Clitic coalescence in San Juan Piñas Mixtec at the syntax-phonology interface

Michelle Yuan (UCLA), Gabriela Caballero (UCSD), and Claudia Juárez Chávez*

NELS 56 @ NYU, Oct. 17-19, 2025

1 Introduction

- In San Juan Piñas Mixtec (henceforth SJPM), the 3SG.F and 3.N determiner/pronominal clitics may optionally coalesce with a left-adjacent element, (1).[1]

(1) $k^wa^3ku^3$ **ɲa$^{\mathbf{1}}$=**le^5e^6 ~ $k^wa^3k^{\mathbf{j}}$**=a$^{\mathbf{31}}$** le^5e^6
IRR.cry 3.N=baby IRR.cry=3.N baby
'The baby will cry.'

- Most prior work on clitic coalescence across different Mixtec varieties has focused on its phonetic/phonological properties (e.g. Penner, 2019; DiCanio et al., 2020; Uchihara and Mendoza Ruiz, 2022; Eischens and Bennett, 2025).

This talk: Two **syntactic correlates** governing the application of clitic coalescence.

1. Clitic coalescence cannot target full DPs, but can target bare Ds and other XPs (a *category restriction*).
2. Though clitics cannot coalesce into a preceding DP, they can coalesce out of their own containing DP (a *directional asymmetry*).

(2) a. *ʃi^{13}ni^{31} ti^5=vi^{3}l**$^{\mathbf{j}}$=a$^{\mathbf{5}}$** Pa^{3}lo^{3}ma^5
COMPL.see 3.AN=cat=3SG.F Paloma
Intended: 'The cat saw Paloma.'

b. ʃi^{13}n**$^{\mathbf{j}}$=a$^{\mathbf{15}}$** Pa^{3}lo^{3}ma^5 ti^5 vi^{3}lu^5
COMPL.see=3SG.F Paloma 3.AN cat
'Paloma saw the cat.'

- **Proposal:** These patterns arise because **DPs are phases** and **phonological processes may be phase-constrained**.
  - Clitic coalescence, which effects segmental/tonal changes, is blocked if the target is in a distinct spell-out domain.
  - This is due to grammatical pressures to preserve the phonological properties of spelled out material (e.g. Newell and Piggott, 2014; McPherson and Heath, 2016; Sande, 2018; van Urk, 2025).
  - We moreover posit that spell-out domains = complements of phase heads (not phases themselves), to explain why determiners may coalesce out of a DP.
  - → Alternative accounts insensitive to syntactic category (e.g., relying solely on prosody/phonology) do not derive the correct empirical generalizations.

- **More broadly:**
  - Our findings support a unidirectional syntax-phonology mapping whereby **syntactic structures delimit phonological domains** (*ibid*, also Sande et al. 2020; Fenger and Weisser 2025, a.o.).
  - Our approach is novel in demonstrating this through mismatches between syntactic and prosodic constituency, given our empirical focus on the stem-clitic juncture.

Roadmap:

*yuanm@ucla.edu, gcaballero@ucsd.edu. In additional to our collaborator Claudia Juárez Chávez, we thank Basi Juárez Chávez, Basilisa Pérez Morelos, Cirila Pérez Morelos, among others, for sharing their language with us. For helpful feedback on various aspects of this project, we thank Claudia Duarte Borquez, Ben Eischens, Sharon Rose, Coppe van Urk, Natalie Weber, audiences at UCSD and UCLA, and anonymous reviewers for NELS 56.

[1] Abbreviations: AN = animate, COMPL = completive, CONT = continuative, EMPH = emphatic, EX = exclusive, F = feminine, IN = inclusive, IRR = irrealis, N = neuter, PL = plural, RND = spherical noun class, SG = singular, WOOD = wood/plant noun class, 1 = 1st person, 2 = 2nd person, 3 = 3rd person,

2 Language overview

2.1 Background

- Mixtec (Oto-Manguean) refers to around 60 language varieties in 18 mutually-unintelligible dialect clusters (2,000 years of internal diversification) (Josserand, 1983; Campbell, 2017).[2]

- San Juan Piñas Mixtec (Tò'ōn Ndá'ví; henceforth 'SJPM') is a previously undocumented variety, belonging to the Southern Baja region of Mixtec (Josserand, 1983), a particularly understudied area (Carroll, 2015).[3]

  - SJPM is spoken in Oaxaca (San Juan Piñas, Santiago Juxtlahuaca) by approximately 900 people and a larger number of speakers in diaspora communities in the US and Mexico.

- **SJPM language project**: In collaboration with Claudia Juárez Chávez and other SJPM speakers in San Juan Piñas, Oaxaca and SoCal (2020-present).

  - **Project goals:** (i) phonetic, phonological, and morphosyntactic documentation and analysis (e.g. Duarte Bórquez and Juárez Chávez, 2022; Van Doren et al., 2024; Yuan and Juárez Chávez, 2025; Caballero et al., to appear); (ii) resources for language revitalization and promotion (Juárez Chávez et al., 2022).

2.2 Key grammatical properties

- The base word order of SJPM is **V(P)-S-O**, regardless of whether the nominals are full (lexical) noun phrases or pronominal clitics.

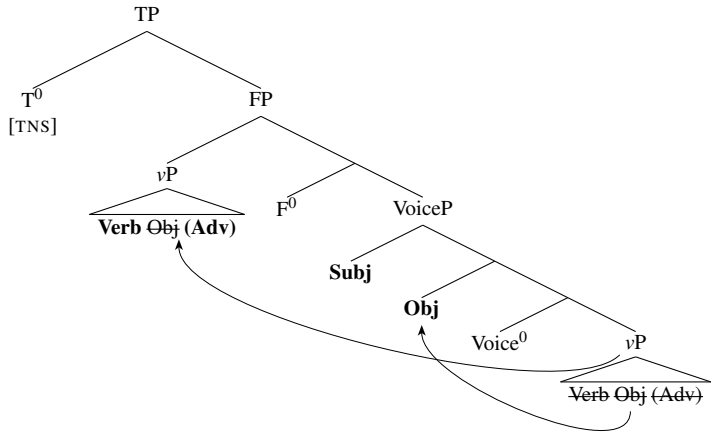
(3) a. ko^{3}ni^{31}=**ɲa^{3}** ti^{5}=vi^{3}lu^{5}
IRR.see=3SG.F 3.AN=cat
'She will see the cat.'
(V=S$_{Cl}$ O)

b. ko^{3}ni^{31} ti^{5}=vi^{3}lu^{5}=**ɲa^{3}**
IRR.see 3.AN=cat=3.N
'The cat will see it.'
(V S=O$_{Cl}$)

- The initial verbal constituent is demonstrably phrasal (§§3-4), so we derive this word order by **object shift + remnant *v*P-movement**, (4) (e.g. Massam, 2001; Lee, 2006).

  - **Note:** If VoiceP is the highest projection of the verbal domain, we assume that it is the clause-medial phase (not *v*P).

(4)



- **Phonological assumptions:**

  - The minimal prosodic word is a **bimoraic foot**, the canonical morphological root, which has the shape (C)V(ˀ)(C)V (see also Carroll 2015, Penner 2019) (= the Mixtec 'couplet'; Pike 1944).

  - The bimoraic foot is right-aligned with the first layer of the prosodic word.

  - Prosodic word (*ω*) (bimoraic foot + prefixes) = domain of progressive nasalization, segmental phonotactic patterns, and certain tonal processes (Duarte Bórquez and Juárez Chávez, 2022).

- Clitics (like most other functional elements) are monomoraic (e.g., (C)V), and they are subcategorized to attach to a linearly adjacent prosodic word (*ω*).

- Some pronominal enclitics are given in (5), with vocalic (V) clitics boxed. In this talk, we focus mainly on the **3SG.F/3.N** forms.[4]

(5) *Human pronominal clitics (non-human and* PL *human forms omitted)*

1SG	2SG	1PL.IN	1PL.EX	2PL	3SG.M	3SG.F	3.N
=e^{1}	=õ^{5}	=e^{5}	=^{n}di$^{(1)}$	=^{n}do^{5}	=ra$^{(1)}$	=ɲa^{5} =ã^{5}	=ɲa$^{(1)}$ =ã$^{(1)}$

[2] Mixtec communities are transnational, in that there are half a million speakers of all varieties in Oaxaca and neighboring states and about 50-100K in California (Kresge, 2007).

[3] SJPM examples are transcribed in this talk using the IPA, plus the Chao numerical system for tones. There are three level tones (H = V^{5}, M = V^{3}, L = V^{1}), plus a right-edge floating L-tone (V$^{(1)}$); the tones 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}, V^{2}) tones are also attested in restricted environments (Caballero et al., 2023). See Caballero et al. to appear and for details about the SJPM tone system.

[4] This is for a few reasons. First, all 1st/2nd person object clitics (of any shape) are categorically ruled out in monotransitives due to a version of the strong PCC (Yuan, 2024; Yuan and Juárez Chávez, 2025), so they are unable to cliticize to preceding DPs for an independent reason. Second, only the 3rd person forms also function as determiners, which are important for our analysis. That being said, the 1st/2nd person vocalic clitics pattern like the 3SG.F/3.N ones in all testable positions.

Important: We take the CV and V 3SG.F/3.N clitic forms to be **contextual allomorphs**, not due to contraction or any other process.

- No synchronic phonological processes explain their forms: no other ɲ-V sequences undergo reduction in SJPM (i.e., this is idiosyncratic to 3SG.F and 3.N clitics).
- The V clitic forms exhibit variable presence/absence of nasality that is speaker-dependent, pointing to further idiosyncracy in phonological form.
- Allomorphy is also more consistent with the cross-Mixtec profile:
  - In some varieties (e.g., San Sebastián del Monte [Mantenuto 2020] and San Martín Peras [Ostrove 2018]), cognate 3SG.F =*ɲa*^{5} lacks a V allomorph but cognate 3.N has both CV and V forms.
  - In other varieties (e.g., Alcozauca [Uchihara and Mendoza Ruiz 2022]), the CV and V allomorphy in the cognate clitic forms does not involve any phonological similarity, e.g. 3SG.F =*ɲa*^{5} and =*e*^{5}.

- We assume that the attachment of a CV clitic forms a recursive prosodic word ($_{\omega 2}$).[5] This is the domain for certain tonal processes, which do not apply across words or within the first layer of the prosodic word.
  - For instance, rightward L-tone spreading targets CV clitics (of a certain tone class):

(6) a. ((ki^{1}si^3)$_\omega$=**na^3**)$_{\omega 2}$
pot=3PL.N
'their pot'

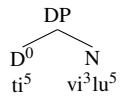
b. ((ki^{3}ni^1)$_\omega$=**na^1**)$_{\omega 2}$
pig=3PL.N
'their pig'

- In contrast, the vocalic (V) clitics are further prosodically integrated: they are parsed into the **same prosodic word** as their hosts, and form a bimoraic foot if they attach to the root (see also Penner 2019, Eischens and Bennett 2025).
- **Coalescence takes place as resolution of vowel hiatus:** V enclitics may coalesce with, replace, or induce gliding of the final V of their host.
  - Note that this could be a preceding root or another enclitic, (7):

(7) a. (ʒu^{3}vi^{15})$_\omega$ + =**a^5** → (ʒu^{3}v^j=**a^{15}**)$_\omega$
petate=3SG.F
'her petate'

b. ((ka^{1}ku^3)$_\omega$=**va^3**)$_{\omega 2}$ + =**a^5** → ((ka^{1}ku^3)$_\omega$=**v**=**a^5**)$_{\omega 2}$
IRR.survive=EMPH =3SG.F
'She will survive!'

- Finally, full DPs bear determiners segmentally identical to their pronominal clitic counterparts.[6] We take determiners and pronominal clitics to be syntactically identical (D^{0}s), differing only in whether they take a nominal complement.

(8) a. 

b. D^0
ti^5

- While CV determiners procliticize to their hosts, the 3SG.F/3.N determiners also have vocalic allomorphs—which encliticize to a linearly preceding element.[7]
- Importantly, **V determiners also undergo the aforementioned leftward coalescence process** (as in (1)-(2)).
  - That vocalic pronouns and determiners behave identically in this respect will inform our analysis in §4.

3 Syntactic distribution of clitic coalescence

3.1 Licit and illicit contexts for coalescence

- The coalescence patterns of vocalic clitics are summarized in (9), though not all representative examples are provided.

(9)

Clitic host	P	N$_{Pssm}$	V	NP-int. Adj	*v*P-int. Adv	Clitic	NP$_{arg}$	DP$_{arg}$
Pron. coalescence?	✓	✓	✓	✓	✓	✓	✓	✗
Det. coalescence?	✓	✓	✓	✓	✓	✓	✓	✗
Ex. in talk	-	(10a)	(1)	(10b)	(11)	(12),(15)	(19c)	(13)

- In addition to coalescence with a preceding verb, seen above, coalescence is permitted with a preceding N(P) (e.g., within a possessive DP), where the host may be the nominal itself or a modifier, (10).

[5] See Gerfen 1999 and Uchihara and Mendoza Ruiz 2022 for arguments for positing the recursive prosodic word in other Mixtec varieties.

[6] Surface tonal differences between determiners and pronominal clitics are fully predictable. It is difficult to find clear phonological evidence that the determiners procliticize to their nominal clitics (as opposed to encliticize to a preceding constituent, like their pronominal clitic counterparts). However, one reason to assume that they are proclitics comes from the fact that they (unlike true enclitics) are available in clause-initial position.

[7] The vocalic determiners are thus descriptively reminiscent of so-called connectives in Tsimshianic languages, which likewise cliticize leftward out of their containing DPs (e.g. Mulder and Sellers, 2010; Davis, 2018).

(10) a. vi^{3}ɟ^{j}=**a^{5}**
cat=3SG.F
'her cat'
(also ✓: vi^{3}lu^{5}=**ɲa^{5}**)

b. le^{5}e^{6} lo$^{3ʔ(w)}$=**a^{5}**
baby small=3SG.F
'her little baby'
(also ✓: le^{5}e^{6} lo^{3ʔ}o^{3}=**ɲa^{5}**)

- If the fronted *v*P is complex, it is the final element that serves as the target for clitic coalescence. This may be a post-verbal Adv (which may also be enclosed in a complex AdvP), (11).
  - Note also that exs. like (11b) support the *v*P-movement analysis of V-S-O word order in §2.2.

(11) a. [$_{vP}$ tu^{5}tu^{5} ʒu^{3ʔ}u^{5} tu^{3}k^{j}]=**ã^{5}**
CONT.whistle mouth again=3SG.F
'She's whistling again.'
(also ✓: …tu^{3}ku^{3}=**ɲa^{5}**)

b. [$_{vP}$ ʃi^{5}ta^{3} [$_{AdvP}$ tʃe^{5ʔ}e^{5} ku^{5}tʃ^{w}]]=**ã^{5}** Pa^{3}lo^{3}ma^{5}
CONT.sing loud very=3SG.F Paloma
'Paloma's singing very loudly.'
(also ✓: …ku^{5}tʃõ^{3} **ɲa^{5}**=Pa^{3}lo^{3}ma^{5})

- Recall from (7b) that clitic coalescence may target a preceding enclitic. Another ex. is provided in (12):

(12) ^{n}da^{15}-kã^{1ʔ}ã^{1}=k=**a^{15}**
NEG.IRR.REP-speak=NPI=3SG.F
'She won't speak again anymore.'
(also ✓: ^{n}da^{15}-kã^{1ʔ}ã^{1}=ka^{1}=**ɲa^{5}**)

- **But clitic coalescence cannot target a preceding DP**, regardless of its internal structure, (13). In such contexts, only the CV variant is available.
  - The badness of these constructions cannot be attributed to morphophonological problems with these particular strings—the same strings are well-formed in other contexts (see (10a), (10b), and (15a)).

(13) a. *ʃi^{13}ni^{31} ti^{5}=vi^{3}ɟ^{j}=**a^{5}** Pa^{3}lo^{3}ma^{5}
COMPL.see 3.AN=cat=3SG.F Paloma
Intended: 'The cat saw Paloma.'
(✓: …ti^{5}=vi^{3}lu^{5} **ɲa^{5}**=Pa^{3}lo^{3}ma^{5})

b. *ko^{3}ni^{3} ɲa^{1}=le^{5}e^{6} lo$^{3ʔ(w)}$=**a^{5}**
IRR.see 3.N=baby small=3SG.F
Intended: 'The little baby will see her.'
(✓: …ɲa^{1}=le^{5}e^{6} lo^{3ʔ}o^{3}=**ɲa^{5}**)

c. *ni^{1}-ta^{3ʔ}vi^{5} ti^{5}=vi^{3}lu^{5}=^{n}d^{j}=**a^{1}** ʒo^{3}o^{3}
COMPL-break 3.AN=cat=1PL.EX=3.N water.jug
Intended: 'Our cat broke the water jug.'
(✓: …ti^{5}=vi^{3}lu^{5}=^{n}di^{1} **ɲa^{1}**=ʒo^{3}o^{3})

- Finally, (14) shows that this pattern is not a restriction inherent to object determiner/pronominal clitics, as coalescence may take place if the preceding DP is extracted.
  - Put differently, coalescence ignores null but syntactically present material, including unpronounced movement copies. Thus, the restriction arises specifically at the DP=Cl juncture.

(14) ^{n}dʒa^{5} ku^{5}u^{3}=na^{1} ʃi^{13}n^{j}=**a^{15}** Pa^{3}lo^{3}ma^{5}?
WH CONT.be=3PL.N COMPL.see=3SG.F Paloma
'Who (is it that) saw Paloma?'
(also ✓: ^{n}dʒa^{5} ku^{5}u^{3}=na^{1} ʃi^{13}ni^{31} **ɲa^{5}**=Pa^{3}lo^{3}ma^{5})

The data provided so far also demonstrate that the clitic coalescence restriction is not due to (purely) prosodic factors.

- Under accounts like Match Theory (Selkirk, 1995, 2011) that map syntactic constituents to prosodic constituents, these units are not distinguishable by syntactic category.[8]
  - e.g., $X^{0} = \omega$ (prosodic word); XP = Φ (phonological phrase).
- But syntactic category does matter here, since elements that are not D in category do tolerate clitic coalescence into them.[9]

★ **The source of this restriction is both phonological and syntactic in nature.**

3.2 DPs vs. other nominal arguments

- At this point, it is also possible to say that the restriction pertains to coalescence with nominal arguments in general (not DPs specifically).
- **However:** This is not a viable alternative, as non-DP arguments are licit targets for clitic coalescence.

[9] In a similar vein, the clitic coalescence restriction does not seem to be about avoiding phonological imbalances, e.g., to satisfy eurhythmic constraints. The fact that determiner clitics (with overt complements) behave like bare pronominal clitics casts doubt on this.

1. Clitic coalescence may target **other pronominal clitics**—bare D^0s rather than full DPs, (15):

(15) a. ka^{3}tõ^{5}=^{n}d^{j}=**a^{1}**
IRR.tie=1PL.EX=3.N
'We'll tie it.'
(also ✓: ka^{3}tõ^{5}=^{n}di^{1}=**ɲa^{1}**)

b. ka^{13}ni^{3} ^{n}dʒa^{3}a^{31}=t^{w}=**ã^{5}**
COMPL.hit crash=3.WOOD=3SG.F
'It (the car) hit her!'
(also ✓: ka^{13}ni^{3} ^{n}dʒa^{3}a^{31}=tõ^{5}=**ɲa^{5}**)

- Like certain other Mixtec varieties (Macaulay, 1987), SJPM has a three-way distinction between pronominal clitics, weak pronouns, and strong pronouns (e.g., 1PL.EX =$^{n}di^{(1)}$ vs. $^{n}du^{1}$ vs. $^{n}du^{1\,?}u^{1}$).
  - Weak pronouns are monomoraic like clitics, but **behave morphosyntactically like strong pronouns**.
  - For instance, the Person-Case Constraint restricts pronominal clitics but not strong or weak pronouns (which are used as repairs) (Yuan and Juárez Chávez, 2025).[10]

2. Importantly, **neither strong nor weak pronouns can be targeted for coalescence**, (16)—another argument against prosodic size as the determinant for clitic coalescence.
   - But if strong and weak pronouns are both DPs, then this explains their uniform behavior wrt the PCC and clitic coalescence.

(16) *ka^{3}tõ^{5} ^{n}du^{1}$^{?w}$=**a^{1}** / *^{n}d^{w}=**a^{1}**
IRR.tie 1PL.EX.PRON(STRONG)=3.N 1PL.EX.PRON(WEAK)=3.N
Intended: 'We'll tie it.'
(✓: ka^{3}tõ^{5} ^{n}du^{1}$^{?}$u^{1}=**ɲa^{1}**, ka^{3}tõ^{5} ^{n}du^{1}=**ɲa^{1}**)

3. Additional evidence comes from **(determinerless, indefinite) N(P)s**, which may appear as internal arguments.
   - These reduced NPs still vacate the fronted *v*P, i.e., are targeted by object shift, (17). (Pseudo noun incorporation, yielding VOS, is only found in selects contexts; **see Appendix**.)

(17) ta^{3}va^{5}=na^{1} **tʃa^{13}ka^{5}**
IRR.take.out=3PL.N fish
'They'll go fishing.' (Lit. 'They'll take out fish.')

[10]This is also true for other Mixtec varieties such as Chalcatongo Mixtec, based on the discussion in Macaulay 1987.

   - SJPM has a handful of double object constructions involving an instrument IO, resulting in V(P)-S-IO-DO order.[11]

(18) kõ^{3}õ^{3}=ra^{1} ʒu^{1}u^{5} ʒo^{5}o^{5}
IRR.hit.with=3SG.M rock 1PL.IN
'He'll hit us (incl.) with a rock.' (V S IO DO)

   - A direct object clitic cannot coalesce with a preceding DP instrument, (19a-b).[12] **But clitic coalescence <u>is</u> possible if the preceding instrument is a determinerless NP, (19c).**

(19) a. *kõ^{13}õ^{3}=ra^{1} ɲa^{1}=ʒu^{1}u^{5} ʒo^{5}$^{?w}$=**a^{5}**
COMPL.hit.with=3SG.M 3.N=rock DEM=3SG.F
Intended: 'He hit her with this rock.'
(✓: ...ɲa^{1}=ʒu^{1}u^{5} ʒo^{5}$^{?}$o^{3}=**ɲa^{5}**

b. *kõ^{13}õ^{3}=ra^{1} k^{w}a^{1}$^{?}$a^{3} ʒi^{1}=**a^{5}**
COMPL.hit.with=3SG.M many rock=3SG.F
Intended: 'He hit her with many rocks.'
(✓: ...k^{w}a^{1}$^{?}$a^{3} ʒu^{1}u^{5}=**ɲa^{5}**)

c. kõ^{13}õ^{3}=ra^{1} ʒi^{1}=**a^{5}**
COMPL.hit.with=3SG.M rock=3SG.F
'He hit her with a rock.'
(also ✓: ...ʒu^{1}u^{5}=**ɲa^{5}**)

★ These data show that coalescence is not blocked with <u>all</u> nominal arguments. **Rather, the clitic coalescence restriction is specifically keyed to DPs.**

4 A phasal analysis

Proposal:

- **DPs are phases** (e.g. Svenonius, 2004; Aravind, 2021), cyclically sent to the interfaces upon being built. Clitic coalescence cannot take place between elements in different spell-out domains.

[11]Such constructions are truly ditransitive in that none of the arguments, including the instrument IO, can be omitted. Therefore, these can be analyzed as having low applicative structures (e.g. Pylkkänen, 2008)

[12]As expected, it is also possible for the direct object clitic to coalesce with an instrument clitic.

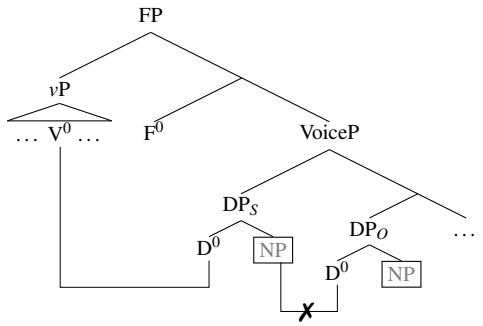
(i) kõõ^{3}=^{n}di^{1}=t^{w}=ã^{5}
IRR.hit.with=1PL.EX=3.WOOD=3SG.F
'We'll hit her with it (e.g., a stick).'
(also ✓: kõõ^{3}=^{n}di^{1}=tõ^{5}=ɲa^{5})

- The spell-out unit is the **phasal complement** (Chomsky 2001, *pace* Boškovič 2016), i.e., excludes the phase head.
- (Besides CP and VoiceP by assumption,) Other XPs are not phases. Put differently, phasality is diagnosed by disruptions to clitic coalescence.

→ These components capture both the **category restriction** (singling out DPs) and **directional asymmetry** (leftward coalescence out of DPs, but no rightward coalescence into them).

- Schematized in (20) (boxes = spelled out material; lines = coalescence):

(20)



4.1 Why no coalescence?

⋆ **Two convergent ideas on the syntax-phonology of phases:** (i) cyclic spell-out of phasal material delimits phonological domains, and (ii) certain phonological processes may not apply across such domains.[13]

- **Caveat:** Most work on this assumes that phases are not fully phonologically impenetrable, but rather selectively so (e.g. Newell and Piggott, 2014; McPherson and Heath, 2016; Newell, to appear). We adopt this position as well, though remain agnostic towards the exact implementation.
- **Application to SJPM, using (21) to illustrate:**

(21) a. ʃi^{13}ni^{31} ti^{5}=vi^{3}lu^{5} **ŋa^{5}**=Pa^{3}lo^{3}ma^{5}
COMPL.see 3.AN=cat 3SG.F=Paloma
'The cat saw Paloma.'

b. *ʃi^{13}ni^{31} ti^{5}=vi^{3}l**ʲ=a^{5}** Pa^{3}lo^{3}ma^{5}
COMPL.see 3.AN=cat=3SG.F Paloma
Intended: 'The cat saw Paloma.'

[13] See, for instance, Marvin 2003, Newell and Piggott 2014, Pak 2008, Kahnemuyipour 2009, McPherson and Heath 2016, Sande 2018, Fenger and Weisser 2025, van Urk 2025, a.m.o.

- The relevant clitic exponents are inserted postsyntactically. Inserting a CV allomorph for the 3SG.F clitic is always an option.[14]
- But if a V allomorph is inserted, it needs to undergo coalescence with its host to resolve **vowel hiatus** (§2.2).
- **We propose spelled-out material is immutable for such tonal and segmental changes, if triggered in a subsequent cycle.**
- But it is equally impossible for coalescence not to apply, as in (22)—as vowel hiatus within prosodic words (including recursive prosodic words) is generally not tolerated either.[15]

(22) *ʃi^{13}ni^{31} ti^{5}=vi^{3}l**u^{5}=a^{5}** Pa^{3}lo^{3}ma^{5}
COMPL.see 3.AN=cat=3SG.F Paloma
Intended: 'The cat saw Paloma.'

- This logic, in turn, suggests that spelled out material is not wholly phonologically invisible[16]—but rather subject to extra faithfulness (McPherson and Heath, 2016).

Thus, when the clitic and host are in separate spell-out domains, the V allomorph is systematically ruled out, regardless of whether coalescence occurs.

- In these contexts, only the CV form may surface.

4.2 Further discussion

- Given our phase-based analysis, we expect other potential phase boundaries (e.g., VoicePs, CPs) to also prevent coalescence across them. **However, these latter configurations are not testable in SJPM.**
  - Given the derivation of VP-S-O word order (§2.2), the verb, subject, and object are all at the VoiceP phase edge or higher.

[14] CV enclitics simply create recursive prosodic words with their hosts (§2.2). We assume this is possible within and across spell-out domains.

[15] This is in contrast to a language like Ojibwe, in which vowel hiatus *is* generally tolerated across spell-out domains (Newell and Piggott, 2014). We assume that this is a matter of cross-linguistic variation.

[16] Indeed, CV clitics attaching to DPs are subject to certain processes stemming from the host. For instance, the L-tone spreading process seen in (6) may apply across DP phase boundaries:

(i) ko^{3}ni^{31} ti^{5}=k^{w}a^{5}ʒu^{1}=**ŋa^{1}**
IRR.see 3.AN=horse=3.N
'The horse will see it.'

- It is likewise impossible to test clitic coalescence across embedded CP phases, as they are necessarily complementizer- or verb-initial.

- **Our phase-based account is thus based entirely on the behavior of DPs.** This raises the question of whether the restriction could be captured with independent (non-phasal) properties of DPs...
  - In lieu of a phase-based account, suppose that DPs have a left vs. right edge asymmetry, where only the right edge is subject to some strong boundary or constraint (e.g. Ito and Mester, 1999; Palancar, 2004; Clemens, 2021).
  - However, this arguably simply restates the empirical generalization without deeper explanation.
  - **In contrast, the phasal account makes sense of why DPs are singled out, and draws parallels with other phase-constrained processes cross-linguistically.**

5 Conclusion

- Besides the novel empirical contributions documenting SJPM, our analysis connects a complex pattern of clitic coalescence in this language to other phonological phenomena shown to be sensitive to syntactic phases.
- On the level of microtypology, we have alluded to the diversity of clitic coalescence patterns within the Mixtec language family.
  - We seek to contribute to further careful documentation and analysis of these patterns, and open new avenues for elucidating the role that syntactic constraints play in their analysis.
  - Indeed, our findings significantly expand on existing work on Mixtec clitic coalescence, which has not considered the role of syntax.

References

Aravind, Athulya. 2021. Successive cyclicity in DPs: Evidence from Mongolian nominalized clauses. *Linguistic Inquiry* 52:377–392.

Bošković, Željko. 2016. What is sent to spell-out is phases, not phasal complements. *Linguistica* 56:25–66.

Caballero, Gabriela, Claudia Duarte Borquez, and Claudia Juárez Chávez. 2023. Tonal upstep and downstep in San Juan Piñas Mixtec (tò'ōn ndā'vī). Presentation at WSCLA 26.

Caballero, Gabriela, Claudia Duarte Borquez, Claudia Juárez Chávez, and Michelle Yuan. to appear. Lexical and grammatical tone in San Juan Piñas Mixtec (Tò'ōn Ndá'ví). *Phonological Data and Analysis* .

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

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

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.

Clemens, Lauren. 2021. The use of prosody as a diagnostic for syntactic structure: The case of verb-initial order. In *Parameters of predicate fronting*, ed. Vera Lee-Schoenfeld and Dennis Ott, 63–95. Oxford: Oxford University Press.

Clemens, Lauren Eby. 2019. Prosodic noun incorporation: The relationship between prosody and argument structure in Niuean. *Syntax* 22:337–377.

Davis, Henry. 2018. Only connect! Determiners, case, and agreement in Tsimshianic. *International Journal of American Linguistics* 84:441–591.

DiCanio, Christian, Caicai Zhang, Douglas H. Whalen, and Rey Castillo García. 2020. Phonetic structure in Yoloxóchitl Mixtec consonants. *Journal of the International Phonetic Association* 50:333–365.

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.

Eischens, Ben, and Ryan Bennett. 2025. The Mixtec couplet is not a consistent grammatical unit. Ms., UCLA and UCSC.

Fenger, Paula, and Philipp Weisser. 2025. Matching domains: The syntax, morphology, and phonology of the verb in Sinhala. *Natural Language and Linguistic Theory* 43:2523–2582.

Ito, Junko, and Armin Mester. 1999. Realignment. In *The prosody-morphology interface*, ed. René Kager, Harry van der Hulst, and Wim Zonneveld, 188–217. Cambridge, UK: Cambridge University 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. URL `https://talkingdictionary.swarthmore.edu/sanjuanpinas/`.

Kahnemuyipour, Arsalan. 2009. *The syntax of sentential stress*. Oxford: Oxford University Press.

Kresge, Lisa. 2007. Indigenous oaxacan communities in California: An overview. *California Institute for Rural Studies* .

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.

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

Marvin, Tatjana. 2003. Topics in the stress and syntax of words. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

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

McPherson, Laura, and Jeffrey Heath. 2016. Phrasal grammatical tone in the Dogon languages. *Natural Language and Linguistic Theory* 34:593–639.

Mulder, Jean, and Holly Sellers. 2010. Classifying clitics in Sm'algyax: Approaching theory from the field. In *Fieldwork and linguistic analysis in indigenous languages of the Americas*, ed. Andrea L. Berez, Jean Mulder, and Daisy Rosenblum, 33–56. Honolulu: University of Hawai'i Press.

Newell, Heather. to appear. Phases and phonology. In *Wiley-blackwell companion to phonology*, ed. Kuniya Nasukawa, Bridget Samuels, Geoff Schwartz, and Miklós Törkenczy. Hoboken, NJ: Wiley-Blackwell.

Newell, Heather, and Glyne Piggott. 2014. Interactions at the syntax-phonology interface: Evidence from Ojibwe. *Lingua* 150:332–362.

Ostrove, Jason. 2018. When ϕ-agreement targets topics: The view from San Martín Peras Mixtec. Doctoral Dissertation, University of California Santa Cruz, Santa Cruz, CA.

Pak, Marjorie. 2008. The postsyntactic derivation and its phonological reflexes. Doctoral Dissertation, University of Pennsylvania, Philadelphia, PA.

Palancar, Enrique L. 2004. Verbal morphology and prosody in Otomi. *International Journal of American Linguistics* 70:251–278.

Penner, Kevin L. 2019. Prosodic structure in Ixtayutla Mixtec: Evidence for the foot. Doctoral Dissertation, University of Alberta, Edmonton, AB.

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

Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge, MA: MIT Press.

Sande, Hannah. 2018. Cross-word morphologically conditioned scalar tone shift in Guébie. *Morphology* 28:253–295.

Sande, Hannah, Peter Jenks, and Sharon Inkelas. 2020. Cophonologies by ph(r)ase. *Natural Language and Linguistic Theory* 38:1211–1261.

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.

Selkirk, Elizabeth. 2011. The syntax-phonology interface. In *The handbook of phonological theory*, ed. John Goldsmith, Jason Riggle, and Alan Yu, 435–485. Blackwell, 2nd edition edition.

Svenonius, Peter. 2004. On the edge. In *Peripheries: Syntactic edges and their effects*, ed. David Adger, Céline de Cat, and George Tsoulas, 259–287. Dordrecht: Kluwer.

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

van Urk, Coppe. 2025. The cycle in a syllable: Phases and impenetrability in Dinka morphology. Ms., QMUL.

Van Doren, Maxine, Claudia Duarte Borquez, Claudia Juárez Chávez, and Gabriela Caballero. 2024. Illustration of San Juan Piñas Mixtec. *Journal of the International Phonetic Association* 54:853–882.

Yuan, Michelle. 2024. More on 'Lexical DP Blocking' effects in the PCC: Evidence from Mixtec. In *Proceedings of NELS 54*, ed. Shaunak Phadnis, Carla Spellerberg, and Brynne Wilkinson, 245–255. Amherst, MA: GLSA.

Yuan, Michelle, and Claudia Juárez Chávez. 2025. Remarks on 'Lexical DP Blocking' in the Person-Case Constraint: Comparative evidence from San Juan Piñas Mixtec. ms., UCLA.

A Further evidence from pseudo noun incorporation

- In SJPM, pseudo noun incorporation of internal arguments is permitted **just for reciprocals**.
- **Assumption:** Reciprocals are bare Ns and are invisible to the phrasal movement operations that can target DPs and NPs.
- Strikingly, if a reciprocal is contained within a larger XP, that entire XP (possibly also containing other nominal arguments) fronts with the *v*P as well.[17]
  - We currently lack a deeper explanation for this fact, but the intuition is that the reciprocal's inability to undergo object shift traps other elements within the containing XP too.

(23) a. $[_{vP}$ no^{3}mi^{3} tã^{5ʔ}ã$^{3}]$=**na^{5}**
IRR.hug each.other=3PL.F
'They will hug each other.' (V O_{Rec}=S_{Cl})

b. $[_{vP}$ ^{n}dʒa^{5}a^{6} se^{1ʔ}e^{3} tã^{5ʔ}ã$^{3}]$=**na^{3}**
CONT.care child each.other=3PL.N
'Everyone takes care of each other's children.' (V Pssr $Pssm_{Rec}$=S_{Cl})

c. $[_{vP}$ kõ^{13}õ^{3} $[_{ApplP}$ ʒu^{1}u^{5} tã^{5ʔ}ã$^{3}]]$=**na^{3}**
COMPL.hit rock each.other=3PL.N
'They hit each other with rocks.' (V IO O_{Rec}=S_{Cl})

- **Clitic coalescence is permitted into such fronted constituents**, (24). (Because reciprocals require a plural antecedent, we illustrate this pattern with 1PL.IN =e^{5}.)
  - These data further demonstrate that the fronted verbal constituent is phrasal (not a series of complex heads).
  - And they show again that clitic coalescence can target a linearly preceding nominal, so long as it is reduced.

(24) a. $[_{vP}$ no^{3}mi^{3} tã$^{5ʔ}]$=**ẽ^{5}**
IRR.hug each.other=1PL.IN
'We will hug each other.' (V O_{Rec}=S_{Cl})

b. $[_{vP}$ ^{n}dʒa^{5}a^{6} se^{1ʔ}e^{3} tã$^{5ʔ}]$=**ẽ^{5}**
CONT.care child each.other=1PL.IN
'We take care of each other's children.' (V Pssr $Pssm_{Rec}$=S_{Cl})

c. $[_{vP}$ kõ^{13}õ^{3} $[_{ApplP}$ ʒu^{1}u^{5} tã$^{5ʔ}]]$=**ẽ^{5}**
COMPL.hit rock each.other=1PL.IN
'We hit each other with rocks.' (V IO O_{Rec}=S_{Cl})

[17] This is the case for other Mixtec varieties as well; see, e.g., Ostrove 2018 on identical facts in San Martín Peras Mixtec.