Clitic coalescence and cyclic spell-out in San Juan Piñas Mixtec

Michelle Yuan (UCLA, yuanm@ucla.edu)
*joint work w/ Gabriela Caballero (UCSD) and Claudia Juárez Chávez**

May 21, 2026, UChicago Linguistics

1 Introduction

- In the San Juan Piñas variety of Mixtec, the 3SG.F and 3.N pronominal and determiner clitics may **optionally coalesce** (segmentally and tonally) with a left-adjacent element, (1).[1]

(1) a. vi^{3}lu^{5}=**ɲa^{5}** ~ vi^{3}l^{j}**a^{5}**
cat=3SG.F cat.3SG.F
'her cat'

b. k^{w}a^{3}ku^{3} **ɲa^{1}**=le^{5}e^{6} ~ k^{w}a^{3}k^{j}**a^{31}** le^{5}e^{6}
IRR.cry 3.N=baby IRR.cry.3.N baby
'The baby will cry.'

- Prior work on clitic coalescence in Mixtec has exclusively 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: The application of clitic coalescence is also governed by **syntactic factors**—providing a new window into the **syntax-phonology interface**.

- **Two related observations:**
  - Clitic coalescence cannot target full DPs, but can target bare D^{0}s and other XPs (a *category restriction*).
  - Though clitics cannot coalesce <u>into</u> a preceding DP, (as determiners) they can coalesce <u>out of</u> their own containing DP (a *directional asymmetry*).

(2) a. *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.'
(✓: … ti^{5}=vi^{3}lu^{5} **ɲa^{5}**=Pa^{3}lo^{3}ma^{5})

b. ʃi^{13}n^{j}**a^{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.'
(also ✓: ʃi^{13}ni^{31} **ɲa^{5}**=Pa^{3}lo^{3}ma^{5}…)

Proposal: These patterns arise because **DPs are phases** and **morphophonological processes may be phase-constrained**.

- Clitic coalescence, which effects segmental/tonal changes, is blocked if the target is in a distinct spell-out domain.
  - 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).
- Spell-out domains = complements of phase heads (<u>not</u> phases themselves), explaining why D^{0}s may coalesce out of DPs.

→ Alternative accounts that do not make reference to spell-out domains fail to derive the correct empirical generalizations.

- More broadly, these findings support a growing body of research showing that **phonological domains are delimited by the syntax** (*ibid*, also Simpson and Wu 2002; Kramer 2010; Sande et al. 2020; Guekguezian 2021; Fenger and Weisser 2025, a.o.).

*In addition to our collaborator Claudia Juárez Chávez, we are grateful to 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, Andrew Hedding, Sharon Rose, Coppe van Urk, Natalie Weber, and audiences at UCSD, UCLA, and NELS 56. Some of the generalizations re: verb-initiality build on Yuan's joint work with Andrew Hedding (UWashington) on the closely related San Martín Peras Mixtec.

[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,

Roadmap:

§2 Language overview
§3 Syntactic distribution of clitic coalescence
§4 DPs vs. non-DP nominals
§5 A phasal analysis
§6 Cliticization and phase (im)penetrability
A Further data from pseudo noun incorporation

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).

  - 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.

  - Closely related to the San Martín Peras variety, on which there has been some syntactic research (e.g. Ostrove, 2018, 2026; Hedding, 2022, to appear; Hedding and Yuan, 2025a,b).

- **SJPM language project**: In collaboration with Gabriela Caballero (UCSD) and various students at UCSD/UCLA, and 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]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).

A note on transcription: SJPM examples are transcribed in this talk using the IPA, plus the Chao numerical system for tones. See Caballero et al. to appear and Duarte Borquez et al. to appear for details about the SJPM tone system.

- There are three level tones (H = V^5, M = V^3, L = V^1), plus a right-edge floating L-tone ($V^{(1)}$) that associates only in certain contexts.
- The tones may combine to form various rising and falling contours (e.g., LH = V^{15}, ML = V^{31}).
- Tones upstepped from H (V^6) and downstepped from H (V^4) and M (V^2) are also attested in restricted environments.

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}=**ŋa^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=**ŋa^3**
IRR.see 3.AN=cat=3.N
'The cat will see it.'
(V S=O_{Cl})

- S-V(P)-O is also commonly attested, and the fronted subject often co-occurs with a resumptive clitic in the subject's base position (such examples occasionally show up in this talk, as in (4)).

(4) **ti^5=vi^{3}lu^5** ko^{3}ni^{31}=**ti^5**=ŋa^3
3.AN=cat IRR.see=3.AN=3.N
'The cat will see it.' (S V=S_{Cl}=O_{Cl})

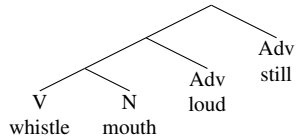
- The initial verbal constituent is **demonstrably phrasal**, given the order of postverbal modifiers that front with the verb (diagnostic from van Urk 2024).[3]

  - Sequences of such modifiers scope <u>right-to-left</u>, which makes sense if they each right-adjoin, (5b).

(5) a. [tu^{5}tu^5 **ʒu^{3}ʔu^5 tʃe^{5}ʔe^5 ʒaa^1**]=ŋa^5
CONT.whistle mouth loud still =3SG.F
'She is still whistling (with mouth) loudly.' (still > loudly > mouth)

[3]This pertains just to *v*P-internal adverbs, as predicted under the present account. Temporal adverbs, for instance, appear only clause-finally or clause-initially.

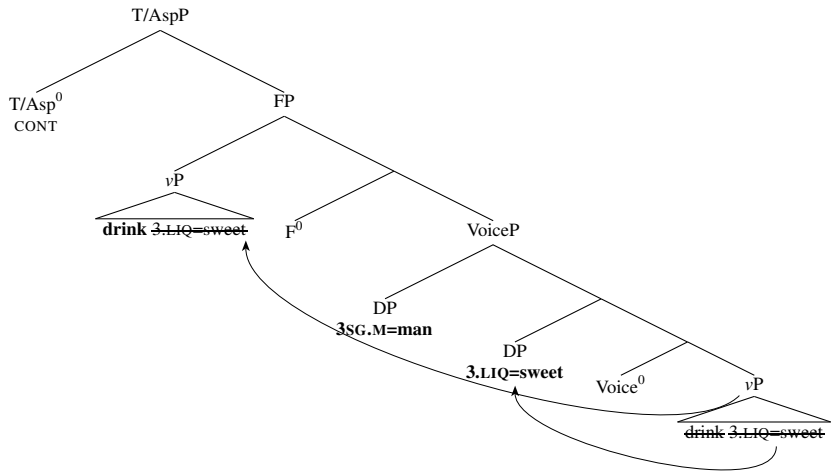
b.



- VP-S-O word order may thus be derived **object shift + remnant *v*P-movement**, (6) (e.g. Massam, 2001; Lee, 2006).
  - In (6), object shift is depicted as tucking into a lower specifier of Voice^0, under the subject (Richards, 2001).
  - *v*P is taken to move to some position between the subject and tense/aspect (which is expressed as either a prefix or a toneme that docks to the left edge of the verb).

(6) a. ʃi^{5}ʔi^{3} ɾa^{1}=tʃa^{1}a^{3} ɾa^{5}=vi^{1}ʃi^{1}
CONT.drink 3SG.M=man 3.LIQ=sweet
'The man is drinking the soda.' (V S O)

b.



2.3 Properties of pronominal clitics

- The minimal prosodic word (also the canonical morphological root) is a **bimoraic foot**, aka the Mixtec 'couplet' (Pike 1944; see also Carroll 2015, Penner 2019). Words in SJPM thus have the shape (C)V(ʔ)(C)V.[4]
- 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).

[4]The bimoraic foot is right-aligned with the first layer of the prosodic word. This first layer of the prosodic word = domain of progressive nasalization, segmental phonotactic patterns, and certain tonal processes (Duarte Bórquez and Juárez Chávez, 2022).

- Some pronominal enclitics are given in (7), which shows that they may be =CV or =V (vocalic clitics are boxed).
  - All of the vocalic enclitics coalesce with a preceding element (but this talk focuses on 3SG.F/3.N).[5]

(7) *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}	=ɾa$^{(1)}$	=ɲa^{5} =ã^{5}	=ɲa$^{(1)}$ =ã$^{(1)}$

Important: We take the CV and V 3SG.F/3.N clitic forms to be **contextual allomorphs** (so words containing coalescence will now be glossed as 'stem=CL').

- 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).
- Although this alternation resembles contraction (e.g., *can't*), allomorphy is more consistent with the cross-Mixtec profile:
  - In the San Sebastián del Monte (Mantenuto, 2020) and San Martín Peras (Ostrove, 2018) varieties, only cognate 3.N =*ɲa*$^{(1)}$, but not cognate 3SG.F =*ɲa*^{5}, has a vocalic variant.
  - In the Alcozauca variety (Uchihara and Mendoza Ruiz, 2022), the CV and V forms not bear any phonological similarity (e.g., 3SG.F =*ɲa*^{5} vs. =*e*^{5}), suggesting suppletive allomorphy.
- Lastly, the 3SG.F/3.N forms behave on par with the other vocalic clitics, which do not have CV variants (so could not be analyzed as optional contraction).

- Different clitics (within a single language) may attach at **different prosodic levels** (e.g. Selkirk, 1995; Ito and Mester, 1999):

(8) a. free clitic: ((host)$_\omega$ clitic)$_\Phi$ ← not relevant to this talk
b. affixal clitic: ((host)$_\omega$ clitic)$_\omega$
c. internal clitic: (host clitic)$_\omega$

[5]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.

- The CV and V clitics in SJPM are **affixal** vs. **internal** clitics, respectively.
  - Put together with the above, we suggest that allomorphs may differ not only in form but also in their prosodic subcategorization properties.
- Affixal CV clitics in SJPM attach to form a **recursive prosodic word**,[6] whose existence is evidenced by phonological processes that do not apply within the first layer of the prosodic word.
  - One such process = rightward L-tone spreading, which targets CV clitics but not other elements in the same minimal prosodic word:

(9) a. $((\underline{so^{1}}kõ^{5}$ / $\underline{ki^{1}}si^{3})_{\omega}$**=na^{3}**$)_{\omega 2}$
neck pot=3PL.N
'their neck/their pot'

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

- Internal V clitics are parsed into the **same prosodic layer** as their hosts (see also Penner 2019, Eischens and Bennett 2025).
  - These clitics may coalesce with, replace, or induce gliding of the final V of their hosts, maintaining the prosodic size of the word (e.g., bimoraic stems remain bimoraic).[7]
  - Coalescence resolves vowel hiatus, which is generally not tolerated in the language.

(10) a. $(so^{1}kõ^{5})_{\omega}$ + **=a^{5}** → $(so^{1}k^{w}$**=ã^{5}**$)_{\omega}$
neck =3SG.F
'her neck'

b. $(i^{3}ta^{3(1)})_{\omega}$ + **=a^{5}** → $(i^{3}t$**=a^{15}**$)_{\omega}$
flower =3SG.F
'her flower'

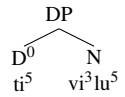
c. $((ka^{1}ku^{3})_{\omega}$**=va^{3}**$)_{\omega 2}$ + **=a^{5}** → $((ka^{1}ku^{3})_{\omega}$**=v=a^{5}**$)_{\omega 2}$
IRR.survive=EMPH =3SG.F
'She will survive!'

- Finally, full noun phrases may bear **classifiers** that are segmentally identical to their pronominal clitic counterparts (seen in several exs. above).

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

[7] In certain cases, coalescence may also affect the penultimate V: e.g., *ku^{3}ʒu^{1}ʔu$^{5(1)}$* 'to fear' → ***ku^{3}ʒi^{15}ʔ=a^{15}*** 'she will fear'.

- We refer to these as **determiners**, following Royer 2019, 2022 on Chuj (Mayan), whose classifier/determiner system is very similar.[8]
  - Moreover, we assume that determiners and pronominal clitics are syntactically identical (D^{0}s), differing only in whether or not they take a nominal complement.
  - Put differently, determiners project while pronominal clitics do not (e.g. Roberts, 2010).[9]

(11) 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.[10]
- Importantly, **V determiners also undergo the aforementioned leftward coalescence process**, as in (1b), repeated as (12).[11]

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

⋆ That vocalic pronouns and determiners are **uniformly enclitic** will inform our analysis.

3 Syntactic distribution of clitic coalescence

- The coalescence patterns of vocalic clitics are summarized in (13) below. Note that the pronominal and determiner coalescence patterns are identical **in every context** tested.

[8] 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.

[9] Or in Roberts's (2010) terms, pronominal clitics are $D^{min/max}$ elements. It is also possible to label these elements as Cl(assifier)^{0}s, with no important consequences for the syntactic analysis in this talk. Recall that this talk takes DPs to be phasal. Under a ClP account of SJPM nominals, we would simply take the highest nominal projection to be phasal, whatever its label.

[10] 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).

[11] Because vocalic determiners are necessarily enclitic, they are unavailable inside DPs in clause-initial position or in isolation.

(13)

Clitic host	V	P	N_{Pssm}	DP-int. Adj/Q.	*v*P-int. Adv	Clitic	NP_{arg}	DP_{arg}
Pron. coalescence?	✓	✓	✓	✓	✓	✓	✓	✗
Det. coalescence?	✓	✓	✓	✓	✓	✓	✓	✗
Ex. in handout	(1b)	(14)	(10a)	(15)	(16)	(17),(18)	(25)	(19)

- In addition to coalescence with a preceding verb, coalescence is permitted with a preceding preposition:

(14) ɲa^{5}=Pa^{3}lo^{3}ma^{51} ĩ^{5}=**ã^{5}** ʃĩ^{ʔ}=**ã^{5}**
3SG.F=Paloma CONT.be.LOC=3SG.F with=3SG.F
'Paloma is with her.'
(also ✓: ĩ^{5}ĩ^{3}=**ɲa^{5}**; ʃĩ^{5ʔ}ĩ^{3}=**ɲa^{5}**)

- Within DPs, coalescence may target a preceding N(P), where the host may be the nominal itself (as in (10a)) or a postnominal modifier, (15a). In quantified expressions, it may also target a quantifier, (15b).

(15) a. 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}**)

b. ^{n}di$^{3ʔ(j)}$=**a^{31}** ʒo^{3}o^{3}
all=3.N water.jug
'all the water jugs'
(also ✓: ^{n}di^{3ʔ}i^{3} **ɲa^{1}**=ʒo^{3}o^{3})

- 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), (16).

(16) a. [$_{vP}$ tu^{5}tu^{5} ʒu^{3ʔ}u^{5} **ti^{5}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 (10c) that clitic coalescence may target a preceding enclitic. Another ex. is provided in (17):

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

- Clitic coalescence may target **other pronominal clitics**—bare D^{0}s rather than full DPs, (18):

(18) 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}**)

- **But clitic coalescence cannot target a preceding (full) DP**, regardless of its internal structure, (19). 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 (1a), (15a), and (18a)).

(19) a. *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.'
(✓: …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})

- Beyond monotransitives, SJPM has a couple of verbs that participate in **double object constructions** involving an instrument IO, producing V(P)-S-IO-DO word order, (20).[12]
  - I assume that these are low applicatives (e.g. Pylkkänen, 2008), and that both IO and DO undergo object shift to individual specifiers of Spec-VoiceP prior to remnant *v*P-fronting (Lee, 2006).[13]

[12] Other ditransitives are formed with PP indirect objects, yielding V(P)-S-PP-DO word order, so they are less informative (though see §3.2).

[13] The output of object shift is moreover order preserving, since the IO necessarily precedes the DO. See Medeiros 2013 for an analysis that captures this idea in another Hawaiian, another VP-S-O language involving object shift and remnant movement.

(20) kõ^{13}õ^{3} [$_{S}$ =ɾa^{1}] [$_{IO}$ li^{5}bro^{6}] [$_{DO}$ sa^{1}ta^{3} ti^{5}=^{n}di^{3}la^{1}^{n}ka^{1}]
COMPL.hit.with =3SG.M book back 3.AN=mosquito
'He hit the mosquito('s back) with a book.'

- Importantly, clitic coalescence is again permitted between two internal arguments if the target is a pronominal clitic, (21a), **but not if it is unambiguously a full DP, (21b-c)**.[14]

(21) a. kõ^{3}õ^{3}=^{n}di^{1}=t^{w}=**ã^{5}**
IRR.hit.with=1PL.EX=3.WOOD=3SG.F
'We'll hit her with it (a wooden object).' (V=S$_{Cl}$=IO$_{Cl}$=DO$_{Cl}$)
(also ✓: kõ^{3}õ^{3}=^{n}di^{1}=tõ^{5}=**ɲa^{5}**)

b. *kõ^{13}õ^{3}=ɾa^{1} ɲa^{1}=ʒu^{1}u^{5} ʒo^{5ʔ}^{w}=**a^{5}**
COMPL.hit.with=3SG.M 3.N=rock DEM.PROX=3SG.F
Intended: 'He hit her with this rock.' (V=S$_{Cl}$ IO=DO$_{Cl}$)
(✓: ...ɲa^{1}=ʒu^{1}u^{5} ʒo^{5ʔ}o^{3}=**ɲa^{5}**)

c. *kõ^{13}õ^{3}=ɾa^{1} k^{w}a^{1ʔ}a^{3} ʒi^{1}=**a^{5}**
COMPL.hit.with=3SG.M many rock=3SG.F
Intended: 'He hit her with many rocks.' (V=S$_{Cl}$ IO=DO$_{Cl}$)
(✓: ...k^{w}a^{1ʔ}a^{3} ʒu^{1}u^{5}=**ɲa^{5}**)

- Finally, coalescence may take place **if the preceding DP is Ā-extracted, (22), or otherwise null, (23)**—in these cases, the object clitic simply coalesces with the element linearly preceding it.

(22) a. ^{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?' (S V <S> O)
(also ✓: ^{n}dʒa^{5} ku^{5}u^{3}=na^{1} ʃi^{13}ni^{31} **ɲa^{5}**=Pa^{3}lo^{3}ma^{5})

b. ni^{3}=ĩĩ^{3}=ti^{5} ko^{15}-tu^{3}v^{j}=**a^{5}**
NEG=one=3.AN NEG.COMPL-sting=3SG.F
'None of them (e.g. wasps) stung her.' (S V <S> O)
(also ✓: ...ko^{15}-tu^{3}vi^{3}=**ɲa^{5}**)

(23) a. no^{3}m^{j}=**a^{5}**
IRR.hug=3SG.F
'Hug her!'
(also ✓: no^{3}mi^{3}=**ɲa^{5}**)

b. [$_{vP}$ ka^{3}tõ^{5} t**i^{3}**k^{j}]=**ã^{3}**
IRR.tie again=3.N
'Tie it again!'
(also ✓: ...tu^{3}ku^{3}=**ɲa^{3}**)

- This rules out an alternative wherein the restriction is inherent to determiner/pronominal clitics in certain structural positions, e.g., object position.[15]

[14]Here, the IO is a complex DP, as opposed to a simple nominal with just a determiner, as my co-author CJC generally disprefers these in IO position (for unclear reasons), regardless of the choice of DO.

[15]For instance, Kurtz (2024) argues for Galician that contracted forms of determiners reflect the occurrence of structural case licensing. This cannot be at play in SJPM, since we would not expect the case licensing of one DP to be impacted by Ā-extraction of a different DP in the clause. However, we could also ask whether the Galician data presented by Kurtz are compatible with the phasal account advanced here.

Interim conclusions:

- The clitic coalescence restriction is not due to purely syntactic factors. Otherwise, coalescence would not be able to ignore null but syntactically present material.
- But it is not purely phonological (e.g., prosodic) either.
  - For instance, under accounts like Match Theory (Selkirk, 1995, 2011), units of a particular prosodic category are not distinguishable by syntactic category (e.g., $X^0 = \omega$; XP = Φ).
  - But syntactic category does matter here, since phrasal elements (Φs) that are **not DPs** do tolerate clitic coalescence into them.

4 DPs vs. non-DP nominals

- This section adds the following:
  1. Some nominal arguments are in fact licit targets for clitic coalescence—if they are **structurally reduced (i.e., not DPs)**.
  2. **Structurally reduced ≠ prosodically reduced**, as seen by comparing strong, weak, and clitic pronouns.
  - (And see the Appendix for [pseudo] noun incorporated objects!)

4.1 DPs vs. reduced nominals

- Although nominal arguments often surface with an overt determiner/classifier, SJPM (like many languages) permits **determinerless internal arguments (DOs and IOs)**.
  - Importantly, these internal arguments are not pseudo incorporated (i.e., still appear in V(P)-S-IO-DO word order), though they are usually interpreted as indefinite, (24).[16]

[16]Determinerless internal arguments are sometimes translated into Spanish by C.J.C. and other speakers with a definite article, but there seem to be some information structure correlates (object backgrounding?) that we haven't fully explored.

(24) a. 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.')

b. ʃa^{13}ʃ=i^{51} **ti^{1}ko^{3}o^{15} u^{1ʔ}va^{1}**
COMPL.eat=1SG tamale salty
'I ate a salty tamale.'

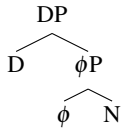
c. tu^{13}ʃi^{5}=ɾa^{1} **la^{5}pi^{6}=ɲa^{4}**
COMPL.poke.with=3SG.M pencil=3SG.F
'He poked her with a pencil.'

- **Since determinerless/indefinite objects are often analyzed as smaller than full DPs, we may ask whether they are licit targets for clitic coalescence in SJPM.**
- Double object constructions are particularly useful, since they permit sequences of internal arguments where the first (the IO) is determinerless/indefinite.
  - Recall from (21) that direct object clitics cannot coalesce with a preceding DP instrument.
  - Now we can see that this <u>is</u> possible if the preceding instrument is determinerless/indefinite, (25):

(25) a. kõ^{13}õ^{3}=ɾa^{1} ʒi^{1}=**a^{5}**
COMPL.hit.with=3SG.M rock=3SG.F
'He hit her with a rock.' (V=S$_{Cl}$ IO=DO$_{Cl}$)
(also ✓: ...ʒu^{1}u^{5}=**ɲa^{5}**)

b. kõ^{13}õ^{3}=na^{1} i^{3}t^{w}=**ã^{5}**
COMPL.hit.with=3PL.N stick=3SG.F
'They hit her with a stick.' (V=S$_{Cl}$ IO=DO$_{Cl}$)
(also ✓: ...i^{3}tõ^{5}=**ɲa^{5}**)

- **Proposal:** These nominals are indeed structurally smaller than DPs, explaining both their indefinite interpretation and licitness as coalescence hosts.
  - For concreteness, I model them as ϕPs in the sense of Dechaine and Wiltschko 2002, (26).
  - The ϕP layer is to differentiate these nominals from (by hypothesis) even smaller ones (e.g., N(P)s) that undergo (pseudo) noun incorporation (**Appendix**).

(26) a.


b.


- These data more broadly show that the clitic coalescence restriction is not tied to all nominal arguments—but rather **nominal arguments of a particular syntactic size/category**.

4.2 Evidence from non-clitic pronouns

- At this point, however, we cannot rule out the restriction being keyed to **nominals of a certain <u>prosodic</u> size**.
  - DPs generally happen to be prosodically more complex than their determinerless counterparts.
- **Non-clitic pronouns** allow us to disambiguate between syntactic vs. prosodic size.
- Like certain other Mixtec varieties (Macaulay, 1987), SJPM has a three-way distinction between strong pronouns, weak pronouns, and pronominal clitics.
  - Strong pronouns are bimoraic, while weak pronouns are monomoraic.
  - Note that this three-way distinction exists only in the 1st/2nd person (3rd person pronouns are clitics only).

	1SG	2SG	1PL.IN	1PL.EX	2PL
Strong pron.	ʒu^{1ʔ}u^{1}	ʒo^{5ʔ}o^{1}	ʒo^{5}o^{5}	^{n}du^{1ʔ}u^{1}	^{n}do^{5ʔ}o^{1}
Weak pron.	ʒu^{1}	ʒo^{51}	ʒo^{5}	^{n}du^{1}	^{n}do^{51}
Clitic	=e^{1}	=õ^{5}	=e^{5}	=^{n}di$^{(1)}$	=^{n}do^{5}

- **Weak pronouns and clitics form a <u>prosodic</u> class:** Both are monomoraic (= prosodically deficient), and accordingly cannot be stressed, uttered in isolation (e.g., in fragment answers), etc..
- **But strong and weak pronouns form a <u>morphosyntactic</u> class:** Independently evident given the Person-Case Constraint (PCC).
  - SJPM displays (a version of) the strong PCC, so 1st/2nd person clitics are banned in DO position, (27a) (Yuan and Juárez Chávez, 2025).
  - As in many other languages, PCC-violating configurations are 'repaired' by expressing the lower argument (the DO) as a non-clitic (e.g. Bonet, 2008; Rezac, 2011). In SJPM, this could be a strong <u>or</u> weak pronoun, (27b).[17]

(27) a. *ʃi^{13}ni^{31}=**ɲa^{5}**=**^{n}di^{3}**
COMPL.see=3SG.F=1PL.EX
Intended: 'She saw us (ex.).'

[17] Macaulay (1987) draws the same conclusion based on Chalcatongo Mixtec, though she does not describe the relevant constructions as related to the PCC.

b. ʃi^{13}ni^{31}=**ɲa^{5}** **^{n}du^{1ʔ}u^{1}** / **^{n}du^{1}**
COMPL.see=3SG.F 1PL.EX(STR) 1PL.EX(WK)
'She saw us (ex.).'

- **Assumption:** Pronominal clitics are D^{0}s, but strong and weak pronouns are both DPs (weak pronouns may just be phonologically reduced strong pronouns; see Eischens 2025).

- **Crucially, strong and weak pronouns also pattern together wrt clitic coalescence:** Unlike pronominal clitics, neither strong nor weak pronouns are licit targets for coalescence, (28) (cf. (18a)).

(28) a. *ka^{3}tõ^{5} **^{n}du^{1ʔw}=a^{1}**
IRR.tie 1PL.EX(STR)=3.N
Intended: 'WE'LL tie it.'
(✓: ka^{3}tõ^{5} **^{n}du^{1ʔ}u^{1}=ɲa^{1}**)

b. *ka^{3}tõ^{5} **^{n}d^{w}=a^{1}**
IRR.tie 1PL.EX(WK)=3.N
Intended: 'We'll tie it.'
(✓: ka^{3}tõ^{5} **^{n}du^{1}=ɲa^{1}**)

- This asymmetry reveals that the clitic coalescence restriction is **insensitive to the prosodic size** of the preceding nominal argument.

- Illicit targets for coalescence are both of category D and syntactically phrasal (i.e., DPs).

5 A phasal analysis

5.1 No coalescence across spell-out domains

- **Phases (a reminder):** Syntactic structures are built in chunks, and are sent (spelled-out) phase-by-phase to the other components of the grammar ("derivation by phase").

- Much work on the syntax-phonology of phases has argued that **cyclic spell-out of phasal material delimits phonological domains**.

  - The recurring observation: Certain morphological/phonological processes fail to apply across certain boundaries that are independently identifiable as phasal in the syntax.
  - See Marvin 2003, Newell and Piggott 2014, Pak 2008, Kahnemuyipour 2009, Kramer 2010, McPherson and Heath 2016, Sande 2018, Talić 2018, Guekguezian 2021, Fenger and Weisser 2025, van Urk 2025, a.o.

- **Two brief illustrations:**

  1. In Ojibwe, complex verbs may be formed across a clause medial phase (= the highest *v*P-level domain), but this boundary blocks vowel hiatus resolution (Newell and Piggott, 2014).

(29) a. aj**a:-a**maw → aj**a:**maw
get-APPL
'get x for y'
(V^0 and $Appl^0$ in same phase)

b. g**i:-a**noki: → g**i:a**noki:
PST-work
'he worked, hunted'
(V^0 and T^0 in different phases)

  2. In Tommo So, L-tone overlays are triggered by certain words (e.g., *nɔ́* 'this') and target the triggers' c-command domains, but cannot target material inside DP phases (McPherson and Heath, 2016).

(30) a. [$_{DP}$ [$_{PossP}$ [$_{DP}$ Sáná] gámmá gém] nɔ́] → Sáná **gàmmà gèm** nɔ́
Sana cat black this
'this black cat of Sana's'

b. [$_{DP}$ [$_{PossP}$ gámmá [$_{DP}$ ḿmɔ] nɔ́] → **gàmmà** ḿmɔ nɔ́
cat mine this
'this cat of mine'

- **Proposal:** The SJPM clitic coalescence restriction is likewise due to the phasal status of DPs—**coalescence cannot take place between elements in different spell-out domains**.

Some necessary assumptions:

- Besides DP (and VoiceP and CP, by assumption—though see §5.2 below), **no other XPs are phases**.
- The spell-out unit is the **phasal complement** (Chomsky 2001, *pace* Boškovič 2016), i.e., excludes the phase head.
- Spell-out of a phasal complement takes place **as soon as the phase is built**, i.e., is not delayed until the construction of the next phase.

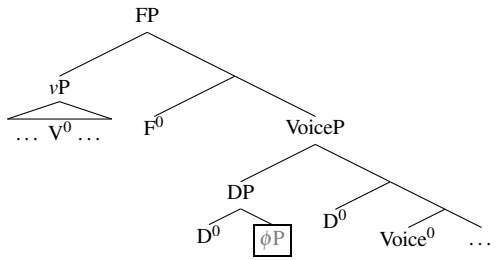
→ Together, these capture the following observations:

  - the **category restriction**—DPs vs. non-DPs (including other nominals)
  - the **directional asymmetry**—why determiner D^0s may coalesce (leftward) out of their containing DPs, but DPs do not allow rightward coalescence into them
  - the **DP vs. D^0 contrast**—clitic D^0s are not phasal because they don't have complements and don't project

How this works:

- Given the clause structure established in §2.2 (ex. (6)), the verb and its DP arguments are all **at the VoiceP edge or higher**, so they are all in the same spell-out domain (= the CP phase).
- But because DPs are phases, the complement of any D^0 is spelled out separately from the above elements.

(31)



- Postsyntactically, the clitics adjoin to their hosts. The relevant clitic exponents are later inserted; inserting either a CV or V allomorph is in principle possible.
  - If a CV allomorph is inserted, then it is ultimately realized within a recursive prosodic word with its host (§2.3).

(32) ko^{3}ni^{31} ti^{5}=vi^{3}lu^{5}=**ŋa^{3}**
IRR.see 3.AN=cat=3.N
'The cat will see it.'
…(vi^{3}lu^{5})$_{\omega}$=ŋa^{3})$_{\omega}$

  - If a V allomorph is inserted, then it should be realized as an internal clitic within the minimal prosodic word, which would then feed coalescence (to resolve vowel hiatus) (§2.3).
  - **This is blocked if the clitic's host has already been spelled out.** Spelled-out material is **immutable** for the tonal/segmental changes involved in coalescence.

(33) *ko^{3}ni^{31} ti^{5}=vi^{3}ɨ^{j}=**a^{5}**
IRR.see 3.AN=cat=3.N
Intended: 'The cat will see it.'
Not possible: …(vi^{3}lu^{5}=a^{3})$_{\omega}$ → [vi^{3}ɨ^{j}a^{5}]

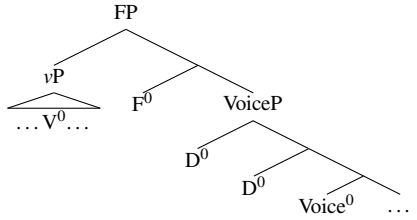
- It is equally impossible for a V clitic to surface without triggering coalescence (e.g., *[vi^{3}lu^{3}a^{5}])—as vowel hiatus within prosodic words is not tolerated in the language.[18]

[18]Tolerance of vowel hiatus in these contexts may be a point of cross-Mixtec variation: Mantenuto 2020 provides data from San Sebastián del Monte Mixtec suggesting (at least orthographically) that vocalic clitics do not coalesce with their hosts.

- In contrast to full DPs, pronominal clitics (bare D^0s) **lack complements altogether** (i.e., spell-out is not triggered). As a result, they are licit hosts for clitic coalescence:

(34) a. 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!'
…(^{n}dʒa^{3}a^{31})$_{\omega}$=tõ^{5}=a^{5})$_{\omega}$ → [^{n}dʒa^{3}a^{31}t^{w}ã^{5}]

b.



- Finally, in all other contexts in Table (13), coalescence is possible because none of these other elements are spelled out separately from a linearly following determiner/pronominal clitic.

5.2 Other phasal domains?

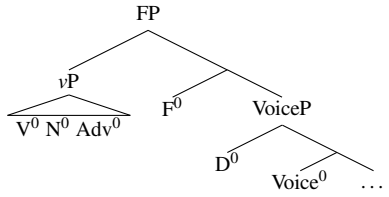
- Given this phase-based analysis, we may expect other posited phase boundaries (e.g., CPs, VoicePs) to also prevent coalescence across thems. **However, these latter configurations are not informative either way.**

1. It is impossible to test clitic coalescence across embedded CPs, as they are necessarily complementizer- or verb-initial.
   - As shown in (35), we simply cannot place a DP in Spec-CP and see whether D^0 can coalesce with a higher element in the matrix clause.

(35) *kã^{5ʔ}ã^{1}=ŋa^{5} [$_{CP}$ **ŋa^{5}=Pa^{3}lo^{3}ma^{5}** a^{5} sa^{5}ka^{3}nã^{5ʔ}ã^{1}=ŋa^{5}]
CONT.think=3SG.F 3SG.F=Paloma C IRR.win=3SG.F
Intended: 'She thinks that Paloma will win.'
(✓: kã^{5ʔ}ã^{1}=ŋa^{5} [$_{CP}$ a^{5} sa^{5}ka^{3}nã^{5ʔ}ã^{1} **ŋa^{5}=Pa^{3}lo^{3}ma^{5}**])

2. Whether there is a clause-medial phase (=VoiceP) to begin with is somewhat contentious (e.g. Keine and Zeijlstra, 2025), but let us suppose that it is.
   - In §2.2, we posited that Voice^{0}'s complement is *v*P. Under the current analysis, *v*P should be opaque—but we have seen that *v*P is a licit target for clitic coalescence, repeated in (36).

(36) a. [$_{vP}$ tu^{5}tu^5 ʒu^{3ʔ}u^5 **ti^5**k^j]=**ã^5**
CONT.whistle mouth again=3SG.F
'She's whistling again.'

b.


- But data like (36) are not problematic for the current analysis, since there are various workarounds (that are hard to disprove).
  - → For instance, the fact that *v*P is able to raise out of the VoiceP domain at all (perhaps transiting through Spec-VoiceP) suggests that it may evade spell-out.
  - → Alternatively, we could also posit that the relevant constituent that moves is not Voice^0's complement, but some smaller XP.

The present phasal account is thus based entirely on the behavior of DPs.

- Is this a weakness of the account? Possibly…
- But it seems empirically and explanatorily superior than any non-phasal alternatives that we can think of…one example given below.

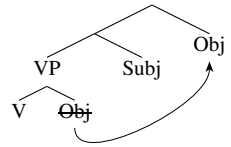
5.3 Against a non-phasal alternative

- Suppose instead that DPs have a left vs. right edge asymmetry, where the right edge is subject to some strong prosodic boundary, (37).

(37) V DP$_S$) DP$_O$)

- This would also prevent clitic coalescence into a preceding DP, while allowing a determiner clitic to coalesce out of a containing DP.
- Palancar (2004) offers a generalization along these lines for San Ildefonso Tultepec Otomi (Western Otomanguean), distantly related to Mixtec.

- But without deeper explanation, this simply restates the empirical generalization—so it is (by itself) perhaps not satisfying.
- **Moreover, applying existing motivations for the prosodic profile in (37) yields the wrong empirical picture for SJPM.**
  - For instance, Clemens 2021 proposes that prosodic structure may diagnose syntactic structure, and that a prosodic phrasing as in (37) would reflect a structure like (38).
  - in (38), both the subject and object are projected as rightward specifiers, and the object additionally raises <u>above</u> the subject (cf. Little 2020 on Chol [Mayan]). But there is no evidence internal to SJPM for either rightward specifiers or O > S inversion.[19]

(38)


In contrast, the phasal account has the following advantages:

- It provides an independent explanation for why DPs are distinct from various other XPs (including reduced nominal constituents).
- It derives the left vs. right edge asymmetry from independent properties of the language (e.g., head-initiality).
- It draws parallels with other phase-constrained processes cross-linguistically (involving DPs or other phasal XPs).

6 Cliticization and phase (im)penetrability

- Much prior work on word-internal phase effects has **taken for granted** that complex words may be formed across phase boundaries (e.g., Newell and Piggott 2014, Fenger and Weisser 2025, van Urk 2025).
  - Recall for instance (39), from Ojibwe (cf. (29)):

(39) g**i:**-**a**noki: → g**i:a**noki:
PST-work
'he worked, hunted' (V^0 and T^0 in different phases)

- This is not a universal assumption! Kramer (2010) argues (for Amharic) that certain kinds of word-formation operations may not apply across phases.[20]

[19] Per Clemens (2021), a VSO language with a VP-fronting derivation (as proposed for SJPM above) should yield a (V)((S)(O)) prosodic phrasing, all things being equal (that is, without the inclusion of additional constraints that may result in a different prosodic phrasing). This would not account for the distribution of the coalescence restrictions in SJPM.

[20] And **?** propose for Inuit languages that individual complex words = individual syntactic phases.

- On the basis of Bosnian-Serbian-Croatian clitic patterns, Talić (2018) argues that **minimal prosodic words** cannot be formed across spell-out domains, (40a).
    - Implicit in this claim (though not discussed explicitly by Talić): non-minimal prosodic words may be formed across such domains, as in (40b).

(40) For two morphemes X and Y originating in different spell-out domains:
a. ✗$(X\ Y)_{\omega}$
b. ✓$((X)_{\omega}\ Y)_{\omega}$

- **Interestingly, this contrast between minimal and non-minimal word-formation across phases is apparent in SJPM!**
    - We have already seen that vocalic clitics form minimal prosodic words with their hosts, but may not do so across a DP phase boundary.
    - In contrast, CV clitics can encliticize to DPs—but they form recursive prosodic words, (41).
    - Recall from §2.3, ex. (9): recursive woodhood diagnosable by L-tone spreading.

(41) a. ko^{3}ni^{31} [$_{DP}$ ti^{5}=vi^{3}lu^{5}]=**ɲa^{3}**
IRR.see 3.AN=cat =3.N
'The cat will see it.'
$\ldots(vi^{3}lu^{5})_{\omega}$=ɲa$^{3})_{\omega}$

b. ko^{3}ni^{31} [$_{DP}$ ti^{5}=k^{w}a^{5}**ʒu^{1}**]=**ɲa^{1}**
IRR.see 3.AN=horse =3.N
'The horse will see it.'
$\ldots(k^{w}a^{5}ʒu^{1})_{\omega}$=ɲa$^{1})_{\omega}$

General observations:

- The occurrence of L-tone spreading suggests that spelled-out material is not invisible, per se—simply **immutable** (e.g., subject to extra faithfulness, as in McPherson and Heath 2016).

- As mentioned, phase (im)penetrability appears to be sensitive to the **prosodic level** of cliticization—an idea that has not really been explored in the literature.
    - How to reconcile this with data like (39)? Perhaps it is worth (re)assessing how such words are formed…
    - e.g., is *gi:-* in (39) really an affix, and does affixation only form minimal prosodic words?

7 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, this talk has alluded to the diversity of clitic coalescence patterns within the Mixtec language family.
    - A broader goal is 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, these findings significantly expand on existing work on Mixtec clitic coalescence, which has not considered the role of syntax.

References

Bonet, Eulàlia. 2008. The Person-Case Constraint and repair strategies. In *Agreement restrictions*, ed. Roberta D'Alessandro, Susann Fischer, and Gunnar H. Hrafnbjargarson, 103–128. Berlin: De Gruyter.

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

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.

Compton, Richard, and Christine Pittman. 2010. Word-formation by phase in Inuit. *Lingua* 120:2167–2192.

Coon, Jessica. 2010. Complementation in Chol: A theory of split ergativity. Doctoral Dissertation, Massachusetts Institute of Technology.

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

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

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.

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

Eischens, Ben. 2025. Laryngeal reduction and mora deletion in Mixtec: Phonetics in phonology. *Phonological Data and Analysis* 7.

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.

Guekguezian, Peter Ara. 2021. Aspectual phase heads in Muskogee verbs. *Natural Language and Linguistic Theory* 39:1129–1172.

Hedding, Andrew. to appear. On the representation of wh-words and foci: Evidence from Mixtec. *Linguistic Inquiry* .

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

Hedding, Andrew A., and Michelle Yuan. 2025a. Distinct pathways to possessor Ā-extraction in Mesoamerican languages. In *Proceedings of NELS 55*, ed. Duygu Demiray, Roger Cheng yen Liu, and Nir Segal, 261–274. Amherst, MA: GLSA.

Hedding, Andrew A., and Michelle Yuan. 2025b. Phase unlocking and the derivation of verb-initiality in San Martín Peras Mixtec. In *Proceedings of WCCFL 41*, ed. Nikolas Webster, Yağmur Kiper, Richard Wang, and Sichen Larry Lyu, 184–193. Somerville, MA: Cascadilla Press.

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.

Keine, Stefan, and Hedde Zeijlstra. 2025. Clause-internal successive cyclicity: Phasality or DP intervention? *Natural Language and Linguistic Theory* 43:1119–1182.

Kramer, Ruth. 2010. The Amharic definite marker and the syntax–morphology interface. *Syntax* 13:196–240.

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

Kurtz, Naomi. 2024. On the loci and timing of case-licensing: Insights from microvariation in Galician. Doctoral Dissertation, University of Chicago, Chicago, IL.

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

Lim, Jun Jie. 2025. (A)symmetries symmetries in the morphosyntax of embedded clauses in Khalkha Mongolian. Doctoral Dissertation, University of California San Diego, La Jolla, CA.

Lim, Jun Jie. 2026. Small nominalized clauses in Khalkha Mongolian: Asymmetries in mixed extended projections. *Natural Language and Linguistic Theory* 44.

Little, Carol Rose. 2020. Mutual dependencies of nominal and clausal syntax in Ch'ol. Doctoral Dissertation, Cornell University, Ithaca, NY.

Lyutikova, Ekaterina, and Asya Pereltsvaig. 2015. The Tatar DP. *Canadian Journal of Linguistics* 60:289–325.

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.

Medeiros, David. 2013. Hawaiian VP-remnant movement: A Cyclic Linearization approach. *Lingua* 127:72–97.

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, 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.

Ostrove, Jason. 2026. Obligatorily overt PRO in San Martín Peras Mixtec. *Linguistic Inquiry* 1:1–48.

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.

Rezac, Milan. 2011. *Phi-features and the modular architecture of language*. Dordrecht: Springer.

Richards, Norvin. 2001. *Movement in language: Interactions and architectures*. New York: Oxford University Press.

Roberts, Ian. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals*. Cambridge, MA: MIT Press.

Royer, Justin. 2019. Domain restriction and noun classifiers in Chuj. In *Proceedings of NELS 49*, ed. Maggie Baird and Jonathan Pesetsky, 87–97. Amherst, MA: GLSA.

Royer, Justin. 2022. Elements of (in)definiteness and binding: A Mayan perspective. Doctoral Dissertation, McGill University, Montreal, QC.

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.

Simpson, Andrew, and Zoe Wu. 2002. IP-raising, tone sandhi and the creation of S-final particles: evidence for cyclic Spell-Out. *Journal of East Asian Linguistics* 11:67–99.

Talić, Aida. 2018. Spelling out enclitics and giving their tone a voice: Cyclic clitic incorporation in BCS and breaking the cycle. *The Linguistic Review* 35:307–370.

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. 2024. Constraining VP-fronting. *Linguistic Inquiry* 55:327–373.

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 data from (pseudo) noun incorporation

- Like many other V(P)-S-O languages (e.g. Massam, 2001; Coon, 2010; Medeiros, 2013; van Urk, 2025), V-O-S word order arises in SJPM if a structurally reduced object remains inside the fronted *v*P.
- In SJPM, this is somewhat restricted to particular constructions:
  1. V-N compounds, e.g., (16a).
  2. Constructions containing reciprocal objects, (42) (I assume that reciprocals are bare Ns too).

(42) a. no^{5}mi^{3}=ra^{1} **na^{1}=va^{5}li^{3}**
CONT.hug=3SG.M 3PL.N=child.PL
'He's hugging the children.' (V=S_{Cl} O)

b. [$_{vP}$ 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})

- Interestingly, complex XPs containing reciprocal objects must also remain in the *v*P, (43).[21]

(43) a. k^{w}a^{1ʔ}a^{3} ɲa^{5}=Pa^{3}lo^{3}ma^{51} i^{3}ta^{31} [$_{PP}$ **^{n}da^{3ʔ}a^{5} ɲa^{5}=Ga^{5}bi^{6}**]
IRR.give 3SG.F=Paloma flower to 3SG.F=Gaby
'Paloma will give flowers to Gaby.' (V S O PP)

b. [$_{vP}$ k^{w}a^{1ʔ}a^{3} [$_{PP}$ **^{n}da^{3ʔ}a^{5} tã^{5ʔ}ã^{3}**]]=na^{5} i^{3}ta^{3}
IRR.give to each.other=3PL.F flower
'They will give flowers to each other.' (V PP_{Rec}=S_{Cl} O)

- And if the complex XP contains additional nouns, those nouns also remain inside the *v*P.[22]

(44) a. [$_{vP}$ ^{n}dʒa^{5}a^{6} [$_{NP?}$ **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 Pssm $Pssr_{Rec}$=S_{Cl})

b. [$_{vP}$ 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})

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

[22] In (44a), the inability for the possessum to undergo object shift without the reciprocal possessor is not surprising, since possessor stranding is otherwise not permitted in SJPM. Note that the label of the PNI-ed nominal constituent in (44a) as NP rather than DP is admittedly a stipulation (in that it is required based on the phasal analysis of DPs), but here I follow Lyutikova and Pereltsvaig 2015 and Lim 2025, 2026 on the idea that possessive nominal constituents may be DPs or smaller (NPs), as signaled (among other things) by the interpretation of the possessive construction. That being said, the pattern in (44b) is genuinely surprising, since applied arguments are otherwise able to undergo object shift. That this genuinely involves PNI of both IO and DO is further demonstrated in (i), which shows that a *v*P-internal right-adjoined adverb may follow both nominals. I do not have an account of these latter data, but the generalization seems to be that the inability to move reciprocal objects may also trap other nominals within the containing XP. See also §4.1 for more discussion.

(i) [$_{vP}$ kõ^{3}õ^{3} [$_{ApplP}$ ʒu^{1}u^{5} tã^{5}ʔã^{3}] tu^{3}k^{w}]=**i^{5}**
IRR.hit rock each.other again=1PL.IN
'We'll hit each other with rocks again.' (V IO O_{Rec} Adv=S_{Cl})

- Crucially, clitic coalescence is permitted into such fronted constituents as well. Because reciprocals require a plural antecedent, this is illustrated in (45) with 1PL.IN =e^5.

(45) 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. [$_{vP}$ ʃa^{1ʔ}a^3 [$_{PP}$ ^{n}da^{3ʔ}a^5 tã^{5ʔ}]]=**ẽ^5** i^{3}ta^3
COMPL.give to each.other=1PL.IN flower
'We gave flowers to each other.' (V PP_{Rec}=S_{Cl} O)

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

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

- Note that PNI of an applied argument (co-occurring with a reciprocal DO) is possible only if it is bare, but not if it is unambiguously a DP:

(46) a. [$_{vP}$ tu^{13}ʃi^5 [$_{ApplP}$ **la^{5}pi^6 tã^{5ʔ}ã^3**]]=na^3
COMPL.poke.with pencil each.other=3PL.N
'They poked each other with pencils.' (V IO O_{Rec}=S_{Cl})

b. *[$_{vP}$ tu^{13}ʃi^5 [$_{ApplP}$ [$_{DP}$ **tõ^5=la^{5}pi^6** **ʒo^{5ʔ}o^3**]
COMPL.poke.with 3.WOOD=pencil DEM.PROX
tã^{5ʔ}ã^3]]=na^3
each.other=3PL.N
Intended: 'They poked each other with these pencils.'

- Although it is puzzling why (45d) and (46a) arise to begin with (fn. 22), the above contrast is again consistent with the idea that internal arguments may come in different sizes, and that this impacts the availability of clitic coalescence.