initiated by H^0 that targets (some projection of) α.

All of this is not enough to make head movement possible, however. Recall: **(i)** probes must search for their goals using a *featural* search criterion (40); **(ii)** the phrasal node ("XP") is, by hypothesis, featurally identical to the head ("X^0"); and **(iii)** the phrasal node is unambiguously closer to the probe than the head is, in terms of the explicit iterative downward search algorithm (IDS) given in (39). So if movement of heads to the exclusion of the rest of their phrases is indeed attested, the impetus for moving the head alone cannot come from the search criterion. It must have a different source.

There is a persistent intuition in the syntactic literature that there is something fundamentally superfluous about phrasal movement. After all, the featural properties of the moving constituent are determined by its head; and so, if movement is a response to the featural needs of some higher attractor, phrasal material outside of the attracted head is not implicated in the mechanism that drives movement in the first place.[25] In line with this intuition, I propose the following condition:

(48) Minimal Remerge: If X^0/X_{min} is movable, move only X^0/X_{min}.

Importantly, the antecedent of the conditional in (48) is often false. Without a previous relation in place between the probe and XP, a relation which would adhere to the A-over-A-like condition and satisfy the PMC (47), there would be no way for X^0/X_{min} to move on its own. The condition in (48) can only wield its influence when a previous relation of this sort is already in place.

In sections 7.5.3–7.5.4, I will discuss in detail how (48) interacts with the A-over-A-like locality condition and the PMC, both in local configurations (where the relevant probe is the immediate sister of XP) and in non-local ones (where the probe is more structurally distant). In the meantime, let us note the following: once a probe H has employed a featural search criterion f to target XP, subsequent relations involving f between H and X will necessarily target the head X^0 alone, because of (48). One consequence of this is that "Agr"-style projections—where a phrase is agreed with by a head and is subsequently attracted, by the same features, to the specifier of that head—are ruled out as a matter of principle. Put another way, a phrase XP that is agreed with by a head H^0 can only move to [Spec,HP] if the movement is triggered by a different feature (or more accurately, by a different featural search criterion) than the one involved in the agreement relation. As an example, consider subjects: if T has agreed with a DP, movement of that DP to [Spec,TP] cannot be driven by φ-features, since that would result by hypothesis in head movement, not phrasal movement; instead, it must involve, e.g., an [EPP]- or [D]-feature.[26] See Starke (2001), among others, for related ideas.

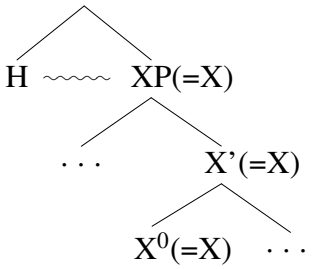
[25] Some examples are Chomsky (1995:262ff.) ("The operation Move [...] seeks to raise just F *[the formal feature being attracted; O.P.]*"), and Donati (2006:29–30) ("Merge just enough material for convergence").

[26] Preminger (2014:129ff.) argues that outside of quirky-subject languages, movement of subjects to [Spec,TP] does not involve its own search procedure at all; instead, movement applies automatically to that XP which was targeted for agreement in φ-features. The dissociation between φ-features and [EPP]- or [D]-features would still be necessary, however, in a quirky-subject language (e.g. Icelandic). It is probably not a coincidence that in the latter type of language, the dissociation between φ-features and the features that drive subject movement is morphologically conspicuous (see the discussion of Icelandic in section 6.2).

7.5.3. Locality, c-selection, and anti-locality

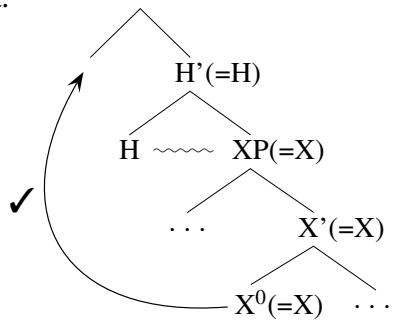
The relation that most often plays the role of satisfying the PMC and rendering X^0 movable, is c-selection. In configurations where a head H attracts a feature borne by its complement, the complement XP already stands in a c-selection relation with H (indicated in (49) with a wavy line):

(49)

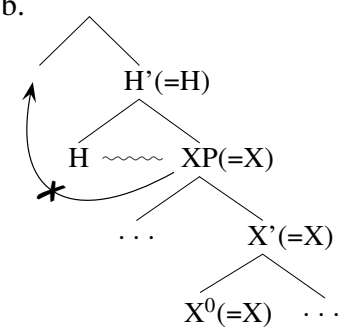


Being a relation between H and XP, c-selection satisfies the A-over-A-like locality condition. The PMC then dictates that subsequent relations between H and anything dominated or c-commanded by XP are licit. This means that X^0 is now moveable, and Minimal Remerge (48) can now wield its influence. And its influence will be to rule out phrasal movement of XP, and only permit head movement of X^0:

(50) a.

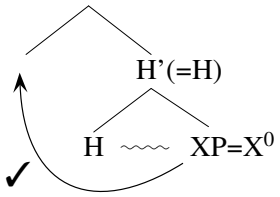


b.



The reader will notice that (50a–b) is an anti-locality effect. In particular, it recapitulates Abels' (2003) version of the constraint, but with one important difference. Abels' version bans movement of an element from [Compl,H] to [Spec,HP], full stop. The system just presented predicts there would be one specific instance in which such movement would be licit, namely, when—unlike in (50)—the constituent in [Compl,H] is non-branching:

(51)

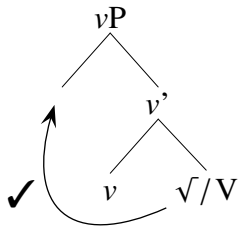


This is because, on the current view, the effect in (50) arises through the interplay of c-selection, the PMC, and Minimal Remerge. What Minimal Remerge (48) mandates is that the minimal movable projection of an X be the constituent that undergoes movement. In a scenario like (51), where the moving X only has one level of projection in the first place, the effects of Minimal

Remerge are vacuous. To put it another way, there is no penalty on moving the entire constituent in [Compl,H] in (51) because there is no smaller projection of X that could have moved.

The question, of course, is whether this deviation from Abels' version of anti-locality is in fact warranted. One consideration that bears on this question involves Matushansky's (2006) theory of head movement. On Matushansky's approach, head movement involves a non-branching constituent undergoing regular regular syntactic movement into a specifier position, followed by m-merger between this specifier and the adjacent head. In light of this, consider movement of an intransitive V^0 (or root) to v^0. In this case, the step prior to m-merger is movement of the lower head to [Spec,*v*P]:

(52)



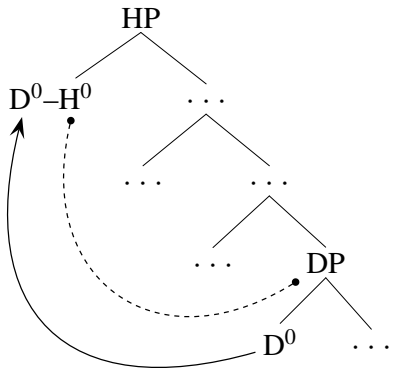
Because the complement of v^0 in (52) is non-branching, this movement would be in violation of Abels' version of the anti-locality constraint: it is movement of the entire [Compl,*v*] constituent into [Spec,*v*P].[27] Thus, if we want to maintain Matushansky's approach to head movement, we cannot maintain Abels' version of the constraint. Importantly, however, the version of anti-locality derived here is compatible with such movement (cf. (51), above).

7.5.4. The clitic doubling caveat: long head movement revisited

Recall that clitic doubling is, by hypothesis, long head movement of D (section 7.1). Given the system laid out in sections 7.5.1–7.5.2, *any* instance in which a probe H triggers movement of a head X must be prefigured by another syntactic relation between H and XP (the maximal projection of X). This is necessary in order to satisfy the PMC with respect to the relevant A-over-A-like locality condition (section 7.3). In cases of maximally local head movement, it was c-selection between H and XP that filled this role. But in cases of long head movement, H and XP do not stand in a sisterhood relation, making c-selection between the two impossible (cf. Chomsky 1994, 1995).

I propose that in cases of clitic doubling, the cliticization host H first enters into agreement with the full DP. It is this agreement relation that satisfies the PMC, enabling subsequent movement of the D head, on its own, to H:

[27] We may rightly ask whether there truly are instances of non-branching verb heads of this sort. Unaccusatives certainly wouldn't fit the bill, since they involve an argument base-generated in [Compl,V]. However, it has been shown that at least some unergatives are truly intransitive, i.e., lack even so much as an implicit object in [Compl,V] (Preminger 2012). Depending on the analysis of weather predicates, they may constitute another example of a complementless V. Finally, if we take seriously the theory of category-less roots undergoing categorization in syntax, roots of result-nominals would stand in the same configuration as (52) relative to their categorizing n^0, and these roots are uncontroversially argumentless.

(53) 

[H = the cliticization host (T / *v* / etc.)]

This explains why it does not matter, for the purposes of PCC distribution, whether a given instance of agreement morphology is an instance of agreement proper (i.e., feature valuation on a functional head), or the result of clitic doubling. That is because clitic doubling still depends on establishing a prior agreement relation of the former kind. Syntactic agreement is therefore implicated in both types of agreement morphology, and as already shown in section 4, it is also the key to understanding PCC effects—especially in light of their sensitivity to finer syntactic hierarchy.

An immediate question raised by this view of clitic doubling is its status with respect to the no-null-agreement generalization discussed in section 5, and repeated here:

(54) THE NO-NULL-AGREEMENT GENERALIZATION *[=(18)]*
There is no such thing as morpho-phonologically undetectable φ-feature agreement.

In many (perhaps, most) cases of clitic doubling, there is no overt morpho-phonological expression of a prior agreement relation. In (55), for example, there is no overt exponence of an agreement relation between the cliticization host (v^0) and the object. (The verb displays agreement with the subject, but there is no agreement with the object independent of the clitic.) And this is paradigm-wide, i.e., it is not a matter of the particular φ-features of *profesor* (“professor”).

(55) Le vi al profesor ayer (Leísta Spanish)
CL I.saw A-the professor yesterday
‘I saw the professor yesterday.’ [*Bleam 1999:45*]

Instances of clitic doubling do exist where, alongside the clitic itself, one finds overt agreement with the doubled argument—for example, clitic doubling of subjects in certain Northern Italian dialects (Poletto 2000). While this may provide circumstantial support for the idea that clitic doubling is prefigured by syntactic agreement as in (53), it does not change the facts of (55) and many cases like it.

There is no way around the fact that if, as argued above, agreement with the object is implicated in cases like (55), then it is an agreement relation that stands in violation of the no-null-agreement generalization. And this, in turn, means that this generalization cannot be a steadfast, combinatorial principle of grammar. But we cannot abandon (54), either; recall that it was a necessary component of any adequate account of the distribution of PCC effects, given the evidence that the PCC was a fundamentally syntactic effect (sections 4–5).

Are we at an impasse, then? In the next section, I suggest that the answer is no—and that this apparent tension can be resolved by viewing (54) not as a grammatical principle unto itself, but as the outcome of a particular kind of acquisition strategy affecting the placement of unvalued φ-features on functional heads.

8. The nature of the no-null-agreement generalization: a conservative acquisition strategy for unvalued φ-features

In earlier sections, we saw evidence in favor of the no-null-agreement generalization, which states that there is generally no such thing as agreement that is morpho-phonologically null across the entire paradigm. Without this generalization, there would be no way to account for the distribution of the PCC, given that the latter is a syntactic effect par excellence (section 4), and yet it comes and goes with the presence of overt agreement morphology (section 5). In section 7, we saw an argument that clitic doubling involves a prior agreement relation between the cliticization host and the full DP of which the clitic is a subpart. Crucially, this agreement relation often goes unexponed, in apparent violation of the aforementioned generalization.

The solution I put forth is to view the aforementioned generalization not as a steadfast principle of grammar, but as the result of a particular kind of acquisition strategy. Before spelling out the strategy in detail, it is worth pointing out that *were* this generalization a principle of grammar per se, it would raise the same kind of modularity issue discussed in relation to the PCC in section 5. Consider: if agreement is a syntactic operation, then it occupies a part of the grammar where reference to the morpho-phonological content of terminals is impossible. The problem would be even more severe, in fact, because the principle would have to be trans-derivational: it is not the morpho-phonological content of a particular terminal in a particular derivation that is at issue, but rather the fact that *some cells in a paradigm* must be overt.

Instead, what I propose is that the no-null-agreement generalization, and its exceptions, arise because of how language acquisition proceeds. Specifically, the learner begins with the assumption that there are no unvalued φ-features on any functional heads (this includes T^0 and v^0). Recall that this does not pose any case-related problems—not even for languages with rich and easily evident case morphology—given the evidence that has accumulated in recent years against structural case being assigned by agreement (Preminger 2014, *a.o.*; see also section 1). There is then a specific and, crucially, limited set of triggers that would cause the learner to revise this hypothesis, and posit unvalued φ-features on a given functional head:

(56) TRIGGERS FOR LEARNER TO POSIT UNVALUED φ-FEATURES ON A HEAD H^0

a. overt morpho-phonological covariance in φ-features between the form of H^0 and a DP

b. long-distance head movement (of a D head) to H^0

Crucially, the list in (56) is anything but open-ended. It absolutely cannot include, for example, the existence of a binding or fake-indexical relation involving H^0. If such phenomena were also triggers for positing unvalued φ-features on H^0, a proper account for the distribution of PCC effects would be rendered impossible. Recall that the PCC arises wherever there is overt agreement or clitic doubling (section 5); the presence of binding and/or fake-indexicals in a given construction does not suffice to give rise to PCC effects. The conclusion, already argued elsewhere on independent grounds, is that phenomena of the latter sort do not involve syntactic agreement in φ-features.

This proposal provides us with a "roadmap" for how a language with PCC effects is acquired. (Or, to be more precise, how a construction with a particular inventory of functional projections that ends up generating PCC effects is acquired.) The learner starts with the assumption that there are no unvalued φ-features on the relevant functional projection—say, v^0. Very quickly, however, she will be driven to revise these assumptions, either because v^0 shows morpho-phonologically overt covariance in φ-features with the direct object (as is the case in Basque, for example; Arregi & Nevins 2008, 2012, Preminger 2009), or because there is a D associated with the direct object that cliticizes to (i.e. undergoes long head movement to) v^0, as in Spanish. In the latter case, the learner can deduce with certainty that there must be a prior agreement relation between v^0 and the the direct object, for the reasons discussed in section 7.5.

Importantly, misidentifying clitic doubling as "pure" agreement, or vice versa, will be fairly innocuous at this stage, since in either case the learner will end up positing unvalued φ-features on the relevant functional head. This is a desirable property: while agreement and clitic doubling are clearly different phenomena, the kind of data that distinguish the two are fairly subtle (see section 6.1 and references therein). It is not unreasonable to assume that, in the course of language acquisition, one may initially be identified as the other; and, in fact, this maybe the etiology of one type of language change, wherein pronominal clitics are reanalyzed as markers of agreement (i.e., valuation of formal features on a probe; see, for example, Gelderen 2011).

Either way, once the learner has posited unvalued features on v^0, the PCC then arises as a direct consequence of agreement and intervention, as discussed in section 4.

On the other hand, the learner acquiring a language *without* PCC effects (i.e., any language lacking agreement morphology cross-referencing internal arguments) will never be driven to posit unvalued φ-features on v^0, meaning their grammar will lack agreement with internal arguments altogether. As discussed in section 5, this is the desired outcome for the grammar of such a language.

9. Conclusion

This paper began by surveying some of the evidence that the Person Case Constraint (PCC) is sensitive to the kind of fine-grained hierarchical distinctions that characterize syntax proper. This means that in any system where there is a meaningful modular distinction between morphology and syntax, the PCC is part of the latter module. I then surveyed, in broad strokes, what a syntactic account of the PCC that is capable of deriving this sensitivity would look like (building on Anagnostopoulou 2003, 2005, Béjar & Rezac 2003, *a.o.*), based on mechanisms of syntactic agreement and intervention.

Next, I turned to the fairly well known fact that the PCC seems coupled to the existence of overt agreement morphology with the arguments involved: as this morphology comes and goes, so does the PCC effect itself. This was shown to hold even intra-linguistically, as demonstrated by the distinction between finite and non-finite environments in Basque. I then juxtaposed this with the earlier results concerning the fundamentally syntactic nature of PCC effects, and its account in terms of agreement and intervention. Assuming that syntax does not make direct reference to the morpho-phonological content of terminals, this led to the conclusion that contexts that do not exhibit the PCC simply lack agreement with internal arguments altogether. I labeled this the no-null-agreement generalization.

One important caveat to these generalizations concerned clitic doubling: for the purposes of the PCC, clitic doubling behaved as though it involved agreement (even though it is by now known

that clitic doubling is not equivalent to agreement). I then showed that, even if clitic doubling is movement, the idea that *all* movement is prefigured by an agreement relation is untenable, and therefore cannot be what underpins this caveat.

I argued that a more promising alternative can be found by investigating the interplay of Bare Phrase Structure (Chomsky 1994), Iterative Downward Search (Kitahara 1994, 1997, Koizumi 1995, Müller 1996, 1998, Takano 1994), and the Principle of Minimal Compliance (Richards 1998, 2001). In particular, the idea is that movement always "strives" to move only the head, but this is seldom possible because Bare Phrase Structure and Iterative Downward Search together yield an A-over-A-like locality condition that demands that the entire phrase be the target of the syntactic relation. Crucially, however, when the attractor already stands in some prior relation with this phrasal node (e.g. c-selection, agreement), this satisfies the relevant locality condition. It follows, given the Principle of Minimal Compliance, that subsequent syntactic operations involving the same relata need not adhere to the same locality condition again, which is what enables movement of the head alone.

Clitic doubling, *qua* long head movement, cannot be prefigured by c-selection because it does not involve a sufficiently local configuration (namely, sisterhood). Some other syntactic relation must therefore be what satisfies the A-over-A-like locality condition in this case. I proposed that syntactic agreement is the relation that plays this role. Clitic doubling thus "counts" as syntactic agreement for the purposes of the distribution of PCC effects because it invariably involves an initial agreement step. The agreement involved in clitic doubling, however, seemed to pose a challenge for the no-null-agreement generalization, since in many clitic-doubling languages, there is no morphology indicating valuation appearing alongside the clitic itself.

I then showed how this picture could arise as the result of a conservative acquisition strategy, where the learner does not posit unvalued φ-features on functional heads unless and until faced with a particular kind of positive evidence. This type of strategy could give rise to the no-null-agreement generalization, as well as its clitic-doubling caveat.

References

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

Abels, Klaus. 2012. *Phases: an essay on cyclicity in syntax*. Linguistische Arbeiten 543, Berlin: de Gruyter.

Albizu, Pablo. 1997. Generalized Person-Case Constraint: a case for a syntax-driven inflectional morphology. *Anuario del Seminario de Filología Vasca Julio de Urquijo (ASJU, International Journal of Basque Linguistics and Philology)* XL:1–33.

Anagnostopoulou, Elena. 2003. *The syntax of ditransitives: evidence from clitics*. Berlin: Mouton de Gruyter.

Anagnostopoulou, Elena. 2005. Strong and weak person restrictions: a feature checking analysis. In *Clitic and affix combinations: theoretical perspectives*, eds. Lorie Heggie & Francisco Ordonez, Linguistics Today 74, 199–235. Amsterdam: John Benjamins.

Anagnostopoulou, Elena. 2006. Clitic doubling. In *The Blackwell companion to syntax*, eds. Martin Everaert & Henk van Riemsdijk, vol. 1, 519–581. Oxford: Blackwell Publishers.

Anagnostopoulou, Elena. 2008. Notes on the Person Case Constraint in Germanic (with special reference to German). In *Agreement restrictions*, eds. Roberta D'Alessandro, Susann Fischer & Gunnar Hrafn Hrafnbjargarson, 15–47. Berlin: Mouton de Gruyter.

Anagnostopoulou, Elena. to appear. Clitic doubling. In *The Blackwell companion to syntax*, eds. Martin Everaert & Henk van Riemsdijk, 2nd Edition, Oxford: Blackwell Publishers.

Andrews, Avery D. 1976. The VP complement analysis in Modern Icelandic. In *Proceedings of the 6th meeting of the North East Linguistic Society (NELS 6)*, eds. Alan Ford, John Reighard & Rajendra Singh, Montreal, QC: McGill University and University of Montreal, 1–21.

Arregi, Karlos & Andrew Ira Nevins. 2008. Agreement and clitic restrictions in Basque. In *Agreement restrictions*, eds. Roberta D'Alessandro, Susann Fischer & Gunnar Hrafn Hrafnbjargarson, 49–86. Berlin: Mouton de Gruyter.

Arregi, Karlos & Andrew Ira Nevins. 2012. *Morphotactics: Basque auxiliaries and the structure of spellout*. Studies in Natural Language and Linguistic Theory 86, Dordrecht: Springer.

Baker, Mark C. 1988. *Incorporation: a theory of grammatical function changing*. Chicago, IL: University of Chicago Press.

Baker, Mark C. 1996. *The polysynthesis parameter*. Oxford: Oxford University Press.

Baker, Mark C. 2008. *The syntax of agreement and concord*. Cambridge Studies in Linguistics 115, Cambridge: Cambridge University Press.

Béjar, Susana & Milan Rezac. 2003. Person licensing and the derivation of PCC effects. In *Romance linguistics: theory and acquisition*, eds. Ana Teresa Perez-Leroux & Yves Roberge, 49–62. Amsterdam: John Benjamins.

Belletti, Adriana. 2005. Extended doubling and the VP periphery. *Probus* 17:1–35.

Bittner, Maria & Ken Hale. 1996. The structural determination of Case and Agreement. *Linguistic Inquiry* 27:1–68.

Bleam, Tonia. 1999. *Leísta Spanish and the syntax of clitic doubling*. Doctoral dissertation, Newark, DE: University of Delaware.

Bobaljik, Jonathan David. 2008. Where's phi? Agreement as a post-syntactic operation. In *Phi Theory: phi-features across interfaces and modules*, eds. Daniel Harbour, David Adger & Susana Béjar, 295–328. Oxford: Oxford University Press.

Bobaljik, Jonathan David & Susi Wurmbrand. 2014. *Questions with declarative syntax tell us what about selection?*, Ms., Storrs, CT: University of Connecticut.

Bonet, Eulalia. 1991. *Morphology after syntax: pronominal clitics in Romance*. Doctoral dissertation, Cambridge, MA: MIT.

Bonet, Eulalia. 1994. The Person-Case Constraint: a morphological approach. In *The morphology-syntax connection*, eds. Heidi Harley & Colin Phillips, MIT Working Papers in Linguistics 22, 33–52. Cambridge, MA: MITWPL.

Borsley, Robert D. & Andreas Kathol. 2000. Breton as a V2 language. *Linguistics* 38:665–710.

Borsley, Robert D., Maria Luisa Rivero & Janig Stephens. 1996. Long head movement in Breton. In *The syntax of the Celtic languages*, eds. Ian Roberts & Robert D. Borsley, 53–74. Cambridge: Cambridge University Press.

Bošković, Željko. 1994. D-structure, θ-theory, and movement into θ-positions. *Linguistic Analysis* 24:247–286.

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

Bošković, Željko & Jairo Nunes. 2007. The Copy Theory of Movement: a view from PF. In *The Copy Theory of Movement*, eds. Jairo Nunes & Norbert Corver, Linguistics Today 107, 13–74. Oxford: Blackwell Publishers.

Brody, Michael. 2000. Mirror Theory: syntactic representation in perfect syntax. *Linguistic Inquiry* 31:29–56, DOI: 10.1162/002438900554280.

Bruening, Benjamin. 2005. *The Algonquian Inverse is syntactic: binding in Passamaquoddy*. Ms.

Carstens, Vicki. 2000. Concord in minimalist theory. *Linguistic Inquiry* 31:319–355, DOI: 10.1162/002438900554370.
Cecchetto, Carlo. 2005. Doubling structures and reconstruction. *Probus* 12:93–126.
Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
Chomsky, Noam. 1994. *Bare phrase structure*. MIT Occasional Papers in Linguistics 5, Cambridge, MA: MITWPL.
Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step by step: essays on minimalist syntax in honor of Howard Lasnik*, eds. Roger Martin, David Michaels & Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: a life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.
Craenenbroeck, Jeroen van & Marjo van Koppen. 2008. Pronominal doubling in Dutch dialects: Big DPs and coordinations. In *Microvariation in syntactic doubling*, eds. Sjef Barbiers, Olaf Koeneman, Marika Lekakou & Margreet van der Ham, 207–249. Bingley: Emerald.
Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.
Diesing, Molly. 1997. Yiddish VP order and typology of object movement in Germanic. *Natural Language & Linguistic Theory* 15:369–427, DOI: 10.1023/A:1005778326537.
Diesing, Molly & Eloise Jelinek. 1993. The syntax and semantics of object shift. *Working Papers in Scandinavian Syntax* 51:1–54.
Donati, Caterina. 2006. On wh-head movement. In *wh-movement : moving on*, eds. Lisa L.-S. Cheng & Norbert Corver, 21–46. Cambridge, MA: MIT Press.
Elordieta, Arantzazu. 2001. *Verb movement and constituent permutation in Basque*. Doctoral dissertation, Leiden: Leiden University. LOT dissertation series.
Emonds, Joseph E. 1970. *Root and structure-preserving transformations*. Doctoral dissertation, Cambridge, MA: MIT.
Emonds, Joseph E. 1976. *A transformational approach to English syntax: root, structure-preserving, and local transformations*. New York, NY: Academic Press.
Etxepare, Ricardo. 2006. Number long distance agreement in (substandard) Basque. In *Studies in Basque and historical linguistics in memory of Robert L. Trask*, eds. Joseba A. Lakarra & Jose Ignacio Hualde, Supplements of the Anuario del Seminario de Filología Vasca Julio de Urquijo (ASJU, International Journal of Basque Linguistics and Philology) XL:1–2, 303–350.
Gelderen, Elly van. 2011. *The linguistic cycle: language change and the language faculty*. Oxford: Oxford University Press.
Gribanova, Vera & Boris Harizanov. 2016. *Whither head movement?*, Paper presented at the Stanford Workshop on Head Movement.
Grimshaw, Jane. 2000. Locality and extended projections. In *Lexical specification and insertion*, eds. Peter Coopmans, Martin Everaert & Jane Grimshaw, 115–134. Amsterdam: John Benjamins.
Grohmann, Kleanthes K. 2003. *Prolific domains: on the anti-locality of movement dependencies*. Amsterdam: John Benjamins.
Harizanov, Boris. 2014. Clitic doubling at the syntax-morphophonology interface: A-movement and morphological merger in Bulgarian. *Natural Language & Linguistic Theory* 32:1033–1088, DOI: 10.1007/s11049-014-9249-5.
Harizanov, Boris. 2016. *Head movement to specifier positions in Bulgarian participle fronting*. Paper presented at the *90th Winter Meeting of the Linguistic Society of America (LSA 90)*. URL: <http://stanford.edu/~bharizan/pdfs/Harizanov_2016_LSA_handout.pdf>.

Harley, Heidi. 1995. Abstracting away from abstract case. In *Proceedings of the 25th meeting of the North East Linguistic Society (NELS 25)*, ed. Jill Beckman, Amherst, MA: GLSA, 207–221.
Harley, Heidi & Elizabeth Ritter. 2002. Person and number in pronouns: a feature-geometric analysis. *Language* 78:482–526, DOI: 10.1353/lan.2002.0158.
Hartman, Jeremy. 2011. The semantic uniformity of traces: evidence from ellipsis parallelism. *Linguistic Inquiry* 42:367–388.
Haspelmath, Martin. 2004. Explaining the ditransitive person-role constraint. *Constructions*.
Holmberg, Anders & Þorbjörg Hróarsdóttir. 2003. Agreement and movement in Icelandic raising constructions. *Lingua* 113:997–1019, DOI: 10.1016/S0024-3841(02)00162-6.
Hornstein, Norbert. 2009. *A theory of syntax: minimal operations and universal grammar*. Cambridge: Cambridge University Press.
Ishii, Toru. 1997. *An asymmetry in the composition of phrase structure and itsconsequences*. Doctoral dissertation, Irvine, CA: University of California.
Ishii, Toru. 1999. Cyclic spell out and that-*t* effects. In *Proceedings of the 18th West Coast Conference on Formal Linguistics (WCCFL 18)*, eds. Sonya Bird, Andrew Carnie, Jason D. Haugen & Peter Norquest, Somerville, MA: Cascadilla Press, 220–231.
Jónsson, Jóhannes Gísli. 1996. *Clausal architecture and Case in Icelandic*. Doctoral dissertation, Amherst, MA: University of Massachusetts.
Jouitteau, Mélanie. 2005. *La syntaxe comparée du Breton*. Doctoral dissertation, Nantes: Université de Nantes.
Kayne, Richard S. 2005. Some notes on comparative syntax with special reference to English and French. In *The Oxford handbook of comparative syntax*, eds. Guglielmo Cinque & Richard S. Kayne, 3–69. New York, NY: Oxford University Press.
Kitahara, Hisatsugu. 1994. Restricting ambiguous rule-application: a unified analysis of movement. In *Formal approaches to Japanese linguistics I*, eds. Masatoshi Koizumi & Hiroyuki Ura, MIT Working Papers in Linguistics 6 24, 179–209. Cambridge, MA: MITWPL.
Kitahara, Hisatsugu. 1997. *Elementary operations and optimal derivations*. Cambridge, MA: MIT Press.
Koizumi, Masatoshi. 1995. *Phrase structure in minimalist syntax*. Doctoral dissertation, Cambridge, MA: MIT.
Kornfilt, Jaklin & Omer Preminger. 2015. Nominative as *no case at all*: an argument from raising-to-ACC in Sakha. In *Proceedings of the 9th Workshop on Altaic Formal Linguistics (WAFL 9)*, eds. Andrew Joseph & Esra Predolac, MIT Working Papers in Linguistics 76, Cambridge, MA: MITWPL, 109–120.
Kratzer, Angelika. 2009. Making a pronoun: fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40:187–237, DOI: 10.1162/ling.2009.40.2.187.
Laka, Itziar. 1996. *A brief grammar of Euskara, the Basque language*. Open-access grammar, ISBN: 84-8373-850-3, Vitoria-Gasteiz: Euskal Herriko Unibertsitatea (University of the Basque Country). URL: <http://www.ehu.eus/en/web/eins/basque-grammar>.
Lambova, Mariana. 2004. On triggers of movement and effects at the interfaces. In *Triggers*, eds. Anne Breitbarth & Henk van Riemsdijk, 231–258. Berlin: Mouton de Gruyter.
Landau, Idan. 1994. *Dative shift and extended VP-shells*. Master's thesis, Tel-Aviv: Tel-Aviv University.
Landau, Idan. 2006. Chain resolution in Hebrew V(P)-fronting. *Syntax* 9:32–66, DOI: 10.1111/j.1467-9612.2006.00084.x.
Lechner, Winfried. 2006. An interpretive effect of head movement. In *Phases of interpretation*, ed. Mara Frascarelli, 45–70. Berlin: Mouton de Gruyter, DOI: 10.1515/9783110197723.2.45, URL: <http://ling.auf.net/lingBuzz/000178>.

Lema, José & Maria Luisa Rivero. 1990. Long head movement: ECP vs. HMC. In *Proceedings of the 20th meeting of the North East Linguistic Society (NELS 20)*, eds. Juli Carter, Rose-Marie Déchaine, Bill Philip & Tim Sherer, vol. 2, Amherst, MA: GLSA, 333–347.

Li, Yafei. 1990. X^0-binding and verb incorporation. *Linguistic Inquiry* 21:399–426.

Mallen, Enrique. 1997. A minimalist approach to concord in noun phrases. *Theoretical Linguistics* 23:49–78, DOI: 10.1515/thli.1997.23.1-2.49.

Marantz, Alec. 1991. Case and licensing. In *Proceedings of the 8th Eastern States Conference on Linguistics (ESCOL 8)*, eds. German Westphal, Benjamin Ao & Hee-Rahk Chae, Ithaca, NY: CLC Publications, 234–253.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37:69–109, DOI: 10.1162/002438906775321184.

McFadden, Thomas. 2004. *The position of morphological case in the derivation: a study on the syntax-morphology interface*. Doctoral dissertation, Philadelphia, PA: University of Pennsylvania.

Merchant, Jason. 2006. Polyvalent case, geometric hierarchies, and split ergativity. In *Proceedings of the 42nd annual meeting of the Chicago Linguistic Society (CLS 42)*, eds. Jackie Bunting, Sapna Desai, Robert Peachey, Chris Straughn & Zuzana Tomkova, vol. 2, Chicago, IL: Chicago Linguistic Society, 57–76.

Müller, Gereon. 1996. A constraint on remnant movement. *Natural Language & Linguistic Theory* 14:355–407, DOI: 10.1007/BF00133687.

Müller, Gereon. 1998. *Incomplete category fronting*. Dordrecht: Kluwer Academic Publishers.

Murasugi, Keiko S. & Mamoru Saito. 1995. Adjunction and cyclicity. In *Proceedings of the 13th West Coast Conference on Formal Linguistics (WCCFL 13)*, eds. Raul Aranovich, William Byrne, Susanne Preuss & Martha Senturia, Stanford, CA: CSLI Publications, 302–317.

Nevins, Andrew Ira. 2007. The representation of third person and its consequences for Person-Case effects. *Natural Language & Linguistic Theory* 25:273–313, DOI: 10.1007/s11049-006-9017-2.

Nevins, Andrew Ira. 2011. Multiple agree with clitics: person complementarity vs. omnivorous number. *Natural Language & Linguistic Theory* 29:939–971, DOI: 10.1007/s11049-011-9150-4.

Nichols, Lynn. 2001. The syntactic basis of referential hierarchy phenomena: clues from languages with and without morphological case. *Lingua* 111:515–537, DOI: 10.1016/S0024-3841(00)00038-3.

Nissenbaum, Jon. 2000. *Investigations of covert phrase movement*. Doctoral dissertation, Cambridge, MA: MIT.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Cambridge, MA: MIT Press.

Pesetsky, David & Esther Torrego. 2001. T-to-C movement: causes and consequences. In *Ken Hale: a life in language*, ed. Michael Kenstowicz, 355–426. Cambridge, MA: MIT Press.

Poletto, Cecilia. 2000. *The higher functional field*. Oxford: Oxford University Press.

Preminger, Omer. 2006. *Argument-mapping and extraction*. Master's thesis, Tel-Aviv: Tel-Aviv University.

Preminger, Omer. 2009. Breaking agreements: distinguishing agreement and clitic doubling by their failures. *Linguistic Inquiry* 40:619–666, DOI: 10.1162/ling.2009.40.4.619.

Preminger, Omer. 2011a. *Agreement as a fallible operation*. Doctoral dissertation, Cambridge, MA: MIT.

Preminger, Omer. 2011b. Asymmetries between person and number in syntax: a commentary on Baker's SCOPA. *Natural Language & Linguistic Theory* 29:917–937, DOI: 10.1007/s11049-011-9155-z.

Preminger, Omer. 2012. The absence of an implicit object in unergatives: new and old evidence from Basque. *Lingua* 122:278–288, DOI: 10.1016/j.lingua.2011.04.007.

Preminger, Omer. 2014. *Agreement and its failures*. Linguistic Inquiry Monographs 68, Cambridge, MA: MIT Press.

Reuland, Eric. 2011. *Anaphora and language design*. Cambridge, MA: MIT Press.

Rezac, Milan. 2008a. Phi-agree and theta-related case. In *Phi Theory: phi-features across interfaces and modules*, eds. Daniel Harbour, David Adger & Susana Béjar, 83–129. Oxford: Oxford University Press.
Rezac, Milan. 2008b. The syntax of eccentric agreement: the Person Case Constraint and absolutive displacement in Basque. *Natural Language & Linguistic Theory* 26:61–106, DOI: 10.1007/s11049-008-9032-6.
Richards, Norvin. 1997. *What moves where when in which language?*, Doctoral dissertation, MIT.
Richards, Norvin. 1998. The Principle of Minimal Compliance. *Linguistic Inquiry* 29:599–629, DOI: 10.1162/002438998553897.
Richards, Norvin. 2001. *Movement in language: interactions and architecture*. Oxford: Oxford University Press.
Rivero, Maria Luisa. 1991. Long head movement and negation: Serbo-Croatian vs. Slovak and Czech. *The Linguistic Review* 8:319–351.
Roberts, Ian. 2004. The C-system in Brythonic Celtic. In *The structure of IP and CP*, ed. Luigi Rizzi, The Cartography of Syntactic Structures 2, 297–328. New York, NY: Oxford University Press.
Roberts, Ian. 2010. *Agreement and head movement: clitics, incorporation, and defective goals*. Linguistic Inquiry Monographs 59, Cambridge, MA: MIT Press.
Rooryck, Johan & Guido Vanden Wyngaerd. 2011. *Dissolving binding theory*. Oxford: Oxford University Press.
Saito, Mamoru & Keiko S. Murasugi. 1999. Subject predication within IP and DP. In *Beyond Principles and Parameters: essays in memory of Osvaldo Jaeggli*, eds. Kyle Johnson & Ian Roberts, 159–182. Dordrecht: Kluwer Academic Publishers.
Schafer, Robin. 1994. *Nonfinite predicate initial constructions in Modern Breton*. Doctoral dissertation, Santa Cruz, CA: University of California.
Sigurðsson, Halldór Ármann. 1989. *Verbal syntax and case in Icelandic in a comparative GB approach*. Doctoral dissertation, Lund: University of Lund.
Sigurðsson, Halldór Ármann & Anders Holmberg. 2008. Icelandic dative intervention: PERSON and NUMBER are separate probes. In *Agreement restrictions*, eds. Roberta D'Alessandro, Susann Fischer & Gunnar Hrafn Hrafnbjargarson, 251–280. Berlin: Mouton de Gruyter.
Sportiche, Dominique. 1998. *Partitions and atoms of clause structure: subjects, agreement, case and clitics*. London: Routledge.
Starke, Michal. 2001. *Move dissolves into Merge: a theory of locality*. Doctoral dissertation, Geneva: University of Geneva. URL: <http://ling.auf.net/lingbuzz/000002>.
Stephens, Janig. 1982. *Word order in Breton*. Doctoral dissertation, London: University College London.
Suñer, Margarita. 1988. The role of agreement in clitic-doubled constructions. *Natural Language & Linguistic Theory* 6:391–434, DOI: 10.1007/BF00133904.
Takano, Yuji. 1994. Unbound traces and indeterminacy of derivation. In *Current topics in English and Japanese*, ed. Masaru Nakamura, 229–253. Tokyo: Hituzi Syobo.
Thráinsson, Höskuldur. 1979. *On complementation in Icelandic*. New York, NY: Garland Publishing, Doctoral dissertation, Harvard University (Cambridge, MA).
Torrego, Esther. 1992. *Case and agreement structure*. Ms., Boston, MA: University of Massachusetts.
Travis, Lisa. 1984. *Parameters and effects of word order variation*. Doctoral dissertation, Cambridge, MA: MIT.
Uriagereka, Juan. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26:79–124.
Vergnaud, Jean-Roger. 1977. *Open letter to Chomsky and Lasnik*. April 1977, available at <http://ling.auf.net/lingBuzz/000461>.

Vicente, Luis. 2007. *The syntax of heads and phrases: a study of verb (phrase) fronting*. Doctoral dissertation, Leiden: University of Leiden.

Vicente, Luis. 2009. An alternative to remnant movement for partial predicate fronting. *Syntax* 12:180–213, DOI: 10.1111/j.1467-9612.2009.00125.x.

Wilder, Chris & Damir Ćavar. 1994. Long head movement? verb movement and cliticization in Croatian. *Lingua* 93:1–58.

Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical functions: the Icelandic passive. *Natural Language & Linguistic Theory* 3:441–483.

Zeijlstra, Hedde. 2004. *Sentential negation and negative concord*. Doctoral dissertation, Amsterdam: University of Amsterdam.

Zeijlstra, Hedde. 2008a. Modal concord. In *Proceedings of the 17th Semantics And Linguistics Theory conference (SALT 17)*, eds. Masayuki Gibson & Tova Friedman, Ithaca, NY: CLC Publications, 317–332.

Zeijlstra, Hedde. 2008b. *Negative concord is syntactic agreement*. Ms., Amsterdam: University of Amsterdam. URL: <http://ling.auf.net/lingBuzz/000645>.

SVN revision code: 8213