Phonological Data & Analysis Volume 1, Article 3: x–xx (xxxx)
https://doi.org/00.0000/nn.0.00

Received 17 January 2018; revised 5 May 2018; accepted 28 January 2019.

Lexical and grammatical tone in San Juan Piñas Mixtec (Tò’ōn Ndá’ví)

Gabriela Caballero[a], Claudia Duarte Borquez[b], Claudia Juárez Chávez[c] & Michelle Yuan[d*]

[a]University of California San Diego – gcaballero@ucsd.edu
[b]University of California San Diego – cduarteb@ucsd.edu
[c]University of California San Diego – cjchavez@ucsd.edu
[d]University of California, Los Angeles – yuanm@ucla.edu

This paper presents the first detailed investigation of the lexical and grammatical tonal patterns of San Juan Piñas Mixtec (Tò’ōn Ndá’ví), a previously undocumented language variety spoken in Oaxaca, Mexico and diaspora communities in the U.S. We show that SJPM exhibits a high number of contrastive tone patterns in stems, resulting from the concatenation of three tonal primitives (H, L and M). The lexical tone system of this language variety also involves a typologically unusual contrast between underlyingly specified M-toned tone-bearing units (TBUs) and underlyingly toneless /Ø/ TBUs. The lexical tone patterns in SJPM moreover exhibit complex interactions with grammatical tone, a system analyzed through the association of morphosyntactically-triggered tonal melodies (L, H and LH), which may replace or accommodate lexical tone. We analyze grammatical tone in this language as *non-dominant neutral* within a dominance theory of grammatical tone (Rolle 2018). The asymmetries observed in lexical-grammatical tone interactions result from the avoidance of phonotactically marked tonal sequences. In these cases, we argue that lexical tonal contrasts are sacrificed in order to preserve the morphological contrasts in paradigms if the phonological grammar prevents concatenation of grammatical and lexical tones.

Keywords: tone; grammatical tone; Oto-Manguean; Mixtec

1 Introduction

This paper presents the first detailed investigation of the tonal patterns of the previously undocumented San Juan Piñas variety of Mixtec (henceforth SJPM; also known as Tò’ōn Ndá’ví), spoken in the Santiago Juxtlahuaca municipality of Oaxaca, Mexico as well as in diaspora communities in Mexico and the United States. We provide a description and analysis of the lexical and grammatical tonal systems of this language, how these systems interact with each other, and how they interact with other grammatical aspects of the

* We would like to express our gratitude to our San Juan Piñas Mixtec language teachers, Maestras Claudia Juárez Chávez, Basi Juárez Chávez, Cirila Pérez Morelos, Basilisa Pérez Morelos, and their families for teaching us about the SJPM language and culture. We also would like to thank two anonymous reviewers for their helpful feedback, as well as Francisco Arellanes, Christian DiCanio, Ben Eischens, Marc Garellek, Sharon Rose, and Hiroto Uchihara for useful comments and suggestions. Lastly, we would like to thank audiences at UCSD, UCSC, UC Berkeley, Uni Leipzig, and Universidad Nacional Autónoma de México, as well as at WCCFL 39, the 2021 Princeton Phonology Forum, WSCLA 26, and the 21st *Réseau Français de Phonologie*.

language. We show that SJPM offers a number of new insights into tonal variation within Mixtec languages, and contributes to our broader typological and theoretical understanding of tone, grammatical tone systems, and lexical-grammatical tone interactions.

A long tradition of documentation of tonal systems of Mixtec language varieties, spanning several decades since Pike's (1944, 1948) seminal work, reveals a great amount of complexity and diversity: typically analyzed as involving a tripartite tone contrast, most Mixtec languages permit a greater number of tones on individual syllables compared to other Oto-Manguean languages (see DiCanio & Bennett (2021) for an overview). These languages also exhibit complex tone sandhi systems, but individual varieties differ in terms of the nature and number of tone sandhi patterns attested (Pike 1948, Pankratz & Pike 1967, Josserand 1983, Dürr 1987, Daly & Hyman 2007). We propose that the tonal system of SJPM is best described with three underlying tonal primitives, a H tone, a L tone and a M tone, which may surface as level tones or in rising contour tones associated on single morae in complex lexical tone patterns. In addition, SJPM exhibits a high number of tonal processes, some of which we detail in this paper. While other Mixtec varieties are analyzed with all surface M tones as underspecified underlyingly as /Ø/ (Daly & Hyman 2007, Carroll 2015), we propose that SJPM possesses a contrast between underlyingly specified M-toned tone-bearing units (TBUs) and underlyingly toneless /Ø/ TBUs that receive a M tone by default or by tone spreading. Tonal systems with a /M/ vs. /Ø/ contrast are typologically highly unusual, but have also been posited for genetically related languages such as Nochixtlán Mixtec (McKendry 2013, 2018) and Sierra Juárez Zapotec (Bickmore & Broadwell 1998), as well as genetically unrelated languages such as Poko Rawo (McPherson & Dryer 2021). With three tone contrasts realized in all positions and only few TBUs that are underlyingly toneless, SJPM exhibits a high degree of tonal density (as defined by Gussenhoven (2004) and Hyman (2009)): the proportion of underlyingly specified (H, L or M) TBUs is high compared to unspecified (/Ø/) TBUs in this language.

The lexical tone patterns in SJPM moreover exhibit complex interactions with grammatical tone, which is used to encode tense[1] and negative polarity on the verb, sometimes in conjunction with segmental morphology. Tone also plays a role in pronominal marking (via fusion of vocalic enclitics to verb stems) and in encoding transitivity alternations and denominal adjective derivation, though these later processes are not fully productive in SJPM. Thus, tone carries a high functional load in the morphological system of SJPM. That being said, the grammatical tone systems of Mixtec languages remain understudied (for an overview, see Palancar et al. 2016), so another goal of this paper is to contribute to our broader understanding of such systems. We propose that the SJPM grammatical tone system is best analyzed through the association of floating tones (L, H, or LH), which generally dock to the leftmost TBU of the morpheme within their (syntactic) scope. These morphosyntactically-triggered tonal melodies exhibit asymmetries in terms of their interaction with lexical tone, with grammatical tone either accommodating the lexical tone of the verb stem (forming tonal contours) or replacing it. We analyze tonal replacement/avoidance patterns in lexical-grammatical tone interactions as resulting from dominance properties of individual tonal exponents (as proposed in Rolle 2018). Specifically, we propose that grammatical tone in SJPM may be classified as *non-dominant neutral* within a dominance theory of grammatical tone, involving primarily concatenation of tonal exponents and lexical tones. In cases where the concatenation of grammatical tone to lexical tone would yield tonotactically illicit patterns, however, grammatical tone replaces lexical tone, among other strategies. Thus, the asymmetries observed in lexical-grammatical tone interactions result from the avoidance of phonotactically marked tonal sequences. In these cases, we argue that lexical tonal contrasts are sacrificed in order to preserve the morphological contrasts in paradigms if the phonological grammar prevents keeping both.

[1] In this paper, we cast the relevant TAM distinctions in terms of tense (past and present, respectively), following Hollenbach 2015. However, the relevant distinctions are also often characterized in terms of aspect (e.g. Bickford 1988). We are largely agnostic about this point, pending a more systematic investigation of the semantic properties of tense vs. aspect in both SJPM and in Mixtec languages more broadly.

We also show that the availability of both segmental and tonal exponents of morphosyntactic categories, as well as the tonal phonology properties of the language that allow complex contours in single morae, allow for the preservation of most lexical and grammatical tonal contrasts in SJPM verbal paradigms. Thus, SJPM instantiates a case of a language with both a high lexical role for tone and a complex grammatical tone system—characteristics also found across the Otomanguean family. This is notable in light of Rolle's (2018) prediction that such systems, if they exist at all, should be rare cross-linguistically (see also Grimm 2022 for further discussion).

The remainder of this paper is organized as follows. In §2, we begin with background information on San Juan Piñas Mixtec and the developing corpus on which our description and analysis are based. In §3, we provide an overview of the segmental phonology and basic morphological properties of the language, illustrating primarily with verbs. In §4, we address lexical tone patterns, including general tonotactic constraints operating in the language. In §5, we describe the grammatical tone system of the language and patterns of segmental and tonal exponence in verb paradigms. In §6, we analyze the interaction between lexical and grammatical tone in SJPM within the dominance framework proposed in Rolle 2018 and situate the SJPM grammatical tone system within a developing typology of grammatical tone. We conclude in §7.

2 The language

2.1 Background

The Mixtec languages belong to the Amuzgo-Mixtecan branch of the Otomanguean language family and are spoken in the *Mixteca* region in the Mexican states of Oaxaca, Guerrero, and Puebla (Smith Stark 1995). SJPM is spoken in the Municipality of Santiago Juxtlahuaca, Oaxaca, Mexico, and diaspora communities in California and other communities in the United States and Mexico. There are approximately half a million speakers of all Mixtec varieties in Mexico (INEGI 2020), and between 50,000 and 100,000 speakers of Mixtec in California (Kresge 2007). There is declining vitality of Mixtec language varieties both in Oaxaca and in diaspora communities; however, there are ongoing language strengthening and reclamation efforts on both sides of the border (Campbell & Reyes Basurto 2023).

There are approximately 60 varieties of Mixtec (Eberhard et al. 2023), which Josserand (1983) groups within 12 dialectal areas. SJPM is located within the Southern Baja Mixtec dialect area within this classification. A long tradition of tonal documentation of Mixtec varieties has revealed that there is a high degree of tonal diversity even among closely related varieties (Campbell 2017). Within this line of work, the Southern Baja varieties remain particularly understudied (though see Carroll 2015 on Ixpantepec Nieves Mixtec and Peters 2018 and Eischens 2022 on San Martín Peras Mixtec). This paper thus contributes to our broader understanding of such varieties in its focus on SJPM.

2.2 Methodology

This paper reports on results of research conducted as part of a broader San Juan Piñas Mixtec language project, which aims to provide comprehensive description and documentation of SJPM, as well as develop a language corpus and other resources for language strengthening and reclamation, as designed and overseen by language expert and co-author, Claudia Juárez Chávez. The data presented in this paper primarily consist of linguistic examples obtained through targeted elicitation, although the broader SJPM documentary corpus also features more naturalistic data from different speech genres.[2] We have found that the particular tonal processes investigated in this paper reliably take place word-internally, regardless of

[2] Audio files of the SJPM data provided in this paper can be found in the supplementary materials, accessible here: X [to be updated before copy-editing stage]. These audio clips are publicly available in UC Berkeley's California Language Archive (Collection X). Archival citations for each example are also provided in the supplementary materials.

whether these forms are produced in isolation or in phrasal contexts. We thus take the findings of this paper to serve as baselines for future work investigating other tonal processes in larger prosodic domains.

Statements about frequency made throughout this paper are based on a lexical database categorized by word class and by various prosodic properties, such as the moraic size of stems and whether the forms are glottalized. The major word classes in the sample (which are minimally bimoraic, as discussed in more detail in §3) include 386 bimoraic stems and 102 trimoraic stems (all monomorphemic) in native vocabulary. In addition, our description and analysis of grammatical tone patterns are based on a sample of partial paradigms of 180 verbs within the database. Though not exhaustive, we take this sample of vocabulary and inflected forms as representative.

3 Phonological and morphological preliminaries

3.1 Segmental phonology

The segmental inventory of SJPM involves 19 consonants, five oral vowels, and three nasal vowels (a description of the SJPM segmental inventory is provided in Van Doren, Duarte Borquez, Juárez Chávez & Caballero to appear). The consonant phonemic inventory of SJPM is provided in Table 1.[3]

Table 1: SJPM consonant phoneme inventory

	Bilabial	**Labiodental**	**Alveolar**	**Post-alveolar**	**Palatal**	**Velar**	**Glottal**
Plosive	p		t ^{n}d			k, k^{w} ^{ŋ}g	
Nasal	m		n		ɲ		
Tap/flap			ɾ				
Fricative		f v	s	ʃ ʒ			h
Affricate				t͡ʃ ^{n}d͡ʒ			
Lateral approximant			l				

SJPM exhibits a simple five-vowel system (/i, e, a, o, u/) and contrastive vowel nasalization for three vowels (/ĩ, ã, õ/), as shown in Table 2.

Table 2: SJPM vowel phoneme inventory

	Front	**Central**	**Back**
Close	i ĩ		u
Mid-close	e		o õ
Open		a ã	

There is no phonemic vowel length contrast in SJPM: as described in the next section, root morphemes are minimally bimoraic with a surface contrast between short vowels in disyllabic words and long vowels in monosyllabic words.

3 The phonemes /p/, /f/, and /^{ŋ}g/ are only present in loanwords and do not seem to be independently attested in the native vocabulary. Note also that the consonant inventory in Table 1 does not include a glottal stop /ʔ/: as discussed below, we analyze glottalization as a feature associated phonetically with vowels rather than as a segmental phoneme.

Throughout this paper, we use both IPA (with Chao tone numerals) and notation conventions developed by language expert and co-author, Claudia Juárez Chávez, which build upon existing orthographic conventions developed for other Mixtec varieties, including those of the *Ve'e Tu'un Savi* (Mixtec Language Academy).[4]

3.2 Word, morpheme, and prosodic structure

SJPM syllables are canonically CV, with optional onsets and no codas, following the pattern attested across Mixtec (Longacre 1957; see also discussion in Macaulay 1996 and Uchihara & Mendoza Ruiz 2022). The minimal prosodic word in SJPM and in Mixtec more generally is bimoraic, otherwise known as a 'couplet' (following Pike 1944, 1948), which corresponds to the canonical morphological root. In both SJPM and other Mixtec varieties, the distribution of tonal, segmental, and supra-segmental contrasts in these language varieties is sensitive to the couplet, and the couplet may be the domain of tonal and segmental processes. We assume that this bimoraic phonological unit is equivalent to a metrical foot, aligned at the right edge of a prosodic word (e.g., Carroll 2015; Penner 2019; Uchihara & Mendoza Ruiz 2022). Specific phenomena that are relevant for this proposal in SJPM include the following (Duarte Borquez 2023):

(1) a. Glottalization is only attested within the couplet, associating with the first vowel of the couplet.

b. In disyllabic couplets, phonemically nasal vowels are only attested in the final syllable of the couplet.

c. Allophonic pre-aspiration of voiceless obstruents is only attested in the medial position of disyllabic couplets.

d. The couplet is the domain of realization of tonal melodies in lexical roots (§4) and a domain relevant for several tonal processes involving grammatical tone patterns (§5).

Thus, and as also argued for other Mixtec varieties (Penner 2019), we take the couplet to correspond to a metrical foot given the templatic function it plays in the synchronic phonology of SJPM (including its minimal prosodic word requirement), as well as the role it serves as the domain for phonotactic patterns and phonological processes. As we discuss in detail in §4 and §5 below, this prosodic domain is crucial for the tonal phonology of the language. In this regard, SJPM resembles other unrelated languages that possess tone patterns and processes sensitive to the metrical foot (e.g., Kera (Chadic; Pearce 2006)).

As shown in (2a), bimoraic roots may be realized as monosyllabic with a long vowel ((C)VV) or disyllabic with two short vowels ((C)VCV). As mentioned above, the bimoraic couplet is the domain of glottalization, which in SJPM is contrastive, as the examples in (2b) illustrate. We follow the analyses of other Mixtec varieties in Macaulay & Salmons 1995, Gerfen 1999, and Cortés et al. 2023 and of SJPM in Duarte Borquez 2023 in analyzing glottalization as a feature of root templates (couplets) associated in the surface form with the couplet-initial vowel. Our transcription of glottalization (ˀ), adopted from Cortés et al. 2023, indicates that glottalization is phonetically associated with the vowel.

[4] Representational conventions for SJPM segments that deviate from the IPA include the following: <r> = /ɾ/, <x> = /ʃ/, <ñ> = /ɲ/, <ku> = /kʷ/, <y> = /ʒ/, <ch> = /tʃ/, <ndy> = /ⁿdʒ/, <'> = /ˀ/. We address the phonological status of glottalization in §3.2. Regarding our representation of tone, we use the Chao system in our IPA transcriptions and diacritics in our orthographic forms. The latter involves the following conventions (discussed in more detail in Caballero, Juárez Chávez & Yuan (to appear)): H tones are encoded with an acute accent (H = á), M tones with a macron (M = ā), and L tones with a grave accent (L = à). When single TBUs host two tones in a contour, each tone is encoded on a single vowel. Because this orthographic convention collapses the distinction between short (monomoraic) and long (bimoraic) vowels (e.g., a sequence such as *àá* could be interpreted as /a^{15}/ or /a^{1}a^{5}/), we also include here the Chao tone numerals to disambiguate between such forms. Finally, pronominal enclitics that are CV in shape are represented separately from host words.

(2)	a.	(C)VV	*nìí*	ni^{1}i^{5}	'blood'	(C)VCV	*ùvì*	u^{1}vi^{1}	'two'
	b.	(C)VˀV	*nì'í*	ni^{1}ˀi^{5}	'hard'	(C)VˀCV	*ù'vì*	u^{1}ˀvi^{1}	'painful'

In addition to the patterns provided above, a significant number of roots in SJPM are trimoraic.[5] We take these forms to be monomorphemic synchronically, though most can be diachronically traced as being derived through the lexicalization of monomoraic prefixes and classifiers attaching to bimoraic stems (see below),[6] but loss of productivity has obscured their morphological structure (for further discussion, see Macaulay 1996, Palancar et al. 2016, and Penner 2019). These forms will be revisited in §4.2, as they are relevant to the distribution of lexical tonal contrasts realized within the metrical foot.

Although roots in SJPM may also be nominal or adjectival/adverbial, this paper primarily investigates the verbal system of the language, given that verbs show the highest amount of inflectional complexity, tonal or otherwise. First, we note that there are broadly two classes of verbs: a 'non-alternating' verb class with a single stem form across inflectional contexts, and an 'alternating' verb class that exhibits stem allomorphy (see Bickford & Marlett 1988, Macaulay 1996, Hollenbach 2015, and Palancar et al. 2016 for discussion of this system in other Mixtec varieties). As exemplified in Table 3 below, in the 'alternating' verb class, two distinct stems are used in the irrealis[7] and realis (indicative) mood contexts, respectively. We assume that these stem changes involve suppletive allomorphy, as they do not figure into our discussion of the lexical and grammatical tonal patterns.[8] Regardless of whether they alternate, irrealis stems are unmarked in inflection (both tonally and segmentally), while realis stems may be marked though tonal and segmental exponents of different inflectional categories (in patterns discussed in detail throughout §5). As shown in Table 3, uninflected irrealis stems are used to encode the future,[9] and, when used in isolation, are understood as citation forms of the verbs (accordingly, they may be translated by speakers into Spanish infinitival forms). In contrast, the realis stems, which bear additional inflection, are used to express the present and past tenses. See also Hollenbach 2015 for discussion.

Table 3: TAM contrasts in non-alternating vs. alternating verb stems

	Non-alternating *kētā* [ke^{3}ta^{3}] 'exit'		**Alternating** ***kā****tā* [**ka**^{3}ta^{3}] 'sing'		
Irrealis	M.M	*kētā* ke^{3}ta^{3}	M.M	***kā****tā* **ka**^{3}ta^{3}	*Future*
Realis	**H.**M	*kétā* ke^{5}ta^{3}	**H.**M	***xí****tā* **ʃi**^{5}ta^{3}	*Present*

[5] From a corpus of 488 monomorphemic roots, 102 (23%) are trimoraic.

[6] These include various derivational prefixes in the case of verbs and classifiers in the case of nouns.

[7] The morphological forms that we describe here as encoding irrealis mood are also described as 'potential' in Macaulay 1996 for Chalcatongo Mixtec and other descriptions of Mixtec language varieties.

[8] Alternating verbs may be classified into sub-classes in SJPM depending on the type of segmental change attested. Irrealis/realis stem contrasts may also involve tone alternations or tone plus segmental alternations in other Mixtec varieties, again typically analyzed as allomorphy (e.g., Chalcatongo Mixtec (Macaulay 1996)). That being said, a question that remains is whether a subset of these alternations is reducible to regular morpho-phonology plus prefixation.

[9] In addition, 2nd person imperatives are also encoded using the irrealis (uninflected) form of the verb. As a result, there are (to our knowledge) no morphological distinctions between the future, the infinitive, and the imperative usages of a given verb, which are all identical to the citation form. To limit our focus to the tense paradigms of SJPM, we set imperatives aside in this paper.

	LM.M *kèētā* ke^{13}ta^{3}	**LM**.M ***xììtā*** **ʃi**^{13}ta^{3}	*Past*

Given the minimal word requirement of the language, most monomoraic elements (setting aside a small number of sentential particles) are generally bound morphemes that concatenate on either side of the root: as with other Mixtec varieties, SJPM has a synthetic morphological system. A partial template of the verb complex is given in (3), along with some representative examples.[10] The preverbal morphemes (represented here as prefixes) may encode various derivational and inflectional distinctions; additionally, the forms in (3) demonstrate that event-modifying morphology (e.g., repetitive and causative morphology) surfaces closer to the root than inflectional morphology encoding tense and polarity. Notably, the irrealis stem of alternating verb stems (e.g., *kātā* ([ka^{3}ta^{3}]) 'sing') is the base for derivational affixation, regardless of the tense of the verbal complex.[11] Finally, these examples also show that negation and tense morphology may be expressed either segmentally or tonally.

(3) NEG-TNS-CAUS-REP-root[12]

a. *ndākātā* ^{n}da^{3}-ka^{3}ta^{3} 'sing again (IRR)'

b. *sándākātā* sa^{5}-^{n}da^{3}-ka^{3}ta^{3} 'make sing again (IRR)'

c. *sàándākātā* sa^{15}-^{n}da^{3}-ka^{3}ta^{3} 'made sing again (PST)'

d. *nìsándākātā* ni^{1}-sa^{5}-^{n}da^{3}-ka^{3}ta^{3} 'made sing again (PST)'

e. *sàándākātā* sa^{15}-^{n}da^{3}-ka^{3}ta^{3} 'not make sing again (NEG.IRR)'

f. *kònísándākātā* ko^{1}-ni^{5}-sa^{5}-^{n}da^{3}-ka^{3}ta^{3} 'did not make sing again (NEG.PST)'

In contrast, postverbal morphology consists of what we take to be clitics, as they attach to either the verb root itself or a following dependent of the verb, such as a postverbal adverb (e.g., Macaulay 1987, Yuan to appear). The enclitics may be pronominal[13] or else express a variety of illocutionary meanings. A non-exhaustive list of the pronominal enclitics of SJPM is given below in Table 4.[14] Although the majority of the enclitics in our inventory are consonantal (CV), there are five attested vocalic (V) enclitics in SJPM (two of which freely alternate with their CV counterparts). Note also that a subset of the enclitics, including

[10] We note that this template schematizes morphotactic generalizations, which are independent of prosodic or other domains. Possibly related to this point, a reviewer asks whether sequences of multiple preverbal morphemes may form their own prosodic domains (for instance, couplets or single phonological words). We are not aware of any evidence in favor of such internal prosodic structure.

[11] For this reason, it appears that the irrealis form is the 'default' realization of a bare verb root, while the realis form may only appear when the verb root is directly modified by tense. This may suggest that the irrealis vs. realis distinction arises as morphosyntactically-conditioned allomorphy (e.g., Bonet & Harbour 2012).

[12] In addition to abbreviations from the Leipzig Glossing Rules, the following are used in this article for language examples: 1 = first person, 2 = second person, 3 = third person, AN = animate noun class, CAUS = causative, CL = classifier, EXCL = exclusive, F = feminine, INCH = inchoative, INCL = inclusive, IRR = irrealis mood, LIQ = liquid noun class, M = masculine, N = neuter, NEG = negation, PL = plural, PST = past tense, PRS = present tense, REP = repetitive, RND = round noun class, SG = singular, TNS = tense.

[13] Pronouns in SJPM may also be expressed with bimoraic tonic pronouns, as well as monomoraic reduced pronouns. As discussed by Macaulay (1987, 1996), the monomoraic forms are distinguishable from the true clitics discussed here based on a number of morphophonological (including tonal) and morphosyntactic diagnostics (*pace* Pike 1944). In this paper, we focus solely on the clitics.

[14] As with other Mixtec varieties, nouns and pronouns in SJPM encode not only person and number, but also several noun class distinctions; however, we focus on human pronouns in this paper. See Macaulay 1987, De León Pasquel 1988, and Cisneros 2019 for discussion of Mixtec noun classes.

a vocalic one, is listed as tonally underspecified but bearing a floating L tone.[15] As we further detail throughout the paper (especially in §4.1 and §5), the enclitics are relevant to understanding the broader tonotactic patterns of the language.

Table 4: SJPM pronominal enclitics

1SG	2SG	3SG.M	3SG.F	3SG.N	1PL.EXCL	1PL.INCL	2PL	3PL.M	3PL.F	3PL.N
=è	=ón	=ra	=ñá / =án	=ña / =an	=ndi	=é	=ndó	=ndya	=ná	=na
=I^{1}	=õ^{5}	=ɾa$^{(+1)}$	=ɲa^{5} / =ã^{5}	=ɲa$^{(+1)}$ / =ã$^{(+1)}$	=^{n}di$^{(+1)}$	=I^{5}	=^{n}do^{5}	=^{n}d͡ʒa$^{(+1)}$	=na^{5}	=na$^{(+1)}$

The consonantal and vocalic enclitics differ in how they attach with their hosts. We assume that cliticization of the consonantal enclitics forms recursive prosodic words, as illustrated in (4) (see Uchihara & Mendoza (2022) for a similar proposal for Alcozauca Mixtec).[16] In these examples, we see that the clitics surface outside of the metrical foot, forming an extended prosodic word ('ω2'). The first cycle of the Prosodic Word, whose right edge coincides with the right edge of the foot, is represented as 'ω1'. In this paper, we represent the prosodic word and recursive prosodic word with square brackets and the metrical foot with parentheses.

(4) *kātā* [ka^{3}ta^{3}] 'sing'
  a. Verb=3SG.M: *kātā rā* [[(ka^{3}ta^{3})]$_{\omega 1}$=ra^{3}]$_{\omega 2}$ 'He will sing.'
  b. Verb=2PL: *kātā ndó* [[(ka^{3}ta^{3})]$_{\omega 1}$=^{n}do^{5}]$_{\omega 2}$ 'You (pl.) will sing.'

In contrast, as with other Mixtec varieties, the vocalic enclitics fuse segmentally and tonally with the final TBU of their hosts (e.g., Carroll 2015; Penner 2019; DiCanio et al. 2020; Uchihara & Mendoza Ruiz 2022), shown in (5) below. We moreover observe that the vocalic enclitics are *prosodically incorporated* into their hosts, meaning that a vocalic enclitic that attaches to a bimoraic root is incorporated into the metrical foot. Although the resulting segmental patterns will not be discussed further, the tonal patterns will be revisited in §4.1.[17]

(5) *kātā* [ka^{3}ta^{3}] 'sing'
  a. V=1SG: *kāt**ēè*** [(ka^{3}t=**e^{31}**)]$_{\omega}$ 'I will sing.'
  b. V=3SG.F: *kāt**á*** [(ka^{3}t=**a^{5}**)]$_{\omega}$ 'She will sing.'
  c. V=3SG.N: *kāt**ā*** [(ka^{3}t=**a^{3}**)]$_{\omega}$ 'It (e.g., the baby) will sing.'

[15] That the underspecified clitics also bear a floating L tone is based on the analysis of Duarte Borquez (2023) (which, in turn, draws on a similar proposal by Carroll (2015) for Ixpantepec Nieves Mixtec) and is included here for completeness. However, in this paper, we abstract away from this property, given that its presence is only observable in particular morphosyntactic contexts not discussed herein. In addition, the enclitic forms for 1SG and 1PL.IN are represented with the archiphoneme /I/, as their realizations are determined by the vowel quality of their hosts.

[16] The distinction between the prosodic word (the bimoraic foot plus prefixes) and the recursive prosodic word (the domain of consonant-initial enclitics) is justified in other varieties of Mixtec on the basis of phonological processes (e.g., accentuation or tonal processes) that apply in these different domains (Gerfen 1999, Uchihara & Mendoza Ruiz 2022). In the case of SJPM, evidence for positing the recursive prosodic word is found in tone sandhi processes that apply within the recursive prosodic word, which are different from other phonological processes that apply within the (first layer of the) prosodic word or across prosodic words. We leave this outside of the scope of this paper.

[17] For purposes of space, we do not discuss the range of segmental patterns that results from fusing the vocalic enclitics, though see Gittlen & Marlett 1985, DiCanio et al. 2020, and others for discussion of comparable data from other Mixtec varieties.

To summarize, this section has provided an overview of the segmental phonology of SJPM, as well as surveyed key prosodic and morphological properties of the language. With these baselines in place, the remainder of this paper focuses on the tone system of SJPM.

4 Lexical tone system

4.1 Lexical tone patterns and tonotactics

The lexical tone system of SJPM involves three level tones, namely H (V^{5}), M (V^{3}), and L (V^{1}), illustrated in the minimal set in (6) and their corresponding f0 pitch traces in Figure 2. We consider these three level tones to be tone feature primitives in SJPM.

(6) a. V=1SG: *kātēè* [(ka^{3}t=**e^{31}**)]$_{\omega}$ 'I will sing.'

b. V=3SG.F: *kātá* [(ka^{3}t=**a^{5}**)]$_{\omega}$ 'She will sing.'

c. V=3SG.N: *kātā* [(ka^{3}t=**a^{3}**)]$_{\omega}$ 'It (e.g., the baby) will sing.'

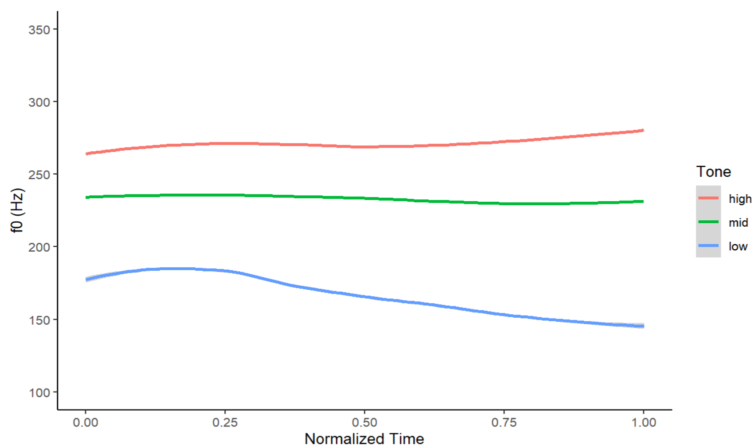


Figure 2: f0 track of SJPM H (*íín* [ĩ^{5}ĩ^{5}] 'hail'), M (*īīn* [ĩ^{3}ĩ^{3}] 'one') and L (*ììn* [ĩ^{1}ĩ^{1}] 'nine') lexical tones[18]

SJPM licenses association of sequences of level tones in a single mora, namely LM (V^{13}) (7a-b) and LH (V^{15}) (7c-e). MH rising contours are not attested underlyingly.[19] Falling contour tones are likewise unattested in lexical tone patterns; they surface only in grammatically derived tonal melodies or in highly restricted contexts involving a floating L tone, as discussed in more detail below.

(7) a **LM**.M *tììnā* ti^{13}na^{3} 'dog'

b. **LM**.H *chààká* t͡ʃa^{13}ka^{5} 'fish'

[18] Tone pitch traces in Figure 2 come from one of the authors of this article, Claudia Juárez Chávez, who was born in and grew up in San Juan Piñas, Oaxaca and is an L1 speaker of SJPM. She uses SJPM with family and community members and also as part of her activities developing language resources; she speaks Spanish (L2) for other daily activities. Pitch traces represent single tokens of H, M and L tones uttered in isolation.

[19] While the SJPM corpus contains a small number of MH contours in single morae in surface forms, these are restricted to only occur in patterns where we posit /HLH/ sequences, where the underlying L tone undergoes anticipatory raising resulting from f0 compression, a type of cross-linguistically common process of vertical assimilation (Hyman 2011). We assume this analysis in a number of monomorphemic forms, where MH contours are only attested in forms where /HLH/ tonal sequences can be posited, e.g., /L.H.LH/ [L.H.**MH**] [tu^{1n}di^{5}ʃi^{35}] 'blue' (alternatively, the second H tone in /HLH/ sequences may undergo downdrift (Duarte Borquez, Juárez Chávez & Caballero to appear)).

c.	**LH**.M	*chììkī*	t͡ʃi^{15}ki^{3}	‘prickly pear’
d.	H.**LH**	*yúchàá*	ʒu^{5}t͡ʃa^{15}	‘tender’
e.	M.**LH**	*yūvìí*	ʒu^{3}vi^{15}	‘petate’

Contours on single morae in SJPM are restricted to two-tone combinations; three-tone contours on single morae do not occur in this language variety.[20] As discussed in §6, this tonotactic constraint plays a role in the analysis of grammatical tone patterns in SJPM.

The underlying tonal melodies of bimoraic root templates include the presence of floating L tones (Caballero, Juárez Chávez & Yuan to appear, Duarte Borquez 2022). Bimoraic root templates exhibit a maximum of four tones in lexical tonal melodies (with up to three linked tones and a floating tone), but more complex tonal melodies are attested in derived environments: bimoraic stems sponsor up to four linked tones, whether monosyllabic (8a) or disyllabic (8b). Monomoraic functional morphemes, on the other hand, sponsor up to two linked tones, as shown in (8c) where the negative marker bears a LH rising tone (*ko^{15}*).

(8)	a.	LH.LH	*ìíìí*	**i^{15}i^{15}**	NEG.delicate	‘not delicate’
	b.	LH.ML	*kìíkuīì*	ki^{15}k^{w}=i^{31}	NEG.sew=1SG	‘I will not sew.’
	c.	LH-M.M	*kòóndīkō*	ko^{15}-^{n}di^{3}ko^{3}	NEG.PST-grind	‘did not grind’

These patterns are taken as evidence that, as proposed for other Mixtec varieties (e.g., Ixpantepec Nieves Mixtec (Carroll 2015), Ixtayutla Mixtec (Penner 2019), Alcozauca Mixtec (Uchihara & Mendoza Ruiz 2022)), the TBU in SJPM is the mora: lexical tone patterns in morphemes show distributions that are sensitive to moraic size, not syllabic structure.

The three tonal primitives of SJPM combine in tone patterns in bimoraic root templates, which are largely mirrored in monosyllabic and disyllabic stems. We assume the bimoraic root template (or ‘couplet’) is the basic unit of analysis for lexical tone distributions and tonal processes in SJPM, following other analyses of Mixtec tone systems (Pike 1947, 1948; Dürr 1982, 1987; Josserand 1983). A total of 18 lexical tone patterns are attested in root templates in the SJPM corpus.[21] We analyze some of these patterns as involving tonally unspecified TBUs (/M.Ø^{L}/, /H.Ø^{L}/, /LM.Ø^{L}/, /LH.Ø^{L}/), the evidence for which is both provided immediately below as well as in §5.1 and §5.3 (recall from §3.2 as well that certain pronominal enclitics are also posited to be tonally unspecified). Representative examples of lexical tonal melodies in bimoraic (glottalized and non-glottalized) monosyllabic and disyllabic root templates in native vocabulary are provided in Table 5, where floating tones are indicated with a superscript. Some lexical tone melodies surface with upstep of the second of a sequence of H stem tones, represented with a level [6] tone in the IPA transcription.[22]

[20] In contrast, some Mixtec varieties belonging to the neighboring Guerrero dialect group, such as Alcozauca Mixtec (Mendoza Ruiz 2016, Swanton & Mendoza Ruiz 2021, Uchihara & Mendoza Ruiz 2022) and Yoloxóchitl Mixtec (DiCanio et al. 2014, Palancar et al. 2016), allow three-tone sequences in single morae.

[21] Additional tone patterns are found only in loanwords from Spanish (for instance, /H.L/ for the (non-glottalized) (C)VV sequence *páàn* (pã^{5}ã^{1}) ‘bread’). These sequences are not included in Table 5 as they are otherwise not present among native words in the language.

[22] SJPM exhibits complex tonal register phenomena, including a phonetic process of upstep, where an extra-high tone surfaces after another H-toned TBU and before a floating L tone (Duarte Borquez, Juárez Chávez & Caballero to appear). Although we include upstepped H tones in our IPA transcriptions, further discussion of upstep is outside of the scope of this paper.

Table 5: SJPM lexical tone patterns in native bimoraic root templates (couplets)

	Monosyllabic				**Disyllabic**			
Tone pattern	**(C)VV**		**(C)V^{ʔ}V**		**(C)VCV**		**(C)VCV**	
/H.H/	*íín* ‘hail’	ĩ^{5}ĩ^{5}	*ché’é* ‘hard’	t͡ʃe^{5ʔ}e^{5}	*yánká* ‘difficult’	ʒa^{5n}ka^{5}	*kuí’ná* ‘thief’	k^{w}i^{5ʔ}na^{5}
/M.M/	*īīn* ‘one’	ĩ^{3}ĩ^{3}	*lō’ō* ‘small’	lo^{3ʔ}o^{3}	*lēsō* ‘rabbit’	le^{3}so^{3}	*yā’vī* ‘expensive’	ʒa^{3ʔ}vi^{3}
/M.Ø^{L}/	*sāā* ‘bird’	sa^{3}a^{3+1}	*ñō’ōn* ‘fire’	ɲõ^{3ʔ}õ$^{3+1}$	*ītā* ‘flower’	i^{3}ta^{3+1}	*ñā’mī* ‘sweet potato’	ɲa^{3ʔ}mi^{3+1}
/L.L/	*ììn* ‘nine’	ĩ^{1}ĩ^{1}	*òn’òn* ‘five’	õ^{1ʔ}õ^{1}	*sùtù* ‘priest’	su^{1}tu^{1}	*ù’vì* ‘painful’	u^{1ʔ}vi^{1}
/H.M/	*káā* ‘there’		*ndí’ī* ‘purple’	^{n}di^{5ʔ}i^{3}	*xíkā* ‘far’	ʃi^{5}ka^{3}	*ká’nō* ‘big’	ka^{5ʔ}no^{3}
/H.L/[23]	—		*chá’à* ‘weak’	t͡ʃa^{5ʔ}a^{1}	*ndíkà* ‘wide’	^{n}di^{5}ka^{1}	—	
/M.L/	*chāàn* ‘tomorrow’	t͡ʃã^{3}ã^{1}	*xā’à* ‘smell’	ʃa^{3ʔ}a^{1}	*yāvì* ‘hole’	ʒa^{3}vi^{1}	*sā’và* ‘frog’	sa^{3ʔ}va^{1}
/M.H/	*ndyēé* ‘thick’	^{n}d͡ʒe^{3}e^{5}	*ñō’ón* ‘earth, dirt’	ɲõ^{3ʔ}õ^{5}	*vīkó* ‘celebration’	vi^{3}ko^{5}	*sā’má* ‘cloth’	sa^{3ʔ}ma^{5}
/H.Ø^{L}/	*kuéé* ‘slowly’	k^{w}e^{5}e^{6+1}	*yá’á* ‘golden’	ʒa^{5ʔ}a^{6+1}	*lúví* ‘turtle’	lu^{5}vi^{6+1}	*ndá’ví* ‘humble’	^{n}da^{5ʔ}vi^{6+1}
/L.H/	*kòó* ‘snake’	ko^{1}o^{5}	*xà’á* ‘lower leg/foot’	ʃa^{1ʔ}a^{5}	*xìní* ‘head’	ʃi^{1}ni^{5}	*vì’ndyá* ‘cactus’	vi^{1ʔn}d͡ʒa^{5}
/L.M/	*chàā* ‘man’	t͡ʃa^{1}a^{3}	*và’ā* ‘good’	va^{1ʔ}a^{3}	*tìxīn* ‘belly’	ti^{1}ʃĩ^{3}	*kù’vā* ‘sibling (op. sex)’	ku^{1ʔ}va^{3}
/M.LH/	*īìín* ‘skin’	ĩ^{3}ĩ^{15}	—		*nāñàá* ‘chayote’	na^{3}ɲa^{15}	—	
/LM.H/	—		—		*chàāká* ‘fish’	t͡ʃa^{13}ka^{5}	—	
/LH.M/	*tìíīn* ‘mouse’	tĩ^{15}ĩ^{3}	*chòó’ō* ‘flea’[24]	t͡ʃo^{15ʔ}o^{3}	*yòósō* ‘metate’	ʒo^{15}so^{3}	*nàá’nā* ‘picture’	na^{15ʔ}na^{3}
/H.LH/	*kuíìí* ‘green’	k^{w}i^{5}i^{15}	*tí’ìín* ‘type of quelite	tĩ^{5ʔ}ĩ^{15}	*yúchàá* ‘tender’	ʒu^{5}t͡ʃa^{15}	*xí’ndyàá* ‘stingy’	ʃi^{5ʔn}d͡ʒa^{15}
/LM.Ø^{L}/	*xàāā* ‘chin’	ʃa^{13}a^{3+1}	*chòō’ō* ‘type of pomegranate’	t͡ʃo^{13ʔ}o^{3+1}	*tìīnā* ‘dog’	ti^{13}na^{3+}	—	
/LH.H/	—		*chòó’ó* ‘flea’	t͡ʃo^{15ʔ}o^{5}	*chòókó* ‘ant’	t͡ʃo^{15}ko^{5}	—	
/LH.Ø^{L}/	—		*ndìí’í* ‘fat’	^{n}di^{15ʔ}i^{6+1}	*xàání* ‘dream’	ʃa^{15}ni^{6+1}	—	

As mentioned above, SJPM exhibits a high degree of *tonal density*, defined as the proportion of underlyingly specified (H, L or M, in the case of SJPM) TBUs compared to unspecified (/Ø/) TBUs in the language (Gussenhoven 2004, Hyman 2009). Given this high tonal density, the restrictions on prosodic size

[23] Native /H.L/ lexical forms are exclusively attested with adjectives. For other Mixtec varieties (e.g., Chalcatongo), a so-called ‘adjectival H’ tone has been posited to derive adjectives from nominal stems (e.g., Hinton et al. 1991). It is not clear at this time if a comparable analysis could be posited for SJPM.

[24] Two forms for ‘flea’, used in free variation, are represented in Table 5.

and shape of stems, and the high number of tone patterns available for bimoraic stems, we conclude that there is a high functional load of tone for lexical contrasts, as also proposed for other varieties of Mixtec (e.g. Palancar et al. 2016).

The contrasts between /M.L/ and /M.Ø^{L}/ and between /H.L/ and /H.Ø^{L}/ (exemplified below in (9a-d) and (9e-h), respectively) suggest that these tonal patterns are not automatically associated to TBUs in this language, i.e., the alignment of these tones to TBUs is not phonologically predictable in lexical tone patterns. These contrasts provide additional evidence against the automatic association of tones, as formulated in the Universal Association Conventions (Goldsmith 1976; see also Hyman & Ngunga 1994, Archangeli & Pullyblank 1994, Tranel 1995, and McPherson & Dryer 2021).

(9) **/M.L/ vs. /M.Ø^{L}/ stems**

a.	/M.L/	ʒa^{3}vi^{1}	'hole'	b.	/M.L/	so^{3}ko^{1}	'shoulder'
c.	/M.Ø^{L}/	i^{3}ta^{3+1}	'flower'	d.	/M.Ø^{L}/	i^{3}su^{3+1}	'deer'

/H.L/ vs. /H.Ø^{L}/ stems

e.	/H.L/	ⁿdi^{5}ka^{1}	'wide'	f.	/H.L/	t͡ʃa^{5ʔ}a^{1}	'weak'
g.	/H.Ø^{L}/	kʷi^{5}ɲo^{6+1}	'jealous'	h.	/H.Ø^{L}/	si^{5ʔ}i^{6+1}	'woman'

The tonal configurations of stems with underspecified TBUs in their underlying forms may have arisen through a diachronic process of rightward tonal shift that affected some tone classes in SJPM, a process claimed to have shaped the tone systems of other Mixtec varieties (Hollenbach 2003, McKendry 2013; see also Penner 2019). According to this proposal, informed by reconstruction of the tone system of Proto-Mixtec (Longacre 1957, Dürr 1987; see also McKendry 2013), the tone patterns of couplets involving two tones, T1 and T2, originally associated with the first and second mora, respectively, have become displaced, with T1 reassociating with the second mora and T2 becoming a floating tone. The comparative evidence supports a diachronic process of rightward tone shift in SJPM: SJPM M-toned stems bearing floating L tones correspond to couplets with *ML tone patterns in Proto-Mixtec. While in other Mixtec varieties this diachronic process is analyzed as resulting in lexical tone melodies aligned with the right edge of the couplet and with initial toneless TBUs (e.g., Southeastern Nochixtlán Mixtec; McKendry 2013), we propose that diachronic rightward shift in SJPM involved displacement of the L tone (T2 in Proto-Mixtec *M.L couplets),[25] leaving the second TBU of these stems toneless (*/M.L/ → /M.Ø^{L}/). In addition, we propose that this diachronic rightward shift may be straightforwardly extended to account for the /H.Ø^{L}/, /LM.Ø^{L}/, and /LH.Ø^{L}/ patterns, in that they likewise may have arisen from rightward shift of the L tone in *H.L, *M.L, and *LH.L forms.[26] We also present further synchronic evidence for underspecification in SJPM from grammatical tone patterns in §5 below.

As seen in Table 5 above, all nine logical possible combinations of level tones are attested in couplets. On the other hand, some three-toned lexical tone patterns are only marginally attested in the SJPM corpus, while others are unattested. Importantly, we observe a systematic restriction with respect to falling contours

[25] We follow Longacre (1957) in labeling the two tonal levels of Proto-Mixtec as *L and *M; these labels correspond *L and *H, respectively, in Dürr's (1987) tonal reconstruction.

[26] Additional evidence for this diachronic development can be found in the tone patterns of Spanish loanwords in SJPM. While in other Mixtec varieties Spanish loanwords were incorporated with a HL tone pattern, with a H tone aligned with the stressed syllable in the source words (Hollenbach 2003), Spanish words with penultimate stress are incorporated into SJPM with a [H.$^{\uparrow}$H] tone pattern and a L floating tone, e.g., Spanish *mesa* 'table' and *paño* 'shawl', with penultimate stress, correspond to SJPM *mésá* [me^{5}sa^{6+1}] and *páñó* [pa^{5}ɲo^{6+1}], respectively (Spanish loanwords with other stress patterns have a HL tone pattern aligned with the stressed syllable of the source word, e.g., *páàn* [pã^{5}ã^{1}] 'bread', from Sp. *pan*). We suggest that Spanish source words with penultimate stress were also incorporated into SJPM with a HL pattern, but the proposed diachronic rightward shift in SJPM moved the L tone of the HL tone pattern one mora rightward in these forms, resulting in the L tone becoming floating. We thus analyze these loanword patterns as involving underspecification in their synchronic underlying representations (*/H.L/ → /H.Ø^{L}/).

on single morae, as these are completely unattested in lexical tone patterns. Though they are present in morphologically complex forms, they are still distributionally highly restricted in such contexts as well, appearing only in couplet-final position (aligned at the right edge of the first cycle of the Prosodic Word). An example of this is observed in the context of 1SG marking, which involves the attachment of a L-toned vocalic enclitic that coalesces with the final vowel of the stem (as introduced in §3.2 above). As shown in (8b) above and in the following examples in (10), the L tone of this vocalic pronominal form concatenates to the lexical tone of the base, yielding a falling contour when cliticizing to M-final (10a) and H-final (10b) stems.

(10)		**Stem**		**Stem=1SG**
a.	[L.M]	*kàkū* [(ka^{1}ku^{3})]$_{\omega}$ 'survive'	[L.**ML**]	*kàkuīì* [(ka^{1}k^{w}=**i**$^{\mathbf{31}}$)]$_{\omega}$ survive=1SG 'I will survive.'
b.	[M.H]	*tūtú* [(tu^{3}tu^{5})]$_{\omega}$ 'whistle'	[M.**HL**]	*tūtíì* [(tu^{3}t=**i**$^{\mathbf{51}}$)]$_{\omega}$ whistle=1SG 'I will whistle.'

Other instances of falling contours are found in derived forms containing a stem bearing a floating L tone. As shown below in (11), when a /M.Ø^{L}/ stem attaches a H-toned enclitic, e.g., the 3SG.F enclitic *=ɲa^{5}*, the floating L tone associates leftward with its host, again creating a falling contour in couplet-final position. In isolation, however, the floating L tone of the stem goes unrealized, leading to a level M-tone pronunciation of the stem, e.g., [M.M] *ɲa^{3ʔ}mi^{3}* 'sweet potato.' We assume that /M.Ø^{L}/ stems otherwise surface as [M.M] in final position due to a process of rightward M tone spreading.

(11)	**/M.Ø^{L}/ stem**		**/M.Ø^{L}/ stem=3SG.F**	
a.	[M.M]	*ñā'mī* [(ɲa^{3ʔ}mi^{3+1})]$_{\omega}$ 'sweet potato'	[M.**ML**=H]	*ñā'mīì ñá* [[(ɲa^{3ʔ}m**i**$^{\mathbf{31}}$)]$_{\omega1}$=ɲa^{5}]$_{\omega2}$ sweet.potato=3SG.F 'her sweet potato'
b.	[M.M]	*ītā* [(i^{3}ta^{3+1})]$_{\omega}$ 'flower'	[M.**ML**=H]	*ītāà ñá* [[(i^{3}t**a**$^{\mathbf{31}}$)]$_{\omega1}$=ɲa^{5}]$_{\omega2}$ flower=3SG.F 'her flower'

Otherwise, derived falling contours are completely unattested in couplet-initial position and in pre-couplet moras in SJPM. We attribute this gap to two interacting tonotactic restrictions operating in SJPM: falling contours are not permitted lexically, and, in derived contexts, are only permitted in the final position of the Prosodic Word. As discussed in §6, this proposal has consequences for the analysis of grammatical tone patterns in SJPM.

Relatedly, while falling melodies in bimoraic couplets (/H.L/, /H.M/ and /M.L/) are attested in SJPM, they are relatively marginal in lexical patterns of bimoraic stems in native vocabulary: from a total of 386 bimoraic stems, /H.L/ melodies account for just 1.81% of the sample,[27] /H.M/ melodies constitute only 4.40% of the sample, and /M.L/ patterns make up 4.40% of the sample. In contrast, the most frequently attested tone patterns include /M.M/ (16.58%), /L.L/ (11%), /M.H/ (15.28%) and other rising melodies

[27] In native SJPM vocabulary, attested /H.L/ couplets are all adjectives, a likely result of a formerly productive process of tonal derivation where adjectives are derived from nouns through attachment of a H tone, a process documented in other varieties of Mixtec, e.g., Chalcatongo (Hinton et al. 1991). We take these tone patterns to be monomorphemic synchronically, having become lexicalized in the language.

(/L.H/ (10.36%) and /L.M/ (7.25%)). There is thus a clear trend in the language favoring rising tonal melodies. This finding is consistent with the generalization that there are no falling contours on single morae in the underlying lexical patterns of SJPM, against a noted cross-linguistic trend where falling melodies are proposed to be more common cross-linguistically (Zhang 2000, 2001; Yip 2002).

Finally, some additional gaps in the inventory of lexical tone patterns include the following: (i) toneless (Ø) TBUs never appear in couplet-initial position, a pattern that may be attributed to the aforementioned diachronic rightward shift that affected only the second TBU in some tone patterns of couplets; (ii) M tone never combines with LM contours in any order; and (iii) L does not occur with toneless (Ø) TBUs or LH contours.

Table 6 summarizes the frequencies of each tone pattern attested in the SJPM corpus.

Table 6. Distribution of lexical tone patterns in native lexical stems

	Monosyllabic		**Disyllabic**	
Tone pattern	**(C)VV**	**(C)VˀV**	**(C)VCV**	**(C)VˀCV**
/H.H/	(5) 6.75%	(1) 1.69%	(5) 2.35%	(1) 2.43%
/M.M/	(14) 18.91%	(4) 6.77%	(40)18.86%	(6) 14.63%
/M.Ø^{L}/	(7) 9.45%	(5) 8.47%	(16) 7.54%	(2) 4.88%
/L.L/	(2) 2.70%	(7) 11.86%	(26) 12.26%	(9) 21.95%
/H.M/	(1) 1.35%	(7) 11.86%	(7) 3.30%	(2) 4.88%
/H.L/	—	(4) 6.77%	(2) 0.94%	(1) 2.43%
/M.L/	(2) 2.70%	(1) 1.69%	(9) 4.24%	(5) 12.20%
/M.H/	(6) 8.10%	(8) 13.56%	(39)18.86%	(6) 14.63%
/H.Ø^{L}/	(2) 2.70%	(3) 5.1%	(10) 4.71%	(1) 2.43%
/L.H/	(13) 17.57%	(6) 10.17%	(19) 8.49%	(2) 4.88%
/L.M/	(9) 12.16%	(9) 15.25%	(9) 4.24%	(3) 7.32%
/M.LH/	(7) 9.46%	—	(3) 1.41%	—
/LH.M/	(1) 1.35%	(2) 3.39%	(7) 3.30%	(1) 2.43%
/LM.H/	(5) 6.75%	—	(4) 1.88%	—
/H.LH/	—	—	(3) 1.41%	(2) 4.88%
/LM.Ø^{L}/	—	—	(8) 3.77%	—
/LH.H/	—	—	(3) 1.41%	—
/LH.Ø^{L}/	—	(2) 3.39%	(2) 0.94%	—
Total	**(74) 19.18%**	**(59) 15.28%**	**(212) 54.92%**	**(41) 10.62%**

4.2 Lexical tone patterns in verb stems

In addition to the phonotactic constraints on tonal distributions noted, the tone patterns of different lexical classes in SJPM (e.g., nouns, adjectives/adverbs, and verbs) exhibit asymmetries.[28] Of particular interest in this paper are the tone patterns of verbs, which, as mentioned above, exhibit the highest degree of complexity in terms of grammatical tone. Notably, verbs have the smallest number of lexical tone patterns of all the major word classes in SJPM: from the 18 possible lexical tonal patterns attested in the language, only 11 are attested in bimoraic verb roots surfacing in the irrealis mood (i.e., uninflected form). One systematic restriction in lexical tone patterns of verbs concerns the fact that H-initial melodies are generally unattested in bimoraic verb stems. The sole exception is the /H.M/ melody, which arises in a single verb, *sa*^{5ʔ}*a*^{3} 'do, make'. This restriction in the lexicon against couplet-initial H tone is also documented in other Mixtec lexical tone systems, and may be attributed to the role that H tone plays in encoding verbal inflection (Carroll 2015, Palancar et al. 2016). The attested lexical tone patterns of bimoraic verb stems are given in Table 7.[29]

Table 7: Lexical tone patterns in bimoraic verb stems

Tone melody	Monosyllabic stem		Gloss	Disyllabic stem		Gloss
/L.L/	*kò'òn*	kõ^{1ʔ}õ^{1}	'go'	*kùsù*	ku^{1}su^{1}	'sleep'
/M.M/	*kōōn*	kõ^{3}õ^{3}	'fall (rain, hail)'	*ndīkō*	^{n}di^{3}ko^{3}	'grind'
/M.Ø^{L}/	*xīī*	ʃi^{3}i^{3+1}	'wilt'	*ndāsī*	^{n}da^{3}si^{3+1}	'close (intr.)'[30]
/L.M/	*yà'ā*	ʒa^{1ʔ}a^{3}	'pass'	*kìkū*	ki^{1}ku^{3}	'weave'
/L.H/	—			*ndàtá*	^{n}da^{1}ta^{5}	'tear'
/M.L/	*xā'àn*	ʃã^{3ʔ}ã^{1}	'smell'	*kōnì*	ko^{3}ni^{1}	'love'
/M.H/	*īín*	ĩ^{3}ĩ^{5}	'peel w/hand'	*kīní*	ki^{3}ni^{5}	'shoot'
/H.M/	*sá'ā*	sa^{5ʔ}a^{3}	'do, make'	—		
/M.LH/	*kōòón*	kõ^{3}õ^{15}	'spread'	—		
/LH.M/	—			*xìítā*	ʃi^{15}ta^{3}	'pull'
/LH.Ø^{L}/	—			*xààní*	ʃa^{15}ni^{6+1}	'dream'

Tonal lexical contrasts in verb stems are realized in both moras of bimoraic stems, whether monosyllabic or disyllabic. In the trimoraic verb stems documented so far, the first mora surfaces with a M tone; the prosodic foot (couplet), aligned at the right edge of the (minimal) prosodic word, bears the lexical tone

[28] Nouns are the most heterogeneous word class in SJPM in terms of lexical tone patterns. This is consistent with the observation that cross-linguistically nouns show phonologically privileged behavior compared to other word classes, especially verbs (Smith 2001).

[29] From the melodies provided in Table 7, three-tone and four-tone melodies are rarely attested in the SJPM corpus: from a sample of 107 bimoraic verb stems, /M.LH/ accounts for 2% of tone patterns, /LH.M/ for 1% and /LH.Ø^{L}/ for 2%.

[30] Certain verb forms in SJPM display partially suppletive causative-inchoative alternations; here, the choice of verb indicates the inchoative (intransitive) usage.

contrasts in trimoraic verb roots.[31] The attested lexical tone melodies of these root types are provided in Table 8. The couplet is indicated in parentheses and mora breaks indicate the alignment of tones to specific moras.

Table 8: Lexical tone patterns in trimoraic verb stems

Tone melody	Trimoraic stem		Gloss
/M.M.M/	*kūndātū*	ku^{3}(^{n}da^{3}tu^{3})	'wait'
/M.M.Ø^{L}/	*ndātōōn*	^{n}da^{3}(tõ^{3}õ$^{3+1}$)	'light'
/M.L.L/	*kīxàà*	ki^{3}(ʃa^{1}a^{1})	'arrive'
/M.H.H/	*kūsíkí*	ku^{3}(si^{5}ki^{5})	'play'
/M.M.H/	*kūndāsí*	ku^{3}(^{n}da^{3}si^{5})	'get scared'
/M.L.M/	*ndākòō*	^{n}da^{3}(ko^{1}o^{3})	'wake up'
/M.LM.Ø^{L}/	*nōtàā'ā*	no^{3}(ta^{13ʔ}a^{3})	'need'
/M.LH.M/	*kūndyèé'ē*	ku^{3}(^{n}d͡ʒe^{15ʔ}e^{3})	'watch'
/M.H.Ø^{L}/	*kūtóó*	ku^{3}(to^{5}o^{6+1})	'like'
/M.H.M/	*kūxínī*	ku^{3}(ʃi^{5}ni^{3})	'eat dinner'
/M.L.H/	*tāndà'á*	ta^{3}(^{n}da^{1ʔ}a^{5})	'get married'
/M.L.H^{L}/	*kūyù'ú*	ku^{3}(ʒu^{1ʔ}u^{5+1})	'fear'

These verb stems are the base for tonal and segmental exponence of both derivational and inflectional morphology, which, as alluded to in §3.2 above, exhibit 'layering' of exponents, as one moves outwards from the root. We turn to grammatical tone patterns next.

5 Grammatical tone patterns

In this paper, we adopt Rolle's (2018) definition of *grammatical tone* as a tonal pattern or process that is not general across the phonological grammar of a language but is instead associated with a specific morpheme or construction or a natural class of morphemes or constructions. While previous research on Mixtec tone systems has focused primarily on lexical tone patterns, the role that tone plays in the encoding

[31] A similar pattern is attested in other Mixtec varieties, e.g., Yoloxóchitl Mixtec (Palancar et al. 2016), where tone melodies of trimoraic stems generally have a M tone in their first TBU. Palancar et al. propose that this pattern can be attributed to the diachronic development of monomorphemic trimoraic verbs stems from lexicalization of bimoraic roots and derivational prefixes (2016:307). A question that arises for SJPM is why all attested couplet-external TBUs in monomorphemic verb forms are M. As proposed for Yoloxóchitl Mixtec, these TBUs may be the product of lexicalization of various monomoraic derivational M-toned prefixes. Since we posit unspecified TBUs elsewhere in the language, an alternative would be to posit that these initial TBUs in trimoraic verb stems are underlyingly unspecified and receive M tone by default, a pattern that would be exclusively attested in verbs but not nouns (in trimoraic monomorphemic nominal stems, the first couplet-external mora may surface with a M or a L tone, e.g., *chīnáká* /t͡ʃi^{3}na^{5}ka^{5}/ 'bat', *tìkuìītí* /ti^{1}k^{w}i^{13}ti^{5}/ 'potato', possibly due to the lexicalization of prenominal classifiers (De León 1988, Macaulay 1996)). However, there is no evidence from tonal processes that the couplet-external TBUs in these verbs are toneless. Furthermore, patterns of L tone dissimilation in trimoraic stems (described in §5.3 below) may provide further evidence that couplet-external TBUs in verbs are underlyingly specified as M.

of morphological contrasts in these language varieties and how these interact with the lexical tone system has remained largely under investigated (cf. Palancar et al. 2016, Swanton & Mendoza Ruiz 2021). As in other Mixtec varieties documented to date, SJPM exhibits a robust grammatical tone system, with tonal exponence playing a crucial role in the encoding of TAM and negative polarity.

We provide the first detailed description of the grammatical tone system of SJPM, focusing on tonal patterns of verbs inflected for tense and negative polarity. There are four dimensions of exponence of these inflectional categories in SJPM: (i) tone may be the sole exponent of inflection (referred to as 'tonal morpheme' in Welmers 1959 and Hyman 2016); (ii) tonal exponents may either replace, accommodate or avoid lexical tones; (iii) there is a choice between segmental and tonal allomorphs for some inflectional categories; and (iv) there is a trend of contrast preservation of paradigmatic contrasts in tonal/segmental exponence. In this section, we organize the description and analysis of grammatical tone in terms of its function within the inflection system.

5.1 Tonal exponents

We analyze grammatical tone patterns in SJPM as resulting from the association of grammatical *floating tones*. Specifically, SJPM possesses three morphosyntactically-triggered tonal melodies that are consistently associated with specific tense and polarity values across paradigms: a H tone encoding present tense, a L tone encoding past tense, and a LH melody encoding negative polarity. The grammatical tone patterns in SJPM can thus be straightforwardly analyzed as involving an additive underlying phonological item (for discussion about criteria for assessing item-based vs. process-based analyses of grammatical tone, see Sande 2022).

A partial template of the SJPM verb complex is shown again in (12) (partially repeated from §3.2 above). The docking patterns of tonal exponents are exemplified in (13) in various derived forms of the verb *kātā* [ka^{3}ta^{3}] inflected for the past tense, which is encoded through a L tone. As these examples demonstrate, tonal exponents dock to the leftmost TBU of the morpheme within their (syntactic) scope. Therefore, the resulting tone patterns may surface directly on the verbal root, as in (13b), or on the leftmost derivational (or inflectional) morpheme of the verb stem prior to the concatenation of the tonal morpheme (13c-d).

(12) TNS-CAUS-REP-root

(13)	a.	M.M	*kātā*	ka^{3}ta^{3}	'sing'	root
	b.	**L**M.M	*xìītā*	ʃi^{13}ta^{3}	'sang'	PST.root
	c.	**L**-M.M	*ndàkātā*	^{n}da^{1}-ka^{3}ta^{3}	'sang again'	PST.REP-root
	d.	**L**H-M-M.M	*sàándākātā*	sa^{15}-^{n}da^{3}-ka^{3}ta^{3}	'made sing again'	PST.CAUS-REP-root

Notably, the addition of the derivational (repetitive (13c) and causative (13d)) prefixes does not trigger tonal alternations in the stem; the only tonal change in these cases is the one brought about by the exponence of past tense. This is consistent with the broader notion of 'layering' of tonal (and segmental) patterns moving outwards from the root.

Below, we illustrate the grammatical tone patterns associated with present tense (H), past tense (L), and negation (LH) in turn.

5.2 Present tense inflected verbs

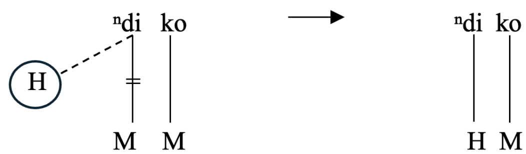
Present tense is encoded via a H tone, which replaces lexical M (14a-e) and L (14f-g) tones of the TBU to which it docks. In (14h), in which H tone applies vacuously to a H-initial verb stem, *sa*^{5ʔ}*a*^{3} 'do, make', the contrast between irrealis and present tense is neutralized. No other contrasts are neutralized with present tense, since bimoraic verb stems are otherwise L- or M-initial.

(14)

	Stem			**Present H tone**			
a.	M.M	*chāā*	t͡ʃa^{3}a^3	**H**.M	*cháā*	t͡ʃa^{5}a^3	'write'
b.	M.M	*ndīkō*	^{n}di^{3}ko^3	**H**.M	*ndíkō*	^{n}di^{5}ko^3	'grind'
c.	M.M	*kātā*	ka^{3}ta^3	**H**.M	*xítā*	ʃi^{5}ta^3	'sing'
d.	M.H	*tī'ví*	ti^{3ʔ}vi^5	**H**.H	*tí'ví*	ti^{5ʔ}vi^5	'sweep'
e.	M.L	*kōnì*	ko^{3}ni^1	**H**.L	*kónì*	ko^{5}ni^1	'love'
f.	L.M	*kìkū*	ki^{1}ku^3	**H**.M	*kíkū*	ki^{5}ku^3	'sew'
g.	L.L	*kàkù*	ka^{1}ku^1	**H**.L	*xákù*	ʃa^{5}ku^1	'laugh'
h.	H.M	*sá'ā*	sa^{5ʔ}a^3	**H**.M	*sá'ā*	sa^{5ʔ}a^3	'do/make'

This process is schematized in (15) using an autosegmental representation (Goldsmith 1990).

(15) Association of H present tense grammatical tone



The grammatical H tone may also dock on stems with derivational prefixes, e.g., repetitive *^{n}da^3-* (16a) or causative *sa^5-* (16b). As with (14h) above, the present H tone vacuously applies in H-initial causative verb stems.

(16)

	Stem		**Present tense (/H/)**	
a.	M-M.M	*ndākātā* ^{n}da^3-ka^{3}ta^3 REP-sing 'sing again'	H-M.M	*ndákātā* ^{n}da^5-ka^{3}ta^3 PRS.REP-sing 'singing again'
b.	H-M.M	*sákātā* sa^5-ka^{3}ta^3 CAUS-sing 'make sing'	H-M.M	*sákātā* sa^5-ka^{3}ta^3 PRS.CAUS-sing 'making sing'

While the present-tense grammatical tone pattern involves replacement of lexical tone in most verb forms, verbs that have a lexical contour tone in the first TBU undergo a different process: in these cases, the grammatical tone melody does not replace the lexical contour, but instead docks to an epenthetic high, front vowel, crucially not found in the citation form of the verb. As a result, the underlying lexical tones are preserved. This is shown in (17).

(17)

a.	LH.Ø^L	*xàání*	ʃa^{15}ni^{6+1}	'dream'
b.	**H**.LH.Ø^L	*íxàání*	i^{5}ʃa^{15}ni^{6+1}	'dreaming'
c.	LH.Ø^L	*yàákón*	ʒa^{15}kõ$^{6+1}$	'peel w/knife'
d.	**H**.LH.Ø^L	*íyàákón*	i^{5}ʒa^{15}kõ$^{6+1}$	'peeling with knife'
e.	LH.M	*xììtā*	ʃi^{15}ta^3	'pull'
f.	**H**.LH.M	*íxììtā*	i^{5}ʃi^{15}ta^3	'is pulling'

Based on our corpus, no verbs with a stem-initial level tone bear this prefix, while all verbs with a stem-initial contour tone do bear it. Therefore we claim this is an instance of tonally-driven epenthesis: the phonological grammar supplies an epenthetic front, high vowel in order to realize a grammatical tone when the grammatical tone would otherwise replace a contour lexical tone, a restriction also documented for Alcozauca Mixtec (Uchihara & Mendoza Ruiz 2022).[32] This phenomenon can be characterized as an instance of *tonological indomitability*, where the set of exceptional targets that resist association of grammatical tone share tonal characteristics (Rolle 2018). Thus, in the case of SJPM, tonal contours on single morae are impervious to replacement by grammatical tones. As described in §5.2.3 below, further support for this characterization comes from the fact that tonal-driven epenthesis is also attested in negative polarity marking with the same verb stems, where the lexical tone contours of the stem's first TBU are again preserved. Tonal-driven epenthesis may therefore be viewed as a general response to the inability to replace lexical tonal contours with grammatical tones. It should also be noted that failure of contour tones to undergo replacement by grammatical tones cannot be explained as avoidance of a marked structure, as we have already seen that sequences of H tones (which would result if tonal replacement would take place) are attested in both lexical tone patterns and in grammatically derived melodies.

Finally, a class of stems, which we posit to be underlyingly /M.Ø^L/, exhibits yet another tone pattern when inflected for present tense. Recall from §4.1 that these stems may be analyzed as diachronically arising from a */M.L/ form in Proto-Mixtec, whereby the L tone has displaced rightward in SJPM, leaving the couplet-final TBU unspecified. Further evidence for this idea comes from the resulting pattern of grammatical tone association in such contexts. Notably, as shown in (18a-d), these stems surface with a [H.$^{\uparrow}$H] tone pattern (with phonetic upstep of the second H-toned TBU of the stem, represented with a level [6] tone in the IPA transcription) in the present tense, rather than the expected [H.M] pattern. Indeed, these stems contrast with /M.M/ stems that do surface as [H.M] in the present tense, as shown in (18e-h).

(18)

	Stem			**Present tense (/H/)**			
a.	[M.M]	*kāvā*	ka^{3}va^3	[H.$^{\uparrow}$H]	*kává*	ka^{5}va^6	'drive'
b.	[M.M]	*kōnī*	ko^{3}ni^3	[H.$^{\uparrow}$H]	*xíní*	ʃi^{5}ni^6	'see'
c.	[M.M]	*kākī*	ka^{3}ki^3	[H.$^{\uparrow}$H]	*xákí*	ʃa^{5}ki^6	'put'
d.	[M.M]	*kā'ndyā*	ka^{3ʔn}d͡ʒa^3	[H.$^{\uparrow}$H]	*xá'ndyá*	ʃa^{5ʔn}d͡ʒa^6	'cut'
e.	[M.M]	*nōmī*	no^{3}mi^3	[H.M]	*nómī*	no^{5}mi^3	'hug'
f.	[M.M]	*ndīkō*	^{n}di^{3}ko^3	[H.M]	*ndíkō*	^{n}di^{5}ko^3	'grind'
g.	[M.M]	*kākā*	ka^{3}ka^3	[H.M]	*xíkā*	ʃi^{5}ka^3	'walk'

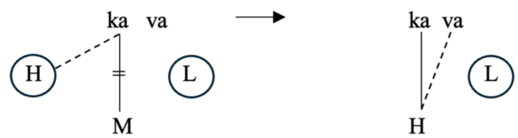
[32] Tone-driven epenthesis is typologically unusual, documented only in a few languages, including Kejom (Babanki; Akumbu et al. 2020), Kifuliiru (van Otterloo 2011), Ghomala' (Bandjoun/Banjun; Nissim 1981; Eichholzer 2010), and Wamey (Santos 1996, Rolle & Merrill 2022) (see Rolle & Merrill 2022 and Rolle 2023 for discussion). Further support for the epenthetic status of [i] comes from the independent observation that an epenthetic [i] is also attested in the adaptation of certain Spanish loanwords, suggesting that high, front unrounded [i] vowels are unmarked in the language. That being said, another possible analysis would be to treat this pattern as involving the affixation of a phonologically-driven empty morph, which hosts grammatical tone only when the stem bears a contour tone in the first TBU. Possible arguments for this alternative are purely diachronic/comparative: an **i*- ~ **y*- durative prefix has been reconstructed for Proto-Mixtec (Kaufman 1988), which can be argued to have become lexicalized in verbal stems in SJPM and other Mixtec varieties in certain cells of verbal paradigms, as it has been proposed for Chalcatongo Mixtec (Macaulay 1996) (see also Hollenbach 2015, Palancar et al. 2016). This formative has been retained in a number of varieties as an aspectual/tense marker, e.g., Alcozauca Mixtec (Swanton & Mendoza Ruiz 2021). Regardless, whether involving epenthesis or an empty morph, the pattern attested is phonologically predictable and results in preservation of lexical tone contrasts in the couplet.

h. [M.M] *kuākū* $k^wa^3ku^3$ [H.M] *xákū* $ʃa^5ku^3$ 'cry'

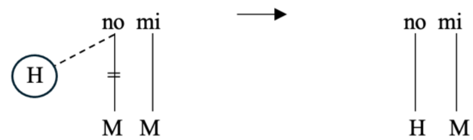
We attribute this contrast to differences in (under)specification of verb stems: in the absence of any derivational prefixes, the grammatical H floating tone docks on the first TBU of the couplet and then spreads rightward onto the second, toneless TBU in $/M.Ø^L/$ stems, as represented in autosegmental terms in (19a). Conversely, the present tense forms of fully specified /M.M/ stems are [H.M], because rightward H tone spreading does not target tonally specified TBUs, shown in (19b).

(19) Present tense tone patterns in $/M.Ø^L/$ and /M.M/ stems

a.

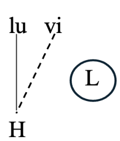


b.



Notably, this analysis generalizes beyond cases of grammatical H tone. For instance, we claim that rightward H tone spreading is also responsible for the surface $[H.^{\uparrow}H]$ tone pattern of monomorphemic (non-verbal) $/H.Ø^L/$ stems, as exemplified in (20) with *lúví* ($[lu^5vi^{6+1}]$) 'turtle'. In §5.4, we will also show that our analysis extends to the grammatical tone patterns of negative polarity inflected forms of $/M.Ø^L/$ verbs.

(20) Rightward H tone spreading in $/H.Ø^L/$ stems



An alternative analysis, which we argue against, also involves positing different underlying representations of the two types of M-toned stems, but without underspecification. Suppose a treatment under which $/M.M^L/$ stems involve a single, doubly-linked M tone (followed by a floating L tone), while /M.M/ stems have each TBU independently linked to a M tone. Upon attachment of the grammatical H tone, this H tone would replace the doubly-linked M tone, yielding a form where both TBUs surface with a H tone ($[H.^{\uparrow}H]$); in contrast, in stems with two independently linked M tones, only the first M tone would be replaced by the grammatical H tone pattern ([H.M]). This alternative, however, can be ruled out, since it incorrectly predicts unattested tonal realizations upon attachment of other grammatical tone patterns, including negative LH. Specifically, a floating LH tone pattern would be predicted to surface as [L.H] in these doubly-linked M stems (assuming a left-to-right mapping of the LH melody). However, as discussed in detail in §5.4 below, attachment of the negative LH tone to such stems results in a surface [LH.H] pattern (with rightward H tone spreading), contrasting with /M.M/ stems, which surface as [LH.M] when inflected for negative polarity. Therefore, the (under)specification analysis yields the correct empirical generalizations and allows us to unify the grammatical tone patterns of present and negative inflected forms, as well as the tone patterns of monomorphemic H-toned stems that bear floating L tones.[33]

[33] In addition, our proposed analysis allows us to unify these grammatical tone patterns with tonal processes attested in the stem-enclitic juncture, where there is evidence of rightward L tone spreading from the stem to toneless enclitics in the language (Caballero, Juárez Chávez & Yuan to appear, Duarte Borquez 2022).

In sum, we have detailed the patterns that arise from the association of a floating H grammatical tone, encoding present tense. This H tone docks to the first TBU of the stem and may: (i) replace L and M stem tones, (ii) apply vacuously to H stem tones, and (iii) dock to an epenthetic vowel if the stem bears a contour tone. We have also illustrated a pattern of rightward H-tone spreading that takes place in inflected /M.Ø^{L}/ stems, which, in turn, provides further evidence for such toneless TBUs.

5.3 Past tense inflected verbs

Past tense in SJPM may also be encoded through a grammatical pattern—a L floating tone that docks on the leftmost TBU of the morpheme within its scope. However, past tense may also be encoded with a *ni*^{1}- prefix, which is described as 'completive' in other Mixtec varieties (Pike 1944, Macaulay 1996), and proposed to be the diachronic source of the floating L grammatical tone across Mixtec (Hollenbach 2003). We analyze the *ni*^{1}- prefix as an allomorph of past tense in SJPM, as the segmental and tonal morphs are both semantically equivalent for native speakers and generally in free variation, except in particular tonal contexts.

As shown in Table 9, both allomorphs are optionally attested for past-inflected verbs in most contexts (e.g., (a-d)). However, in cases where the first TBU of the stem is L, the tonal exponent of past tense generally may not attach, as in (e-g)—unless the L-initial stem displays an irrealis/realis stem alternation, as in (h). Lastly, the tonal exponent may not attach to stems with a rising tone on the first TBU, as in (i).[34] We take this to be a type of phonologically conditioned allomorphy (e.g., see also Mendoza Ruiz (2016) and Uchihara & Mendoza Ruiz (2022) for Alcozauca Mixtec). We illustrate these patterns with bimoraic verb stems.

Table 9: Tonal and segmental allomorphs of past tense in bimoraic stems

	Verb tone		**Past *ni*^{1}- allomorph**		**Past L allomorph**		
a.	M.M	*ndīkō* ^{n}di^{3}ko^{3}	L-M.M	*nìndīkō* ni^{1}-^{n}di^{3}ko^{3}	**LM.**M	*ndìīkō* ^{n}di^{13}ko^{3}	'ground'
b.	M.H	*kē'é* ke^{3ʔ}e^{5}	L-M.H	*kē'é* ni^{1}-ke^{3ʔ}e^{5}	**LM.**H	*kèē'é* ke^{13ʔ}e^{5}	'checked (by touch)'
c.	M.L	*kū'vì* ku^{3ʔ}vi^{1}	L-M.L	*nìkū'vì* ni^{1}-ku^{3ʔ}vi^{1}	**LM**.L	*kùū'vì* ku^{13ʔ}vi^{1}	'got sick'
d.	H.M	*sá'ā* sa^{5ʔ}a^{3}	L-H.M	*nìsá'ā* ni^{1}-sa^{5ʔ}a^{3}	**LH**.M	*sàá'ā* sa^{15ʔ}a^{3}	'made/did'
e.	L.M	*kàkū* ka^{1}ku^{3}	L-L.M	*nìkàkū* ni^{1}-ka^{1}ku^{3}		NA	'survived'
f.	L.L	*kà'àn* kã^{1ʔ}ã^{1}	L-L.L	*nìkà'àn* ni^{1}-kã^{1ʔ}ã^{1}		NA	'spoke'
g.	L.L	*chà'ndyà* t͡ʃa^{1ʔn}dʒa^{1}	L-L.L	*nìchà'ndyà* ni^{1}-t͡ʃa^{1ʔn}dʒa^{1}		NA	'got cut'
h.	L.L	*kuàkù* k^{w}a^{1}ku^{1}	L-L.L	*nìxàkù* ni^{1}-ʃa^{1}ku^{1}	**L**.L	*xàkù* ʃa^{1}ku^{1}	'laughed'

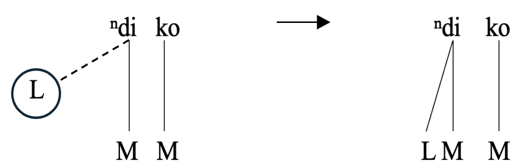
[34] As also pointed out by a reviewer, the epenthetic vowel that surfaces in comparable present and negative contexts does not occur here, presumably because of the independently available segmental past tense allomorph.

i.	LH.H^{L}	*xààní* ʃa^{15}ni^{6+1}	L-LH.H	*nìxààní* ni^{1}-ʃa^{15}ni^{6+1}	NA	'dreamt'

As seen in the examples in (e)-(g), the L grammatical tone pattern may undergo tonal blocking: native speakers' metalinguistic judgments are that only a segmentally-marked alternative exists to encode past tense in these contexts. No such blocking effect occurs in (h), however, as the segmental stem allomorphy allows for the disambiguation between irrealis (*kuàkù* [k^{w}a^{1}ku^{1}]) and past (*xàkù* [ʃa^{1}ku^{1}]). This blocking pattern is moreover in contrast to the vacuous application of H tone in causative stems, exemplified in (16b) in §5.2 above, where future and present tense interpretations are available for the same (homophonous) forms.

In the examples shown so far, the past tense floating L tone concatenates with surface M and H tones, yielding rising contours in the first TBU of bimoraic couplets. We analyze this process as concatenation of the grammatical tone to the lexical tone, a process represented in (21) with *ndīkō* ([^{n}di^{3}ko^{3}]) 'grind'.

(21) Association of L past tense grammatical tone in bimoraic stems



In contrast to bimoraic stems, trimoraic stems display multiple phonologically-conditioned tonal association patterns in the past tense. As we show in this section, the metrical foot or 'couplet' is crucial to the analysis of the tonal patterns in these longer stems. Importantly, these patterns are predictable. First, as shown in (22a-c), the L past tense tone may *replace* (rather than concatenate to) a couplet-external M tone, which surfaces as a level L tone when the immediately following, couplet-initial TBU bears a H or a M tone. Alternatively, it may instead concatenate to a surface M, forming a contour, as in (22d-f): the contour pattern is only attested when the immediately following couplet-internal TBU bears a L tone followed by a M or H tone (including if the L forms a contour with the following tone, as in (22f)).

(22) Grammatical tone patterns of past-inflected trimoraic stems

		Stem		Past tense (/L/)
a.	M(H.H)	*kūsíkí* ku^{3}(si^{5}ki^{5}) 'play'	L(H.H)	*sàsíkí* sa^{1}(si^{5}ki^{5}) 'played'
b.	M(H.M)	*sāchíñōn* sa^{3}(t͡ʃi^{5}ɲõ^{3}) 'work'	L(H.M)	*sàchíñōn* sa^{1}(t͡ʃi^{5}ɲõ^{3}) 'worked'
c.	M(M.M)	*kāndītā* ka^{3}(^{n}di^{3}ta^{3}) 'jump'	L(M.M)	*kàndītā* ka^{1}(^{n}di^{3}ta^{3}) 'jumped'
d.	M(L.M)	*nōkò'ōn īnī* no^{3}(kõ^{1ʔ}õ^{3}) i^{3}ni^{3} 'remember'	**LM**(L.M)	*nòōkò'ōn īnī* no^{13}(kõ^{1ʔ}õ^{3}) i^{3}ni^{3} 'remembered'
e.	M(L.H)	*tāndà'á* ta^{3}(^{n}da^{1ʔ}a^{5}) 'get married'	**LM**(L.H)	*tàāndà'á* ta^{13}(^{n}da^{1ʔ}a^{5}) 'got married'

f.	M(LM.Ø^{L})	*nōtàā'ā* no^{3}(ta^{13ʔ}a^{3+1}) 'ought to, need to'	**LM**(LM.Ø^{L})	*nòōtàā'ā* no^{13}(ta^{13ʔ}a^{3+1}) 'needed to'

The set of patterns shown in monomorphemic trimoraic forms is also found in derived (multimorphemic) trimoraic contexts. A M-toned derivational morpheme inflected for the past tense similarly surfaces with a level L tone when concatenating with a H- or M-initial bimoraic verbal stem (23a-b) but again surfaces with a LM contour if the first TBU of the bimoraic stem bears a L tone followed by a M or H tone (23c-d). Moreover, the grammatical L tone systematically concatenates with the H tone of derivational prefixes, e.g., causative *sa*^{5}- (23e). This is consistent with the patterns attested elsewhere in the language where L tones do not replace H or M tones.

(23) Grammatical tone patterns of past-inflected derived trimoraic stems

		Stem		**Past tense (/L/)**
a.	M-(H.M)	*ndūñó'ōn* ^{n}du^{3}-(ɲõ^{5ʔ}õ^{3}) IRR.INCH-lost 'get lost'	L-(H.M)	*ndùñó'ōn* ^{n}du^{1}-(ɲõ^{5ʔ}õ^{3}) PST.INCH-lost 'got lost'
b.	M-(M.M)	*ndākātā* ^{n}da^{3}-(ka^{3}ta^{3}) IRR.REP.sing 'sing again'	L-(M.M)	*ndàkātā* ^{n}da^{1}-(ka^{3}ta^{3}) PST.REP-sing 'sang again'
c.	M-(L.M)	*ndākuà'ā* ^{n}da^{3}-(k^{w}a^{1ʔ}a^{3}) IRR.REP-deliver 'deliver again'	**LM-**(L.M)	*ndàākuà'ā* ^{n}da^{13}-(k^{w}a^{1ʔ}a^{3}) PST.REP-deliver 'delivered again'
d.	M-(L.H)	*ndānòná* ^{n}da^{3}-(no^{1}na^{5}) IRR.REP-open 'open again'	**LM-**(L.H)	*ndàānòná* ^{n}da^{13}-(no^{1}na^{5}) PST.REP-open 'opened again'
e.	H-(M.M)	*sákuīsō* sa^{5}-k^{w}i^{3}so^{3} CAUS-carry 'load'	**LH-**(M.M)	*sàákuīsō* sa^{15}-(k^{w}i^{3}so^{3}) PST.CAUS-carry 'loaded'

Finally, a third pattern emerges in past-inflected trimoraic forms when *both* TBUs of the bimoraic couplet are L-toned. In such contexts, the initial M-toned TBU outside of the couplet is realized with a level L tone—but the first TBU of the couplet surfaces as M rather than the expected L. Examples of this third pattern are provided in (24) below. Again, this pattern holds regardless of whether the trimoraic form is monomorphemic, as in (24a-b), or if it contains a derivational morpheme attached to a bimoraic root, as in (24c-d).

(24)		**Stem**		**Past /L/**
a.	M(L.L)	*kīkòtò* ki^{3}(ko^{1}to^{1}) IRR.visit 'visit'	<u>**L**</u>(<u>**M**</u>.L)	*xìkōtò* **ʃi^{1}**(**ko^{3}**to^{1}) PST.visit 'visited'

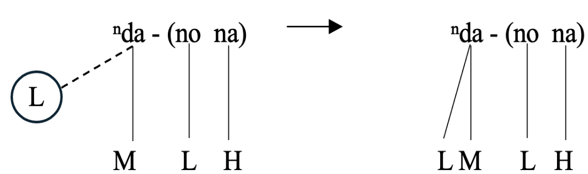
b.	M(L.L)	*kīxàà* ki^{3}ʃa^{1}a^{1} IRR.arrive 'arrive'	<u>L(**M**</u>.L)	*kìxāà* **ki^{1}**(**ʃa^{3}**a^{1}) PST.arrive 'arrived'
c.	M-(L.L)	*kūñò'mà* ku^{3}-(ɲo^{1ʔ}ma^{1}) IRR.INCH-smoke 'get smoked'	<u>L-(**M**</u>.L)	*kùñō'mà* **ku^{1}**-(**ɲo^{3ʔ}**ma^{1}) PST.INCH-smoke 'got smoked'
d.	M-(L.L)	*ndākùsù* nda^{3}-(ku^{1}su^{1}) IRR.REP-sleep 'sleep again'	<u>L-(**M**</u>.L)	*ndàkūsù* **^{n}da^{1}**-(**ku^{3}**su^{1}) PST.REP-sleep 'slept again'

Altogether, these surface tone patterns are fully predictable based on both tonal context and prosodic domain. We argue that the placement and distribution of grammatical tone is partially dependent on prosodic structure, including the metrical foot, which is the domain of various phonological processes and static phonological constraints in the language, as discussed in §3.2 above. We offer three concrete observations based on the above examples. First, whereas the L grammatical tone systematically forms LM contours when associating *within* the couplet, these LM contours are restricted *outside* the couplet to occur only when a sequence of two L tones, separated by a couplet boundary, would be expected instead. Second, whether the resulting LML patterns involve a LM contour or a sequence of level L and M tones depends on the tone of the second TBU within the couplet. Finally, sequences of L tones are tolerated in various other contexts, such as within the couplet and in non-derived contexts, given the existence of L-L.L forms such as *nìkà'àn* ([ni^{1}-kã^{1ʔ}ã^{1}]) 'slept'.

We argue that these observations are interrelated, arising from the interaction of two different constraints: (i) a general constraint against LM contours outside of the metrical foot[35] and (ii) an Obligatory Contour Principle (OCP) constraint militating against sequences of L tones. The former constraint accounts for why grammatical L tone generally replaces the leftmost M tone in trimoraic stems. However, tonal replacement is blocked in certain phonological environments: LM contours are attested outside of the metrical foot only when their presence helps avoid the OCP violation.

First, consider the derivation in (25) below. Here, replacing the lexical tone of the first TBU with grammatical L tone would introduce an OCP violation, given that the adjacent TBU, the first TBU of the couplet, bears L tone as well (arising in a [L(L.X)] configuration). Therefore, the grammatical L tone instead concatenates with the lexical M tone of the initial TBU resulting in a contour that prevents the OCP violation.

(25) L tone dissimilation though contour formation

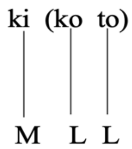


In the case of trimoraic stems containing L-toned couplets (/M(L.L)/), however, contour formation is not attested (the past form is [L(M.L)], not *[LM(L.L)]). We claim this is due to the status of L-toned couplets

[35] We note that, while we have not yet conducted an acoustic study (which we believe is outside of the scope of this paper), we do anticipate that syllables preceding the metrical foot may be shorter in duration, thus precluding tonal contours outside of this prosodic unit and giving rise to a *LM constraint.

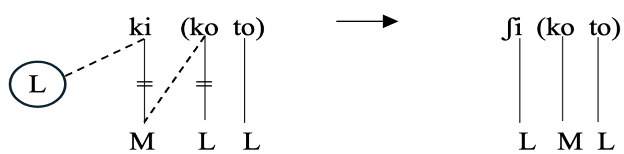
as having two adjacent, independently linked L tones, incurring a violation of the OCP in their underlying representation, as shown in (26). As we argue below, violations of the OCP are tolerated in underlying representations but banned in phonologically derived environments.

(26) Underlying representation of L-toned couplets



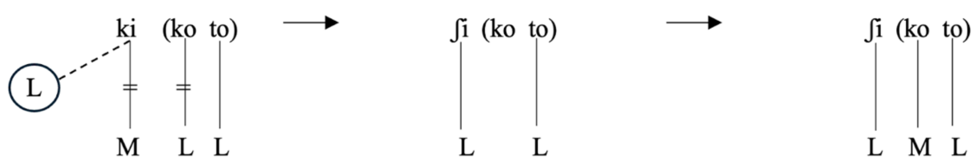
We are agnostic about the precise mechanism by which the [L(M.L)] past form is achieved, and outline two possible analyses below; both are compatible with tonal processes attested elsewhere in the language. First, the relevant process may be analyzed as a tone shift: under this approach, upon docking of the grammatical L tone pattern, the lexical M tone of the initial TBU spreads rightward, delinking the initial L tone of the couplet from the initial mora. The lexical M tone is itself delinked from the initial TBU. This is exemplified in (27). As in the cases of L tone dissimilation through contour formation schematized in (25), this tone shift would avoid a violation of the OCP while preserving a lexical tone contrast: the lexical tone of the first TBU of the stem is realized within the couplet.

(27) L tone dissimilation through tone spreading and delinking



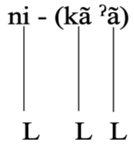
Alternatively, these forms may be analyzed as involving delinking of the L tone of the couplet-initial TBU; the newly toneless TBU would then acquire surface pitch through default M tone insertion, a process that is independently attested in the language (Caballero, Juárez Chávez & Yuan to appear, Duarte Borquez 2022). This alternative is exemplified in (28). Delinking of the L tone of the couplet-initial TBU repairs the OCP-L violation, while allowing maintenance of the lexical L tone pattern in the second TBU of the couplet.

(28) L tone dissimilation through tone delinking and default M tone insertion



Finally, it is important to note that this violation of the OCP is tolerated in *non-derived environments*—it is only ruled out when the association of floating L tones (and their replacement of lexical tone) creates the environment for the OCP restriction against two consecutive L tones in the relevant prosodic context. That is, we propose that repairs to OCP-L violations in SJPM involve a type of *phonologically-derived environment effect*, a phenomenon whereby a phonological pattern arises by virtue of another phonological alternation applying first, creating a new phonological environment (for discussion, see Inkelas 2014, *inter alia*). Indeed, as mentioned above, forms such as *nìkà'àn* [ni^{1}-kã^{1ʔ}ã^{1}] 'spoke' are permitted, as L tones linked to their respective morphemes do not trigger L tone dissimilation, as illustrated in (29). In other words, sequences of L tones are attested across the couplet boundary in the absence of the application of a grammatical tone process.

(29) No dissimilation of underlyingly linked L tones



Overall, the variable realization of the grammatical tone patterns in past inflected verbs is fully phonologically predictable, with both concatenation with and replacement of lexical tones by grammatical tones being driven by tonotactic well-formedness constraints in the language.

Thus far, we have documented the patterns that result from the application of H and L grammatical tones, encoding present and past tense, respectively. These reveal two asymmetries in the grammatical tonal system of SJPM. First, grammatical H generally replaces the first mora of stem tones, while grammatical L may replace or concatenate with them, depending on the tone pattern and prosodic size of the stem. We aim to further account for these facts in §6 below. However, we turn to the tonal patterns seen in negative inflected verbs in §5.4 next, as these both extend and corroborate the generalizations we have formed thus far.

5.4 Negative inflected verbs

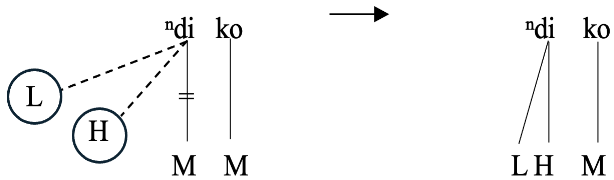
Finally, as with present and past tense, negative polarity in SJPM may be encoded tonally, through a grammatical LH tonal allomorph in the irrealis mood. Negative polarity may also be encoded segmentally, through a marker *ko*^{15} (bearing the same melody as the tonal allomorph), in realis contexts (this alternation is also attested elsewhere in Mixtec; see Hollenbach 2015, Palancar et al. 2016, Swanton & Mendoza Ruiz 2021, and Uchihara & Mendoza Ruiz 2022).

Like the tonal exponents of present and past tense, the negative polarity tonal melody associates with the first TBU of the morpheme within its scope. This LH grammatical tone pattern replaces the lexical tone of the first TBU of verb stems, as shown in (30). Moreover, tonal replacement of lexical tones by grammatical tones is attested in all examples below, regardless of whether the first TBU of the base bears a M (30a-d), a L (30e-f), or a H (30g) tone, and regardless of whether the stem is bimoraic (30a-c, e-g) or trimoraic (30d).

(30)	**Stem**			**Negative irrealis (/LH/)**			
a.	M.M	*ndīkō*	^{n}di^{3}ko^{3}	**LH**.M	*ndìíkō*	^{n}di^{15}ko^{3}	'grind'
b.	M.L	*kū'vì*	ku^{3ʔ}vi^{1}	**LH**.L	*kùú'vì*	ku^{15ʔ}vi^{1}	'get sick'
c.	M.H	*kātón*	ka^{3}tõ^{5}	**LH**.H	*kààtón*	ka^{15}tõ^{5}	'tie'
d.	M.M.M	*kāndītā*	ka^{3n}di^{3}ta^{3}	**LH.**M.M	*kààndītā*	ka^{15n}di^{3}ta^{3}	'jump'
e.	L.L	*kùsù*	ku^{1}su^{1}	**LH.**L	*kùúsù*	ku^{15}su^{1}	'sleep'
f.	L.M	*kìkū*	ki^{1}ku^{3}	**LH**.M	*kìíkū*	ki^{15}ku^{3}	'weave'
g.	H.M	*sá'ā*	sa^{5ʔ}a^{3}	**LH**.M	*sàá'ā*	sa^{15ʔ}a^{3}	'do/make'

We will show throughout this section that the negative grammatical tone pattern involves a sequence of two tonal primitives, /L/ and /H/, which both associate to the left edge of the stem, rather than a true /LH/ contour. The representation of this process is shown in (31) below.

(31) Association of LH negative polarity grammatical tone to *ndīkō* (/^{n}di^{3}ko^{3}/) ‘grind’



As also attested in (positive) present tense-inflected forms, the negative LH melody associates with an epenthetic vowel if the stem bears a lexical contour tone in the first TBU, as shown in (32). This, again, can be understood as a response to the inability for grammatical tones to replace lexical contour tones. Notably, the segmental allomorph of the negative marker is not deployed in these contexts, which we attribute to its morphosyntactic conditioning, i.e., the segmental negative marker *ko^{15}* is precluded to occur in irrealis contexts.[36]

(32)	**Stem**			**Negative irrealis (/LH/)**			
a.	LH.$^{\uparrow}$H^{L}	*xàání*	ʃa^{15}ni^{6+1}	**LH**.LH.$^{\uparrow}$H^{L}	*ìíxàání*	i^{15}ʃa^{15}ni^{6+1}	‘dream’
b.	LH.$^{\uparrow}$H^{L}	*yàákón*	ʒa^{15}kõ$^{6+1}$	**LH**.LH.$^{\uparrow}$H^{L}	*ìíyàákón*	i^{15}ʒa^{15}kõ$^{6+1}$	‘peel with knife’
c.	LH.M	*tìívī*	ti^{15}vi^{3}	**LH**.LH.M	*ìítìívī*	i^{15}ti^{15}vi^{3}	‘lose’

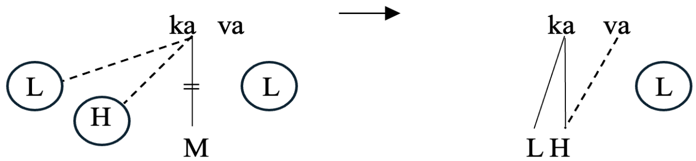
Recall also that a subclass of verb stems are posited in this paper to be underlyingly /M.Ø^{L}/, and that these verbs are realized with a [H.↑H] tone pattern in the present tense due to H tone spreading. The examples in (33a-d) below show that a comparable pattern arises when the hypothesized /M.Ø^{L}/ verb forms are in the negative irrealis: the first TBU of the verb stem bears the LH negative grammatical tone melody, with the H tone then spreading to the second TBU (again, the second H is upstepped). Once again, on par with their present-inflected counterparts, the negative inflected /M.M/ forms in (33e-f) do not display H tone spreading to the second TBU.

(33)	**Stem**			**Negative irrealis (/LH/)**			
	/M.Ø^{L}/ verb forms						
a.	[M.M] *kāvā*		ka^{3}va^{3+1}	[LH.$^{\uparrow}$H]	*kàává*	ka^{15}va^{6+1}	‘drive’
b.	[M.M] *kōnī*		ko^{3}ni^{3+1}	[LH.$^{\uparrow}$H]	*kòóní*	ko^{15}ni^{6+1}	‘see’
c.	[M.M] *kākī*		ka^{3}ki^{3+1}	[LH.$^{\uparrow}$H]	*kàákí*	ka^{15}ki^{6+1}	‘put’
d.	[M.M] *kā’ndyā*		ka$^{3\text{ʔn}}$d͡ʒa^{3+1}	[LH.$^{\uparrow}$H]	*kàá’ndyá*	ka$^{15\text{ʔn}}$d͡ʒa^{6+1}	‘cut’
	/M.M/ verb forms						
e.	[M.M] *nōmī*		no^{3}mi^{3}	[LH.M]	*nòómī*	no^{15}mi^{3}	‘hug’
f.	[M.M] *ndīkō*		^{n}di^{3}ko^{3}	[LH.M]	*ndìíkō*	^{n}di^{15}ko^{3}	‘grind’

H tone spreading in negative inflected forms of /M.Ø^{L}/ stems such as *kāvā* ([ka^{3}va^{3+1}]) ‘drive’ is represented in (34). As mentioned above, that only H spreads can be taken as evidence that the LH melody is not a true contour.

[36] In contrast, and as noted by an anonymous reviewer, in Yoloxóchitl Mixtec, verb stems with a tonal contour in the first mora are marked with a segmental exponent (a negative adverb *kwa^{14}*, cognate of SJPM *ko^{15}*), even though this marker is used in irrealis contexts elsewhere (Palancar et al. 2016). In incompletive stems (equivalent to SJPM present tense verb forms), verbs with a contour tone in the first mora attach a high front vowel [i] with a H tone, which Palancar et al. analyze as an archaic incompletive prefix that has a tonally restricted distribution (2016:17).

(34) Rightward spreading of H tone after attachment of negative LH



The examples shown thus far all involve negative irrealis forms, which involve the grammatical LH tone pattern directly interacting with the lexical tones of verb stems. Additional tonal and segmental patterns obtain in other negative-inflected verbs in other tense configurations. Crucially, the interactions between the LH grammatical tone melody and the tonal exponents of tense (L and H for past and present respectively) are restricted, since the tonal and segmental allomorphs of negation have morphosyntactically conditioned distributions. As illustrated in Table 10, the tonal LH allomorph is used with irrealis stems, while the segmental allomorph *ko*^{15}- is used with realis stems.

Table 10: TAM and polarity contrasts in the verb stem /M.M/ *ndīkō* [^{n}di^{3}ko^{3}] 'grind'

	Positive polarity		**Negative polarity**		
Irrealis	M.M	*ndīkō* ^{n}di^{3}ko^{3}	**LH**.M	*ndìíkō* ^{n}di^{15}ko^{3}	*Future (irrealis)*
Realis	**H**.M	*ndíkō* ^{n}di^{5}ko^{3}	**L**-**H**.M	*kòndíkō* ko^{1}-^{n}di^{5}ko^{3}	*Present*
	L-M.M	*nìndīkō* ni$^{1\text{-}n}$di^{3}ko^{3}	**L**-**H**-M.M	*kòníndīkō* ko^{1}-ni^{5}-^{n}di^{3}ko^{3}	*Past ni*^{1}-
	LM.M	*ndìīkō* ^{n}di^{13}ko^{3}	**LH**-M.M	*kòóndīkō* ko^{15}-^{n}di^{3}ko^{3}	*Past*

Multiple distinct tonal patterns are attested in the realis forms. First, the negative present is encoded with L tone on the segmental allomorph *ko*- and H tone on the first TBU of the verb stem. The negative past reveals two additional patterns. If the segmental allomorph attaches to the segmental past tense prefix *ni*^{1}-, *ko*- again only bears L tone and the H tone of the negative polarity tonal melody is realized on the latter prefix. If the *ko*- prefix attaches directly to the verb stem, it bears the full LH tonal melody and the verb stem appears without the grammatical L tone.

What is notable about the former two patterns in the present and past, respectively, is the loss of the H tone on the negation segmental morph. We analyze these patterns as involving a phonologically general process of (progressive) *tonal absorption*, where, in underlying /LH-H/ sequences across the couplet boundary or in TBUs preceding the couplet, there is simplification of the LH contour, becoming L before the H tone, /LH-H/ → [L-H]. We propose that this process takes place straightforwardly in the negative present, e.g., /**LH**-**H**.M/ /ko^{15}-^{n}di^{5}ko^{3}/ → **L**-**H**.M [ko^{1}-^{n}di^{5}ko^{3}] 'not grinding'.[37] In the case of negative past stems with the past *ni*^{1}- prefix, we posit that the H tone spreads from the segmental marker *ko*^{15}- onto the

[37] An anonymous reviewer suggests that an alternative to our proposed tonal absorption process is to simply posit that the negative prefix has a L-toned allomorph (*ko*^{1}-) that arises in the context of the present tense. However, this morphological alternative misses important phonological and prosodic generalizations (discussed below) that can capture the distributions of these forms without appealing to allomorphy.

stem, replacing the lexical tone of the *ni^{1}*- prefix (LH-L → LH-H), as shown in (35a). This process then yields the context for tonal absorption (LH-H → L-H), as shown in (35b).

(35)	a.	H-tone spreading			
		LH-L-M.M	→	**LH-H**-M.M	
		ko^{15}-ni^{1}-(^{n}di^{3}ko^{3})	→	ko^{15}-ni^{5}-(^{n}di^{3}ko^{3})	'did not grind'
	b.	Tonal absorption			
		LH-H-M.M	→	**L-H**-M.M	
		ko^{15}-ni^{5}-(^{n}di^{3}ko^{3})	→	ko^{1}-ni^{5}-(^{n}di^{3}ko^{3})	'did not grind'
				kòníndīkō	

We note that tonal absorption is a tonal process that is prosodically restricted. Specifically, this process is attested in the boundary between the negative segmental prefix and the stem, whether across the couplet boundary as in the negative present (/**LH-(H**.X)/ → **L-(H**.X)), or preceding the couplet as in the negative past with the *ni^{1}*- prefix (/LH-L-(X.X)/ → **L-H-**(X.X)). In other prosodic domains, e.g., within the metrical foot (equivalent to the couplet), there is no tonal absorption. For instance, recall that when a LH grammatical tone pattern docks to the first TBU of bimoraic stems with a final H tone, the LH-H pattern persists, e.g., /M.H/ /(ka^{3}tõ^{5})/ → LH.H [(ka^{15}tõ^{5})] 'won't tie.' Likewise, there are lexical LH.H patterns in non-derived forms, e.g., *chòó'ó* [t͡ʃo^{15ʔ}o^{5}] 'flea.' This pattern further exemplifies the crucial role that metrical structure plays in the analysis of tonal processes in SJPM.

The third negative pattern, attested in the past tense, exhibits a LH contour in the segmental negative marker *ko^{15}*-, followed by the stem bearing its lexical tones. Notably, this pattern does not exhibit the expected layering of tonal markers, given the absence of the L tonal exponent in the first TBU of the stem (as otherwise attested in the positive polarity), e.g., LH-M.M *ko^{15}-^{n}di^{3}ko^{3}* rather than *LH-**L**.M **ko^{15}-^{n}di^{13}ko^{3}*). These forms could be argued to exhibit a paradigm uniformity effect between the two types of negative past forms (with and without the past *ni^{1}*- prefix), though we leave further investigation of this idea for future work.[38]

Finally, for the verb stems that bear a lexical contour tone in the first TBU, the epenthetic vowel seen above again surfaces in the negative present. As shown in (36a-b), the segmental morph *ko*- bears only a L tone rather than the expected LH melody, while the epenthetic vowel bears a H tone. This pattern is consistent with the analysis of the negative present developed above, in that a posited underlying LH-H sequence is simplified via (progressive) tonal absorption (e.g. /**LH-H**-LH.M/ /ko^{15}-i^{5}-ti^{15}vi^{3}/ → **L.H**.LH.M [ko^{1}-i^{5}-ti^{15}vi^{3}]).

(36)		**Irrealis**			**Negative present**		
a.	'dream'	LH.H^{L}	*xàání*	ʃa^{15}ni^{6+1}	L.H.LH.H^{L}	*kòíxàání*	ko^{1}-i^{5}-ʃa^{15}ni^{6+1}
b.	'lose'	LH.M	*tììvī*	ti^{15}vi^{3}	L.H.LH.M	*kòítììvī*	ko^{1}-i^{5}-ti^{15}vi^{3}

The negative past forms of these verbs are given in (37a-b); these too follow from the analysis of the negative past presented above: the H tone of the negative LH melody is either realized on the prefix *ni*-, if present, or on the negative morph itself.

(37)		**Irrealis**			**Negative past**		
a.	'dream'	LH.H^{L}	*xàání*	ʃa^{15}ni^{6+1}	L.H.LH.H^{L}	*kòníxàání*	ko^{1}-ni^{5}-ʃa^{15}ni^{6+1}

[38] In particular, the stem tone patterns in both types of forms are identical, whether they attach the past prefix or not, e.g., L-H-(**M.M**) [ko^{1}-ni^{5}-^{n}di^{3}ko^{3}], LH.(**M.M**) [ko^{15}-^{n}di^{3}ko^{3}] 'did not grind' (in both cases, the stem tone patterns are the lexical tones of roots). We note that these effects are not attested elsewhere in the language, but only in negative past cells of verbal paradigms, and could have arisen through a diachronic process of paradigm leveling. Alternatively, the input to negative polarity marking is a tense-unmarked stem, which would explain the morphological distribution of the tonal and segmental allomorphs of negation, with the former restricted to future-tense marked forms and the latter to past and present forms.

					LH.LH.H^{L}	*kòóxàání*	ko^{15}-ʃa^{15}ni^{6+1}
b.	'lose'	LH.M	*tìívī*	ti^{15}vi^{3}	L.H.LH.M	*kònítìívī*	ko^{1}-ni^{5}-ti^{15}vi^{3}
					LH.LH.M	*kòótìívī*	ko^{15}-ti^{15}vi^{3}

5.5 Interim summary

This section has detailed the grammatical tone patterns that arise in the present and past tense, as well as in the negative polarity in SJPM. These grammatical tones are analyzed as floating tones that associate to the leftmost TBU of the verb stem, i.e., as concatenative tonal morphemes that exhibit prefix-like behavior. The tonal association patterns are summarized below in (38):

(38) SJPM grammatical tone patterns: summary

- Present tense is encoded with a H tone which may replace M and L tones on the verb stem and may apply vacuously to stems that already bear a H tone in the initial TBU of the stem. However, they may not associate with TBUs that contain a contour, resulting in the insertion of an epenthetic vowel.
- Past tense is encoded with a L tone or with a segmental allomorph *ni^{1}*-. The tonal morph yields different patterns depending on the phonological and prosodic context. In couplet-initial position, it concatenates with M and H, creating contours; (vacuous) association with a L TBU is blocked, resulting in a segmental allomorph. However, in pre-couplet position, the relevant TBU may either surface with a contour or simply with L, depending on the tonal pattern of the stem.
- Finally, negative polarity has a LH tonal melody, which we have decomposed into a sequence of L and H. The LH melody may associate to the verb stem, a segmental prefix, or both, depending on the tense specification of the stem.

Throughout this section, we have also made a number of analytical claims, on the basis of the exact tonal association patterns that arise in different contexts. For instance, we provided evidence for positing the existence of tonally unspecified TBUs, based on a pattern of H tone spreading across couplets. Likewise, we identified several interacting tonal constraints and processes that occur in some prosodically defined environments but not others, revealing the significance of prosodic domains in this language—specifically, the couplet (equivalent to a metrical foot) and whether individual TBUs are internal or external to this domain. SJPM can thus be added to the list of languages where the tonal system is sensitive to metrical structure, as posited for both related (e.g., Ixtayutla Mixtec (Penner 2019)) and unrelated (Kera (Chadic; Pearce 2006)) languages.

SJPM is representative of Mixtec grammatical tone systems, which have been characterized in the literature as involving a high functional load (Pankratz & Pike 1967, Palancar et al. 2016, DiCanio et al. 2021). This assessment is not made on the basis of the number of tonal exponents available in the language, but rather in terms of their *systematicity* in verb paradigms:[39] all forms of inflected verbs have distinctive tone patterns involving these exponents, which are crucial for retrieving grammatical information (mood and/or tense). Furthermore, some segmental allomorphs for some grammatical categories are tonally conditioned, contributing to the centrality of tone for lexical and grammatical contrasts in this language.

In the remainder of this paper, we provide an account of these lexical-grammatical tone interactions. We argue that the various patterns that arise result from dominance properties of individual tonal exponents, constrained by the general phonotactic restrictions of tonal patterns in the language.

[39] For similar arguments regarding the functional load of grammatical tone in Gyeli (Northwestern Bantu), see Grimm 2022.

6 Lexical-grammatical tone interactions

We have now seen that both lexical and grammatical tone have a high functional load in SJPM, and that grammatical tone patterns exhibit different behaviors in terms of their interaction with lexical tone in different prosodic and tonal contexts. But is there a unifying mechanism that governs the interaction between lexical and grammatical tone in this language? In this section, we examine this question through the lens of a dominance framework of grammatical tone (Rolle 2018). In §6.1, we provide a brief overview of this theory, and propose that the grammatical tones of SJPM are best characterized as *non-dominant*, and that the appearance of dominance of some tonal exponents (associated with replacement of lexical tones by grammatical tone patterns) in particular environments may be responses to various tonotactic constraints—not idiosyncratic properties of individual tonal exponents. In §6.2, we then show that SJPM offers an empirical challenge to a key prediction made by Rolle (2018) regarding how conflicts arising in lexical-grammatical tone interactions are resolved.

6.1 Non-dominant neutral grammatical tone in SJPM

In the dominance framework of grammatical tone of Rolle 2018,[40] the behavior of grammatical tone patterns results from idiosyncratic properties of grammatical tonal exponents in terms of their (in)ability to delete the underlying tone of the target (where the target is understood to be the morpheme(s) where grammatical tone associates): dominant grammatical tone involves systematic deletion of the underlying tone of the target, while no such systematic deletion is attested in non-dominant grammatical tone. Under this view, the typological parameters that shape variation in grammatical tone (GT) systems cross-linguistically can be understood to result from the following tension, from Rolle 2018:10: "[w]ithin dominant GT all outputs have a uniform tone shape which has the advantage of providing a more consistent cue for the grammatical category of the trigger, but sacrifices the lexical contrast of the target. In contrast with non-dominant GT, outputs do not have a uniform form and thus maintain lexical contrast unambiguously, but at the cost of having a less delimited cue for the trigger".[41] Importantly, (non-)dominance is *idiosyncratic*, in that whether tones are dominant or non-dominant cannot be predicted by any other properties of the grammar.

We propose that the SJPM grammatical tone system can be broadly classified as **non-dominant neutral**, which is defined as follows:[42]

(39) Neutral non-dominant grammatical tone: "the lack of automatic replacement/deletion of the underlying tone of the target or automatic non-application of the grammatical tone (i.e. simple toneme concatenation)" (Rolle 2018:4)

In other words, we propose that the grammatical tones in SJPM—present H, past L, and negative LH—are *all* concatenative in nature. Given the particular patterns detailed throughout §5, however, this proposal requires some elaboration. Although we see a clear illustration of this concatenative behavior in cases of attachment of the grammatical L tone pattern encoding past tense in bimoraic stems (where the grammatical L tone creates a contour with the lexical stem tone), this is less immediately obvious for the present H and negative LH tonemes, as they replace lexical tones in most (but not all) contexts. Nonetheless, we argue that the association patterns of present H and negative LH—replacement, blocking, and tonal conditioned epenthesis—are all phonologically predictable rather than idiosyncratic, and thus should not be viewed as dominant. This, in turn, suggests a *uniformly non-dominant* (i.e., concatenative) approach.

[40] This proposal is couched within the more general theoretical treatments of dominance effects in Kiparsky & Halle 1977, Kiparsky 1984, and Inkelas 1998, among others.

[41] In this framework, a 'trigger' is defined as the morpheme or construction which licenses the tonological operation.

[42] Within Rolle's (2018) framework, neutral non-dominant systems stand in contrast to 'recessive' ones, wherein grammatical tones may only associate with tonally unspecified TBUs. This is clearly not the case in SJPM, hence our classification of the SJPM system as neutral.

We posit that the complete range of surface grammatical tone patterns in SJPM can be attributed to a single idiosyncratic pattern of concatenation for *all* grammatical floating tones, with all deviations from this pattern surfacing in phonologically specific contexts as avoidance of tonotactically illicit sequences. We have already argued that some patterns of *concatenation* of grammatical tone with lexical tone can be analyzed as involving avoidance of tonotactically illicit sequences (i.e. LM contours in pre-couplet moras of trimoraic stems, where concatenation avoids an OCP-L violation across the couplet boundary). We add here that the patterns of H and LH grammatical tone association may likewise be analyzed in terms of the phonological grammar: specifically, concatenation of H and LH to form contours is not possible because this would yield either falling or complex contours, both of which are independently ruled out in SJPM (except in very restricted environments), as already established in §4.1. Table 11 summarizes the different types of potential and attested interaction between lexical and grammatical tone in SJPM for each tonal exponent.

Table 11: SJPM lexical-grammatical tone interactions and relationship to phonological grammar

Tonal exponent	**Expected pattern under concatenation**	**Attested interaction with lexical tone**	**Repair of illicit tonal sequences**
Present tense H	*HM, *HL	Replacement of M and L	Avoidance of falling contours in couplet-initial position
	*HLM	Grammatical tone conditioned epenthesis	Inability to alter an input monomoraic tonal contour (tonal indomitability)
Past tense L	*LM(X.X)	Replacement of M	Avoidance of LM contours outside of the couplet
Negative LH	*LHM, *LHL	Replacement of M and L	Avoidance of complex contours in single morae[43]
	*LHLM	Grammatical tone conditioned epenthesis	Inability to alter an input monomoraic tonal contour (tonal indomitability)

This table has a number of components. It specifies what tonotactic violation would result from the expected pattern of concatenation of different tonal exponents to lexical tone: for instance, falling contours would arise in base-initial TBUs with M and L tone upon attachment of present H tone (*HM, *HL), and complex contours would arise in single morae in the case of attachment of negative LH to M and L lexical tones (*LHM, *LHL). This table also illustrates that different tonotactically illicit sequences may be repaired through a single mechanism (e.g., tonal replacement is attested in different contexts involving H, L and LH grammatical tones), though some types of repairs are exclusively attested in specific phonological environments (e.g., grammatical tone conditioned epenthesis is only attested in cases of tonological indomitability of contour tones). There are thus a variety of constraints imposed by the phonological grammar and a variety of repairs available to avoid them.

As mentioned above, it would not be viable to take the SJPM grammatical tone system to involve *both* dominant and non-dominant tonal exponents, with replacement vs. concatenation in grammatical tone patterns resulting exclusively from the status of individual grammatical tones as dominant or recessive. In

[43] In addition, these expected patterns would also involve illicit falling contours in couplet-initial position.

this treatment, negative LH and present H would be dominant and past L would be non-dominant—all idiosyncratic properties of the tonal exponents in question. However, recall that concatenation, replacement and other processes in grammatical tone patterns in SJPM have clear phonological distributions, as shown above. We therefore contend that these distributions are better captured by a uniformly non-dominant approach for all tonal exponents, coupled with general tonotactic constraints in the language.

6.2 Grammatical tone patterns reference morphological properties

With our classification of the grammatical tones of SJPM in place, we now turn to the factors that condition the *choice* of specific repair in each case (e.g., deletion of lexical tone) vs. other logically possible alternatives (e.g., deletion of the grammatical tone, maintaining an unmarked tone, etc.). Within the dominance framework, these results are predicted to be determined entirely by the phonological grammar of each language in neutral grammatical tone systems. That is, while some languages may allow both the grammatical tone and underlying lexical tone to surface, in other languages either the lexical tone or the grammatical tone may be deleted due to general tonal markedness constraints. This leads Rolle to formulate the following predictions if conflicts arise in these systems: "(i) there is no influence based on the morphological identity of the sub-constituents to resolve the conflict, and consequently (ii) the least phonologically marked output will surface to resolve it. The strongest version of this principle is … [the **n]eutral G[rammatical] T[one] conflict resolution**: conflicts in neutral patterns are *entirely resolved via reference to phonological markedness, and never to morphological properties*" (2018:192) (emphasis ours).

Given that the SJPM grammatical tone system is neutral, we now assess whether the strong version of this prediction is borne out in this language. In fact, we propose that SJPM presents an empirical challenge to this principle: the choice of repair in cases where the interaction between lexical and grammatical tone leads to a tonotactically illicit configuration *is* in fact resolved in terms of the morphological properties of the tonemes in conflict. That is, we posit that SJPM is a language where the tonal grammar and availability of both tonal and segmental exponents of inflection allows for preservation of most lexical and grammatical tonal contrasts in verb paradigms—but the preservation of grammatical tone is prioritized over lexical tone if one of the two must be deleted and if there are no morphological or phonological mechanisms to keep both. This logic is evident throughout the lexical-grammatical tone interactions summarized in Table 11 above, which showed the replacement of L and M lexical tones in various inflected environments.

As a concrete example, consider again the tone patterns of /M.M/ verb stems when inflected for present tense, repeated in (40) below: in these cases, the concatenation of grammatical H tone would yield an illicit falling [HM] contour in the first stem mora; grammatical H tone replaces lexical M tone instead.

(40)	**Stem**				**Present tense**		
a.	M.M	*nōmī*	no^{3}mi^{3}	'hug'	**H**.M	*nómī*	no^{5}mi^{3}
b.	M.M	*ndīkō*	^{n}di^{3}ko^{3}	'grind'	**H**.M	*ndíkō*	^{n}di^{5}ko^{3}
c.	M.M	*tīīn*	tĩ^{3}ĩ^{3}	'grab'	**H**.M	*tíīn*	tĩ^{5}ĩ^{3}
d.	M.M	*ñōō*	ɲo^{3}o^{3}	'drag'	**H**.M	*ñóō*	ɲo^{5}o^{3}

Thus, in a context where either the grammatical H tone or the lexical M tone could be deleted, it is the grammatical H tone that survives. Similar asymmetries were shown throughout §5 with grammatical L and LH surviving over lexical tones. Crucially, we take such asymmetries to be *morphological* in nature, in that grammatical tones are morphemes (mappings between individual tones to particular inflectional categories) while lexical tones are not. In other words, this asymmetry cannot be resolved based on purely phonological properties of L, M and H tones, divorced from their lexical vs. grammatical status.

To strengthen this claim, we offer and reject an alternative that would be consistent with Rolle's prediction that the conflict between lexical and grammatical tones should be resolved on entirely phonological grounds. This alternative involves appealing to the relative markedness of tones in SJPM.

Specifically, under this alternative, the tone replacement patterns attested in verb paradigms would be predictable on the basis of the relative markedness of each tonal value, which would be assessed in terms of the (in)ability of each tone to be overwritten in cases of tonal conflict, with marked tones being preserved in outputs over unmarked tones (Pulleyblank 2004). In (40), M is replaced by H, which would suggest a H > M markedness hierarchy under this view. Given that M tones are replaced by L tones in some past tense marked forms, we might also conclude under this alternative that L tones are more marked than M tones, yielding a L > M ranking. Thus, this alternative would allow us to attribute the tonal deletion patterns to the relative markedness of H, L and M tones, which in SJPM would involve a H > L > M scale.

We argue, however, that the consideration of additional tonal processes in SJPM significantly weakens the arguments for positing a markedness analysis of tonal replacement patterns. In particular, a phonological argument for distinguishing marked from unmarked tonal values comes from additional asymmetries in trigger/target patterning in tonal spreading. L, M and H tones may spread in SJPM in different contexts, and each process exhibits differences in terms of possible targets for spreading. Crucially, such patterns reveal a contradictory markedness behavior of M tones in SJPM. In (41), we summarize the properties of L and M tone spreading in SJPM.

(41) Tone spreading of L and M tones in SJPM

a. L tones may spread onto toneless TBUs in the stem-enclitic domain, but not onto tonally specified M and H TBUs (Caballero, Juárez Chávez & Yuan to appear, Duarte Borquez 2022).

b. M tones may spread onto toneless TBUs in /M.Ø^{L}/ stems and other stems with toneless TBUs (§4), but there is no evidence that M ever spreads onto tonally specified TBUs (e.g., L and H).

Given that both L and M tone can only spread onto toneless TBUs, this would suggest that both are relatively unmarked. A third spreading pattern, namely that of H tone spreading, is especially relevant for the assessment of tonal markedness. We have already seen in §4 and §5.2 that H tones may spread onto toneless TBUs, both in monomorphemic /H.Ø^{L}/ stems and in present forms of /M.Ø^{L}/ stems. However, we now show that H tone spreading cannot target M tones. This is demonstrable in the tone patterns below, in which a H tone from a nominal classifier proclitic may spread onto the initial TBU of its host. As shown in (42), H tone may spread onto a L-toned stem TBU—but not a M-toned stem TBU.[44]

(42) a. *rá víxì*
/ɾa^5=vi^{1}ʃi^1/ → [ra^5=**vi**^{5}ʃi^1]
CL.3.LIQ=sweet
'soda' (lit. 'sweet liquid')

tí chínkì
/ti^5=t͡ʃĩ^{1}ki^1/ → [ti^5=**t͡ʃĩ**^{5}ki^1]
CL.3.RND=pine cone
'pine cone'

b. *rá lō'ō*
/ɾa^5=lo^{3ʔ}o^3/ → [ra^5=lo^{3ʔ}o^3]
CL.3.LIQ=small
'small liquid (e.g., puddle)'

tí kīnī
/ti^5=ki^{3}ni^3/ → [ti^5=ki^{3}ni^3]
CL.3.AN=ugly
'bad/ugly animal'

The H tone spreading pattern seen in (42) is informative. Although the ability for H to spread onto L is consistent with a markedness ranking that emerges in tonal replacement patterns, the inability for H to spread onto M contradicts this ranking. Specifically, while the tonal replacement pattern in present tense forms would suggest a H > M hierarchy, the inability for H to spread onto M in (42) would instead suggest M > H.

However, this contradiction could be resolved by simply acknowledging that H-tone has different grammatical sources across these examples: in (42), it is a lexical tone, while, in (40), it is a grammatical tone. We thus conclude that it is indeed the morphological status of H-tone that allows it to survive in (40), rather than relative phonological markedness.

[44] This pattern is also attested in other constructions in the language, including negative adjectives (Duarte Borquez 2022).

In sum, we propose that the availability of both segmental and tonal exponents of morphosyntactic categories, as well as the tonal properties of the language that allow complex contours in single morae, allows preservation of most lexical and grammatical tonal contrasts in SJPM verbal paradigms. Thus, SJPM instantiates a case of a language with both a high lexical role for tone and a complex grammatical tone system. This is notable in light of Rolle's (2018) prediction that such systems, if they exist at all, should be rare cross-linguistically (see also Grimm 2022 for another case and further discussion).

7 Conclusion

This paper has provided the first comprehensive description and analysis of the lexical and grammatical tone systems in SJPM. First, we have identified a number of tonotactic asymmetries in the language, such as the general absence of (non-derived) falling contours and the inability for individual TBUs (morae) to host more than two tones. These are not only revealed by the lexical tone patterns in the language, but are also mirrored by the grammatical tone system, which result from the association of grammatical floating tones. We have argued that tonal concatenation, replacement, and avoidance patterns in lexical-grammatical tone interactions are determined jointly by the status of grammatical tones as non-dominant (in the sense of Rolle 2018), the aforementioned tonotactic restrictions, and various tonal processes that operate in the language. Close examination of lexical-grammatical tone patterns furthermore contributes relevant evidence for positing tonal representation, including tonal (under)specification: we propose SJPM has a typologically unusual /M/ vs. /Ø/ contrast in its tone system, previously proposed for genetically related languages such as Nochixtlán Mixtec (McKendry 2013, 2018) and Sierra Juárez Zapotec (Bickmore & Broadwell 1998), as well as genetically unrelated languages such as Poko Rawo (McPherson & Dryer 2021).

We have argued that the tone system of SJPM is of interest to a developing typology of grammatical tone for a number of reasons: (i) it instantiates a system whereby a variety of tonal patterns emerge in inflectional paradigms with clear phonological conditioning; (ii) while the asymmetries observed in lexical-grammatical tone interactions result from the avoidance of phonotactically marked tonal sequences, we argue that the choice of repair is sensitive to the morphological properties of the tones in conflict, with grammatical tones winning over lexical tones; (iii) it exhibits grammatical tone driven epenthesis, previously thought to be unattested in grammatical tone systems (Rolle 2018; cf. Rolle & Merrill 2022).

Finally, on the level of micro typology, we have alluded to the high degree of diversity attested in Mixtec tone systems, with Mixtec varieties exhibiting differences in terms of how many tones may surface on an individual TBU, and phonological analyses differing in terms of whether tonal representation involves underspecification and the number and nature of tonal sandhi phenomena attested (Daly & Hyman 2007, Carroll 2015, Penner 2019, McKendry 2013, DiCanio & Bennett 2021). Furthermore, the grammatical tone systems and the interaction between lexical and grammatical tone in Mixtec varieties is still largely understudied (cf. Palancar et al. 2016). We seek to contribute to further careful documentation and analysis of these systems, as their systematic comparison will deepen our understanding of the tonal variation within the Mixtecan language family and, more generally, what are possible interactions between grammatical tone and the phonological grammar across Mixtec and beyond.

References

Archangeli, Diana B., & Douglas G. Pulleyblank. 1994. *Grounded phonology*. Cambridge, MA: MIT Press.

Akumbu, Pius, Larry M. Hyman & Roland Kiessling. 2020. The segmental and tonal structure of verb inflection in Babanki. *Language Data and Analysis* 2(2).1-33.

Bonet, Eulàlia & Daniel Harbour. 2012. Contextual allomorphy. In Jochen Trommer (ed.), *The morphology and phonology of exponence*, 195-235. Oxford: Oxford University Press.

Bickford, Albert J. & Stephen A. Marlett. 1988. The semantics and morphology of Mixtec mood and aspect. *Work Papers of the Summer Institute of Linguistic*s, University of North Dakota 32.1-39.

Bickmore, Lee & George Aaron Broadwell. 1998. High tone docking in Sierra Juárez Zapotec. *International Journal of American Linguistics* 64(1). 37–67.

Bradley, C. Henry. 1970. A linguistic sketch of Jicaltepec Mixtec. (Summer Institute of Linguistics Publications in Linguistics and Related Fields, 25.)

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

Campbell, Eric W. 2017. Otomanguean historical linguistics: Past, present, and prospects for the future. *Language and Linguistics Compass* 11(4). e12240.

Campbell, Eric W. & Griselda Reyes Basurto. 2023. El Tu'un Savi (mixteco) en California: documentación y activismo lingüístico. In Marcela San Giacomo, Fidel Hernández Mendoza and Michael Swanton (eds.), Estudios sobre lenguas mixtecanas. México, D.F.: Seminario Permanente de Lenguas Mixtecanas, Instituto de Investigaciones Antropológicas, Universidad Nacional Autónoma de México.

Carroll, Lucien S. 2015. *Ixpantepec Nieves Mixtec word prosody.* Doctoral dissertation, UC San Diego.

Castillo García, Rey. 2007. *La fonología tonal del mixteco de Yoloxóchitl, Gro*. MA thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social, Mexico City, Mexico.

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

Cortés, Felix., Iara Mantenuto & Jeremy Steffman. 2023. San Sebastián del Monte Mixtec. *Journal of the International Phonetic Association* 53(3). 1-22.

Daly, John P. 1978. Notes on Diuxi Mixtec tone. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session* Vol. 22(3).

Daly, John P. & Larry Hyman. 2007. On the representation of tone in Peñoles Mixtec. *International Journal of American Linguistics* 73(2).165-207.

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.

DiCanio, Christian, Jonathan Amith & Rey Castillo García. 2014. The phonetics of moraic alignment in Yoloxóchitl Mixtec. *Proceedings of the 4th Tonal Aspects of Language Symposium*, Nijmegen, the Netherlands. 203–210.

DiCanio, Christian, Joshua Benn & Rey Castillo García. 2021. Disentangling the effects of position and utterance-level declination on the production of complex tones in Yoloxóchitl Mixtec. *Language and Speech* 64(3). 515-557.

DiCanio, Christian & Ryan Bennett. 2021. Prosody in Mesoamerican languages. In Ajou Chen & Carlos Gussenhoven (eds.), *The Oxford Handbook of Language Prosody*. Oxford University Press.

DiCanio, Christian, Caicai Zhang, Douglas H. Whalen & Rey Castillo García. 2020. Phonetic structure in Yoloxóchitl Mixtec consonants. *Journal of the International Phonetic Alphabet* 50(3). 333-365.

Duarte Borquez, Claudia. 2022. The representation of tone in San Juan Piñas Mixtec. Comprehensive research paper, UC San Diego.

Duarte Borquez, Claudia. 2023. Tone, prosodic structure and grammatical structure in
San Juan Piñas Mixtec (Tò'ōn Ndá'ví). Qualifying paper, UC San Diego

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: D. K. E. Reisinger (ed.), *Proceedings of WSCLA 26*, UBC Working Papers in Linguistics.

Dürr, Michael. 1987. A Preliminary Reconstruction of the Proto-Mixtec Tonal System. Indiana 11. 19-62. Berlin.

Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2023. *Ethnologue: Languages of the World*. Twenty-sixth edition. Dallas, Texas: SIL International. http://www.ethnologue.com.

Eichholzer, Erika. 2010. Dictionnaire ghomálá'–français. Manuscript. University of Zürich.
Eischens, Benjamin. 2022. *Tone, phonation, and the phonology-phonetics interface in San Martín Peras Mixtec.* Doctoral dissertation, University of California Santa Cruz.
Gerfen, Chip. 1999. *Phonology and phonetics in Coatzospan Mixtec*. In *Studies in natural language and linguistic theory.* Vol.48. Dordrecht: Springer.
Gittlen, Laura & Stephen A. Marlett. 1985. Ñumí Mixtec syllable structure and morphology. *Work Papers of the Summer Institute of Linguistics*, University of North Dakota 29. 175-194.
Goldsmith, John. 1976. *Autosegmental phonology.* Cambridge, MA: MIT dissertation.
Goldsmith, John A. 1990. Autosegmental and metrical phonology. Oxford: Basil Blackwell.
Grimm, Nadine. 2022. Exponence and the functional load of grammatical tone in Gyeli. *Phonology* 39(3). 1-29.
Gussenhoven, Carlos. 2004. *The phonology of tone and intonation.* Cambridge: Cambridge University Press.
Hinton, Leanne. 1991. An accentual analysis of tone in Chalcatongo Mixtec. In James E. Redden (ed.). *Papers from the American Indian Languages Conference, University of California, Santa Cruz, July and August 1991*, 173-182. Occasional Papers in Linguistics 16. Carbondale: Southern Illinois University.
Hinton, Leanne, Gene Buckley, Marv Kramer & Michael Meacham. 1991. Preliminary analysis of Chalcatongo Mixtec tone. In James E. Redden (ed.), *Papers from the American Indian Languages Conferences*. 147-155.
Hollenbach, Barbara. 2003. The historical source of an irregular Mixtec tone-sandhi pattern. In M. R. Wise, T. Headland, & R. Brend (Eds.), *Language and life: essays in memory of Kenneth L. Pike,* 535–552. Dallas: SIL International.
Hollenbach, Barbara. 2015. Notes on tense, mood, and negation in Mixtec languages. Manuscript. *Summer Insititute of Linguistics (SIL) International*.
Hunter, Georgia & Eunice Pike. 1969. The phonology and tone sandhi of Molinos Mixtec. *Linguistics* 47. 24–40.
Hyman, Larry M. & Armindo Ngunga. 1994. On the non-universality of tonal association 'conventions': Evidence from Ciyao. *Phonology* 11(1). 25–68. https://doi.org/10.1017/S0952675700001834.
Hyman, Larry M. 2001. Privative tone in Bantu." In Shigeki Kaji (ed.), Cross-linguistic studies of tonal phenomena: Tonogenesis, Japanse accentology, and other topics, 237-257. Tokyo: Institute for the Study of Languages and Cultures.
Hyman, Larry M. 2006. Word-prosodic typology. *Phonology* 23(2). 225-257.
Hyman, Larry M. 2009. How (not) to do phonological typology: the case of pitch-accent. *Language Sciences* 31(2-3). 213–238.
Inkelas, Sharon. 1988. The theoretical status of morphologically conditioned phonology: A case study of dominance effects. In Geert Booij & Jaap Van Marle (eds.), *Yearbook of Morphology 1997,* 267-290. Dordrecht: Springer.
Inkelas, Sharon. 2014. *The interplay of morphology and phonology*. Oxford: Oxford University Press.
Josserand, Judy Kathryn. 1983. Mixtec Dialect History. PhD. dissertation, Tulane University.
Kaufman, Terrence. 1988. Otomanguean tense, aspect, mood, voice, and nominalization markers. Manuscript. University of Pittsburgh.
Kiparsky, Paul. 1973. Abstractness, opacity, and global rules. In Osamu Fujimura (ed.), *Three Dimensions of Linguistic Theory*, 57-86. TEC: Tokyo.
Kiparsky, Paul. 1984. A Compositional Approach to Vedic Word Accent. In S.D. Joshi (ed.), *Amrtadhara: Prof R. N. Dandeka Felicitation Volume*. Ajanta Publications, Jawahar Nagar Delhi.
Kiparsky, Paul & Morris Halle. 1977. Towards a Reconstruction of the Indo-European Accent. In L. Hyman (ed.), *Studies in Stress and Accent*. Department of Linguistics, University of Southern California, Los Angeles.

Kresge, Lisa 2007. Indigenous Oaxacan communities in California: An overview. *California Institute for Rural Studies* 1107.

León Vásquez, Octavio. 2017. *Sandhi tonal en el mixteco de Yucuquimi de Ocampo.* Masters thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS).

Longacre, Robert E. 1957. *Proto-Mixtecan.* Publication 5. Indiana University Research Center in Anthropology, Folklore and Linguistics and International Journal of American Linguistics 23(4).

Macaulay, Monica. 1987. Cliticization and the morphosyntax of Mixtec. *International Journal of American Linguistics* 53(2).119-135.

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

Macaulay, Monica & Joseph C. Salmons. 1995. The phonology of glottalization in Mixtec. *International Journal of American Linguistics*61(1). 38-61.

Mak, Cornelia 1953. A Comparison of two Mixtec Tonemic Systems. *International Journal of American Linguistics* 19(2). 85-110.

Mak, Cornelia. 1958. A Tonal system of a third Mixtec dialect. *Summer Institute of Linguistics.* 61-70.

McKendry, Inga. 2013. *Tonal association, prominence and prosodic structure in South-Eastern Nochixtlán Mixtec*. Doctoral dissertation, University of Edinburgh.

McKendry, Inga. 2018. Evidence for underlying mid tones in South-eastern Nochixtlán Mixtec. *Proceedings of Sixth International Symposium on Tonal Aspects of Languages*. 173-177.

McPherson, Laura & Matthew Dryer. 2021. The Tone system of Poko-Rawo (Skou). *Phonological Data and Analysis* 3(1). 1–32.

Mendoza Ruiz, Juana. 2016. Fonología segmental y patrones tonales del Tu'un Savi de Alcozauca de Guerrero. MA Thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social.

Myers, Scott. 1998. Surface underspecification of tone in Chichewa. *Phonology* 15(3). 367– 391.

Nissim, Gabriel M. 1981. *Le bamileke-ghomálá' (parler de Bandjoun, Cameroun): phonologie, morphologie nominale, comparaison avec des parlers voisins.* Paris: Société d'études linguistiques et anthropologiques de France.

North, Joanne & Shields, Jäna K. 1977. Silacayoapan Mixtec phonology. *Summer Institute of Linguistics and the University of Texas at Arlington* 54. 21-33.

Odden, David. 1999. Typological issues in tone and stress in Bantu. In Shegeki Kaji (ed.) , *Cross-linguistic studies of tonal phenomena: Tonogenesis, typology, and related topics*, 187-215. Tokyo: ILCAA.

van Otterloo, Karen. 2011. *The Kifuliiru language, volume 1: Phonology, tone, and morphological derivation*. Dallas, TX: SIL International.

Overholt, Edward. 1961. The Tonemic System of Guerrero Mixtec. In Benjamin Elson & Juan Comas (eds.), *A William Cameron Townsend en el vigésimoquinto aniversario del Instituto Lingüístico de Verano*, 597-626. Mexico: Instituto Lingüístico de Verano.

Palancar, Enrique L., Jonathan D. Amith & Rey Castillo García. 2016. Verbal inflection in Yoloxóchitl Mixtec. In Enrique L. Palancar & Jean-Léo Léonard (eds.), *Tone and Inflection*, 295–336. Berlin/Boston: Walter De Gruyter.

Pankratz, Leo & Eunice V. Pike. 1967. Phonology and morphotonemics of Ayutla Mixtec. *International Journal of American Linguistics*33(4). 287-299.

Paster, Mary & Rosemary Beam de Azcona. 2004a. A phonological sketch of the Yucunany Dialect of Mixtepec Mixtec. *Workshop on American Indigenous Languages (WAIL), UC Santa Barbara.* 61–75.

Paster, Mary & Rosemary Beam de Azcona. 2004b. Aspects of tone in the Yucunany dialect of Mixtepec Mixtec. Paper presented at the Conference on Otomanguean and Oaxacan Languages, University of California, Berkeley.

Pearce, Mary. 2006. The interaction between metrical structure and tone in Kera. *Phonology* 23(2). 259-286.

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

Pike, Eunice V. & Priscilla Small. 1974. Downstepping terrace tone in Coatzospan Mixtec. In Ruth Brend (ed.), *Advances in Tagmemics*, 105-134. Amsterdam: North-Holland Publishing Company.

Pike, Eunice V. and Thomas Ibach. 1978. The phonology of the Mixtepec dialect of Mixtec. Linguistic and Literary Studies in Honor of Archibald A. Hill, Volume 2: Descriptive Linguistics, ed. by Mohammed Jazayery, Edgar C. Polomé, and Werner Winter, 271-85. The Hague: Mouton.

Pike, Eunice V. & Kent Wistrand. 1974. Step-up terrace tone in Acatlán Mixtec. In Ruth Brend (ed.), *Advances in Tagmemics*, 83-104. Amsterdam: North-Holland Publishing Company.

Pike, Eunice V. & Joy Oram. 1976. Stress and tone in the phonology of Diuxi Mixtec. *Phonetica* 33(5). 321–333.

Pike, Kenneth l. 1944. Analysis of a Mixteco text. *International Journal of American Linguistics* 10(4). 113–138.

Pike, Kenneth L. 1947. Grammatical prerequisites to phonemic analysis. *Word* 3(3). 155-172.

Pike, Kenneth L. 1948. *Tone languages*. Ann Arbor, MI: University of Michigan Press.

Pulleyblank, Douglas. 1986. *Tone in lexical phonology.* Dordrecht: Reidel.

Rice, Keren. 1999. How phonetic is phonology? Evidence from tones in Athapaskan languages. Paper presented at the *Conference on Distinctive Feature Theory, ZAS Berlin*.

Rolle, Nicholas. To appear. A tonological rarity: Tone-driven epenthesis in Ghomala'. In Cormac Anderson, Natalia Kuznetsova & Shelece Easterday (eds.), *Rarities in phonetics and phonology*. Language Science Press.

Rolle, Nicholas. 2018. *Grammatical tone: Typology and theory*. Doctoral dissertation, University of California, Berkeley.

Rolle, Nicholas & John T. Merrill. 2022. Tone-driven epenthesis in Wamey. *Phonology 39*(1). 113-158.

Sande, Hannah. 2022. Is grammatical tone item-based or process-based? *Phonology* 39(3). 1-44.

Santos, Rosine Sachine. 1996. *Le mey: langue ouest-atlantique de Guinée.* Doctoral dissertation, University of the Sorbonne.

Smith, Jennifer L. 2001. Lexical category and phonological contrast. In *PETL 6: Proceedings of the Workshop on the Lexicon in Phonetics and Phonology,* 61-72. Edmonton: University of Alberta.

Swanton, Michael & Juana Mendoza Ruiz. 2021. Observaciones sobre la diacronía del tono en el Tu'un Savi (mixteco) de Alcozauca de Guerrero. In Francisco Arellanes & Lilián Guerrero (eds.), *Estudios lingüísticos y filológicos en lenguas indígenas mexicanas. Celebración por los 30 años de Seminario de Lenguas Indígenas*, 309-346. Instituto de Investigaciones Filológicas, UNAM.

Tranel, Bernard. 1995. Rules vs. constraints: a case study. Ms. University of California, Irvine. Paper presented at the conference on Current Trends in Phonology: Models and Methods, Royaumont, France, June 19-21, 1995.

Uchihara, Hiroto & Juana Mendoza Ruiz. 2022. Minimality, maximality and Perfect Prosodic Word in Alcozauca Mixtec. *Natural Language and Linguistic Theory* 40(2). 599-649.

Van Doren, Maxine, Claudia Duarte Borquez, Claudia Juárez Chávez, and Gabriela Caballero. To appear. Illustration of San Juan Piñas Mixtec. *Journal of the International Phonetic Association.*

Welmers, Wm. E. 1959. Tonemics, morphotonemics, and tonal morphemes. *General Linguistics* 4. 1-9.

Yip, Moira. 2002. *Tone*. Cambridge, UK: Cambridge University Press.

Yuan, Michelle. To appear. Two restrictions on pronominal clitics in San Juan Piñas Mixtec (Tò'ōn Ndá'ví). In: D. K. E. Reisinger (ed.), *Proceedings of WSCLA 26*, UBC Working Papers in Linguistics.

Zhang, Jie. 2000. Phonetic duration effects on contour tone distribution. In M. Hiratoni, A. Coetzee, N. Hall and J.-Y Kim (eds.), *Proceedings of the North East Linguistics Society 30* New Brunswick NJ: Rutgers University, Vol. 2. 775-85.

Zhang, Jie. 2001. *The Effects of Duration and Sonority on Contour Tone Distribution - Typological Survey and Formal Analysis*. PhD Dissertation, UCLA.

Zimmermann, Eva. 2017. Morphological H-tones in Macuiltianguis Zapotec (2017), In: Jessica Kantarovich, Tran Truong, and Orest Xherija (eds.), *Proceedings of the Chicago Linguistics Society (CLS)* 52, Chicago Linguistic Society, Chicago. 527-541.

Zylstra, Carol F. 1980. Phonology and Morphophonemics of the Mixtec of Alacatlazala, Guerrero. *Summer Institute of Linguistics,* Inc. 15-42.

Gabriela Caballero
Department of Linguistics
University of California San Diego
9500 Gilman Dr., #0108
La Jolla, CA 92093 USA
gcaballero@ucsd.edu

Claudia Duarte Borquez
Department of Linguistics
University of California San Diego
9500 Gilman Dr., #0108
La Jolla, CA 92093 USA
cduarteb@ucsd.edu

Claudia Juárez Chávez
Department of Linguistics
University of California San Diego
9500 Gilman Dr., #0108
La Jolla, CA 92093 USA
cjchavez@ucsd.edu

Michelle Yuan
Department of Linguistics
University of California, Los Angeles
Campbell Hall 3125
Los Angeles, CA 90095
yuanm@ucla.edu