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

Omer Preminger∗
University of Maryland

1. What this talk is about

• When you talk to syntacticians, you’ll often hear things like:
  “Sure, language \(L\) doesn’t have object agreement on the surface; but that’s just a fact about morpho-phonology.”

⇒ My first goal today is to show you that this kind of thinking, when it comes to agreement, is almost always wrong:
  Generally speaking, there is no agreement where you can’t see agreement.

• In the course of this investigation, we’ll run into an interesting wrinkle:
  ◦ as far as the relevant diagnostics are concerned, clitic doubling behaves as though it were agreement
  ◦ This is surprising, because clitic doubling is an instance of movement —
    ◦ and because, contra Chomsky (2000, et seq.), movement does not generally involve an initial agreement step

⇒ In figuring out why it is that clitic doubling does have to be prefigured by agreement, we will uncover some interesting things about:
  ◦ locality
  ◦ head movement
  ◦ the locality of head movement
  ◦ the anti-locality of phrasal movement

2. What this talk is not about

• Recent years have seen a cottage industry of reductions-to-agreement —
  ◦ i.e., attempts to reduce various other linguistic phenomena to the same formal operation hypothesized to underpin agreement
    ◦ e.g. Chomsky’s (2000, 2001) Agree
  Incl.: Binding Theory (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 even the formation of in-situ questions (Bobaljik & Wurmbrand 2014)

• This talk is not about these (attempted) reductions.
• When I say ‘agreement’ I mean:
  \(\varphi\)-features covarying between a noun phrase and some verbal head
  ◦ where ‘\(\varphi\)-features’ is some non-empty subset of \{person, number, gender/noun-class\}

3. Outline

§4: A quick refresher on, or very brief introduction to, the Person Case Constraint (PCC)

§5: Why the PCC is fundamentally a syntactic phenomenon (Albizu 1997, Rezac 2008b)

§6: The sensitivity of the PCC to overtness, and the consequences of that sensitivity for linguistic theory (spoiler: no null agreement!)

§7: The clitic-doubling caveat

§8: A non-explanation for the caveat: the fallacy of agreement as a precursor to all DP movement

§9: A more promising alternative:
  ◦ an A-over-A explanation of why head movement tends to be maximally local (Hornstein 2009, Roberts 2010)
  ◦ c-selection, agreement, and the Principle of Minimal Compliance (Richards 1998, 2001)

§10: Learning the distribution of unvalued \(\varphi\)-features: no null agreement as an acquisition strategy

§11: Conclusion

∗For comments and helpful discussion, I thank Ümit Atlamaz, Jessica Coon, Vera Gribanova, Boris Harizanov, Norbert Hornstein, Ruth Kramer, Howard Lasnik, Theodore Levin, Jeff Lidz, David Pesetsky, Carolina Petersen, and Maria Polinsky; audiences at UMD’s S-Lab, NELS 46 (Concordia), ZAS (Berlin), and the Stanford Workshop on Head Movement (September 2016); and the participants in the fall 2016 seminar on Head Movement, Noun Incorporation, and Clitic Doubling at the University of Maryland. All errors are my own.
4. A quick refresher on, or very brief introduction to, the Person Case Constraint

- The Person Case Constraint (PCC; a.k.a., the “*me-lui Constraint”):
  - a family of restrictions limiting the person features of different arguments in relation to one another
    - usually affecting combinations of multiple internal arguments of a single predicate
      ⇒ most commonly illustrated using the direct and indirect objects of a ditransitive verb
  - Cross-linguistically, the constraint comes in (at least) four varieties:
    - Strong, Weak, Me-First, and UltraStrong
  - see Anagnostopoulou (2005) and Nevins (2007), and references therein

Example:

(1) **Strong PCC in Basque ditransitives:**
    In finite clauses, the direct object of a ditransitive verb must be 3rd person.

(2) a. Zuk niri liburu-a saldu d-i-f-da-zu
    ‘You have sold the book to me.’

b. *Zuk harakin-ari ni saldu n-(a)i-f-o-zu
    ‘You have sold me to the butcher.’ [Laka 1996]

- As you can see from (2a–b), the PCC is asymmetric —
  - (1) is restriction on the features of the DO in the presence of an IO;
  - but there is no corresponding restriction on the features of the IO in the presence of a DO

- The PCC is also person-specific —
  - it is a restriction on the person features of the DO in the presence of an IO;
  - but there is no corresponding restriction on the number features of the DO in the presence of an IO (see Nevins 2011, a.o.)

5. The PCC is syntactic (Albizu 1997, Rezac 2008b)

- Albizu (1997) and Rezac (2008b) show that the PCC is a fundamentally syntactic effect
  - and not, say, a morphological filter (cf. Bonet 1991, 1994, a.o.)

- They focus on two-place unaccusatives —
  - (verbs that take an ABS DP and a DAT DP, but no Erg DP)
    - which Rezac calls “applicative unaccusatives.”

- It turns out that there are two classes of applicative unaccusatives in Basque:
  - DAT ≫ ABS verbs
  - ABS ≫ DAT verbs

(for reasons that I won’t get into here, all true ditransitives, i.e., triadic verbs, are DAT ≫ ABS in Basque; see Elordieta 2001, Rezac 2008b)

(3) **DAT ≫ ABS:**

a. Kepa-ri bere buru-a gusta-tzen zako
   Kepa-DAT his head-ARTsg(ABS) like-HAB AUX
   ‘Kepa likes himself.’

b. * Kepa bere buru-a-ri gusta-tzen zako
   Kepa(ABS) his head-ARTsg-DAT like-HAB AUX

(4) **ABS ≫ DAT:**

a. * Kepa-ri bere buru-a ji-ten zako isplu-a-b
   Kepa-DAT his head-ARTsg(ABS) come-PROG AUX mirror-ARTsg(ABS)-LOC
   Intended: ‘Kepa is approaching himself in the mirror.’

b. Miren bere buru-a-ri mintzatu zai
   Miren(ABS) his/her head-ARTsg-DAT talk-PRT AUX
   ‘Miren talked to herself.’ [Rezac 2008b:75]

1But see Coon et al. (2017) for a more nuanced view—as well as evidence, from copular clauses in German, for the existence of a corresponding number effect. Importantly, Coon et al.’s results are fully compatible with the approach taken here. That is because the number effect that they find only arises in configurations where the intervener stays in place after person agreement has been attempted. In canonical PCC configurations, like the ones discussed here, the intervener (in this case, the indirect object) undergoes clitic doubling right after person probing, and ceases to be an intervener; see Béjar & Rezac (2003) and Preminger (2009), a.o., for details.

This compares favorably with approaches such as Nevins 2007, where the person-specificity of the PCC is derived from an ontological difference between person and number features. The latter would have trouble accommodating data of the sort Coon et al. discuss.
Crucially, dat ≫ abs verbs show the PCC, while abs ≫ dat ones don’t:

(5) dat ≫ abs:
   a. Miren-i gozoki-ak gusta-tzen φ-zai-zki-o
      Miren-DAT sweet-Art2(abs) like-IMPF 3.abs- √ pl.abs-3sg.dat
      ‘Miren likes candy.’
   b. * Nime gusta-tzen me(abs) Miren-DAT like-IMPF 1.abs- √ sg.abs-3sg.dat
      ‘Miren likes me.’

(6) abs ≫ dat:
   Ni Peru-ri hurbildu me(abs) Peru-DAT approach 1.abs- √ sg.abs-3sg.dat
   ‘I approached Peru.’

Things to note:
   ◦ the ‘target forms’ in (5b) and (6) are identical
     – they are not merely phonologically identical; they are morpho-
       syntactically identical:
       · the two express the same set of features, {1sg.abs, 3sg.dat}
   ◦ the distinction between the two cases is only in the finer hierarchical
     organization of the relevant arguments

NB: This is also a good argument against ‘usage’- or ‘grammaticalization’-based
approaches to the PCC (e.g. Haspelmath 2004). Whatever you want to say about
the target form in (5b), it is clearly not missing from the grammatical vocabulary
of the language.

This is a distinction that lives in the module of grammar known as syntax.

In particular, the PCC as it applies to ditransitives (see (1–2), above) is a subcase
of a broader pattern:

(7) “Generalized” PCC in Basque:
   In finite clauses that have a dat argument located higher than the abs
   argument, the abs argument must be 3rd person.
   (Recall that all triadic ditransitive verbs are dat ≫ abs.)

A sketch of a syntactic account of the PCC (following Béjar & Rezac 2003):

(8)

   • roughly: a 1st/2nd person argument that could have, case-wise, been
targeted for ϕ-agreement—but wasn’t—is cause for ungrammaticality
   • lots more to say here… but the crucial point for us right now is this:
     – syntactic accounts like (8) are the ones equipped to predict (5–6)
       because the interruption of agreement shown in (8) will only obtain if the
dat DP is structurally higher than the theme DP

6. The PCC is sensitive to the overtness of ϕ-agreement
   • The PCC is famously absent in environments that lack overt agreement
     morphology with the internal arguments of the verb
   • This is true cross-linguistically —
     ◦ Hebrew lacks agreement morphology with internal arguments; and Hebrew
does not exhibit the PCC:

(9) a. dat ≫ acc . . .
   ha-mehapnet-etk/ΦM la-cofei et acmoi
   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.’)

b. . . . but no PCC
   ha-menhel-et ta-cig la-hem oti<1sg.DO>
   the-manager-F fut.3sg.F-introduce dat.the-them acc.me
   ‘The manager will introduce me<1sg.DO> to them.’
• And it is also true intra-linguistically —
  o non-finite environments in Basque (incl. nominalizations) lack agreement
    morphology; and these environments do not exhibit the PCC:

(10) *Zuk harakin-ari ni saldu n-(a)i-φ-o-zu
  you.e butcher-ARTg.D me(A) sell 1.A-√'-sg.A-3sg,D-2sg,E
  ‘You have sold me to the butcher.’
  \[=(2b)\]

(11) Gaizki irudi-tzenlook -zai-φ-t
  k nime
  [ zuk ni harakin-ari sal-tze-a ]
  wrong look-impf 3.A-√'-sg.A-1sg,D
  you.e me(A) butcher-ARTg.D sold-NMZ-ARTg(A)
  ‘It seems wrong to me for you to sell me to the butcher.’
  \[Laka 1996\]
  o the embedded non-finite clause in (11) contains the same verb, and the same
    combination of arguments, as the ungrammatical (10)
  ↳ but this embedded clause lacks agreement morphology, and the PCC
    does not arise

Now let us juxtapose this with the results of §5:
• We have, in the PCC, a syntactic effect par excellence—which nevertheless
  only arises in the presence of overt agreement morphology

⇒ How can something in narrow syntax be sensitive to the overtness of agreement
    morphology?

As best I can tell, the only possible answer is this:²

The mechanisms of agreement & intervention, implicated in the PCC, are only in
place when we can see them.

• To put it another way:
  o there is generally no such thing as “abstract” agreement, null across the
    entire paradigm
  (There is no prohibition against particular cells being null in what is otherwise an overt
    paradigm: the PCC still arises in such cases.)

⇒ The PCC goes away in the absence of overt agreement morphology not because
  it is a morphological filter —
    (we already saw in §5 that the PCC cannot be a morphological filter)
  — but because, in the absence of overt agreement morphology, there is no
    agreement there, not even “abstract” agreement.

7. The clitic-doubling caveat
• Our characterization so far of the intra- and cross-linguistic distribution of PCC
  effects has been idealized in one important respect:
  o it abstracted away from the distinction between ϕ-agreement and clitic
    doubling, treating them both as “agreement morphology”

A quick refresher:

(12) a. \[Kat\]e tis kid\[her\]t\]
    sinodhepse accompanied [the\[his\]] mother\] [to every\[child\]]
    [every\[mother\]] accompanied [the\[child\] his\] ACC
    ‘[Every mother] accompanied [her\], child\].’
  b. ?*\[Kat\]e tis kid\[her\]t\]
    sinodhepse accompanied [the\[his\]] mother\] [to every\[child\]]
    [every\[mother\]] accompanied [the\[child\] his\] ACC
    ‘[His\] mother\] accompanied [every\, child\].’
  c. \[Kat\]e tis kid\[her\]t\]
    to\[the\] sinodhepse accompanied [the\[his\]] mother\] [to every\[child\]]
    [every\[mother\]] accompanied [the\[child\] his\] ACC
    ‘[Every mother\] accompanied [her\], child\].’
  d. \[Kat\]e tis kid\[her\]t\]
    to\[the\] sinodhepse accompanied [the\[his\]] mother\] [to every\[child\]]
    [every\[mother\]] accompanied [the\[child\] his\] ACC
    ‘[His\] mother\] accompanied [every\, child\].’

- \( \varphi \)-agreement:
  
  - there is no particular reason to expect that the exponents of \( \varphi \)-agreement will resemble the free-standing pronouns of the language\(^3\)
  
  - moreover, it is possible for these exponents to exhibit allomorphy, and even suppletion, based on the (other) features of the head \( H^0 \)
    
    - a common example would be the agreement exponents in one tense/aspect differing from those found in another tense/aspect
    
    - e.g. \textit{am - are - is} vs. \textit{was - were - was}

- clitic doubling:
  
  - doubled clitics do not, in contrast to \( \varphi \)-agreement, exhibit allomorphy based on specific features of their hosts
  
  - we may expect that at least in some cases, doubled clitics will bear morphophonological resemblance to the free-standing pronouns of the language

Another example of clitic doubling:\(^4\) (Basque; Laka 1996)

\begin{equation}
\begin{array}{lr}
\text{(14)} & \text{Guraso-e-k} \quad \text{ni-ri} \quad \text{barritako errak} \quad \text{erosi} \\
& \text{parent(s)-ART_{pl}}^{-ERG} \quad \text{me-DAT} \quad \text{earring(s) beautiful-ART_{pl}}^{(ABS)} \quad \text{bought} \\
& \text{d-i-zki-da-te} \\
& \text{3.ABS-\text{\text{-}}\text{pl} \text{LABS-1SG.DAT-3PL.ERG}}
\end{array}
\end{equation}

\text{‘(My) parents have bought me beautiful earrings.’}

Further issues:

(i) apparent sensitivity to referential properties of the doubled nominal

- clitic doubling is not, generally speaking, conditioned by factors like animacy, specificity, definiteness, etc. etc.

- nor is it optional, in the general case

- clitic doubling in (14), for example, is entirely obligatory, irrespective of the properties of the doubled nominals
  
  - they can be definite or indefinite; quantificational or not; rigid designators or not; etc.

\Rightarrow \text{ therefore, even in languages where clitic doubling looks like it is sensitive to such nominal properties — (Portuñol Spanish; Suñer 1988:396)}

\begin{equation}
\begin{array}{lr}
\text{(15) a.} & \text{La i} \quad \text{niñag} \quad \text{e-} \quad \text{niños} \\
& \text{cl hear.PAST.3PL} \quad \text{Paca} / \text{the girl} \quad \text{la gata} / \text{the cat} \\
& \text{‘They listened to Paca / the girl / the cat.’}
\end{array}
\end{equation}

\begin{equation}
\begin{array}{lr}
\text{b.} & \text{(*La i)} \quad \text{(*niñag)} \\
& \text{buscaban} \quad \text{los ayudara)} \\
& \text{(*CL) search.PAST.3PL} \quad \text{a somebody comp cl.pl help.SUBJUNCT} \\
& \text{‘They were looking for somebody who could help them.’}
\end{array}
\end{equation}

- it would be a mistake to build this sensitivity into the clitic-doubling operation itself

- 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, Sportiche 1998, Merchant 2006, \textit{a.o.})

- importantly, animacy/definiteness/specificity are known to regulate A-movement of DPs, even in languages that lack clitic doubling entirely
  

- it would therefore be redundant to build this sensitivity into the clitic-doubling operation itself

(ii) locality

- for the purposes of locality, doubled DPs behave like traces of A-movement (Anagnostopoulou 2003, \textit{a.o.})
  
  - which are known to be non-interveners, at least for \( \varphi \)-agreement / further A-movement operations (Holmberg & Hróarsdóttir 2003, \textit{a.o.})

\footnotesize\text{\(^3\)Diachronically, \( \varphi \)-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 \( \varphi \)-agreement rather than clitic doubling, there is no longer any principled reason to expect such similarity. Indeed, it is possible that sound changes affecting doubled clitics could serve as a catalyst for the diachronic realanalysis of clitics into \( \varphi \)-agreement in the narrow sense.}

\footnotesize\text{\(^4\)For extensive argumentation that these are indeed instances of clitic doubling, see Arregi & Nevins (2008, 2012), as well as Preminger 2009.}
example:5

(16) a. [[Miren-entzat]_pp [harri horiek]_abs altxa-tze-n] probatu
   Miren-ben stone(s) those_pot ABS lift-NMZ-LOC attempted
   [d-it-u-zte]_aux
   3.ABS-pl.ABS-V-3pl.ERG
   ‘They have attempted to lift those stones for Miren.’

b. [[Lankide-e-i]_DAT [liburu horiek]_abs irakur-tze-n] probatu
   colleague(s)-ART_pl-DAT book(s) those_pot(ABS) read-NMZ-LOC attempted
   [d-/it-u-(z)te]_aux
   3.ABS-sg/pl.ABS-V-3pl.ERG
   ‘They have attempted to read those books to the colleagues.’

⇒ dative DPs are interveners in Basque
   • but they aren’t interveners when they’ve been clitic-doubled;
     compare (16b) with (14):

(14) 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-ic.
    3.ABS-V-pl.ABS-1sg.DAT-3pl.ERG
    ‘(My) parents have bought me beautiful earrings.’
    – in (14) (a monoclausal ditransitive), the dative DP has been
      clitic-doubled, and agreement with the plural absolutive DP goes
      through

(iii) clitic doubling vs. cliticization

• since clitic doubling exists, and pro exists, I am going to assume there is
  no third thing called (syntactic) ‘cliticization’;
• rather, there is simply clitic doubling of pro

⇒ But if clitic doubling is not agreement, why is its occurrence enough to give rise
  to PCC effects?

8. A non-explanation: agreement as a precursor to all
   DP movement

Here’s an answer that doesn’t work: “Because clitic doubling is DP movement, and
all DP movement is prefigured by an agreement relation.”

(17) A two-step approach to DP movement (Chomsky 2000, 2001)

1. H0 enters into an Agree relation in \( \varphi \)-features with a DP
   \( \rightarrow \) and subsequently/consequently:
2. the DP moves to the domain of H0 (= [Spec,HP])

We know that (17) is wrong (at least as a general requirement). Here’s why:

• Overt agreement in \( \varphi \)-features is double-dissociable from DP movement
  • e.g. in Icelandic, there are sentences where agreement targets a nominative
    non-subject, while a non-nominative DP (e.g. a dative) moves to subject
    position despite failing to control \( \varphi \)-agreement:

(18) [Einhverjum stúdent]_sg.finnast
    some student.SG.DAT find.pl.
    t[ölcvunarn] computers.the.PL.NOM ugly
    ljótar
    ‘Some student finds the computers ugly.’ [Holmberg & Hróarsdóttir 2003:999]

• The standard retort to this is the abstractness gambit:

---

5The data in (16a–b) are from ‘substandard’ varieties of Basque; see Etxepare (2006:303n2).
all DP movement is still prefigured by agreement, it's just “abstract”
agreement (lacking in morpho-phonological expression)
   – and so, non-nominative subjects are still targeted for agreement prior to
undergoing DP movement

But we just saw that this cannot be the case…
   o there is never any overt \( \varphi \)-agreement with datives in Icelandic
   o and, as we just saw in §6:
     – there is no such thing as syntactic \( \varphi \)-agreement that never receives any
morpho-phonological expression
⇒ there can be no \( \varphi \)-agreement with datives in Icelandic.
   o to wit, Icelandic doesn’t have person restrictions affecting DAT DPs
     – just as we would expect, given that NOM DPs are targeted for
agreement (as is obvious from the overt morphology)
   – see Preminger 2011b, Sigurðsson & Holmberg 2008, and references
therein, for discussion

Overall, the conclusion is that datives are never agreed with in Icelandic
(overtly or otherwise)
   – and yet they can, and sometimes do, undergo DP movement to subject
position
⇒ Chomsky’s (2000, et seq.) proposal that all DP movement is prefigured by
agreement is simply wrong.

What we’re looking for, then, is a reason why clitic doubling must have to be
prefigured by an agreement relation
   o even though movement in general, and even DP movement in particular,
do not have to be

In particular, it is head movement that “skips” at least one c-commanding head
in its path, thus violating Travis’ (1984) Head Movement Constraint (HMC)

To see why this is, consider what it would look like if clitic doubling /
cliticization did comply with the HMC…
   o if clitic doubling respected the HMC, we would expect the constituent
structure in (19):

(19) \{auxiliary/TAM, \{(transitivity/voice), \{clitic, \(y/X\}\}\}\}

whereas what we actually find looks like (20):

(20) [L’as]-tu fait?
    [cl.-have]-you done
    ‘Have you done it?’
⇒ clitic doubling / cliticization, qua head movement, is movement of D at least
as far as \( v \) (thus, necessarily skipping over \( V^0 \))
   – and often further still: e.g. movement of D to T, skipping over \( V \) and \( v \)
   (and Asp, and . . . )

One might be tempted to take this incompatibility with the HMC as evidence
against a head-movement account of clitic doubling
   – but: we know violations of the HMC exist, even independent of clitic
doubling / cliticization

(21) a. Lenn a ra Anna al levr.
    [cl.-inf]-read does Anna the book
    ‘Anna reads the book.’
   b. Lennet en deus Anna al levr.
      [cl.-p-prt has] has Anna the book
      ‘Anna has read the book.’

9. Towards an account of the clitic-doubling caveat

9.1. Background: clitic doubling as long head movement
   - Clitic doubling (incl. syntactic cliticization) is non-local head movement
     o see Rezac (2008a), Roberts (2010)
What the PCC tells us about “abstract” agreement, head movement, and locality

February 2017

Preminger

(22)

[Roberts 2010:194ff.]

– for discussion and, crucially, for arguments that these Breton data indeed involve long head movement (rather than, say, remnant VP-fronting) —

– and for evidence of long head movement in other empirical domains —

• So here’s the structure of clitic doubling:

(23)

• At this juncture, you might be asking: why, then, is D\textsuperscript{0} pronounced twice?

(24) [I mitera tu\textsubscript{k}i, \textit{to k} sinodhepse \textsubscript{t} t\textsubscript{i} t\textsubscript{s} \textit{kathe pedhi}]\textsubscript{k} [the mother his], SOM cl.ACC accompanied [the every child],ACC

‘His mother accompanied every child,’

[Modern Greek; = (13d)]

(25) li-kro et ha-sefer, hi kvar kar’a inf-read acc the-book, she already past.read-3sgF

‘She has already read the book.’ (lit. ‘Read the book, she already has.’)

(26)

• ANSWER:
Under certain conditions, head movement gives rise to “double pronunciation”
  ○ as an example, consider predicate clefting in Hebrew
    (see Landau 2006 for details on this construction)

(27) CONDITIONS ON PHONOLOGICAL CHAIN REDUCTION OF HEAD MOVEMENT

Let X\textsuperscript{0} be a head that undergoes movement to Y\textsuperscript{0}, and let \( \alpha \) be the lower copy of X\textsuperscript{0}. \( \alpha \) will be phonologically deleted iff either of the following conditions is met:

(i) \( \alpha \) and Y\textsuperscript{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\textsuperscript{0} c-commands \( \alpha \) in the surface structure (i.e., no constituent containing \( \alpha \) but not Y\textsuperscript{0} has undergone subsequent movement to a position above Y\textsuperscript{0})

○ for reasons of time, I won’t go into this in more detail here
  – see Preminger (to appear)
• and for ways in which (something like) (27.i–ii) might be derivable from more basic assumptions, see Gribanova & Harizanov 2016

To recap, what we’re looking for is this:

• a reason why, unlike phrasal movement, non-local head movement (as in (23)) does require a prior agreement relation
  – because that would explain why clitic doubling triggers the PCC (which, you’ll recall, is contingent on syntactic agreement)

9.2. Head movement & locality

• Here, I build on an idea by Hornstein (2009:72–74) and Roberts (2010:33–40):
  o Bare Phrase Structure + iterative downward search conspire to yield an A-over-A-like effect —
    – ruling out most, but not all, instances of head movement

• Here’s how it works:
  o Bare Phrase Structure (Chomsky 1994) tells us that the label of the entire phrase (“XP”) is nothing but the head itself
    – in other words, the distinction between $X^0/X_{\text{min}}$ and $XP/X^{\text{max}}$ can be defined relationally —
      • but it cannot be defined featurally

  o if movement is viewed from the perspective of the attractor (or in earlier versions of the theory, the perspective of the landing site) —
    – an iterative downward search algorithm, looking for a node to move (or remerge), will encounter the phrasal node first

  (28)

  ![Diagram](image)

  o and because the two are literally one and the same object:
    – there is no possible featural search criterion that would result in skipping the phrasal label (“XP”) in favor of the head (“$X^0$”)

As it stands, this predicts that head movement should never be possible, which is obviously too strong …

9.3. The Principle of Minimal Compliance (Richards 1998, 2001)

(29) a. * [Which book]$_k$ did the journalist spread the rumor that the senator wanted to ban $t_i$?
   b. ? [Which journalist], $t_i$ spread the rumor that the senator wanted to ban [which book]$_k$?

---

*Here is an explicit algorithm for iterative downward search:

(i) a. Let $P$ be a syntactic probe, and let $XP$ be $P$’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”

Looking at (29a–b), one might be tempted to proffer an explanation along the following lines:

- “(29b) is better because, in that example, the wh-phrase generated inside the island does not try to move out”

What Richards (1998, 2001) shows is that such an explanation is at best insufficient:

(30) a. * [Koja kniga]k razprostranjava-še žurnalista-t [mávlata če senator-ati da zabrani tiš]?
   [Which book]k did the journalist spread the rumor that the senator
   wanted to ban
   "[Which book]k did the journalist spread the rumor that the senator
   wanted to ban [which book]k?"

b. ? [Koja žurnalista]t [koja kniga]k tš razprostranjava-še [mávlata če
   which journalist which book spread rumor that
   senator wanted to ban
   "[Which journalist]tš spread the rumor that the senator wanted to
   ban [which book]k?"

Once one wh-phrase has moved to [Spec,CP] of the interrogative clause —
- other potential movers to the same position are exempt from Subjacency
  (or whatever else is responsible for strong islands)

[nb: This is pretty bad news for the idea that syntactic islands can be derived
from the spellout (or Transfer) of phases.]

Importantly, these effects pertain to multiple relations established with the same
probe / landing-site; accordingly —
- the amelioration exemplified by (30b) does not arise if the island-violating
  movement targets a C0 that no other, well-formed movement has targeted:

(31) * [Kakvo]k kazva tozi služitel na [žurnalista, kojto tš razsledvati tiš], če
   what tells this official to journalists who investigate that
   communists have deceived their(editors) their
   Intended: "[What]k does this official tell journalists who, tš are
   investigating tiš] that the communists have deceived their editors?"

[Richards 1997:256]

(32) **Principle of Minimal Compliance (PMC)**

[my version; slightly revised from Richards 1998, 2001]

Once a probe \( P \) has successfully targeted a goal \( G \), any other goal \( G' \) that
meets the same featural search criteria, and is dominated or c-commanded
by \( G (= \text{dominated by the mother of } G) \), is accessible to subsequent probing by \( P \)
irrespective of locality conditions.

### 9.4. Minimal Remerge, head movement, and Anti-Locality

- Given the PMC, only the first syntactic relation targeting \( X \) will be subject to the
  aforementioned A-over-A-like locality condition
  - in other words, only the first syntactic relation is obligated to
    target \( XP/X_{\text{max}} \)

- But this, on its own, is not enough to give rise to head movement
  - that’s because, as we already pointed out, there is no featural search criterion
    that could favor \( X_{\text{0}}/X_{\text{min}} \) over \( XP/X_{\text{max}} \)

\( \Rightarrow \) So the question remains: how does head movement (i.e., movement of \( X_{\text{0}}/X_{\text{min}} \)
alone) ever arise?

**Proposal:**

(33) **Minimal Remerge**

If \( X_{\text{0}}/X_{\text{min}} \) is movable, move only \( X_{\text{0}}/X_{\text{min}} \).

- Recall: in many (perhaps most) cases, \( X_{\text{0}}/X_{\text{min}} \) is not movable
  - because of the A-over-A-like locality condition

- Only when some prior syntactic relation has already targeted \( XP/X_{\text{max}} \) does
  the PMC nullify this locality condition, rendering \( X_{\text{0}}/X_{\text{min}} \) movable.

\[ \Rightarrow \] Now it’s time to ask: what are examples of such “prior syntactic relations” that
would give rise to the possibility of head movement?

---

7There is a certain affinity (in spirit, though, crucially, not in technical detail) between (33) and
Chomsky’s (1995:262ff.) suggestion that—all else being equal—feature-movement is the preferred
mode of syntactic movement. See also Donati (2006:29–30).
Here’s one example:

- Let H be a movement attractor, and let X be the head of H’s complement
  ⇒ it follows that H is in a c-selection relation with X(P)
  (indicated here as a “wavy line”)

\[(34)\]

\[
\begin{array}{c}
H \rightarrow \rightarrow \rightarrow X P(=X) \\
\cdots \text{ X' (=X) } \\
\text{ X'(0) (=X) } \cdots
\end{array}
\]

- We therefore have in place a structural relation between H and X(P), which conforms to the aforementioned A-over-A-like condition

**It Note:** This point, concerning c-selection being implementable under pure sisterhood, was one of the selling points of Bare Phrase Structure in the first place.

⇒ Given the PMC (32), subsequent relations between H and X are no longer subject to this locality condition
  - meaning it is now possible for H to attract the head of X alone (35a)

- And, crucially, in situations where both phrasal movement and head movement are in principle possible, Minimal Remerge (33) can exert its force
  - ruling out phrasal movement (35b) in favor of head movement (35a)

\[(35)\]

a. \[
\begin{array}{c}
H'(=H) \\
\text{ H } \rightarrow \rightarrow \rightarrow X P(=X) \\
\cdots \text{ X' (=X) } \\
\text{ X'(0) (=X) } \cdots
\end{array}
\]

b. \[
\begin{array}{c}
H'(=H) \\
\text{ H } \rightarrow \rightarrow \rightarrow X P(=X) \\
\cdots \text{ X' (=X) } \\
\text{ X'(0) (=X) } \cdots
\end{array}
\]

- You’ll notice that (35b) is essentially Abels’ (2003) Anti-Locality condition —
  - now derived from the PMC (32) + Minimal Remerge (33)

- Importantly: on this view, there is nothing intrinsically wrong with moving the complement of H to [Spec,HP]
  - it is only a problem when there is something smaller that could have moved

- In the case where the complement of H is non-branching, Minimal Remerge is trivially satisfied
  ⇒ meaning we expect no Anti-Locality effects in this case

- This is arguably what’s going on in cases of head movement of a complement-less head (Matushansky 2006)

\[(36)\] the Matushansky exception to Anti-Locality

- Importantly: on this view, there is nothing intrinsically wrong with moving the complement of H to [Spec,HP]
  - it is only a problem when there is something smaller that could have moved

- In the case where the complement of H is non-branching, Minimal Remerge is trivially satisfied
  ⇒ meaning we expect no Anti-Locality effects in this case

- This is arguably what’s going on in cases of head movement of a complement-less head (Matushansky 2006)

\[(36)\] the Matushansky exception to Anti-Locality

10. Clitic doubling, locality, and acquisition

Recall that clitic doubling is non-local head movement (section 9.1) —

- In this case, the probe (e.g. v<sup>0</sup>) is not related to the goal (the DP in [Compl,V]) via c-selection
  - because the goal DP is the complement of V (which is probably never the cliticization host; see section 9.1, above)

- If clitic doubling has successfully obtained, there must have been a prior agreement relation between the probe and (the phrasal projection of) the goal
  - otherwise, the A-over-A-like condition would have prohibited head movement of D alone
• It is by virtue of this prior agreement relation that the $A$-over-$A$-like condition has already been satisfied once —
  ◦ and, due to the PMC, subsequent relations are no longer subject to it

(37)

• But doesn’t this violate the earlier dictum on “no null agreement”?  
  ◦ after all, in many cases of clitic doubling, there seems to be no overt morpho-phonological expression of a prior agreement relation:

(38) Levi al profesor ayer
  cl. I saw A-the professor yesterday
  ‘I saw the professor yesterday.’

  ◦ in (38), for example, there is no overt exponent of an agreement relation between $v^0$ and the object —
    (the verb displays subject-agreement, but that is irrelevant here; the point is that there is no object-agreement independent of the clitic)
  — and this is paradigm-wide, i.e., it is not a matter of the particular $\varphi$-features of profesor

NB: There are instances of clitic doubling 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 (37), it does not change the facts of (38) and many cases like it.

⇒ The answer, obviously, is “yes”:
  ◦ this does violate the aforementioned dictum (“no null agreement”)
  ⇒ meaning this dictum cannot be an actual, combinatorial principle of grammar

• Before moving on, it is worth pointing out that were this dictum an actual principle of grammar, it would raise the same modularity issue discussed in §6
  ◦ as a grammatical principle, a ban on null agreement would require simultaneous reference to syntax and morpho-phonology

• 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;
  ◦ the requirement is that some cells in the paradigm must be overt

My proposal is that the no null agreement dictum is not a principle of grammar at all; it is an acquisition strategy.

• The learner starts with the assumption in (39):

(39) initial state of acquisition
  There are no unvalued $\varphi$-features on any functional heads.

• There is a specific set of triggers that can prompt the learner to revise (39), and posit unvalued $\varphi$-features on a particular functional head:

(i) overt morpho-phonological covariance in $\varphi$-features between a functional head and a DP
(ii) long-distance head movement (of a D head)

• But crucially, the list of things that may prompt the learner to revise (39) is anything but open-ended —

  for example, the list absolutely cannot include “there’s a binding / fake-indexical / … relation between F and a DP”
  ◦ that would render it impossible to properly account for the distribution of PCC effects
    – recall that binding / fake-indexicals / … do not give rise to the PCC;
    – it only arises in the presence of overt $\varphi$-agreement or clitic doubling (and recall that this cannot be a morphological matter)
  ◦ the conclusion, as already argued elsewhere—and on independent grounds—is that these other relations do not involve syntactic agreement in $\varphi$-features
    – see also Preminger 2013, 2015, Preminger & Polinsky 2015
Returning to the main point... we now have an acquisition roadmap for how a “PCC language” is acquired:

- The learner starts with the assumption that there are no unvalued $\varphi$-features on $v$ (as mandated by (39))
- Very quickly, however, she will be driven to revise this assumption:
  - either because $v$ shows overt covariance in $\varphi$-features with the direct object, as in Basque
  - or because there is a $D$ associated with the direct object that cliticizes to $v$ (i.e., $D$ undergoes long-distance head movement), as in Spanish
    - meaning there has to be a prior agreement relation between $v$ and $DP$
      - satisfying the A-over-A-like condition once, and allowing (through the PMC) for subsequent operations to ignore it
  - and, importantly, misidentifying one of these ($\varphi$-agreement or clitic doubling) as the other will be completely innocuous, at this stage
    - since, in both cases, syntactic agreement in $\varphi$-features will have been posited
  - this is good news because agreement vs. clitic doubling is a frequent nexus of diachronic change (~ misanalysis)
    - see, e.g., van Gelderen (2011) and references therein
- Once unvalued $\varphi$-features on $v$ have been posited, the PCC arises as a consequence of the mechanisms and agreement $&$ intervention discussed in §5

11. Conclusion
In this talk, we have seen:

- That the PCC is a fundamentally syntactic phenomenon (§5)
- That it is nevertheless sensitive to the presence of overt $\varphi$-agreement, meaning there is no agreement where we cannot see agreement (no null agreement; §6)
- That there is a caveat to this, whereby clitic doubling behaves as though it involved agreement (even though clitic doubling is itself not agreement; §7)
- That we cannot maintain that all movement (or even just all DP movement or A-movement) is prefigured by $\varphi$-agreement (§8)
  - and so this cannot explain the clitic-doubling caveat
- That a more promising alternative can be found by investigating the interplay of Bare Phrase Structure, iterative downward search, and the Principle of Minimal Compliance (§9)
  - in particular, the idea is that movement always “wants” to move only the head, but this is seldom possible due to an A-over-A-like locality condition
  - but if the landing site stands in some prior syntactic relation (e.g., c-selection or $\varphi$-agreement) to the relevant phrase, this satisfies the locality condition once
    - and, given the PMC, subsequent syntactic operations need not adhere to the locality condition
    $\Rightarrow$ thereby enabling head movement
- Clitic doubling, qua long-distance head movement, cannot be prefigured by c-selection (a strictly local relationship)
  - and therefore, requires a prior agreement relation to have satisfied the A-over-A-like locality condition
  $\Rightarrow$ explaining why clitic doubling always entails a syntactic agreement relation
    - and thus, why clitic doubling “counts” for the purposes of the PCC (§10)
- Finally, I showed how this picture could arise through a conservative acquisition strategy regarding the distribution of unvalued $\varphi$-features on functional heads
  - which has, as its consequence, the no null agreement generalization, as well as its clitic-doubling caveat

References

svx revision code: 9480