Obligatorily overt PRO in San Martín Peras Mixtec

Jason Ostrove

Abstract

This article presents obligatory control constructions in San Martín Peras Mixtec, a language in which PRO must be exponed with an overt pronoun. I propose a morphological analysis of this phenomenon in which this language lacks a null allomorph for bound minimal pronouns (Kratzer 2009, Safir 2014, Landau 2015, 2018), posited to underlie silent PRO in other languages. This suggests that null exponence ought not be ontologically tethered to PRO's distribution or interpretation, but can rather be reduced to the routine functions of language-specific contextual allomorphy.

Keywords: control, copy control, PRO, bound variable anaphora, minimal pronouns

1 Introduction

A robust observation is that the subject of control clauses, e.g., PRO in 'Everyone$_i$ tried [PRO$_i$ to win the race],' is phonologically null in many languages. This silence is taken to be so common that deriving it is a desideratum of many theories. For example, one prominent analysis is that control clauses entirely lack a subject position (Thomason 1976, Bresnan 1978, Chierchia 1984, Wurmbrand 1998, 2001, 2004, inter alia) or require that their subject position be suppressed (Wurmbrand and Shimamura 2017). Such approaches are only plausible, of course, without an overt subject. PRO's presumed silence is also presented as circumstantial evidence for movement theories of control (e.g., Hornstein 1999:82, Boeckx and Hornstein 2006:603-604), as movement traces are also routinely null.

In contrast, other work posits that the distinctive properties of control depend solely on syntactic configurations, with null exponence epiphenomenal (Kratzer 2009, Safir 2014, Landau 2015, 2018, McFadden and Sundaresan 2018, Allotey 2021, Sulemana 2021). The appeal of these approaches is that, indeed, there is no a priori reason for PRO's silence within a minimalist framework as was first pointed out by Hornstein 1999, without perhaps some additional theoretical machinery, e.g. the Avoid Pronoun principle (Chomsky 1981) or a pragmatic requirement that silence be interpreted as anaphoric (Levinson 1987, 1991). Conceptually severing PRO from silence can do away with such conditions.

These two views generate opposite predictions. If null morphophonology is required for control, then we predict that the subject of control clauses (i.e., the bracketed clause in 'Everyone wants [PRO to win the race]') in all languages will be silent. If silence is not required for obligatory control, though, then we expect the subject of control clauses to be exponed overtly in some languages. Stated different, is it true that PRO must be null?

This article argues for the latter view: silence is not required for control. The argument for this position comes from San Martín Peras Mixtec (SMPM), an Oto-Manguean language spoken by about 12,000 people in Oaxaca, Mexico (ISO: JMX). In this language, the subject of a control clauses must be exponed as an overt clitic pronoun, with silence, i.e. PRO, strongly ungrammatical. This is shown in (1). Note that this article focuses primarily on exhaustive, obligatory control, though other types of control are presented.[1]

(1) a. Ntùkú Juân$_i$ [ka'ani *(=rà$_{i,\ *j}$) iin ntsìbá'yi].
try:COMP Juan kill:IRR =he one coyote
'Juan tried to kill a coyote.'

b. Nàkú'ún ini Maria$_i$ [ku'ún *(=ñá$_{i,\ *j}$) *tienda* koni].
remember:COMP in Maria go:IRR =she store yesterday
'Maria remembered to go to the store yesterday.'

c. Sàkwă'a Julio$_i$ [kani'i *(=rà$_{i,\ *j}$) *carro*].
learn:COMP Julio drive:IRR =he car
'Julio learned how to drive.'

Section 4 argues that the bracketed clauses in (1) demonstrate the Obligatory Control Signature of Landau 2013. In addition, Section 6 shows that the pronominal subjects of such clauses, like *=rà* 'he' in (1a), are base-generated pronouns and not overt tails of movement dependencies (Lee 2003, Boeckx et al. 2009). As such, SMPM instantiates what may be thought of as obligatorily overt PRO. Section 5 further shows that 'overt PRO' in SMPM differs from most other instances reported in the literature in that it cannot bear focus (Szabolcsi 2009, inter alia) and occurs outside of attitude reports (Grano and Lotven 2018, 2019, Li 2021), but rather patterns with supposedly typologically rare Ghanaian languages (Allotey 2021, Sulemana 2021).

Section 7 provides an analysis of overt PRO in SMPM by using the minimal pronoun framework of bound variable anaphora (Kratzer 2009, Safir 2014, Landau 2015, 2018, 2021, Sulemana 2021). In this approach, the exponence of any bound variable, including PRO, is due solely to language-specific contextual allomorphy. As such, overt PRO in SMPM is derived simply by positing that SMPM lacks a silent allomorph for bound variable anaphora of the sort presumed to derive null PRO in other languages. In this way, the interpretive and syntactic properties of PRO are divorced from its morphophonology, with silence, like all exponence in non-lexicalist frameworks of morphology (e.g., distributed morphology, Halle and Marantz 1993), derived post-syntactically without any special treatment for PRO. This section concludes by reexamining the premise that null PRO is the cross-linguistic norm, suggesting that overt PRO is more common than previously thought.

2 Syntactic background on San Martín Peras Mixtec

SMPM is predicate-initial and VSO in matrix clauses, as shown in (2a-c) for verbal, nominal, and adjectival predicates respectively.

(2) a. Kàsa'a Juân be'e.
build:COMP Juan house
'Juan will build the house.'

b. *Doctora* bà'a kúu Maria.
doctor good COP:CONT Maria
'Maria is a good doctor.'

c. Ntákŭ koo =ñá sí'i se'e Maria.
strong COP:IRR =she female child Maria
'Maria's daughter will be strong.'

All clauses in SMPM must be marked with one of three aspectual categories. In other words, SMPM lacks morphologically nonfinite clauses. Within the descriptive Mixtecan literature (e.g., Bradley and Hollenbach 1988-1992), the three morpho-aspectual categories are the 'completive (COMP),' 'continuous (CONT),' and irrealis (IRR) or 'potential.[2]'

In addition to requiring aspect marking, all clauses in SMPM require overt subjects, and all transitive clauses require overt objects: SMPM does not allow *pro*-drop, not even of expletives (3d) or impersonal pronouns (3e).[3] Several examples of this are provided in (3).

(3) a. Koni tàjan *(=yá).
yesterday quake:COMP =it
'Yesterday there was an earthquake.'

b. Kôni =ì tát =ì. David náni *(=rà).
love:CONT =I father =my David be.called:CONT =he
'I love my father. He is called David.'

c. Xìin Juân iin kwâyì. Kôni *(=rà) *(=rí).
buy:COMP Juan one horse love:CONT =he =it.AML
'Juan bought a horse. He loves it (the horse).'

d. Tásabà xíniñu'u *(=à) koo miĭ *(=é).
sometimes need:CONT =it COP:IRR alone we.INCL
Provided for 'It is important to be alone sometimes.'

e. Kwa'ă kwé'e *hora* xínchichi *(=é) ini *metro* NYC.
many very hour be.standing:CONT.SG =we.INCL in metro NYC
'One stands for a long time in New York City on the subway.' (Ostrove 2022)

SMPM is rich in pronouns, distinguishing clitic and non-clitic pronouns in local persons and six grammatical genders in non-local persons, as shown in (4) and (5). Note that pronouns in SMPM do not show case distinctions but are glossed with the appropriate English case following the descriptive tradition. There is also no number distinction for most non-local pronouns, e.g., *=rà* 'he, they (all male group).' These are glossed as singular, again following the descriptive tradition. The distinction between clitic and non-clitic pronouns will be discussed in section 5.

(4) Local persons

	SG		PL
1	yù'u/=ì	INCL	=(y)é
		EXCL	ndú'ú/=ndú
2	yô'o/=ú		ndó'ó/=ndó

(5) Non-local persons

	NEUTRAL	FEM	MASC	LIQ	WD	AML
SG	=ñà/=(y)à	=ñá	=rà	=rá	=tún	=rí
PL	=nà	=ná				

To conclude this section, let us briefly consider the main exception to predicate initiality, as it occurs throughout. In SMPM, quantified nominals optionally, though preferentially, move to a position to the left of the predicate. This quantifier fronting process is demonstrated in (6).

(6) a. Iin ntsi'bǐ$_i$ kí'i =ñá __$_i$.
one egg pick.up:CONT =she
'She's picking up an egg.'

b. Kwa'ă ntsi'bǐ kí'i =ñá __$_i$.
many egg pick.up:CONT =she
'She's picking up many egg.'

c. Ni'iin ntsi'bǐ kò kí'i =ñá __$_i$.
no egg NEG pick.up:CONT =she
'She's not picking up any eggs.'

If multiple quantified nominals are present, the fronted one preferentially, but not obligatorily, takes scope; fronted quantified nominals can reconstruct for scope. This is shown

in (7). Fronting multiple quantified nominals is ungrammatical.

(7) a. Ntsi'i =rà tsyàja$_i$ tsìi __$_i$ iin tsyaká.
every =they man catch:COMP one fish
'Every man caught a fish.' (∀>∃ preferred, ∃>∀ possible)

b. Iin tsyaká$_i$ tsìi ntsi'i =rà tsyàja __$_i$.
one fish catch:COMP every =they man
'Every man caught a fish.' (∃>∀ preferred, ∀>∃ possible)

Syntactically, quantifier fronting targets a position immediately below the landing site of wh-movement, i.e., Spec,CP (Hedding 2022). This can be seen in (8).

(8) a. Yóó kwa'ă chútun xìni?
who many cat see:COMP
'Who saw many cats?'

b. Nă ntsi'i =nà tsyàja xìni?
what every =they man see:COMP
'What did every man see?

With this cursory introduction to the language, let us begin our analysis of control in SMPM with a discussion of embedded clauses.

3 Embedded clauses in SMPM

Going back to at least the distinction in Bresnan 1972 between finite English embedded clauses selected by 'that,' nonfinite embedded clauses with a non-referent subject with the complementizer 'for' (e.g., 'I want for Mary to go') or ECM, and nonfinite embedded nonfinite clauses that require control (e.g., under 'try'), various typologies of embedded complement clauses similarly identify three core types.[4] In recent work, these three are often identified as finite embedded clauses, tensed nonfinite embedded clauses (e.g., under 'want'), and untensed nonfinite embedded clauses (e.g. under 'try'), existing a continuum of finiteness (e.g., Givón 1980). See Landau 2004, 2015, Ramchand and Svenonius 2014,

Wurmbrand 2014, 2015, Wurmbrand and Lohninger 2019, and Wurmbrand et al. 2020 and the citations therein for recent discussions. As we shall see, despite lacking morphologically nonfinite predicates, correlates of all three types occur in SMPM, and distinguishing them is necessary in the analysis of control in this language.

Let us start with what are commonly thought of as finite embedded clauses. Cross-linguistically, these are distinguished by several features. First are two related but logically separable features: morphological TAM inflection and the presence of an independent semantic tense referent (e.g., Stowell 1982:562). Recall from (2) that SMPM distinguishes three morpho-aspectual categories in matrix clauses: the completive, the continuous, and the irrealis. Embedded clauses selected under predicates like *ka'án* 'think,' *kà'àn* 'say,' or *kusijǐ ini* 'be happy' may also be marked with each, as provided in (9).

(9) a. Ká'án =ì [{ xìin, xíin, kwiin } Ana iin be'e].
think:CONT =I buy:COMP buy:CONT buy:IRR Ana one house
'I think Ana { bought, is buying, will buy } a house.'

b. Nì- kà'àn Maria [{ kìxi, kíxi, kusi } kwé'e =ñá].
COMP- say Maria sleep:COMP sleep:CONT sleep:IRR much =she
'Maria said that she { slept, sleeps, will sleep } a lot.'

c. Kùsijǐ ini =rà [{ nì- nì'ǐ, nî'ǐ, ni'ǐ } =rà ña'á].
be.happy:COMP in =he COMP- get get:CONT get:IRR =he thing
'He was happy that he { got, is getting, will get } presents.'

Similarly, embedded clauses selected by these predicates may have a semantic tense referent that differs from that of the matrix clause. This can be seen in their ability to host temporal adverbs like *koni* 'yesterday,' *bitsìn* 'now, today,' and *itsyààn* 'tomorrrow,' that differs in tense referent to that of the matrix clause. Several examples are provided in (10).

(10) a. Bitsìn kâ'àn Maria [xìin =ñá iin be'e xàa koni].
now say:CONT Maria buy:COMP =she one house new yesterday
'Maria is saying (right) now that she bought a new house yesterday.'

b. Itsyààn ka'án Raúl [xí'i kwé'e =rà bitsìn].
tomorrow think:IRR Raul drink:CONT a.lot =he now
'Tomorrow Raul will think that he is drinking too much (right) now.'

The next diagnostic for finite embedded clauses, seemingly originating in observations made by Bresnan 1972 and further elaborated in Chomsky and Lasnik 1977:456-460, is that they freely allow subjects that are non-coreferent with a matrix argument, whereas nonfinite clauses either disallow non-coreferent subjects or allow them only under certain predicates and with distinct marking. For instance, finite embedded clauses in English with non-coreferent subjects are morphologically identical to finite embedded clauses with coreferent subjects, as in (11a). In contrast, non-coreferential subjects of nonfinite embedded clauses are allowed only under certain predicates with distinct marking, such as 'want' with the complementizer 'for,' as in (11b). This strategy is not available for all predicates, though, as seen in the ungrammaticality of this strategy in (11c) under 'try.'

(11) a. John$_i$ told Bill$_j$ that he$_{i,\,j,\,k}$ was a genius.

b. John wants *(for) his daughter to be a genius.

c. *John tried (for) his daughter to be a genius.

In SMPM, embedded clauses selected by the predicates in (9-10) similarly allow subjects that are not coreferential with a matrix argument without any special marking. One example was provided in (9a), and several more are in (12).[5]

(12) a. Káchi Raúl$_i$ xí'in Pablo$_j$ [xìxi kwé'e =rà$_{i,\,j,\,k}$ ndùchì koni].
say:COMP Raul with Pablo eat:COMP many =he bean yesterday
'Raul$_i$ told Pablo$_j$ that h$_{i,\,j,\,k}$ ate a lot of beans yesterday.'

b. Kúsijǐ ini Eva$_i$ [kèba'a =ñá$_{i,\,j}$ *carrera*].
be.happy:CONT in Eva win:COMP =she race
'Eva$_i$ is happy that she$_{i,\,j}$ won the race.'

While SMPM lacks morphologically nonfinite predicates as discussed in section 2, it

has a class of embedded clauses that correspond to nonfinite clauses under predicates that select nonfinite clauses in other languages, such as *kòni* 'want,' *ntatu* 'hope,' *ntukú* 'try (lit. look for),' *xiniñu'u* 'need,' and *kixă* 'start.' In lieu of nonfinite morphology, clauses selected by these predicates must be marked irrealis.

(13) a. Ntátu =rà [{ *xòná, *xóná, koná } rà ña'á].
hope:CONT =he open:COMP open:CONT open:IRR he thing
'He hopes to open presents (lit. things).'

b. Kôni =ì [{ *xòná, *xóxá, koná } =ì mí yùye'e].
want:CONT =I open:COMP open:CONT open:IRR =I the door
'I want to open the door.'

c. Ntùkú Ana [{ *xìxi, *xíxi, kuxi } kwa'ă =ñá *ntstika*].
try:COMP Ana eat:COMP eat:CONT eat:IRR more =she banana
'Ana tried to eat more bananas.'

d. Xíniñu'u =ì [{ *xì'i, *xí'ì, ko'o } kwa'ă =ì tskwi̋ì].
need:COMP =I drink:COMP drink:CONT drink:IRR more =I water
'I need to drink more water.'

Following the discussion of similar clauses in Hebrew and several Balkan languages in Landau 2004, I call such obligatorily irrealis clauses subjunctives. Importantly for our purposes, and similarly to counterparts in Balkan languages (Terzi 1992, 1997, Dobrovie-Sorin 1994, Krapova and Petkov 1999; Krapova 2001, Landau 2004, Wurmbrand and Lohninger 2019 inter alia and et sequitur), SMPM contrasts two distinct types of subjunctive clause that correspond to tensed and untensed nonfinite embedded clauses in other languages. See especially the discussion in Landau 2004:826-833 of 'C-subjunctives,' i.e., untensed subjunctives, versus 'F-subjunctives,' i.e., tensed subjunctives.

Let us first consider semantic tense referent. In languages like English, it has been observed since at least Stowell 1982 that some nonfinite clauses, such as those selected under 'want,' may take a different tense referent than the matrix clause. See as well the discussion in Landau 2004 of "anaphoric" and "dependent" tense in obligatorily irrealis

clauses in Hebrew and Balkan languages, and Wurmbrand 2014 and the citations therein for more recent discussion and analysis. Unlike the finite clauses in (10), though, this non-identical embedded tense cannot occur before the matrix time, but rather must occur after. This is shown in (14).

(14) a. Yesterday I wanted to go to the market { today, tomorrow }.

b. ??Tomorrow I will want to have gone to the market { yesterday, today }.

This non-identical tense referent is not available for all nonfinite clauses in English, though. For instance, it is not available under 'try' or 'remember (to).' Rather, under these predicates, the embedded tense referent must be identical or 'simultaneous' (Wurmbrand 2014) to the matrix tense referent. This is provided in (15).

(15) a. I tried to go the store { yesterday, *tomorrow }.

b. I remembered to buy bananas { yesterday, *tomorrow }.

These same two classes of tensed and untensed nonfinite embedded clauses occur in SMPM, albeit with subjunctives. Examples of such clauses are provided in (16-17).

(16) Tensed subjunctive

a. Bitsìn ntátu =ì [kusi bà'a =ì itsyààn].
today hope:CONT =I sleep:IRR well =I tomorrow
'Today I hope to sleep well tomorrow.'

b. Koni kò nì- kòni Maria [ku'ún =ñá *tienda* bitsìn].
yesterday NEG COMP- want Maria go:IRR =she store today
'Yesterday Maria did not want to go to the store today.'

(17) Untensed subjunctive

a. *Bitsìn nákú'ún ini Maria [ku'ún =ñá *tienda* itsyààn].
today remember:CONT in Maria go:IRR =she market tomorrow
Intended: 'Today Maria remembers to go to the store tomorrow.'

b. *Koni ntùkú Juân [ka'ani =rà iin ntsìbá'yi itsyààn].
yesterday try:COMP Juan kill:IRR =he one coyote tomorrow
Intended: 'Yesterday Juan tried to kill a coyote tomorrow.'

Like English, tensed subjunctives in SMPM allow subjects that are not coreferent with a matrix argument, though, also like English, special marking is required. To see this, first consider (18a). Here, the subject of the embedded clause =*ñá* 'she' matches the matrix subject *Maria* in ϕ-features, and only a coreferent reading is available. However, a morpheme *ná* may be used, which forces disjoint reference, as in (18b). If the subject of the embedded clause does not match the matrix subject in ϕ-features, *ná* is optional, as in (18c-d).[6] Note as well that non-pronominal embedded subjects are grammatical in tensed subjunctives, as in (18d).

(18) a. Kôni Maria$_i$ [kusi =ñá$_{i,\ *j}$].
want:CONT Maria sleep:IRR =she
'Maria wants to sleep.'

b. Kôni Maria$_i$ [ná kusi =ñá$_{*i,\ j}$].
want:CONT Maria NÁ sleep:IRR =she
'Maria$_i$ wants her$_j$ to sleep.'

c. Kôni Maria$_i$ [(ná) kusi =rí$_j$].
want:CONT Maria NÁ sleep:IRR =it.AML
'Maria wants it (an animal) to sleep.

d. Kôni Maria$_i$ [(ná) kusi lěe =ñá].
want:CONT Maria NÁ sleep:IRR baby =her
'Maria wants her baby to sleep.'

Ná only occurs with tensed subjunctives; untensed subjunctives require a subject that is coreferent with a matrix argument, and *ná* is ungrammatical. This is shown in (19).[7]

(19) a. *Ntùkú Maria$_i$ [(ná) ku'un { =ñá$_j$, =rà$_j$ } *tienda*].
try:COMP Maria NÁ go:IRR =she =he store
Intended: 'Maria$_i$ tried for { her$_j$, him$_j$ } to go to the store.'

b. *Nàkú'ún ini Maria$_i$ [(ná) kata bà'a { =ñá$_j$, =rà$_j$ }].
remember:CONT in Maria NÁ sing:IRR well =she =he
Intended: 'Maria$_i$ remembered for { her$_j$, him$_j$ } to sing well.'

An additional difference between tensed and untensed subjunctives is related to restructuring behavior. Restructuring, or clause union, refers to so-called "monoclausality" (Grano 2012:xiv, 35), or "the lack of clause-boundedness" effects (Wurmbrand 2004:991) with respect to syntactic operations that otherwise demonstrate them. To see restructuring in SMPM, consider quantifier fronting (c.f. 6-8), which is normally clause-bounded, as seen in its inability to occur out of finite embedded clauses. This is shown in (20).

(20) No quantifier fronting out of finite embedded clauses

a. *Kwa'ă *carro$_i$* ndàtŭ'un Maria [nàkatsya Pedro __$_i$].
many car talk:COMP Maria wash:COMP Pedro
Intended: 'Maria said that Pedro washed a lot of cars.'

b. *Ni'iin =nà ká'á =rà *doctor* [kò kú'u].
no =they think:CONT =he doctor NEG be.sick:CONT
Intended: 'The doctor thinks no one is sick.'

Rather, if a quantified nominal that enters the derivation in a finite embedded clauses undergoes quantifier fronting, it must target an "intermediate" position within the embedded clause to the left of the embedded verb. This is provided (21).

(21) a. Káchi Maria [kòjmĭ xìta$_i$ xìxi =ñá __$_i$].
say:COMP Maria four tortilla eat:COMP =she
'Maria said that she ate four tortillas.'

b. Kúntàà ini =ì [ntsikû mí tsĭnà$_i$ ntá'yi __$_i$].
believe:CONT in =I all the dog cry:CONT
'I believe that all the dogs are barking.'

Tensed subjunctives show identical behavior; quantifier fronting may not occur from a tensed subjunctive clause to the matrix clause, but rather quantifier fronting must target a position to the left of the embedded verb.[8] This is provided in (22-23).

(22) No quantifier fronting from tensed subjunctives

a. *Ntsi'i *carro*$_i$ ntátu Pedro [nakatsya =rà __$_i$].
every car hope:CONT Pedro wash:IRR =he
'Pedro hopes to wash every car.'

b. *Kwa'ă yà'a í'ní$_i$ chìkàà ini Juân [kaxi =rà __$_i$].
many chili hot put.in:COMP in Juan eat:IRR =he
Intended: 'Juan thought to (lit. put inside him) eat many hot chilis.'

(23) Quantifier fronting within tensed subjunctives

a. Ntátu =rà lo'o [ntsi'i ña'á$_i$ koná =rà __$_i$].
hope:CONT =he little every thing open:IRR =he
'The boy hopes to open every present.'

b. Chìkàà ini Juân [kwa'ă yà'ă$_i$ í'ní kaxi =rà __$_i$].
put.in:COMP in Juan many chili hot eat:IRR =he
'Juan thought to eat many hot chilis.'

In contrast, quantifier fronting cannot occur within untensed subjunctives, but rather must target the matrix clause. This mandatory restructuring is shown in (24-25).[9]

(24) Quantifier fronting from untensed subjunctives

a. Kwa'ă ko'ŏ$_i$ nàntŏso Mateo [nakatsya =rà __].
many plate forget:COMP Mateo wash:IRR =he
'Mateo forgot to wash many plates.'

b. Kòjmĭ yà'ă$_i$ ntùkú Maria [kaxi =ñá __].
four chili try:COMP Maria eat:IRR =she
'Maria tried to eat four chilis.'

(25) No quantifier fronting within untensed subjunctives

a. *Nàntŏso Mateo [kwa'ă ko'ŏ$_i$ nakatsya =rà __].
forget:COMP Mateo many plate wash:IRR =he
Intended: 'Mateo forgot to wash many plates.'

b. *Ntùkú Maria [kòjmĭ yà'ă$_i$ kaxi =ñá __].
try:COMP Maria four chili eat:IRR =she
Intended: 'Maria tried to eat four chilis.'

In this way, SMPM distinguished between three types of embedded clause: finite embedded clauses, tensed subjunctives, and untensed subjunctives, the latter two of which correspond to tensed and untensed nonfinite clauses in other languages respectively. These findings are summarized in (26). Lists of predicates which have been found to select each type of embedded clause are also provided in (27).[10]

(26)

	Unrestricted TAM morphology?	Non-coreferent subject?	Restructuring?
Finite	✓	✓	**x**
Tensed subjunctive	**x**	✓	**x**
Untensed subjunctive	**x**	**x**	✓

(27) a. Predicates that select finite embedded clauses: *ka'án* 'think,' *nakanini* 'believe,' *kuntàà ini* 'wonder,' *kònì* 'know,' *kà'àn* 'say,' *ntatŭ'un* 'chat,' *káchi* 'said,' *kusijĭ ini* 'be happy,' *ntsi'i ini* 'be sad,' *ntsiko ini* 'regret'

b. Predicates that select tensed subjunctives: *kòni* 'want,' *síso ini* 'hate (lit. boil inside),' *iyĭ'bi* 'fear, be afraid,' *kuntasí* 'be scared,' *nakwatu* 'pray,' *ntatu* 'hope,' *xiinka* 'agree,' *xĭinka* 'refuse (lit. not agree),' *chikàà ini* 'get the idea to' (lit. put inside)

c. Predicates that select untensed subjunctives: *ntukú* 'try,' *nakú'ún ini* 'remember' *nantŏso* 'forget,' *kutô* 'like to,' *kixă* 'start, begin,' *ntsi'i* 'finish,' *xikwîn* 'stop,' *kò xikwîn* 'continue (lit. not stop)' *xiniñu'u* 'need,' *kònì (xá kasa)* 'know (how to),' *sakwă'a* 'learn (how to)' *kò ntaa* 'not bother' (lit. not climb)

4 Obligatory control in San Martín Peras Mixtec

As was discussed briefly in section 3, untensed subjunctives require that their subject be coreferent with a matrix argument. In this section, I show that untensed subjunctives in fact display the distinctive properties of control, specifically obligatory, exhaustive control, whereas tensed subjunctives do not. See Landau 2004 for an identical conclusion in Hebrew and several Balkan languages. Stated differently, untensed subjunctives demonstrate what Landau 2013 calls 'the OC signature,' defined in (28).[11]

(28) The OC Signature: In a control construction $[\ldots X_i \ldots [_S \text{ PRO}_i \ldots] \ldots]$, where X controls the PRO subject of the clause S:

a. The controller(s) X must be (a) co-dependent(s) of S.

b. PRO (or part of it) must be interpreted as a bound variable.

In contrast, tensed subjunctives do not show properties associated with control. Instead, the subject of a tensed subjunctive patterns as an uncontrolled pronoun, not as a bound variable. From this also comes the observation that the subject of a tensed subjunctive does not require a local, c-commanding antecedent.

4.1 Behavior as a bound variable

Since Morgan 1970, Fodor 1975, and Partee 1975 it has been known that the subject of an obligatory control clause must be interpreted as a bound variable. For instance, obligatory control clauses allow only sloppy readings under VPE, as in (29a). In this way, the subject position of an obligatory control clause patterns with other bound variable anaphora, like reflexives (29b), but not with uncontrolled pronouns, which allow strict readings (29c).

(29) a. John wants PRO to win, and Mary does too. (✓ sloppy, *strict)

b. Maria loves herself, and Julio does too. (✓ sloppy, *strict)

c. Juan$_i$ thinks that he$_i$ is a genius, and Maria does too (✓ sloppy, ✓ strict)

Similar to the uncontrolled pronoun in (29c), pronouns in embedded finite clauses allow both strict and sloppy interpretations under VPE in SMPM. This is provided in (30).

(30) a. Káchi Julia [bà'a kwé'e xìjnǐ =ñá], să =ti Mateo =ba.
say:COMP Julia good very head =she so =too Mateo =EMPH
'Julia said that she is very smart, and so does Mateo.' (✓ sloppy, ✓strict)

b. Kà'án Juan [ni'í xìnu =rà], să =ti Sergio =ba.
think:COMP Juan fast run:COMP =he so =too Sergio =EMPH
'Juan thought he ran fast, and Sergio did too.' (✓ sloppy, ✓strict)

Likewise, as with the reflexive anaphor in (29b), reflexive anaphors[12] and bound variable pronouns in SMPM both only allow sloppy readings under VPE. These are provided in (31). See as well section 6 for more on reflexives in this language. The context used to attempt to elicit a strict judgment for (31b) is also provided.

(31) a. Kôni kwé'e Raúl$_i$ mí =rà$_i$, să =ti Maria =ba.
love:CONT very Raul the =him so =too Maria =EMPH
'Raul loves himself, and Maria does too.' (✓ sloppy, *strict)

b. Ni'iin =rà bálí$_i$ kò kútô mástro =rà$_i$, să =ti =ná bálí =ba.
no =he little.PL NEG like:CONT teacher =his so =too =they.FEM little.PL =EMPH
'No boy$_i$ likes his$_i$ teachers, and no girl does either.' (✓ sloppy, *strict)

Context for strict: There's a school where boys and girls have different teachers, but everyone knows everyone else. No one likes the boys' teachers.

With this in mind, let us consider the behavior of subjunctives under VPE. Tensed subjunctives, just like the uncontrolled pronouns in (30), readily allow both strict and sloppy interpretations under VPE. This is provided in (32).[13]

(32) Tensed subjunctive VPE

a. Kôni Marco [keba'a =rà *carrera*], să =ti Maria =ba.
want:CONT Marco win:IRR =he race so =too Maria =EMPH
'Marco wants to win the race, and Maria does too.' (✓ sloppy, ✓ strict)

Context for strict: Maria is not in the race, but she wants Marco to win.

b. Ntátu Raúl [ku'ún =rà *Olímpico*], să =ti nána =rà =ba.
hope:CONT Raul go:IRR =he Olympics so =too mother =his =EMPH
'Raul hopes to go to the Olympics, and his mother does too.' (✓ sloppy, ✓ strict)

Context for strict: Raul's mother is not an athlete, but she hopes her son will go to the Olympics.

In some ways, the ability of tensed subjunctives to support strict readings under VPE is not surprising. As we have seen throughout, tensed subjunctives have overt pronominal subjects, such as *=rà* 'he' in the bracketed clauses in (32). Therefore, that (32b) is interpreted more like 'Raul hopes that he will go to the Olympics and his mother does too (✓sloppy, ✓strict)' rather than 'Raul hopes to go to the Olympics, and his mother does too (✓sloppy, *strict)' is a priori not strange.

In contrast, untensed subjunctives only support sloppy readings under VPE, as shown in (33), with the contexts that were provided to speakers to support possible strict readings.

(33) Untensed subjunctive VPE

a. Xînì Ana [ixutsya =ñá], să =ti Laura =ba.
know:CONT Ana swim:IRR =she so =too Laura =EMPH
'Ana knows (how) to swim, and Laura does too.' (✓ sloppy, *strict)

Context for strict: Laura cannot swim, but she knows that Ana can.

b. Nàkú'ún ini Maria [ku'ún =ñá *tienda*], să =ti Juân =ba.
remember:COMP in Maria go:IRR =she store so =too Juan =EMPH
'Maria remembered to go to the store, and Juan did too.' (✓ sloppy, *strict)

Context for strict: Juan did not go to the store, but he remembers Maria did.

In contrast to the tensed subjunctives in (32), untensed subjunctives are not interpreted in the expected way; (33b) does not and cannot mean 'Maria remembered that she went to the store,' as a word-for-word translation might indicate. Rather, in only allowing sloppy readings under VPE, subjects of untensed subjunctives behave like bound variable anaphora in this language, such as in (31). Furthermore, as was provided in (33), this bound variable interpretation is obligatory. In this way, we begin to see that untensed subjunctives pattern with obligatory control, whereas tensed subjunctives do not.

Another feature of bound variable anaphora is exhaustive binding (Postal 1974, Chomsky 1981, Lasnik 1989, Charnavel and Sportiche 2016, inter alia and et sequitur). For instance, a reflexive anaphor cannot nonexhaustively overlap in reference with its binder. This is provided in (34). (34a) demonstrates a reflexive anaphor that is a subset of its binder, while in (34b) the reflexive anaphor is only partially bound by its binder.

(34) (Adapted from den Dikken et al. 2001:137)

a. *We voted for myself.

b. *I voted for ourselves.

These patterns hold for bound variable anaphora in SMPM as well, as shown in (35) for reflexive anaphors.

(35) a. *Xìni [Ana$_i$ ra Pablo$_j$] mí =rà$_j$.
see:COMP Ana and Pablo the =him
Intended: 'Ana and Pablo$_i$ saw himself$_i$.'

b. *Xìni Pablo$_i$ mí =nà$_{i+j}$.
see:COMP Pablo the =them
Intended: 'Pablo$_i$ saw themselves$_{i+j}$'

Within the control literature, it has been found that not all instances of PRO behave the same with respect to exhaustive binding. As was first shown by Landau 2000, 2004, tensed nonfinite embedded clauses allow partial control, similar to the ungrammatical (34b) in

which whereas untensed nonfinite embedded clauses do not.[14] This is provided in (36) for English through the behavior of inherently plural predicates like 'meet' with a singular matrix subject and adverbs like 'together,' which require a plural antecedent.

(36) a. The boss wants to meet at 11 o'clock tomorrow morning.

b. John hopes to work on the project together tomorrow.

c. *The boss starts to meet at 11 o'clock every day.

d. *John is trying to work on the project together.

Like the untensed C-subjunctives in Balkan languages reported by Landau 2004, untensed subjunctives in SMPM disallow partial control. This is provided in (37).[15]

(37) a. *Nàntǒso =rà *patrón* [nakitá'àn =nà kâ uxi iin].
forget:COMP =he boss meet:IRR =they hour ten one
Intended: 'The boss forgot to meet at 11 o'clock.' (Untensed subjunctive)

b. *Xíniñu'u Maria [ku'ún ntíbi =nà bìjkǒ].
need:CONT Maria go:IRR together =they party
Intended: 'Maria needs to go to the party together.' (Untensed subjunctive)

In contrast, tensed subjunctives in SMPM do allow nonexhaustive binding, although, as tensed subjunctives do not seem to involve control at all, this would not be a true instance of 'partial control.' See the discussion in Sevdali and Sheehan 2021 of the confound of identifying partial control in languages of this type.[16] Examples of this nonexhaustive binding are provided in (**??**).

In this way, subject pronouns of untensed subjunctives consistently pattern with and are interpreted as bound variables in this language, satisfying the second condition of the OC signature from (28b). In contrast, the subjects of tensed subjunctives do not.

4.2 Untensed subjunctives require a local, c-commanding antecedent

As is well established (see Landau 2013 for an overview), the binder of complement obligatory control PRO must be a local c-commander, as in (38a). The same holds for other bound variable anaphora, as shown in (38b) for reflexives.[17]

(38) a. Julio$_j$'s mother$_i$ wanted [PRO$_{i,*\,j,\,*k}$ to see the puppy].

b. Every boy$_j$'s father$_i$ loves himself$_{i,\,*j,\,*k}$.

Beginning with local binding, as was briefly discussed in Section 2, the subjects of some subjunctives in SMPM need not be bound. Importantly, as was discussed above, uncontrolled subjects are available only for tensed subjunctives, as shown in (18b-c), repeated in (39). In contrast, the subjects of untensed subjunctives must be bound, as shown in (40).

(39) Non-locally bound subjects in tensed subjunctive

a. Kôni Maria$_i$ [ná kusi =ñá$_{*i,\,j}$].
want:CONT Maria NÁ sleep:IRR =she
'Maria$_i$ wants her$_j$ to sleep.'

b. Kôni Maria$_i$ [(ná) kusi =rí$_j$].
want:CONT Maria NÁ sleep:IRR =it.AML
'Maria wants it (an animal) to sleep.

(40) Locally bound subjects only in untensed subjunctive

a. *Xíniñu'u Maria$_i$ [(ná) kwiin { =ñá$_j$, =rà$_j$ } iin koto xàá].
need:CONT Maria NÁ buy:IRR =she =he one shirt new
Intended: 'Maria$_i$ needs { her$_j$, him$_j$ } to buy a new shirt.'

b. *Kìxă Maria$_i$ [(ná) kata bà'a { =ñá$_j$, =rà$_j$ }].
start:CONT Maria NÁ sing:IRR well =she =he
Intended: 'Maria$_i$ started for { her$_j$, him$_j$ } to sing well.'

Tensed subjunctives also allow non-coreferent non-pronominal subjects, as in (41a), while untensed subjunctives do not allow this, as in (41b).

(41) a. Ntátu Maria [bà'a ná kusi ìjǐ̌ =ñá].
hope:CONT Maria well NÁ sleep:IRR husband =her
'Maria hopes for her husband to sleep well.' (Tensed subjunctive)

b. *Xíniñu'u Maria [kwiin Juân iin koto xàá].
need:CONT Maria buy:IRR Juan one shirt new
Intended: 'Maria needs for Juan to buy a new shirt.' (Untensed subjunctive)

Likewise, subjects of tensed subjunctives do not require a c-commanding antecedent, whereas subjects of untensed subjunctives do. This can be seen in considering the behavior of possessors. In SMPM, possessors do not seem to c-command out of their containing DP, as evidenced by their inability to license a reflexive in object position, just like in (38b). This is shown in (42). See Section 6 for further discussion of reflexives in SMPM.

(42) Xìni ísi'i$_i$ Marcario$_j$ { mí =ñá$_i$, *mí =rà$_j$ } ini yùtátá.
see:COMP wife Marcario the =her the =him in mirror
'Marcario$_j$'s wife$_i$ saw { herself$_i$, *himself$_j$ } in the mirror.'

Despite this lack of c-command, a matrix possessor can antecede the subject of a tensed subjunctive, but not the subject of an untensed subjunctive, as shown in (43) and (44).

(43) Non-c-commanding antecedent: tensed subjunctive

a. Kôni nána Julio$_i$ [ná keba'a =rà$_i$ mí *carrera*].
want:CONT mother Julio NÁ win:IRR =he the race
'Julio$_i$'s mother wants him$_i$ to win the race.'

b. Ntátu *amiga* ña'à Marco$_i$ [ná ku'u =rà$_i$ mí *Olimpico*].
hope:CONT friend.FEM POSS Marco NÁ go:IRR =he the Olympics
'Marco$_i$'s friend hopes for him$_i$ to go to the Olympics.'

(44) No non-c-commanding antecedent: untensed subjunctive

a. *Nàntǒso táta Maria$_i$ [ku'un =ñá$_i$ *tienda*].
forget:COMP father Maria go:IRR =she store
Intended: 'Maria$_i$'s father forgot for her$_i$ to go to the store.'

b. *Xíniñu'u nána Julio$_i$ [kwiin =rà$_i$ iin koto xàá].
need:CONT mother Julio buy:IRR =he one shirt new

Intended: 'Julio$_i$'s mother needs him$_i$ to buy a new shirt.'

More extreme examples of non-c-commanding antecedents for subjects of tensed subjunctives also occur, such as embedding a possessor antecedent in an adjunct in (45). Similar data are not replicable for untensed subjunctives, though, as in (46).

(45) Tensed subjunctive

a. Bìjkǒ ña'à Marco$_i$ xìndatu Ana [kaxi =rà$_i$ yù'u =ñá].
party POSS Marco hope:IMP:COMP Ana eat:IRR =he mouth =her
'At Marco$_i$'s party, Ana hoped for him$_i$ to kiss her.'

b. Bìjkǒ ña'à Marco$_i$ kòni Ana [kaxi =rà$_i$ yù'u =ñá].
party POSS Marco want:COMP Ana eat:IRR =he mouth =her
'At Marco$_i$'s party, Ana wanted for him$_i$ to kiss her.'

(46) Untensed subjunctive

a. *Bìjkǒ ña'à Marco$_i$ xìniñu'u Ana [kaxi =rà$_i$ yù'u =ñá].
party POSS Marco need:COMP Ana eat:IRR =he mouth =her
Intended: 'At Marco$_i$'s party, Ana needed for him$_i$ to kiss her.'

b. *Bìjkǒ ña'à Marco$_i$ ntùkú Ana [kaxi =rà$_i$ yù'u =ñá].
party POSS Marco try:COMP Ana eat:IRR =he mouth =her
Intended: 'At Marco$_i$'s party, Ana tried for him$_i$ to kiss her.'

This behavior aligns subjects of untensed subjunctives with other bound variable anaphora in SMPM. Like the reflexive anaphor in (42), bound variable pronouns also seem to require c-command in this language. This is shown in (48). Note that only *ni'iin* 'no' is presented to avoid confounds with e-type readings (Evans 1980:340).

(47) No bound variable interpretations without c-command

a. Kò kôni nána ni'iin =nà báli$_i$ =nà$_{*i,\,j}$.
NEG love:CONT mother no =they little.PL =them
Intended: 'No child$_i$'s mother loves him$_i$.'

b. Kò kôni se'e =ná$_i$ ni'iin =ná sí'i$_{*i,\,j}$.
NEG love:CONT child =their.FEM no =they.FEM female

Intended: 'Their$_i$ children love no women$_i$.'

c. Bìjkǒ ña'à ni'iin =rà tsyàja$_i$ nǐ- xìni Sofia =rà$_{*i,\ j}$.
party POSS no =he man COMP.NEG- see Sofia =them.FEM
Intended: 'At no man$_i$'s party Sofia saw him$_i$.'

(48) Bound variable interpretations with c-command

a. Kò táxá'á ni'iin =nà bálí$_i$ xí'in nána =nà$_i$.
NEG dance:CONT no =they little.PL with mother =their
'No child$_i$ is dancing with their$_i$ mother.'

b. Nǐ- taxi ni'iin =nà bálí$_i$ skwéla ijta nda'ă mástro =nà$_i$.
COMP.NEG- give no =they little.PL school flower to teacher =their
'No student$_i$ gave a flower to their$_i$ teacher.'

c. Kò ká'án ni'iin =rà tsyàja$_i$ kěba'a =rà$_i$ mí *carrera*.
NEG think:CONT no =he man win:IRR.NEG =he the race
'No man$_i$ thinks that he$_i$ will not win the race.'

Like subjects of tensed subjunctives in (43), though, pronominal coreference without quantificational variable binding does not require c-command, as shown in (49).

(49) a. Kò nì- i'ín chútun sâna Julio$_i$ =rà.$_i$
NEG COMP- scratch cat POSS.AML Julio =him
'Julio$_i$'s cat did not scratch him$_i$.'

b. Kôni kwé'e nána Ana$_i$ =ñá$_i$.
love:CONT a.lot mother Ana =her
'Ana$_i$'s mother loves her$_i$ a lot.'

c. Bìjkǒ ña'à Marco$_i$ xìni Ana =rà$_i$.
party POSS Marco see:COMP Ana =him
'At Marco$_i$'s party Ana saw him$_i$.'

In this way, the subjects of untensed subjunctives fit the profile of bound variable anaphora in this language, whereas the subjects of tense subjunctives pattern as unbound, possibly coreferential, pronouns.

4.3 Intermediate summary

In this section, we saw that untensed subjunctives demonstrate all of the core features of control: their subjects are interpreted as bound variables, which, in this language, require a c-commanding binder. Crucially, these interpretive properties must occur with an overt pronoun, rather than with null PRO. This was first shown in (1), repeated as (50), and which we can now recognize as involving untensed subjunctives.

(50) a. Ntùkú Juân$_i$ [ka'ani *(=rà$_{i,\ *j}$) iin ntsìbá'yi].
try:COMP Juan kill:IRR =he one coyote
'Juan tried to kill a coyote.'

b. Nàkú'ún ini Maria$_i$ [ku'ún *(=ñá$_{i,\ *j}$) *tienda* koni].
remember:COMP in Maria go:IRR =she store yesterday
'Maria remembered to go to the store yesterday.'

c. Sàkwă'a Julio$_i$ [kani'i *(=rà$_{i,\ *j}$) *carro*].
learn:COMP Julio drive:IRR =he car
'Julio learned how to drive.'

In this way, SMPM can be thought of as a language in which obligatorily control PRO is always overt. As an anonymous reviewer observes, this conclusion may not be surprising in some respects given the typological profile of SMPM. As discussed above, SMPM is a non-*pro*-drop language, and all obligatory control clauses are finite.[18] Therefore, the appearance of overt pronominal subjects in obligatory control clauses is prima facie expected.

Caution is merited here, though, as it is not the case that PRO is always overt in non-*pro*-drop languages with finite control. For instance, Landau 2004 reports that *pro*-drop is ungrammatical in Hebrew with third-person subjects in all tenses, as shown in (51).

(51) *(hu) { diber, yedaber, medaber }.
he talked.3SG will.talk.3SG talks.3SG
'He { talked, will talk, talks }.' (Landau 2004:815)

Despite this lack of *pro*-drop, Hebrew requires null subjects, i.e., PRO, in finite object

control constructions such as (52).

(52) Hem kivu še- { PRO, # hem } yelxu ha-bayta mukdam.
they hoped that- they will.go.3PL home early
Intended: 'They hoped to go home early.' (Landau 2004:816)

Overt *pro* good without OC, e.g. 'They$_i$ hoped that they$_{i,\,j}$ will go home early.'

Likewise, it is not the case that if a language has overt PRO then it will have finite control. While we discuss this language further in section 5, Sulemana 2021 reports that Bùlì has obligatory overt PRO in non-finite clauses, briefly shown in (53).

(53) Asouk$_i$ tìerì wà { dā gbáŋ, # dà gbăŋ }.
Asouk remember 3SG buy.NONFIN book buy.PAST book
'Asouk remembered to buy a book.' (Adapted from Sulemana 2021:75)

Finite verb good without OC, e.g. 'Asouk$_i$ remembered that he$_{i,\,j}$ bought a book.'

As such, we cannot straightforwardly conflate overt PRO with a combination of finite control and non-*pro*-drop. That said, there may be a weaker correlation between overt PRO and non-*pro*-drop. The two best studied languages with obligatorily overt PRO in complement clause obligatory control, Gã (Allotey 2021) and Bùlì (Sulemana 2021), each discussed further in section 5, are reported to be non-*pro*-drop languages like SMPM. This means that we can posit the tentative one-way implicational universal in (54), although more typological work is, of course, necessary.[19]

(54) If a language requires the subject of obligatory control clauses (i.e., PRO) to be overt, then that language will not allow *pro*-drop.

5 Overt PRO cross-linguistically

As mentioned briefly above, SMPM is not the first language to be described as having overt PRO. The phenomenon of having an overt subject in a control clause is referred to by Polinsky and Potsdam 2006 as copy control, defined in (55).

(55) Copy control: The phenomenon in which the subject of a control clause is a phonologically overt copy of its controller.

Setting aside adjunct control (see Haddad 2009, 2011), copy control is rarely reported in the literature, and indeed believed to be rare (Potsdam and Haddad 2017:9). The attested copy control systems fall into two broad categories. I term the first full copy control, as the subject of a control clause is exponed as a full copy of its controller, as in (56).

(56) Full copy control

a. B- yennlààa'z bxuuhahz$_i$ [ny- ahcnèe bxuuhahz$_i$ Gye'eihlly].
PERF- forget priest SUBJ- help priest Mike
'The priest forgot to help Mike.' (San Lucas Quievaní Zapotec, Lee 2003).

b. Me rá Juán$_i$ [ca- 'na' Juán$_i$].
want Juan POT- come Juan
'Juan wants to come.' (Copala Triqui, Broadwell 2019)

Full copy control is not available in SMPM, though, as shown in (57). Instead, a pronoun must be used, as we have seen throughout.

(57) a. Nàntǒso Juana$_i$ [nakatsya { =ñá$_i$, *Juana$_i$ } mí tsyàà].
forget:COMP Juana wash:IRR =she Juana the clothes
'Juana forgot to wash the clothes.'

b. Kìxă mí leso$_i$ [taxá'á { =rí$_i$, *mí leso$_i$ }].
start:COMP the rabbit dance:IRR =it.AML the rabbit
'The rabbit started to dance.'

Patterns superficially like SMPM in which the subject position of the controlled clause

is exponed as a pronoun matching its controller in ϕ-features are also attested. I call this broad patten pronominal copy control. In the literature, three types of pronomial copy control are discussed. The first, which I term logophoric pronominal copy control, seems to be found only in attitude reports. Several examples are provided in (58).

(58) Logophoric copy control

a. Ámã$_i$ dʒí [bé jè$_{i,\ *j}$ lá ɖù nũ].
Ama want C LOG POT eat thing
'Ama wants to eat.' (Gengbe, Grano and Lotven 2018)

b. Lisi dasuan$_i$ [xia ke yihou ta$_i$ yao qu kan-kan].
Lisi plan end class after he will go look-look
'Lisi plans to go take a look after class.' (Mandarin, Li 2021)

While theoretically significant in similar ways, logophoric pronominal copy control differs from copy control in SMPM in several respects. First, as we saw in Section 4, pronominal copy control in SMPM does not only occur in attitude reports; rather, attitude predicates like *ntatu* 'hope' or *kòni* 'want' select tensed subjunctives, which were shown to not display the OC signature. Second, as Li 2021 shows, logophoric pronominal copy control in Mandarin does not seem to involve PRO, distinguishing it from pronominal copy control in SMPM. See Section 4.

The second sort of pronominal copy control seems to have first been discussed by Szabolcsi 2009, who observed that if PRO is modified by a scope-taking operator, particularly focus, then it often must be realized as an overt pronoun. See as well Livitz 2014, Herbeck 2015, Landau 2015, Barbosa 2018, and Ganenkov To appear. I call this scope-sensitive pronominal copy control, examples of which are provided in (59).[20]

(59) Scope-sensitive pronominal copy control

a. Ogni ragazzo$_i$ odierebbe [andare solo lui$_i$ a Milano].
every boy would.hate.3SG go.INF only he to Milan
'Every boy would hate it if only he went to Milan.' (Italian, Szabolcsi 2009)

b. *pro*$_i$ Utálna [mindig csak ő$_i$ kapni büntetést].
hate.COND.3SG always only he get.INF punishment.ACC
'He would hate if always only he got punished.' (Hungarian, Livitz 2014)

c. Só o João detestou [resolver ele$_i$ o problema].
only the João hated. solve.INF PRO the problem
'Only João hated it that only him solved the problem.' (European Portuguese, Barbosa 2009 via Livitz 2014)

Pronominal copy control in SMPM differs from scope-sensitive copy control in two ways. First, the languages in (59) all require null PRO without focus or a scope-taking operator, as shown in (60). As was shown in (50), this is not available in SMPM.

(60) a. Ogni ragazzo$_i$ odierebbe [andare { PRO$_i$, *lui$_i$ } a Milano].
every boy would.hate.3SG go.INF PROR he to Milan
'Every boy would hate to go to Milan.' (Italian, Szabolcsi 2009)

b. *pro*$_i$ Utálna [mindig { PRO$_i$, *ő$_i$ } kapni büntetést].
would-hate.3SG always PRO he get.INF punishment.ACC
'He would hate to always get punished.' (Hungarian, Livitz 2014)

c. Só o João$_i$ detestou [resolver { PRO, # ele$_i$ } o problema].
only the João hated. solve.INF PRO he the problem
'Only João hated solving the problem.' (European Portuguese, Barbosa 2009 via Livitz 2014)

Second, without an overt operator, the pronominal subjects of untensed subjunctives in SMPM would presumably need to be interpreted as bearing focus, comparable to the European Portuguese in (59c). This is not a possibility for this language, though. To see this, recall from (4) that SMPM distinguished between clitic and non-clitic pronouns. The local person clitic/non-clitic pairs are provided (61) in (62).

(61) Local person clitic pronouns

	SG		PL
1SG	=ì	Incl.	=(y)é
		Excl.	=ndú
2SG	=ú		=ndó

(62) Local person nonclitic pronouns

	SG		PL
1SG	yù'u	Incl.	-
		Excl.	ndú'ú
2SG	yô'o		ndó'ó

If a distinct non-clitic form is not present, as with the first plural inclusive and non-local persons, several 'strengthening' strategies exist to make non-clitic pronouns. These include modifying a clitic pronoun with a demonstrative, e.g. *yé yo'o* 'we (INCL) here,' or adding the definite article *mí*, also used to form reflexive anaphors (see section 6), yielding forms like *mí =rà* 'himself.' See McCloskey and Hale (1984) for similar phenomena in Irish.

Clitic pronouns in SMPM pattern in the usual way (Cardinaletti and Starke 1999): clitic pronouns cannot be coordinated (63a), cannot occur on their own (63b), and may have impersonal readings (63c). These properties do not hold for non-clitic pronouns, as in (64).

(63) a. **Cita* ku'ún =ñá ra =rà itsyààn.
date go:IRR =she and =he tomorrow
Intended: 'Her and him are going on a date tomorrow.'

b. *Yô =nà xàxi *galleta*? =Ñá!
who =they eat:COMP cookie =she
Intended: 'Who ate the cookies? Her!'

c. Ñûù Ká'nu chí'i =nà ndòjo.
town big sow:CONT =they sugar.cane
'They (in general) grow sugar cane in San Martín Peras (lit. the big town).'

(64) a. *Cita* ku'ún mí =ñá ra mí =rà itsyààn.
date go:IRR the =she and the =he tomorrow
'Her and him are going on a date tomorrow.'

b. Yô =nà xàxi *galleta*? Ñá kan!
who =they eat:COMP cookie she there
'Who ate the cookies? Her!'

c. #Ñûù Ká'nu chí'i nà kan ndòjo.
town big sow:CONT they there sugar.cane
Intended: 'They (in general) grow sugar cane in San Martín Peras.'

Good as 'They (a specific group) grow sugar cane in San Martín Peras.'

Importantly for our purposes, clitic pronouns cannot bear focus. In such contexts, as with *íntàà* 'only' in (65), non-clitic pronouns must be used.[21]

(65) a. Íntàà { yù'u, * =ì } kachûn itsyààn.
only I =I work:IRR tomorrow
'Only I will work tomorrow.'

b. Íntàà { =ra kan, * =rà } kachûn itsyààn.
only =he there =he work:IRR tomorrow
'Only he will work tomorrow.'

Overt operators are not required for focus interpretations of pronouns, though, as shown in contrasting the non-focused interpretation of the clitic pronoun in (66) with the focus interpretation of the corresponding non-clitic.

(66) Á chíntsyĕ Javier { =ndú, ndú'ú } bitsìn?
Q help:CONT Javier =you.PL you.PL today
'Will Javier help { you all, YOU ALL } today?'

Speaker comment: *Ndú'ú* invokes alternatives, "it's like we're contrasting you with other people" i.e., it bears focus. *=Ndú* does not.

Crucially, the pronominal subjects of untensed subjunctives must be clitics, with any non-clitic forms judged as sharply ungrammatical. This is shown in (67).

(67) a. *Saá nántŏso =ndó$_i$ [koná ndó'ó$_i$ yùye'e].
always forget:CONT =you.PL open:IRR you door
Intended: 'You always forget for YOU to open the door (not other people).'

b. *Xíni Marco$_i$ [ixutsya mí =rà$_i$].
know:CONT Marco swim:IRR the =he
Intended: 'Marco knows how to swim.'

c. *Kìxă mí leso$_i$ [taxá'á rí kan$_i$].
start:COMP the rabbit dance:IRR it.AML there
Intended: 'The rabbit started to dance.'

In contrast, pronominal subjects of tensed subjunctives, argued above to not involve control, do not need to be clitics, as shown in (68).

(68) a. Kôni Jûan$_i$ [keba'a mí =rà$_i$ *carrera*].
want:CONT Juan win:IRR the =he race
'Juan wants to win the race.'

b. Ntátu Maria$_i$ [ku'ún mí =ñá$_i$ yòjo].
hope:CONT Maria go:IRR the =she moon
'Maria hopes to go to the moon.'

As such, it seems to be the case that subjunctives interpreted as controlled clauses, e.g., untensed subjunctives, require clitic pronominal subjects. This, coupled with the inability of clitic pronouns to occur with focus, indicates that pronominal copy control in SMPM cannot be analyzed as scope-sensitive pronominal copy control on par with the languages in (59).

Rather, SMPM seems to instantiate an apparently rare type of pronominal copy control that I term obligatory pronominal copy control. I am aware of only two other languages argued to be of this type: Gã (Allotey 2021) and Bùlì (Sulemana 2021).[22]

(69) Obligatory pronominal copy control

a. Gbekebii$_i$ lɛ nye ni { ame$_i$, *PRO } he shia.
children DET managed C 3PL buy home
'The children managed to buy a home.' (Gã, Allotey 2021)

b. Asouk$_i$ tìerì { wà, *PRO } dā gbáŋ.
Asouk remember 3SG buy book
'Asouk remembered to buy a book.' (Bùlì, Sulemana 2021)

Allotey 2021 notes that obligatory pronominal copy control may be an areal feature of Ghana and surrounding countries, describing it in Igbo, Akan, and Ewe[23] as well as Gã. To

my knowledge, SMPM is the first language outside this area to be argued to have obligatory pronominal copy control.

6 Against a movement analysis

Having established that untensed subjunctives in SMPM involve obligatory control, let us start to derive 'overt PRO' in SMPM, beginning with its syntactic derivation.

The literature provides two broad derivations of control. The first, termed the base-generated analysis, shown in (70a), involves two distinct positions, one in the embedded clause and one in the matrix clause, that are each occupied by distinct elements, e.g. a matrix controller and PRO. Under this analysis, the two positions are linked via a non-movement relation such as variable binding. The second, shown in (70b), has the matrix controller enter the derivation in the subject position of the embedded clause, and then move to the matrix clause. As such, this analysis involves two copies of just one DP, rather than two distinct DPs. See Landau 2013 and Hornstein and Polinsky 2010 for overviews and arguments for each.

(70) a. $^{\text{matrix}}$[Controller$_i$ … $^{\text{control}}$[Subject$_i$ …]] (Base generation)

b. $^{\text{matrix}}$[Controller … $^{\text{control}}$[~~Subject~~ …]] (Movement)

While movement theories of control were built to handle languages in which the controlled subject is null (Hornstein 1999), such analyses have been extended to copy control under the copy theory of movement (Chomsky 1995). See Lee 2003 and Boeckx et al. 2009. Likewise, a movement analysis of pronominal copy control in SMPM could procede along similar lines, with the pronominal subject of the untensed subjunctive a kind of reduced copy (e.g., Kandybowicz 2006 or Harizanov 2014).

My goal in this discussion is not to argue for or against the theoretical merits of either approach. Rather, I modestly argue that a movement analysis is not appropriate for SMPM. The basis of this argument comes from the distribution of exempt anaphors. See Barbosa 2009:109-110 for a similarly structured argument. To begin, consider the behavior of locally bound reflexive anaphors. As shown in (71) and presented briefly above in (42) and (65), reflexive anaphora are composed of the definite article *mí* with a clitic pronoun.

(71) a. Xìni Juân$_i$ mí =rà$_{i,\ *j}$ ini yùtátá.
see:COMP Juan the =him in mirror
'Juan saw himself in the mirror.'

b. Saá kâ'àn Maria$_i$ xa'ă mí =ñá$_{i,\ *j}$.
always talk:CONT Maria about the =her
'Maria always talks about herself.'

A reflexive interpretation is unavailable without *mí*. Rather, only a non-coreferent interpretation is available. This can be seen in contrasting (71) with (72).

(72) a. Xìni Juân$_i$ =rà$_{*i,\ j}$ ini yùtátá.
see:COMP Juan =him in mirror
'Juan$_i$ saw him$_j$ in the mirror.'

b. Saá ká'àn Maria$_i$ xa'ă =ñá$_{*i,\ j}$.
always talk:CONT Maria about =her
'Maria$_i$ always talks about her$_j$.'

As will be relevant to our discussion below, quantified nominals readily bind locally bound reflexive anaphors as well, similar to (71). This is shown in (73).

(73) a. Tá'iin'iin =nà bálí$_i$ xìni mí =nà$_i$ ini yùtátá.
each =they little.PL see:COMP the =them in mirror
'Every child saw themselves in the mirror.'

b. Ni'iin =nà$_i$ nǐ- kà'àn xa'ă mí =nà$_i$ bìjkŏ.
no =they COMP:NEG- talk about the =them party
'No one talked about themselves at the party.'

Pronouns with *mí* are required for reflexive interpretations with direct objects and objects of prepositions, as in (71), but not with possessors (e.g., 'Mary$_i$ saw her$_i$ father'). If present, though, speakers translate these items into Spanish as *su propio* 'its own,' as in (74). I will refer to these optional uses of reflexive anaphors outside of their usual binding domain as exempt anaphors (Pollard and Sag 1992).

(74) a. Tsìi tsǐnà$_i$ ndò'ò { =rí$_i$, mí =rí$_i$ }.
bite:COMP dog tail =its the =its
'The dog bit { its, its own } tail.'

b. Xìni Maria$_i$ táta { =ñá$_i$, mí =ñá$_i$ }.
see:COMP Maria father =her the =her
'Maria saw { her, her own } father.'

Unlike locally bound reflexive anaphors (c.f. 73), exempt anaphors cannot be anteceded by a quantified nominal. This is provided in (75).

(75) a. *Tá'iin'iin tsǐnà$_i$ tsìi ndò'ò mí =rí$_i$.
each dog bite:COMP tail the =its
Intended: 'Each dog bit its own tail.'

b. *Ni'iin =ná báli$_i$ nǐ- xini táta mí =ná$_i$.
no =they.FEM little.PL COMP.NEG see father the =their.FEM
Intended: 'No girl saw her own father.'

Exempt anaphors do not only occur as possessors, and this restriction against quantified antecedents extends to other positions. This can be seen in contrasting (76) with (77).

(76) a. Káchi =nà xí'in Julio$_i$ [bà'a sásiki mí =rà$_i$ *pelota*].
say:COMP =they with Julio well play:CONT the =he soccer
'They told Julio$_i$ that himself$_i$ plays soccer very well.'

b. Káchi Juân$_i$ [kà'àn nuù Maria xa'ă mí =rà$_i$].
say:COMP Juan say:COMP insult Maria about the =him
'Juan$_i$ said that Maria insulted himself$_i$.'

(77) a. Ni'iin =nà bálí$_i$ nǐ- ka'án [santañù'u (*mí) =nà$_i$].
no =they little.PL COMP:NEG- think lose:IRR the =they

'No child$_i$ thought that himself$_i$ would lose.'

b. Ntàtŭ'un 'iin'iin =nà bálí$_i$ [kàni Maria (*mí) =nà$_i$].
talk:COMP each =they little.PL hit:COMP Maria the =them
'Each child$_i$ said that Maria hit himself$_i$.'

As such, it seems that SMPM obeys a restriction like (78).

(78) The antecedent of an exempt anaphor cannot be quantified.

Setting aside the origin of (78) or how it is enforced in the grammar[24], its repair is to use left dislocation with a clitic pronoun in argument position, as in (79).

(79) a. Tá'iin'iin tsǐnà$_i$ tsìi =rí$_i$ ndò'ò mí =rí$_i$.
each dog bite:COMP =it.AML tail the =itsAML
'Each dog, they bit their own tail.'

b. Ni'iin =ná báli$_i$ nǐ- xini =ná$_i$ táta mí =ná$_i$.
no =they.FEM little.PL COMP:NEG see =they.FEM father the =their.FEM
'No girl, she didn't saw her own father.'

I argue that this "clitic left dislocation" can circumvent (78) because it does not involve movement in this language. This can be seen in two ways. First, like French (de Cat 2007), clitic left dislocation in SMPM is not island sensitive. To see this, compare the (a) examples in (80-81), which demonstrate attempted wh-movement out of adjunct and wh-islands respectively, with the grammatical clitic left dislocation (b) examples.

(80) Adjunct island

a. *Yó$_i$ ntsí'i kwé'e ini Julio [chi nì- xì'ì __$_i$]?
who sad very in Julio because COMP- die
Intended: 'Who is Julio sad because __ died?'

b. Raúl$_i$ kúsijǐ ini Maria [chi kachûn =rà$_i$ itsyààn].
Raul be.happy:CONT in Maria because work:IRR =he tomorrow
'Raul$_i$, Maria is happy because he$_i$ will work tomorrow.'

(81) Wh-island

a. *Nă$_i$ xǐnì =ú [năchûn xìxi Diego __$_i$ bìjkŏ]?
what know:CONT.NEG =you why eat:COMP Diego party
Intended: 'What do you not know why Diego ate __ at the party?'

b. Mí tsǐnà$_i$ kò kúntàà ini Javier [năchûn ntá'yi =rí$_i$].
the dog NEG believe:CONT in Javier why bark:CONT =it.AML
'The dog$_i$, Javier doesn't know why it$_i$ is barking.'

Second, indefinites that undergo clitic left dislocation fail to reconstruct for scope. Rather, only a wide scope, *de re* interpretation is available, similar to Modern Greek as reported by Alexopoulou and Folli (2019). This is shown in (82), where we see that the clitic left dislocated indefinite in (82a) can be followed by the *de re* context in (82b), but not the *de dicto* context in (82c).

(82) a. Iin *dragón$_i$* kôni =rà tsyàja ka'ani =rà =rí$_i$...
one dragon want:CONT =he man kill:IRR =he =it.AML
'A dragon, a man wants to kill it...'

b. ... ra yúkŭ íyo =rí.
and forest be:CONT =it.AML
'...and it is in the forest.' (✓ *de re*, ∃ > want)

c. ?... so kò xá'anì =rà mí =rà, ntsyân =rí nikuù =ba.
but NEG care:CONT =he the =he which =it.AML ever =EMPH
'...but he doesn't care which one.' (**de dicto*, want > ∃)

Together, (80-82) indicate that clitic left dislocation in SMPM does not involve movement. Rather, the pronoun in argument position (e.g. *=rí* 'it (an animal)' in 82a) seems to be base generated and linked to the dislocated nominal via a process other than movement. This allows us to understand why clitic left dislocation constitutes a repair to (78): the base generated pronoun, and not the dislocated quantified DP, antecedes the exempt anaphor.

With this in mind, consider the behavior of exempt anaphors in untensed subjunctives. To begin, recall our two analyses of control, repeated in (83).

(83) a. $^{\text{matrix}}$[Subject$_i$ ... $^{\text{control}}$[Subject$_i$ ...]] (Base generation)

b. $^{\text{matrix}}$[Subject … $^{\text{control}}$[~~Subject~~ …]] (Movement)

If we adopt the movement analysis in (83b), then we expect exempt anaphors to not be available in an untensed subjunctive with a quantified matrix controller. This is because the quantified subject would syntactically be in both matrix and embedded subject position, with the pronominal subject of the untensed subjunctive a sort of reduced, overt copy (e.g., Kandybowicz 2006, Harizanov 2014). As such, an exempt anaphor in the embedded clause would have no licit antecedent[25], as (78) would be violated. A sample derivation is provided in (84).

(84) $^{\text{Matrix}}$[Q DP$_i$ … $^{\text{untensed subj.}}$[~~Q DP~~ub*i* … *exempt anaphor*$_i$]]

Prediction: ungrammatical, violates (78)

In contrast, a base generation account posits two distinct syntactic items, the matrix controller and the pronominal subject of the untensed subjunctive. This means that an exempt anaphor should be possible within the untensed subjunctive; the pronominal subject would serve as the antecedent, just like the base generated pronoun under clitic left dislocation in (79). This is shown in (85).

(85) $^{\text{Matrix}}$[Q DP$_i$ … $^{\text{untensed subj.}}$[D$_i$ … *exempt anaphor*$_i$]]

Prediction: grammatical, D antecedes exempt anaphor

Given these opposite predictions, the availability of exempt anaphors in untensed subjunctives with quantified controllers can diagnose the derivation of control in SMPM: if exempt anaphors are available in this configuration, this suggests a base generation analysis is most appropriate, but if they are not available, this would suggest a movement analysis. As it turns out, exempt anaphors are grammatical in untensed subjunctives with quantified controllers without clitic left dislocation. This can be seen in comparing (86) with (75).

(86) a. Tá'iin'iin tsǐnà$_i$ kìxă [tsii =rí$_i$ ndò'ò mí =rí$_i$.]
each dog start:COMP bite:IRR =it.AML tail the =its.AML
'Each dogs started to bite their own tails.'

b. Ni'iin =rà bálí$_i$ kò xíniñu'u [kònì =rà$_i$ táta mí =rà$_i$].
no =he little.PL NEG need:CONT see:IRRR =he father the =his
'No boy needs to see his own father.'

Note as well that the quantifier fronting in (86) is not necessary, as shown in (87).

(87) a. Kìxă iin tsǐnà$_i$ [tsii =rí$_i$ ndò'ò mí =rí$_i$].
start:COMP one dog bite:IRR =it.AML tail the =its.AML
'A dog started to bit its own tail.'

b. Kò xíniñu'u ni'iin =rà bálí$_i$ [kònì =rà$_i$ táta mí =rà$_i$].
NEG need:CONT no =he little.PL see:IRRR =he father the =his
'No boy needs to see his own father.'

In this way, a movement analysis of control makes incorrect predictions for SMPM with respect to the distribution of exempt anaphors.[26] Stated informally, exempt anaphors require that the pronominal subject of a untensed subjunctive be a bona fide pronoun in order to have an antecedent that does not violate (78). The copy theory of movement cannot provide this pronoun, as syntactically the apparent pronoun would be a copy of the infelicitous quantified antecedent. As such, let us move forward under a base generation approach to control and see how it can derive pronominal copy control in SMPM.

7 A morphological analysis

We begin with the desiderata for any analysis of pronominal copy control in SMPM. First, we saw in Section 4.1 that the pronominal subject of untensed subjunctives must be interpreted as a bound variable. Second, Section 6 showed that this pronoun must be base generated, not derived via movement. From this perspective, then, deriving pronominal copy control in SMPM becomes a question of the exponence of bound variable anaphora.

The exponence of bound variable anaphora varies widely cross-linguistically. For example, bound variable anaphora in English are exponed as *-self* reflexives when locally bound, but as PRO or as pronouns in other positions. This is shown in (88).

(88) a. Jose saw { *herself, *PRO, her }. (Referential pronoun)

b. No girl$_i$ saw { herself$_i$, *PRO$_i$, *her$_i$ } (Reflexive)

c. No girl$_i$ started { *herself$_i$, PRO$_i$, *her$_i$ } to dance. (Control)

d. No girl$_i$ saw { *herself$_i$, *PRO$_i$, her$_i$ } mother. (Bound variable pronoun)

The pattern in (88) is not universal, though. For example, reflexive anaphors, controlled subjects (e.g., PRO), and bound variable pronouns are all syncretic with referential pronouns in Quiegolani Zapotec, shown in (89), with the bound variables underlined.[27]

(89) Quiegolani Zapotec (Black 1994)

a. S- ya men ru x- yuu men. (Referential pronoun)
PR- go 3RD mouth POSS- house 3RD
'He is going to his house.' (pg. 64)

b. R- wii men$_i$ <u>men$_i$</u>. (Reflexive)
HAB- see 3RD 3RD
'She/he/they$_i$ see(s) him/her/themself$_i$.' (pg. 95)

c. Zem r- laan Jose$_i$ ts- a <u>men$_i$</u> Laa. (Control)
seem HAB- want Jose POT- go 3RD Oaxaca
'Jose$_i$ seems to want PRO$_i$ to go to Oaxaca.' (pg. 80)

d. R- a txup tson wnaa$_i$ r- ka <u>men$_i$</u> gyus. (Bound variable pronoun)
HAB- go two three woman HAB- buy 3RD pot
'A few women$_i$ went and (they$_i$) bought a pot.' (pg. 103)

In other languages with syncretism among bound variables, not all bound variable anaphora are isomorphic like Quiegolani Zapotec in (89). For example, reflexive anaphors in Haitian are syncretic with bound variable pronouns, but not with controlled subjects, which are null (i.e., PRO),[28], as in (90).[29]

(90) Haitian

a. Jak$_i$ tchwe l$_j$. (Referential pronoun)
Jak kill 3SG
'Jak$_i$ killed him/her/it$_j$.' (Déchaine and Manfredi 1994:209)

b. Li$_i$ wè l$_i$ nan glas la. (Reflexive)
3SG see 3SG in mirror the
'S/he$_i$ saw himself/herself$_i$ in the mirror.' (Déchaine and Manfredi 1994:203)

c. Ti medam yo pa vle [PRO$_i$ koute msye]. (Control)
girl the.PL NEG want PRO listen gentleman
'The girls$_i$ do not want PRO$_i$ to listen to the gentleman.' (Targète and Urciolo 1993:19)

d. Chak moun$_i$ dwe defann peyi 'l$_i$. (Bound variable pronoun)
every person should defend country 3SG
'Every person$_i$ should defend his$_i$ country.' (Targète and Urciolo 1993:43)

From this perspective, pronominal copy control in SMPM can be understood as syncretism similar to that in Haitian. Rather than syncretism between reflexive anaphors and bound variable pronouns, though, in SMPM syncretism occurs between controlled subjects and bound variable pronouns. This can be seen in (91b-c).

(91) San Martín Peras Mixtec

a. Iin tsǐnà$_i$ xìin =ì. Rí chînù =rí$_i$. (Referential pronoun)
one dog buy:COMP =I it.AML curly =it.AML
'I bought a new dog$_i$. It$_i$ has curly fur.'

b. Ni'iin tsǐnà$_i$ kò nì- tsìi mí =rí$_i$. (Reflexive)
no dog NEG COMP- bite:COMP the =it.AML
'No dog$_i$ bit itself$_i$.'

c. Ni'iin tsǐnà$_i$ kò nì- kixă [nta'yi =rí$_i$]. (Control)
no dog NEG COMP- start cry:IRR =it.AML
'No dog$_i$ started PRO$_i$ to bark.'

d. Ni'iin tsǐnà$_i$ kò nì- tsìi ndò'ò =rí$_i$. (Bound variable pronoun)
no dog NEG COMP- bite tail =its.AML
'No dog$_i$ bit its$_i$ tail.'

These findings are summarized in the table in (92), in which '=' signifies that the relevant cells are syncretic, while 'X' signifies that they are not.[30]

(92)

		Referential pronoun	Reflexive	Controlled subject (e.g. PRO)	Bound variable pronoun
a.	English	=	X	X	=
b.	Quegolani Zapotec	=	=	=	=
c.	Haitian	=	=	X	=
d.	SMPM	=	X	=	=

In order to account for morphological diversity and syncretism among bound variable anaphora, let us adopt the framework of Kratzer 2009, Safir 2014, Landau 2015, and Sulemana 2021. Building off Kratzer 1998, this literature proposes that all instances of bound variable anaphora are syntactically represented by a single element, termed a minimal pronoun by Kratzer 2009 and Landau 2015 and D-bound by Safir 2014. For concreteness, let us assume the specific formulation of Kratzer 2009 and Landau 2015, 2018. In this system, all bound variables are syntactically represented as *minimal pronouns*, bare heads of category D that enter the derivation with only a set of unvalued ϕ-features, as in (93).

(93) X is a minimal pronoun if and only if X = [D,$u\phi$] (Landau 2015:23)

From this single element, "all anaphoric diversity is morphological" (Safir 2014:93). Stated differently, "[a minimal pronoun] can become a reflexive, a bound lexical pronoun. . .or indeed. . .a controlled PRO. The choice among these options is determined by a combination of the syntactic context and the lexical inventory of the language" (Landau 2018:12). See

as well the UPRO of McFadden and Sundaresan 2018 for a similar idea. Specifically, this literature proposes that different exponents of bound variable anaphora are determined post-syntactically via the routine tools of contextual allomorphy applied to a minimal pronoun.[31] Several sample vocabulary items are provided in (94) for English. Note that the notation [$\not\pi\phi$] indicates that these vocabulary items are meant to apply at a stage in the derivation after the minimal pronoun has received a ϕ-value from its antecedent, and are intended to be a shorthand for a more articulated series of vocabulary items, e.g., "$D_{[+female],\,[+sing]} \rightarrow$ herself / when locally bound" etc. See the sample derivation in Safir 2014:95 for precedent for this type of shorthand.[32]

(94) Minimal pronoun vocabulary items for English

a. $D_{[\not\pi\phi]} \rightarrow \emptyset$ (i.e., PRO) / ...

"A minimal pronoun with valued ϕ-features is exponed as null (i.e., PRO) in subject position of a controlled clause, e.g., Spec,FinP (Landau 2015), etc."

b. $D_{[\not\pi\phi]} \rightarrow$ reflexive anaphor / ...

"A minimal pronoun with valued ϕ-features is exponed as a (ϕ-appropriate) reflexive when locally bound (Kratzer 2009, Safir 2014, etc.)."

c. $D_{[\not\pi\phi]} \rightarrow$ pronoun

"A minimal pronoun with valued ϕ-features is exponed as a (ϕ-appropriate) pronoun."

The relationship among allomorphs in (94) is important. In particular, observe that (94c) does not have a licensing environment. Rather, so long as one of the specific licensing contexts in (94a-b) is not met, a minimal pronoun will be exponed as a pronoun identical to a referential pronoun, deriving the common syncretism between bound and referential pronouns in (92). See as well Kratzer 2009 and Safir 2014:92.

Safir 2014 argues that this 'pronoun as Elsewhere' phenomenon in a cross-linguistic

universal, citing speculative functional pressure related to anaphor resolution. While its universality remains to be verified, there is considerable evidence that this idea may be on the right track. See Rooryck and Vanden Wyngaerd 2011:19 for a similar proposal specific to reflexive anaphors, as well as the cross-linguistic investigation in Middleton 2020. As such, while its *raison d'être* remains somewhat murky, the proposal that pronouns are the universal default exponence for bound variable anaphora has empirical support and is less hypothetical than may seem at first glance.[33]

To see how this system works, consider the sample derivation in (95). Let us assume that the minimal pronoun, denoted as X following Landau 2015, is locally bound by 'John,' without adopting a specific notion of locality. Let us assume as well that ϕ-features have been transmitted from 'John' to X.

(95) John$_i$ saw X$_i$ in the mirror.

X in (95) will be subject to the vocabulary item in (94b) because it is locally bound, however defined, yielding "John saw himself in the mirror" in English. Recall, though, that vocabulary items such as (94) are not universal, however common (Safir 2014:119), but subject to language-specific variation. This means we allow for a language without a vocabulary item like (94b). Haitian Creole, as discussed in (90), seems to be precisely such a language. Sample vocabulary entries are givens in (96).

(96) Minimal pronoun vocabulary items for Haitian Creole

a. $D_{[\cancel{\pi\phi}]} \rightarrow \emptyset$ (i.e., PRO) / ...

"A minimal pronoun with valued ϕ-features is exponed as null (i.e., PRO) in subject position of a controlled clause, e.g., Spec,FinP (Landau 2015), etc."

b. $D_{[\cancel{\pi\phi}]} \rightarrow$ pronoun

"A minimal pronoun with valued ϕ-features is exponed as a (ϕ-appropriate)

pronoun."

Without a vocabulary item corresponding to (94b), a locally bound minimal pronoun in Haitian Creole will be exponed as a pronoun via the hypothesized universal Elsewhere item in (96b). This derives the pattern observed in (90b). See also Zribi-Hertz 1995, Rooryck and Vanden Wyngaerd 2011, and Déchaine and Wiltschko 2017 for a similar idea.

By this token, consider a sample derivation like (97) in which the minimal pronoun is in subject position of a control clause.

(97) Every girl$_i$ started [X$_i$ to sing].

Abstracting away from the exact licensing context (e.g., caseless subject of nonfinite clause (Chomsky and Lasnik 1993), Spec,FinP (Landau 2015), etc.), let us assume that X in (97) is in a position where it is subject to the vocabulary item in (94a) or (96a). As such, X will correctly be exponed as PRO in English and Haitian, c.f. (88c) and (90c).

Just as we proposed in (96) that Haitian lacks a vocabulary item like to (94b), though, we also can imagine a language that lacks a vocabulary item like to (94a). I analyze SMPM as just such a language. (98) provides the proposed vocabulary items for this language.

(98) Minimal pronoun vocabulary items for SMPM

a. $D_{[\#\phi]} \rightarrow$ *mí* + pronoun / ...

"A minimal pronoun with valued ϕ-features is realized as a (ϕ-appropriate) pronoun preceded by *mí* when locally bound."

b. $D_{[\#\phi]} \rightarrow$ pronoun

"A minimal pronouns with valued ϕ-features is exponed as a (ϕ-appropriate) pronoun."

Without a vocabulary item corresponding to (94a), SMPM must use the Elsewhere item

in (98b) for non-locally bound contexts like (97), yielding an overt pronoun in subject position of the control clause. Indeed, and unsurprisingly given what we have seen throughout, this is what we see. Compare (97) to (99).

(99) Ni'iin =ná bálí$_i$ kò nì kixă [kata =ná$_i$].
no =they.FEM little.PL NEG COMP- start sing:IRR =they.FEM
'No girl started to sing.'

As such, (98) successfully and straightforwadly derives pronominal copy control in SMPM, using only tools routine to morphological analysis. Furthermore, this analysis meets both of our desiderata at the start of this section. First, there is no question that a minimal pronoun will be interpreted as a bound variable, as this is what minimal pronouns were designed to do (Kratzer 2009, Safir 2014). Second, minimal pronouns are a subcategory of pronoun and are not derived via a movement chain. As such, there is no a priori reason why they should not be able to antecede exempt anaphors, as we saw was necessary in Section 6.

Proposing that SMPM lacks a null allomorph for bound variable anaphora further predicts that other types of PRO will never be silent, not just obligatory control PRO bound by subjects into complement clauses examined so far. This turns out to be the case. First, both adjunct clauses that require control, such as with the goal clauses in (100), and those that do not, such as the temporal clauses in (101), require overt pronominal subjects. See Landau 2021 for a recent summary of obligatory and non-obligatory control into adjunct clauses.

(100) a. Nì- xà'à =rà$_i$ nùya'abi [kwiin { =rà$_{i,\ *j}$, *PRO$_i$ } xuxa yáa].
COMP- go =he market buy:IRR =he wax white
'He went to the market (in order) to buy white wax.'

b. Nàki'i mí leso$_i$ kwà'an =rí$_i$ [kuxi { =rí$_i$, *PRO$_i$ }].
exit:COMP the rabbit go:PROG =it.AML eat:IRR =it.AML
'The rabbit went out (in order) to eat.'

(101) a. Tá ntsí'i { =nà, *PRO } ki'i =nà nijǐ, sáko'oyǒ =nà =ñà.
when end:CONT =they get:IRR =they corn shuck =they =it
Provided for: 'After harvesting the corn, they shuck it.'

b. Kôni ku'un { =ñá, *PRO } nùya'abi, tsyàa Maria iin *carta*.
after go:IRR =she market write:COMP Maria one letter
'Before going to the market, Maria wrote a letter.'

Non-obligatory control also requires an overt pronoun in place of PRO_{Arb} as well, specifically an impersonal pronoun: *=nà* 'they' for exclusive impersonal readings, as in (102), and *=é* 'we (INCL)' for inclusive impersonal readings, as in (103).[34]

(102) a. Á kundê { =nà, *PRO_{Arb} } iin *mil* kwîyà nujǔ yòjo.
Q be.PL:IRR =they one thousand year on moon
Provided for 'It will be possible PRO_{Arb} to live on the moon in 1000 years.'

b. Tá xina'á, xìniñu'u =á [kachûn miǐtún { =nà, *PRO_{Arb} }].
in antiquity need:COMP =it work:IRR alone =they
Provided for 'In antiquity, it was necessary PRO_{Arb} to work alone.'

(103) a. Ná kà'àn { =é, *PRO_{Arb} } ñà ntàà.
NÁ speak:IRR we.INCL it true
Provided for 'It is necessary PRO_{Arb} to tell the truth.'

b. Xíniñu'u =á koo miǐtún { =é, *PRO_{Arb} } tásaba.
need:CONT =it COP:IRR alone =we.INCL sometimes
Provided for 'It is important PRO_{Arb} to be alone sometimes.'

Both of these patterns have a ready analysis under the present proposal: without a null allomorph of the minimal pronoun, overt pronouns must be used. This reasoning also extends to object control, as nothing in the proposal here distinguishes subject from object control. Indeed, this pans out: object control requires the same overt pronominal subjects as subject control, as shown in (104).

(104) a. Nì- kà'àn Maria xí'in Juân$_i$ [kasa'a { =rà$_i$, *PRO } ixìjni].
COMP- speak Maria with Juan make:IRR he hat
'Maria told Juan to make hats.'

b. Sànă'a Maria Juân$_i$ [ixutsya { =rà$_i$, *PRO }].
teach:CONT Maria Juan swim:IRR =he
'Maria taught Juan how to swim.'

This analysis leaves open an interesting question: if deriving overt PRO is so theoretically straightforward, why are languages like SMPM so rare, apparently only being described in several West African languages (Allotey 2021, Sulemana 2021)? Why is a null allomorph so common for PRO that it was believed to be universal? *Are* languages like SMPM rare? While I do not have answers to these questions, I suggest that null controlled subjects are as pervasive as was once thought. For instance, 'overt PRO' is ubiquitous Oto-Manguean, the family of SMPM and Quiegolani Zapotec in (89).[35]

(105) Pronominal copy control across Oto-Manguean

a. Jíní tēe$_i$ ún cahu de$_i$ tutu.
know:CONT man the read:IRR he paper
'The man knows how to read.' (Atatlahuca Mixtec, Alexander 1980:167)

b. Ne'en^{3} cha^{3}nà$^{1}{}_{i}$ canànj^{2} no'$^{3}{}_{i}$ a.
know:CONT woman weave:IRR she DECL
'The woman knows how to weave.' (Copala Trique, Erickson de Hollenbach 2008:82)

c. R- yenlaz Jwáyn$_i$ gu- segw =an$_i$ ru'næz.
HAB- forget Juan POT- close =3SG.IF entrance
'Juan forgets to close the main entrance.' (Teotitlán del Valle Zapotec, Gutiérrez Lorenzo 2021:245)

d. Nchka 'in =ne' tas'a$_i$ =ăn ku'ni =ne'$_i$ tñan kan'.
be.able:CONT of =they brother =my do:IRR =they work that
'My brothers are able to do that work.' (Highland Chatino, Pride and Pride 2010:387)

e. Hi^{2}meh$^{2}{}_{i}$ lang^{12} ma^{2}- mi^{2}tan^{12} ha^{2} láh^{2} høa^{12} =tsih$_i$ si^{2}.
boy this ASP- learn how read:IRR =he.YOUNG letter
'This boy is learning how to read.' (San Pedro Tlatepuzco Chinantec, Merrifield and Anderson 2007:16)

f. Ma- tseijndeii ñetyjò$_i$ na ma- tseiwee =$^{n}_{i}$ cwii xuu
CONT.SG- strive my.brother C CONT.SG- lift =3SG one load
t'ma^{n}.
big
'My brother is striving to lift a heavy load.' (Xochistlahuaca Amuzgo, Buck 2018:304)

Such patterns are readily found outside Mesoamerica as well, with logophoric pronominal copy control under 'want' of the sort discussed by Grano and Lotven 2018, 2019 in section 5 particularly well-attested. Several examples are provided in (106).

(106) a. Aɣa$_i$= bbáaɣa sa aɣ$_i$= tə- didá.
I= want COMP I= FUT- walk
'I want to go.' (Tadaksahak, Songhay, Christiansen-Bolli 2010:175)

b. Hasaga awta$_i$ dm yeltg =t$_i$ da lax galts'ap.
want porcupine FUT return =3 to place village
'Porcupine wanted to return to the mainland.' (Sm'algyax, Tshimshianic, Mulder 1988)

c. Ávyéjuubéváa$_i$ imíléh wájyamu i$_i$ újcune.
chief want cloth self get
'The chief wants to get cloth.' (Bora, Bora-Witoto, Thiesen and Weber 2012:482)

d. Jái uniábi$_i$ wówai ǐli$_i$= wína wayai.
now winter want 3SG.MASC= catch us
'Now winter wants to surprise us.' (Achagua, Arawak, Wilson 1992:155)

e. Duu$_i$ to=xo=bo mamaa rö iduu$_i$ ra alolo.
3DU 3NSG =PAST =HAB want for 3DU PURP marry
'They (two) wanted to get married.' (Barok, Austronesian, Du 2010:120)

While less common, apparent pronominal copy control is also found outside of attitude reports, similar to SMPM and Oto-Manguean in (105). (107) provides several examples.

(107) a. Mo$_i$ mandi tauhunju mo$_i$ hanhani.
3SG simply start 3SG eat.REDUP
'He simply started to eat.' (Tamambo, Austronesian, Jauncey 2011)

b. Ra$_i$= mo tavuigi ra$_i$= mo mwasigi.
they= REAL start they= REAL bear.fruit
'They are starting to bear fruit.' (Ambae, Austronesian, Hyslop 2001:456)

c. Sàlàd è ālde$_i$ dōòs -so ē$_i$ wàj.
hyena with fox start -COMP they go
'A hyena and fox started to go.' (Gaahmg, Nilo-Saharan, Orta Adaw et al. 2015:227)

d. A$_i$= cakôc to a$_i$= crà.
you= speak C you= continue
'You continue to speak.' (Canela-Krahô, Jê, Popjes and Popjes 1986:181)

e. Nuyá$_i$ jó- ka éewa nu$_i$= íinua -ka bándola.
I NEG -AFF able I= play -AFF bandola
'I cannot/don't know how to play bandola.' (Achagua, Arawak, Wilson 1992:156)

f. Aɣa$_i$= b- hurrú sa aɣ$_i$= tə- kár a= ka huurú.
I= IMPERF- try COMP I= FUT- hit 3SG= LOC fire
'I try to make fire on it.' (Tadaksahak, Songhay, Christiansen-Bolli 2010:211)

Whether the patterns in (105-107) involve control, and if so its derivation, is unclear. Taken at face value, then, any theory which universally requires PRO to be silent or for control clauses to lack a subject position seems to be simply untenable.

8 Conclusion

This article considered control constructions in SMPM. In this language, we saw that one variety of subjunctive clause, untensed subjunctives, display the characteristic properties of control despite requiring overt clitic pronouns as subjects. From this, we concluded that silence cannot be a necessary property of controlled subjects, e.g. PRO.

We also saw that this pattern falls out naturally from recent approaches to control that treat PRO as a minimal pronoun whose exponence is determined by language-specific vocabulary items (Kratzer 2009, Safir 2014, Landau 2015, 2018). In this system, the pattern

in SMPM was derived by simply not positing a null allomorph. Part of the appeal of this approach is its compatibility with a variety of syntactic analyses of control, so long as PRO is analyzed as a minimal pronoun and the morpho-syntactic context of controlled subjects is unique enough to be referenced by a contextual allomorph. Likewise, this is accomplished without any special treatment for PRO or any additional ad hoc mechanism except, perhaps, the claim that the default allomorph for bound variables is a pronoun.

References

Alexander, Ruth María. 1980. Gramática Mixteca: Mixteco de Atatlahuca. Number 2 in Series de gramáticas de lenguas indígenas de México. Distrito Federal, Mexico: Instituto Lingüístico de Verano.

Alexopoulou, Theodora, and Raffaella Folli. 2019. Topic strategies and the internal structure of nominal arguments in Greek and Italian. Linguistic Inquiry 50.

Allotey, Deborah. 2021. Overt pronouns of infinitival predicates in Gã. Western Papers in Linguistics/Cahiers linguistiques de Western 4.

Barbosa, Pilar. 2009. A case for an Agree-based theory of control. Ms., Universidade do Minho/CEHUM.

Barbosa, Pilar. 2018. Controlled overt pronouns as specificational predicates. In Complement Clauses in Portuguese: Syntax and acquisition, ed. Ana Lúcia Santos and Anabela Gonçalves, volume 17 of Issues in Hispanic and Lusophone Linguistics, 129–186. John Benjamins.

Barker, Chris. 2012. Quantification binding does not require c-command. Linguistic Inquiry 43:614–633.

Black, Cheryl. 1994. Quiegolani Zapotec syntax. Doctoral Dissertation, University of California, Santa Cruz.

Bobaljik, Jonathan. 2012. Universals in Comparative Morphology: Suppletion, superlatives, and the structure of words. Cambridge, MA: MIT Press.

Boeckx, Cedric, and Norbert Hornstein. 2004. Movement under Control. Linguistic Inquiry 35:431–452.

Boeckx, Cedric, and Norbert Hornstein. 2006. Control in Icelandic and theories of control. Linguistic Inquiry 37:591–606.

Boeckx, Cedric, Norbert Hornstein, and Jairo Nunes. 2009. Copy-reflexive and copy-control constructions: A movement analysis. Linguistic Variation Yearbook 8:61–100.

Bradley, C. Henry, and Barbara (eds) Hollenbach. 1988-1992. Studies in the Syntax of Mixtecan Languages. The Summer Institute of Linguistics and The University of Texas at Arlington. Volumes 2-4 are a part of the Publications in Linguistics Series.

Bresnan, Joan. 1972. Theory of complementattion in English syntax. Doctoral Dissertation, MIT.

Bresnan, Joan. 1978. A Realistic Transformational Grammar. In Linguistic Theory and Psychological Reality, ed. Morris Halle, Joan Bresnan, and George Miller, 1–60. Cambridge, MA: MIT Press.

Broadwell, George Aaron. 2003. Optimality, Complex Predication, and Parallel Structures in Zapotec. In Proceedings of the LFG03 Conference, ed. Miriam Butt and Holloway King, 75–91. Stanford, CA: CSLI Publications.

Broadwell, George Aaron. 2019. Documenting control and pseudo-control in Copala Triqui. Ms. available here.

Buck, Marjorie. 2018. Gramática del amuzgo de Xochistlahuaca, Guerrero. Number 16 in Series de gramáticas de lenguas indígenas de México. Mexico City: Instituto Lingüistico de Verano.

Bui, Thuy. 2019. Binding and Coreference in Vietnamese. Doctoral Dissertation, UMass Amherst.

Caha, Pavel. 2018. Notes on insertion in Distributed Morphology and Nanosyntax. In Exploring Nanosyntax, ed. Lena Baunaz, Liliane Haegeman, Karen De Clercq, and Eric Lander, Oxford Studies in Comparative Syntax, 57–87. New York: Oxford University Press.

Cardinaletti, Anna, and Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Clitics in the Languages of Europe, ed. Henk van Riemsdijk, 145–290. Berlin: Mouton de Gruyter.

de Cat, Cécile. 2007. French Dislocation without Movement. Natural Language & Linguistic Theory 25:485–534.

Charnavel, Isabelle, and Dominique Sportiche. 2016. Anaphor Binding: What French Inanimate Anaphors Show. Linguistic Inquiry 47:35–87.

Chierchia, Gennaro. 1984. Topics in the Syntax and Semantics of Infinitives and Gerunds. Doctoral Dissertation, UMass, Amherst.

Chierchia, Gennaro. 1990. Anaphora and attitudes *de se*. In Semantics and Contextual Expression, ed. Renate Bartch, Johan van Benthem, and Peter van Emde Boas, 1–33. Dordrecht: Foris.

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

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, ed. Roger Martin, David Michaels, and Keyser Samuel Jay Uriagereka, Juan, 89–156. 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, and Howard Lasnik. 1977. Filters and Control. Linguistic Inquiry 8:425–504.

Chomsky, Noam, and Howard Lasnik. 1993. The Theory of Principles and Parameters. In Syntax: An international handbook of contemporary research, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 506–569. Berlin: Mouton de Gruyter.

Christiansen-Bolli, Regula. 2010. A Grammar of Tadaksahak, a Northern Songhay Language of Mali. Doctoral Dissertation, Universiteit Leiden.

Chung, Sandra. 1989. On the notion 'null anaphor' in Chamorro. In The Null Subject Parameter, ed. Osvaldo Jaeggli and Ken Safir, 143–184. Dordrecht: Reidel.

Chung, Sandra. 1998. The design of agreement: Evidence from Chamorro. Chicago: University of Chicago Press.

Chung, Sandra. 2019. Freedom within Control. Talk presented at the Linguistics Association of Great Britain conference, Ulster University. Handout available here.

Cisneros, Carlos. 2019. Definiteness in Cuevas Mixtec. In Definiteness across languages, ed. Ana Aguilar-Guevara, Julia Pozas Loyo, and Violeta Vázquez-Rojas Maldonado, number 25 in Studies in Diversity Linguistics, 39–82. Berlin: Language Science Press.

Cline, Laura. 2018. Some Features of Participant Reference in Xochapa Mixtec. Master's thesis, University of North Dakota.

Cole, Peter, Gabriella Hermon, and Yanti. 2015. Grammar of Binding in the Languages of the World: Innate or Learned? Cognition 141:138–160.

Déchaine, Rose-Marie, and Victor Manfredi. 1994. Binding Domains in Haitian. Natural Language & Linguistic Theory 12:203–257.

Déchaine, Rose-Marie, and Martina Wiltschko. 2017. Bound variable anaphora. In The Wiley Blackwell Companion to Syntax, Second Edition, ed. Martin Everaert and Henk van Riemsdijk, 1–33. https://onlinelibrary.wiley.com/doi/book/10.1002/9781118358733: John Wiley & Sons, Inc.

den Dikken, Marcel. 1995. Binding, Expletives, and Levels. Linguistic Inquiry 26:347–354.

den Dikken, Marcel, Anikó Lipták, and Zsófia Zvolenszky. 2001. On inclusive reference anaphora: New perspectivs from Hungarian. In WCCFL 20, ed. Karine Megerdoomian and Leora Anne Bar-el, 137–149. Somerville, MA: Cascadilla Press.

Dobrovie-Sorin, Carmen. 1994. The syntax of Romanian. Comparative Studies in Romance. Berlin: Mouton de Gruyter.

Doliana, Aaron, and Sandhya Sundaresan. 2022. Proxy control: A new species of control in grammar. Natural Language & Linguistic Theory 40:43–101.

Du, Jingyi. 2010. Towards a Grammar of the Usen Dialect of the Barok language of New Ireland, Papua New Guinea. Doctoral Dissertation, La Trobe University.

Eischens, Ben, and Andrew Hedding. m.s. San Martín Peras Mixtec. Available here.

Embick, David. 2010. Localism versus Globalism in Morphology and Phonology, volume 60 of Linguistic Inquiry Monograph. Cambridge, MA: MIT Press.

Evans, Gareth. 1980. Pronouns. Linguistic Inquiry 11:337–362.

Fodor, Jerry. 1975. The Language of Thought. Cambridge, MA: Harvard University Press.

Ganenkov, Dmitry. To appear. Partial control with overt embedded subjects in Chirag. Language .

García Baltazar, Leonardo. 2019. Gramática Popular Mazateca: Mazateco de San Jerónimo Tecóatl, Oaxaca. Instituto Lingüistico de Verano. Available here.

Givón, Talmy. 1980. The binding hierarchy annd the typology of complements. Studies in Language 4:333–377.

Grano, Thomas. 2012. Control and restructuring at the syntax-semantics interface. Doctoral Dissertation, University of Chicago.

Grano, Thomas, and Samson Lotven. 2018. Control, logophoricity, and harmonic modality in gengbe desire reports. In Proceedings of *Sinn und Bedeutung* 21, ed. Robert Truswell, Chris Cummins, Caroline Heycock, Brian Rabern, and Hannah Rohde, 481–498. Published online through semanticsarchive.net.

Grano, Thomas, and Samson Lotven. 2019. Control of logophoric pronouns in Gengbe. In African linguistics across disciplines: Selected papers from the 48th Annual Conference on African Linguistics, ed. Samson Lotven, Silvina Bongiovanni, Philip Weirich, Robert Botne, and Samuel Gyasi Obeng, 325–337. Berlin: Language Science Press.

Gutiérrez Lorenzo, Ambrocio. 2021. A description and analysis of the syntax and functions of subordinate clausess in Teotitlán del Valle Zapotec. Doctoral Dissertation, University of Texas, Austin.

Haddad, Youssef. 2009. Copy Control in Telugu. Journal of Linguistics 45:69–109.

Haddad, Youssef. 2011. Control into Conjunctive Participle Clauses: The case of Assamese. Berlin: De Gruyter.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger, ed. Ken Hale and Samuel Jay Keyser, 111–176. Cambridge, MA: MIT Press.

Harizanov, Boris. 2014. Clitic doubling at the syntax-morphophonology interface: A-movement and morphological merger in Bulgarian. Natural Language & Linguistic Theory Linguistic Theory 32:1033–1088.

Hedding, Andrew. 2022. How to move a focus: The syntax of alternative particles. Doctoral Dissertation, University of California, Santa Cruz.

Herbeck, Peter. 2015. Overt PRO in Romance: Towards a unification of PRO and pro. In Hispanic Linguistics at the Crossroads: Theoretical linguistics, language acquisition, and language contact. Proceedings of the Hispanic Linguistics Symposium 2013, ed. Rachel Klassen, Juana M. Liceras, and Elena Valenzuela, 25–48. Amsterdam: John Benjamins.

Higginbotham, James. 1980. Pronouns and bound variables. Linguistic Inquiry 11:679–708s.

Hollenbach, Barbara. 1984. Reflexives and Reciprocals in Copala Trique. International Journal of American Linguistics 50:272–291.

Erickson de Hollenbach, Elena. 2008. Gramática popular del triqui de Copala. Number 11 in Series de gramáticas de lenguas indígenas de México. Mexico City: Instituto Lingüistico de Verano.

Holmberg, Anders. 2005. Is there a little pro? evidence from Finnish. Linguistic Inquiry 36:533–564.

Holmberg, Anders. 2010. The null generic subject pronoun in Finnish: a case of incorporation in T. In Parametric variation: Null subjects in minimalist theory, ed. Theresa Biberauer, Anders Holmberg, Ian Roberts, and Michelle Sheehan, 200–230. Cambridge: Cambridge University Press.

Holmberg, Anders. 2010b. Null subject parameters. In Parametric Variation: Null subjects in minimalist theory, ed. Theresa Biberauer, Anders Holmberg, Ian Roberts, and Michelle Sheehan, 88–124. Cambridge University Press.

Holmberg, Anders, and Michelle Sheehan. 2010. Control into finite clauses in partial null-subject languages. In Parametric Variation: Null subjects in minimalist theory, ed. Theresa Biberauer, Anders Holmberg, Ian Roberts, and Michelle Sheehan, 125–152. Cambridge: Cambridge University Press.

Hornstein, Norbert. 1999. Movement and Control. Linguistic Inquiry 30:69–96.

Hornstein, Norbert, and Maria Polinsky. 2010. Movement Theory of Control. Amsterdam: John Benjamins.

Hyslop, Catriona. 2001. The Lolovoli dialect of the North-East Ambae language, Vanuatu. Number 515 in Pacific Linguistics. Canberra, Australia: The Austalian National University.

Jaeggli, Osvaldo. 1986. Arbitrary plural pronominals. Natural Language & Linguistic Theory 4:43–76.

Jauncey, Dorothy. 2011. Tamambo, the language of west Malo, Vanuatu. Canberra, Australia: Pacific Linguistics.

Kandybowicz, Jason. 2006. Conditions on multiple spell-out and the syntax-phonology interface. Doctoral Dissertation, UCLA.

Kiparsky, Paul, and Carol Kiparsky. 1970. Fact. In Progress in Linguistics: A collection of papers, ed. Manfred Bierwisch and Karl Erich Heidolph, volume 43 of Janua Linguarum. Series Maior, 143–173. The Hague: Mouton.

Kramer, Ruth. 2014. Clitic doubling or object agreement: The view from Amharic. Natural Language & Linguistic Theory 32:593–634.

Krapova, Iliyana. 2001. Subjunctives in Bulgarian and Modern Greek. In Comparative Syntax of Balkan Languages, ed. María Luisa Rivero and Angela Ralli, 105–126. Oxford: Oxford University Press.

Krapova, Iliyana, and Vassil Petkov. 1999. Subjunctive Complements, Null Subjects and Case Checking in Bulgarian. In Proceedings of the 7th Meeting of Formal Approaches to Slavic Linguistics, ed. Katarzyna Dziwirek, Herbert Coats, and Cynthia M. Vakareliyska, 267–285. University of Michigan, Ann Arbor: Michigan Slavic Press.

Kratzer, Angelika. 1998. More structural analogies between pronouns and tenses. In Proceedings of Semantics and Linguistic Theory VIII, ed. Devron Strolovitch and Aaron Lawson, 92–109. Cornell University: CLC Publications.

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

Landau, Idan. 2000. Elements of Control: Structure and meaning in infinitival constructions. Studies in Natural Language and Linguistic Theory. Dordrecht: Kluwer.

Landau, Idan. 2004. The Scale of Finiteness and the Calculus of Control. Natural Language & Linguistic Theory 22:811–877.

Landau, Idan. 2013. Control in Generative Grammar: A Research Companion. Cambridge: Cambridge University Press.

Landau, Idan. 2015. A Two-Tiered Theory of Control. Number 71 in Linguistic Inquiry Monographs. MIT Press.

Landau, Idan. 2018. Direct Variable Binding and Agreement in Obligatory Control. In Pronouns in Embedded Contexts at the Syntax-Semantics Interface, ed. Pritty Grosz-Patel, Patrick Georg Grosz, and Sarah Zobel, volume 99 of Studies in Linguistics and Philosophy, 1–42. Dordrecht: Springer International Publishing.

Landau, Idan. 2021. A Selectional Theory of Adjunct Control. Number 84 in Linguistic Inquiry Monographs. Cambridge, MA: MIT Press.

Lasnik, Howard. 1989. Essays on anaphora. Dordrecht: Reidel.

Lee, Felicia. 2003. Anaphoric R-expressions as bound variables. Syntax 6:84–114.

Levinson, Stephen. 1987. Pragmatics and the grammar of anaphora. Journal of Linguistics 23:379–434.

Levinson, Stephen. 1991. Pragmatic reduction of the Binding Conditions revisited. Journal of Linguistics 27:107–161.

Li, Daoxin. 2021. Controlling overt subjects in Mandarin. In Proceedings of the Linguistics Society of America, volume 6, 303–316. Linguistics Society of America.

Livitz, Inna. 2014. Deriving Silence through Dependent Reference: Focus on Pronouns. Doctoral Dissertation, New York University.

Macaulay, Monica. 1987. Morphology and Cliticization in Chalcatongo Mixtec. Doctoral Dissertation, University of California, Berkeley.

Macaulay, Monica. 1993. Argument Status and Constituent Structure in Chalcotongo Mixtec. In Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society: Special Session on Syntactic Issues in Native American Languages, ed. David Peterson, 73–85. Berkeley, CA: Berkeley Linguistics Society.

Macaulay, Monica. 2005. The syntax of Chalcotongo Mixtec: Preverbal and postverbal. In Verb First: On the syntax of verb-initial languages, ed. Andrew Carnie, Heidi Harley, and Sheila Ann Dooley, 341–366. Amsterdam: John Benjamins.

Madigan, Sean. 2008. Obligatory Split Control into Exhortative Complements in Korean. Linguistic Inquiry 39:493–502.

Mantenuto, Iara. 2020. 'Other' and 'More' in San Sebastián del Monte Mixtec and Beyond. Doctoral Dissertation, University of California, Los Angeles.

McCloskey, Jim, and Ken Hale. 1984. On the Syntax of Person-Number Inflection in Modern Irish. Natural Language & Linguistic Theory 1:487–533.

McFadden, Thomas, and Sandhya Sundaresan. 2018. Reducing *pro* and PRO to a single source. The Linguistics Review 35:463–518.

Merchant, Jason. 2015. How much context is enough? Two cases of span-conditioned stem allomorphy. Linguistic Inquiry 46:273–303.

Merrifield, William, and Alfred Anderson. 2007. Diccionario Chinanteco de la diáspora del pueblo antiguo de San Pedro Tlatepuzco. Number 39 in Serie de vocabularios y diccionarios indígenas. Mexico City: Instituto Lingüistico de Verano.

Middleton, Hannah Jane. 2020. *ABA syncretism patterns in pronomial morphology. Doctoral Dissertation, University College London.

Morgan, Jery. 1970. On the Criterion of Identity for Noun Phrase Deletion. In Proceedings of CLS 6, ed. Mary Ann et al. Campbell. Chicago: Chicago University Press.

Mulder, Jean Gail. 1988. Ergativity in Coast Tsimshian (Sm'algyax). Doctoral Dissertation, UCLA.

Nevins, Andrew. 2011. Multiple agree with clitics: Person complementarity vs. omnivorous number. Natural Language & Linguistic Theory 29:939–971.

North, Joanne. 1987. Partiticpant tracing and prominence in Silacayoapán Mixtec. In SIL - Mexico Workpapers, volume 8, 100–128. Summer Institute of Linguistics.

Orta Adaw, Hashi, Sadig Bashir Tingala, Huda Daldoum, Anur Omar Salva Adam, Musa Abdul hKim, Ezadin Abir Adik, and Yaasirs Khamis Konjur. 2015. Gaahmg Grammar Book: Trial Edition. Nairobi: Sudan Workshop Programme.

Ostrove, Jason. 2018. Stretching, spanning, and linear adjacency in Vocabulary Insertion. Natural Language & Linguistic Theory 36:1263–1289.

Ostrove, Jason. 2022. Impersonals in San Martín Peras Mixtec. In Proceedings of the 25th Workshop on Structure and Constituency in Languages of the Americas, ed. Marianne Huijsmans. University of British Columbia Working Papers in Linguistics.

Palancar, Enrique. 2012. Clausal juxtaposition and subordination: New light from Otomi. In Clause Combining in Cross-linguistics Perspective: Data-driven approaches to cross-clausal syntax, ed. Volker Gast and Holger Diessel, volume 249 of Trends in Linguistics: Studies and Monographs. Berlin: De Gruyter Mouton.

Partee, Barbaba. 1975. Deletion and Variable Binding. In Formal Semantics of Natural Languages, ed. Edward Keenan, 16–34. Cambridge: Cambridge University Press.

Patel-Grosz, Pritty. 2020. Pronominal typology and the *de se*/*de re* distinction. Linguistics and Philosophy 43:537–587.

Patel-Grosz, Pritty, and Patrick Grosz. 2017. Revisiting pronominal typology. Linguistic Inquiry 48:259–297.

Pearson, Hazel. 2013. The Sense of Self: Topics in the semantics of *de se* expressions. Doctoral Dissertation, Harvard University.

Peters, Simon. 2017. Inventario y distribución tonal en el mixteco de San Martín Peras. In Memorias del VIII Congreso de Idiomas Indígenas de Latinoamérica. University of Texas, Austin.

Polinsky, Maria, and Eric Potsdam. 2006. Expanding the scope of control and raising. Syntax 9:171–192.

Pollard, Carl, and Ivan Sag. 1992. Anaphors and the scope of binding theory. Lingustic Inquiry 23:261–303.

Popjes, Jack, and Jo Popjes. 1986. Canela-Krahô. In Handbook of Amazonian Languages, ed. Desmond Derbyshire and Geoffrey Pullum, volume 1, 128–1999. Berlin: Mouton de Gruyter.

Postal, Paul. 1974. On raising. Cambridge, MA: MIT Press.

Potsdam, Eric, and Youssef Haddad. 2017. Control Phenomena. In The Wiley Blackwell Companion to Syntax, ed. Martin Everaert and Henk van Riemsdijk. John Wiley & Sons, Inc., second edition edition.

Pride, Kitty, and Leslie Pride. 2010. Diccionaio Chatino de la Zona Alta: Panixtlahuaca, Oaxaca y otros pueblos. Number 47 in Serie de vocabularios y diccionarios indígenas. Mexico City: Instituto Lingüistico de Verano.

Ramchand, Gillian, and Peter Svenonius. 2014. Deriving the functional hierarchy. Language Sciences 46:152–174.

Rezac, Milan. 2010. ϕ-Agree versus ϕ-feature movement: Evidence from floating quantifiers. Linguistic Inquiry 41:496–508.

Rooryck, Johan, and Guido Vanden Wyngaerd. 2011. Dissolving binding theory. Number 34 in Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press.

Safir, Ken. 2014. One True Anaphor. Linguistic Inquiry 45:91–124.

Satık, Deniz. 2020. Logophoric pronouns are not (always) inherently logophoric. In Proceedings of the 55th Annual Meeting of the Chicago Linguistics Society (CLS 55). Chicago, IL: Chicago Linguistics Society.

Satık, Deniz. 2022. An economy theory of PRO. Available at https://ling.auf.net/lingbuzz/006833.

Schwarz, Florian. 2009. Two Types of Definites in Natural Language. Doctoral Dissertation, University of Massachusetts, Amherst.

Sevdali, Christina, and Michelle Sheehan. 2021. Rethinking partial control: New evidence from finite control clauses. In Syntactic architecture and its consequences III: Inside syntax, ed. András Bárány, Theresa Biberauer, Jamie Douglas, and Sten Vikner, 323–340. Berlin: Language Science Press.

Stewart, Donald. 1966. Gramática de Mazahua. Instituto Lingüistico de Verano. Avavilable here.

Stiebels, Barbara. 2007. Towards a typology of complement control. In Studies in complement control, number 47 in ZAS Papers in Linguistics, 1–59. Berlin: Zentrum Allgemeine Sprachwissenschaft.

Stowell, Tim. 1982. The Tense of Infinitives. Linguistic Inquiry 13:561–570.

Sulemana, Abdul-Razak. 2021. Non-finite Complementation: A case study of Bùlì. Doctoral Dissertation, MIT.

Szabolcsi, Anna. 2009. Overt Nominative Subjects in Infinitival Complements: Data, Diagnostics, and Preliminary Analyses. In NYU Working Papers in Linguistics, Vol. 2: Papers in Syntax, ed. Patricia Irwin and Violeta Vasquéz Rojas Maldonado. New York: NYU.

Targète, Jean, and Raphael G. Urciolo. 1993. Haitian Creole - English Dictionary. Kensington, MD: Dunwoody Press.

Terzi, Arhonto. 1992. PRO in Finite Clauses: A study of the inflectional heads of the Balkan languages. Doctoral Dissertation, CUNY Graduate Center.

Terzi, Arhonto. 1997. PRO and Null Case in Finite Clauses. The Linguistics Review 14:335–360.

Thiesen, Wesley, and David Weber. 2012. A Grammar of Bora with Special Attention to Tone. Number 148 in Publications in Linguistics. Dallas, TX: SIL International.

Thomason, Richard. 1976. Some Extensions of Montague Grammar. In Montague Grammar, ed. Barbaba Partee, 77–118. New York: Academic Press.

Volkova, Anna, and Eric Reuland. 2014. Reflexivity without reflexives? The Linguistic Review 31:587–633.

Wagers, Matt, Sandra Chung, and Manuel Borja. To appear. Processing reflexive pronouns when they don't announce themselves. Glossa Psycholinguistics .

Wilson, Peter. 1992. Una descripción preliminar de la gramática del Achagua (Arawak). Bogotá, Colombia: Instituto Lingüístico de Verano.

Wurmbrand, Susi. 1998. Infinitives. Doctoral Dissertation, MIT.

Wurmbrand, Susi. 2001. Infinitives: Restructuring and clause structure. Berlin: Mouton de Gruyter.

Wurmbrand, Susi. 2002. Syntactic vs. Semantic Control. In Studies in Comparative Germanic Syntax: Proceedings of the 15th Workshop on Comparative Germanic Syntax, ed. Jan-Wouter Zwart and Werner Abraham, 95–129. Amsterdam: John Benjamins.

Wurmbrand, Susi. 2004. Two types of restructuring-Lexiical vs. functional. Lingua 114:991–1014.

Wurmbrand, Susi. 2014. Tense and aspect in English infinitives. Linguistic Inquiry 45:403–447.

Wurmbrand, Susi. 2015. Restructuring cross-linguistically. In Proceedings of the 45th meeting of the North Eastern Linguistics society, ed. Thuy Bui and Deniz Özyıldız, 227–240. Amherst: University of Massachusetts, GLSA.

Wurmbrand, Susi. 2017. Icelandic as a partial null subject language - Evidence from fake indexicals. In Order and Structure in Syntax II: Subjecthood and argument structure, ed. Michelle Sheehan and Laura Bailey, 339–345. Berlin: Language Science Press.

Wurmbrand, Susi, Iva Kovač, Magdalena Lohninger, Caroline Pajančič, and Neda Todorović. 2020. Finiteness in South Slavic complement clauses: Evidence for an implicational finiteness universal. Linguistica 60:119–137.

Wurmbrand, Susi, and Magdalena Lohninger. 2019. An implicational universal in complementation: Theoretical insignts and emprical progress. In Propositional Arguments in Cross-Linguistic Research: Theoretical and Empirical Issues, ed. Jutta Hartmann and Angelika Wöllstein. Tübingen: Gunter Narr Verlag.

Wurmbrand, Susi, and Koji Shimamura. 2017. The features of the voice domain: Actives, passives, and restructuring. In The verbal domain, ed. Roberta D'Alessando, Irene Franco, and Ángel Gallego, 179–204. Oxford: Oxford University Press.

Zribi-Hertz, Anne. 1995. Emphatic or Reflexive? On the Endophoric Character of French *lui-même* and Similar Complex Pronouns. Journal of Linguistics 31:333–374.

Zribi-Hertz, Anne, and Liliane Mbolatianavalona. 1999. Towards a Modular theory of Linguistic Deficiency: Evidence from Malagasy Personal Pronouns. Natural Language & Linguistic Theory 17:161–218.

Author: Jason Ostrove
Affiliation: University of California, Santa Cruz, Linguistics Department
Email: jostrove@ucsc.edu

My thanks first and foremost to the San Martín Peras Mixtec speakers who shared with me their time, expertise, and friendship: Natalia Gracida-Cruz, Roselia Duran Cruz, Rocelia Cruz, and Gabriel Mendoza. Any mistakes are due solely to my own misunderstanding of what you taught me. This work began as a final paper for a class with Ivy Sichel as a graduate student in 2017, and I thank her for her early support and continued interest. Nascent aspects of this project also appeared in my UC Santa Cruz doctoral dissertation, chaired by Jim McCloskey with Sandy Chung, Ryan Bennett, and Ruth Kramer. My thanks to each of them for their mentorship and guidance. My thanks as well to Matt Wagers for his permission to use UC Santa Cruz affiliation for this work. I also had the privilege of presenting this project at various stages at WSCLA 22, Encuentro de Lingüística Formal (ELF) III, and SULA 12. My thanks to the audiences of each of those conferences for their thoughtful and encouraging feedback. Lastly, I would like to thank the two anonymous LI reviewers whose thorough, firm, and supportive comments made this work significantly better.

Notes

[1]All SMPM data come from the author's own fieldwork with four speakers carried out from 2014 to the present. SMPM is a largely unwritten language. Letter values are the same as Spanish, though *b* = [β], *x* = [ʃ], and *'* = [ʔ]. Tone is indicated through accents, and any mistakes in tonal transcription are entirely my own.

Abbreviations: ACC = accusative, ADV = untranslated adverb, AML = animal gender, AN = animate, APPL = applicative C = complementizer, COMP = completive aspect, CONT = continuous aspect, COP = copula, DAT = dative, DECL = declarative marker, EMPH = emphatic clitic, EXCL = exclusive, FEM = feminine, FUT = future, HAB = habitual aspect, HUM = human, IF = informal, INCL = inclusive, INF = infinitive, IRR = irrealis, NEG = negation, NOM = nominative, PERF = perfective aspect, POSS = alienable possession, POT = potential aspect, PL = plural, PRES = present, PROG = progressive aspect, Q = polar question marker, REDUP = reduplication, SG = singular, TNS = untranslated tense marker, WOOD = wooden gender, *(X) = ungrammatical if X is absent, (* X) = ungrammatical if X is present

[2]Aspectual morphology is primarily tonal, with completive aspect marked by a low tone on the first vowel or a prefix *nì-*, the continuous by a high tone on the first vowel, and irrealis by a mid/unmarked tone or stem changes for certain verbs. See Peters 2017 and Eischens and Hedding m.s.. In the discussions of finiteness and control, I use such stem-changing verbs wherever possible to avoid confounds of tonal mistranscription.

[3]See Holmberg 2005, 2010,b, Holmberg and Sheehan 2010, and Wurmbrand 2017 for such partial null subject languages. See also Ostrove 2022 for impersonal pronouns in this language.

[4]Subtypes are, of course, attested as well, such as the distinction between presuppositional/factive and non-presuppositional/non-factive finite embedded clauses (e.g., Kiparsky and Kiparsky 1970, et sequitur). For our purposes these more fine-grained distinctions are not relevant.

[5]*Káchi* 'said' in (12a) is a defective verb that occurs only in the completive aspect.

[6]This strategy seems to occur throughout Mixtec languages. See Macaulay 1987:103, 116-122.

[7]Unsurprisingly considering that *ná* does not occur with coreferent embedded subjects, c.f. (18a-b), *ná* cannot occur in untensed subjunctives with coreferent subjects either. This is provided in (108).

(108) Ntùkú Maria$_i$ [(* ná) ku'un =ñá$_i$, *tienda*].
try:COMP Maria NÁ go:IRR =she store
Intended: 'Maria$_i$ tried to go to the store.'

[8]The one exception that I am aware of to this generalization is *kòni* 'want,' which shows restructuring effects in allowing quantifier fronting out of tensed subjunctives. This is shown in (109). Note that 'want' is frequently exceptional cross-linguistically. See Grano 2012:137-181.

(109) Iin *dragón$_i$* kôni =rà tsyàja [ka'ani =rà __$_i$].
one dragon want:CONT =he man kill:IRR =he
'The man wants to kill a dragon.'

[9]Recall from section 2 that quantifier fronting targets a position below Spec,CP, the site of wh-movement, but to the left of the finite verb. If this position is identified as Spec,TP, then this split between tensed and untensed subjunctives may indicate that tensed subjunctives possess a TP layer or an additional modal layer such as *woll* of Wurmbrand 2014, while untensed subjunctives have such a layer. See Grano 2012 and Wurmbrand 2015.

[10]An anonymous reviewer observes that the predicates in (27) are expected from a semantic perspective, though a full investigation of these correspondences would take us too far afield. See Givón 1980, Ramchand and Svenonius 2014, Wurmbrand 2015, Wurmbrand and Lohninger 2019 and Wurmbrand et al. 2020.

[11]Another traditional property of OC not included in (28) is obligatory *de se* ascriptions in attitude contexts (e.g., Chierchia 1990, inter alia). I do not discuss *de se* ascriptions for two reasons. First, Landau 2013 convincingly argues that obligatory *de se* ascriptions is not criterial for OC (Landau 2013:32). Second, none of the untensed subjunctive selecting predicates are attitude predicates, with the possible exception of *nantǒso*

'forget to,' which Pearson 2013 classifies as a "non-canonical attitude predicate." Whether such predicates truly are attitude predicates is contested, though see Landau 2015:19,89 for arguments that they are not. Therefore, we likely cannot test the availability of *de re* ascriptions with untensed subjunctives in SMPM at all.

[12]An anonymous reviewer pointed out that the formation of reflexive anaphors with a definite article is typologically rare. Reflexives formed with cognates of *mí* are well-attested in Mixtecan, though, but it is likely that the semantics of these cognates are not identical. See, for instance, Copala Triqui (Hollenbach 1984), Silacayoapan Mixtec (North 1987:104), Chalcatongo Mixtec (Macaulay 2005), Xochapa Mixtec (Cline 2018:23), Cuevas Mixtec (Cisneros 2019:58), and San Sebastian del Monte Mixtec (Mantenuto 2020:74). A full investigation of the semantics of *mí* in SMPM is beyond the scope of this work.

[13]Note that *sǎ* 'so' cannot be used as a pronominal anaphor like 'it' in 'John wanted to win the race, and Mary wanted it too.' In attempting to elicit such constructions, speakers tend to simply provide (110a), without the pronominal anaphor. When the ambiguity in Spanish is pointed out to speakers (c.f. Chierchia 1984, Wurmbrand 2002), they respond with (110b), using the *ná* strategy discussed above.

(110) Provided for Spanish: 'Juan quiere ganar la carrera. Maria también lo quiere.'

a. Kôni Juân [keba'a =rà *carrera*]. Kôni =ti Maria [keba'a =ñá =ñà].
want:CONT Juan win:IRR =he race want:CONT =too Maria win:IRR =she =it
'Juan wants to win the race. Maria also wants to win the race.'

b. Kôni Juân [keba'a =rà *carrera*]. Kôni =ti Maria [ná keba'a =rà =ñà].
want:CONT Juan win:IRR =he race want:CONT =too Maria NÁ win:IRR =he =it
'Juan wants to win the race. Maria also wants him to win it.'

[14]Partial control in this literature refers to a subset to containing superset binding relation. The inverse, "sub-group" control which would involve a superset to subset binding relationship, as in (34a), never seems to occur in obligatory control. See Doliana and Sundaresan 2022:60 for recent discussion.

[15]Doliana and Sundaresan 2022 report a further type of nonexhaustive control they term proxy control, which only seems to occur under predicates involving illocutionary and permissive modality (Doliana and Sundaresan 2022:48). As was shown in (27), no such predicates, like 'ask,' select for subjunctive clauses in SMPM. This means that, unfortunately, proxy control cannot be tested in this language.

[16]Sevdali and Sheehan 2021 do identify what seem to be true instances of partial control in Greek and Romanian, though these seem to only involve translations of the ability modal 'can' in contexts such as 'I can meet tomorrow.' Unfortunately, SMPM's translation of this modal, *kuu*, does not straightforwardly select an untensed subjunctive, making such examples not possible to replicate in this language. Several examples of this modal are provided in (111), though the syntax of these constructions is left to future work.

(111) a. Kuu kuntsye'ě Juân *tele* ndìbi.
can watch:IRR Juan television daytime
'Juan can watch TV during the day.'

b. Ntákʻu kwé'e David. Kuu ntaani'i =rà iin *ciento* *kilogramo*!
strong very David can lift:IRR =he one hundred kilogram
'David is very strong. He can lift one hundred kilograms!'

[17]This claim is well established for exhaustive obligatory control PRO and reflexives, but more controversial for bound variable pronouns. See Barker 2012 for a view against the c-command requirement, and Déchaine and Wiltschko 2017 for a defense. Note, though, that Barker 2012 recognizes other languages behave differently, citing the inability of possessors in Chamorro to bind pronouns (Chung 1998) in contrast with English (Higginbotham 1980, inter alia). See as well (48) below.

[18]See Satık 2022 for a more thorough theory of the connection between finiteness/clause size and the morphophonology of PRO.

[19]This tentative universal is one-way, of course, as there are non-*pro*-drop languages without overt PRO,

e.g., English. If this universal does prove to be robust, though, it may be the case that it provides evidence for PRO and *pro* being exponents of the same lexical item. See McFadden and Sundaresan 2018.

[20]Landau 2015 suggests that overt anaphora in subject position of control clauses, as reported for Korean, Japanese, and Mandarin, also involve a focus interpretation and would therefore constitute an example of scope-sensitive pronominal copy control. See Madigan 2008 for Korean.

[21]Note that foci modified by *íntàà* 'only' must move to the left periphery. See Hedding 2022.

[22]Two more possible languages of this type are reported by Stiebels 2007: Jakaltek (Mayan) and Mangap-Mbula (Oceanic). It is unclear, though, that the relevant constructions in these languages display the diagnostic properties of obligatory control.

[23]See as well Satık 2020 for Ewe. Overt PRO in Ewe is a logophoric pronoun, making it difficult to distinguish from logophoric pronominal copy control. Satık argues convincingly, though, that these pronouns in Ewe are not inherently logophoric in OC constructions. I do not replicate his arguments for reasons of space, but Ewe is likely also another case of obligatory pronominal copy control.

[24]Speculatively, (78) may provide a hint as to the structure of exempt anaphors in SMPM. As mentioned above, exempt anaphors are composed of a definite article *mí* and a pronoun. This is reminiscent of 'strong' demonstrative pronouns as analyzed by Patel-Grosz and Grosz 2017, which also cannot be anteceded by quantified nominals. See Patel-Grosz 2020:553 and the citations therein. The details of such an analysis, in particular the character of the definite article *mí* as either a strong or weak definite in the sense of Schwarz 2009, remains a question for future research.

[25]This assumes a copy theory of movement (Chomsky 1995, Boeckx et al. 2009) and that movement chain reduction is a post-syntactic process. See Kandybowicz 2006, Harizanov 2014 and the citations therein.

[26]An anonymous reviewer asked if overt subjects of untensed subjunctives in this language could be analyzed in terms of Wurmbrand and Shimamura 2017. In this system, controlled subjects are suppressed via AGREE between matrix *v*, which enters into an AGREE relation with the controller, and the embedded *v* of the control clause. If we were to adapt this system to SMPM, we could analyze overt subjects of untensed subjunctives in SMPM as agreement morphemes on embedded *v*. These data suggest, though, that such an analysis is not tenable for SMPM either. This is because agreement morphemes cannot antecede pronouns or affect binding relations (den Dikken 1995, Rezac 2010, Kramer 2014). As such, this analysis would also incorrectly predict that exempt anaphors would not be available in untensed subjunctives, as an agreement morpheme would not be able to antecede the exempt anaphor.

[27]It is unclear whether Quiegolani Zapotec distinguishes between tensed and untensed subjunctives like SMPM (see Section 4). Therefore, (89b) may not truly involve control. That said, predicates like *runt gan* 'not know how to' that select untensed subjunctives cross-linguistically show the same pattern, as in 112.

(112) R- un -t noo$_i$ gan y- tsaa <u>noo$_j$</u> leter.
HAB- do -NEG 1EX able POT- write 1EX letter
'I$_i$ didn't know how PRO$_i$ to write a letter.' (Black 1994:271)

[28]Chamorro (Chung 1989, Chung 2019, Wagers et al. To appear), Khanty (Volkova and Reuland 2014), Jambi Malay (Cole et al. 2015), and Vietnamese (Bui 2019) also seem to follow this pattern.

[29]Pronouns like *li* in (90) are generally ambiguous between a bound and referential reading. As Déchaine and Manfredi 1994 report, though, some predicates like *tchwe* 'kill' in (90a) disallow coreferential readings. As we are only interested in the bound reading, the referential interpretation is not marked in (90b).

[30]The emergent typology in (92) leaves open possibilities. For example, we might expect a language in which all bound variable anaphora are syncretic to the exclusion of referential pronouns, or a language in which bound variable pronouns are different from reflexives and controlled subjects. I am unaware of attested examples of the first type, though San Lucas Quievaní Zapotec may instantiate the latter (Lee 2003). Also unattested, to my knowledge, is a language in which PRO is syncretic with a reflexive anaphor but not a pronoun, e.g. 'The boys knows how himself to ride a bike.'

[31]Exactly how this sort of allomorphy fits into the broader morphological literature is unclear. For example, licensing contexts for minimal pronoun allomorphs seem to be less local than what work on contextual

allomorphy posits. See Embick 2010, Bobaljik 2012, Merchant 2015, Caha 2018, and Ostrove 2018 for recent examples that, while different, employ a highly restrictive syntax for contextual allomorphy.

[32]Proposals differ in how this feature transmission occurs, e.g. Chomskyian AGREE (Chomsky 2000, 2001) in Kratzer 2009 or variable binding (Landau 2015). The exact mechanism is orthogonal to our concerns here.

[33]More specifically, the universal 'pronoun as Elsewhere' seems to apply to clitic pronouns, not non-clitic pronouns. See Zribi-Hertz and Mbolatianavalona 1999 and the discussion of these facts in Déchaine and Wiltschko 2017, as well as, perhaps, Cardinaletti and Starke 1999:156-159.

[34]That PRO_{Arb} ought to be connected to impersonal pronouns like English 'one' or German *man* has a long history. See Jaeggli 1986, Chomsky and Lasnik 1993:37, Hornstein 1999:91, and Boeckx and Hornstein 2004:441-442, inter alia. This overt impersonal provides support for this connection. See as well Ostrove 2022 for more on impersonal readings of pronouns in SMPM specifically.

[35]Only predicates that select untensed subjunctives, either in SMPM or cross-linguistically, were selected for (105), though it is unclear if these languages make this distinction. I also selected examples without pronominal antecedents where available to avoid confounds raised by Broadwell 2003, who observed that copy control in San Dionicio Ocotepec Zapotec occurs only with pronominal antecedents. Also excluded were languages in which third person clitic pronouns are null (c.f., Tecóatl Mazatec, García Baltazar 2019:41 and Chalcatongo Mixtec, Macaulay 1993:74), as it is impossible to distinguish between PRO and *pro* with the data available. See, though, McFadden and Sundaresan 2018. Comparable constructions may also occur in Otomi and Mazahua, shown in (113), though it is unclear if the relevant morphemes are pronominal as they exhibit tense variance (Nevins 2011).

(113) a. Dá$_i$= jwá =hé dá$_i$= n¨é =hé.
1.PAST= finishs =PL.EXCL 1.PAST= dance =PL.EXCL
'We (EXCL) finished dancing.' (San Ildefonso Tultepec Otomí, Palancar 2012:43)

b. Xo rí$_i$ pjëch =tjo rá$_i$ xörü.
also I.PRES know =ADV I.FUT read
'I also know how to read.' (Central Mazahua, Stewart 1966:34)