Binding and coreference in Mayan: Evidence for object raising

Justin Royer
McGill University — justin.royer@mail.mcgill.ca
Berkeley SSCircle — November 13th 2020

---

1 Context

Given standard views on binding (e.g. Reinhart 1976, 1983; Chomsky 1981, 1986), the distribution of co-indexed nominals depends in part on c-command.[1]

- Since subjects c-command objects, Condition C will enforce that only the subject DP should be an R-expression in a VOS configuration like (1)—regardless of surface word order:

  (1) verb $[_{\text{OBJ}} \ldots [_{\text{POSS}} \text{pronoun}_1]]$ $[_{\text{SUBJ}} \underline{\text{R-expression}_1}]$

- Some VOS languages, like Ch'ol, pattern as expected given binding principles:

  (2) Tyi i-chok-o $[_{\text{OBJ}}$ i-tyuñ $[_{\text{POSS}}$ Ø $]_1]$ **abi** $[_{\text{SUBJ}}$ jiñi alob $]_1$.
  PFV A3-throw-TV A3-stone his yesterday DET boy
  'The boy$_i$ threw his$_i$ stone yesterday.' Ch'ol (Mayan)

- The subject (R-expression) binds the possessor of the object (pronoun); linear order is irrelevant.

But not all VOS languages are "well-behaved" like Ch'ol.

- In Chuj, another Mayan language, **linear order—and not c-command**—is what seems to govern the distribution of co-indexed nominals (sometimes in apparent violation of Condition C):

  (3) verb $[_{\text{OBJ}} \ldots [_{\text{POSS}} \underline{\text{R-expression}_1}]]$ $[_{\text{SUBJ}} \text{pronoun}_1]$

- Similar data in closely-related Popti' (Jakaltek) led Craig (1977), Hoekstra (1989), and Aissen (2000) to suggest that binding principles do not apply in the same way across languages.

[1]I would like to thank the Chuj consultants who generously and patiently contributed to this project: Matal Torres and Mateo Pablo. For judgements and help with other cited Mayan languages, I thank Morelia Vázquez Martínez, Jessica Coon, and Carol-Rose Little (Ch'ol); Gilda Ixem Cum (Kaqchikel), Henry Sales and Tessa Scott (Mam); Pedro Mateo Pedro and Jorge Juárez (Q'anjob'al); Hugo Héctor Vázquez López (Tojol-ab'al), and Jaime Pérez González (Tseltal). I am especially thankful to Jessica Coon and Aron Hirsch for valuable guidance, and I also thank Dan Brodkin, Peter Jenks, Carol-Rose Little, Pedro Mateo Pedro, Martina Martinović, Jaime Pérez González, Rodrigo Ranero, Tessa Scott, Hugo Héctor Vázquez López, the McGill syntax-semantics reading group and Fieldwork lab, and the audience of NELS 51 for discussion.

Plan today:

- Novel data showing that linear precedence governs co-indexed nominals in Chuj.
- Argue that the role of linear precedence in Chuj and Popti' is not due to the absence of binding principles (contra previous work), but is conditioned by a syntactic configuration, namely:

  (i) Objects consistently move over subjects in ("HIGH-ABS") languages like Popti' and Chuj, but not in ("LOW-ABS") languages like Ch'ol, as independently-proposed in Coon et al. 2014.

  (ii) Object raising bleeds c-command relations between co-indexed DPs, which I will argue explains why binding principles don't seem operative.

- Argue that binding under c-command is actually still needed in these languages to account for the distribution of reflexive objects—in such cases linear precedence is irrelevant.
- Hypothesize that when binding under c-command is impossible, two coreferring DPs are generated and PF elides the one that comes linearly second.

  ➽ *Theoretical consequence* (building on Aissen 2000): Indices are needed in syntax.

  ➽ *Typological consequence*: Binding as a new correlate of HIGH-ABS vs. LOW-ABS Mayan languages.

Roadmap:

2 The puzzle from Chuj: Linear precedence matters

2.1 Chuj

- Chuj belongs to the Q'anjob'alan branch of Mayan languages and is spoken in Guatemala and Mexico by approximately 70,000 speakers (Piedrasanta 2009; Buenrostro 2013).
- It is a close relative of Popti', the main language of study in Craig 1977 and Aissen 2000.
- Data were collected in Canada using a hypothesis-driven fieldwork methodology (Matthewson 2004, Davis, Gillon, and Matthewson 2014).

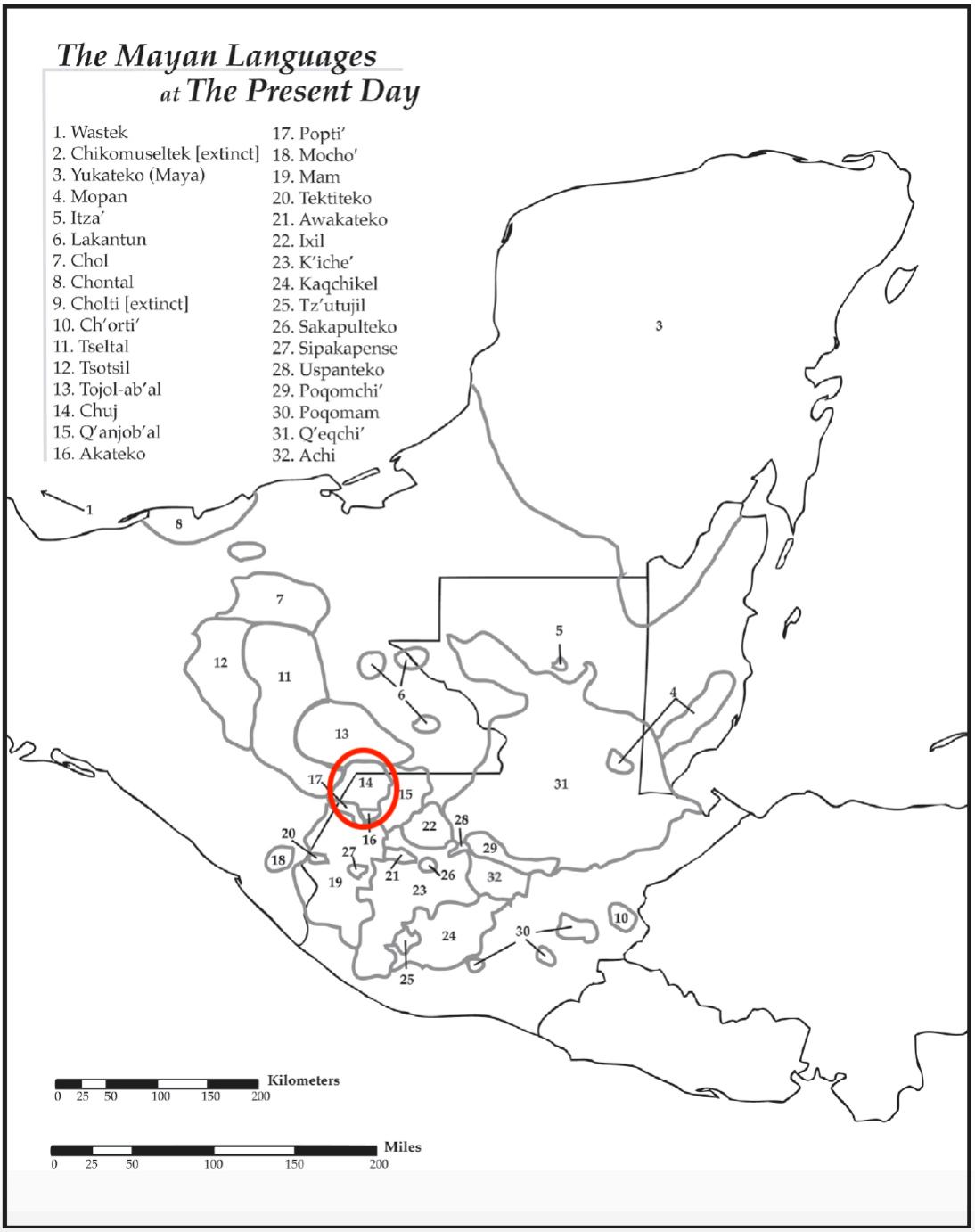


Figure 1: Current-day Mayan-speaking area (Law 2014, p. 25)

2.2 Background on the grammar of Chuj

- VOS word order:[2]

(4) Ix-y-il [$_{OBJ}$ winh winak] [$_{SUBJ}$ ix ix].
PFV-A3-see CLF man CLF woman
'The woman saw the man.'

- Postnominal possessors:

(5) y-unin [$_{POSS}$ ix ix].
A3-child CLF woman
'the woman's child'

- Subjects and possessors trigger identical "Set A" prefixes (bold), on the verb and noun respectively.

(6) **A**-VERB [OBJECT] [SUBECT]

(7) [**A**-POSSESSEE [POSSESSOR]]

- Chuj also features a set of noun classifiers, which are used as third person pronouns in certain environments (Buenrostro et al. 1989; Royer 2019).

(8) Ix-w-il [**ix** unin].
PFV-A1S-see **CLF** girl
'I saw the girl.'

(9) Ix-w-il [**ix**].
PFV-A1S-see **CLF.PRON**
'I saw her.'

- These classifier pronouns alternate with null pronouns in certain environments described in Table 1.

Table 1: Generalization about classifier pronouns vs. null pronouns

Pronoun type	*Generalization*
classifier pronoun	must **not** be co-indexed with another classifier pronoun or R-expression within a minimal CP including relative clauses
null pronoun (Ø)	**must** be be co-indexed with another nominal expression within a minimal CP including relative clauses

(10) [$_{CP}$ Ix-y-al **waj Xun** [$_{CP}$ to ix-lolon **winh**/*Ø yet' ix s-nun Ø/***winh**]]
PFV-A3-say CLF Xun COMP PFV-speak CLF.PRON with CLF.PRON A3-mother PRON
'Xun$_1$ said that he$_1$ spoke with his$_1$ mother.'

- We'll be focusing on the distribution of null pronouns (Ø).
- I'll argue, as was argued for Popti', that the distribution of Ø is largely governed by linear precedence:

(11) Generalization on the realization of co-indexed DPs in Chuj (and Popti')
If there are two co-indexed DPs within a certain domain, the linearly second DP is Ø.

[2]For more on Chuj grammar, see Hopkins 1967, Maxwell 1981, García Pablo 2007, and Buenrostro 2013.

2.3 Evidence for the role of linear precedence from possessors

- First consider (12), where the subject and possessor have obligatory disjoint reference:

(12) Ix-s-chonh [$_{\text{OBJ}}$ s-wakax [$_{\text{POSS}}$ **ix**]] [$_{\text{SUBJ}}$ ix Ana].
PFV-A3-sell A3-cow CLF.PRON CLF Ana
'Ana$_1$ sold her$_{2/*1}$ cow.'

(13) sold [$_{\text{OBJ}}$ cow [$_{\text{POSS}}$ her$_2$]] [$_{\text{SUBJ}}$ Ana$_1$]

- Consider now an example where the subject and possessor are co-indexed, dubbed the "extended reflexive" in Mayanist literature (Aissen 1999).

(14) Ix-s-chonh s-wakax ix Ana.
PFV-A3-sell A3-cow CLF Ana
'Ana$_1$ sold her$_{1/*2}$ cow.'

- As per (12)/(13), a null expression must either occupy the possessor or subject position in (14):

(15) a. sold [$_{\text{OBJ}}$ cow [$_{\text{POSS}}$ Ø$_1$]] [$_{\text{SUBJ}}$ Ana$_1$] (lit: Ana$_1$ sold her$_1$ cow)
b. sold [$_{\text{OBJ}}$ cow [$_{\text{POSS}}$ Ana$_1$]] [$_{\text{SUBJ}}$ Ø$_1$] (lit: She$_1$ sold Ana$_1$'s cow)

- Given basic assumptions about binding, (15b) should be out on the basis of Condition C.
- Nevertheless, there's evidence that the right parse is (15b)—i.e. only *linear precedence* matters and standard binding principles are not operative.

2.3.1 Adverb placement

- Adverbs can normally intervene between subjects and objects:

(16) S-b'o' tek {ewi} **waj Xun** {ewi}.
A3-make meal yesterday CLF Xun yesterday
'Xun made the meal yesterday.'

- But in minimal pairs in which the object is possessed, and the possessor co-indexed with the subject, the adverb placement options change (notice contrast with Ch'ol in (2)):

(17) S-b'o' s-tek {*ewi} **waj Xun** {ewi}.
A3-make A3-meal yesterday CLF Xun yesterday
'Xun$_k$ made his$_k$ meal yesterday.'

- Assuming adverb placement options remain constant, this suggests that in (17), the possessor is overt and the subject is null, as schematized in (18b).

(18) a. made [$_{\text{OBJ}}$ meal] {yesterday} [$_{\text{SUBJ}}$ Xun] {yesterday} = (16)
b. made [$_{\text{OBJ}}$ meal [$_{\text{POSS}}$ Xun$_i$]] {yesterday} [$_{\text{SUBJ}}$ Ø$_i$] {yesterday} = (17)

- In other words, (17) literally translates as *He$_i$ made Xun$_i$'s meal yesterday.*

2.3.2 Object extraction

- When an object A-bar extracts, as in the focus example in (19), the possessor is overt and the co-indexed subject is null:

(19) [$_{\text{OBJ}}$ Ha s-mam [$_{\text{POSS}}$ **waj Xun**]]$_i$ ix-y-il-a' ___$_i$ [$_{\text{SUBJ}}$ Ø].
FOC A3-father CLF Xun PFV-A3-see-TV PRON
'Xun$_1$ saw *his$_1$ father*.'
Lit: 'He$_1$ saw *Xun$_1$'s father*

- The opposite configuration is ungrammatical:

(20) *[$_{\text{OBJ}}$ Ha s-mam [$_{\text{POSS}}$ Ø]]$_i$ ix-y-il-a' ___$_i$ [$_{\text{SUBJ}}$ waj Xun].
FOC A3-father PFV-A3-see-TV CLF Xun
Intended: 'Xun$_1$ saw his$_1$ father.'

– And note that in LOW-ABS Ch'ol, the pattern is opposite:

(21) CH'OL
[$_{\text{OBJ}}$ I-wakax [$_{\text{POSS}}$ Ø]]$_i$ tyi i-choñ-o ___$_i$ [$_{\text{SUBJ}}$ **aj-Ana**].
A3-cow PRON PFV A3-sell-TV CLF-Ana
Lit: 'Ana$_{i/*j}$ sold *her$_i$ cow*.' (compare with Chuj (19))

2.4 VSO/VOS alternations

- VSO order is exceptionally possible—and optional—in Chuj when the subject is a classifier pronoun:[3]
- In such cases, the subject is overtly realized and the possessor is null (in opposition with the two cases seen above, or (23) below):

(22) Ix-s-chonh [$_{\text{SUBJ}}$ **winh**] [$_{\text{OBJ}}$ s-wakax [$_{\text{POSS}}$ Ø]].
PFV-A3-sell CLF.PRON A3-COW PRON
'He$_i$ sold his$_i$ cow.' (VSO)

(23) Ix-s-chonh [$_{\text{OBJ}}$ s-wakax [$_{\text{POSS}}$ **winh**]] [$_{\text{SUBJ}}$ Ø].
PFV-A3-sell A3-COW CLF.PRON PRON
'He$_i$ sold his$_i$ cow.' (VOS)

[3] In the San Sebastián Coatán dialect, which is not under study here, word order is rigidly VSO (see e.g. Maxwell 1981).

2.5 Evidence for linear precedence from relative clauses

- Consider first a VOS sentence without any co-indexed DPs:

(24) Man y-ojtak-ok laj [$_{OBJ}$ ni unin [$_{RC}$ ix-il-an **ix** t'a parke]] [$_{SUBJ}$ **ix** **Ana**].
NEG A3-know-IRR NEG CLF boy PFV-see-AF CLF.PRON PREP park CLF Ana
'Ana$_1$ doesn't know the boy who saw her$_{2/*1}$ in the park.'

- When the subject is co-indexed with a DP inside the object, it's *the DP inside the object* that gets realized in VOS sentences:

(25) Man y-ojtak-ok laj [$_{OBJ}$ ni unin [$_{RC}$ ix-il-an **ix** **Ana** t'a parke]] [$_{SUBJ}$ Ø] .
NEG A3-know-IRR NEG CLF boy PFV-see-AF CLF Ana PREP park PRON
'Ana$_1$ doesn't know the boy who saw her$_1$ in the park.'
Lit: 'She$_1$ doesn't know the boy who saw Ana$_1$ in the park.'

- The opposite configuration is again impossible:

(26) *Man y-ojtak-ok laj [$_{OBJ}$ ni unin [$_{RC}$ ix-il-an Ø$_i$ t'a parke]] [$_{SUBJ}$ **ix** **Ana**]
NEG A3-know-IRR NEG CLF boy PFV-see-AF PRON PREP park CLF Ana
Intended: 'Ana$_1$ doesn't know the boy who saw her$_1$ in the park.'

2.6 Summary: Linear precedence matters

- Linear precedence governs the distribution of co-indexed nominals in Chuj, as Craig (1977), Hoekstra (1989), and Aissen (2000) proposed for Popti'.
- These facts are surprising from a cross-linguistic perspective, especially since they are not observed in other Mayan languages like Ch'ol (see e.g. (2)).

Crossroads — Either...

1. Binding principles are not universal; or
2. Binding principles *are* universal, but there's something special about the syntax of Mayan languages like Chuj and Popti'.

- Craig and Aissen took the first route:

► Aissen, building on Craig, proposed that binding in Popti' is conditioned by prosody (see appx. A).[4]

- In what follows, we will take the second route: binding principles are universal.

[4] As discussed in Aissen 2000, Hoekstra (1989) argues that the Popti' facts can be explained if binding principles are parameterized such that in some languages they apply at Deep Structure (e.g. English), while in others at Surface Structure (Popti'). However, Aissen (2000) shows that this analysis is not viable for Popti'. In Chuj, this analysis also runs into a number of issues. For instance, it cannot account for basic data like (17).

3 Analysis: Object raising bleeds c-command relations

Main proposal: Mayan languages for which *linear precedence* seems to matter (like Chuj) exhibit a different syntax than languages where only *structure* seems to matter (like Ch'ol).

§3.1 Background: Objects consistently raise to a position above the subject in a subset of Mayan languages, like Chuj and Popti' (Coon et al. 2014).

§3.2 Proposal (part 1): Object raising bleeds c-command relations between objects and subject—this explains why binding principles are not operative.

3.1 Background: The LOW-ABS/HIGH-ABS Mayan languages and object raising

- Chuj and Ch'ol differ on another syntactic level: While Chuj features a subject extraction constraint known as the "Ergative Extraction Constraint" (EEC, Aissen 2017), Ch'ol doesn't (Coon et al. 2014).

(27) Chuj → EEC

a. Ix-ach-y-il ix ix.
PFV-B2S-A3-see CLF woman
'The woman saw you.'

b. *Mach$_j$ ix-ach-y-il-a' t_j?
who PFV-B2S-A3-see-TV
'Who saw you?'

(28) Ch'ol → no EEC

a. Tyi y-il-ä-yety x-ixik.
PFV A3-see-DTV-B2 CLF-woman
'The woman saw you?'

b. Maxki tyi y-il-ä-yety?
who PFV A3-see-DTV-B2
'Who saw you?'

- Coon, Mateo Pedro, and Preminger (2014), Assmann et al. (2015), and Coon, Baier, and Levin (2020) propose that the presence or not of the EEC maps to a deep syntactic difference among two types of Mayan languages, so-called LOW-ABS and HIGH-ABS languages:[5]

(29) Objects do *not* raise in LOW-ABS languages like Ch'ol
[$_{v\text{P}}$ SUBJECT [$_{\text{VP}}$ V OBJECT]]]

(30) Objects raise in HIGH-ABS languages like Chuj
[$_{v\text{P}}$ OBJECT [SUBJECT [$_{\text{VP}}$ V <OBJECT>]]]

- They propose that raising of the object in HIGH-ABS languages creates an **intervention problem** (formalized differently in different works, but not relevant here):

(31) Raising of object in HIGH-ABS languages blocks subject extraction
[$_{\text{CP}}$ __ ... [$_{v\text{P}}$ OBJECT [SUBJECT [$_{\text{VP}}$ V ~~OBJECT~~]]]]

[5] The literature on subject extraction asymmetries across Mayan is rich, and won't be discussed here. For relevant overview and list of references, see Aissen 2017 and Coon et al. 2020.

A few notes:

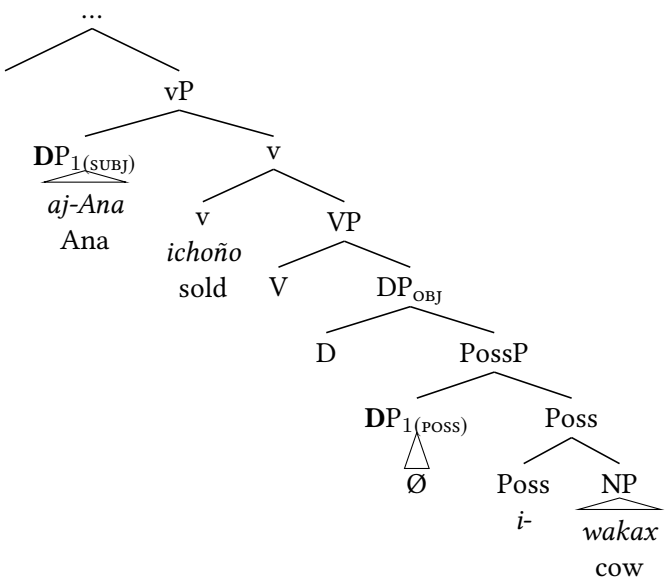
- Object raising is proposed to be driven by an EPP feature on v in Coon et al. 2020; for reasons discussed below, I take this to be an instance of A-movement.
- Object raising correlates with the position of ABS inside the verb stem, see (27) and (28) above.
- Object raising in HIGH-ABS languages does not necessarily correlate with VOS word order (e.g. some previous work argues it is derived post-syntactically, e.g. Clemens and Coon 2018).[6]

3.2 Proposal: Object raising bleeds c-command relations

LOW-ABS-languages: In LOW-ABS languages like Ch'ol, which abides by standard binding principles, the object does **not** raise.

- This means that the subject will necessarily c-command co-indexed expressions inside the object.
- Given Condition C—the R-expression will have to be in subject position:

(32) Structure for 'Ana$_1$ sold her$_1$ cow' in Ch'ol'

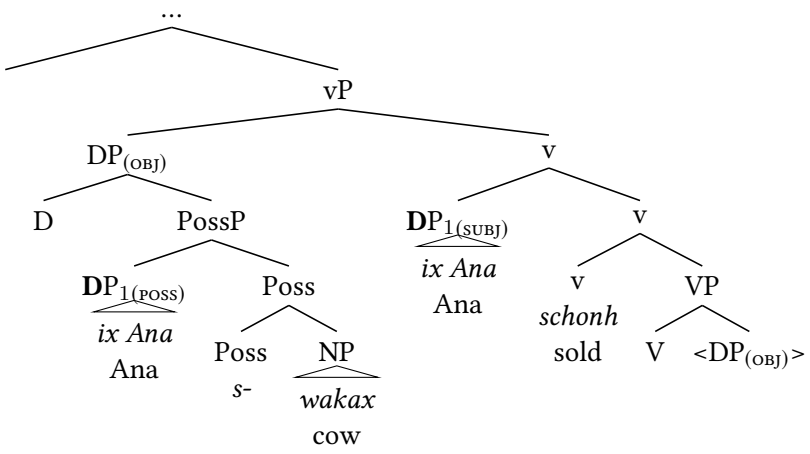


[6] Throughout the talk, I will intentionally ignore how word order is derived here for purposes of illustration. See England 1991, Aissen 1992, Coon 2010, Clemens and Coon 2018, and Little 2020 for varying accounts of VOS/VSO order in Mayan..

HIGH-ABS-languages: Object raising in HIGH-ABS languages will have crucial consequences for binding:

(33) *Consequence of object raising for binding*:
Object raising bleeds c-command relations between subjects and co-indexed DPs inside the object.

(34) Tree for 'Ana$_1$ sold her$_1$ cow' in Chuj



- The subject DP in (34) won't c-command the possessor—and vice-versa; standard binding principles won't rule out the possessor from being overtly realized.

Reconstruction? To account for the apparent absence of traditional binding in HIGH-ABS languages, it will be important that either:

- ► A-movement does not reconstruct for binding (Chomsky 1995, Lasnik 1999, Baltin 2010, etc.);
- ► A-movement reconstruction sometimes happens (e.g. Boeckx 2001; Takahashi 2010), but not in (34).

. . . . . .

Interim summary:

- We saw in §2 that binding principles aren't operative in Chuj; Instead, linear precedence seems to govern the distribution co-indexed nominals.
  - ► Proposal: Object raising bleeds c-command relations between subjects and co-indexed DPs inside the object, and so binding principles become irrelevant.
  - ► Benefit: We no longer need to deny the universality of binding principles.

- But we still want answers to the following questions:

1. Is binding under c-command ever necessary in Chuj? → §4
2. Why does linear precedence matter in the absence of binding under c-command? → §5

4 C-command matters for binding, even in Chuj

Question: Is binding under c-command ever necessary in Chuj?

Answer: Yes! And in such cases linear precedence is irrelevant.

- For binding under c-command (between subjects and objects) to take place in Chuj, either:
  - (i) The object exceptionally does not raise (in which case it can be c-commanded by the subject).
  - (ii) A-movement reconstruction is exceptionally possible.
- Coon et al. (2020) argue for (ii) when a DP inside the object must be semantically bound by a non-referring subject DP, for instance a *wh*-word.
- I'll argue for (i): reflexives need to be bound (Condition A), and that for this reason they don't raise.

4.1 Reflexive objects don't raise — linear order becomes irrelevant

Reflexives across Mayan pattern like possessed nouns in appearing with Set A agreement and serving as the thematic object of transitive verbs (Aissen 2017):

(35) Ix-y-il s-**b'a** waj Xun.
PFV-A3-see A3-self CLF Xun
'Xun$_1$ saw himself$_1$.'

(36) Ix-y-il s-**tz'i'** waj Xun.
PFV-A3-see A3-dog CLF Xun
'Xun$_1$ saw his$_1$ dog.'

- But there's reason to think that (35) and (36) are structurally different.
- Recall that adverbs can normally intervene between subjects and objects in Chuj:

(37) Y-il nok' tz'i' {ewi} waj Xun {ewi}.
A3-see CLF dog yesterday CLF Xun yesterday
'Xun$_1$ saw the dog yesterday.' (see also (16) above)

- And that when a possessor is co-indexed with the subject, adverb placement options are different:

(38) Y-il s-**tz'i'** {*ewi} waj Xun {ewi}.
A3-see A3-dog yesterday CLF Xun yesterday
'Xun$_1$ saw his$_1$ dog yesterday.' (see also (17) above)

- Reflexive objects pattern differently from examples like (38) in allowing identical adverb placement options as regular transitive clauses like (37):

(39) Y-il s-**b'a** {ewi} waj Xun {ewi}.
A3-see A3-self yesterday CLF Xun yesterday
'Xun$_1$ saw himself$_1$ yesterday.'

- These data suggest that the subject is overt and possessor null in reflexive constructions (40b):

(40) a. see [$_{OBJ}$ dog [$_{POSS}$ Xun$_i$]] {yesterday} [$_{SUBJ}$ Ø$_i$] {yesterday} = (38)
b. see [$_{OBJ}$ self [$_{POSS}$ Ø]] {yesterday} [$_{SUBJ}$ Xun] {yesterday} = reflexive (39)

Proposal:

- ▶ Reflexive objects don't raise because they must be c-command by subjects (i.e. Condition A is operative).
- ▶ When the subject c-commands the object, standard binding principles prevail and *linear precedence becomes irrelevant*, as seen (40b).

- That reflexives exceptionally don't raise has been independently-suggested in the literature on the Ergative Extraction Constraint in Mayan (see e.g. Ordóñez 1995; see also Aissen 2017).[7]
- Let's see some evidence:

4.2 More evidence for the lack of object raising with reflexives

First, ergative subject extraction is exceptionally possible in all HIGH-ABS Mayan languages when the object is reflexive (see e.g. Craig 1977; Mondloch 1981; Ordóñez 1995; Aissen 1999, 2017).[8]

(41) *Reflexive* → No EEC
Mach ix-y-il s-**b'a**?
who PFV-A3-see A3-self
'Who saw themself?'

(42) *Regular transitive clause* → EEC
*Mach ix-y-il ix ix?
who PFV-A3-buy CLF woman
Int: 'Who saw the woman?'

- Recall that previous work proposes that object raising blocks subject extraction.
- The absence of the EEC in (41) can be understood if reflexive objects don't raise.

Second, in rigidly *VSO* HIGH-ABS Mayan languages, reflexives exceptionally trigger *VOS* word order (for similar facts in Popti', see Craig 1977: 217):

(43) Max y-il s-**b'a** ix ix.
PFV A3-see A3-self CLF woman
'The woman$_1$ saw herself$_1$.' (Q'anjob'al: Coon et al. 2014, (77b)).

(44) Max y-il [$_{SUBJ}$ ix ix] [$_{OBJ}$ naj winaq].
PFV A3-see CLF woman CLF man
'The woman saw the man.' (Q'anjob'al)

[7] It is worth noting that in other syntactically-ergative languages, raised absolutive objects can bind ergative subjects, in which case the ergative argument is realized as a reflexive. This is the case, for instance, in some Austronesian languages (see e.g. Richards 2000 (Tagalog), Pearson 2001 (Malagasy); Brodkin 2020). Mayan languages don't seem to allow such configurations: Reflexives and reciprocals have to be in object position, and don't appear to raise.

[8] In some HIGH-ABS Mayan languages, including Chuj and Popti', extended reflexives (sentences in which the possessor of an object and subject are co-indexed) can also optionally circumvent the EEC. Coon et al. (2020) argue that in such cases the object does raise, but that it subsequently reconstructs to get bound. According to them, reconstruction feeds subject extraction. Alternatively, we could explore the possibility that it's the possessor that A-bar extracts instead of the subject. I hope to explore this possibility in future work.

In sum: There is evidence that binding under c-command is sometimes necessary in HIGH-ABS languages, and so that binding principles (like Principle A) are sometimes operative.

- Crucially, in such cases, linear precedence no longer seems to govern the distribution of co-indexed nominal expressions.

(45) With reflexive objects, linear precedence is irrelevant

a. verb $[_{\text{OBJ}}$ self $[_{\text{POSS}}$ Ø_1]] $[_{\text{SUBJ}}$ R-expression$_1$]

b. *verb $[_{\text{OBJ}}$ self $[_{\text{POSS}}$ R-expression$_1$]] $[_{\text{SUBJ}}$ Ø_1]

Generalization: Linear precedence governs the distribution of co-indexed nominal expressions *only when binding under c-command is not possible.*

5 Why linear precedence? A constraint against cataphora

Question: Why does linear precedence govern the distribution of co-indexed nominals in Chuj *only in the absence of c-command relations*?

Hypothesis: Without c-command between co-indexed DPs, a constraint against cataphora kicks in.

5.1 Formulating the hypothesis: Distinguishing bound from free pronouns

- One way to formalize this is a PF constraint that distinguishes bound vs. *accidentally coreferring* DPs.
- Distinguishing bound from free DPs is already common practice in semantics (e.g. Reinhart 1983, Heim and Kratzer 1998, Grodzinsky and Reinhart 1993; Büring 2005).[9]
- Let's assume following Reinhart (1983) the Coreference Rule:

(46) Coreference rule (adapted from Büring 2005: 260)
If a DP can be interpreted as a bound variable, it must be.

- With (46), co-indexed DPs inside the object will necessarily receive a bound variable interpretation in LOW-ABS languages like Ch'ol and likewise whenever the object doesn't raise in HIGH-ABS languages like Chuj (i.e. with reflexive objects):

(47) With c-command, Ø is a bound variable[10] (as in Ch'ol (32) or Chuj (39))
[DP_{SUBJ} [λ1 … $[_{\text{DP(OBJ)}}$ … [Ø_1] …]]]

- But when the object raises in HIGH-ABS languages, no c-command relations will hold, and so the two co-indexed DPs will only accidentally corefer:

(48) Without c-command, Ø is a referring pronoun: (as in Chuj (34))
[$[_{\text{DP(OBJ)}}$ … [DP_1] …] … $\text{Ø}_{1/\text{SUBJ}}$]

[9] See Déchaine and Wiltschko 2002 and Bjorkman 2017 for evidence that bound/free pronouns differ morphosyntactically.

[10] One way to handle variable binding could be via movement of the possessor to subject position (e.g. Heim 1993; Hornstein 2001; Kayne 2002; Zwart 2002). See Newman 2020 for a proposal specific to Mayan languages that is compatible with movement of the possessor to subject position.

Idea: The distinction between accidentally coreferring pronouns and bound variables is syntactically encoded *and* visible to PF.[11]

(49) *PF constraint against cataphora in* HIGH-ABS *languages:*
When two DPs in a given domain *accidentally corefer*, delete the linearly second DP.

- (49) will guarantee that when the possessor of an object is co-indexed with the subject in HIGH-ABS languages, the possessor gets overtly realized (34) and the subject is null/elided:

(50) [[[$DP_{1/POSS}$] …] … ~~$DP_{1/\overline{SUBJ}}$~~]

- Two paths to "Ø":
  - Ø can be a bound variable, as in (47); in this case linear precedence is *not* relevant.
  - Ø can be the reflex of ellipsis, as in (50); in this case linear precedence *is* relevant.

5.2 Accidental coreference and cataphora across languages

- The PF rule in (49) is perhaps surprising (at least coming from English), given the reported possibility of "backwards pronominalization" with accidentally coreferring DPs (e.g. Ross 1967; Kayne 2002).

(51) a. [The woman that saw $Felipe_1$] scolded him_1.
b. [The woman that saw him_1] scolded $Felipe_1$.

- But as Kayne (2002) points out, the availability of backwards pronominalization is subject to cross-linguistic variation: many languages actually don't allow it (see e.g. Huang 1982 on Chinese).
- In French, for instance, cataphora is often not allowed:

(52) French
a. La mère de $Léa_1$ l_1'-aime.
the mother of Léa her-likes
'Lea's mother loves her.'
b. *Sa_1 mère aime $Léa_1$.
her mother likes Léa
Intended: 'Her mother loves Léa.'

(53) French
a. La femme qui a parlé à $Jeanne_1$ l_1'-aime.
the woman who has talked to Jeanne her-likes
'The woman who spoke to Jeanne loves her.'
b. ??La femme qui lui_1 a parlé aime $Jeanne_1$.
the woman who her has talked.to like Jeanne
Intended 'The woman who spoke to her loves Jeanne.'

[11] One possibility is that accidentally coreferring subject and possessor DPs in Chuj are "repetitions" which are both externally merged (in the sense of e.g. Chomsky 2013, Collins and Groat 2018, and Chomsky et al. 2019). In Ch'ol, on the other hand, the co-indexed possessor could be the copy of a DP which has re-merged in subject position (if variable binding is done through movement). If PF can distinguish repetitions from copies, then we can make sense of the generalization in (49).

5.3 No cataphora with other pronouns

- Backwards pronominalization is always judged as ungrammatical in Chuj.[12]
- Other pronominal expressions—like classifier pronouns—also appear to be subject to (49).

(54) a. [$_{CP}$ Tato tz-b'at **waj Xun**$_1$ t'a San Mateo], te-junk'o'olal ol-aj **winh**$_1$.
If IPFV-go CLF Xun to San Mateo INTS-happy PROSP-be CLF.PRON
'If Xun$_1$ goes to San Mateo, he$_1$ will be very happy.'

b. *[$_{CP}$ Tato tz-b'at **winh**$_1$ t'a San Mateo], te-junk'o'olal ol-aj **waj Xun**$_1$.
If IPFV-go CLF.PRON to San Mateo INTS-happy PROSP-be CLF Xun
Intended: 'If he$_1$ goes to San Mateo, Xun$_1$ will be very happy.'

In sum: In the absence of binding under c-command, cataphora is generally not tolerated in Chuj.

- So we need a PF constraint like (49) that bans cataphora.
- One way to do this is to say that the PF constraint only targets free DPs (and not also bound DPs).

6 Discussion: Typological and theoretical consequences

- In Mayan languages like Chuj and Popti', we find surprising patterns of nominal coreference, where on the surface, structure seems to be ignored and only linear precedence seems to matter.
- I argued that the "linear precedence effect" is conditioned by a syntactic configuration, namely in HIGH-ABS languages, objects raise above subjects (Coon et al. 2014), with pervasive effects on grammar:

1. Transitive subjects can't extract (Coon et al. 2014, Assmann et al. 2015 and Coon et al. 2020).
2. **Object raising bleeds the possibility of binding under c-command, so two accidentally coreferring DPs are generated, and PF elides the one that comes linearly second.**

- The proposal is conceptually appealing, since it allows to maintain the universality of binding principles (see e.g. Grodzinsky and Reinhart 1993; Reuland 2010; 2011), in opposition to previous accounts.
- I finally discuss two consequences of my proposal, one theoretical and one typological.

[12] Exception: Right-side topic constructions in Chuj tolerate cataphora (see e.g. Royer to appear). This is the only exception I was able to find. Interestingly, right-side topics in French also exceptionally tolerate cataphora.

6.1 Theoretical consequence: On the necessity of indices in syntax:

- Aissen (2000) noticed that the coreference facts had implications for the status of indices in grammar—PF rules like (49), repeated below, require that PF *have access to information about indices*:

  (55) *Linear precedence generalization in* HIGH-ABS *languages:*
  When two DPs in a given domain co-refer, delete the linearly second DP.

- And if PF sees indices, then they must be syntactically-represented, in violation of Chomsky's (1995, 2001) Inclusiveness condition, an assumption that has played a major role in recent theories of binding, e.g. Hornstein 2001, Safir 2004, Rooryck and vanden Wyngaerd 2011; Reuland 2001, 2011:

  (56) *Inclusiveness* (Chomsky 2001, 2-3) (cited from Collins and Groat 2018).
  [Inclusiveness] bars introduction of new elements (features) in the course of computation: **indices**, traces, syntactic categories or bar levels, and so on.

- So HIGH-ABS Mayan facts provide a challenge for Inclusiveness: PF needs to know which DPs are co-indexed in order for the linear precedence generalization in (55) to kick in.

- And they provide support for recent work arguing indices are syntactically-represented.

- See e.g. Heim 1993; Rezac 2004; Hicks 2009; Kratzer 2009; Grosz 2015; Collins and Stabler 2016; Deal 2017; Collins and Groat 2018; and Arregi and Hanink 2018 and Clem 2019 in the context of switch-reference, and Hanink 2020 and Jenks 2020 in the context of anaphoric definites.

6.2 Typological consequence: A new correlate of the LOW-/HIGH-ABS parameter

- Since the proposal is tied to the LOW-/HIGH-ABS parameter and object raising, we make the following prediction:

> **Prediction:** HIGH-ABS languages should pattern with Chuj in terms of binding and coreference, while LOW-ABS languages should pattern with Ch'ol.

- We should be able to test this across Mayan, for instance, in cases of object extraction, when the possessor and subject are co-indexed:

  (57) a. LOW-ABS languages (possessor should be null and subject overt):
  $[_{\text{OBJ}} \ldots [_{\text{POSS}} \text{Ø}_1]]$ VERB $[_{\text{SUBJ}}$ R-expression$_1$ $]$
  b. HIGH-ABS languages (possessor should be overt and subject null):
  $[_{\text{OBJ}} \ldots [_{\text{POSS}}$ R-expression$_1$ $]]$ VERB $[_{\text{SUBJ}} \text{Ø}_1]$

- Preliminary investigation support this prediction.

- Languages treated as LOW-ABS in Coon, Mateo Pedro, and Preminger 2014 that were surveyed behave like Ch'ol in terms of binding and coreference:

(58) Ch'ol

[$_{OBJ}$ I-wakax [$_{POSS}$ Ø]]$_i$ tyi i-choñ-o ___$_i$ [$_{SUBJ}$ **aj-Ana**].
A3-cow PRON PFV A3-sell-TV CLF-Ana
'Ana$_{i/*j}$ sold *her$_i$ cow*.'

(59) Tojol-ab'al

[$_{OBJ}$ Ja' ja s-wakax [$_{POSS}$ Ø]] x-chon-a [$_{SUBJ}$ ja Jwan-i'].
FOC DET A3-cow PRON A3-sell-TV DET Jwan-DET
'Jwan$_1$ sold *his*$_{1/*2}$ *cow*.'

(60) Tseltal

[$_{OBJ}$ Ja' x-wakax [$_{POSS}$ Ø]] la x-chon [$_{SUBJ}$ te j-Wan-e].
FOC A3-cow PRON PFV A3-sell DET CL-Wan-DET
'Wan$_1$ sold *his*$_{1/*2}$ *cow*.'

- Languages listed as HIGH-ABS in Coon, Mateo Pedro, and Preminger 2014 that were surveyed seem to behave like Chuj in the relevant binding patterns (Popti' is also like Chuj, as shown in Aissen 2000):

(61) Q'anjob'al

[$_{OBJ}$ A no' s-wakax [$_{POSS}$ naq Xhunik]] max s-txon-o' [$_{SUBJ}$ Ø].
FOC CLF A3-cow CLF Xhunik PFV A3-sell-IV PRON
'Xhunik$_1$ sold *his*$_1$ *cow*.'

(62) Mam (thanks to Henry and Tessa!)

[$_{OBJ}$ A t-chej [$_{POSS}$ Xwan]] o tz'-ok t-b'yo-'n [$_{SUBJ}$ Ø] .
DET A3S-horse Xwan PFV B3S-DIR A3S-hit-DS PRON
'Xwan$_1$ hit *his*$_1$ *horse*.'

(63) Kaqchikel

[$_{OBJ}$ Ja ri ru-wakx [$_{POSS}$ ri xta Ana]] x-u-k'ayi-j [$_{SUBJ}$ Ø] .
FOC DET A3S-cow DET CLF Ana PFV-A3-sell-DTV PRON
'Ana$_1$ sold *her*$_1$ *cow*.'

. . . . . .

- Btw, Mam is actually even more interesting! In regular VSO sentences, both the possessor and subject get realized (Tessa Scott, p.c.):

(64) Mam

O tz'-ok t-b'yo-'n [$_{SUBJ}$ Xwan] [$_{OBJ}$ t-chej [$_{POSS}$ *(Xwan)]].
PFV B3S-DIR A3S-hit-DS Xwan A3S-horse Xwan
Lit: 'Xwan$_1$ hit *Xwan*$_1$ *horse*.'

- I think this supports the view that in HIGH-ABS languages two accidentally co-referring DPs get generated.
- And Mam just shows us that sometimes both of these DPs can surface.

References

Aissen, Judith. 1992. Topic and focus in Mayan. *Language* 68:43–80.

Aissen, Judith. 1999. Agent focus and inverse in Tzotzil. *Language* 75:451–485.

Aissen, Judith. 2000. Prosodic conditions on anaphora and clitics in Jakaltek. In *The Syntax of Verb Initial Languages*, eds. Andrew Carnie and Eithne Guilfoyle. Oxford University Press.

Aissen, Judith. 2017. Correlates of ergativity in Mayan. In *Oxford Handbook of Ergativity*, eds. Jessica Coon, Diane Massam, and Lisa Travis. New York, NY: Oxford University Press.

Arregi, Karlos, and Emily A. Hanink. 2018. Switch reference in Washo as multiple subject agreement. In *NELS 48: Proceedings of the Forty-Eighth Annual Meeting of the North East Linguistic Society*, eds. Sherry Hicklebridge and Max Nelson, volume 1, 39–48. GLSA.

Assmann, Anke, Doreen Georgi, Fabian Heck, Gereon Müller, and Philipp Weisser. 2015. Ergatives move too early: On an instance of opacity in syntax. *Syntax* 18:343–387.

Baltin, Mark. 2010. The copy theory of movement and the bindingtheoretic status of a-traces: You can't get there from here. Ms. NYU.

Bjorkman, Bronwyn. 2017. Singular *they* and the syntactic representation of gender in English. *Glossa: A journal of general linguistics* 2:80.

Boeckx, Cedric. 2001. Scope reconstruction and A-movement. *Natural Language and Linguistic Theory* 19:503–548.

Brodkin, Dan. 2020. Antiagreement is syntactic: Evidence from south sulawesi. Paper presented at NELS 51.

Buenrostro, Cristina. 2013. La voz en Chuj de San Mateo Ixtatán. Ph.D. dissertation, El Colegio de México, Mexico City.

Buenrostro, Cristina, José Carmen Díaz, and Roberto Zavala. 1989. Sistema de clasificación nominal del Chuj. In *Memorias del Segundo Coloquio Internacional de Mayistas*, volume II. Mexico City: UNAM.

Büring, Daniel. 2005. *Binding Theory*. Cambridge: Cambridge University Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

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

Chomsky, Noam. 1995. *The Minimalist Program*. 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.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33–49.

Chomsky, Noam, Ángel Gallego, and Dennis Ott. 2019. Generative grammar and the faculty of language: Insights, questions, and challenges. *Catalan Journal of Linguistics* 229–261.

Clem, Emily. 2019. Agreement, case, and switch-reference in Amahuaca. Ph.D. dissertation, University of California, Berkeley.

Clemens, Lauren Eby, and Jessica Coon. 2018. Deriving verb initial order in Mayan. *Language* 94:237–280.

Collins, Chris, and Erich Groat. 2018. Copies and repetitions. Ms. NYU.

Collins, Chris, and Edward Stabler. 2016. A formalization of minimalist syntax. *Syntax* 19:43–78.

Coon, Jessica. 2010. VOS as Predicate Fronting in Chol. *Lingua* 120:354–378.

Coon, Jessica, Nico Baier, and Theodore Levin. 2020. Mayan agent focus and the ergative extraction constraint: Facts and fictions revisited. Ms. McGill.

Coon, Jessica, Pedro Mateo Pedro, and Omer Preminger. 2014. The role of case in A-bar extraction asymmetries: Evidence from Mayan. *Linguistic Variation* 14:179–242.

Craig, Colette Grinevald. 1977. *The Structure of Jacaltec*. Austin, TX: University of Texas Press.

Davis, Henry, Carrie Gillon, and Lisa Matthewson. 2014. How to investigate linguistic diversity: Lessons from the Pacific Northwest. *Language* 90:180–226.

Deal, Amy Rose. 2017. Towards an etiology of outer indices. In *A schrift to fest Kyle Johnson*, eds. Nicholas LaCara, Keir Moulton, and Anne-Michelle Tessier, volume 1. Linguistics Open Access Publications.

Déchaine, Rose-Marie, and Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33:409–442.

England, Nora. 1991. Changes in basic word order in Mayan languages. *International Journal of American Linguistics* 57:446–86.

García Pablo, Pascual Martín Domingo Pascual, Gaspar. 2007. *Gramática Descriptiva Chuj*. Guatemala: Academia de Lenguas Mayas de Guaatemala.

Grodzinsky, Yosef, and Tanya Reinhart. 1993. The innateness of binding and coreference. *Linguistic Inquiry* 24:69–101.

Grosz, Patrick. 2015. Movement and agreement in right-node raising constructions. *Syntax* 18:1–38.

Hanink, Emily A. 2020. DP structure and internally headed relatives in Washo. *Natural Language & Linguistic Theory* .

Heim, Irene. 1993. Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. In *The interpretive tract*, eds. Uli Sauerland and Orin Percus. MIT Working Papers in Linguistics.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell Publishers.

Henderson, Robert. 2012. Morphological alternations at the intonational phrase edge: The case of K'ichee'. *Natural Language and Linguistic Theory* 30:741–787.

Hicks, Glyn. 2009. *The derivation of anaphoric relations*. Linguistik Aktuell.

Hoekstra, Eric. 1989. A parameter for anaphor binding: The case of Jacaltec. In *Configurationality: The Typology of Asymmetries*, eds. L. Marácz and P. Muysken. Foris.

Hopkins, Nicholas. 1967. The Chuj language. Ph.D. dissertation, University of Chicago, Chicago, IL.

Hornstein, Norbert. 2001. *Move! A minimalist theory of construal*. Oxford: Blackwell Publishers.

Huang, Cheng-Teh James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT, Cambridge, MA.

Jenks, Peter. 2020. Are indices syntactically represented? Paper presented at *McGill University Colloquium Series*.

Kayne, Richard. 2002. Pronouns and their antecedents. In *Derivation and Explanation*, eds. Samuel David Epstein and T. Daniel Seeley. Oxford: Blackwell Publishers.

Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40:187–237.

Lasnik, Howard. 1999. Chains of arguments. *Current Studies in Linguistics Series* 32:189–216.

Law, Danny. 2014. *Language contact, inherited similarity and social difference: The story of linguistic interaction in the Maya lowlands*. Amsterdam: John Benjamins.

Little, Carol-Rose. 2020. Mutual dependencies of nominal and clausal syntax in ch'ol. Ph.D. dissertation, Cornell University.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70:369–415.

Maxwell, Judith. 1981. How to talk to people who talk *chekel* 'different': The chuj (mayan) solution. Ph.D. dissertation, University of Chicago, Chicago, IL.

Mondloch, James. 1981. Voice in Quiche-Maya. Ph.D. dissertation, SUNY Albany.

Newman, Elise. 2020. On the movement/(anti)agreement correlation in Romance and Mayan. Ms. MIT.

Ordóñez, Francisco. 1995. The antipassive in Jacaltec: A last resort strategy. *CatWPL* 4:329–343.

Pearson, Matthew. 2001. The clause structure of Malagasy: A minimalist approach. Ph.D. dissertation, UCLA.

Piedrasanta, Ruth. 2009. *Los Chuj, Unidad y rupturas en su espacio*. Guatemala City, Guatemala: Amrar Editores.

Reinhart, Tanya. 1976. The syntactic domain of anaphora. Ph.D. dissertation, MIT, Cambidge, MA.

Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. London: Croom Helm.

Reuland, Eric. 2001. Primitives of binding. *Linguistic Inquiry* 32:439–492.

Reuland, Eric. 2010. The universality of binding principles. In *Structure Preserved: Studies in syntax for Jan Koster*, eds. Jan-Wouter Zwart and Mark de Vries, 277–287. John Benjamins Publishing Company.

Reuland, Eric. 2011. *Anaphora and Language Design (Linguistic Inquiry Monographs)*. MIT Press.

Rezac, Milan. 2004. Elements of Cyclic Syntax: Agree and Merge. Doctoral dissertation, University of Toronto. URL http://ling.auf.net/lingBuzz/000050.

Richards, Norvin. 2000. Another look at Tagalog subjects. In *Formal Issues in Austronesian Linguistics*, eds. Ileana Paul, Vivianne Phillips, and Lisa Travis, 105–116. Dordrecht: Kluwer.

Rooryck, Johan, and Guido vanden Wyngaerd. 2011. *Dissolving Binding Theory*. Oxford University Press.

Ross, John R. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT, Cambridge, MA.

Royer, Justin. 2019. Domain restriction and noun classifiers in Chuj. In *NELS 49: Proceedings of the forty-ninth annual meeting of the North-East Linguistics Society*, eds. Maggie Baird and Jonathan Pestesky, volume 3, 87–97.

Royer, Justin. to appear. Prosody as syntactic evidence: The view from mayan. *Natural Language & Linguistic Theory* .

Safir, Ken. 2004. *The syntax of anaphora*. Oxford: Oxford University Press.

Takahashi, Shoichi. 2010. Traces or copies, or both: Part I: Characterizing movement properties. *Language and Linguistics Compass* 4:1091–1103.

Zwart, Jan-Wouter. 2002. Issues related to a Derivational Theory of Binding. In *Derivation and explanation in the Minimalist Program*, eds. Samuel David Epstein and T. Daniel Seeley, 269–304. Oxford: Blackwell Publishers.

Appendix

A Aissen 2000: An account based on precedence

Below, I summarize Aissen's (2000) prosodic account of Ø in Popti'.

- Building on data in Craig 1977 and Hoekstra 1989, Aissen (2000) argues that the distribution of Popti' Ø is governed exclusively by phonological factors:

  (65) Conditions on Ø (Aissen 2000, (20))
  The anaphor Ø must be co-indexed with a nominal which precedes it within the same intonational phrase.

- The domain of the intonational phrase is in turn determined via a syntax-prosody mapping algorithm, based on Aissen 1992 (which also seems right for Chuj, see Royer to appear):

  (66) (TOPICS)$_{\text{INTP}}$ (main CP + RC)$_{\text{INTP}}$ (CP ADJUNCTS / COMPLEMENTS)$_{\text{INTP}}$

- As Aissen shows, this analysis successfully derives the distribution of Ø in Popti'.

- Aissen's analysis makes almost all the right predictions for Chuj, e.g.:

  (67) a. Ix-y-il [$_{\text{OBJ}}$ s-mam [$_{\text{POSS}}$ waj Xun]] [$_{\text{SUBJ}}$ Ø].
  PFV-A3-see A3-father CLF Xun PRON
  'Xun$_1$ saw his$_{1/*2}$ father.'

  b. *Prosody:* (ixyil smam waj Xun$_1$ Ø$_1$)$_{\text{INTP}}$

  (68) a. Ix-y-al waj Xun [$_{\text{CP}}$ to ol-b'at winh/*Ø].
  PFV-A3-say CLF Xun COMP PROSP-go CLF.PRON
  'Xun$_1$ said that he$_{1/2}$ will go.'

  b. *Prosody:* (ixyal waj Xun$_1$)$_{\text{INTP}}$ (to olb'at winh/*Ø$_1$)$_{\text{INTP}}$

- However, it does not account for reflexive data (see (39) above).

- And suffers a conceptual issue: under this theory, binding can no longer be considered universal, despite its cross-linguistic prevalence (e.g. Reuland 2010; 2011), including in other Mayan languages (like Ch'ol).