Chapter 5

Constituency in Ayautla Mazatec

Shun Nakamoto
Universidad Nacional Autónoma de México

This study reports the result of 30 constituency diagnostics applied to Ayautla Mazatec, a Popolocan (Otomanguean) language from Oaxaca, Mexico, following the methods laid out by Tallman (2020, 2021). This language shows a high convergence rate of morphosyntactic diagnostics, while phonological diagnostics rarely converge.

1 Introduction

This study aims at providing a comprehensive list of phonological and morphosyntactic diagnostics which define syntagmatic domains in the Ayautla Mazatec sentence with a verbal predicate, following the methods laid out by Tallman (2020, 2021). §1.1 introduces the language background, and §1.2 lays out the methodology, before examining the planar structure in Ayautla Mazatec (§2) and the constituency diagnostics (§3). §4 concludes this study.

1.1 Language background

Ayautla Mazatec (ISO 639-3: vmy; Glottocode: ayau1235) is a Lowland variety of Mazatec spoken mainly in San Bartolomé Ayautla in the northernmost part of the state of Oaxaca, Mexico, by approximately 3,000 individuals of all ages. Mazatec, along with Ngiwa (Chocholtec-Popoloca) and moribund Ixcatec, belongs to the Popolocan branch of Otomanguean linguistic family (Fernández de Miranda 1951, Gudschinsky 1958, Hamp 1958, 1960). San Martín Duraznos Mixtec (Sandra Auderset & Ventayol-Boada 2024 [this volume]), Teotitlán del Valle Zapotec (Gutiérrez & Uchihara 2024 [this volume]) and Zenzontepec Chatino (Campbell

Shun Nakamoto. Constituency in Ayautla Mazatec. To appear in: Adam J.R. Tallman, Sandra Auderset & Hiroto Uchihara (eds.). Constituency and convergence in the Americas. Berlin: Language Science Press. [preliminary page numbering]

2024 [this volume]) are its distant relatives. Despite the genetic distance, Ayautla Mazatec shows striking similarities in its planar structures and, to some extent, the results of constituency diagnostics, with Teotitlán del Valle Zapotec and Zenzontepec Chatino. The data for this study come from my own fieldwork since 2017. Recorded textual examples from unpublished sources are indicated by a unique identifier. For the sake of illustration, however, many examples were elicited.

Phonologically, Ayautla Mazatec is a heavily tonal language—/1/ being the lowest tone and /4/ the highest—, with a boundary-sensitive phonological process of tone sandhi caused by floating /4/, transcribed as /(4)/ (Nakamoto 2020: 171–196; see also §3.12-3.14 on sandhi-related constituency diagnostics). Morphologically, inflectional exponence of person/number and aspect/mood involves extensive weak or strong suppletion of the verb roots and suppletive allomorphy of the prefixes derived from verb roots (Nakamoto 2020: 267–319). Morphosyntactically, Ayautla Mazatec is a strongly head-marking language (cf. Nichols 1986) with split intransitivity, and consistently shows the typological correlates of VO languages (cf. Dryer 2007). A root-level distinction between verbs, nouns, adjectives and positionals — mostly used as dependent components of compound verbs (cf. Cowan & Cowan 1947; Kalstrom Dolson et al. 1995: 343)—is well motivated (Nakamoto 2020: 27–63). As other Popolocan languages, Ayautla Mazatec has a relatively small number of verb roots, approximately 150 (Nakamoto 2020: 267–268; see also Pike 1948: 161; Pike 1967: 328; Krumholz et al. 1995). This paucity of verb roots is compensated by an extensive use of verb compounding and derivational prefixes that originated in verb roots.

1.2 Planar structure, constituency diagnostics and wordhood

In order to explicitly represent the linear nature of speech and to avoid *a priori* definitions of word, clitic or any other syntagmatic units of analysis, I adopt a radically flat representation of clause structure (planar structure) advocated by Tallman (2020, 2021, this volume).[1]

A planar structure is made up of **positions**, which may be a **slot** filled with an **element**—i.e. a morpheme, a compound stem, a phrase or a group of mor-

[1]For expositional reasons, however, I follow the segmentation and boundary signs used in Nakamoto (2020), where a plus sign (+) represents the compound boundary; a hyphen (-) the affix boundary; and an equal (=) the boundary of proclitics, second position clitics, enclitics and focus marker.

Spaces are written between orthographic words, which treat proclitics (with or without second position clitics) as separate words.

See §2 on proclitics, second position clitics and enclitics in Ayautla Mazatec.

phemes/stems/phrases in a paradigmatic relationship—, or a **zone**, where more than one element can occur without restrictions on their ordering.

Once the planar structure is established, language-specific **constituency tests** are mapped onto it around the **core** (root or compounded/derived stem). A constituency test is a phonological and/or morphosyntactic phenomenon which defines a syntagmatic domain, such as the domain of stress assignment, the domain of free occurrence, and so on. If a given test is ambiguous and delimits different spans according to interpretation, **test fracturing** is applied following Tallman (2021). For example, if the positive evidence and the negative evidence of a phenomenon define different domains, they are treated as two **constituency diagnostics**. Every constituency diagnostic then is well-defined, in the sense that it has a beginning position and an ending position.

According to this methodology, wordhood is not defined in accordance with some preestablished criteria (see Haspelmath 2011 for a critique on the universality of common wordhood criteria). Instead, wordhood can be understood as one of such domains on which two or more constituency diagnostics **converge**.

2 Ayautla Mazatec planar structures

In this section, first I present the planar structure of sentences with a verbal predicate (§2.1). Then I explore internally complex positions in the verbal planar structure, i.e. adverbs (§2.2) and noun complexes (§2.3), followed by a clarification about extra-clausal operations of topicalization and afterthoughts (§2.4).

Note that subordinate clauses, including relative clauses, complement clauses and adverbial clauses, will be treated in §3.6 as repeated subspans of the planar structure.

2.1 Verbal predicate

Table 1 shows the planar structure of the Ayautla Mazatec sentence with a verbal predicate, which is built around the verb root (position 19). Non-verbal predicates are outside the scope of this study (but see Nakamoto 2020: 55–57, 252).

Some constructions involve two positions. The existence of a second formative is indicated by three dots (…) after the first formative. The position for the latter is indicated in the template as "XX 2", such as "negation + ant/post 2" of the position 13.

Morphemes that occur in certain positions are arbitrarily called "clitics", for expositional reasons. Each class of clitics is defined as a group of bound morphemes with distributional properties in common. Morphemes in positions 6-8

Table 1: Verbal planar structure of Ayautla Mazatec

Pos.	Type	Elements	Forms
(1)	slot	connector	*ha¹, ʔba¹, ʔa²sa¹, tu¹ⁿka²*, etc.
(2)	zone	sentence adv., topicalization	
(3)	slot	polar Q, 'even', 'only'	*tu¹, su²ba⁴, na²se³⁽⁴⁾, ʔa²*, etc.
(4)	zone	noun complex, adverb	
(5)	slot	focus	*=³⁽⁴⁾*
(6)	slot	anterior/posterior	*he²⁽⁴⁾=, khe²⁴=(...=⁴hĩ⁴)*
(7)	slot	'almost'	*me²he⁴=*
(8)	slot	negation + ant/post	*bi⁴=, bi⁴=...(ⁿte¹=), ni¹=*
(9)	slot	temporal sequence	*=hba⁴, =hba⁴ni²³, =hba⁴ra², =ra²*
(10)	slot	modality, evidentiality	*=ʃu³⁽⁴⁾, =hĩ⁴, =ru¹*
(11)	slot	'a little'	*=¹tɕi⁴*
(12)	slot	past habitual, 'always'	*=ʔĩ³⁽⁴⁾, =ⁿtshɛ⁴*
(13)	slot	negation + ant/post 2	*ⁿte¹=*
(14)	slot	adverb, pronoun (P/R)	*ʔã², hi²³, kʷi², hĩ¹, ɲã³⁽⁴⁾, hũ⁴*, etc.
(15)	slot	progressive	*ti²⁽⁴⁾-, te²-*, etc.
(16)	slot	aspect/mode	*b-, t-, kʷ¹-*, etc.
(17)	slot	associated motion	*hi²-, ⁿte²⁽⁴⁾-, e¹-, hɛ²ʔɛ²-*, etc.
(18)	slot	causative, inchoative	*tsi²⁽⁴⁾-, a²-*, etc.
(19)	slot	verb root(s)	
(20)	slot	comitative	*-ko¹³*
(21)	slot	focus	*=³⁽⁴⁾*
(22)	slot	temporal sequence	*=hba⁴, =hba⁴ni²³, =hba⁴ra², =ra²*
(23)	slot	modality, evidentiality	*=ʃu³⁽⁴⁾, =hĩ⁴, =ru¹*
(24)	slot	'a little'	*=¹tɕi⁴*
(25)	slot	past habitual, 'always'	*=ʔĩ³⁽⁴⁾, =ⁿtshɛ⁴*
(26)	zone	adverbial/quantifier clitics	*=hⁿku²³, =je²he², =ⁿka²ɲi³⁽⁴⁾*, etc.
(27)	slot	modal clitics	*=⁴hĩ⁴, =ɲi²³, =ni⁴ɲi²³*
(28)	slot	pronominal clitics	*=a² ~ =a¹, =i, =i¹, =a³⁽⁴⁾, =u⁴*, etc.
(29)	zone	noun complex, adverb	
(30)	slot	attitudinal particles	*ja²ʔa², je²he², tsa²kɛ²⁴, ⁿte¹*, etc.
(31)	zone	afterthoughts, vocative words	

and 13 are called proclitics. A group of morphemes occur in position 9–12 if any of 4 or 6-8 is occupied; if not, they occupy 22–25; these elements can be considered as Wackernagel or second position clitics (cf. Wackernagel 2020). Focus marker (position 5 or 21) is also segmented as a clitic. Morphemes in positions 26-28 are called enclitics. Inside the positions for clitics around the verb root, i.e. positions 15-18 and 20, are called affixes. Together with the clitics, "connectors" (position 1) and "attitudinal particles" (position 30) are bound morphemes too. However, given that they do not interact distributionally with the rest of the sentence, this study will pay little attention to them.

2.2 Adverbs

The adverb has the following template (Table 2). If there are "adverb 1" and "adverb 2" at the same time, they constitute only one free form. Positions 2-4, 6-8 or 11-13 are occupied if there are no other elements in the positions 4–8 of the verbal template. Note that temporal, modality/evidentiality and 'a little' clitics have a scope over the whole verbal predicate and not on the adverb.

Table 2: Planar structure for adverbs

Pos.	Type	Elements	Forms
(1)	slot	negation	*bi⁴=, ni¹=*
(2)	slot	temporal sequence	*=hba⁴, =hba⁴ni²³, =hba⁴ra², =ra²*
(3)	slot	modality, evidentiality	*=ʃu³⁽⁴⁾, =hĩ⁴, =ru¹*
(4)	slot	'a little'	*=¹tɕi⁴*
(5)	slot	adverb	*ʔba¹…(=ⁿte¹), ʔba¹…(kʷã¹³)*, etc.
(6)	slot	temporal sequence	*=hba⁴, =hba⁴ni²³, =hba⁴ra², =ra²*
(7)	slot	modality, evidentiality	*=ʃu³⁽⁴⁾, =hĩ⁴, =ru¹*
(8)	slot	'a little'	*=¹tɕi⁴*
(9)	slot	adverb 2	*=ⁿte¹, kʷã¹³*, etc.
(10)	slot	adverb type marker	*=tsa², =ni¹, =¹*
(11)	slot	temporal sequence	*=hba⁴, =hba⁴ni²³, =hba⁴ra², =ra²*
(12)	slot	modality, evidentiality	*=ʃu³⁽⁴⁾, =hĩ⁴, =ru¹*
(13)	slot	'a little'	*=¹tɕi⁴*

2.3 Noun complex

The noun complex has the following template (Table 3). Positions 2-4, 7-9 or 16-18 of this template are occupied if there are no other elements in the positions 4–8 of the verbal template. Absolute state marker $=^{1}$ occurs at the end of a noun complex (i) without a possessor, (ii) without a demonstrative, (iii) not in vocative function, and (iv) without a floating tone /4/ immediately before it (Nakamoto 2020: 241–250).

Table 3: Planar structure for noun complexes

Pos.	Type	Elements	Forms
(1)	slot	'more', 'also'	$^{n}ki^{2}sa^{1}$, ko^{13}
(2)	slot	temporal sequence	$=hba^{4}$, $=hba^{4}ni^{23}$, $=hba^{4}ra^{2}$, $=ra^{2}$
(3)	slot	modality, evidentiality	$=ʃu^{3(4)}$, $=hĩ^{4}$, $=ru^{1}$
(4)	slot	'a little'	$=^{1}tɕi^{4}$
(5)	slot	additive, 'entire'	$^{n}ki^{2}$-, $^{n}ka^{2}$-
(6)	slot	numeral, quantifier	$h^{n}ku^{23}$, $^{n}khĩ^{4}$, etc.
(7)	slot	temporal sequence	$=hba^{4}$, $=hba^{4}ni^{23}$, $=hba^{4}ra^{2}$, $=ra^{2}$
(8)	slot	modality, evidentiality	$=ʃu^{3(4)}$, $=hĩ^{4}$, $=ru^{1}$
(9)	slot	'a little'	$=^{1}tɕi^{4}$
(10)	slot	noun root(s)	
(11)	slot	adjective	
(12)	slot	possesssor	$=na^{1}$, $=^{1}ri^{2}$, $=re^{1}$, $=ni^{1}$, $=na^{3(4)}$, $=^{1}nu^{4}$
(13)	slot	demonstrative	$=bi^{1}$, $=bju^{1}$
(14)	slot	relative clause	
(15)	slot	absolute state marker	$=^{1}$
(16)	slot	temporal sequence	$=hba^{4}$, $=hba^{4}ni^{23}$, $=hba^{4}ra^{2}$, $=ra^{2}$
(17)	slot	modality, evidentiality	$=ʃu^{3(4)}$, $=hĩ^{4}$, $=ru^{1}$
(18)	slot	'a little'	$=^{1}tɕi^{4}$

2.4 On extra-clausal operations

Within the verbal planar structure, I do not break down topicalization or left-dislocation (position 2) and afterthoughts or right-dislocation (position 31) in the planar structure, because they show no distributional interactions with the rest of the sentence. Specifically, the modality/evidentiality clitic used in the main clause appears duplicated in a topicalized constituent—as illustrated by the first instance

of reported information *=ʃu³⁽⁴⁾* in (1)—or an afterthought, without altering the definition of the second position of the verbal predicate clause which begins with *su²ba⁴*.[2]

(1) *ˈhⁿku²³ʃu³ ja²ni⁴ˈtɕa²¹ su²ˈba⁴ ˈɕi¹re¹³ʃu³ ˈkʷã²⁴*

hⁿku²³ =ʃu³⁽⁴⁾ ja³⁽⁴⁾+ni²tɕa² =¹ su²ba⁴ ɕi¹ =re¹ =³⁽⁴⁾

v: 2[ ] 3 4[ ] 5

n: 6 8 10 15 - 10 12 -

TOP[one =REP tree+ocote =ABS.ST] only [piece =POSS3] =FOC

=ʃu³⁽⁴⁾ kʷ- ã²⁴

10 16 19

- - -

=REP PFV- PFV:become

'an ocote tree reportedly fell apart into mere pieces.' (Sánchez Díaz & Nakamoto 2020: 139, English by SN)

3 Constituency diagnostics

In this section I will describe the following 14 tests and 30 constituency diagnostics I have so far identified for Ayautla Mazatec. Tests 1-6 are treated generally as morphosyntactic tests, while tests 8-14 are phonological tests. Pause (test 7) is sometimes considered a morphosyntactic test (cf. Haspelmath 2011) and sometimes a phonological test (cf. Dixon & Aikhenvald 2002, Gerdts & Werle 2014). The syntagmatic spans identified by the diagnostic appear in parentheses. If test fracturing applies, the smaller span appears first.

1. Free occurrence, or minimum free form (19-19, 15-28)
2. Deviation from biuniqueness (15-19, 15-28)
3. Ciscategorial selection (15-19, 15-28)
4. Non-interruptability (15-28, 6-28)
5. Fixed order or non-permutability (15-20, 13-21)

[2] In interlinearized examples, the first line corresponds to the surface (or post-sandhi) form, the second line to the underlying (or pre-sandhi) form, the third to the verbal planar structure and the fourth and the fifth, if necessary, to the nominal and adverbial planar structure. The lines for planar structures are indicated with 'v:', 'n:' and 'adv:', respectively. The last two lines correspond to morpheme-by-morpheme gloss and free translation with a unique identifier for textual examples.

6. Subspan repetition (15-19, 15-19, 3-20, 6-28, 3-29, 2-29)
7. Pauses and fillers (15-28)
8. Stress assignment (19-20, 15-20)
9. *ɛ.j constraint (19-19, 13-25)
10. *3.(2)4 constraint (19-19, 15-21)
11. Syllable-internal segmental interactions (16-19, 16-28)
12. Disyllabic sandhi-blocking tone sequences (15-19, 15-21)
13. Obligatory sandhi (15-28)
14. Possible sandhi (15-28, 2-31)

In the following subsections, I will describe each one of these constituency tests.

3.1 Free occurrence (19-19, 15-28)

Being a minimum free form, that is, possibly constituting an utterance (and not two), has been an oft-cited criterion of wordhood (cf. Haspelmath 2011: 39). The minimal and the maximal extension of a free form in Ayautla Mazatec verbal predicates delimit different spans, thus providing two constituency diagnostics.

On the one hand, the MINIMAL MINIMUM FREE FORM consists only of the verb root (position 19) if the verb is a non-derived stative verb in third person form, such as (2).

(2) *ja²ˀʔa²³*
ja²ʔa²³
v: 19
carry:3
'he has, holds, carries.'

On the other hand, MAXIMAL MINIMUM FREE FORM includes the maximal range of elements which can occur as a single free-standing form, which spans from 15 to 28, as illustrated in (3).

(3) *te⁴ˈhbja²³¹*
te²- *b-* *hi²³* *=a¹*
v: 15 16 19 28
PROG:1- HAB- go:1 =1SG
'I am going.'

Before position 15 for progressive prefix, position 14 is occupied by adverbs and independent pronouns which can be free-standing forms. After position 28 for pronominal enclitics, position 29 for noun complexes and/or adverb also consists of one or more free standing forms. Therefore, the maximal definition of this test is 15-28.

3.2 Deviation from biuniqueness (15-19, 15-28)

Deviation from biuniqueness (or one-to-one correspondence between form and function) has often been referred to as a characteristic of words but not phrases (cf. Haspelmath 2011: 54). Ayautla Mazatec shows many cases of non-automatic allomorphy, i.e. many-to-one correspondences between form and function (cf. Pike 1948: 132).

Specifically, progressive (position 15; Nakamoto 2020: 50–52), aspect/mode (position 16; Nakamoto 2020: 39–50, 288-306), associated motion (position 17; Nakamoto 2020: 52–53), voice (position 18; Nakamoto 2020: 29–30), verb root (position 19; Nakamoto 2020: 270–288) and bound pronouns (position 28; Nakamoto 2020: 236–239) show allomorphy not conditioned by phonology. Some allomorphies in these positions are illustrated in the following pair of examples. In (4), the progressive *ti²⁽⁴⁾-* ~ *te²-*, the andative *hi⁴-* ~ *ʔi²-*, the causative *tsi²⁽⁴⁾k-* ~ *ni²k-* and the verb root *i²se³⁽⁴⁾* ~ *i²se⁴* show different allomorphs conditioned by the agent person, in addition to the habitual *b-* ~ *m-* conditioned phonologically as well as lexically.[3]

(4) a. *ti²hbi⁴tsi²ki²se²ˈthẽ⁴¹*
ti²⁽⁴⁾- *b-* *hi⁴-* *tsi²⁽⁴⁾k-* *i²se³⁽⁴⁾+thẽ¹*
v: 15 16 17 18 19
PROG:3- HAB- ANDT:3- CAUS:3- rise:3
'he is going there to raise (something)'

[3]The allomorphs listed above are not at all exhaustive. See corresponding sections in Nakamoto (2020) cited above.

b. *te²ʔmi²ni²ki²se⁴'thẽ¹³*

te²-	*m-*	*ʔi²-*	*ni²k-*	*i²se⁴+thẽ¹*	*=i³*
v: 15	16	17	18	19	28
PROG:2-	HAB-	ANDT:2-	CAUS:2-	rise:2	=2SG

'you are going there to raise (something)'

Given the discontinuity of positions which show deviation from biuniqueness, this test can be fractured into two constituency diagnostics. The minimal interpretation of this test includes the positions 15-19, where all positions show deviation from biuniqueness (**minimal deviation from biuniqueness**). The maximal interpretation of this test includes the positions 15-28, which covers all positions outside which deviation from biuniqueness is known not to be observed (**maximal deviation from biuniqueness**).

3.3 Ciscategorial selection (15-19, 15-28)

Whether a given morpheme occurs exclusively with certain lexical categories or not has been a major criterion for distinguishing clitics from affixes (cf. Haspelmath 2011: 45). In Ayautla Mazatec verbal predicates, progressive (position 15), aspect/mode (position 16), associated motion (position 17), voice (position 18), verb roots (position 19)—root-level distinction of lexical categories is clear in this language—, adverbial/quantifier clitic (position 26) and bound pronouns (position 28) are limited to verbal predicates.[4] For example, in (5), independent pronouns *kʷi²* 'PRONOM3' and *ʔã²* 'PRONOM1SG' as well as comitative *ko¹³* are used in non-verbal predicates.

(5) *ma²'sẽ² ta¹² bi⁴ ⁿte¹ **ko¹³ 'kʷi²** he²⁽⁴⁾ tu¹ 'ʔã²*
ma²sẽ² ta¹² bi⁴= ⁿte¹ ko¹³= kʷi² he²= tu¹ ʔã²
half but NEG= anymore with= PRONOM3 already= only PRONOM1SG
'some (lit. half) but not with them anymore, now it's only me'
(180624-002 08:56)

As with the previous test, ciscategorial selection can be interpreted in different manners. On the one hand, all elements in positions 15-19 occur exclusively in verbal predicates (**minimal ciscategorial selection**). On the other hand, no position outside 15-28 shows word-class selectivity (**maximal ciscategorial selection**).

[4] Although a similar set of bound pronouns is used to indicate the possessor of some bodypart and kinship terms, it has unpredictable differences between the one used for predicates the one used for possession. I excluded comitative (position 20) from this list, because it can be used as preposition in noun complexes.

3.4 Non-interruptability (15-28, 6-28)

Another common definition of word is that of the uninterruptible string of morphemes (cf. Haspelmath 2011: 44). However, following previous critiques (cf. Haspelmath 2011, Tallman 2021), I provide two constituency diagnostics, namely, NON-INTERRUPTIBLE BY A SINGLE FREE FORM and NON-INTERRUPTIBLE BY COMPLEX FREE FORM(S) or MORE THAN ONE FREE FORM, which give different results in Ayautla Mazatec.

The maximum span of non-interruptible elements by a single free form begins at the progressive (position 15). Immediately before it (position 14), a non-focused adverb (6a) or an emphatic pronoun in P/R function (6b)—both of which are free forms—can occur.

(6) a. *bi⁴ʃu³ ʔba¹ tsɛ̃²ˈʔɛ̃³ kʷe¹ˈʔe⁴ri²*

	bi⁴=	*=ʃu³⁽⁴⁾*	*ʔba¹*	*tsɛ̃²ʔɛ̃³⁽⁴⁾*	*kʷ-*	*e¹ʔe⁴*	*=ri²*
v:	8	10	14[	]	16-	19	=28
adv:	-	-	5	9	-	-	-
	NEG=	=REP	[so	do:3]	POT-	POT:beat	=3/2SG

'he won't beat you like that.'

b. *bi⁴ʃu³ ˈhi²³ kʷe¹ˈʔe⁴ri²*

	bi⁴=	*=ʃu³⁽⁴⁾*	*hi²³*	*kʷ-*	*e¹ʔe⁴*	*=ri²*
v:	8	10	14	16-	19	=28
	NEG=	=REP	PRONOM2SG	POT-	POT:beat	=3/2SG

'he won't beat you.'

If position 14 is filled with two free forms, or complex free form, like (7a, 7b), the result is ungrammatical.

(7) a. *bi⁴ʃu³ ʔba¹ tsɛ̃²ˈʔɛ̃³ ˈhi²³ kʷe¹ˈʔe⁴ri²

	bi⁴=	*=ʃu³⁽⁴⁾*	*ʔba¹*	*tsɛ̃²ʔɛ̃³⁽⁴⁾*	*hi²³*	*kʷ-*	*e¹ʔe⁴*	*=ri²*
v:	8	10	14[	-]	14	16	19	28
adv:	-	-	5	9	-	-	-	-
	NEG=	=REP	[so	do:3]	PRONOM2SG	POT-	POT:beat	=3/2SG

intended: 'he won't beat you like that.'

b. *bi⁴ʃu³ ˈhi²³ ʔba¹ tsɛ̃²ˈʔɛ̃³ kʷe¹ˈʔe⁴ri²

	bi⁴=	*=ʃu³⁽⁴⁾*	*hi²³*	*ʔba¹*	*tsɛ̃²ʔɛ̃³⁽⁴⁾*	*kʷ-*	*e¹ʔe⁴*	*=ri²*
v:	8	10	14	14[	-]	16	19	28
adv:	-	-	-	5	9	-	-	-
	NEG=	=REP	PRONOM2SG	[so	do:3]	POT-	POT:beat	=3/2SG

intended: 'he won't beat you like that.'

On the other hand, the maximum span non-interruptible by more than one free form covers positions 6-28, delimited by two zones which may have noun complexes and adverbs (positions 4, 29), possibly filled with complex free forms.

In sum, non-interruptability by a free form spans from position 15 to position 28, while non-interruptability by more than one free forms covers positions 6-28.

3.5 Fixed order or non-permutability (15-20, 13-21)

Fixed order or non-permutability of morphemes has typically been considered characteristic within a word but not a phrase (cf. Dixon & Aikhenvald 2002: 19–20). However, the ambiguity of this test is notorious, since strict ordering of syntactic elements and variable ordering of affixes have also been reported. In this study, following the critique by Tallman (2021: §5.4), I divide the non-permutability test in two versions. One is STRICT NON-PERMUTABILITY, which entails a span of positions whose elements always occur in a fixed order. The other is NON-PERMUTABILITY WITHOUT SCOPAL DIFFERENCE, which, in addition to the previous one, includes positions with variably ordered elements, where this variable ordering corresponds to a difference in scope.

Strict non-permutability defines positions 15-20 as its span, since morphemes in these positions 15-20 in this order. Just outside progressive (position 15), there is a position for adverbs and independent pronouns (position 14). Adverbs in this position have scope over the predicate, while adverbs in position 2 have scope over the whole sentence. For that reason, some adverbs cannot occur in one of the two positions. For example, *ʔba¹ tsɛ̃²ʔɛ̃³* 'that way' can occur in position 14 (8a) but cannot in position 2 (8b).

(8) a. ***bi⁴ ʔba¹ tsɛ̃²ˈʔɛ̃³*** *e²ˈhɲu⁴na¹*

bi⁴=	*ʔba¹*	*tsɛ̃²ʔɛ̃³⁽⁴⁾*	*e²hɲu⁴*	*=na¹*
v: 8	14		16:19	28
NEG=	that.way		PFV:deceive:3	=3/1SG

'he didn't deceived me that way.'

b. ***ʔba¹ tsɛ̃²ˈʔɛ̃³*** bi⁴ e²ˈhɲu⁴na¹

ʔba¹	*tsɛ̃²ʔɛ̃³⁽⁴⁾*	*bi⁴=*	*e²hɲu⁴*	*=na¹*
v: 2		8	16:19	28
that.way		NEG=	PFV:deceive:3	=3/1SG

intended: 'that way he didn't deceived me.'

Just before the preverbal adverb or independent pronoun (position 14) is for the second part of the bipartite proclitic *bi⁴…ⁿte¹* 'not…anymore' (position 13),

which is exclusive to this position.[5] Therefore, 13 is included in the less strict definition of non-permutability.

After the comitative (position 20), the focus marker (position 21) can also occur in another place (position 5). This difference too corresponds to different focalized constituents. Therefore, 21 is included in the broader interpretation of non-permutability.

Outside these positions (13-21) are Wackernagel clitics (positions 9-12 and 22-25), which occur in different positions within the clause, but do not involve differences in scope. Therefore, all these positions involve permutable elements.

In sum, strict non-permutability defines 15-20 as its span, while non-permutability without scopal difference defines 13-21.

3.6 Subspan repetition (15-19, 15-19, 3-20, 6-28, 3-29, 2-29)

Some constructions specify a span of positions which can be repeated and thus can be employed as constituency tests of SUBSPAN REPETITION (cf. Tallman 2021). So far I have identified total reduplication, verbal parallelism (both at §3.6.1), grammatical nominalization with absolute state marker *=¹* (§3.6.2), and coordination (§3.6.3) as distinct subspan repetition constructions.

3.6.1 Total reduplication (15-19) and verbal parallelism (15-19)

The smallest subspan repetition constructions in Ayautla Mazatec are TOTAL REDUPLICATION and VERBAL PARALLELISM, which specify positions 15-19.

Total reduplication in Ayautla Mazatec, illustrated in (9) below, repeats a subspan of the verbal predicate, regardless of phonological conditions such as the syllable structure and the number of syllables. This process indicates the exhaustivity of the action expressed by the verb and therefore it is limited to dynamic verbs. In (9a), position 14 is excluded from the repeated subspan. In (9b), position 20 is excluded from the repeated subspan.

(9) a. *he² ˈkʷi² bo²ʔo²ˈja⁴ bo²ʔo²ˈja⁴ʔĩ³*

	he²⁽⁴⁾=	*kʷi²*	*b-*	*o²ʔo²+ja⁴*	*b-*
v: 6	14	$_{\text{redup1}}$[16	19]		$_{\text{redup2}}$[16
already=	PRONOM3	$_{\text{redup1}}$[HAB-	hit+POS:inside]		$_{\text{redup2}}$[HAB-

[5] Probably the second morpheme of this construction *ⁿte¹* comes from the same morpheme used as a sentence-final attitudinal particle *ⁿte¹* 'thus'. However, I consider this construction undecomposable, therefore I do not consider *ⁿte¹* as a permutable morpheme.

o²ʔo²+ja⁴ *=ʔĩ³⁽⁴⁾*
19] 25
hit+POS:inside] =PST.HAB
'he already used to beat and beat him.'

b. *pa²³'la¹ te²khe²'ʔⁿki³ te²khe²ʔⁿki²'ko⁴ʔĩ³ɲa³²*

pa²³la¹	*te²-*		*khe²+ʔⁿki³*	*te²-*	*khe²+ʔⁿki³⁽⁴⁾*
v: 4	redup1[15		16:19]	redup2[15	16:19]
spade	redup1[PROG-		HAB:pull:1+dig]	redup2[PROG-	HAB:pull:1+dig]

-ko¹³ *=ʔĩ³⁽⁴⁾* *=ɲi³⁽⁴⁾* *=a²*
20 25 27 28
-COM =PST.HAB =ASR =1SG
'I used to be digging and digging with a spade.'

Verbal parallelism construction also repeats from progressive (position 15) to verb root (position 19). However, unlike total reduplication, each repeated subspan has a different positional root, which is part of the position 19 for verb roots. This construction expresses the distributivity of an action, therefore is only available for dynamic verbs. In (10), subspan 15-19 is repeated each followed by *+tsha³⁽⁴⁾* 'sideways' and *+niʔɲa²* 'quadrupedal'. Note that combinations of the positional roots show considerable flexibility, reflecting each speaker's expressivity.

(10) *'hⁿku²³ ku¹ⁿtu¹tɕu¹'tsĩ³ he² ti²thu⁴'tsha³ ti²thu⁴niʔɲa²'ko¹³*

hⁿku²³	*ku¹ⁿtu¹tɕu¹tsĩ³*	*he²⁽⁴⁾=*	*ti²⁽⁴⁾-*	
v: 4[	]	6	vpar1[15	
n: 6	10	-	-	
one	bottle	already=	vpar1[PROG-	

thu⁴+tsha³⁽⁴⁾ *ti²⁽⁴⁾-*
16:19] vpar2[15
- -
HAB:come.out+POS:sideways] vpar2[PROG-
thu⁴+niʔɲa² *-ko¹³*
16:19] 20
- -
HAB:come.out+POS:quadrupedal] -COM

'he is already staggering with a bottle.'

3.6.2 Nominalization (3-20, 3-29)

Grammatical NOMINALIZATION in Ayautla Mazatec[6] targets predicates and derives noun- or adverb-like constituents which mean events, participants or circumstantial situations. Syntactically, nominalization may function as arguments or adjuncts in positions 4 or 29, in addition as the optative form in a main clause by insubordination, which is in complementary distribution to imperative (cf. Nakamoto 2020: 248, 47-50).

Within the planar structure, nominalization is indicated at two positions: (i) the complementizer/adverbial subordinator *ⁿka² ~ :²* or the relativizer *ʨi² ~ :²* at the beginning, and (ii) absolute state marker *=¹* (cf. §2.3) at the end.[7] In the example (11) below, nominalization *ⁿka² tu¹ tʨa²'ⁿtu⁴¹ ki²tsi²ka²'ʔbi³ 'tõ²⁴¹* 'that only Antonio distributed money' begins, except for the nominalizer (subordinator here) itself, at position 3 for focus introducer *tu¹* 'only'.

(11) *tsa²'be²⁴ʃu³ ⁿtsʔja³² ⁿka² **tu¹** tʨa²'ⁿtu⁴¹ ki²tsi²ka²'ʔbi³ 'tõ²⁴¹*

	tsa²-	*be²⁴*	*=ʃu³⁽⁴⁾*	*ⁿtsʔe³⁽⁴⁾*		*=a²*	*ⁿka²*	*tu¹*	*tʨa³⁽⁴⁾+ⁿtu⁴*	*=¹*
v:	16	19	23	29[-		-]	29[1	3	4[-	-]
n:	-	-	-	10		12	-	-	10	15
	PFV-	see:3	REP	[brother:SAP		=2SG]	SUB	only	[Antonio	=ABS.ST]
	ki²-	*tsi²k-*	*a²ʔbi³⁽⁴⁾*		*tõ²⁴*	*=¹*				
	16	18	19		29[-	-]]				
	-	-	-		10	15				
	PFV-	CAUS-	be.distributed		[money	=ABS.ST]				

'my brother saw that only Antonio distributed money.'

Due to the distributional ambiguity of absolute state marker, however, nominalization test is fractured into minimal and maximal interpretations, corresponding to positions 3-20 and 3-29, respectively.

MINIMAL NOMINALIZATION ends at the final position where the absolute state marker is unambiguously observed at the clause (and not the noun complex) level. For example, in (12), the nominalization *ⁿka² tõ²⁴¹ʃu² ki⁴ski²ne²* 'that it ate money' ends with a verb root (position 19) followed by an absolute state marker *=¹*. Similarly, in (13), the nominalization *ⁿka² ka²hbi²ko¹³* 'when he took it' ends

[6] Abstract noun formation by *kʷha¹-* also targets some verb forms in habitual (or neutral) aspect with corresponding segmental prefix. However, given its limited productivity, I do not discuss here the abstract noun formation.

[7] Nominalization with *ⁿka²* and absolute state *=¹* is not exclusive to verbs; numerals can be nominalized too (Nakamoto 2020: 329–330).

with a comitative *-ko*^{13} (position 20) followed by an absolute state marker *=*^{1}. The syllable with the absolute state marker in question is emphasized in boldface.

(12) *ta*^{12} *ˈkʷi*^{2}*ʃu*^{2} *ˈkʷha*^{41} *ⁿka*^{2} *ˈtõ*^{241}*ʃu*^{2} *ki*^{4}*skiˀˈ**ne**^{21}*

	ta^{12}	*kʷi*^{2}	*=ʃu*$^{3(4)}$	*kʷha*^{1}	*ⁿka*^{2}	*tõ*^{24}	*=*^{1}	*=ʃu*$^{3(4)}$	*ki*^{2}*s-*
v:	1	-	-	-	-[1	4[-	-]	10	16
n:	-	10	-	10	10(NMLZ)[-	10	15	-	-
	but	PRONOM3	=REP	matter	[SUB	[money	=ABS.ST]	=REP	PFV-

ki^{2}*ne*^{2} **=**^{1}
19] -
-] 15
eat:3] **=ABS.ST**

'but the matter is that it [the donkey] ate money.' (180816-002 00:55)

(13) *ha*^{1} *kaˀˈʔbja*^{231} *ⁿka*^{2} *ka*^{2}*hbiˀˈ**ko**^{131}*

	ha^{1}	*ka*^{2}*-*	*ʔbe*^{23}	*=a*^{1}	*ⁿka*^{2}	*ka*^{2}*b-*	*hi*^{2}	*-ko*^{13}	*=*^{1}
v:	1	16	19	28	29[1	16	19	20	-]
n:	-	-	-	-	10(NMLZ)[-	-	-	-]	15
	well	PST-	see:1	=1SG	[[SUB	PST-	go:3	-COM]	**=ABS.ST**]

'I saw (him) when he took it.' (Nakamoto 2020: 249)

In contrast, absolute state marker *=*^{1} does not occur in the subsequent positions until the post-verbal noun complex (position 29). On the one hand, focus marker *=*$^{3(4)}$ (position 21) has a floating tone /4/, which blocks the occurrence of absolute state marker (Nakamoto 2020: 248–250). On the other hand, absolute state marker *=*^{1} does not cooccur with second position clitics and enclitics (positions 22-28), as illustrated by (14). The syllable on which the absolute state marker would occur is indicated in boldface.

(14) *niˀˈʃthĩ*^{23} *ⁿka*^{2} *kʷhe*^{1}*ⁿti*^{2}*ˈ**bɛ**^{4}*

	niˀʃthĩ^{23}	*ⁿka*^{2}	*kʷ-*	*he*^{1}*ⁿti*^{2}*ba*^{4}	*=i*	
v:	-	NMLZ[1	16	19	28]	
n:	10	14[-	-	-	-]	
	day	[SUB	POT-	come	=2SG]	

'the day you will come.' (Nakamoto 2020: 245)

This cooccurrence restriction between the absolute state marker *=*^{1} and the clitics (positions 22-27) plausibly has a historical explanation. Given the broader distribution of the absolute state marker *=*^{1} in nouns, it is safe to attribute its origin to the nominal morphosyntax. Within noun complexes, however, this morpheme

does not occur if the noun is possessed, while the possessor is indicated by the nearly identical dependent pronouns used in verbal predicates. I suggest that this parallelism between the nominal template and the verbal template plays a role in blocking the absolute state marker after clitics in the verbal planar structure.

Maximal nominalization includes the post-verbal noun phrase (position 29), which is the last position where the absolute state marker =¹ is found. However, it is indeterminate as to whether the absolute state marker in this position is due to the nominalization, the noun phrase, or both. For example, in (15), the end of the nominalization *ⁿka² he² ⁿti²ba⁴ ɕi²hⁿku²³¹* 'that the other already came back' coincides with the end of the noun phrase *ɕi²hⁿku²³¹* 'the other' within it. Note that the final absolute state marker =¹ is inside the inner bracket if it occurs at the noun phrase level, and outside the inner bracket if it occurs at the clause level.

(15) *ʔba¹ ka²ʔⁿta² bi⁴ʃu³ tsa²ʾbe²⁴ ⁿka² he² ⁿti²ʾba⁴ ɕi²ʾ**hⁿku²³¹***
ʔba¹ ka²ʔⁿta² bi⁴= *=ʃu³⁽⁴⁾* *tsa²-* *be²⁴* *ⁿka²* *he²⁽⁴⁾=* *ⁿti²ba⁴*
v: 1 4 8 10 16 19 29[1 6 16:19
n: - - - - - - 10[- - -
and even NEG= =REP PFV- know:3 [SUB already= PFV:come:3
ɕi²hⁿku²³ *=¹*
29[- -]]
10 15]
[the.other =ABS.ST]]
'and he didn't even notice when the other already came back.'
(Sánchez Díaz & Nakamoto 2020: 132, English by SN)

Absolute state marker =¹ does not occur after the attitudinal particles (position 30), such as *ja²ʔa²* 'well' in example (16). I suggest that this is because the attitudinal particle (position 30) is found outside the nominalization.

(16) *ⁿka¹tʔa² tu¹ ti²ma⁴ⁿka²tsa⁴ʾⁿka:³² tsu²ʾʔba²ʔĩ² ja⁴**ʔa²***
ⁿka¹tʔa² *tu¹* *ti²⁽⁴⁾-* *m-* *a²ⁿka²tsa⁴ⁿka²* *=³⁽⁴⁾* *=:²* *tsu²ʔba²*
v: 1 3 15 16 19 21 29[1 16:19
because only PROG- HAB- run:3 =FOC [SUB HAB:wander:3
=ʔĩ³⁽⁴⁾ *ja²ʔa²*
25] 30
=PST.HAB] well
'because he only ran when he used to go around.' (180629-002 1:02)

In sum, minimal nominalization includes positions 3-20 and maximal nominalization 3-29.

3.6.3 Coordination (6-28, 2-29)

Coordination is also fractured into minimal and maximal interpretation. Minimal coordination includes all positions which cannot be elided, or if elided, the semantic scope changes. In (17), *ʃthe*^{3} 'garbage' is the only omitted or optional element in the second coordinated constituent. Thus, positions 6-28 correspond to the minimal interpretation of coordination test.

(17) *tsa*$^{2'}$*ʔbja*^{231} *ⁿka*^{2} *he*^{2} *ki*^{4}*ka*$^{2'}$*kɛ*^{3} *ˈʃthe*^{3} *ʔba*^{1} *he*^{2} *ki*^{4}*ka*^{2}*te*$^{2'}$*tɕe*^{2}

	tsa^{2}-	*ʔbe*^{23}	=*a*^{1}	*ⁿka*^{2}	*he*$^{2(4)}$=	*ki*^{2}*k*-	*a*^{2}*ka*$^{3(4)}$	=*i*	*ʃthe*$^{3(4)}$	*ʔba*^{1}
v:	16	19	28	1	[6	16	19	28	29]	1
	PFV-	know	=1SG	SUB	[already=	PFV-	burn:2	=2SG	garbage]	and

he$^{2(4)}$=	*ki*^{2}*k*-	*a*^{2}*te*^{2}*tɕa*^{2}	=*i*	
[6	16	19	28]	
[already=	PFV-	sweep:2	=2SG]	

'I know that you already burned the garbage and swept it.'

Maximal coordination, on the other hand, includes an entire sentence except for the conjunction itself (position 1), the attitudinal particles (position 30) and afterthoughts (position 31), i.e. positions 2-29.

3.6.4 Summary of subspan repetition

In sum, both total reduplication and verbal parallelism define positions 15-19 as their repeated subspans; nominalization specifies positions 3-20 minimally and 3-29 maximally; and coordination test covers positions 6-28 minimally, and 2-29 maximally.

3.7 Pauses and fillers (15-28)

Pausability, or possibility of having a pause, is defined here as the contiguous positions containing the verb core not interrupted by any pausable juncture.

Although this test is difficult to elicit—a speaker of a language may divide a string of speech and pronounce syllable by syllable, even if it has extralinguistic function, such as the clarification of pronunciation—, this test has been used as a constituency diagnostic (cf. Gerdts & Werle 2014: 609).

In this study, I identify pausable junctures from the transcribed instances of filler *hu*^{1}*ni*^{2} 'er' in naturally occurred speech. Specifically, I have observed that pauses with a filler *hu*^{1}*ni*^{2} may occur after a conjunction (position 1, example 18a), a proclitic such as negation *bi*^{4}= (position 8, example 18b), a second position clitic

in pre-predicate position such as *=hba⁴ni²³* 'at once' (position 9, example 18a) or reported information *=ʃu³⁽⁴⁾* (position 10, example 18c), an independent pronoun in patient-like or recipient-like function (position 14, example 18d), or between a person/number enclitic (position 28) and a following noun phrase (position 29), as in (18e). However, in a sample of 13 short texts with 129 transcribed instances of *hu¹ni²* 'er', none intrudes on the positions 15-28.

(18) a. *ʔba¹, **hu¹ni²**, tu¹ khja²ˈʔa⁴³ʃu³ ˈhⁿku²³hba⁴ni²³, **hu¹ni²**, ki⁴tsi²ˈⁿka⁴.*
ˈhⁿku²hba⁴ni²³ʃu³ ha²ˈne⁴

ʔba¹,	*hu¹ni²,*	*tu¹*	*khja²ʔa⁴*	*=³⁽⁴⁾*	*=ʃu³⁽⁴⁾*	*hⁿku²³*	*=hba⁴ni²³*	*hu¹ni²*	*ki⁴-*
v: 1	-	3	4	5	10	4	9	-	16
and	FILL	only	when	=FOC	=REP	one	=at.once	FILL	PFV-

tsi²ⁿka⁴	*hⁿku²*	*=hba⁴ni²³*	*=ʃu³⁽⁴⁾*	*ha²ne⁴*
19	4	9	10	16:19
burst	one	=at.once	=REP	PFV:sound

'and, er, suddenly one (thunderclap) at once, er, bursted, one roared at once.' (Sánchez Díaz & Nakamoto 2020: 138, English by SN)

b. *bi⁴, **hu¹ni²**, bi⁴ ˈtsha²¹nu⁴² ⁿtsu¹ˈʔba³²*

bi⁴=	*hu¹ni²*	*bi⁴=*	*tsha²*	*=¹nu⁴²*	*ⁿtsu¹ʔba³⁽⁴⁾*	*=a²*
v: 8	-	8	16:19	28	29[	]
n: -	-	-	-	-	10	12
NEG=	FILL	NEG=	HAB:give:1	=1SG/2PL	[mouth	=1SG]

'I don't, er, I don't give you my words (lit. my mouth).' (180624-002 15:44)

c. *ʔba¹ he²ʃu³, **hu¹ni²**, ʔba¹ he²ʃu³ kjo¹ tse²kʔe⁴ˈⁿtu²ⁿka²ni³ ⁿte¹*

ʔba¹	*he²⁽⁴⁾=*	*=ʃu³⁽⁴⁾*	*hu¹ni²*	*ʔba¹*	*he²⁽⁴⁾=*	*=ʃu³⁽⁴⁾*	*kjo¹*	*tse²k-*	*ʔe⁴ⁿtu²*
v: 1	6	10	-	1	6	10	14	16	19
and	already=	=REP	FILL	and	already=	=REP	there	PFV-	sit:PL

=ⁿka²ɲi³⁽⁴⁾	*ⁿte¹*	
26	30	
=again	thus	

'and already, er, and they already established themselves there again.' (180811-001-e2 04:08)

d. *ha¹ ˈkʷi²ru¹ nɛ¹ʔɛ²ni¹ˈsti²³na¹ ʔba¹ ni¹ˈma¹³ thi²ˈʔmi⁴re¹ tsa² ˈkʷi², **hu¹ni²**, ki²sʔe²ˈne:⁴¹ pre²si²den²³ˈte¹*

ha¹	*kʷi²*	*=ru¹*	*nɛ¹ʔɛ²+ni¹sti²³*	*=na¹*	*ʔba¹*	*ni¹ma¹*
v: 1	4[	]	4[	]	4	4
n: -	10	17	10	12	-	-
well	[PRONOM3	=ASSM]	[man+child	=POSS1SG]	like.that	much

=$^{3(4)}$	*thi^{2}-*	*m-*	*ʔĩ^{4}*	*=re^{1}*	*tsa^{2}*	*k^{w}i^{2}*	*hu^{1}ni^{2}*	*ki^{2}-*	*s-*
5	15	16	19	28	1	14	-	16	18
-	-	-	-	-	-	-	-	-	-
=FOC	PROG-	HAB-	be.told	=3/3	if	PRONOM3	FILL	PFV-	IMPERS-

ʔe^{2}ne^{4}	*=:^{1}*	*pre^{2}si^{2}den^{23}te^{1}*	
19	28	29	
-	-	-	
impose	=3/3	president	

‘well, I assume that they are telling it to my husband if they had given him, er, the *cargo* of president.’ (180630-001 16:32)

e. *…ʔba^{1} ^{n}ka^{2}, ^{n}ka^{2} si$^{1'}$khĩ^{2}re^{1},* ***hu^{1}ni^{2}****, ɕi^{2} he^{2} he$^{2'}$sun^{41}*

ʔba^{1}	*^{n}ka^{2}*	*^{n}ka^{2}*	*si^{1}+khĩ^{2}*	*=re^{1}*	*hu^{1}ni^{2}*	*ɕi^{2}*	*he$^{2(4)}$=*	*he^{2}sun^{4}*
v: 1	1	1	16:19	28	-	29[1	6	16:19
n: -	-	-	-	-	-	10	-	-
and	SUB	SUB	POT:make:3+far	=3/3	FILL	[REL	already=	PFV:die:PL

=^{1}			
-]			
15			
=ABS.ST]			

‘... and (he said) that, that it would keep him away from, er, those who already died.’ (180816-002 04:43)

In addition, during my participant observation as an Ayautla Mazatec learner, I have noticed that many Ayautla Mazatec speakers find difficult to follow my utterances if I have any interruption in positions 15-28. Therefore, I infer that a pause in these positions yields infelicitous utterances which require additional task of processing. Hence, the impossibility of having a pause defines a domain which consists of positions 15-28.

3.8 Stress assignment (19-20, 15-20)

Stress in Ayautla Mazatec is phonetically semi-long with an increased intensity, and is obligatory, culminative and predictably assigned on the right edge (or the end) of the stress domain. In order to determine the left edge (or the beginning) of stress domain, however, test fracturing is applied. According to the positive evidence, all stressable positions from the verb root onward are included (MINIMAL STRESS ASSIGNMENT). According to the negative evidence, all unstressed positions from the stressed syllable backward until the first unstressed position are included (MAXIMAL STRESS ASSIGNMENT).

Minimal stress assignment includes positions 19-20, which are verb roots or contiguous to verb roots and have the possibility to bear stress. In (19a), the final syllable of position 19 is stressed. In (19b), position 20 is stressed instead of the final syllable of position 19. Elements after position 20 do not shift the stress, as partially illustrated in (19c).

(19) a. *ba²'sẽ⁴*
b- *a²sẽ⁴*
v: 16 19
HAB- stand:3
'he stands.'

b. *ba²se²'ko¹³*
b- *a²sẽ⁴* *-ko¹³*
v: 16 19 20
HAB- stand:3 -COM
'he helps (lit. stands with)'

c. *ⁿku¹ ha¹ ba²se²'ko¹³hĩ⁴ni²³re¹ je²he²*
ⁿku¹ *ha¹* *b-* *a²sẽ⁴* *-ko¹³* *=hĩ⁴* *=ɲi²³* *=re¹* *je²he²*
v: 1 1 16 19 20 23 27 28 30
you.know well HAB- stand:3 -COM =INFR =ASR =3/3 anyway
'well, you know, he should help them anyway.' (181118-002 38:55)

Maximal stress assignment covers positions 15-20, where the stress is found only in 19-20, i.e. the domain of minimal stress assignment. Outside this domain, independent pronouns (position 14) have their own stress, as illustrated in (20).

(20) *bi⁴ 'kʷi² ni²'kʷɛ¹³*
bi⁴= *kʷi²* *ni²* *-ko¹³* *=i*
v: 8 14 16:19 20 =28
NEG= PRONOM3 HAB:do:2 -COM =2SG
'don't touch it.'

Aside from the phonetic correlates of duration and intensity, the stressed syllable has several phonotactic traits as its phonological correlates. Specifically, the stressed syllable tends to have more phonological contrasts. When a lexical root is found in unstressed syllables by suffixation or compounding, it tends to undergo denasalization (Nakamoto 2020: 110–111), deaspiration (Nakamoto 2020: 111–113), monosyllabification of disyllabic roots (Nakamoto 2020: 113–114) and tone neutralization (Nakamoto 2020: 154–161). In this study, however, I do not

treat these phonotactic traits as separate constituency tests. These neutralization processes are morphophonological in nature and the same phonotactic traits in stressed syllables may be found outside the minimal stress assignment domain, thus they cannot provide well-defined constituency diagnostics.

In summary, the domains of stress assignment can be positively defined as positions 19-20 and negatively as positions 15-20.

3.9 *ɛ.j constraint (19-19, 13-25)

*ɛ.j, or constraint against a sequence of /ɛ/ and /j/ at the syllable boundary, is remedied by alternating (or dissimilating) underlying /ɛ/ to /a/ when such a sequence occurs as a result of morpheme concatenation, i.e. ɛ > a / _ j. It is the only segmental constraint across the syllable boundary I have so far identified in Ayautla Mazatec (Nakamoto 2020: 97–98). Minimal *ɛ.j defines the position 19 as a domain within which this alternation takes place (21).

(21) *ʔba²ⁿ**tha⁴ˈ**ja²*
b- *ʔa²ⁿthɛ⁴+ja²*
v: 16 19
hab- change:3+pos:inside
'it (state, situation) changes.'

Maximal *ɛ.j can be defined by skipping over the junctures where this alternation cannot take place until one finds its initial and final positions. For example, between the third and the fourth syllables of (22a), or between the sixth and the seventh syllables of (22b). This domain includes positions 13-25. Note that no morpheme ends with /ɛ/ or begins with /j/ between positions 13-15 or 20-25.

(22) a. *ⁿkɛ²ˈʔɛ¹ⁿ**tshɛ⁴** ja²ˈkhã⁴*
ⁿkɛ²ʔɛ¹ *=ⁿtshɛ⁴* *j-* *a²khã⁴*
v: 4 12 16 19
here =always pfv- break
'he always broke it here.'

b. *ki²tsi²tɕi²kũ²tɛ̃²**ʔɛ̃²³**je²he²na¹*
ki²- *tsi²+tɕi²kũ²+tɛ̃²ʔɛ̃²³* *=je²he²* *=na¹*
v: 16 19 26 28
pfv- do:3+sacred+word(?) =all.inan =3/1sg
'he blessed all of them for me.'

Therefore, *ɛ.j defines 19 and 13-25 as its minimal and maximal interpretations.

3.10 *3.(2)4 constraint (19-19, 15-21)

Nakamoto (2020: 154–161) described that lexical tones except /1/ and /2/ tend to neutralize in pretonic syllables. Among such instances, neutralization of /3/ before /24/ and /4/ takes place obligatorily within certain domain, i.e. ***3.(2)4**. According to positive evidence, MINIMAL *3.(2)4 is obligatorily found in compound verbs (position 19). For example, in (23), the underlying /3/ neutralized obligatorily with /2/ before /4/. The syllable which undergoes neutralization is emphasized in boldface.

(23) *ba²**ne²'**sũ⁴*
b- *a²ne³⁽⁴⁾+sũ⁴*
v: 16 19
HAB- wash:3+POS:above
'he washes (the surface of).'

MAXIMAL *3.(2)4 can be established in positions 15-21 where negative evidence of *3.24 and *3.4 is available. In (24), the underlying sequence of /3/ in position 14 followed by /4/ does not undergo neutralization. In (25), the underlying sequence of /3/ followed by /4/ in position 22 does not undergo neutralization.

(24) *bi⁴ '**ɲa³** ja⁴'tʔa²na³*
bi⁴= *ɲa³⁽⁴⁾* *j-* *a⁴+tʔa²* *=na³⁽⁴⁾*
v: 8 14 16 19 28
NEG= PRONOM1INCL PFV- lay:3+POS:stuck =3INCL
'he didn't registered us.'

(25) *ba²'**ne³**hba⁴ni²³*
b- *a²ne³⁽⁴⁾* *=hba⁴ni²³*
v: 16 19 =22
HAB- wash:3 =at.once
'he washes at once.'

Therefore, *3.(2)4 constraint defines position 19-19 as its minimal interpretation and positions 15-21 as its maximal interpretation.

3.11 Syllable-internal segmental interactions (16-19, 16-28)

Given that every free form in Ayautla Mazatec begins with a consonant and ends with a vowel (cf. Nakamoto 2020: 83–85), the existence of SYLLABLE-INTERNAL

SEGMENTAL INTERACTIONS suggests some grade of fusion between morphemes.[8] Specifically, aspect/mode (position 16), associated motion (position 17), voice (position 18) and verb root (position 19) have consonant-initial morphemes, while associated motion (position 17), voice (position 18), verb roots (position 19) and pronominal enclitics (position 28) include vowel-initial morphemes. Example (26) illustrates some syllable-internal segmental interactions around the verb root: habitual prefix is syllabified with andative; causative prefix is syllabified with verb root; and verb root is syllabified with pronominal enclitic.

(26) *hbi²tsi²k?o⁴ʲja²³¹*

	b-	*hi²-*	*tsi²⁽⁴⁾k-*	*?o²³+ja²³*	*=a¹*
v:	16	17	18	19	28
	HAB-	ANDT:1-	CAUS-	go.out:1+POS:inside	=1SG

'I put out, switch off.'

MINIMAL SYLLABLE-INTERNAL SEGMENTAL INTERACTIONS can thus be defined as 16-19, the span in which all elements are known to show syllable-internal segmental interactions, while MAXIMAL SYLLABLE-INTERNAL SEGMENTAL INTERACTIONS covers 16-28, outside which syllable-internal segmental interactions are not found.

3.12 Disyllabic sandhi-blocking tone sequences (15-19, 15-21)

Tone sandhi in Ayautla Mazatec is a phonological process which consists of a progressive association of floating /4/ across syllables (= tone bearing units), /4/ being the highest tone and /1/ the lowest. As a result of tone sandhi, the syllables with underlying /1/, /13/, /2/ or /23/ are generally substituted by /4/. However, the applicability, the obligatoriness, and the output of tone sandhi are subject to tonal and prosodic conditions of the syllable receiving the floating /4/ (Nakamoto 2020: 171–196).

Among several phonological and non-phonological conditions which block the application of tone sandhi (Nakamoto 2020: 184–191), DISYLLABIC SANDHI-BLOCKING TONE SEQUENCES constitute one of the tonal and prosodic conditions (the other being '1possible sandhi", see §3.14). If the syllable which receives the floating /4/ is the first syllable of a /1.24/, /1.4/, /2.24/ or /2.4/ sequence within the

[8]The situation is different with tonal morphemes, because the stem or the host to such tonal morphemes always has its own tone(s) and thus is pronounceable. Therefore, the fusion between the stem or host and the tonal morpheme is a phenomenon local to each juncture.

positions 15-19, the application of sandhi is blocked.[9] For example, in (27a) and (27b), the second syllable of the example is part of the underlying /2.4/ sequence and tone sandhi fails to apply, while in (27c), tone sandhi does apply to the second syllable which is not part of a /2.4/ sequence.

(27) a. *he² ti²tsi⁴'the²*

	he²⁽⁴⁾=	*ti²⁽⁴⁾-*	*tsi⁴-*	*the²*
v:	6	15	16:18	19
	already=	PROG:3-	HAB:CAUS:3-	cough

'he is already clearing his throat.'

b. *he² ti⁴tsi⁴'the²

c. *he² ti⁴ma²'hɲu⁴*

	he²⁽⁴⁾=	*ti²⁽⁴⁾-*	*m-*	*a²-hɲu⁴*
v:	6	15	16	19
	already=	PROG:3-	HAB-	INCH-night

'it is already getting dark.'

In contrast, such underlying sequences fail to block tone sandhi if one of the two syllables is found outside the positions 15-21. For example, the second and the third syllables in (28a), in positions 14 and 16, have an underlying /1.4/ sequence, but it does not block tone sandhi. The same is true in (28b), where the underlying /1.4/ sequence in positions 19 and 22 cannot block the application of tone sandhi.

(28) a. *bi⁴ʔĩ² 'hĩ⁴¹ tsu⁴'ja²ni²³*

	bi⁴=	*=ʔĩ³⁽⁴⁾*	*hĩ¹*	*tsu⁴+ja²*	*=ɲi²³*
v:	8	12	14	16:19	27
	NEG=	=PST.HAB	PRONOM1INCL	HAB:say:3+POS:inside	=ASR

'he did not used to explain it to us.'

b. *skhe²'ɕi⁴¹hba⁴ni²³*

	s-	*khe³⁽⁴⁾+ɕi¹*	*=hba⁴ni²³*
v:	16	19	22
	POT-	pull+piece(?)	=at.once

'he blows his nose at once.'

Therefore, minimal disyllabic sandhi-blocking tone sequences spans 15-19, while maximal disyllabic sandhi-blocking tone sequences covers 15-21.

[9] Similar blocking conditions have been reported for other Mazatec varieties with tone sandhi, such as Soyaltepec (Pike 1956: 63–64) and Chiquihuitlán (Nakamoto 2018).

3.13 Obligatory sandhi (15-28)

Tone sandhi in Ayautla Mazatec, or progressive association of a floating /4/, is obligatory within the positions 15-28 and is optional outside this domain. Obligatory sandhi is illustrated in (29). Sandhi from progressive (position 15) to habitual and inchoative (positions 16 and 18) as well as sandhi from verb root (position 19) to pronominal clitic (28) are obligatory. If sandhi does not apply in any of these positions, the result is ungrammatical (29b, 29c, 29d), where the syllables with underapplication of sandhi are highlighted.

(29) a. *ti²ma⁴ni²ˈhɲa²na⁴²*

	ti²⁽⁴⁾-	*m-*	*a²-*	*ni²hɲa³⁽⁴⁾*	*=na¹*
v:	15	16	18	19	28
	PROG:3-	HAB-	INCH-	be.sleepy	=3/1SG

‘I’m getting sleepy.’

b. *ti²**ma²**ni²ˈhɲa²na⁴²

c. *ti²ma⁴ni²ˈhɲa³**na¹**

d. *ti²**ma²**ni²ˈhɲa³**na¹**

However, tone sandhi from position 14 or to position 29 is optional, as illustrated in the following examples. In (30a), tone sandhi from the second morpheme (position 14) may or may not apply; if applied, the underlying tone /3/ alternates with /2/ (cf. Nakamoto 2020: 142–143). We can observe the same in (30b) with the first floating /4/ associated with the third morpheme (position 23). So far I have not been able to identify what else conditions the application of sandhi outside the domain of obligatory sandhi in Ayautla Mazatec.

(30) a. *bi⁴ ˈɲa³ **su¹**ˈba¹na³*
*~ bi⁴ ˈɲa² **su⁴**ˈba¹na³*

	bi⁴=	*ɲa³⁽⁴⁾*	*su¹ba¹*	*=na³⁽⁴⁾*
v:	8	14	16:19	28
	NEG=	PRONOM1INCL	POT:catch	=3/1INCL

‘he won’t catch us.’

b. *ja²te²ˈɲa²³ʃu³ **thju¹**na²ˈɲa²re⁴²*
*~ ja²te²ˈɲa²³ʃu² **thju⁴**na²ˈɲa²re⁴²*

	j-	*a²te²ɲa²³*	*=ʃu³⁽⁴⁾*	*thju¹na³ɲa³⁽⁴⁾*	*=re¹*
v:	16	19	23	29[	]
n:	-	-	-	10	12
	PFV-	sell:3	=REP	[dog	=POSS3]

‘(reportedly) he sold his dog.’ (Nakamoto 2020: 179)

Therefore, the obligatoriness of sandhi defines 15-28 as its domain.

3.14 Possible sandhi (15-28, 2-31)

In addition to the disyllabic sandhi-blocking tone sequences (§3.12), the other prosodic condition which impedes the application of tone sandhi is that of POSSIBLE SANDHI domain. Nakamoto (2020: 176–177) reported that tone sandhi in Ayautla Mazatec is blocked across the boundary of coordination and verbal parallelism construction. For example, in (31a), tone sandhi is blocked between asyndetically coordinated verbs, i.e. *kʷʔi³⁽⁴⁾* 'he will drink' and *ski¹ⁿta²ˈja⁴* 'he will cry', as well as *se³⁽⁴⁾* 'he will sing' and *ste¹* 'he will dance'; in (31b), tone sandhi is blocked between two verb forms in a verbal parallelism construction (cf. §3.6.1), i.e. *thu²tsha³⁽⁴⁾* and *thu²ni²ɲa²*. In both cases, the rest of phonological conditions (tone and stress) for tone sandhi to be realized are satisfied (cf. Nakamoto 2020: 180–184); the syllable which would receive a floating tone /4/ is highlighted.

(31) a. *ˈkʷʔi³, **ski¹ⁿ**ta²ˈja⁴, ˈse³, ˈ**ste¹**, kʷha¹hɲa²ˈhbe⁴*
kʷʔi³⁽⁴⁾ *ski¹ⁿta²ja⁴* *se³⁽⁴⁾* *ste¹*
v: [16:19] [16:19] [16:19] [16:19]
[POT:drink] [POT:cry] [POT:sing] [POT:dance]
kʷha¹hɲa²+hbe⁴
[16:19]
[POT:lie.down+POS:asleep]
'he will drink, cry, sing, dance and sleep.' (Nakamoto 2020: 176)

b. *thu²ˈtsha³ **thu²**ni²ˈɲa²*
thu²+tsha³⁽⁴⁾
v: vpar1[16:19]
vpar1[HAB:come.out+POS:sideways]
thu²+ni²ɲa²
vpar2[16:19]
vpar2[HAB:come.out+POS:quadrupedal]
'he staggers.' (Nakamoto 2020: 177)

This fact can be interpreted as follows: tone sandhi cannot extend over two planar structures. However, Ayautla Mazatec has additional prosodic restrictions as to its application: within the verbal planar structure, even though the rest of phonological conditions (i.e. tone and stress) are met, connectors (position 1), focus introducers (polar question, 'even', 'only'; position 3) and adverbs (positions 4, 14, 29) do not undergo sandhi. For example, in (32), even though *tu¹* 'only'

satisfies the rest of tonal and stress-related conditions, it never undergoes tone sandhi.

(32) *he² ti²ma⁴ˈba²ʔĩ²na⁴², nɛ¹ʔɛ³, ˈɲa⁴ ˈ**tu¹** kʷhɛ²ˈʔɛ¹ni²³*

he²⁽⁴⁾=	*ti²⁽⁴⁾-*	*m-*	*a⁴-ba²*	*=ʔĩ³⁽⁴⁾*	*=na¹*	*nɛ¹ʔɛ³⁽⁴⁾*	*ɲa⁴⁽⁴⁾*	*tu¹*
v: 6	15	16	19	25	28	31	4	4
already=	PROG-	HAB-	INCH-sad	=PST.HAB	=3/1SG	sir	where	only

kʷhɛ²ʔɛ¹ =ni²³
16:19 27
be.from =ASR
'I was already worrying, sir, where the hell do you come from?'
(Sánchez Díaz & Nakamoto 2020: 141, English by SN)

Since the positions without restrictions on applying tone sandhi are discontinuous, the possible sandhi test is fractured into maximal and minimal interpretations: MINIMAL POSSIBLE SANDHI spans positions 15-28—where no position includes morphemes which block tone sandhi not by underlying tone or stress—, while MAXIMAL POSSIBLE SANDHI spans positions 2-31 where all positions without such prosodic restrictions are included.

4 Summary and discussions

Figure 1 summarizes the constituency diagnostics in Ayautla Mazatec described in this study, sorted by domain size. Looking at the domains, the span 15-28 ("Layer 9" in the Figure) has the highest number of convergences of 7, followed by 15-19 ("4") with 5 convergences, followed by 19 ("1") with 3 convergences, and 15-20 ("5"), 15-21 ("6") as well as 6-28 ("11") with 2 convergences.

Looking at the individual positions, progressive (position 15) is the initial position for 16 diagnostics. As for the final position, the bound pronouns (position 28) has the highest convergence rate with 10 diagnostics, followed by the verb root (position 19) with 9 diagnostics.

A closer look at individual diagnostics reveals that the convergences are mainly observed among morphosyntactic diagnostics (see also figures 2 and 3 below for convergence of morphosyntactic and phonological domains, respectively). For example, out of the 5 diagnostics which converge at the span 15-19, 4 are morphosyntactic diagnostics while only one of them is phonological; in the same vein, 4 out of the 7 diagnostics at the span 15-28 are morphosyntactic diagnostics while two are phonological and the other concerns the pause. Therefore, at least based on morphosyntactic criteria, we can recognize two strong candidates

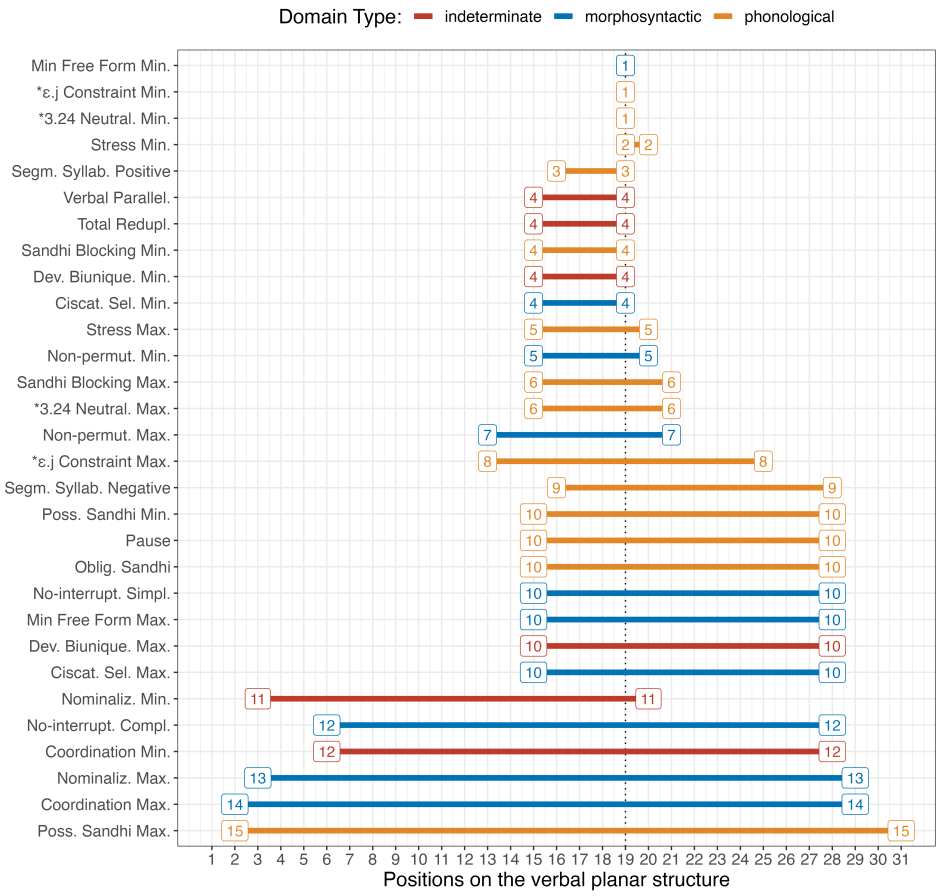

Figure 1: Convergence of domains

for words or phrases. Contrasting these candidates with my working concept of affix-clitic distinction, the span 15-19 differs from it in excluding the comitative (position 20).[10]

However, phonological diagnostics (cf. Figure 3) tend not to converge with other phonological diagnostics in Ayautla Mazatec. Out of 13 diagnostics, we have convergences of at most two phonological constituency diagnostics, namely, minimal *ɛ.j constraint and that of minimal *3.(2)4 constraints at position 19; maximal *3.(2)4 constraint and maximal sandhi-blocking tone sequences at positions

[10] I included comitative in my working concept of word, because it is inside the stress assignment domain (see §3.8).

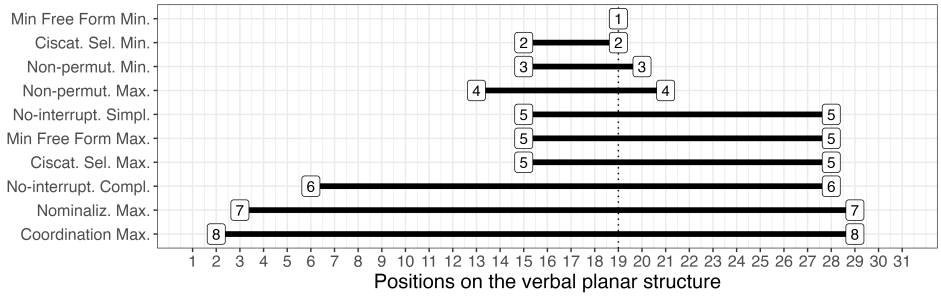

Figure 2: Convergence of morphosyntactic domains

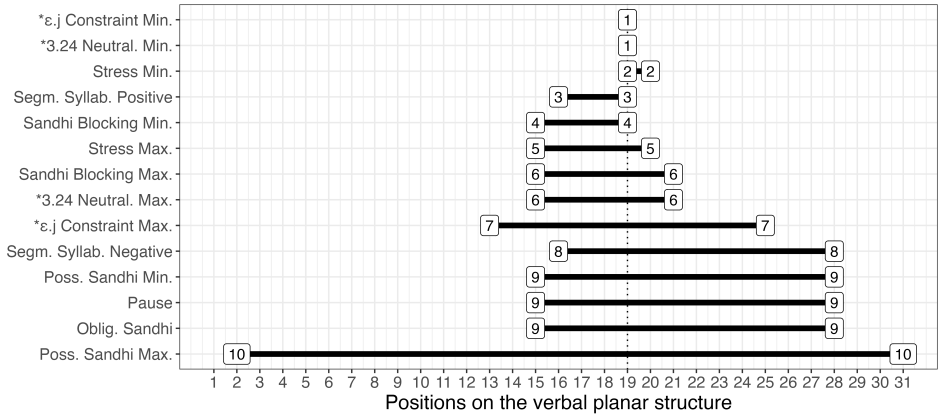

Figure 3: Convergence of phonological domains

15-21; and obligatory sandhi and minimal possible sandhi at positions 15-28. It is partly due to my analytical decision of not treating the correlates of stress as separate constituency diagnostics (see §3.8). It is also partly due to the fact that some positions do not have adequate phonological context for a given diagnostic. For example, the focus marker *=³(⁴)* (positions 5, 21) consists of tones and cannot provide any positive evidence for segmental processes; comitative *-ko¹³* is the only morpheme in position 20 and is not amenable to some tonal tests.[11]

[11]A reviewer suggested ignoring positions for tonal morphemes, which would correspond to focus marker (positions 5 and 21), some of the aspect/mode markers (position 16) and some of the person/number markers (position 28), as well as absolute state marker in noun complexes (position 15). However, tones in Ayautla Mazatec are basically concatenated in linear order, such as /1/ + /3/ + /1/ > /131/ and /3/ + /1/ + /3/ > /313/ (cf. Nakamoto 2020: 196–207). Therefore, it is important to represent the linear order of these morphemes.

Diachronically, these variations of prosodic constituents in one or two positions, namely, between positions 15 and 16 and between positions 19 and 20, seems to be accounted for by recent grammaticalizations. Both progressive (position 15) and comitative (position 20) have their etymologies identifiable outside the verb complex, namely, posture verbs and preposition 'with', respectively. A possible interpretation is to see from aspect/mood (position 16) to verb root(s) (position 19) as a historically stable and more established constituent, and to see the prosodically indeterminate status of progressive (position 15) and comitative (position 20) as a result of their recent grammaticalization on the way to cohere with the inner constituent.

Synchronically, however, Ayautla Mazatec situation supports the position of Schiering et al. (2010: 704) who state that 'prosodic domains are conceived of as language-particular, intrinsic and highly specific properties of individual phonological rules or constraints'. Therefore, as same as in Tibeto-Burman language Limbu studied in Schiering et al. (2010), we cannot accommodate Ayautla Mazatec prosodic constituents into some allegedly universal hierarchy of phonological stem, word or phrase.

In conclusion, Ayautla Mazatec verbal predicates show six domains where two or more constituency diagnostics converge. The convergences are mainly observed among morphosyntactic diagnostics, and phonological domains tend not to converge in this language. This result supports the non-universality of prosodic domains suggested by Schiering et al. (2010) and, due to the lack of a strong candidate for a phonological word, is against the word bisection thesis advocated by Dixon & Aikhenvald (2002).

Acknowledgements

Ayautla Mazatec speakers, especially Josefina Sánchez; Hiroto Uchihara, Adam J. R. Tallman, Sandra Auderset, Naonori Nagaya, and an anonymous reviewer for their comments; Sandra Auderset for her help in plotting constituency diagnostics; errors and shortcomings are exclusively mine.

Abbreviations

ABS.ST	absolute state	CAUS	causative
ANDT	andative	COM	comitative
ASR	assertive	FILL	filler
ASSM	assumed	FOC	focus

HAB	habitual	POT	potential
IMPERS	impersonal	PROG	progressive
INAN	inanimate	PRONOM	pronominal element
INCH	inchoative	PST	past
INCL	inclusive	Q	question particle
INFR	inferred	REL	relative
NEG	negative	REP	reportative
NMLZ	nominalizer	SAP	speech act participant
PFV	perfective	SG	singular
PL	plural	SUB	subordinator
POS	positional	TOP	topic
POSS	possessive		

References

Campbell, Eric W. 2024. Constituency in Zenzontepec Chatino. In Adam J.R. Tallman, Sandra Auderset & Hiroto Uchihara (eds.), *Constituency and convergence in the Americas*, 363–413. Berlin: Language Science Press. DOI: ??.

Cowan, George & Florence Cowan. 1947. Mazateco: Locational and directional morphemes. *Aboriginal Linguistics* 1. 3–9.

Dixon, R. M. W. & Alexandra Y. Aikhenvald. 2002. Word: A typological framework. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: A Cross-linguistic Typology*, 1–41. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511486241.002.

Dryer, Matthew S. 2007. Word order. In Timothy Shopen (ed.), *Language typology and syntactic description*, 61–131. Cambridge: Cambridge University Press.

Fernández de Miranda, María Teresa. 1951. Reconstrucción del protopopoloca. *Revista mexicana de estudios antropológicos* 12. 61–93.

Gerdts, Donna B. & Anna Werle. 2014. Halkomelem clitic types. *Morphology* 24. 245–281.

Gudschinsky, Sarah C. 1958. *Proto-Popotecan: A comparative study of Popolocan and Mixtecan*. Baltimore: Waverly Press.

Gutiérrez, Ambrocio & Hiroto Uchihara. 2024. Words as emergent constituents in Teotitlán del Valle Zapotec. In Adam J.R. Tallman, Sandra Auderset & Hiroto Uchihara (eds.), *Constituency and convergence in the Americas*, 303–361. Berlin: Language Science Press. DOI: ??.

Hamp, Eric. 1958. Protopopoloca internal relationships. *International Journal of American Linguistics* 24. 150–153.

Hamp, Eric. 1960. Chocho-Popoloca innovations. *International Journal of American Linguistics* 26. 62.

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80. DOI: 10.1515/flin.2011.002.

Kalstrom Dolson, Marjorie, Jeanne Austin Krumholz & Doris Bartholomew Ewan. 1995. Gramática del popoloca de San Juan Atzingo. In Jeanne Austin Krumholz, Marjorie Kalstrom Dolson & Miguel Hernández Ayuso (eds.), *Diccionario popoloca de San Juan Atzingo, Puebla*, 283–390. Tucson: Summer Institute of Linguistics.

Krumholz, Jeanne Austin, Marjorie Kalstrom Dolson & Miguel Hernández Ayuso. 1995. *Diccionario popoloca de San Juan atzingo, Puebla*. Tucson: Instituto Lingüístico de Verano.

Nakamoto, Shun. 2018. Tone and tonal processes in Chiquihuitlán Mazatec. Presentation at Sound Systems of Latin America III, University of Massachusetts at Amherst. Oct. 19-21.

Nakamoto, Shun. 2020. *Tonología mazateca: San Bartolomé Ayautla*. MA thesis, Universidad Nacional Autónoma de México. Universidad Nacional Autónoma de México. (MA thesis).

Nichols, Johanna. 1986. Head–marking and dependent–marking grammar. *Language* 62(1). 56–119.

Pike, Eunice V. 1956. Tonally differentiated allomorphs in Soyaltepec Mazatec. *International Journal of American Linguistics* 22(1). 57–71.

Pike, Eunice V. 1967. Huautla de Jiménez Mazatec. In Norman A. McQuown (ed.), *Handbook of Middle American Indians. Volume 5: Linguistics*, 311–330. Austin: University of Texas Press.

Pike, Kenneth L. 1948. *Tone languages: A technique for determining the number and type of pitch contrasts in a language, with studies in tonemic substitution and fusion*. Ann Arbor: The University of Michigan Press.

Sánchez Díaz, Josefina & Shun Nakamoto. 2020. Na'ain Cho'on/El señor del trueno: Un texto en mazateco de Ayautla. *Tlalocan* 25. 115–145.

Sandra Auderset, Carmen Hernández Martínez & Albert Ventayol-Boada. 2024. Constituency in Tù'un Ntá'ví (Mixtec) of San Martín Duraznos. In Adam J.R. Tallman, Sandra Auderset & Hiroto Uchihara (eds.), *Constituency and convergence in the Americas*, 263–301. Berlin: Language Science Press. DOI: ??.

Schiering, René, Balthsar Bickel & Kristine A. Hildebrandt. 2010. The prosodic word is not universal, but emergent. *Journal of Linguistics* 46(3). 657–709.

Tallman, Adam J. R. 2020. Beyond grammatical and phonological words. *Language and Linguistics Compass* 14(2). 1–14. DOI: 10.1111/lnc3.12364.

Tallman, Adam J. R. 2021. Constituency and coincidence in Chácobo (Pano). *Studies in Language* 45(2). 321–383. DOI: 10.1075/sl.19025.tal. (20 April, 2023).
Wackernagel, Jacob. 2020. *On a law of Indo-European word order: über ein Gesetz der indogermanischen Wortstellung*. Berlin: Language Science Press.