Two restrictions on pronominal clitics in San Juan Piñas Mixtec (Tò'ōn Ndā'ví)[*]

Michelle Yuan
University of California, San Diego

Abstract: San Juan Piñas Mixtec (Tò'ōn Ndā'ví) is an underdocumented Mixtec variety spoken primarily in San Juan Piñas, Santiago Juxtlahuaca, Oaxaca. This paper reports on research from an ongoing collaborative project documenting, analyzing, and developing linguistic resources for the language. I investigate two restrictions on the pronominal clitics of SJPM and show that they have distinct grammatical sources. The first restriction is a ban on 1st/2nd person object clitics in the presence of a subject (of any person, clitic or full DP)—an extension of the strong Person Case Constraint (PCC) (see also Sichel and Toosarvandani to appear). I show that contemporary syntactic accounts of the PCC (e.g. Deal to appear) may be straightforwardly extended to the SJPM PCC. The second restriction is a ban on certain 3rd person object clitics, which I argue is morphophonological in nature: it only arises with vocalic clitics, is sensitive to linear adjacency, and may be prosodically-driven.

Keywords: Mixtec, syntax, morphology, pronouns, clitics, Person-Case Constraint

1 Introduction

Cross-linguistically, pronominal elements are often variably realized as independent tonic pronouns or as prosodically dependent pronominal clitics. Both options are illustrated in (1), from the San Juan Piñas variety of Mixtec (henceforth, 'SJPM'):

(1) **ndu¹ʔu¹**   ʃi¹³ka³=**ndi³**
     PRON.1PL.EX PST.walk=CL.1PL.EX[1]
     'We walked.'

However, much prior work has shown that the distribution of pronominal clitics may be constrained compared to their tonic pronoun counterparts. This paper documents two restrictions on subject/object pronominal clitic combinations in SJPM, and provides preliminary analyses of both

[1] **Abbreviations:** ANIM = (non-human) animate, ARB = arboreal, CL = clitic, COP = copula, EMPH = emphatic, EX = exclusive, F = feminine, H = human, IN = inclusive, IRR = irrealis, LIQ = liquid, M = masculine, N = neuter, NEG = negation, PL = plural, PRES = present, PRON = pronoun, PST = past, RND = round, SG = singular, WH = wh-element, 1 = first person, 2 = second person, 3 = third person. **Note on transcription:** This paper uses the Chao system for tonal transcription, as well as a modified IPA for segmental transcription (for readability, prenasalization and labialization of stops are not superscripted (e.g. *ndV*, *kwV*). SJPM has three level tones (H = V^5, M = V^3, L = V^1), plus a floating L-tone, which may combine to form various rising and falling contours (e.g. LH = V^{15}, ML = V^{31}). Upstepped (e.g. V^6) and downstepped (e.g. V^4) tones are also attested in restricted environments (Caballero, Duarte Bórquez, and Juárez Chávez to appeara).

restrictions. I demonstrate that these restrictions have distinct grammatical profiles, and conclude that they hold at different components of the grammar.

The first restriction concerns a constraint on 1st/2nd person (participant) object clitics, (2-a); the second is a constraint on certain 3rd person object clitics, (2-b).

(2)   a.   *ko⁵ni¹=ɲa⁵=ndi³
              PRES.love=CL.3SG.F=CL.1PL.EX
              'She loves us (ex.).'

        b.   *ʃi⁵ni⁶=ndi¹=a⁵
              PRES.see=CL.1PL.EX=CL.3SG.F
              'We (ex.) see her.'

I propose that the participant clitic restriction is *morphosyntactic* in nature. Concretely, it is a manifestation of the Person-Case Constraint, and should be modeled in terms of syntactic Agree with goals bearing the feature [PART]. The 3rd person restriction, in contrast, is *morphophonological*, and as such is insensitive to hierarchical structure and syntactic features. Rather, I suggest that the latter specifically affects vowel-initial clitics due to their prosodic properties.

Overall, this paper makes three broad contributions. First, it provides a detailed look at the pronominal system of SJPM, which may, in turn, aid subsequent (micro-)comparative research across Mixtec and Oto-Manguean. Additionally, it offers an illustration of how phenomena that seem surface-similar may be distinguished using theoretically-motivated diagnostics (see also Foley, Kalivoda, and Toosarvandani 2019). Finally, as the PCC pattern in SJPM displays a number of properties that are cross-linguistically rare (though found elsewhere in Oto-Manguean; see Sichel and Toosarvandani to appear, in particular), the paper extends our existing typology of PCC effects (e.g. Anagnostopoulou 2017; Deal to appear; Nevins 2007).

The rest of the paper is organized as follows. In §§2-3, I present key grammatical properties of SJPM, focusing on the pronominal clitics in the language. In §4, I provide evidence that the restriction on participant object clitics is indeed a variety of the PCC, and offer a preliminary analysis, following Deal (to appear). Finally, §5 turns to the 3rd person restriction and shows that it is morphophonological in nature.

2   Overview of SJPM

Mixtec belongs to the Eastern branch of the Oto-Manguean language family and consists of around 60 language varieties (Campbell 2017; Josserand 1983). It is spoken primarily in Oaxaca, Mexico and neighbouring states. The variety investigated in this paper is known as San Juan Piñas Mixtec or as Tò'ōn Ndā'ví (its endonym), though henceforth SJPM here. SJPM is spoken in the town of San Juan Piñas (pop. ~900), part of the larger municipality of Santiago Juxtlahuaca, Oaxaca, and is also spoken in diaspora communities in California and beyond.

This paper reports on morphosyntactic and morphophonological research from a larger, collaborative language project, undertaken in San Diego and in San Juan Piñas.[2] The project includes the documentation and analysis of SJPM's grammatical properties, as well as the development of

---

[2] The project began through field methods courses offered at UC San Diego, first taught by Claudia Juárez Chávez and Justin McIntosh in 2018, and then by Claudia Juárez Chávez, Gabriela Caballero, and Michelle Yuan in 2020. Other project members have included language teachers Basilisa Juárez Chávez, Basilisa Pérez Morales, Came Juárez Chávez, as well as students Claudia Duarte Bórquez, José Armando Fernández Guerrero, Colin Gazaui, Ray Huaute, Akil Iyer, Nico Tedeschi, Maxine van Doren, and Tiffany Wu.

linguistic resources for language reclamation purposes.[3]

The base word order of SJPM (as well as other Mixtec varieties and related Oto-Manguean languages) is V(erb)-S(ubject)-O(bject), as in (3-a). In SJPM, full nominals are often accompanied by a classifier or determiner (assumed here to be a D^0). In (3-b), we additionally see that pronouns often surface as *enclitics* that attach to a preceding host, and that they are also found in VSO order. Finally, note that there are no case distinctions; nouns in SJPM are morphologically invariant regardless of syntactic position or grammatical function.

(3) a. ni^1-ta^3ʔvi^5 **ɲa⁵**   **Ga⁵bi⁶¹ ɲa¹**   **ʒoo³**
         PST-break CL.3SG.F Gaby   CL.3.N water.jug
         'Gaby broke the water jug.'   (V S O)

     b. ni^1-ta^3ʔvi^5=**ɲa⁵**=**ɲa³**
         PST-break=CL.3SG.F=CL.3.N
         'She broke it.'   (V=S$_{Cl}$=O$_{Cl}$)

Following Yuan (2022) and Hedding and Yuan (to appear), VSO word order in SJPM is derived by movement of the object out of the VP, followed by (remnant) VP-fronting.[4] For concreteness, the structure in (4) situates the landing site of the object and VP in Spec-*v*P and Spec-TP, respectively, and the base position of the subject in Spec-VoiceP, though nothing in this paper crucially hinges on these assumptions. The VP-fronting analysis is relevant because pronominal clitics attach to the *linearly rightmost* element of the fronted VP, which may not necessarily be the verb. Many examples of this sort will be given throughout this paper. Moreover, it will be useful in §4.2, when we develop our analysis of the PCC in SJPM.

(4)
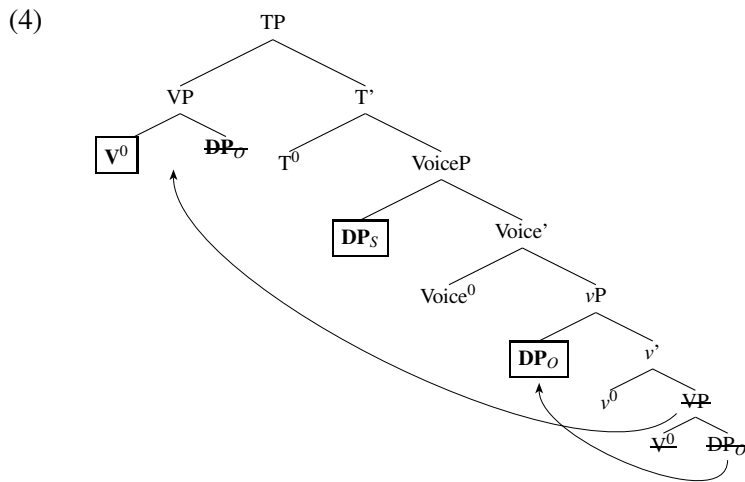
Finally, SVO is also attested in SJPM and, in such contexts, the preverbal subject often co-occurs with a postverbal pronominal clitic, as in (5-a). Notably, however, extracted negative indefinites such as neg-words and clefted wh-words must leave a gap, (5-b-c). I assume that these SVO con-

---

[3] See, for instance: Caballero, Juárez Chávez, and Yuan (to appearb); Duarte Bórquez, Caballero, Juárez Chávez, and Yuan (2023); Duarte Bórquez and Juárez Chávez (2022); Juárez Chávez, Caballero, Duarte Bórquez, Guerrero, Huaute, Iyer, Simmons, Tedeschi, Doren, and Yuan (2022); Tedeschi (2023).

[4] See also Massam (2001), Lee (2006), and others.

structures are derived from the base VSO structure in (3), with the subject undergoing Ā-movement to Spec-CP. The pattern in (5-a) is the product of clitic-doubling (alternatively, the doubled clitic may be viewed as a resumptive pronoun, since it is the tail of the movement chain[5]). This also captures the absence of the clitic in (5-b-c), given that non-specific indefinites, including neg-words and (non-D-linked) wh-words, cannot be clitic-doubled cross-linguistically (e.g. Baker and Kramer 2018; Suñer 1988; Yuan 2021).

(5)   a.   **ɲa⁵**   **Ga⁵bi⁶¹** ni¹-ta³ʔvi⁵=$\boxed{\textbf{ɲa}^5}$ ɲa¹   ʒoo³
              CL.3SG.F Gaby PST-break=CL.3SG.F CL.3.N water.jug
              'Gaby broke the water jug.'   (S V=Cl$_S$ O)

    b.   **ni³** **ĩ³=na¹**   ta¹⁵ʔvi⁵ $\boxed{}$ ki¹si³
              NEG one=CL.3PL.N NEG.IRR.break   pot
              'No one will break the pot.'   (S V(*=Cl$_S$) O)

    c.   **ndʒa⁵ ku⁵=na¹**   ʃi¹³ni³¹ $\boxed{}$ ɲa⁵   Pa⁵lo⁵ma⁶?
              WH   PRES.COP=CL.3PL.N PST.see   CL.3SG.F Paloma
              'Who (is it that) saw Paloma?'   (S V(*=Cl$_S$) O)

To summarize, I have shown that full DPs, tonic pronouns, and pronominal clitics (including doubled-clitics) all surface in base VSO order.

3   Properties of pronominal clitics in SJPM

Clitics in SJPM are monomoraic (e.g., (C)V), in contrast to the bimoraic (e.g., (C)V(ʔ)(C)V) minimal word requirement in the language (e.g. Carroll 2015; Pike 1944; Uchihara and Mendoza Ruiz 2022). Like other clitics cross-linguistically, they cannot be fronted to Ā-position or surface in coordination and fragment constructions (Cardinaletti and Starke 1999); their tonic pronoun counterparts must be used instead, (6).

(6)   a. $\boxed{\textbf{ʒoo}^5}$ ʃi⁵tʃ=**i⁵**
              PRON.1PL.IN PRES.swim=CL.1PL.IN
              'We (in.) are swimming.'

    b. $\boxed{\textbf{ʒo}^5\textbf{ʔo}^1}$ a⁵ $\boxed{\textbf{ʒu}^1\textbf{ʔu}^1}$ c.   sa³⁵tu³ $\boxed{\textbf{ʒu}^1\textbf{ʔu}^1}$=va³
              PRON.2SG or PRON.1SG   also   PRON.1SG=EMPH
              'you or I'   'Me too!'

They are also relatively unpicky about their hosts, attaching to whatever element immediately precedes them. In (7-a), the subject clitic attaches to a right-adjoined VP-internal adverb; in (7-b), the object clitic attaches to a subject DP.

(7)   a.   [$_{VP}$ tu⁵tu⁵   ʒu³ʔu⁵ tʃe⁵ʔe⁵ **ʒaa¹**]=$\boxed{\textbf{ɲa}^5}$
                 PRES.whistle mouth loud   still =CL.3SG.F

              'She is still whistling loudly.'   (V N Adv Adv=$\boxed{\text{Cl}_S}$)

---

[5] This distinction is not important here, especially given accounts that unify the two (e.g. Boeckx 2003; Harizanov 2014a). See also McCloskey (2017) for discussion.

b. ko³ni³¹ [$_{DP}$ ti⁵ **kwa⁵ʒu¹**]=ɲa¹
   IRR.see CL.3.ANIM horse =CL.3SG.N
   'The horse will see it (e.g. the baby).' (V S=Cl$_O$)

We have already seen that 1st/2nd person pronouns may be realized as tonic pronouns or as pronominal clitics. A full paradigm is given in (8). In contrast, there are no tonic pronouns in the 3rd person, as in (9); these forms are all monomoraic enclitics.[6] The 3rd person clitics make several noun class distinctions, and additionally encode number and gender distinctions for human clitics.[7]

(8)

	1SG	**2SG**	**1PL.IN**	**1PL.EX**	**2PL**
Pronoun	ʒu¹ʔu¹	ʒo⁵ʔo¹	ʒoo⁵	ndu¹ʔu¹	ndo⁵ʔo¹
Clitic	=e¹	=õ⁵	=e⁵	=ndi	=ndo⁵

(9)

3SG.M	**3SG.F**	**3.N**	**3PL.M**	**3PL.F**	**3PL.N(H)**	**3.ANIM**	**3.RND**	**3.ARB**	**3.LIQ**
=ra	=ɲa⁵	=ɲa	=ndʒa	=na⁵	=na	=ti⁵	=ti⁵	=tõ⁵	=ra⁵

Note that some clitics in both paradigms are tonally underspecified. Following Duarte Bórquez and Juárez Chávez (2022) and Caballero et al. (to appearb), they are phonetically [M] by default but can be targeted by a word-internal rightward [L]-tone spreading process. This is shown in (10). This property also serves to illustrate the prosodic integration of the clitics with their hosts.

(10) a. tu³tu⁵=**ndi³**   b. ko³ni¹=**ndi¹**
        IRR.whistle=CL.1PL.EX   IRR.love=CL.1PL.EX
        'We'll whistle.'   'We'll love.'

The vocalic (=V) clitics (e.g., 1SG, 2SG, 1PL.IN in (8)) are even more prosodically integrated, coalescing both segmentally and tonally with their hosts, (11):

(11) a. tu³t=**i⁵¹**   b. ka³k=**i³¹**=ɲa¹
        IRR.whistle=CL.1SG   IRR.put=CL.1SG=CL.3.N
        'I'll whistle.'   (cf. tu³tu⁵ 'whistle')   'I'll put it.'   (cf. ka³ki³ 'put')

Prior cross-linguistic work has shown that clitics may attach at different prosodic levels, even within a single language (Harizanov 2014b; Ito and Mester 2012; Selkirk 1995); this seems to be the case in SJPM, as well. Since =V clitics coalesce with their hosts and surface within the bimoraic foot, they presumably adjoin internal to the minimal prosodic word (ω), as in (12-a). Conversely, =CV clitics presumably adjoin outside of the minimal prosodic word, yielding recursive prosodic words, (12-b-c). We will return to this distinction in §5.

---

[6] In syntactic environments that disallow (bare) clitics, the pronominal meanings are often instead expressed using a classifier + deictic, e.g. ɲa⁵ ʒo⁵ʔo³ 'she (here)', ɲa⁵ ka⁵a³ 'she (there)'.

[7] The noun classes differ across Mixtec varieties. See e.g. De León Pasquel (1988); Macaulay (1987), and Cisneros (2019) for discussion of certain varieties. In SJPM, the noun classes distinguish {M/F/N} humans, animals (ANIM), wood-based elements (ARB), round elements (RND), liquids (LIQ), and miscellaneous inanimate elements (N). Only human forms have dedicated SG and PL variants. There is also variation across Mixtec concerning how the classifiers are used. I treat them as D⁰s in SJPM, though nothing hinges on this. Finally, the 3rd person clitics are identical to the classifiers/determiners in the language (surface tonal differences, when they appear, are phonologically predictable). I will set this fact aside in this paper, however.

(12) a. $[(\text{tu}^3\text{t}=\textbf{i}^{\textbf{51}}{}_{\text{FT}})]_\omega$ b. $[[(\text{ka}^3\text{ka}^3)_{\text{FT}}]_\omega=\textbf{ɲa}^{\textbf{5}}]_{\omega2}$ c. $[[[(\text{ko}^5\text{ni}^1)_{\text{FT}}]_\omega=\textbf{ra}^{\textbf{1}}]_{\omega2}=\textbf{ɲa}^{\textbf{5}}]_{\omega3}$

   IRR.whistle=CL.3SG.F   IRR.walk=CL.3SG.F   PRES.love=CL.3SG.M=CL.3SG.F

   'She will whistle.'   'She will walk.'   'He loves her.'

With these properties in place, the next two sections turn to the two restrictions mentioned in §1.

4 A morphosyntactic restriction: The PCC in SJPM

4.1 The basic PCC pattern

The Person-Case Constraint (PCC) is typically defined as a family of restrictions on combinations of phonologically weak arguments, usually pronominal clitics (Perlmutter 1971, et seq.). Although the PCC is most commonly discussed in the context of combinations of IO and DO arguments, it has also been shown to hold among phonologically weak *subject/object* combinations—as is the case in SJPM. The PCC in SJPM is a ban of 1st/2nd person (participant) object clitics, regardless of the choice of subject. Notably, this is reminiscent of the *Strong PCC*, summarized in (13):

(13) *The Strong PCC:* In combinations of two clitics, Cl_1 and Cl_2, where Cl_1 is structurally higher than Cl_2, Cl_2 must be 3rd person.

Cl_1	2	1	3	3	2	1	3
Cl_2	1	2	1	2	3	3	3
	✗	✗	✗	✗	✓	✓	✓

In SJPM, this amounts to a ban on participant (but not 3rd) person clitics in object position. This is shown in (14)-(15). For consistency, I primarily illustrate the PCC throughout this paper with 1PL.EX =*ndi*, though the restriction holds for all participant clitics and there is no internal ranking between 1st and 2nd person.[8]

(14) a. $*\text{ʃi}^{13}\text{ni}^{31}=\textbf{ndo}^{\textbf{5}}=\boxed{\textbf{ndi}^{\textbf{3}}}$ b. $*\text{ko}^5\text{ni}^1=\textbf{ɲa}^{\textbf{5}}=\boxed{\textbf{ndi}^{\textbf{3}}}$

   PST.see=CL.2PL=CL.1PL.EX   PRES.love=CL.3SG.F=CL.1PL.EX

   'You (pl.) saw us (ex.).'   'She loves us (ex.).'

   $(✓\text{ʃi}^{13}\text{ni}^{31}=ndo^5\ ndu^1\ ʔu^1)$ $(✓\text{ko}^5\text{ni}^1=ɲa^5\ ndu^1\ ʔu^1)$

(15) $\text{ko}^5\text{ni}^1=\textbf{ndi}^{\textbf{1}}=\boxed{\textbf{na}^{\textbf{1}}}$ / $=\textbf{ra}^{\textbf{1}}=\boxed{\textbf{na}^{\textbf{1}}}$

   PRES.love=CL.1PL.EX=CL.3PL.N =CL.3SG.M=CL.3PL.N

   'We (ex.) love them.' / 'He loves them.'

That being said, there is a twist to the PCC pattern in SJPM: participant object clitics are still banned when the subject is a *full DP*. In other words, it is not the case that the PCC holds only when both elements are phonologically weak (contra, e.g., Bonet 1991). Although cross-linguistically unusual, this property is also featured elsewhere in the Oto-Manguean language family, such as in Southeastern Sierra Zapotec (Sichel and Toosarvandani to appear).[9] In addition, a similar pattern is

---

[8] In §5, I also discuss =V participant clitics, which require further elaboration.

[9] This property also possibly holds in other varieties of Mixtec, based on cursory mentions by Macaulay (1996:140-141) for Chalcatongo Mixtec and Mantenuto (2020:65-72) for San Sebastián del Monte Mixtec. Note also that the exact effect found in Zapotec differs from that seen in SJPM; in Zapotec, 3rd person object clitics are also banned in the presence of a full DP subject (Sichel and Toosarvandani to appear).

found in French causative constructions, as noted by Sheehan (2020) and Deal (to appear).

(16) a. *ko⁵ni¹   [ti⁵   vi³lu⁵]= **ndi³**
      PRES.love   CL.3.ANIM cat   =CL.1PL.EX
      'The cat loves us (ex.).'   (✓*ko⁵ni¹ ti⁵ vi³lu⁵ ndu¹ ʔu¹*)

     b. ko⁵ni¹   [ti⁵   vi³lu⁵]= **ra³**
      PRES.love   CL.3.ANIM cat   =CL.3SG.M
      'The cat loves him.'

Finally, the examples in (17) demonstrate that the PCC still holds even when the subject is Ā-extracted and does not leave a postverbal pronominal clitic (this is again also seen in Zapotec). Importantly, these data point towards a syntactic analysis of this restriction. Unpronounced movement copies (that are presumably deleted postsyntactically) are treated the same way as overt DPs. This is, conversely, unexpected under a purely morphological account of the PCC (e.g. Bonet 1991), since the postverbal subject is no longer overtly present.

(17) a. ***ni³** **ñ̃³=na¹**   ko¹⁵-ʃi³ni³¹= **ndi¹**
      NEG one=CL.3PL.N NEG.PST-see=CL.1PL.EX
      'No one saw us (ex.).'   (✓*ni³ ñ̃³=na¹ ko¹⁵-ʃi³ni³¹ ndu¹ ʔu¹*)
      → *****ni³ ñ̃³=na¹** ko¹⁵-ʃi³ni³¹ ~~ni³ ñ̃³=na¹~~ =**ndi**

     b. **ni³** **ñ̃³=na¹**   ko¹⁵-ʃi³ni³¹= **na¹**
      NEG one=CL.3PL.N NEG.PST-see=CL.3PL.N
      'No one saw them.'
      → ✓**ni³ ñ̃³=na¹** ko¹⁵-ʃi³ni³¹ ~~ni³ ñ̃³=na¹~~ =**na**

In sum, SJPM displays a version of the Strong PCC, with the addition that full DP subjects (whether overt or null) also induce the effect.

4.2 A syntactic analysis

There are many different accounts of the PCC, though one common ingredient is that it involves a single syntactic head Agreeing with—or attempting to Agree with—multiple clitics. I now apply this logic to the PCC in SJPM. I illustrate with one recent account to the PCC, developed by Deal (to appear) as part of a larger theory of Agree (Deal 2015, et seq.). That being said, the objective of this paper is *not* necessarily to adjudicate between existing analyses;[10] rather, I simply show here that the pattern found in SJPM fits cleanly within the typology of PCC patterns cross-linguistically and may thus be straightforwardly analyzed using existing syntactic approaches.

     In Deal's conception of Agree, Agreeing probes are specified with *interaction* and *satisfaction* conditions, rather than unvalued or uninterpretable features. These conditions determine, respec-

---

[10] See, for instance, Anagnostopoulou (2003); Béjar and Rezac (2003); Coon and Keine (2021); Foley and Toosarvandani (2022); Nevins (2007); Pancheva and Zubizarreta (2018); Sichel and Toosarvandani (to appear); Stegovec (2020) for different analyses of the PCC. These analyses vary in the exact mechanics of the Agree operation, the placement of the agreeing probe, and whether both clitics are successfully agreed with.

tively, which features are copied back to the probe and which features halt the probing process.

Deal (to appear) proposes that the PCC arises from the following interrelated premises. First, pronominal clitics are generated by Agree (e.g. Anagnostopoulou 2003), so the presence of a pronominal clitic entails the occurrence of successful probing. Cliticization of multiple arguments may be derived by a single head that Agrees with both arguments. Second, the probe interacts with the *lower argument* before interacting with the higher one (deemed the *direct object preference* hypothesis by Deal). The PCC arises if the features of the lower argument satisfy the probe, preventing it from interacting with the higher argument.

Putting these together, consider a basic schema like (18), meant to capture a typical strong PCC pattern in ditransitive constructions. In this structure, $Appl^0$ is specified for [INT:ϕ,SAT:PART]. This means that it may copy any ϕ-feature, but probing will be halted once it finds a goal bearing [PART]. Moreover, the probe's object preference is derived via *cyclic expansion* (Béjar and Rezac 2009): $Appl^0$ first searches within its c-command domain, and then the constituent Appl' searches within its own c-command domain, which includes its specifier.

(18)
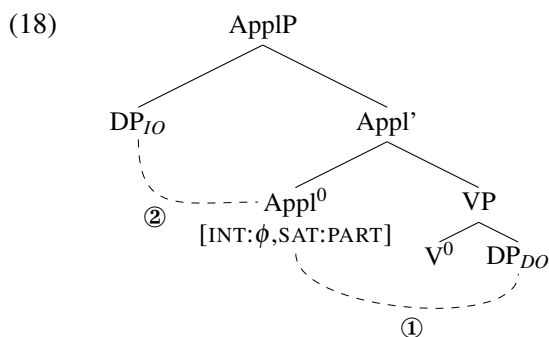
In a well-formed configuration (e.g., PART/3 or 3/3), assuming that 3rd person is featurally simply [ϕ], the probe may interact with both the direct object and the indirect object in turn. However, in a PCC-violating configuration (e.g., 3/PART or PART/PART), $Appl^0$ is satisfied by the [PART] direct object, so it does not continue probing for the indirect object.

Let us now extend this approach to the strong PCC in SJPM. The strong PCC arises in generally the same way as described above. However, because full DP subjects also participate in the restriction, I follow Sichel and Toosarvandani (to appear) in taking all nominals to bear an additional feature [δ], resulting in the feature geometries in (19):

(19) a. Full DP:
δ

b. 3rd person pron.:
δ
|
ϕ

c. Participant pron.:
δ
|
ϕ
|
PART
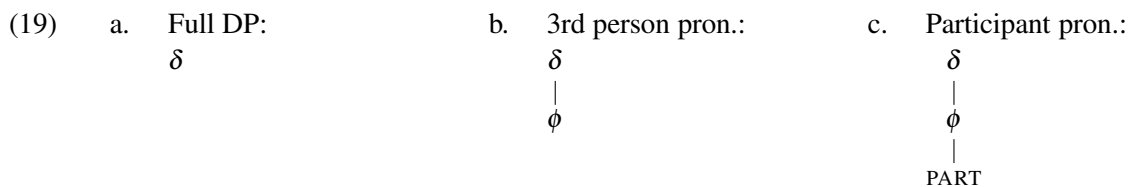
As shown in (20), since the PCC affects subject/object combinations in SJPM, we may situate the probe in $Voice^0$, based on (4) from §2. Recall that the subject is in Spec-VoiceP and the object is in Spec-vP after object shift. Moreover, the probe in $Voice^0$ is specified for [INT:δ,SAT:PART], which allows it to interact with full DPs and pronouns alike. Once again, in configurations not containing a [PART] object, the probe is able to Agree with both arguments, generating both clitics. If the object

bears [PART], however, the probe is satisfied upon finding it and does not probe further.

(20)

VoiceP
- DP$_S$
- Voice'
  - ② ----- Voice^0 [INT:δ,SAT:PART]
  - vP
    - DP$_{DO}$
    - v'
      - v^0 …
  - ①
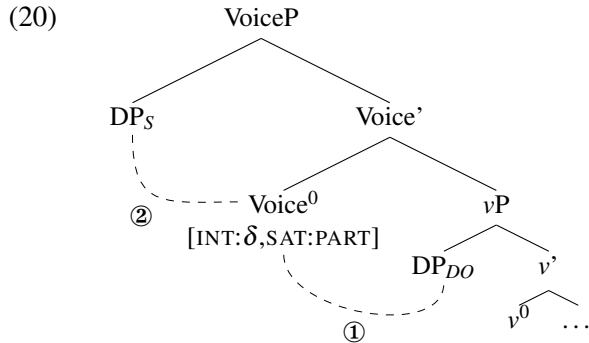
One possible complication of this particular analysis concerns the PCC 'repair' in SJPM—namely, the use of a tonic pronoun in object position, as in (21). Following Deal (to appear), I assume that such constructions are not derivationally related to the PCC-violating constructions. Deal raises one possibility: perhaps tonic pronouns in SJPM are properly contained within additional structure, such that they featurally behave like DPs. Within the present account, this would mean that only the feature [δ] is accessible to the probe. This type of treatment would allow the probe to continue searching for potential goals after its initial encounter with the tonic pronominal object.

(21) ko⁵ni¹=ɲa⁵   ⌈ndu¹ʔu¹⌉
      PRES.love=CL.3SG.F PRON.1PL.EX
      'She loves us (ex.).'

As potential support for this approach, I note that this hypothesized additional structure may be overtly present in other Mixtec varieties. In San Sebastián del Monte Mixtec (SSMM), for instance, [PART] object clitics must attach to what Mantenuto (2020) (descriptively) calls a pronominal base, creating a complex DP structure. This is shown in (22-a). Crucially, in (22-b) we see that the same strategy is also used for reflexive objects of *all persons* (even 3rd person, for which no PCC effect arises).[11] Enclosing a reflexive within a larger structure is notably a common strategy that many languages employ to prevent it from being targeted for ϕ-agreement (Rizzi 1990; Woolford 1999; Yuan 2023). It may be that the [PART] pronoun in (22-a) is similarly rendered invisible within this complex structure, thus allowing the probe in Voice^0 to interact with other goals.

(22) a. Kàni=rà   ⌈*(mee)=ó⌉   b. Kóni   Chuchi ⌈*(meé)=rà⌉
          PST.hit=CL.3SG.M BASE=CL.2SG   PRES.love Chuchi BASE=CL.3SG.M
          He hit you.'   'Chuchi loves himself.'
          (SSMM, Mantenuto 2020)

I leave a deeper assessment of this idea for future work. The broader point is that the PCC pattern in SJPM (and other Mixtec varieties) is syntactic in nature, and thus requires a syntactic analysis. I have provided an illustration of how the exact effect found in SJPM fits within existing theoretical approaches to the PCC.

---

[11] In SJPM, the cognate base *mi³i⁵* is also used for reflexive objects of all persons, suggesting a common analysis with SSMM. I have not checked if this base may be used in lieu of a tonic pronoun in PCC-violating configurations, but the account developed here predicts that it may.

5 A morphophonological restriction: A ban on =V object clitics

In contrast to participant clitics, the fact that 3rd person clitics are permitted in object position is because they lack the [PART] feature. Thus, any restrictions that they do exhibit should be *non-syntactic*. I now show that this is indeed borne out. Recall from §1 that there is a second restriction that appears to affect certain 3rd person object clitics. I argue that this restriction is grammatically distinguishable from the PCC, and is morphophonological in nature. I also observe that this restriction does not actually target object clitics, despite surface appearances. Rather, it is a constraint on *subject nominals*, which in most—though not all—contexts immediately precede objects.

In SJPM, the 3SG.F and 3.N clitics take two forms, which are seemingly in free variation in most contexts. They may either be realized with their =CV form (=$\textipa{ɲa}^5$ and =$\textipa{ɲa}$, respectively) or with a vocalic form (=$\tilde{a}$ and =$\tilde{a}$). These alternatives are illustrated in (23), an update of (9).

(23)

3SG.M	3SG.F	3.N	3PL.M	3PL.F	3PL.N(H)	3.ANIM	3.RND	3.ARB	3.LIQ
=ra	=$\textipa{ɲa}^5$	=$\textipa{ɲa}$	=ndʒa	=na^5	=na	=ti^5	=ti^5	=tõ^5	=ra^5
	=$\tilde{a}^5$	=$\tilde{a}$							

Like the participant =V pronominal clitics, shown in (11), the 3SG.F/3.N =V clitics coalesce segmentally and tonally with the immediately preceding element. In (24-a-b), these clitics are attached to a verb and a VP-internal adverb, respectively:

(24) a. nda^{3}ko^1=$\boxed{\tilde{a}^{15}}$
          IRR.wake.up=CL.3SG.F
          'She will wake up.'
          (also ✓: *nda^{3}koo^{13}=ɲa^5*)

    b. [nda^{13}koo^{13} nã^{5}ʔ]=$\boxed{\tilde{a}^{15}}$
          PST.wake.up early =CL.3SG.F
          'She woke up early.'
          (also ✓: *nã^{3}ʔã^{31}=ɲa^5*)

The examples in (25) show that these clitics may also attach to a non-pronominal clitic (illustrated here with the emphatic VP-internal clitic =*va^3*), and to a possessed nominal within a complex DP.[12]

(25) a. i^{5}ʒo^6 va^{1}ʔa^3=**v**=$\boxed{\mathbf{a^5}}$
          PRES.COP good=EMPH=CL.3SG.F
          'She's fine.'
          (also ✓: *va^1 ʔa^3=va^3=ɲa^5*)

    b. **vi^{3}ʎ**=$\boxed{\mathbf{a^5}}$
          cat=CL.3SG.F
          'her cat'
          (also ✓: *vi^{3}lu^5=ɲa^5*)

However, there is one environment in which the =CV and =V forms are not in free variation: *the vocalic forms may not cliticize to (i.e., coalesce with) subjects*. Importantly, this holds regardless of the choice of subject—for instance, whether it is a pronominal clitic, (26-a), or a full DP, (26-b). The example in (26-b) moreover contrasts neatly with (25-b) above, since the resulting morphological form (*vi^{3}ʎ=a^5*) should be identical in both cases.

(26) a. *ʃi^{13}ni^{31}=**ndo^5**=$\boxed{\mathbf{a^5}}$
          PST.see=CL.2PL=CL.3SG.F
          'You (pl.) saw her.'
          (✓ *ʃi^{13}ni^{31}=ndo^5=ɲa^5*)

    b. *ʃi^{13}ni^{31} **ti^5**   **vi^{3}ʎ**=$\boxed{\mathbf{a^5}}$
          PST.see CL.3.ANIM cat=CL.3SG.F
          'The cat saw her.'
          (✓ *ʃi^{13}ni^{31} ti^5 vi^{3}lu^5=ɲa^5*)

---

[12] Likewise, as an extension of fn. 11, the 3SG.F and 3SG.N =V clitics may also coalesce with the pronominal base *mi^{3}t^5*, producing the sequence *mi^3=a^5*.

In (27), I provide all the (im)possible hosts for 3SG.F/3.N =V elements that I have tested thus far.

(27)

	P^0	N^0(POSS)	V^0	Adv^0	Adj^0	VP-int. clitic	Cl_S	DP_S
=V?	✓	✓	✓	✓	✓	✓	✗	✗

At first blush, this looks like a restriction on 3rd person *object* clitics. In actuality, however, the restriction pertains to *subjects*—in that they are not able to serve as hosts for vocalic clitics. Crucially, there are certain environments in which such clitics *are* permitted in object position. Recall from §§3-4 that neg-word and wh-subjects obligatorily leave gaps when they are Ā-extracted, and that fronting such subjects does not alleviate the PCC. With these facts in mind, consider (28-a-b) below. These show that 3SG.F (and 3SG.N) =V object clitics are in fact licit; they simply cliticize to the linearly preceding element (e.g., the verb), past the unpronounced copy of the subject.

(28) a. **ni³ ĩ³=ti⁵**   ko¹⁵-tu³vʲ=⟨**a⁵**⟩

     NEG one=CL.3PL.N NEG.PST-sting=CL.3SG.F

     'None of them (e.g. wasps) stung her.'   (also ✓: *tu³vi³=ɲa⁵*)

     →   ✓**ni³ ĩ³=ti⁵** ko¹⁵-tu³vi³ ~~ni³ ĩ³=ti⁵~~ =ã⁵

b. **ndʒa⁵ ku⁵=na¹**   ʃi¹³nʲ=⟨**a¹⁵**⟩?

     WH   PRES.COP=CL.3PL.N PST.see=CL.3SG.F

     'Who (is it that) saw her?'   (also ✓: *ndʒa⁵ ku⁵=na¹ ʃi¹³ni³¹=ɲa⁵*)

     →   ✓**ndʒa⁵ ku⁵=na¹** [$_{CP}$ ~~ndʒa⁵~~ ʃi¹³ni³¹ ~~ndʒa⁵~~ =ã⁵]

These data also highlight the fact that this restriction is fundamentally morphophonological in nature. This is perhaps already obvious, given that the relevant clitics do not form a natural class based on morphosyntactic features (3SG.F and 3.N), but do form an obvious morphophonological class (=V). However, (27)-(28) also reveal that this restriction cares only about *overt material*, given that it ignores unpronounced subjects. Assuming a framework in which the syntax precedes (and feeds) a postsyntactic grammatical module (e.g. Embick and Noyer 2001; Halle and Marantz 1993), this means that the restriction must hold only after the syntactic structure has been "flattened" and only after the abstract elements in the syntax are mapped to phonological forms.

We may also contrast the 3rd person =V clitics with participant ones; recall from (8) that 1SG, 2SG, and 1PL.IN clitics are =V. The data in (29) first show that [PART] =V clitics, such as the 2SG clitic =õ⁵, are banned in object position. These data are in principle compatible with both the syntactic explanation put forth above (the PCC) and the morphophonological ban on =V coalescence.

(29) a. *ʃi¹ni³¹=ɲa⁵=⟨õ⁵⟩   b. *ʃi¹ni³¹ ɲa⁵   Ga⁵bʲ=⟨õ¹⁵⟩

     PST.see=CL.3SG.F=CL.2SG   PST.see CL.3SG.F Gaby=CL.2SG

     'She saw you (sg.)   'Gaby saw you (sg.).'

     (✓ *ʃi¹ni³¹ ɲa⁵ ʒo⁵ ʔo¹*)   (✓ *ʃi¹ni³¹ ɲa⁵ Ga⁵bi⁵¹ ʒo⁵ ʔo¹*)

However, the examples in (30-a-b) allow us to tease these options apart: unlike their 3rd person counterparts, participant =V clitics in object position *cannot* cliticize past an unpronounced subject—which we attributed in §4 to the syntactic nature of the PCC. This reveals that the ill-formedness of (29) above is due to the PCC, after all; the [PART] clitics all behave like syntactically.

(30) a. *__ni³__ __ĩ³=na¹__ ko¹⁵-ʃi³n=$\boxed{\text{õ¹⁵}}$

NEG one=CL.3PL.N NEG.PST-see=CL.2SG

'No one saw you.' (✓*ni³ ĩ³=na¹ ko¹⁵-ʃi³ni³¹ ʒo⁵ʔo¹*)

→ *__ni³ ĩ³=na¹__ ko¹⁵-ʃi³ni³¹ ~~ni³ ĩ³=na¹~~ =õ⁵

  b. *__ndʒa⁵ ku⁵=na¹__ ʃi¹³n=$\boxed{\text{õ¹⁵}}$?

WH PRES.COP=CL.3PL.N PST.see=CL.2SG

'Who saw you?' (✓*ndʒa⁵ ku⁵=na¹ ʃi¹³ni³¹ ʒo⁵ʔo¹*)

→ *__ndʒa⁵ ku⁵=na¹__ [$_{CP}$ ~~ndʒa⁵~~ ʃi¹³ni³¹ ~~ndʒa⁵~~ =õ⁵

In sum, whereas all participant clitics behave uniformly *regardless of* their morphological form, the 3rd person clitics behave differently *depending on* their morphological form (i.e., =CV vs. =V). This follows if there is *both* a syntactic operation in the language that targets the [PART] feature *and* a morphophonological rule that picks out =V forms.

Before concluding, I tentatively speculate on why 3SG.F and 3SG.N =V clitics are not able to attach to overt subjects. As mentioned in §3, the =V clitics attach at a lower prosodic level than the =CV clitics do, given that they are integrated into the minimal prosodic word (the bimoraic foot). It may be that subjects (both pronominal clitics and full DPs) have particular parsing or prominence-related requirements, perhaps formulable as a constraint on prosodic integrity. Such a constraint presumably rules out word-internal cliticization to subjects.

6 Conclusion

I have presented two restrictions on pronominal clitics in SJPM, and have argued that they have distinct grammatical sources. The restriction on participant object clitics is a manifestation of the PCC and is syntactic in nature: it references abstract morphosyntactic features (e.g., [PART]), is sensitive to structural notions such as c-command, and considers both overt and unpronounced syntactic material. In contrast, the restriction on certain 3rd person clitics is morphophonological in nature: it specifically prevents the coalescence of vowel-initial clitics to certain hosts, is insensitive to unpronounced syntactic material, and is likely motivated by prosodic requirements.

Finally, I have noted above that similar constraints on clitic combinations appear to be attested in related Oto-Manguean languages. However, prior work on these topics has focused on Zapotec (Foley et al. 2019; Foley and Toosarvandani 2022; Sichel and Toosarvandani to appear; Toosarvandani to appear), distantly related to Mixtec; conversely, there are (to my knowledge) no detailed descriptions or analyses of comparable phenomena in other Mixtec varieties (see fn. 9). The findings of this paper thus form an empirical basis for subsequent (micro-)comparative research, which may, in turn, yield new insights into the pronominal systems of these languages.

References

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

Anagnostopoulou, Elena. 2017. The person-case constraint. In *The wiley-blackwell companion to syntax*, ed. Martin Everaert and Henk van Riemsdijk, 1–47. Hoboken, NJ: Wiley-Blackwell, 2nd edition edition.

Baker, Mark, and Ruth Kramer. 2018. Doubled clitics are pronouns: Amharic objects and beyond. *Natural Language and Linguistic Theory* 36:1035–1088.

Béjar, Susana, and Milan Rezac. 2003. Person licensing and the derivation of PCC effects. In *Romance linguistics: Theory and acquisition*, ed. Ana-Teresa Perez-Leroux and Yves Roberge, 49–61. Amsterdam: John Benjamins.

Béjar, Susana, and Milan Rezac. 2009. Cyclic Agree. *Linguistic Inquiry* 40:35–73.

Boeckx, Cedric. 2003. *Islands and chains: Resumption as stranding*. Amsterdam: John Benjamins.

Bonet, Eulalia. 1991. Morphology after syntax: Pronominal clitics in Romance. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Caballero, Gabriela, Claudia Duarte Bórquez, and Claudia Juárez Chávez. to appeara. Tonal upstep and downstep in San Juan Piñas Mixtec (Tò'ōn Ndā'vī). In *Proceedings of WSCLA 26*. Vancouver, BC: UBCWPL.

Caballero, Gabriela, Claudia Juárez Chávez, and Michelle Yuan. to appearb. The representation of tone in San Juan Piñas Mixtec (Tò'ōn Ndā'ví): Phonological and orthographic implications. In *Proceedings of WCCFL 35*, ed. Gabriela de la Cruz Sanchez, Ryan Walter Smith, Luis Irizarry, Tianyi Ni, and Heidi Harley. Somerville, MA: Cascadilla Press.

Campbell, Eric. 2017. Otomanguean historical linguistics: Past, present, and prospects for the future. *Language and Linguistics Compass* 11:e12240.

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, 145–233. New York: Mouton de Gruyter.

Carroll, Lucien S. 2015. Ixpantepec Nieves Mixtec word prosody. Doctoral Dissertation, University of California San Diego, La Jolla, CA.

Cisneros, Carlos. 2019. Definiteness in Cuevas Mixtec. In *Definiteness across languages*, ed. Ana Aguilar-Guevara, Julia Pozas Loyo, and Violeta Vázquez-Rojas Maldonado, 39–81. Berlin: Language Science Press.

Coon, Jessica, and Stefan Keine. 2021. Feature gluttony. *Linguistic Inquiry* 52:665–710.

De León Pasquel, Maria de Lourdes. 1988. Noun and numeral classifiers in Mixtec and Tzotzil: A referential view. Doctoral Dissertation, University of Michigan Ann Arbor, Ann Arbor, MI.

Deal, Amy Rose. 2015. Interaction and satisfaction in ϕ-agreement. In *Proceedings of NELS 45*, ed. Thuy Bui and Deniz Ozyildiz, 179–192. Amherst, MA: GLSA.

Deal, Amy Rose. to appear. Interaction, Satisfaction, and the PCC. *Linguistic Inquiry* .

Duarte Bórquez, Claudia, Gabriela Caballero, Claudia Juárez Chávez, and Michelle Yuan. 2023. Propiedades tonales y morfosintácticas de los enclíticos pronominales del mixteco de San Juan Piñas (tò'ōn ndā'ví). Presentation at COLOV IX.

Duarte Bórquez, Claudia, and Claudia Juárez Chávez. 2022. The representation of tone in San Juan

Piñas Mixtec: The role of underspecification. Presentation at SSILA 2022.

Embick, David, and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–595.

Foley, Steven, Nick Kalivoda, and Maziar Toosarvandani. 2019. Forbidden clitic clusters in Zapotec. In *Proceedings of CLS 53*, ed. Daniel Edmiston, Marina Ermolaeva, Emre Hakgüder, Jackie Lai, Kathryn Montemurro, Brandon Rhodes, Amara Sankhagowit, and Michael Tabatowski, 87–102. Chicago: Chicago Linguistics Society.

Foley, Steven, and Maziar Toosarvandani. 2022. Extending the Person–Case Constraint to gender: Agreement, locality, and the syntax of pronouns. *Linguistic Inquiry* 53:1–40.

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. Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge, MA: MIT Press.

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

Harizanov, Boris. 2014b. The role of prosody in the linearization of clitics: Evidence from Bulgarian and Macedonian. In *Proceedings of FASL 22: The McMaster meeting 2013*, ed. Cassandra Chapman, Olena Kit, and Ivona Kučerová, 109–130. Ann Arbor, MI: Michigan Slavic Publications.

Hedding, Andrew A., and Michelle Yuan. to appear. Phase unlocking and the derivation of verb-initiality in San Martín Peras Mixtec. In *Proceedings of WCCFL 41*. Somerville, MA: Cascadilla Press.

Ito, Junko, and Armin Mester. 2012. Recursive prosodic phrasing in Japanese. In *Prosody matters: Essays in honor of Elisabeth Selkirk*, ed. Toni Borowsky, Shigeto Kawahara, Mariko Sugahara, and Takahito Shinya, 280–303. London: Equinox Press.

Josserand, Judy Kathyrn. 1983. Mixtec dialect history. Doctoral Dissertation, University of Michigan Ann Arbor, Ann Arbor, MI.

Juárez Chávez, Claudia, Gabriela Caballero, Claudia Duarte Bórquez, José Armando Fernández Guerrero, Ray Huaute, Akil Iyer, Mark Simmons, Nico Tedeschi, Maxine Van Doren, and Michelle Yuan. 2022. *Tò'ōn Ndā'ví San Juan Piñas Mixtec Talking Dictionary*. Living Tongues Institute for Endangered Languages.

Lee, Felicia. 2006. *Remnant raising and VSO clausal architecture*. Dordrecht: Springer.

Macaulay, Monica. 1987. Cliticization and the morphosyntax of Mixtec. *International Journal of American Linguistics* 53:119–135.

Macaulay, Monica. 1996. *A grammar of Chalcatongo Mixtec*. Berkeley, CA: University of California Press.

Mantenuto, Iara. 2020. 'other' and 'more' in San Sebastián del Monte Mixtec and beyond. Doctoral Dissertation, UCLA, Los Angeles, CA.

Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19:153–197.

14

McCloskey, Jim. 2017. Resumption. In *Wiley Blackwell companion to syntax*, ed. Martin Everaert and Henk van Riemsdijk, 3809–3838. Malden, MA: Wiley-Blackwell, 2nd edition edition.

Nevins, Andrew. 2007. The representation of third person and its consequences for Person-Case effects. *Natural Language and Linguistic Theory* 25:273–313.

Pancheva, Roumyana, and Maria Luisa Zubizarreta. 2018. The Person Case Constraint: The syntactic encoding of perspective. *Natural Language and Linguistic Theory* 36:1291–1337.

Perlmutter, David. 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.

Pike, Kenneth L. 1944. Analysis of a Mixteco text. *International Journal of American Linguistics* 10:113–138.

Rizzi, Luigi. 1990. On the Anaphor-Agreement Effect. *Rivista di Linguistica* 2:27–42.

Selkirk, Elizabeth. 1995. The prosodic structure of function words. In *Papers in optimality theory*, ed. Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 439–470. Amherst, MA: GLSA.

Sheehan, Michelle. 2020. The Romance person case constraint is not about clitic clusters. In *Dative constructions in Romance and beyond*, ed. Anna Pineda and Jaume Mateau, 143–171. Berlin: Language Science Press.

Sichel, Ivy, and Maziar Toosarvandani. to appear. The featural life of nominals. *Linguistic Inquiry* .

Stegovec, Adrian. 2020. Taking case out of the Person-Case Constraint. *Natural Language and Linguistic Theory* 38:261–311.

Suñer, Margarita. 1988. The role of agreement in clitic-doubled constructions. *Natural Language and Linguistic Theory* 6:391–434.

Tedeschi, Nicolas. 2023. Frames of reference in San Juan Piñas Mixtec: A methodology for remote elicitation. Presentation at WAIL 25.

Toosarvandani, Maziar. to appear. The interpretation and grammatical representation of animacy. *Language* .

Uchihara, Hiroto, and Juana Mendoza Ruiz. 2022. Minimality, maximality and perfect prosodic word in Alcozauca Mixtec. *Natural Language and Linguistic Theory* 40:599–649.

Woolford, Ellen. 1999. More on the Anaphor Agreement Effect. *Linguistic Inquiry* 30:257–287.

Yuan, Michelle. 2021. Diagnosing object agreement vs. clitic doubling: An Inuit case study. *Linguistic Inquiry* 52:153–179.

Yuan, Michelle. 2022. Deriving VSO in San Juan Piñas Mixtec. Presentation at Syntax Workshop, UMass Amherst.

Yuan, Michelle. 2023. Case as an Anaphor Agreement Effect: Evidence from Inuktitut. *Linguistic Inquiry* 54:413–428.