

Chapter 1

Words as emergent constituents in Teotitlán del Valle Zapotec

Ambrocio Gutiérrez^a & Hiroto Uchihara^b

^aUniversity of Colorado Boulder ^bTokyo University of Foreign Studies

This chapter reports the results of the application of 21 constituency tests to the verbal complex and 18 tests to the nominal complex in Teotitlán del Valle Zapotec (TdVZ). These tests are both morphosyntactic and phonological. Assuming “words” are identified as domains of structure where constituency diagnostics converge (e.g. Matthews 2002), our goal is to assess what type of word constituents, if any, are motivated in TdVZ. In contrast to recent work emphasizing the ubiquity of wordhood domain divergence (Haspelmath 2011; Bickel & Zúñiga 2017; Tallman 2020), we argue that the data of TdVZ provide support for at least one type of word constituent in this language. A novel method for estimating the chance probability of constituency diagnostic convergence supports the main thesis of this chapter.

1 Introduction

Linguists do not agree with each other with respect to whether speaker intuitions or some set of established wordhood criteria should be used to discern word boundaries (Sapir 1921: 33-34; Aronoff & Fudeman 2005: 36; Mithun 2014: 74 vs. Bloomfield 1914). In fact, there is no joint set of necessary and sufficient linguistic criteria for identifying words (Haspelmath 2011) and speakers do not always have consistent intuitions regarding word boundaries. For instance, the following example illustrates that the definition of the ‘words’ based on the speaker intuition (1a) can be quite different from that based on one of the traditional linguistic criteria (1b) of ‘grammatical word’ assumed in much of the literature on Zapotec (Broadwell 2000, Munro 2004, Beam de Azcona & Cruz Santiago 2022). This will be discussed in detail in this chapter on Teotitlán del Valle Zapotec.

- (1) a. [speaker intuition]
 (kēd'gwæ: 'zlyā' dízyán) _{Word} ('llwâ') _{Word}
 b. [wordhood criteria]
 (kēd'gwæ:) _W ('zlyā') _W (dí) _W (zī) _W (án) _W ('llwâ') _W
kēd=*gw-æ:* *zlyā'*=*dí=zī=an* *llwâ'*
 NEG=COMPL-go in.vain=NEG=only=3SG.INF Oaxaca
 'S/he did not go only in vain to Oaxaca.'

Notably, speaker intuitions do not always associate the word boundary with a stress domain: orthographic practices of speakers vary in terms of how synthetic they represent the language. On the other hand, linguistic wordhood criteria tends to see this language as much less synthetic because of the high degree of promiscuity of bound morphemes ("clitics"), which are regarded as syntactically placed.

In order to resolve divergences, linguists have proposed a distinction between a phonological and a morphosyntactic word(s) (for instance Dixon & Aikhenvald 2002). However, as Matthews (2002) comments: "there are divergences within these two domains. It seems that no criterion is either necessary or sufficient, but they are relevant insofar as, in particular languages, they tend to coincide" (Matthews 2002: 274). Also, Haspelmath (2011) remarks that "in order to show that a fuzzy concept of word is theoretically significant, one would have to demonstrate that grammatical units are not randomly distributed over the continuum between fully bound and fully independent units, but that they cluster significantly".

In this chapter we assess whether a word constituent can be motivated in TdVZ based on the notion of convergence beyond chance stated in the introduction of this book. We apply 21 constituency tests to the verbal complex and 18 tests to the nominal complex to TdVZ. These tests are both morphosyntactic and phonological. Our goal is to assess what type of (word) constituents, if any, are motivated in this language.

We argue that the data of TdVZ provide support for at least one type of word constituent in this language, in contrast to recent work emphasizing the ubiquity of wordhood domain divergence (Schiering et al. 2010; Haspelmath 2011; Bickel & Zúñiga 2017; Tallman 2020). This study thus suggests that the high degree of misalignments found in Bickel & Zúñiga (2017) may result from the consideration of an arbitrarily low number of diagnostics (they only consider 6). Furthermore, the misalignments found in Chácobo (Tallman 2021) may be something peculiar to this language and not necessarily general. The upshot of our results is that they show that languages vary in terms of the degree to which wordhood diagnostics

cluster, even within the same language in different domains, in this specific case, verbal vs nominal.

After this introduction, the following section summarizes some features of Zapotec, the third section then provides an overview of the verbal template in TdVZ. The fourth part of this chapter provides an overview of divergences and convergences motivated by constituency diagnostics. In both categories (verbal (§4) and nominal (§5)), we show that in the morphosyntactic domain divergences are just as common as convergences. Methodological, theoretical and typological implications are discussed in §6.

2 Teotitlán del Valle Zapotec

Teotitlán del Valle Zapotec (ISO 639-3: [zab]) is a Central Valley Zapotec variety that belongs to the Zapotecan language family spoken in the Mexican state of Oaxaca. The Zapotecan language family in turn is one branch of the Otomanguean language stock. Zenzontepec Chatino (Campbell 2024 [this volume]), Mazatec (Nakamoto 2024 [this volume]), and Duraznos Mixtec (Sandra Auderset & Ventayol-Boada 2024 [this volume]) also belong to this language stock. This section gives an overview of the basics of phonology (§2.1) and morphosyntax (§2.2) of TdVZ.

2.1 Phonology

TdVZ has six vowels (*i*, *æ*,¹ *a*, *o*, *u*, *ɨ*), of which *ɨ* is marginal. Vowel length is mostly predictable from the position of prominence and the consonant that follows: in prominent syllables the vowel is usually long when followed by a lenis consonant or no consonant, and otherwise the vowel is short; this is due to the requirement that demands that a prominent syllable be at least bimoraic. However, this is not always the case since all loanwords and some native words have a long vowel even though the prominent vowel is followed by a fortis consonant (e.g. *llú:py* ‘Guadalupe’, *ʃú:k* ‘kiss’, *gá:ti* ‘not yet’). For this reason, we consider vowel length in TdVZ to be marginally contrastive, and thus we mark vowel length with a colon (:) even when it is predictable, as in *ru:* ‘cough’.

TdVZ also contrasts three phonation types, modal (*a*), creaky (*ǁ*) and glottalized (*aʔ*); the contrast is justified by a triplet contrasting only in the phonation

¹In TdVZ /*æ*/ is realized as [e] in certain phonological contexts (Uchihara & Gutiérrez 2020a). Given that the distribution of these allophones depends on complex phonological factors, i.e., height of the adjacent consonant, syllable structure and accent, we will represent both allophones throughout this chapter.

types, such as *ru*: ‘cough’, *ry*: ‘carry’, and *ru*’ ‘mouth’. As in other Otomanguean languages, TdVZ is tonal; five contrastive tonal patterns are found on one syllable: low (*a*), mid (*ā*), high (*á*), falling (*â*) and rising (*ǎ*).

TdVZ has 24 native consonants (*b*, *p*, *d*, *t*, *g*, *k*, *gw* [g^w], *kw* [k^w], *m* [m:], *n*, *nn* [n:], *r*, *z*, *s*, *ʒ*, *f*, *j* [x], *l*, *ll* [l:], *ts*, *čʒ*, *ʃ*, *y* [j], *w*), most of which come in pairs of lenis and fortis consonants, such as *d* vs. *t* and *n* vs. *nn* (as in other Zapotec languages: cf. Nellis & Hollenbach 1980). Fortis obstruents are voiceless, never fricated if they are stops, and relatively long. Lenis obstruents are often voiced, variably fricated, and relatively short. Duration is the main difference between lenis and fortis sonorants.

2.2 Morphosyntax

The basic morphosyntactic characteristics of Teotitlán del Valle Zapotec are as follows. The basic word order is VSO, and this language exhibits most of the predicted correlations cited by Dryer (1992) for VO languages. As in other Zapotecan languages (cf. Lee 1999), TdVZ is a head marking language and no case markings appear on NPs. Below we discuss in detail the morphosyntax of verbs, and then we discuss the nouns.

In Teotitlán del Valle Zapotec, a verb stem may be morphologically simple (*-a:w* ‘eat’) or complex. A morphologically complex verb stem may consist of a root plus the diminutive and the comitative suffix (*-ak-nā̃*: [be.done-COMIT] ‘help’) or a verb root plus another lexical root. A verb root can be compounded with a noun root (*-nnyab-dī:čʒ* [ask.for-word] ‘ask (a question)’), adjective root (*-ak-nall* [occur/be.done-cold] ‘feel cold’), or a numeral (*-ka-tap* [sound-four] ‘be at four o’clock’).

A verb stem obligatorily takes a tense/aspect/modal (TAM) prefix, as in (2–4), and some verbs exhibit one of the voice prefixes, the causative *u-* or the “restorative” *a-*,² as shown in (5–6). These prefixes, however, are generally glossed as part of the TAM prefix throughout this work since in many cases they are lexicalized.

- (2) 'rya:b
 r-ya:b
 HAB-fall.down
 ‘(S/he) falls down.’

²The main function of “restorative”, at least in Teotitlán Zapotec, is to encode middle voice rather than the restorative function proposed by Smith-Stark (2002). However, since this terminology is widely used among Zapotecanists, we adhere to ‘restorative’.

- (3) 'bya:b
b-ya:b
 COMPL-fall.down
 '(S/he) fell down.'
- (4) ri'ɕi'
ri-ɕi'
 HAB-get.clogged
 '(It) gets jammed.'³
- (5) ra'ɕi'
r-a-ɕi'
 HAB-REST-get.clogged
 '(It) gets clogged.'⁴
- (6) ru'tʃi'
r-u-tʃi'
 HAB-CAUS-cover
 '(S/he) covers.'

When the subject is not expressed as a full NP, a pronominal enclitic encoding this grammatical role (and 3rd person object in some cases) occurs after the verb stems.

- (7) 'rya:ban
r-ya:b=an
 HAB-fall.down=3SG.INF
 'S/he falls down.'
- (8) 'ryā:bdān
r-ya:b=dān
 HAB-fall.down=3PL.FOR
 'They fall down.'

A verb may optionally take proclitics and adverbial enclitics. Here, we consider that clitics in TdVZ are morphemes that are phonologically defective and must join with another syntactic terminal element to form a prosodic word (Sapir 1930: 71), following the traditional analysis of Zapotec. In addition to this, clitics in TdVZ, especially enclitics, can occur in various positions within the clause.

³Agent is assumed, such as when a road gets clogged/jammed with people or cars.

⁴Agent is not assumed, such as when a nose gets clogged with mucus.

Proclitics occur before the TAM prefix while enclitics come between the stem and the pronominal enclitics. This is illustrated in (9):

- (9) ni'ra:wzátū
ni=r-a:w=zá=tū
SUB=HAB-eat=also=2PL.INF
'What y'all also eat.'

Uninflected nouns may be morphologically simple (*gæ't* 'tortilla') or complex. A morphologically complex noun may consist of a fossilized prefix plus root (*bi-zi:n* 'mouse') or of compounded roots (*kʲe-ʲu'* [head.of-building] 'roof'). Nouns may take a plural marker *d'* (*d=bénny* [PL=person] 'people')⁵ and the diminutive suffix (*bækw-æ'n* [dog-DIM] '(a) nice/little dog').

Nouns can be inflected for a possessor. Alienable nouns take the possessive prefix *f-*, which may provoke fortification of the root (*f-kæt* [POSS-tortilla] 'tortilla of'), and pronominal enclitics when the possessor is not expressed as a full NP (*f-kæt=an* [POSS-tortilla=3SG.INF] 'her/his tortilla'). Inalienable nouns do not require the prefix *f-*. In addition, when the noun is possessed by a 1PL possessor, the noun takes proclitic *dū=* as well as the 1PL pronominal enclitic *=un* (INCL) or *=ūn* (EXCL) (*dū=f-kæt=un* [1PL=POSS-tortilla=1PL.INCL] 'our tortilla').

3 Data presentation and concepts: planar structure(s) and constituency tests

In this section we briefly lay out various concepts related to the methodology followed. These have been broadly discussed and defined in the introduction of this book. Thus, here we only highlight specific notions for our chapter.

3.1 Planar structure(s)

A planar structure is a templatic structure that represents all elements of some (verbal or nominal) domain regardless of constituency structure, motivated or not. Thus syntactic, and morphological elements will be displayed on the same level in this structure. Planar structures are built out of a number of parts: elements, positions, slots, and zones.

Positions are organized into a template that captures their linear order. Variable ordering of elements is captured, in the first place by placing the elements

⁵Various elements in TdVZ occur together with a floating high tone that triggers Tone Sandhi (§4.2.6, §5.2.6). This floating tone is represented with an acute accent throughout this work.

in zones, and in the second place by allowing elements to occur in more than one position in the planar structure if it is necessary. A more detailed exposition of the distinction between slots and zones can be found in the introduction of this book.

Tables 1 and 2 present the planar structure of the TdVZ verbal predicate construction and the nominal construction. These will be referred to throughout the rest of this chapter. A detailed defense of the relative ordering and identification of positions in the TdVZ verb and nominal complex is found in Gutiérrez & Uchihara (n.d.).

We define the verb base as the stem minus the tense/aspect/mood (TAM) prefix. As mentioned above, the TAM prefix is segmented in such way that it usually contains the fossilized causative (*u-*) or restorative (*a-*) vowel morphemes. This means that the position of these fossilized prefixes is not represented in the template. In the same vein, the (possibly) fossilized prefix in certain nouns will not be segmented either. The verb and noun base constitute the semantic head of the phrase insofar as the phrase they head is an example of a verbal or nominal construction. It is also important to highlight that adverbial clitics in the preverbal positions 6–8 will only occur if they have the right host (e.g. the negative marker) and otherwise they do not occur. Also, the second morpheme in monoclausal negation (= *di*) in position 18 does not occur without the negative proclitic *kēd=* in position 5.

3.2 Constituency tests

Following Tallman (2021), we assume that a constituency test is a generalization within or across constructions that targets or crucially refers to some subspan of a planar structure. The constituency tests considered are shown below. These are applied to both verbal (§4) and nominal (§5) domains.

(10) Morphosyntactic and indeterminate diagnostics

- Minimum free form
- Non-interruptability
- Subspan repetition in serialization
- Nonpermutability
- Deviations from biuniqueness
- Ciscategorical selection

Table 1: Planar structure for verb in Teotitlán del Valle Zapotec

| Position | Type | Elements |
|----------|------|---|
| (1) | slot | subordinators |
| (2) | slot | NP (A/s, O), ADV |
| (3) | slot | focus marker (=ēn) |
| (4) | slot | ADV á= ‘already’ |
| (5) | slot | clausal negator: <i>kēd</i> = |
| (6) | slot | adverbial enclitics of frequency and manner: = <i>pkā</i> , = <i>kā</i> , = <i>zī</i> |
| (7) | slot | adverbial enclitic of equality: = <i>zá</i> |
| (8) | slot | adverbial enclitic of comparison: = <i>rú</i> |
| (9) | slot | indefinite, interrogative pronouns <i>tū</i> =, <i>fī</i> =, <i>kālī</i> =; plural im- perative <i>gūl</i> = |
| (10) | slot | tense/aspect/mood |
| (11) | slot | motion: andative <i>e</i> -, venitive <i>ēd</i> - |
| (12) | slot | verb base |
| (13) | slot | second element in a compound: nominal, adjectival root |
| (14) | slot | comitative - <i>nā</i> : |
| (15) | slot | diminutive: - <i>æ’ny</i> /- <i>i’ny</i> |
| (16) | slot | manner adverbs: <i>ʒlyā</i> ‘in vain’, <i>dʒí</i> : ‘quietly’ |
| (17) | slot | intensifiers <i>tæ</i> :, <i>dā:n</i> |
| (18) | slot | negation: = <i>dī</i> |
| (19) | slot | adverbial enclitics of frequency and manner: = <i>pkā</i> , = <i>kā</i> , = <i>zī</i> |
| (20) | slot | adverbial enclitic of equality: = <i>zá</i> |
| (21) | slot | adverbial enclitic of comparison: = <i>rú</i> |
| (22) | slot | reciprocal <i>sa</i> ’ |
| (23) | slot | pronominal enclitic; NP (A/s) |
| (24) | slot | pronominal enclitic; NP (p) |
| (25) | slot | pronominal enclitic; NP (R) |
| (26) | slot | adverb (lexical) |
| (27) | slot | subordinate clause |
| (28) | slot | discursive enclitics |

Table 2: Planar structure for noun in Teotitlán del Valle Zapotec

| Position | Types | Elements |
|----------|-------|---|
| (1) | slot | prepositions: <i>ftê:n</i> ‘of’; <i>nez</i> ‘by’ |
| (2) | zone | quantifiers (QR): <i>zyē:n</i> ‘various’, <i>tubru</i> ‘some’, INDEF <i>te</i> = etc; relational nouns: <i>low</i> ‘face of’, <i>kwæ</i> ‘side of’ etc. |
| (3) | slot | adverbial enclitics of frequency and manner: <i>=pkā</i> , <i>=kā</i> , <i>=zī</i> |
| (4) | slot | adverbial enclitic of equality: <i>=zá</i> |
| (5) | slot | adverbial enclitic of comparison: <i>=rú</i> |
| (6) | slot | plural <i>d’</i> =; 1PL <i>dū</i> =; <i>dā</i> = ‘Mr.’; <i>tyú</i> = ‘uncle’ |
| (7) | slot | possessive: <i>f-</i> |
| (8) | slot | noun base |
| (9) | slot | second element of a compound |
| (10) | slot | adjective(s) (up to three) |
| (11) | slot | diminutive: <i>-æ’ny/-i’ny</i> |
| (12) | slot | intensifiers: <i>dâ:n</i> ; <i>tæ</i> : |
| (13) | slot | adverbial enclitics of frequency and manner: <i>=pkā</i> , <i>=kā</i> , <i>=zī</i> |
| (14) | slot | adverbial enclitic of equality: <i>=zá</i> |
| (15) | slot | adverbial enclitic of comparison: <i>=rú</i> |
| (16) | slot | demonstratives: <i>=kī</i> DEM.TEMP, <i>=kán</i> DEM.MED, <i>=ráé</i> DEM.PROX <i>=ræ</i> DEM.DIST |
| (17) | slot | prepositional phrase, pronominal enclitic, NP Possessor |
| (18) | slot | relative clause |
| (19) | slot | demonstrative: <i>=kī</i> DEM.TEMP <i>=kán</i> DEM.MED, <i>=ráé</i> DEM.PROX, <i>=ræ</i> DEM.DIST |
| (20) | slot | focus marker <i>=ēn</i> |

- (11) Phonological diagnostics
- Glottal Dissimilation
 - Accentuation
 - Syllabification
 - Rising Tone Levelling
 - Mid Tone Spreading
 - Tone Sandhi
 - Final Glottalization

An oft-neglected aspect of constituency and wordhood tests is that they can provide ambiguous results (Tallman 2020, 2021). The definition and results of constituency test often depends on what type of elements or constructions are being considered. Thus, wordhood or constituency tests, stated abstractly, can often be ambiguous with respect to which string they identify (see Osborne 2018 as well on this issue). Test fracturing refers to the practice of adding special conditions on constituency tests such that they provide discrete results. In doing so, a constituency test is fractured into more than one result. In the following section, we report all the tests and their respective fracturing to avoid being opportunistic.

4 Verbal domain

In this section we discuss the 21 tests applied to the verbal domain. We first discuss six morphosyntactic tests (fractured into ten). We then focus on the seven which are phonological (fractured into eleven).

4.1 Morphosyntactic diagnostics

4.1.1 Minimum free form (10-12, 4-22)

There are at least two ways of applying the minimum free form test depending on what constraints we impose on adding and subtracting bound elements on the positions flanking obligatory elements in a sentence. The ambiguity in test application is illustrated by the fact that the test can be fractured into at least two distinct interpretations depending on whether we consider the smallest possible span that is a single free form or the largest possible span that is a single free form. The distinction of interpretations is listed below.

- (12) Free occurrence (minimum): The free form that contains elements from positions with the shortest distance from each other with respect to positions in the planar structure.

- (13) Free occurrence (maximal): The free form that contains elements from positions with the largest distance from each other with respect to positions in the planar structure.

The smallest possible span that is contiguous on its edges with a minimal free form consists of just the verb and a TAM prefix,⁶ positions 10–12. This is shown by a complete utterance below. Neither the verb base, as shown in (15), nor the TAM prefix can occur as free forms (by themselves).

- (14) bi'zu:
bi-zu:
 v:10-12
 COMPL-tremble
 '(It) trembled.'

- (15) *'zu:⁷
zu:
 v:12
 'tremble'
 Intended reading: '(It) tremble(s).'

The maximal free occurrence span is from 4–22, as shown below. This span contains all those elements with the largest distance from one another and that cannot stand as a single free form. Note that the right edge of this span (position 23) corresponds to the position of the subject of the sentence; since this element can be a nominal (which can stand as a free form), it is not included within this span.

- (16) ákēdrú baká'næ:dí 'sa'dán
á=kēd==rú ba-kánæ:=di sa'='dán
 v:4=5==8 10-12-14 22=23
 already=NEG==more COMPL-fight=NEG RECP=3PL.INF
 'They didn't fight with each other anymore.'

⁶In some stative and potential forms, the TAM prefix is zero, e.g. *ø-zu:* 'STAT-stand' or *ø-dâ:* 'POT-pour'. However, in such cases we consider that there is still a zero prefix, rather than that there is no TAM prefix.

⁷Without the prefix, *zu:* means 'earthquake'; thus, it is grammatical, but not as a verb.

4.1.2 (Non-)interruptability (10-15, 4-22)

(Non-)interruptability identifies a span of positions that cannot be interrupted by some interrupting element. Traditionally the test has been articulated such that the interrupting element is a word (Bloomfield 1933), but this definition is obviously circular to the extent that non-interruptability is supposed to form the basis for identifying words at the onset (Mugdan 1993: 2552). One solution is to fix the definition of interrupting element based on some testable criterion. Haspelmath (2011) proposes that the interrupting element should be a free form. However, this choice is arbitrary and fails to capture the fact that interruptability (or conversely “contiguity”) is a matter of degree (Croft 2001: 190-191; Tallman 2018: 117-120). An approach that can be used to capture and report more fine-grained details in linguistic structure is to fracture (non-)interruptability into a number of subtests depending on the criterial wordhood properties of the interrupting element. For TdVZ, we apply two interruptability tests to the verbal domain.

- (17) Non-interruption1: A span that cannot be interrupted by an element that can occur in more than one position in the planar structure (e.g. second position clitic, a free function word or indefinite pronouns).
- (18) Non-interruption2: A span that cannot be interrupted by any free form (e.g. a noun phrase or interjection).

The non-interruption1 test identifies the span 10–15. This span is shown in (19). Determining the precise span which cannot be interrupted by an element that can occur in more than one position in the planar structure is achieved by intending to place a second position clitic right after the minimum free form (position 10–12) as shown below, in this case with the adverbial enclitic of equality, =*zá*, in position 20. As noted, this is not possible when the comitative -*næ* in position 14 or the diminutive -*i'ny* morpheme in position 15 occurs.

- (19) *resut'ně'nzán'lä:n*
r-e-sut-næ:-i'ny=zá=an *lä:n*
v:10-11-12-14-15=20=23 24
HAB-AND-AND:play-COM-DIM=**also**=3SG.INF 3SG.INF
‘He also goes to play with him/her (how nice!).’
- (20) **resutzá'ně'nyan'lä:n*
r-e-sut=zá-næ:-i'ny=an *lä:n*
v:10-11-12=20-14-15=23 24
HAB-AND-AND:play=**also**-COM-DIM=3SG.INF 3SG.INF
Intended reading: ‘He also goes to play with him/her (how nice!).’

On the other hand, the non-interruption² test identifies the span 4–22. That is, the only slot where an NP may occur is either in position 2 or 23. In (22)–(23), we show that when an NP is placed between any elements within this span, the construction is ungrammatical. This span cannot be interrupted by an interjection either. Position 3 is not included within this span since the occurrence of a morpheme in this position (i.e., the focus marker) requires the occurrence of the NP in position 2.

- (21) ákēdrú re'llē'wdí Jwá:yn 'la:dy
 á=kēd==rú r-e-llē'w=di Jwá:ny la:dy
 v:4=5==8 10-11-12=18 23 24
 already=NEG==more HAB-AND-rinse=NEG Juan clothes
 'Juan doesn't go to/and rinse the clothes anymore.'
- (22) *kēd' Jwá:yn re'llē'wdí 'la:dy
 kēd= Jwá:ny r-e-llē'w=di la:dy
 v:5= 23 10-11-12=18 24
 NEG= Juan HAB-AND-rinse=NEG clothes
 Intended reading: 'Juan doesn't go to/and rinse the clothes.'
- (23) *kēdre'llē'w Jwá:yndi 'la:dy
 kēd=r-e-llē'w Jwá:ny=di la:dy
 v:5=10-11-12 23=18 24
 NEG=HAB-AND-rinse Juan=NEG clothes
 Intended reading: 'Juan doesn't go to/and rinse the clothes.'

4.1.3 Subspan repetition in serialization (10-12, 10-23)

Subspan repetition in serialization refers to subspans of the verbal planar structure that are repeated because they cannot be interpreted unless they are present in the subspan itself. For TdVZ we consider the two most prototypical serialization constructions (Gutiérrez 2014): typical and motion serialization. The latter differs from the former in that the second verb in the construction gets more modifying morphemes (i.e., the second verb has a broader expansion). Therefore, in this section we are not only looking at one type of serialization but two. For each type we fracture the constituency test as follows.

- (24) Minimal repeated subspan: the subspan of positions whose elements cannot be interpreted unless they are present in the subspan itself.

- (25) Maximal (repeated) subspan: the subspan of positions whose elements can occur in each of the coordinated constituents without reference to whether some of these elements can be elided or interpreted via wide scope of one element over the repeated subspans.

Typical Serialization occurs with a verb that indicates an effect on the subject followed by a verb that introduces the cause of the effect, as in (26). The minimum span in typical serialization occurs in the same span identified by the minimum free form (minimal) test (§4.1.1): positions 10–12. In (27), note that the verb base (position 12) cannot occur by itself in this construction; thus, the TAM prefix of the first verb does not have scope over the second.⁸

- (26) *kēdtūbíḏḡi:bydi gu'nnā: 'lū:y*
kēd= tū= bi- ḡi:by =di gu- nnā: lu:y
 v:5= 9= 10- 12 =18 10- 12 24
 NEG= INDF.PRON= COMPL- be.afraid= NEG COMPL- witness 2SG.INF
 ‘Nobody was afraid of you.’ #Wide scope: ‘nobody was afraid, nobody saw you.’

- (27) **kēdtūbíḏḡi:bydi 'nnā: 'lū:y*
kēd= tū= bi- ḡi:by =di nnā: lu:y
 v:5= 9= 10- 12 =18 12 24
 NEG= INDF.PRON= COMPL- be.afraid =NEG witness 2SG.INF
 Intended reading: ‘Nobody was afraid of you.’

The verbs in typical serialization share the subject, which must be encoded on each verb. Thus, the maximal span that must be repeated in these constructions is 10–23, as shown in the example below.

- (28) *riḡi:by 'Yá:n ri'nnā:s 'mwǎ:s*
ri- ḡi:by Yá:n ri- nnā: =an mwǎ:s
 v:10- 12 23 10- 12 =23 24
 HAB- be.afraid Ana HAB- witness =3SG.INF teacher
 ‘Ana is afraid of the teacher.’

Serial verb construction (SVC) whose first verb in the construction is the verb *-æ:* ‘go’ or *-ǎ:d* ‘come’ differs from typical serialization since the second verb

⁸In fact, in typical serialization none of the modifying elements for one verb can have wide scope over the serialized (repeated) spans. In (26), the indefinite pronoun only has scope on the first verb of the construction.

in the construction can be modified by second position clitics (in Typical Serialization this triggers ungrammaticality). Therefore, we assume they are different constructions. In motion serialization, the minimum and maximum repeated subspan is the same as in typical serialization. That is, the minimal span is 10–12 while the maximum is 10–23, as shown in the example below. Given that both types of serialization cover the same spans, these are included only once in table and the figures that summarize the convergences in the verbal domain.

- (29) 'ræ:pkādán re'ti:ʒpkādán 'ʃtʃā'
 r- æ: =pkā =dán r- æ- ti:ʒ =pkā =dán ʃtʃā'
 v:10- 12 =19 =23 10- 11- 12 =19 =23 24
 HAB- go =always =3PL.INF HAB- AND- AND:pay =always =3PL.INF light
 'They always go to/and pay the electricity (bill).'

4.1.4 (Non-)permutability (10-15)

In TdVZ the constituency test of (non-)permutability identifies a span of positions whose elements cannot be permuted. That is, the ordering of elements is 'fixed' in this span. This test identifies the span 10–15, as in (30). This is the same span identified by the non-interruptability test (§4.1.2). This test is not fractured in TdVZ since variable affix ordering has not been attested in this language, as shown in (31).

- (30) resut'ně'nyan 'lǎ:n
 r- e- sut -nǎ: -i'ny =an lǎ:n
 v:10- 11- 12 -14 -15 =23 24
 HAB- AND- AND:play -COM -DIM =3SG.INF 3SG.INF
 'He goes to play with him/her (it is nice).'

- (31) *resutæ'nǎ:n lǎ:n
 r- e- sut -i'ny -nǎ: =an lǎ:n
 v:10- 11- 12 -15 -14 =23 24
 HAB- AND- AND:play -DIM -COM =3SG.INF 3SG.INF
 Intended reading: 'He goes to play with him/her (it is nice).'

Thus, besides not being permutable, elements from position 10–15 cannot occur in a different position in the template. However, elements outside this span may occur in different positions in the template. Although the elements outside of this span can occur in a different position in the template, they cannot occur in a random order; that is, these elements follow a strict sequence (wherever they

occur), as shown in the example below where 'ʒlyā' always precede =zī whether they occur post or preverbally.

- (32) ba'llē'w 'ʒlyá' zī Yá:n 'la:dy
 ba- llē'w ʒlyā' =zī Yá:n la:dy
 v:10- 12 16 =19 23 24
 COMPL- rinse in.vain =only Ana clothes
 'In vain Ana rinsed the clothes.'
- (33) 'ʒlyā' zī bá'llē'w Yá:n 'la:dy
 ʒlyā' =zī ba- llē'w Yá:n la:dy
 v:2 =6 10- 12 23 24
 in.vain =only COMPL- rinse Ana clothes
 'In vain Ana rinsed the clothes.'

4.1.5 Deviations from biuniqueness (10-13)

In Teotitlán Zapotec, all morphemes between the positions 10 – 13 manifest non-automatic allomorphy; that is, where the alternation is not due to phonological processes or phonologically conditioned. This defines the span of deviations from biuniqueness. This test is not further fractured since morphemes outside of this span do not show any non-automatic allomorphy.

Table 3: Non-automatic allomorphy across habitual, completive and potential forms

| gloss | HAB | COMPL | POT |
|-------------|-----------|-----------|----------|
| 'get lost' | ri-'dyu'n | bi-'dyu'n | Ø-'dyû'n |
| 'live' | ri-bā:yn | gu-bā:yn | Ø-bā:yn |
| 'shake off' | ri-bi:by | gu-bi:by | 'kwí:by |

First, the completive and potential prefixes that occur in position 10 manifest allomorphy, which has motivated verbal classification in Zapotecan linguistics (Kaufman 1989; Smith-Stark 2002; Campbell 2011; Pérez Báez & Kaufman 2016; Beam de Azcona 2019; among others). Such allomorphy is illustrated below. As seen below, the completive prefix can either be *bi-* or *gu-*, and the potential prefix can either be zero and or fortition of the stem-initial consonant, as illustrated in Table 3, both with a tonal effect on the stem. The distribution of such allomorphs cannot be predicted by the phonological environments.

Secondly, the motion prefixes in the position 11, namely andative *e-* and the venitive *ēd-* display suppletion conditioned by the agent persons. Compare the forms in (34) without the agent person and (35) with the 1st person plural agent, which has the suppletive allomorph *yóp-* for the venitive:

- (34) *rēd'tá:w*
r- ēd- ta:w
v:10- 11- 12
HAB- VEN- eat
‘Comes to eat.’
- (35) *ryóp'tō:n*
r- yóp- tō:w =un
v:10- 11- 12 -23
HAB- VEN:1PL- eat:1PL =1PL.INCL
‘We come to eat.’

The verb base in position 12 also displays suppletive allomorphy according to the agent persons or tense/aspect/mood. Thus, in the 1st person forms the agentive verbs undergo stem alternation (Uchihara & Gutiérrez 2020b). For instance, the verb ‘come’ undergoes suppletion according to the agent persons:

- (36) *'rǣ:d*
r- ǣ:d
v:10- 12
HAB- come
‘(S/he) comes.’
- (37) *'rǣllá*
r- ǣll =a
v:10- 12 =23
HAB- come:1SG =1SG
‘I come.’
- (38) *'ryópún*
r- yóp =un
v:10- 12 =23
HAB- come:1PL =1PL.INCL
‘We come.’

Some verbs undergo suppletion, weak or strong, according to tense/aspect/mood categories. Thus, the verb *akw* ‘put on shirt’ undergoes suppletion in the completive aspect:

- (39) 'rakw
 r- *akw*
 v:10- 12
 HAB- put.on.shirt
 'puts on shirt'
- (40) 'gut
 gu- *Vt*
 v:10- 12
 COMPL- COMPL:put.on.shirt
 'put on shirt'

A noun root in position 13 can also undergo alternation (mostly tonal) according to the agent person. Thus, in (42), the incorporated noun root *dya:g* 'ear' undergoes tonal alternation, in addition to the tonal alternation of the verb base *-kwā:* 'throw':

- (41) rukwa'dya:g
 ru- *kwā:* *+dya:g*
 v:10- 12 +13
 HAB- throw +ear
 'listens'
- (42) rukwá'dyā:gá
 ru- *kwâ:* *+dyǎ:g* =*a*
 v:10- 12 +13 =23
 HAB throw:1SG +ear:1SG =1SG
 'I listen.'

Morphemes in the positions outside of the span 10–13 do not show non-automatic allomorphy; they do manifest alternations, but all of such cases are phonologically conditioned.

4.1.6 Ciscategorial selection (10–12, 4–14)

All of the morphemes in positions 10–12 are unique to the verbs; it is ungrammatical to attach the tense/aspect/modal prefix or the motion prefix to any parts of speech other than verbs. Thus, this defines the minimal span of ciscategorial selection.

In addition, the morphemes in positions 4 (*á=* 'already') and 14 (*-nǣ:* comitative) are also unique to verbs, although there are morphemes within this span

that are not unique to verbs, such as adverbial enclitics in positions 6–8 (as can be seen in the nominal planar structure in Table 2) and the incorporated noun root in position 13. This is illustrated in the following examples. In (43), the adverbial clitic =*zī* ‘only’ in position 6 occurs with a noun. In (44), the compounded noun root in position 9 in the nominal planar structure (Table 2) corresponds to the incorporated noun root in position 13 in the verbal planar structure.

- (43) *ʃkæ̃tzi Jwá:yn*
ʃ- gæ̃t =zī ʃwá:ny
 n:7- 8 =13 17
 POSS- tortilla =only Juan
 ‘Only / just Juan’s tortilla(s)’

- (44) *bæll'yu:*
bæll +yu:
 n:8 +9
 snake +soil
 ‘worm’

Thus, the span 4–14 defines the maximal domain of ciscategorial selection; in other words, no morpheme outside of this domain is unique to verbs.

4.2 Phonological diagnostics

This section reviews the phonological constituents that could be supported by processes that change the segmental forms of elements in the subspan.

4.2.1 Glottal dissimilation (3-15)

TdVZ has a glottal dissimilation rule such that a glottalized syllable is deglottalized when followed by another glottalized syllable (CV'CV' > CVCV'). A similar process is reported for the Miahuatec variety of Zapotec (Hernández Luna 2021). Glottal Dissimilation is illustrated in the following example, where the glottalized vowel in the verb root -*ta'w* loses its glottalization before the diminutive suffix -*æ'n* which also has a glottalized vowel.

- (45) *ruta'wæ'nānbā*
ru- ta'w -æ'ny =ān =bā
 v:10- 12 -15 =23 =28
 HAB- sell -DIM =3SG.FOR =then
 ‘So s/he sells (it is good!).’

It is not the case that Glottal Dissimilation is always observed between any sequences of two adjacent syllables with glottalized vowels. Thus, between the verb root *gāʔ* ‘lay’ in position 12 and the compounded noun root *kwæʔ* in the example below no Glottal Dissimilation is observed.

- (46) *ragāʔkwæʔn*
ra- gāʔ-kwæʔ =ān
 v:10- 12 -13 =23
 HAB- lay -side =3SG.INF
 ‘S/he lays on his/her side.’

The other positions between the span of positions 3–15 do not have any morphemes with a glottalized vowel to see if this process is applied or not.⁹

A morpheme in position 16, *ʒlyāʔ*, has a glottalized vowel but it does not undergo Glottal Dissimilation as shown in the following example; here, the glottalized vowel in the verb base in position 12 and *ʒlyāʔ* in position 16 are adjacent, but Glottal Dissimilation is not observed. Thus, position 16 is outside of the domain of Glottal Dissimilation.

- (47) *baʔtaʔwʔʒlyāʔtæ:n gûʔn*
ba- taʔwʔʒlyāʔ tæ: =an gûʔn
 v:10- 12 16 17 =23 24
 COMPL- sell in.vain INTENS =3SG.INF bull
 ‘S/he sold the bull very in vain.’

Before the verb base, a morpheme with a glottalized vowel is not attested between positions 3 to 11. However, some morphemes with a glottalized vowel are attested in position 2, and such morphemes do not undergo Glottal Dissimilation, as shown in the following example. Here, the syllables ‘*lāʔ*’ in position 2 and ‘*dæʔ*’ in position 12 are adjacent but glottalization is kept in both syllables. Taken together, the domain of Glottal Dissimilation is 3–15; that is, Glottal Dissimilation may (but not always) apply within this domain, but it is never applied outside of this domain.¹⁰

⁹This means that in the verbal domain Glottal Dissimilation is observed only between the verb base in position 12 and the diminutive suffix in position 14, which make it seem like a process specific to the diminutive suffix. However, as we will see in §5.2.1, this process applies between the other morphemes in the nominal domain, and we consider that this process is general enough to be included as a phonological diagnostic.

¹⁰Since positions immediately preceding and following the verb base in position 12, namely positions 11 and 13, do not have any morpheme with a glottalized vowel, we cannot know if any of these positions constitute the minimal domain of Glottal Dissimilation.

- (48) 'lâ' 'dæ' nna'dzi:
 lâ' Ø- dæ' nna'dzi
 v:2 10- 12 26
 leucaena POT- be.picked today
 'Leucaena (*guaje*) will be picked today.'

4.2.2 Accentuation (10-15, 3-15)

In the verbal domain, the accent (or prominence) is assigned to the last syllable of the span 10–15, which is the minimal domain of Accentuation.¹¹ The syllables to which the morphemes in these positions belong to may bear an accent. First, when the prefix and the root form one single syllable, as in (49), all of which are in the position 10–12, the prominence is assigned to this single syllable of the root.

- (49) 'gâ:
 Ø'- ga
 v:10- 12
 POT- get.stretched
 '(It) will get stretched.'

When the prefix (in the position 10) constitutes its own syllable and the root another (position 12), the prominence is still assigned to the root syllable, as shown below.

- (50) ri'za:
 ri- za:
 v:10- 12
 HAB- walk
 '(S/he) walks.'

¹¹The prominent syllable in TdVZ is the position of more phonological contrasts, and segmental and suprasegmental contrasts are neutralized in the non-prominent syllables (Smith-Stark 2003: 25, 32; Chávez-Peón 2015). Segmentally, the marginal contrast between *e* and *æ* is generally neutralized to *e* in many non-prominent syllables. Vowel duration contrast is also neutralized to a short vowel in non-prominent positions. All the prefixes and clitics have a short vowel; a long vowel is shortened when it loses its prominence due to suffixation or compounding. Also, in non-prominent positions, the contrast between modal and creaky vowels is often neutralized to a modal vowel. Thus, all the prefixes have a modal vowel, and enclitics can only contrast modal and non-modal phonation types, while a prominent syllable can contrast modal, creaky and glottalized vowels, as mentioned above. Lastly, only level tones can occur on non-prominent syllables typically. Thus, no prefix or clitic has a contour tone. In non-prominent syllables, a rising and a falling tone neutralizes with a high tone (along with the neutralization of the vowel duration and phonation contrasts).

In compounds (within the positions 10–13), the prominence is assigned to the last root of the compound, as in the following: the preceding syllables (the verb base in position 12) are not assigned prominence.

- (51) rinnyab'di:ɕ
 ri- nnyɔ:b +di:ɕ
 v:10- 12 +13
 HAB- ask.for +word
 ‘(S/he) asks (a question).’

When the verb base and a suffix in the positions 14 (comitative) and 15 (diminutive) constitute independent syllables, the prominence is assigned to the syllable of the suffix while the root preceding these suffixes do not have prominence, as shown in the following examples.

- (52) rusēd'næ:
 ru- sæ:d -næ:
 v:10- 12 -14
 HAB- practice -COM
 ‘(S/he) studies with.’

- (53) rigi'tæ'n
 ri- git -æ'ny
 v:10- 12 -15
 HAB- play -DIM
 ‘(S/he) plays (nicely).’

The morphemes outside of this domain are not within the domain of Accentuation that includes the verb base in position 12. Thus, in (54) the prominence is assigned to the syllable of the comitative suffix in the position 12, and not to the syllable of the adverbial enclitic =pká in position 19.

- (54) row'næ:pkán'lā:n
 r- a:w -næ: =pká =(a)n lā:n
 v:10- 12 -14 =19 =23 24
 HAB- eat -COM =always =3SG.INF 3SG.FOR
 ‘He always eats with him.’

Outside of the minimal span of Accentuation, namely 10–15, morphemes may or may not have their own prominence. When they do not have their own prominence, they never bear prominence, unlike the morphemes in positions 10–15

which can bear prominence. Thus, in the preverbal positions, the morphemes in position 2 have their own prominence, while others do not.¹² For instance, in the following example, the interrogative *ʃá=* in position 9 does not have its own prominence, and this morpheme never bears prominence.

- (55) *ʃába'ki:nyan 'bæll?*
ʃá= ba- ki:ny =an bæll
v:9= 10- 12 =23 24
 how= COMPL- consume =3SG.INF fish
 'How/in which way did s/he eat (the) fish?'

In the postverbal positions, the morphemes in positions 16, 17, 22, 26, and 27 have their own prominence, while the morphemes in positions 23–25 may have their own prominence. For instance, in the following example, the morpheme *ʒlyá'* 'in vain' in position 16 has its own prominence, followed by morphemes in positions 19 (=zī) and 23 (=an) which do not have their own prominence, again followed by the 2SG pronoun *lu:y* which has its own prominence.

- (56) *gunni'nā:ʒlyá'zyán 'lu:y*
gu- nni: -nā: ʒlyá' =zī =an lu:y
v:10- 12 -14 16 =19 =23 24
 COMPL- say -COM in.vain =only =3SG.INF 2SG.INF
 'S/he spoke with you in vain.'

The maximal domain of accentuation can thus be defined as the span of positions 3–15; there is only one prominence within this span, and outside of this span there *can* be morphemes with their own prominence.

4.2.3 Syllabification (10-12, 1-28)

In TdVZ, a canonical syllable structure is CV(:)(C). The onset is obligatory except for very few native words (*i:z* 'year') and more recent versions of loanwords (*á:n* 'Ana'); in older loans, onset is inserted when the source form has no onset: *gú:r* 'hour', *yá:n* 'Ana'. Unlike other Central Zapotec varieties (Chávez-Peón 2010: 13-16), onset clusters are not common in Teotitlán Zapotec and mostly restricted to the sequences of a consonant + *y* (*gyæ:* 'flower'), a nasal + a lenis consonant

¹²Unless the adverbial enclitics attach to the negative proclitic, in which case they optionally acquire accent, as in *kēd=* NEG + *=pká* 'always' → *'kē:dpká*. This is possibly due to the requirement that the combination of a proclitic and the adverbial clitic constitute a prosodic word.

(*nga* ‘purple/blue’, *ndɔ:w* ‘amarillo dish’) or a sibilant + a consonant (*ʃtyé:ʒy* ‘garlic’, *stú:y* ‘another/ once more’). Any consonant may occur in the coda position, and coda clusters are uncommon except for a consonant + *y* sequences (*jālyy* ‘twenty’).

Segments are resyllabified in the sequence of a motion prefix (position 11) + the base, as in (57), and of a (TAM) prefix (position 10) + the base (position 12), as in (58); this (positions 10–12) defines the minimal span of syllabification. The syllable boundary is indicated with a dot in this subsection.

- (57) *rē.dyú:n*
r- ēd- yū:n
 v:10- 11- 12
 HAB- VEN- cry
 ‘comes to cry’

- (58) *ʼra:w*
r- a:w
 v:10- 12
 HAB- eat
 ‘eat’

The morphemes beyond this span may or may not participate in syllabification. Thus, syllabification is applied in the sequence a base + (diminutive) suffix (position 15) as in (59). However, positions between the base (position 12) and diminutive (position 15), namely the positions 13 (compounded root) or 14 (comitative), do not contain any morpheme that begins with a vowel. Thus, we cannot tell if syllabification applies or not between the verb base in position 12 and these positions.

- (59) *gu.zu.tæʼn*
gu- zut -æʼny
 v:10- 12 -15
 COMPL- COMPL:play DIM
 ‘(S/he) played nicely.’

Segments are also syllabified in the sequences of the verb base + certain enclitics after position 19. The following is an example with a base (position 12) ending with a consonant and a pronominal enclitic (position 23) beginning with a vowel; here, the final consonant of the root *d* is syllabified as the onset of the following syllable:

- (60) 'bǣ:.dán
 b- ǣ:d =án
 v:10- 12 =23
 COMPL- come =3SG.INF
 'S/he came.'

However, when the positions 23 or 24 is occupied by a noun phrase, resyllabification does not take place, as in the following example. Here, the final consonant *d* of the first word is not resyllabified as the onset consonant of the syllable *Á:n*. This different behavior of a bound vs free morpheme that occupy the same position in the planar structure is a recurrent issue in TdVZ (cf. §5.2.4, §5.2.5). This could be resolved by fracturing the test according to whether these positions are occupied by a noun phrase or a bound morpheme, but this is not done in this chapter for the sake of space.

- (61) 'bǣ:d 'Á:n
 b- ǣ:d á:n
 v:10- 12 23
 COMPL- come Ana
 'Ana came.'

The following is an example with a stem ending with a vowel and a final clitic (position 28) which consists solely in a consonant; here, the final clitic =*f* is syllabified as the coda of the syllable of the host.

- (62) 'rǣ:f
 r- ǣ: =f
 v:10- 12 =28
 HAB- go =then
 '(S/he) goes, then.'

On the other hand, not all enclitics appear to be within the domain of syllabification. Thus, adverbial enclitics (position 19) which begin with a consonant cluster (= *pkā* 'always'; = *ʒgá* 'first') do not resyllabify with the preceding V-final root, either.

- (63) 'rǣ:.pkā
 r- ǣ: =pkā
 v:10- 12 =19
 HAB- go =always
 '(S/he) always goes.'

Thus, any morphemes within the whole span of 1–28 *may* participate in syllabification, but not all the positions contain morphemes which would allow us to judge if syllabification is applied or not. Therefore, we cannot know whether syllabification is applied. The first morpheme has to end in a consonant and the second morpheme has to begin with a vowel (or a glide) to see if syllabification is applied (or if the morpheme only consists in a consonant or begins with a consonant cluster, we could see if syllabification applies or not, as we saw above). Since the verb stem that minimally consists of the tense/aspect/modal prefix in position 10 and the verb base in position 12 always begins with a consonant, we cannot tell if syllabification applies between the morphemes in positions before 10. On the other hand, after the verb base in position 12, only positions 15, 23, and 24 may contain a morpheme beginning with a vowel, and we have seen above that syllabification may apply between these positions. Thus, the maximal domain of syllabification is the span of positions 1–28.

4.2.4 Rising tone levelling (10-24)

Rising Tone Levelling is a tonal process in which a rising tone (which we analyze as a sequence of a mid tone and a high tone) is split into a mid tone on one syllable and a high tone (or a falling tone, when this syllable is prominent lexically has a low tone) on the next syllable with a low or mid tone, as it is illustrated in Figure 1 (a similar process is reported in the Miahuattec variety of Zapotec; Hernández Luna 2021).

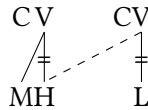


Figure 1: Rising Tone Levelling

In order for Rising Tone Levelling to apply, a few structural requirements have to be met. First, the first syllable has to have a lexical rising tone. Secondly, the rising tone needs to be in a syllable which is syllabified as an open syllable; if the following morpheme begins with a consonant, Rising Tone Levelling does not take place. Since positions immediately preceding and following the verb base in position 12, namely positions 11 and 13, do not contain any elements that meet these structural requirements, we cannot identify the minimal span of this test, and thus we do not fracture this test.

The span whose positions contain elements that display positive evidence for Rising Tone Levelling covers positions 10 to 24. Rising Tone Levelling is observed between the tense/aspect/modal prefix in position 10 and the verb base in position 12 as in the following. Here, the rising tone on the potential prefix splits into a mid tone (which is indistinguishable with a low tone in an atonic position) and a high tone on the following syllable:

- (64) gũ'tǎé'
 gũ- tǎé'
 v:10- 12
 POT- gather
 'will gather'

If the verb base has a rising tone, it can undergo Rising Tone Levelling when the following morpheme belongs to position 15 (diminutive) or positions 23 and 24, only if they are occupied by pronominal enclitics and not independent NPs. This is illustrated in the following examples. First, in (65) we show the application of Rising Tone Levelling between the verb base in position 12 and the diminutive suffix in position 15.

- (65) gudĩ'bǎé'nan
 gu- dǐ:b -ǎé'ny =an
 v:10- 12 -15 =23
 COMPL- COMPL:sew -DIM =3SG.INF
 'S/he sewed (how nice!).'

The following example illustrates the application of Rising Tone Levelling between the verb base in position 12 and a pronominal enclitic in position 23:

- (66) ri'gĩ:bú
 ri- gĩ:b =u
 v:10- 12 =23
 HAB- sew =2SG.INF
 'S/he sews.'

When position 23 or 24 is occupied by an independent NP, Rising Tone Levelling is not applied, even if other structural requirements are met. This is illustrated in the following example. Here, the first syllable has a rising tone, and the morpheme *i:z* begins with a vowel, but the process is not applied. Again, this different behavior could be captured by fracturing the test.

- (67) 'gāk i:z *gāk í:z
 g' - ak i:z
 v:10- 12 23
 POT- become year
 'will be a (new) year'

Outside of the 10–24 span, Rising Tone Levelling is not observed, because such positions do not contain any morpheme that satisfies the structure requirement for this process to apply. Thus, in the preverbal position, no morpheme with a rising tone is attested except for position 2. Even when position 2 is occupied by a morpheme that has a rising tone, such as *mǎ:yn* ‘animal’, the verb stem (that is, the minimal combination of a tense/aspect/modal prefix in position 10 and the verb base in position 12) always begins with a consonant. Thus, we cannot tell if Rising Tone Levelling would be applied or not. In the postverbal position, only positions 23 or 24 may have a morpheme which begins with a vowel, and in such cases Rising Tone Levelling is not applied, as we saw above.

4.2.5 Mid tone spreading (11-14, 1-28)

Mid Tone Spreading is a tone spreading process where a mid tone spreads to the preceding syllable when its lexical tone is low, cf. Figure 2.

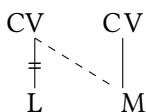


Figure 2: Mid Tone Spreading

This process can be illustrated by the following examples. The verb bases of the forms in the following are a minimal pair in terms of low vs. mid tone, *-zæ:by* ‘get hanged’ and *-zæ̃:by* ‘sink’. This tonal contrast is neutralized when the 3sg formal enclitic =*ān*, with a mid tone, follows and then its mid tone spreads to the root:

- (68) a. ri'zæ:byān
 ri- zæ:by =ān
 v:10- 12 =23
 HAB- get.hanged =3SG.FOR
 ‘He hangs (off/on a tree).’

- b. ri'zæ:byān
 ri- zæ:by =ān
 v:10- 12 =23
 HAB- sink =3SG.FOR
 'He sinks'.

In order for this process to be applied, we need a sequence of low tone on one syllable and a mid tone on the next. Mid Tone Spreading is known to apply between the andative prefix in position 11 and the verb base in position 12 as in (69); between the verb base and the compounded root in position 13 as in (70), and between the verb base and the comitative suffix in position 14 as in (71). This defines the minimal domain of Mid Tone Spreading (positions 11–14), which is the shortest span which includes positions, all of which manifest positive evidence.

- (69) rē'gāf
 r- e- gaf
 v:10- 11- 12
 HAB- AND- pull.out
 'goes to pull out.'

- (70) r-yēpy'fġā'
 r- yepy +fġā'
 v:10- 12 +13
 HAB- go.up +light
 'has a chill.'

- (71) rusæd'næ:
 ru- sæ:d -næ
 v:10- 12 -14
 HAB- study -COM
 '(S/he) studies with.'

Outside of this domain, in the whole span of positions 1–28, Mid Tone Spreading may apply, but positive evidence is not always available, since not all the positions have a morpheme that would satisfy the structural requirement for this process to apply. Mid Tone Spreading applies in the sequences of a subordinator in position 1 + a base (position 12) as in (72), of a base + an adverbial clitic in position 18 as in (73), of a base + a pronominal clitic in position 23 as in (74), and of a base + a discursive clitic in position 28, as in (75).

- (72) ʃi'ræ:d
ʃi= r- æ:d
v:1= 10- 12
when= HAB- come
'When (s/he) comes.'
- (73) ri'zæ:byi
ri- zæ:by =zi
v:10- 12 =19
HAB- get.hanged =only
'(It) just gets hanged (without motives).'
- (74) rā:wān
r- a:w =ān
v:10- 12 =23
HAB- eat =3SG.FOR
'S/he (formal) eats.'
- (75) ba'llū:bbā
ba- llū:b =bā
v:10- 12 =28
IMP- sweep =then
'(Go ahead and) sweep, then.'

On the other hand, Mid Tone Spreading is not observed between independent phonological words. First, when positions 23 or 24 are occupied by an independent NP, instead of a pronominal enclitic, Mid Tone Spreading is not applied between the verb base, as shown below:

- (76) ri'gats bēnny *rigāts bēnny
ri- gats bēnny
v:10- 12 23
HAB- be.buried person
'People are buried.'

Similarly, Mid Tone Spreading is not observed between the verb base in position 12 and a free function word in position 16, which have their own prominence (that is, they constitute their own prosodic words).

- (77) gu'do:w 'ʒlyā' bækw
 gu- do:w ʒlyā' bækw
 v:10- 12 16 23
 COMPL- COMPL:eat in.vain dog
 '(The) dog ate in vain.'

Again, this difference between free vs bound forms could be captured by fracturing the test.

4.2.6 Tone sandhi (11-19, 1-28)

Tone Sandhi is a process whereby a mid tone (and one class of high tone which is derived from a mid tone) assigns a falling or high tone to the following syllable which lexically has a low or mid tone, cf. Figure 3. This is because the mid tone in Teotitlán Zapotec is always associated with a floating high tone, since historically it comes from a rising tone (as in Quiaviní Zapotec, cf. Uchihara 2016).

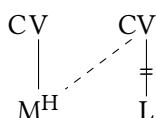


Figure 3: Tone Sandhi

The minimal pair in the following illustrates this process; in (78a), the verb base -zæ:by 'get hanged' has a low tone and does not trigger any tone change on the following vowel of the 2SG.INF enclitic, =u. In (78b), on the other hand, the verb root -zæ:by 'fall into' has a mid tone, and thus the following vowel of the 2SG.INF enclitic is assigned a high tone.

- (78) a. ri'zæ:byu
 ri- zæ:by =u
 v:10- 12 =23
 HAB- get.hung =2SG.INF
 'you hang (from somewhere/something)'
 b. ri'zæ:byú
 ri- zæ:by =u
 v:10- 12 =23
 HAB- sink =2SG.INF
 'you sink (into a hole)'

In the preverbal position, Tone Sandhi applies between the motion prefix in position 11 and the verb base in position 12, as in (79). In the postverbal position, Tone Sandhi is known to apply between the verb base and the compounded root in position 13 as in (80); comitative in position 14 as in (81); diminutive in position 15 as in (82); manner adverbs in position 16 as well as adverbs of frequency and manner in position 19 as in (83); intensifiers in position 17 as in (84), and the negative enclitic =*di* in position 18 as in (85). This domain (positions 11- 19) defines the smallest span of Tone Sandhi.

- (79) *rēd'gáty*
r- ēd- gaty
 v:10- 11- 12
 HAB- VEN- VEN:die
 'comes to die'
- (80) *ribē'lâ:*
ri- bǣ: +la:
 v:10- 12 +13
 HAB- take.out +name
 'names in (an)other way'
- (81) *rusyā'nǣ:*
ru- syā: -nǣ:
 v:10- 12 -14
 HAB- clean -COM
 'cleans with'
- (82) *gú'nǣ'n*
g- ' ū:n -ǣ'ny
 v:10- 12 -15
 POT- cry -DIM
 'will cry (a little)'
- (83) *ba'fūll 'zlyá'zyán 'be:dy*
ba- fūll zlyá' =zī =an be:dy
 v:10- 12 16 =19 23 24
 COMPL- peel in.vain =only =3SG.INF chicken
 'He plucked the chicken in vain.'

- (84) ba'ʒū:ʒ tæ:
 ba- ʒū:ʒ tæ:
 v:10- 12 17
 COMPL- fray INTENS
 'frayed/scratched (something) a lot'

- (85) kēdbæ:ddí
 kēd= b- æ:d =di
 v:5= 10- 12 -18
 NEG= COMPL- come =NEG
 'did not come'

Outside the span of 11–19, Tone Sandhi may apply anywhere within the whole verbal plan structure, namely positions 1–28. The example in (78b) above illustrates the application of Tone Sandhi between the verb base in position 12 and a pronominal enclitic in the position 23.

If the sequence is found between independent spans, Tone Sandhi is not observed. In the following example, the final syllable of the first utterance ends in a mid tone (*pān*), which would assign a high tone to the following syllable. However, no tone change is observed in the following syllable with a low tone, *zit*, since these syllables belong to different planar structures.

- (86) 'zit'tæ: 'mē:dy 'gūpān, 'zit'tæ: 'mē:dy 'gūpān
 zit tæ: mē:dy gu-(ā)p=ān zit tæ: mē:dy
 v:2 10-12=23 2
 much INTENS money COMPL-have=3SG.FOR much INTENS money
 gu-(ā)p=ān
 10-12=23
 COMPL-have=3SG.FOR
 'He had a lot of money, he had a lot of money!'

4.2.7 Final glottalization (1-28)

Finally, Final Glottalization is a process whereby vowel-final atonic syllables with low or high tone are glottalized at the final position of the whole span (1 - 28); elsewhere, the glottalization is not found. A similar process is reported for Southern Zapotec (Beam de Azcona 2004; Hernández Luna 2021). In (87), the # represents the juncture of the positions 28 and 1.

(87) Final Glottalization

- a. =V_(with low or high tone) → =V^ʔ / _]#
 b. =V / elsewhere

The following examples illustrate Final Glottalization. In (88), the last syllable which the 1SG enclitic =a belongs to is at the final position and thus is accompanied by a glottalization. In (89), on the other hand, the same 1SG enclitic is within the whole span and thus Final Glottalization is not applied:

(88) 'nistrú ri'kā:zá'

nis =*rú* *ri-* *kā:z* =*a*
 v:2 =8 10- 12 =23
 water =more HAB- want:1SG =1SG
 'I want more water.'

(89) ri'kā:zá 'gâ: 'kyæ:

ri- *kā:z* =*a* *ø'* *ga:* *kyæ:*
 v:10- 12 =23 27 -
 HAB- want:1SG =1SG POT- trim head.of:1SG
 'I want to get a haircut.'

4.3 Coincidence and convergence in the verbal domain

In this section, we examine all the tests applied to the verbal domain. Below we show a summary of the convergence of these tests.

Table 4: Tests and convergence in the verbal domain in TdVZ

| Test | Left | Right | Size | Conv. | |
|-----------------------------------|------|-------|------|-------|---|
| Syllabification (minimal) | 10 | 12 | 3 | 4 | The shortest span of positions that contains morphemes that are known to undergo resyllabification. |
| Serialization 1 (minimal) | 10 | 12 | 3 | 4 | The smallest span that must be repeated in typical serialization. |
| Minimum free form (minimal) | 10 | 12 | 3 | 4 | The smallest possible span that is contiguous on its edges with a minimal free form. |

1 *Words as emergent constituents in Teotitlán del Valle Zapotec*

| | | | | | |
|-------------------------------------|----|----|----|---|---|
| Ciscategorical selection (minimal) | 10 | 12 | 3 | 4 | The contiguous span that contains elements that are unique to verbs. |
| Deviation from biuniqueness | 10 | 13 | 4 | 1 | A span where non-automatic allomorphy is observed. |
| Mid Tone Spreading (minimal) | 11 | 14 | 4 | 1 | The contiguous span which contains the morphemes which show positive evidence for Mid Tone Spreading. |
| Accentuation (minimal) | 10 | 15 | 6 | 3 | The contiguous span which is attested to be assigned prominence, which includes the verb base. |
| Non-permutability | 10 | 15 | 6 | 3 | Elements in this span cannot be permuted or variably ordered. |
| Non-interruption1 (movable element) | 10 | 15 | 6 | 3 | Elements in this span cannot be interrupted by an element that can occur in more than one position in the planar structure (e.g., a 2 nd position clitic). |
| Tone Sandhi (minimal) | 11 | 19 | 9 | 1 | The contiguous span which contains elements which show positive evidence for Tone Sandhi. |
| Ciscategorical selection (maximal) | 4 | 14 | 11 | 1 | Outside of this span no element is unique to verbs. |
| Glottal Dissimilation | 3 | 15 | 13 | 2 | The span of positions which contain elements which display no evidence against Glottal Dissimilation. |
| Accentuation (maximal) | 3 | 15 | 13 | 2 | Within this domain, there can only be one prominent syllable. |
| Serialization1 (maximal) | 10 | 23 | 14 | 1 | The largest span that can be repeated in typical serialization. |
| Rising Tone Levelling | 10 | 24 | 15 | 1 | The span whose positions contain elements that display positive evidence for Rising Tone Levelling |

| | | | | | |
|---------------------------------|---|----|----|---|--|
| Non-interruption2 (by an NP) | 4 | 22 | 19 | 2 | Elements in this span cannot be interrupted by a free form (by an NP) |
| Minimum free form (maximal) | 4 | 22 | 19 | 2 | A span that is contiguous on its edges with the minimal free form that contains elements with the largest difference from one another. |
| Syllabification (maximal) | 1 | 28 | 28 | 4 | The maximal span where elements of adjacent positions may resyllabify |
| Tone Sandhi (maximal) | 1 | 28 | 28 | 4 | The maximal span where Tone Sandhi may apply. |
| Mid Tone Spreading (maximal) | 1 | 28 | 28 | 4 | The span that contains elements where Mid Tone Spreading may apply. |
| Final Glottalization | 1 | 28 | 28 | 4 | The span where the final atonic vowel-final syllable is glottalized. |

Figure 4 provides an overview of the results of the constituency variables applied to TdVZ verbs in terms of layers.¹³ Assuming that words are areas of convergence, the best candidates for the word are the 10–12 span (4 diagnostics) and the 1–28 span (4 diagnostics).

The following Figures 5 and 6 show an overview of the results of the morphosyntactic and phonological constituency variables applied to TdVZ in terms of layers, respectively.¹⁴

As we can observe, according to the words = convergence/clustering assumption, a morphosyntactic word is the 10 - 12 span (4 morphosyntactic diagnostics) and the best candidate for the phonological word is the 1 - 28 span (4 phonological diagnostics). Both categories appear to be motivated in TdVZ verbal domain, but they cover spans that are not very intuitive for linguists. In other words, TdVZ contains very small (maximum 3 morphs) morphosyntactic words and very large (covering the sentence) phonological words. One cannot really get a more radical misalignment, but Post (2009) obtained a similar result in Galo (a Sino-Tibetan

¹³These figures are from Sandra Auderset and Adam Tallman.

¹⁴Note that for the purposes of this chapter we interpret ‘indeterminate’ domains as morphosyntactic tests.

1 Words as emergent constituents in Teotitlán del Valle Zapotec

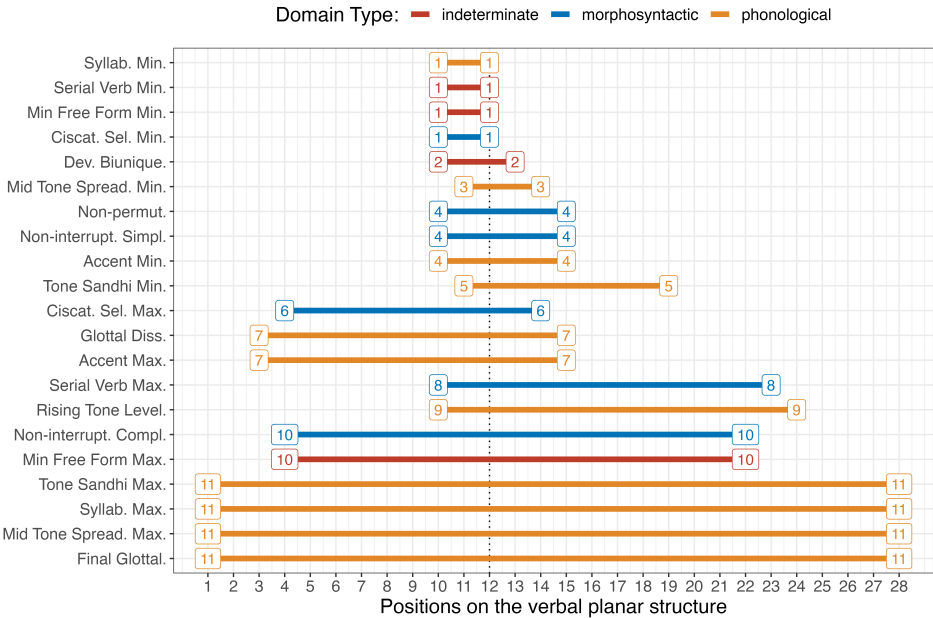


Figure 4: Constituency domains organized by converging layers in TdVZ verbs

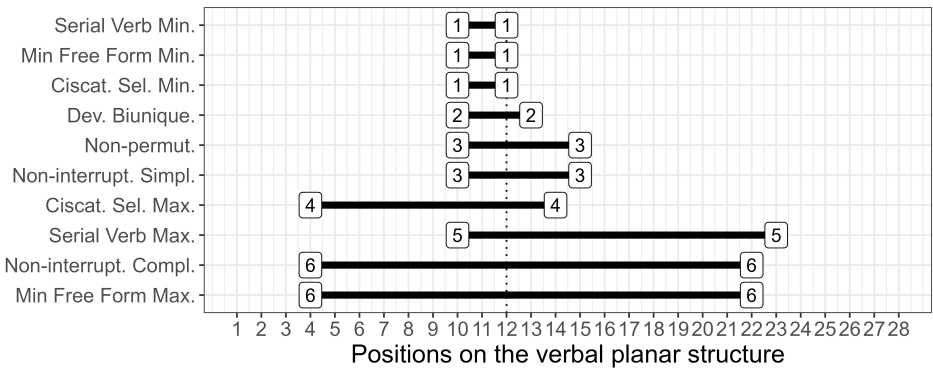


Figure 5: Morphosyntactic and indeterminate constituency domains organized by converging layers in TdVZ verbs

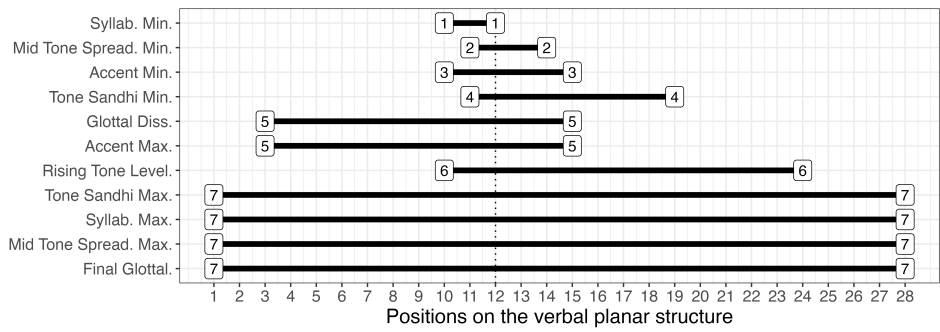


Figure 6: Phonological constituency domains organized by converging layers in TdVZ verbs

Language). One solution to these results could be to drop the convergence criteria for wordhood from consideration and assume that phonological words are just those phonological domains that are closest to morphosyntactic words.

5 Nominal domain

In this section, we discuss the tests applied to the nominal domain. Most of them are similar to those shown above for the verbal domain, with some important differences.

5.1 Morphosyntactic diagnostics

5.1.1 Minimum free form test (8-8, 6-9)

Just as in the verbal domain, the free occurrence test for the nominal domain is fractured in two. Below we explain this fracturing.

- (90) Minimum free form (minimum): The free form that contains elements from positions with the shortest distance from each other with respect to positions in the planar structure.
- (91) Minimum free form (maximal): The free form that contains elements from positions with the largest distance from each other with respect to positions in the planar structure. If one adds morphemes outside of this span, the sequence no longer constitutes a single free-standing form.

The smallest possible span that is contiguous on its edges with a minimal free form consists of just the nominal base: position 8. This is shown below.

- (92) 'nis 'water'

nis

n:8

'water'

- (93) 'gi: 'fire'

gi:

n:8

'fire'

The maximal free occurrence is the span containing the positions 6–9, as shown below. That is, the only bound morpheme that can be added on the right side without generating two free forms is the second element of a compound in position 9; if a (free form) adjective in position 10 is added, we will have two independent free forms. On the left side, the morphemes that can be added without generating two free forms occupy positions 6 (plural *d=*) and 7 (possessive marker *f-*) as in (94). Thus, the maximal domain of the free form is 6–9.

- (94) 'dʃgæ's'gi:b 'Llú:py

d' = f- gæ's +gi:b Llú:py

n:6= 7- 8-9 17

PL= POSS- pot +metal Guadalupe

'Guadalupe's pots'

5.1.2 Non-interruptability (6-12, 1-9)

We fracture the (non-)interruptability test to TdVZ nominal domain as defined below.

- (95) Non-interruption1: A span that cannot be interrupted by an element that can occur in more than one position in the planar structure (i.e. a second position clitic).
- (96) Non-interruption2: A span that cannot be interrupted by any free form (e.g. an adjective).

The non-interruption1 test identifies the span 6–12. This is shown in the following examples. In (97), note that the second position clitic may occur on the relational noun *kwæ* 'side of' in position 2. On the right side, the second position clitic must occur after the intensifier (position 12) as in (98), but not before unless the elements in position 10–12 do not occur, as seen in (99). Recall that second

position clitics can only occur on the nominal domain when the nominal moves to the preverbal focus position. Also, the occurrence of the intensifier(s) depends on the occurrence of adjectives. That is, if adjectives do not occur, intensifiers do not occur either.

- (97) 'kwæ'zá 'dmě:ʒræ 'zɥ:bdēn
 kwæ' =zá d' = mē:ʒ =ræ ø- zɥ:b =dēn
 v:2 10- 12 =23
 n:2 =4 6= 8 =16
 side.of =also PL= table =DEM.DIST STAT- POS.V.placed =3PL.INAN
 'Those (things) are also placed next to those tables.'
- (98) 'kwæ' 'dmě:ʒ rô'w tæ:záræ 'zɥ:bdēn
 kwæ' d' = mē:ʒ rô'w tæ: =zá =ræ ø- zɥ:b
 v:2 10- 12
 n:2 6= 8 10 12 =14 =16
 side.of PL= table big INTENS =also =DEM.DIST STAT- POS.V.placed
 =dēn
 =23
 =3PL.INAN
 'Those (things) are also placed next to those very big tables.'
- (99) *'kwæ' 'dmě:ʒ rô'wzá tæ:ræ 'zɥ:bdēn
 kwæ' d' = mē:ʒ rô'w=zá tæ: =ræ ø- zɥ:b
 v:2 10- 12
 n:2 6= 8 10=14 12 =16
 side.of PL= table big=also INTENS =DEM.DIST STAT -POS.V.placed
 =dēn
 =23
 =3PL.INAN
 Intended reading: 'Those (things) are also placed next to those tables.'

The non-interruption₂ test, on the other hand, identifies a span from positions 1–9, as illustrated in the following example. This span cannot be interrupted by a free form element, such as an adjective. In (101), we show that placing an adjective between positions 2–8 triggers ungrammaticality.

- (100) 'ftɛ:n 'dæts 'yu' 'ngĩtskánēn
 ftɛ:n dæts yu' ngĩts =kán =ēn
 n:1 2 8 10 =16 =20
 PREP.of back.of building white =DEM.MED =FOC
 'It is from/of behind the white building.'
- (101) *'ftɛ:n 'dæts 'ngĩts 'yu'kánēn
 ftɛ:n dæts ngĩts yu' =kán =ēn
 n:1 2 10 8 =16 =20
 PREP.of back.of white building =DEM.MED =FOC
 Intended reading: 'It is from/of behind the white building.'

5.1.3 Nonpermutability (6-20)

(Non-)permutability identifies spans where the ordering of elements cannot be permuted. In the nominal domain, this test identifies the span 6–20, as in (102).¹⁵ In (103), we show that the order of the elements in this span is strict. Outside of this domain, a second position clitic that attaches to the noun (base) when moved to the focus position may occur on the noun or on the preceding element in position 2, as shown in (104).

- (102) 'dʃkú'n gwě'ní'n 'tæ:rwánkĩ bā:nĩn
 d' = f- gũ'n gwě:n -i'ny tæ: =rú =an =kĩ
 v:2
 n:6= 7- 8 10 -11 13 =15 =17 =19
 PL= POSS- bull small -DIM INTENS =more =3SG.INF =DEM.TPRL
 b-a:ny =ĩn
 10-12 =24

 COMPL-do =3SG.INAN
 'Those smallest bulls of her/his did it.'
- (103) *'dʃkú'næ'n 'gwě:n 'tæ:rwánkĩ bā:nĩn
 d' = f- gũ'n -æ'ny gwě:n tæ: =rú =an =kĩ b-
 v:2 10-
 n:6= 7- 8 -11 -10 13 =15 =17 =19
 PL= POSS- bull -DIM small INTENS =more =3SG.INF =DEM.TPRL COMPL-

¹⁵When two adjectives occur in position 10, they usually follow the order of: color + size + human propensity, although the order of color and size may vary with some specific colors (black and white). However, since this does not occur with most color terms, we consider that adjectives still have a non-permutable pattern.

a:ny = *īn*

12 =24

do =3SG.INAN

Intended reading: ‘Those smallest bulls of her/his did it.’

(104) *kwæ'zá* 'dʃkũ'n *gwē'nī'n* 'tæ:rwánkī *zû:m*

kwæ' = *zá* *d'* = *ʃ-* *gũ'n* *gwē:n* -*i'n* *tæ:* = *rú* = *an*

v:2

n:2 =4 6= 7- 8 10 -11 12 =15 =17

side.of =also PL= POSS- bull small -DIM INTENS =more =3SG.INF

=*kī* *ø-* *zu:* =(*u*)*m*

10- 12 =23

=19

=DEM/TPRL STAT- stand =3SG.ANML

‘It (the animal) is/was also standing next to those smallest bulls of her/his.’

5.1.4 Deviations from biuniqueness (8-8)

The only deviations from biuniqueness present in the noun complex are multiple forms to one meaning mappings. In nominal planar structure, only the noun base in position 8 may show non-automatic allomorphy. First, noun bases may undergo suppletion (strong, in the cases of (106) and (107)) according to their possessive status: the possessed and unpossessed forms are segmentally unrelated. The following are such examples:

(105) a. 'ge:ɖʒ 'village'

b. 'la:ɖʒ 'village of'

(106) a. 'yu' 'building'

b. 'li:z 'building of'

(107) a. 'la:dy 'clothes'

b. 'ʃa:b 'clothes of'

(108) a. 'byā:g 'shirt'

b. 'zyā:g 'shirt of'

(109) a. 'bækw 'dog'

b. 'ʃikw 'dog of'

Furthermore, noun bases may undergo unpredictable tonal (and in some rare cases, phonation) alternation when the possessor is the 1SG. In the following example, the base form without possessor is *l̥aːz* ‘essence’, with a low tone and a creaky vowel. However, when the possessor is 1SG, the base tone alternates with a rising tone and the creaky vowel with the modal:

- (110) *l̥aːzá*
l̥ǎːz =a
 n:8 =17
 essence:1SG =1SG
 ‘my essence/center’

5.1.5 Ciscategorical selection (7-8)

In the nominal planar structure, the only positions unique to nouns are the noun base in position 8 and the possessive prefix in position 7. This defines the span of ciscategorical selection. The morphemes in the positions outside of this span may occur with other parts of speech. For instance, the plural *d=* in position 6 may occur with the subordinator or a numeral; the postponed elements in positions 9–17 may also occur with adjectives or verbs, as we have seen in §4.

5.2 Phonological diagnostics

5.2.1 Glottal dissimilation (8-9, 3-15)

As mentioned above in §4.2.1, TdVZ has a Glottal Dissimilation rule such that a glottalized syllable is deglottalized when followed by another glottalized syllable (CV^ʔCV^ʔ > CVCV^ʔ). The effect of this rule is evident with the diminutive (in the position 11 in the nominal template), which itself has a glottalized vowel. When the noun base has a glottalized vowel, the glottalization of the root vowel is lost when it is followed by the diminutive suffix, as shown in the following example.

- (111) *gú'næʔn*
gúʔn -æʔny
 n:8 -11
 bull -DIM
 ‘little / nice bull’

For the nominal domain, this test has been fractured in the following fashion:

- (112) Glottal Dissimilation (minimal): The minimum contiguous span overlapping the noun stem where this process is attested.
- (113) Glottal Dissimilation (maximal): the span of positions which contain elements which display no evidence against Glottal Dissimilation. Outside of this domain Glottal Dissimilation is never observed.

The minimal span where Glottal Dissimilation is attested is between position 8–9. That is, within this span, all morphemes undergo Glottal Dissimilation when they meet the structural requirement, as shown in the following.

- (114) ru'ryu'
 ru' +ryu'
 n:8 +9
 mouth +HAB.enter
 'entrance'

Outside of this domain, within the span of positions 3–15, this process may or may not be observed. As we have shown above, it is applied in the sequence of position 8 (noun base) and position 11 (diminutive), but it is not applied between positions 8 and 10, as shown in the following. The generalization is that Glottal Dissimilation is not applied between independent prosodic words (but Glottal Dissimilation is not always applied between the noun base and a bound morpheme either, as we will see in (116) below).

- (115) tu'bru' 'ri' 'ngâ' (*tu'bru 'ri 'ngâ')
 tubru' ri' ngâ'
 n:2 8 10
 QR.some pitcher blue
 'Some blue pitchers'

In other cases, we simply cannot tell if Glottal Dissimilation is applied or not, since some positions do not contain any morphemes with a glottalized vowel. This is the case with the positions 1, 3, 4, 5, 6, 7, 12, and 20.

Outside of the span of 3–15, Glottal Dissimilation is known not to be applied; this defines the maximal span of Glottal Dissimilation. For instance, this process is not applied between a quantifier in position 2 and the noun base in position 8, or between the noun base and the proximal demonstrative in position 16, as shown below.

- (116) tu'bru' 'bæ' ræ'¹⁶
 tubru' *bæ'* =*ræ'*
 n:2 8 =16
 QR.some mushroom =DEM.PROX
 'Some of this mushroom'

5.2.2 Accentuation (8-9, 7-11)

We have fractured Accentuation in the following manner to test constituency:

- (117) Accentuation (minimal): The minimal contiguous span where prominence is attested to be assigned.
- (118) Accentuation (maximal): Elements in this span interact with stress assignment but not necessarily. Outside of this domain, each element has its own prominence or never has prominence.

Prominence is assigned to the last syllable of the minimal span of Accentuation. In the case of a simple root as in (119), or when the noun base contains a fossilized prefix as in (120), all of which are in position 8, the prominence is assigned to the last syllable.

- (119) *'bi:*
 n:8
 'air'
- (120) *gu'tip*
 n:8
 'wasp'

In compounds which consist of the morphemes in the positions 8–9, prominence is assigned to the last syllable, as in the following example; the preceding (syllable) root does not have prominence. This is the span where prominence is attested, thus, the minimum domain.

- (121) *diʒ'za:*
 di:ɕ +*za:*
 n:8 +9
 language +Zapotec?
 'Zapotec language'

¹⁶The final glottalization in the first line is due to final glottalization (§4.2.7).

The morphemes outside of this domain, within the span of 8–11, may or may not be assigned prominence. For instance, the diminutive suffix in position 11 is within the domain of Accentuation. When a noun root in position 8 and the diminutive suffix in position 11 constitute independent syllables, the prominence is assigned to the syllable of the suffix and the root preceding these suffixes does not have prominence.

- (122) gubá'ni'n
 gubâ'ny -i'ny
 n:8 -11
 broom -DIM
 'little/nice broom'

However, the adjectives in position 10 project prominence independent from the noun, and thus constitute a separate domain of Accentuation from the noun base:

- (123) te'yu' na'zé:n dŭftæ:
 te= yu' nazé:n dŭftæ:
 n:2= 8 10 12
 INDF= building narrow INTENS
 'a very, very wide building'

On the right side, the diminutive suffix is the last element that can be assigned prominence; on the left side, we know that stress assignment does not interact with positions 1–6, but position 7 only contains a morpheme (*f*-) that does not form a syllable. Thus, it is included within the maximal domain. Elements outside of this domain are never assigned prominence (as in positions 3, 4, 5, 6, 13, 14, 15, 16, 17 (when enclitic), 19, and 20) or have their own prominence (1, 2, 12, 17 (when NP), and 18).

5.2.3 Syllabification (6–8; 1–20)

Segments in the nominal domain are syllabified in the sequences of the noun base in position 8 and the compounded root in position 9 as in the following. Here, the coda consonant of *ya:g* 'tree' is realized as part of the onset cluster in the following syllable.

- (124) ya.'gyu:
 ya:g +yu:
 n:8 -9
 tree +soil
 'soil tree (a type of tree from the region)'

To the left side of the nominal planar structure, the possessive prefix *f-* in position 7 syllabifies with the noun base in position 8 as in (125); this is also the case with the plural proclitic in position 6 *d=* as in (126). Thus, the span 6–8 is the minimal domain of syllabification.

- (125) 'f̃ā:r.mán
 f- *ǎ:rm* =an
 n:7- 8 =17
 POSS- container.used.for.measure =3SG.INF
 'his container used for measure.'

- (126) 'dî:z
 d= *i:z*
 n:6= 8
 PL= year
 'years'

Beyond this minimal span, syllabification may or may not apply within the whole span of 1–20. A sequence of morphemes are also resyllabified between a base + diminutive suffix in position 11, as in (127), or between the noun base and the pronominal enclitic in position 17, as in (128). Syllabification also applies between the noun base and the focus enclitic in position 20, as in (129). Thus, the right edge of the domain of syllabification includes all post-nominal elements.

- (127) bæ.'kwæ'n
 bækw -æ'ny
 n:8 -11
 dog -DIM
 'little/nice dog'

- (128) 'fpā:.yú
 f- *bǎ:y* =u
 n:7- 8 =17
 POSS- scarf =2SG.INF
 'your scarf'.

- (129) 'gĩ:.dín
 gĩ:dy =ĩn
 n:8 =20
 hen/female =FOC
 'It is (a) hen.'

However, other positions in this domain (6 - 10) are known not to participate in syllabification. Thus, resyllabification is not observed between the noun base and the following adjective in position 10:

- (130) 'i:z.'yuf (*i.zyuf)
 i:z +yuf
 n:8 +10
 year +old
 'old year, the year that ended'

In addition, syllabification is not observed between independent prosodic words (this is also the case with (130) above). Thus, when position 17 is occupied by a possessor NP, syllabification is not applied, as in (131). Syllabification is also not observed between independent prosodic words in position 1 and the noun base as in (132) or between position 2 and the noun base as in (133).

- (131) 'ʃkæt 'Á:n
 ʃ- gæt á:n
 n:7- 8 17
 poss- tortilla Ana
 'Ana's tortilla'

- (132) 'ʃt̬ɛ:n 'Á:n
 ʃt̬ɛ:n á:n
 n:1 8
 possession Ana
 'Ana's/of Ana'

- (133) 'zyē:n ĩ:z
 zyē:n ĩ:z
 n:2 8
 various year
 'various years'

Other positions within this span do not contain any morphemes that would allow us to see if syllabification is applied with the noun base. Additionally, the first morpheme has to end in a consonant and the following with a vowel or a glide (which is not common), one of the morphemes has to be bound, since as we have seen above, syllabification is not applied between independent prosodic words.

5.2.4 Rising tone levelling (8-20)

As mentioned above in §4.2.4, Rising Tone Levelling is a tonal process where a rising tone is split into a mid tone and a high (or falling) tone on the following syllable. In the nominal domain, Rising Tone Levelling is known to apply between the noun base in position 8 and the diminutive suffix in position 11:

- (134) ʒi'tæ'n
 ʒi't -æ'ny
 n:8 -11
 cat -DIM
 'little/ nice cat'

Rising Tone Levelling is also applied between the noun base in position 8 and a pronominal enclitic in position 17 as in (135) or a focus enclitic in position 20 as in (136):

- (135) ʃpā:yú
 ʃ- bǎ:y =u
 n:7- 8 =17
 POSS- scarf =2SG.INF
 'your scarf.'

- (136) gī:dín
 gī:dy=īn
 n:8 =20
 hen/female =FOC
 'It is (a) hen.'

Outside of this span, Rising Tone Levelling is not attested, thus Rising Tone Levelling identifies the span 8–20. This is possibly due to the structural requirement of this process: for this process to apply, the first morpheme has to have a rising tone and is closed (all underived morphemes with a rising tone are closed),

and the next morpheme has to begin with a vowel. It is not applied when the following morpheme begins with a consonant, even if the sequence falls within the span of 8–20, as in the following:

- (137) 'fũ:bdán
f- *ʒũ:b* =*dán*
 n:7- 8 =17
 POSS- corn =3PL.INF
 'their corn'

5.2.5 Mid tone spreading (8-9, 1-20)

As we saw above in §4.2.5, Mid Tone Spreading is a process whereby a mid tone spreads to the preceding syllable with a lexical low tone. Mid Tone Spreading is observed between positions 8 and 9, as shown below. This is the shortest span where mid tone is always observed.

- (138) *gæ'sgē'w*
gæ's + gē'w
 n:8 +9
 pot +lime
 'lime container'

Mid Tone Spreading also applies in the sequences of a base (position 8) + a pronominal clitic indicating the possessor (position 17), as in (139). Also, this process is observed in the sequence of a base + the focus marker (position 20), as in (140).

- (139) 'ʃkĩtsdān
f- *gĩts* =*dān*
 n:7- 8 =17
 POSS- chapter =3PL.F
 'their paper.'

- (140) 'gĩtsēn
gĩts =*ēn*
 n:8 =20
 paper =FOC
 'It is paper.'

On the left side, Mid Tone Spreading is applied between the indefinite proclitic *te=* and the noun base as in (141); however, this process is not applied when this position is occupied by an independent prosodic word, as in (142).

- (141) *tēbēnny*
te= bēnny
 n:2= 8
 INDF= person
 ‘a person’
- (142) *tu'bru' gǣll* (**tu'brū' gǣll*)
tubru' gǣll
 n:2 8
 a.little anonas.fruit
 ‘Some of (the) anonas (fruit)’

There is no way to verify if this process would apply in positions 1, 3, 4, 5, 6, and 7 since no morpheme in this position has a low tone; recall that the syllable has to have a lexical low tone to undergo Mid Tone Spreading. Thus, we conclude that the maximal domain of Mid Tone Spreading is the span of positions 1–20; within this span, some morphemes may undergo Mid Tone Spreading, but not always.

5.2.6 Tone sandhi (8–11, 1–20)

As we saw in §4.2.6, Tone Sandhi is a process where a syllable (with a low or mid tones) following one with a mid tone is assigned a high or falling tones. In the post-nominal positions, Tone Sandhi applies between the noun base in position 8 and the morphemes in position 9 (compounded roots) as in (143), position 10 (adjectives) as in (144), and position 11 (diminutive) as in (145). The intensifiers in position 12 are dependent on the occurrence of an adjective as in (146), thus, (empirically) we cannot test if Tone Sandhi applies between position 8 and 12. Therefore, we consider that the minimal domain for Sandhi is 8–11.

- (143) *bēn'gî:w*
bēnny +ngî:w
 n:8 +9
 person +man
 ‘male person / señor’

- (144) 'bēnny ná'da:w
bēnny nada:w
 n:8 10
 person patient
 'patient person'
- (145) bǣ'llǣ'n
bǣll -ǣ'ny
 n:8 -11
 woman's.sister -DIM
 '(nice) sister'
- (146) tē'zā *(ngǣs) 'tǣ:
tē= zā ngǣs tǣ:
 n:2= 8 10 12
 INDF= cloud black INTENS
 'a very dark cloud'

Beyond this domain, not all the morphemes can be shown to participate in Tone Sandhi. Thus, the enclitics in positions 13, 14, 15, 16, 19, and 20 have mid or high tones on an atonic syllable, which cannot participate in Tone Sandhi. The possessor in position 17 participates in Tone Sandhi, whether it is occupied by an enclitic or a free-standing NP:

- (147) 'bǣllú
bǣll =u
 n:8= 17
 woman's.sister =2SG.INF
 'your sister'
- (148) 'bǣll yá'né:t
bǣll yané:t
 n:8 17
 woman's.sibling Janet
 'Janet's sister'

In the pre-nominal positions, Tone Sandhi applies within positions 2–8, as in (149), between an adverbial enclitic in position 3 and the noun base as in (150), and between a proclitic in position 6 and the noun base as in (151).

- (149) 'zyē:n 'bénny
 zyē:n bēnny
 n:2 8
 QR.various person
 'Various people'
- (150) 'rá:zī fīnéky
 rá: =zī fīnéky
 n:2 =3 8
 all =only thing
 'Almost all the things / Any object around'
- (151) dākré'sě:nsy
 dā= kresě:nsy
 n:6= 8
 Mr.= Crescencio
 'Mr. Crescencio'

Positions 1, 4, 5, and 7 do not have any morpheme with a mid tone, thus we would not know if Tone Sandhi is applied between these positions and the noun base in position 8. Thus, the maximal domain of Tone Sandhi is 1-20.

5.3 Coincidence and convergence in the nominal domain

In this section, we display all the tests applied to the nominal domain. Below we show a summary of the convergence of these tests.

Table 5: Tests and convergence in the nominal domain in TdVZ

| Test | L | R | Size | Conv. | Definition |
|---------------------------------|---|---|------|-------|---|
| Minimum free form (minimal) | 8 | 8 | 1 | 2 | A span that is contiguous on its edges with a minimal free form that contains elements from positions with the shortest distance from each other. |
| Deviation from biuniqueness | 8 | 8 | 1 | 2 | A span where non-automatic allomorphy is observed. |
| Glottal Dissimilation (minimal) | 8 | 9 | 2 | 3 | The contiguous span where Glottal Dissimilation is attested. |

| | | | | | |
|--------------------------------------|---|----|----|---|--|
| Mid Tone Spreading (minimal) | 8 | 9 | 2 | 3 | The span that contains contiguous positions where Mid Tone Spreading applies. |
| Accentuation (minimal) | 8 | 9 | 2 | 3 | The contiguous span where prominence is attested to be assigned. |
| Ciscategorical selection | 7 | 8 | 2 | 1 | A span that contains elements that are unique to nouns. |
| Syllabification (minimal) | 6 | 8 | 3 | 1 | The contiguous span where elements of adjacent positions interact in syllabification. |
| Minimum free form (maximal) | 6 | 9 | 4 | 1 | A span that is contiguous on its edges with the minimal free form that contains elements with the largest difference from one another. |
| Tone Sandhi (minimal) | 8 | 11 | 4 | 1 | The contiguous span where tone sandhi is observed. |
| Accentuation (maximal) | 7 | 11 | 5 | 1 | Elements in this span interact with stress assignment, but not necessarily. Outside of this domain, each element has its own prominence or is never assigned prominence. |
| Non-Interruption1 (moveable element) | 6 | 12 | 7 | 1 | Elements in this span cannot be interrupted by a moveable element. |
| Non-Interruption2 (by an NP) | 1 | 9 | 9 | 1 | Elements in this span cannot be interrupted by a free form element. |
| Glottal Dissimilation (maximal) | 3 | 15 | 13 | 1 | The span of positions which contain elements which display no evidence against Glottal Dissimilation. |
| Rising Tone Levelling | 8 | 20 | 13 | 1 | The span whose positions contain elements that display positive evidence for Rising Tone Levelling. |

| | | | | | |
|------------------------------|---|----|----|---|--|
| Non-permutability | 6 | 20 | 15 | 1 | Elements in this span cannot be permuted or variably ordered. |
| Syllabification (maximal) | 1 | 20 | 20 | 3 | The longest span where elements of adjacent positions interact in syllabification. |
| Tone Sandhi (maximal) | 1 | 20 | 20 | 3 | The span where Tone Sandhi may apply. |
| Mid Tone Spreading (maximal) | 1 | 20 | 20 | 3 | The span containing elements which display no evidence against Mid Tone Spreading. |

Figure 7 provides an overview of the results of the constituency variables applied to TdVZ nouns in terms of layers. We can observe that the convergences are found in the spans 8-9 and 1-20, both of which have the convergence of three diagnostics, which are the best candidates for the word in TdVZ nouns.

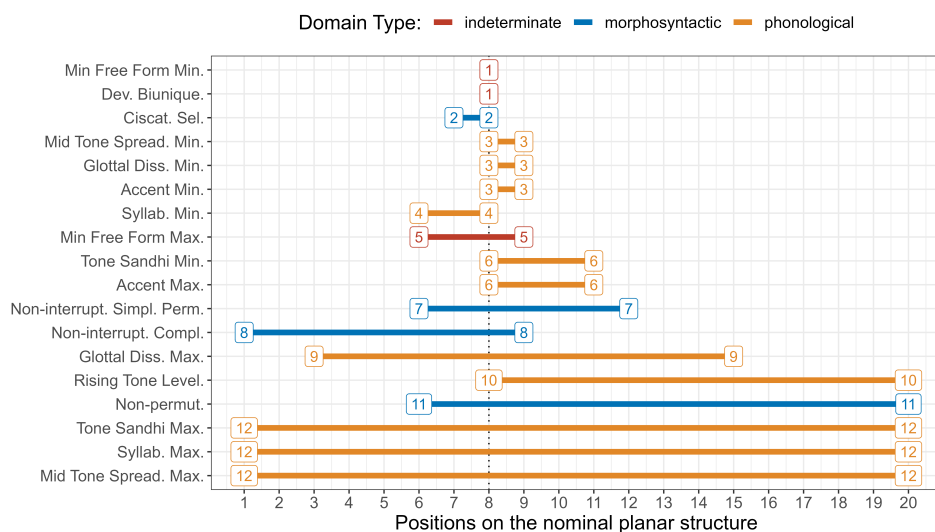


Figure 7: Constituency domains organized by converging layers in TdVZ nouns

The following Figures 8 and 9 show an overview of the results of the morphosyntactic and phonological constituency variables applied to TdVZ in terms of layers, respectively. As we can observe, taking the assumption that words are

areas of convergence, there is no clear candidate for a morphosyntactic word in the nominal domain in TdVZ. On the other hand, the best candidate for the phonological word is the 8–9 span and the 1–20 span, where three diagnostics converge each.

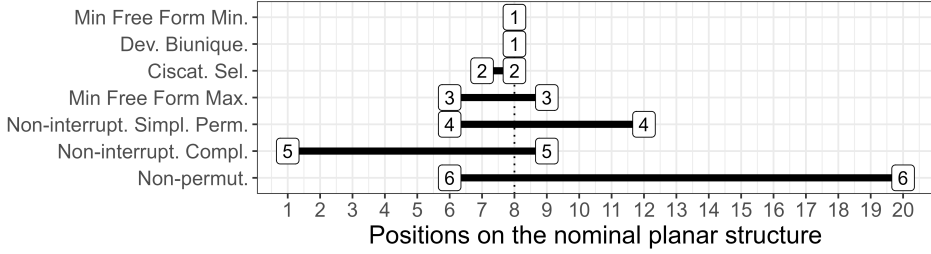


Figure 8: Morphosyntactic and indeterminate constituency domains organized by converging layers in TdVZ nouns

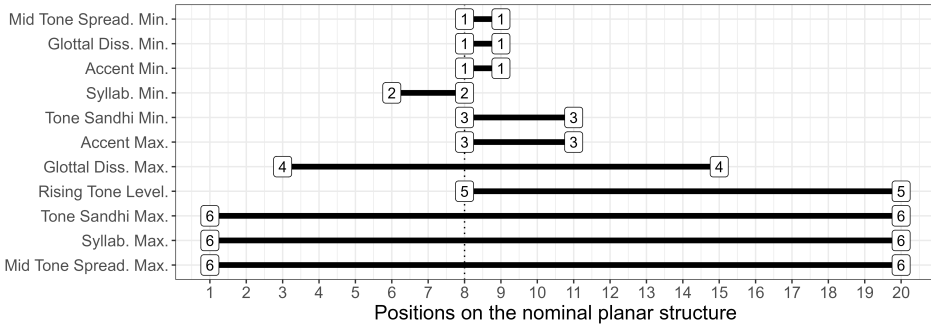


Figure 9: Phonological constituency domains organized by converging layers in TdVZ nouns

6 Conclusions and further research

In this chapter, we have reported the results of the application of 21 constituency tests to the verbal complex and 18 tests to the nominal complex in Teotitlán del Valle Zapotec (TdVZ). Assuming “words” are identified as domains of structure where constituency diagnostics converge (e.g. Matthews 2002), our goal was to assess what type of word constituents, if any, are motivated in TdVZ. In contrast to recent work emphasizing the ubiquity of wordhood domain divergence (Haspelmath 2011; Bickel & Zúñiga 2017; Tallman 2021), we showed that the data of

TdVZ provide support for at least some types of word constituent in this language: the span of positions 10–12 at the morphosyntactic level (convergence of 4 diagnostics) and 1–28 at the phonological level (4 diagnostics) in the verbal domain, and the spans of positions 8–9 and 1–20 at the phonological level (3 diagnostics each) in the nominal domain.

This study thus suggests that the high degree of misalignments found in Bickel & Zúñiga (2017) may result from the consideration of an arbitrarily low number of diagnostics; Bickel & Zúñiga (2017) only consider 6. Given that we applied the same methodology Tallman (2021) applied to Chácobo, the results suggest that languages vary in terms of the degree to which wordhood diagnostics cluster. In fact, we showed that such divergence/convergence varies language internally when we compare the morphosyntactic structure of part of speech categories.

As we showed, there are some particularities in the application of tests that need to be included in the methodology proposed by Tallman (2020). The planar structure strategy used here has advantages for comparison and for testing convergences, but we encountered some difficulties in applying this methodology. For instance, there are elements that show inter-dependencies; these elements will only occur if an element they attach to occurs. For instance, the focus marker in position 3 in the verbal domain will only occur when a phrase takes the pre-verbal position in position 2. This could be captured with the free occurrence test or through deviation from biuniqueness test.

Another difficulty we encountered basing purely on the linear order of the morphemes is capturing the difference in the behavior of the bound and free morphemes that occupy the same position. For instance, positions 23 - 25 in the verbal domain and position 17 in the nominal domain can be occupied either by a pronominal enclitic or an independent NP. When these positions are occupied by the enclitics, they undergo various phonological processes, such as Syllabification, Rising Tone Levelling, and Mid Tone Spreading, but they do not when they are occupied by an independent NP. This difference could be captured by test fracturing, which is not done in this chapter for the sake of space.

Finally, the question remains if words should really be defined just based on clustering of diagnostics in the way we have done here. If not, then how? Is there a non-ad-hoc way to define morphosyntactic and phonological words such that they correspond more closely to “intuitive words” employed by speakers and linguists?

Abbreviations

| | | | |
|-------|---------------|--------|---------------------|
| 1 | first person | INTENS | intensifier |
| 2 | second person | MED | medial |
| 3 | third person | NEG | negative |
| ANML | animal | PL | plural |
| CAUS | causative | POS | positional |
| COM | comitative | POSS | possessive |
| COMPL | completive | POT | potential |
| DEM | demonstrative | PREP | preposition |
| DIM | diminutive | PRON | independent pronoun |
| DIST | distal | PROX | proximal |
| F | feminine | QR | quantifier |
| FOC | focus | RECP | reciprocal |
| FOR | formal | REST | restorative |
| HAB | habitual | SG | singular |
| IMP | imperative | STAT | stative |
| INAN | inanimate | SUB | subordinator |
| INCL | inclusive | TPRL | temporal |
| INDF | indefinite | V | verb |
| INF | infinitive | VEN | venitive |

References

- Aronoff, Mark & Kirsten Fudeman. 2005. *What is morphology?* Oxford: Blackwell.
- Beam de Azcona, Rosemary. 2004. *A Coatlán-Loxicha Zapotec grammar*. University of California, Berkeley. (Doctoral dissertation).
- Beam de Azcona, Rosemary. 2019. Southern Zapotec verb classes. In Enrique L. Palancar, Matthew Baerman & Timothy Feist (eds.), *Inflectional class complexity in the oto-manguean languages*, 121–166. Special issues of Amerindia 41.
- Beam de Azcona, Rosemary & Emiliano Cruz Santiago. 2022. Los compuestos verbales y las expresiones idiomáticas del zapoteco miahuateco de San Bartolomé Loxicha. In Francisco Arellanes, Mario Ernesto Chávez Peón & Rosa María Rojas Torres (eds.), *Estudios descriptivos de lenguas zapotecas: Fonética-fonología, morfosintaxis y semántica léxica*, 91–138. CDMX: UNAM.
- Bickel, Balthasar & Fernando Zúñiga. 2017. The ‘word’ in polysynthetic languages: Phonological and syntactic challenges. In Michael D. Fortescue, Marianne Mithun & Nicholas Evans (eds.), *The Oxford handbook of polysynthesis*, 158–185. Oxford: Oxford University Press.

- Bloomfield, Leonard. 1914. Sentence and word. *Transactions and Proceedings of the American Philological Association* 45. 65–75.
- Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart & Winston.
- Broadwell, George A. 2000. Coordination, clitic placement, and prosody in Zapotec. In *Berkeley linguistic society*, 251–260. Berkeley: University of Berkeley.
- Campbell, Eric W. 2011. Zenzontepec Chatino aspect morphology and Zapotecan verb classes. *International Journal of American Linguistics* 77(2). 219–246.
- 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: ??.
- Chávez-Peón, Mario E. 2010. *The interaction of metrical structure, tone and phonation types in Quiaviní Zapotec*. The University of British Columbia. (Doctoral dissertation).
- Chávez-Peón, Mario E. 2015. Morfología prosódica en el zapoteco de Quiaviní. In Esther Herrera Zendejas (ed.), *Tono, acento y estructuras métricas en lenguas mexicanas*, 207–234. CDMX: El Colegio de México.
- Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- 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.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68. 81–138. DOI: 10.1353/lan.1992.0028.
- Gutiérrez, Ambrocio. 2014. *Construcciones de verbos seriales en el zapoteco de teotitlán del valle*. Centro de Investigaciones y Estudios superiores en Antropología Social. (MA thesis).
- Gutiérrez, Ambrocio & Hiroto Uchihara. N.d. A grammar of Teotitlán del Valle Zapotec.
- 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.
- Hernández Luna, Mario. 2021. La palabra fonológica en miahuateco. *Cuadernos De Lingüística De El Colegio De México* 8. 1–37.
- Kaufman, Terrence. 1989. The phonology and morphology of Zapotec verbs. <https://islandora-ailla.lib.utexas.edu/es/islandora/object/ailla%3A247379>.
- Lee, Felicia. 1999. *Antisymmetry and the syntax of San Lucas Quiaviní Zapotec*. University of California, Los Angeles. (Doctoral dissertation).

- Matthews, Peter H. 2002. What can we conclude? In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: A cross-linguistic typology*, 266–281. Oxford: Oxford University Press.
- Mithun, Marianne. 2014. What’s in a word? In Carol Genetti (ed.), *How languages work*, 71–99. Cambridge: Cambridge University Press.
- Mugdan, Joachim. 1993. Morphological units. In Robert E. Asher & James M. Y. Simpson (eds.), *The encyclopedia of language and linguistics*, 2543–2553. Oxford: Pergamon Press.
- Munro, Pamela. 2004. “Second position” in San Lucas Quiavini Zapotec. In *The Linguistic Association of the Southwest (LASSO), annual meeting*, 251–260. New Orleans.
- Nakamoto, Shun. 2024. Constituency in Ayautla Mazatec. In Adam J.R. Tallman, Sandra Auderset & Hiroto Uchihara (eds.), *Constituency and convergence in the Americas*, 229–261. Berlin: Language Science Press. DOI: ??.
- Nellis, Donald G. & Barbara E. Hollenbach. 1980. Fortis versus lenis in Cajonos Zapotec phonology. *International Journal of American Linguistics* 46. 92–105. DOI: 10.1086/465639.
- Osborne, Timothy J. 2018. Tests for constituents: What they really reveal about the nature of syntactic structure. *Language Under Discussion* 1(5). 1–41.
- Pérez Báez, Gabriela & Terrence Kaufman. 2016. Verb classes in Juchitán Zapotec. *Anthtopological Linguistics* 58(3). 217–257.
- Post, Mark W. 2009. The phonology and grammar of Galo “words”: A case study in benign disunity. *Studies in Language* 33(4). 934–974. DOI: 10.1075/sl.33.4.05pos.
- 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: ??.
- Sapir, Edward. 1921. *Language*. New York: Harcourt, Brace & World.
- Sapir, Edward. 1930. Southern Paiute, a Shoshonean language. *Proceedings of the American Academy of Arts and Sciences* 65(1). 1–296.
- Schiering, René, Balthasar Bickel & Kristine A. Hildebrandt. 2010. The prosodic word is not universal, but emergent. *Journal of Linguistics* 46(3). 657–709.
- Smith-Stark, Thomas. 2002. Las clases verbales del zapoteco de Chichicapan. In *Proceedings of the VI encuentro internacional de lingüística en el Noroeste*, 165–212. Hermosillo: Editorial UniSon.
- Smith-Stark, Thomas. 2003. Tipos prosódicos de sílabas en el zapoteco de San Baltazar Chichicapan. In Esther Herrera Zendejas & Pedro Martín Butragueño

- (eds.), *La tonía: Dimensiones fonéticas y fonológicas*, 111–139. Mexico City: El Colegio de México.
- Tallman, Adam J. R. 2018. *A grammar of Chácobo, a southern Pano language of the northern Bolivian Amazon*. University of Texas at Austin. (Doctoral dissertation).
- 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.
- Uchihara, Hiroto. 2016. Tone and registrogenesis in Quiavini Zapotec. *Diachronica* 33(2). 220–254. DOI: 10.1075/dia.33.2.03uch.
- Uchihara, Hiroto & Ambrocio Gutiérrez. 2020a. Open and closed mid-front vowels in Teotitlán del Valle Zapotec. *Phonological Data and Analysis* 2(7). 1–22. DOI: 10.3765/pda.v2art7.42.
- Uchihara, Hiroto & Ambrocio Gutiérrez. 2020b. Subject and agentivity in Teotitlán Zapotec. *Studies in Language* 44(3). 548–605. DOI: 10.1075/sl.18025.uch.