

Phonemic Description of San Miguel del Valle Zapotec

by

Isabelle Altman

B.A., University of California, Los Angeles

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado Boulder in partial fulfillment

Of the requirement for the degree of

Master of Arts

Department of Linguistics

2025

Committee Members:

Ambrocio Gutiérrez Lorenzo

Rebecca Scarborough

J Calder

Abstract

Altman, Isabelle (M.A., Linguistics University of Colorado Boulder)

Phonemic Description of San Miguel del Valle Zapotec

Thesis directed by Assistant Professor Ambrocio Gutiérrez Lorenzo

San Miguel del Valle Zapotec (SMdVZ) is a variety of Central Valley Zapotec spoken in the village of San Miguel del Valle (SMdV). It is spoken by the 3000-3500 members of the San Miguel del Valle community. This variety has not been documented nor described as far as I am aware. I visited the community over the summer of 2024 to collaborate with speakers and work on describing the language. There are 22 phonemic consonants and 7 phonemic vowels in this language. Similar to other Central Valley Zapotec varieties, there are three tones, including low, rising, and falling, and three phonation types, including modal, laryngealized, and glottalized. This thesis aims to describe this variety of Zapotec while also comparing it to other varieties, aiming to place it in relation to these other varieties. It also presents questions for further research. It should be noted that this thesis is a description of the language and is not an experimental work.

Acknowledgements

This thesis would not have been possible without the help of many individuals. Firstly, I would like to thank the community of San Miguel del Valle and the speakers who collaborated with me on this project. I am especially grateful to Carmen Santiago, Margarita García Hernández, Francisca García Hernández, Humberto Hernández García, Asunción López, Teresa López Antonio, and Rigoberto Hernández López for their patience and guidance during elicitation sessions.

I am extremely grateful to my advisor Dr. Ambrocio Gutiérrez Lorenzo for introducing me to the Zapotec languages and guiding me in my fieldwork and analysis of San Miguel del Valle Zapotec. He has been an incredible mentor and support throughout my time at the University of Colorado, Boulder.

I am also grateful to the other members of my committee, Dr. Rebecca Scarborough and Dr. J Calder. They've been so helpful throughout the thesis writing process and as professors and mentors during my time at CU Boulder. Their input was invaluable and the courses they offered extremely beneficial.

I would also like to thank the Department of Linguistics at the University of Colorado, Boulder for creating a wonderfully encouraging environment in which I could pursue my interests and grow as a documentary and descriptivist linguist. The work in this thesis was possible due to the funds allocated to me from the Department of Linguistics during my time at the University of Colorado, Boulder as a master's student. Other funds came from the Latin American and Latinx Studies Center at the University of Colorado, Boulder and the IIE Study Abroad Scholarship, which made my fieldwork trip to San Miguel del Valle possible.

Lastly, I would like to thank my wonderful family and friends for their unwavering support in my academic endeavors. I am so grateful that I was raised in an environment where

pursuing my interests and learning were encouraged and celebrated. Thank you to my loving brothers, Quinn and Blake, my supportive step-mother, Marni, my amazing father, Steve, and my incredible mother, Kirstin.

Table of contents	
1. Introduction	1
1.1 San Miguel del Valle and language situation	1
1.2 Relationship to other Zapotec varieties	4
2. Methodology	6
3. Segmental phonology	7
3.1 Consonants	7
3.1.1 Fortis and lenis distinction	11
3.1.2 Non-native consonants and the influence of Spanish	15
3.1.3 Palatalization	16
3.2 Vowels	17
4. Suprasegmental phonology	19
4.1 Vowel quality (phonation)	19
4.1.1 Lack of a phonemic breathy phonation type	26
4.2 Diphthongs	30
4.3 Stress	34
4.4 Tone	36
5. Conclusions	45
6. Further research	45
References	48

List of tables

Table 1. Stops in word-initial, word-medial, and word-final positions	8
Table 2. Fricatives in word-initial, word-medial, and word-final positions	9
Table 3. Glides in word-initial, word-medial, and word-final positions	10
Table 4. Consonant phonemic inventory	10
Table 5. Vowel phonemic inventory	18
Table 6. Vowel quality (phonation) contrasts	21
Table 7. Possible glide-vowel and vowel-glide sequences	31
Table 8. Tone contrasts	44

List of figures

Figure 1. Map of the Central Valley in Oaxaca, México	4
Figure 2. Spectrogram of a lengthened modal vowel before a lenis consonant in /'ɣu:lj/ 'to be born, <i>nacer</i> ' by female speaker Carmen Santiago	13
Figure 3. Spectrogram of a non-lengthened modal vowel before a fortis consonant in /'nal:/ 'cold, <i>frio</i> ' by female speaker Francisca García Hernández	14
Figure 4. Spectrogram of a lengthened modal vowel word-finally in /'βdi:/ 'ant, <i>hormiga</i> ' by male speaker Humberto Hernández García	15
Figure 5. Spectrogram of a palatalized consonant in /e.'dʒopj/ 'to suck, <i>chupar</i> ' by female speaker Francisca García Hernández	17
Figure 6. Continuum of phonation types (after Ladefoged, 1971)	20
Figure 7. Spectrogram of a laryngealized vowel in /'βǣl/ 'meat, <i>carne</i> ' by female speaker Carmen Santiago	21
Figure 8. Spectrogram of a modal vowel in /'βæɫ/ 'fire, <i>fuego</i> ' by female speaker Carmen Santiago	22
Figure 9. Spectrogram of a laryngealized vowel in /'βjǎ/ 'nopal, <i>nopal</i> ' by female speaker Carmen Santiago	23
Figure 10. Spectrogram of a laryngealized vowel in /'βjǎʔ/ 'dance, <i>baila</i> ' by female speaker Francisca García Hernández	24
Figure 11. Spectrogram of a laryngealized vowel followed by a lenis consonant in /do.'βaŋ/ 'leather strap, <i>mecapal</i> ' by female speaker Carmen Santiago	25
Figure 12. Spectrogram of a glottalized vowel followed by a fortis consonant in /'ɣuʔn:/ 'bull, <i>toro</i> ' by female speaker Francisca García Hernández	26

Figure 13. Spectrogram and waveform of a modal vowel in /'βets/ 'brother, <i>hermano</i> ' by female speaker Francisca García Hernández in SMdVZ	27
Figure 14. Spectrogram and waveform of a breathy vowel in /'bɛts/ '(man's) brother, <i>hermano (de un hombre)</i> ' by female speaker in SLQZ (Pérez Báez and Uchihara 2016)	28
Figure 15. Spectrogram of a phonetic breathy vowel in /'βets/ 'brother, <i>hermano</i> ' by female speaker Francisca García Hernández in SMdVZ	29
Figure 16. Spectrogram of a glide-vowel sequence in /'zju/ 'long, <i>largo</i> ' by female speaker Francisca García Hernández	32
Figure 17. Spectrogram of a glide-vowel sequence in /'lwa/ 'my face, <i>mi cara</i> ' by female speaker Carmen Santiago	32
Figure 18. Spectrogram of a low tone in a glide-vowel sequence in /'dʝip/ 'strong, <i>fuerte</i> ' by female speaker Francisca García Hernández	34
Figure 19. Spectrogram of the word /i.'dʒa/ 'animal, <i>animal</i> ', demonstrating stress that falls on the ultimate syllable, by speaker Francisca García Hernández	35
Figure 20. Spectrogram of a low tone in an open-syllable in the word /'βi/ 'air, <i>aire</i> ' by speaker Humberto Hernández García	38
Figure 21. Spectrogram of the phonetic mid tone represented in the word /'zã/ ['zã] 'cloud, <i>nube</i> ' by speaker Carmen Santiago, a female speaker	41
Figure 22. Spectrogram of the phonemic low tone represented in the word /'za/ 'fat, <i>grasa</i> ' by speaker Carmen Santiago, a female speaker	42
Figure 23. Spectrogram of the phonemic high tone represented in the first syllable of the word /tú.'nɛ/ 'who, <i>quién</i> ' by speaker Carmen Santiago, a female speaker	43

Figure 24. Spectrogram of the phonemic rising tone represented in the word /'dãnj/ 'hill, *cerro*'
by speaker Francisca García Hernández, a female speaker 43

Figure 25. Waveform and spectrogram of the phonemic falling tone represented in the word /'nî/
'sour, *agrio*' by speaker Carmen Santiago, a female speaker 44

Abbreviations

IMP Imperative

SLQZ San Lucas Quiavini Zapotec

SMdV San Miguel del Valle

SMdVZ San Miguel del Valle Zapotec

SPGZ San Pablo Güilá Zapotec

TdVZ Teotitlán del Valle Zapotec

1. Introduction

The goal of this thesis is to give a basic phonemic description of San Miguel del Valle Zapotec, a variety of Zapotec native to the village of San Miguel del Valle, Oaxaca, in southern Mexico. It will provide evidence for phonemic parts of the language using evidence collected from word lists and narratives over three different visits to the village from seven different collaborators.

This work begins with an overview of the language situation of San Miguel del Valle and an attempt at placing the variety in relation to other Zapotec varieties in order to give background on the language and its speakers. Next, methodology is covered in detail to describe how the data for this language was collected for this thesis, followed by an overview of the segmental phonology, including consonants and vowels. Suprasegmental phonology follows, including information on vowel quality (phonation), diphthongs, stress, and tone. Finally, conclusions are presented, placing San Miguel del Valle Zapotec in relation to other Zapotec varieties, and questions for further research are posed.

1.1 San Miguel del Valle and language situation

Zapotec is a family of languages spoken mainly in the state of Oaxaca, Mexico. San Miguel del Valle is a village located in the state of Oaxaca in southern Mexico. It is located 26 miles away from the capital city of Oaxaca. Its location is highlighted below in Figure 1 with a blue arrow. It has a population of approximately 3000-3500 people, many of whom are speakers of San Miguel del Valle Zapotec (SMdVZ), a variety of Zapotec unique to the village. A map of Central Oaxaca is provided below with San Miguel del Valle highlighted in Figure 1, along with other villages and landmarks in the area. From my own observations, there is a significant portion of the population of the village that is SMdVZ dominant or even monolingual, with little

to no Spanish competence. This generation of speakers is older; generally 70 years and older. The middle age population is considered to be the last “fluent” generation of speakers of SMdVZ. This generation (30-70 years old) generally speaks both SMdVZ and Spanish fluently. The youngest generation of speakers (0-30 years old) includes some fluent speakers, some of whom I have collaborated with, but is made up of mainly Spanish-dominant speakers. This generation may have passive knowledge of SMdVZ, but doesn’t speak the language.

There could be a few reasons why this younger generation of speakers is slowly phasing out of speaking SMdVZ. For one, Spanish is considered to be the language of prestige and opportunity to the SMdV community. To these parents, Spanish competency will allow their children to do well in school, find a job, and be successful in a life outside of the village. Because of this, parents and the older generation choose not to speak SMdVZ to their children in hopes that they will learn Spanish instead (Beck 2008).

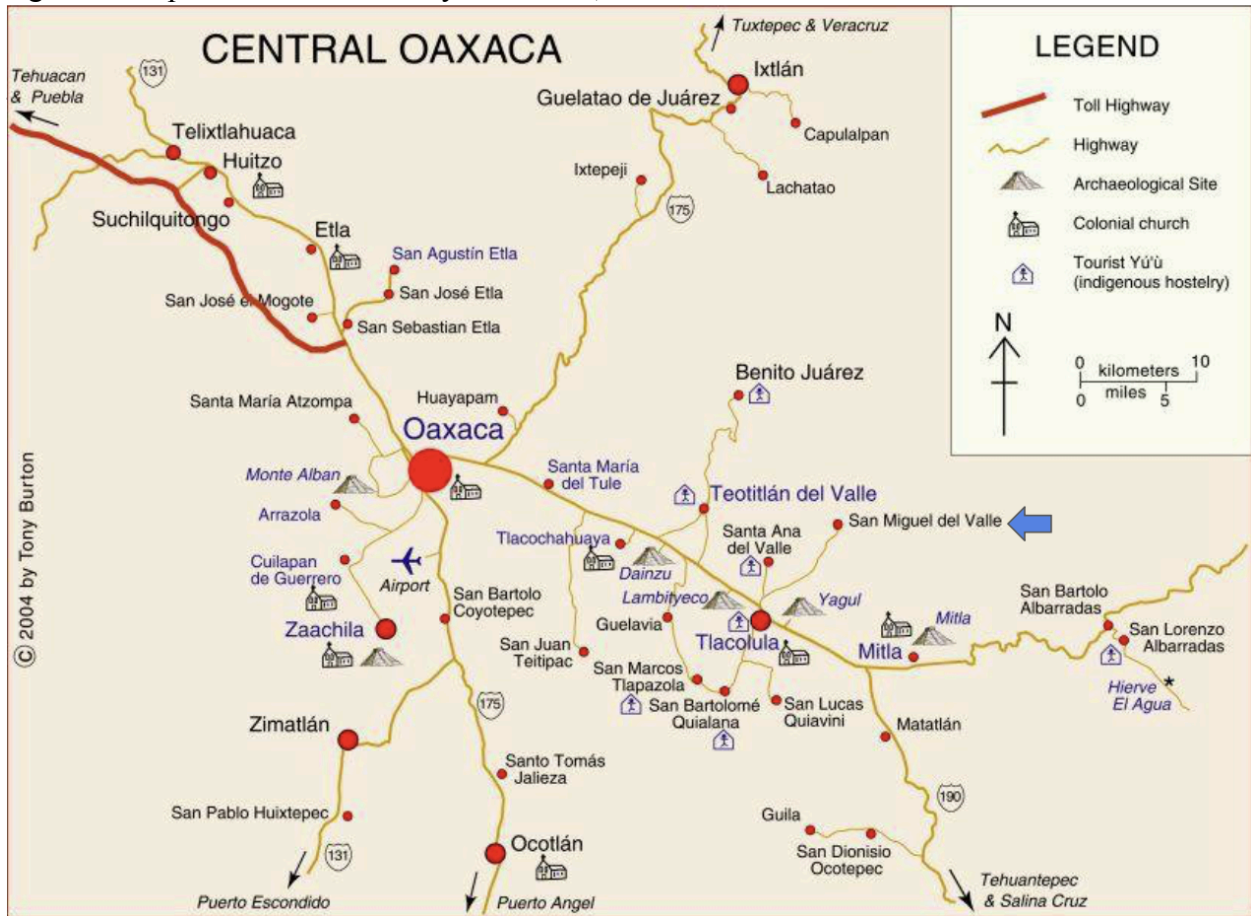
Another potential reason that the younger generation of SMdVZ speakers are Spanish-dominant is because the older generation can be purist in their views of language. From my observations, this older generation of speakers will hear the younger generation speak SMdVZ and disapprove of the way in which they speak, trying to give feedback and correct the speech of this younger generation. Because of this, the younger generation feels as though they don’t speak the language the “right” way and decide to not speak the language at all to avoid being corrected and accused of speaking the language in the “wrong” way. Of course, the influence of Spanish on SMdVZ over the years has led to some linguistic changes in the way younger Spanish bilingual speakers of SMdVZ produce their speech. This unfortunately is not an uncommon phenomenon, as other Indigenous communities have seen a decline in the younger

generation's use of their languages due to ridicule and judgement from older generations of speakers.

A final possible reason the youngest generation of SMdVZ speakers is dwindling is due to the language not being taught in schools (Hamel 2008). Many of the public schools in the area claim to teach SMdVZ, but are only bilingual in name and not in practice. This is due to the complicated public school teacher allocation system used throughout the state and in the rest of the country. In this allocation system, teachers can indicate that they speak Zapotec, but are often incorrectly placed into Zapotec-speaking communities that speak a different variety than the variety that they speak. Because of this, teachers are mismatched with communities that speak different varieties of Zapotec, resulting in the local variety of the language not being taught in schools.

The language situation of this community wouldn't necessarily be considered dire yet, but younger generations are definitely not using SMdVZ as much as their older counterparts. Because of this, there has been a movement in the community toward creating learning materials in SMdVZ, teaching SMdVZ in schools, and moving toward a "standard" orthography for the community to use.

Figure 1. Map of the Central Valley in Oaxaca, México



SMdVZ is an Otomanguean language, specifically a Central Valley Zapotec language, so it has VSO word order (Munro 2004; Gutiérrez Lorenzo 2021; Beam de Azcona 2023). Like many other Zapotec languages, it is both tonal and has different contrastive voice qualities, meaning that modal vowels contrast with non-modal vowels. As far as I am aware, there is no documentation or description of this particular variety of Zapotec so far.

1.2 Relationship to other Zapotec varieties

Based on my findings, SMdVZ seems to be closely related to Teotitlán del Valle Zapotec (TdVZ), a variety spoken in the nearby village of Teotitlán del Valle. Both varieties have most of the same consonants and vowels, sharing twenty-two consonants and six vowels (Gutiérrez

2021). Like TdVZ, SMdVZ has three contrastive tones, three distinct voice qualities (phonations), and VSO word order. Central Valley Zapotec languages are known to have either four or five contrastive tones, with varieties such as San Lucas Quiavini Zapotec (SLQZ) having four tones and others such as TdVZ having five tones. These languages are also known to have up to four distinct phonation types, including modal, laryngealized (creaky), glottalized, and breathy, and virtually all Zapotec languages are analyzed to have tone (Jaeger and Van Valin 1982). Varieties such as San Lucas Quiavini have all four of these phonation types, while other varieties have just three phonation types, such as TdVZ and San Pablo Güilá Zapotec.

Here I will briefly review the phonation types and voice qualities of different Zapotec varieties in order to understand where SMdVZ is in relation to these varieties.

Cajonos Zapotec, a Northern Zapotec Variety, has four phonemic vowels /i, o, e, a/ and three different phonation types, including modal, laryngealized, and glottalized (Nellis and Hollenbach 1980). There are four tones, including high, low, mid, and a “downglide from high to low” (Nellis and Hollenbach 1980).

Teotitlán del Valle Zapotec has six phonemic vowels /i, i, æ, o, u, a/ and three different phonation types, including modal, laryngealized, and glottalized (Gutiérrez Lorenzo 2021). There are five tones, including high, low, mid, rising, and falling.

San Pablo Güilá Zapotec (SPGZ) has six phonemic vowels /i, i, e, o, u, a/ and three different phonation types, including modal, laryngealized, and glottalized (Arellanes 2009). There are four tones, including high, low, rising, and falling.

Santa Ana del Valle Zapotec has six phonemic vowels /i, i, e, o, u, a/ and three different phonation types, including modal, breathy, and laryngealized (Esposito 2003). Tone operates in a slightly different manner for this Zapotec variety in that “Tone is contrastive on modal vowels,

which can have either a high or rising tone. Breathy and creaky vowels both have a falling tone” (Esposito 2003, 2004a).

San Lucas Quiaviní Zapotec (SLQZ) has six phonemic vowels /a, e, i, o, u, ĩ/ and four different phonation types, including modal, laryngealized, glottalized, and breathy (Munro and Lopez et al. 1999). There are four tones, including high, low, rising, and falling.

Mitla Zapotec has six phonemic vowels /a, e, ε, i, o, u/ and four different phonation types, including modal, laryngealized, glottalized, and breathy (García 2015). There are four tones, including high, low, rising, and falling.

Through understanding the different phonation types, phonemic inventories, and tone systems of each Zapotec variety, we can conclude that the Zapotec languages are varied and complex in their relationships to each other.

2. Methodology

The collection of data for this work has occurred in various stages. My advisor, Dr. Ambrocio Gutiérrez Lorenzo, visited San Miguel del Valle twice; once in the summer of 2022 and once in the summer of 2023. In the summer of 2024, I joined Dr. Gutiérrez Lorenzo on a fieldwork trip to San Miguel del Valle. The data recorded comes from word lists (of 500 words) and narratives from speakers of SMdVZ. All of the data used in this paper, however, come from the word lists. In total, there are seven speakers who consented to collaborate through recordings, all fluent speakers of SMdVZ with Spanish competence. To record collaborators, I used the Zoom H4n Pro 4 Track Recorder (2020 model). Collaborators were given a word in Spanish and were asked to repeat the Zapotec translation back twice. Collaborators were compensated accordingly. Praat was used to verify phonetic details and Microsoft Excel was used to store and compare data.

The seven collaborators varied in age and gender. Five of them were female, including Carmen Santiago, a bilingual Spanish and Zapotec speaker around 54 years of age, Margarita García Hernández, a bilingual Spanish and Zapotec dominant speaker around 70 years of age, Francisca García Hernández, a bilingual Spanish and Zapotec speaker around 50 years of age, Asunción López, a bilingual Spanish and Zapotec dominant speaker around 70 years of age, and Teresa López Antonio, a bilingual Spanish and Zapotec speaker around 45 years of age. The remaining two were male, including Rigoberto Hernández López, a bilingual Spanish and Zapotec speaker around 19 years of age, and Humberto Hernández García, a bilingual Spanish and Zapotec speaker around 58 years of age. All of the collaborators are native speakers of SMdVZ. The collaborators oftentimes recorded in a group setting, with up to two to three people participating in a single recording.

3. Segmental phonology

3.1 Consonants

SMdVZ has a total of twenty-two phonemic consonants, shown below in Figure 2. ‘F’ refers to ‘fortis’ and ‘L’ to ‘lenis’. Analytic contrasts were determined through minimal and near minimal pairs and the phonemic inventory of other varieties, namely Teotitlán del Valle Zapotec. Instead of a voiced and voiceless distinction, Zapotec varieties distinguish consonants based on a fortis and lenis distinction. Fortis consonants are generally voiceless and longer in duration, while lenis consonants are generally voiced and shorter in duration. This fortis and lenis distinction will be covered in depth in a later section. There are four stops, including /p/, /t/, /d/, and /k/, which can occur word-initially, word-medially, and word-finally. These are shown below in the following table in bold font:

Table 1. Stops in word-initial, word-medial, and word-final positions

	Word-initial	Word-medial	Word-final
/p/	/'psǎ'n/ 'leaves, <i>deja</i> '	/ɣu.'tǐp.ʒâ/ 'hornet, <i>avispón</i> '	/'tǒp/ 'pile (up), <i>amontona</i> '
/t/	/'tǒp/ 'accumulate, <i>amontona</i> '	/ɣu.'tǐp.ʒâ/ 'hornet, <i>avispón</i> '	/'yět/ 'below/down, <i>abajo</i> '
/d/	/'djæ/ 'squeeze, <i>aprieta</i> '	/áβ.'di/ 'finished, <i>acabado</i> '	/'tsǐd/ 'make a hole, <i>agujero</i> '
/k/	/'kǔtʃ.'zɪt/ 'armadillo, <i>armadillo</i> '	/'ʃkǔʃ/ 'sharpen, <i>afilado</i> '	/'reʔk/ 'burn, <i>arde</i> '

Three of these stops are fortis and one is lenis. There are three nasal consonants, including /m/, /n:/, and /n/. Interestingly, unlike some other Zapotec varieties, there does not seem to be a fortis bilabial nasal (Gutiérrez 2021). It should be noted that the tilde diacritic /Ṽ/ indicates laryngealization and the hook diacritic /Vʔ/ indicates glottalization, phonation types that will be discussed in further detail in later sections. All three of these nasal consonants can occur word-initially, word-medially, and word-finally. Like other lenis consonants, the two lenis nasals can occur only after lengthened vowels, while the single fortis nasal occurs after normal-length vowels. There is one tap consonant, the alveolar tap, which is a lenis consonant. This tap consonant can occur word-initially and word-medially, but not word-finally. There are six fricative consonants including /β/, /f/, /ʒ/, /s/, /z/, and /ɣ/, all of which can occur word-initially, word-medially, and word-finally. This is illustrated in the table below with the fricative in question in bold:

Table 2. Fricatives in word-initial, word-medial, and word-final positions

	Word-initial	Word-medial	Word-final
/β/	/βa.'sêɾ/ 'bee, <i>abeja</i> '	/áβ.'di/ 'finished, <i>acabado</i> '	/'ɣiβ/ 'iron, <i>fierro</i> '
/ʃ/	/'ʃkũʃ/ 'sharpen, <i>afilado</i> '	/'βʃiɭ/ 'open, <i>abre</i> '	/'jeʃ/ 'avocado, <i>aguacate</i> '
/ʒ/	/'ʒiɭj/ 'cotton, <i>algodón</i> '	/næ.'ʒin/ 'wide/open space, <i>amplio</i> '	/'juʒ/ 'sand, <i>arena</i> '
/s/	/stal.'dũʃ/ 'more, <i>más</i> '	/βa.'sêɾ/ 'bee, <i>abeja</i> '	/'nʝis/ 'water, <i>agua</i> '
/z/	/zú.'sa/ 'open, <i>abierto</i> '	/βzi.'lo/ 'finish, <i>acaba</i> '	/βă.'dziʒ/ 'squirrel, <i>ardilla</i> '
/ɣ/	/'ɣiβ/ 'iron, <i>fierro</i> '	/ɣĩɾf.'ɣôɭ/ 'cane, <i>cana</i> '	/'jaɣ/ 'plant, <i>planta</i> '

There are two lateral consonants including the fortis and lenis counterparts. Both of these lateral consonants can occur word-initially, word-medially, and word-finally. Like other lenis consonants, the lenis lateral consonant can occur only after lengthened vowels, while the fortis lateral consonant occurs after normal-length vowels. There are four affricates which can all occur word-initially, word-medially, and word-finally. Finally, there are two glides, the palatal glide /j/ and the labio-velar glide /w/, both of which can occur word-initially, word-medially, and word-finally. This is illustrated in the table below, followed by a complete table of these consonants:

Table 3. Glides in word-initial, word-medial, and word-final positions

	Word-initial	Word-medial	Word-final
/j/	/'jeʃ/ 'avocado, <i>aguacate</i> '	/lo.'ju/ 'below/down, <i>abajo</i> '	/'ɣaj/ 'five, <i>cinco</i> '
/w/	/wá.'zi/ 'spider, <i>araña</i> '	/'ɣwe/ 'go! (IMP), <i>anda</i> '	/'βsæw/ 'close, <i>cierra</i> '

Table 4. Consonant phonemic inventory

	Labial		Alveolar		Post-alveolar		Palato-alveolar		Palatal	Velar		Labio-velar	
	F	L	F	L	F	L	F	L		F	L	F	L
Stop	/p/		/t/	/d/						/k/			
Nasal		/m/	/n:/	/n/									
Tap								/ɾ/					
Fricative		/β/			/ʃ/	/ʒ/	/s/	/z/			/ɣ/		
Lateral			/l:/	/l/									
Affricate			/ts/	/dz/			/tʃ/	/dʒ/					
Glide									/j/				/w/

3.1.1 Fortis and lenis distinction

There is a distinction between fortis and lenis consonants in SMdVZ (as in other Zapotec languages: cf. Nellis & Hollenbach 1980; Jaeger 1983; Arellanes 2009; Chavez-peón 2010). Instead of a voiced/voiceless distinction, Zapotec languages have been claimed to have fortis and lenis pairs. Fortis obstruents are voiceless and relatively long, while lenis obstruents are (usually) voiced and shorter in duration. The main difference between fortis and lenis sonorants is duration, with fortis being longer than lenis, which is why these consonants are split up into fortis and lenis rather than voiced and voiceless. Though vowel length is not a contrastive feature in this language, vowels are lengthened before lenis consonants and not lengthened before fortis consonants. Vowels are also lengthened in open syllables. Some minimal and near minimal pairs are provided below to illustrate the contrast between given consonant pairs. The phonemes in question are in bold in the example words below. Minimal/Near minimal pairs for fortis/lenis pairs, fortis/lenis fricatives, and fortis/lenis affricates are given below. It should be noted that the low tone (J) is usually represented in IPA with the diacritic /à/ but in this thesis, it will be represented without a diacritic /a/. More on tone and its orthographic and IPA transcription representation will be discussed in later sections. Because stress is a relevant prosodic feature of this language, it is represented in the transcriptions below. Further information on stress will be provided in later sections.

/t/ vs /d/: /'β**ta**?/ 'crushed, *aplastado*', /'d**o**/ 'eat (IMP), *come*'

/s/ vs /z/: /'β**sa**/ 'flat (land), *llano*', /'β**za**/ 'bean, *frijol*'

/ts/ vs /dz/: /'t**si**/ 'ten, *diez*', /'d**zit**/ 'teeth, *dientes*'

/n:/ vs /n/: /'m**n**:i/ 'speaks, *habla*', /'n**î**/ 'sour, *agrio*'

/ʃ/ vs /ʒ/: /'j**ā**.gá/ '(my) cheek, (*mi*) *mejilla*', /'ʒ**jā**.gá/ '(my) grandson/daughter, (*mi*) *nieto/a*'

/tʃ/ vs /dʒ/: /'(β)t**ʃa**/ 'full, *llena*', /'d**ʒâ**/ 'I am full, (*estoy*) *lleno*'

/l:/ vs /l/: /'βl:ǎs/ 'lift, *levanta*', /'lǎs/ 'slim, *delgado*'

It should be noted that in much of the literature, fortis sonorants are frequently written as a double consonant /nn/ or /ll/ instead of /n:/ or /l:/. In this paper, these fortis consonants will be written as /n:/ and /l:/ respectively instead to provide a more accurate phonetic representation of the phenomenon. The preceding consonant does not affect the length of the following vowel; only the following vowel.

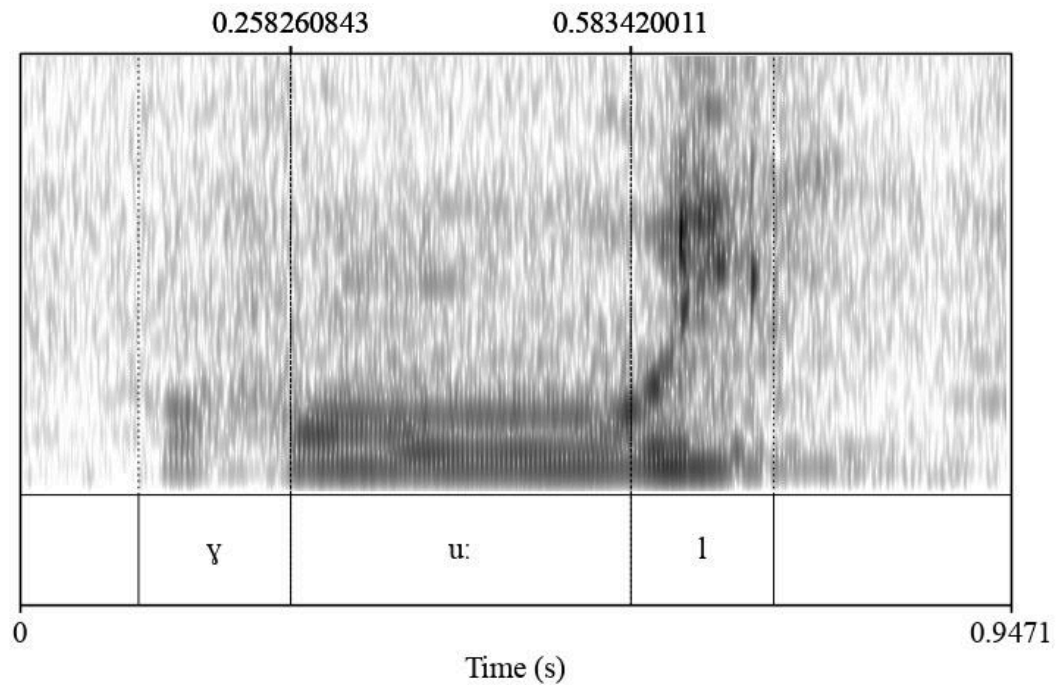
The complementary distribution of long and short vowels in stressed syllables with fortis and lenis consonants and in open syllables is shown below:

V: / C_ C_[+lenis]

V / C_ C_[+fortis]

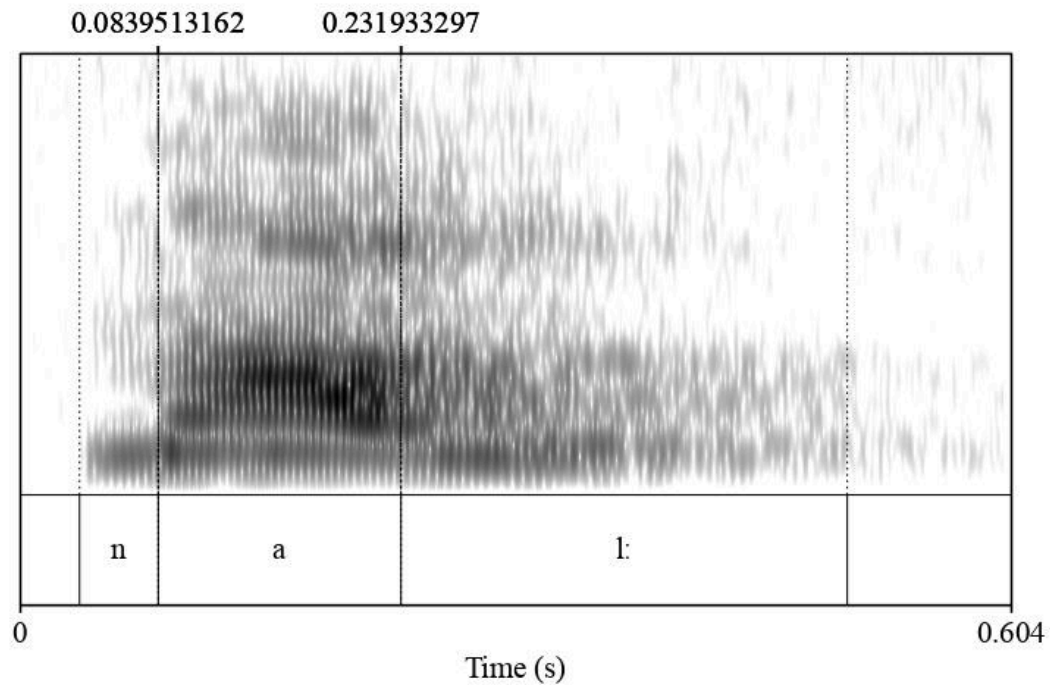
V: / C_]σ

Figure 2. Spectrogram of a lengthened modal vowel before a lenis consonant in /'yulj/ 'to be born, *nacer*' by female speaker Carmen Santiago



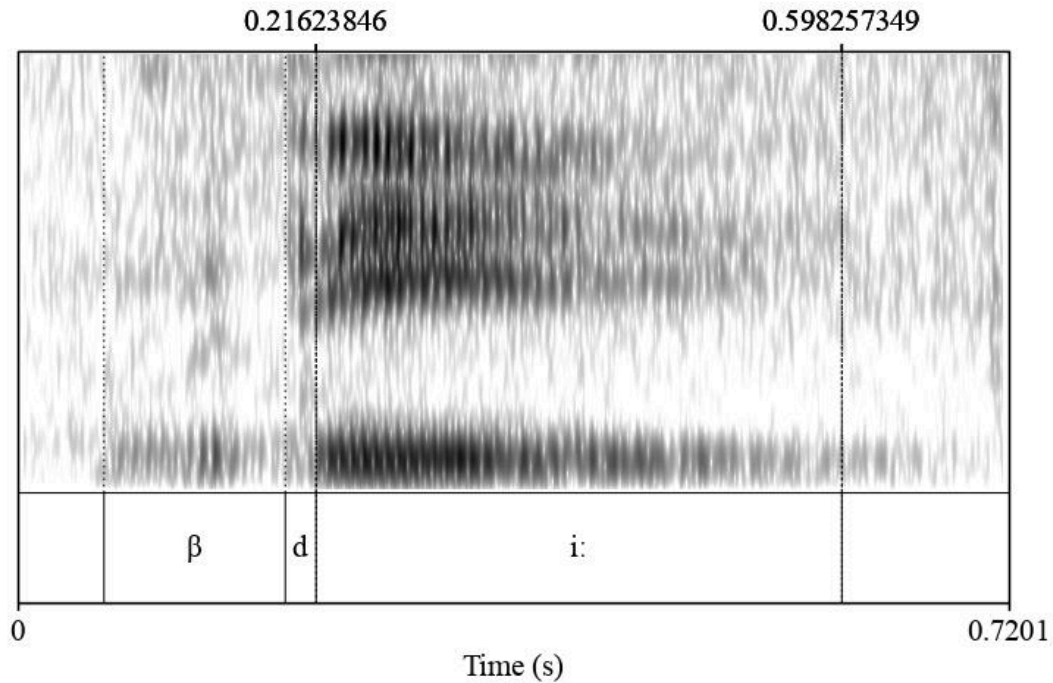
In the figure above, we can see the modal vowel /u/ before the lenis consonant /l/. This word is narrowly transcribed as ['yu:lj]. Because this vowel occurs before a lenis consonant, based on the environments written above, it is lengthened to /u:/. The length of this vowel according to Praat is 0.32516 seconds, which is longer than the vowel in the figure shown below:

Figure 3. Spectrogram of a non-lengthened modal vowel before a fortis consonant in /'nal:/ 'cold, *frio*' by female speaker Francisca García Hernández



In the figure above, we can see the modal vowel /a/ before the fortis consonant /l:/. Because this vowel occurs before a fortis consonant, based on the rule written above, the vowel will not be lengthened and will stay a “normal” length. The length of this vowel according to Praat is 0.14798 seconds, which is shorter than the lengthened vowels in Figure 3 above and Figure 4 below:

Figure 4. Spectrogram of a lengthened modal vowel word-finally in /'βdi/ 'ant, *hormiga*' by male speaker Humberto Hernández García



In the figure above, we can see the modal vowel /i/ occurring word-finally. This word is narrowly transcribed as [ˈβdi:]. Because this vowel occurs without a consonant in coda position, based on the rules written above, it is lengthened to /i:/. The length of this vowel according to Praat is 0.38202 seconds, making it comparable in length to the vowel in Figure 2 above and significantly longer than the vowel in Figure 3 above.

3.1.2 Non-native consonants and the influence of Spanish

The phonemes /f/ and /r/ occur in the language as well, but only in loanwords from Spanish such as /'flêtʃ/ 'arrow, *flecha*' and /re'βâs/ 'left/the other way around, *izquierda*/(al) *revés*'.

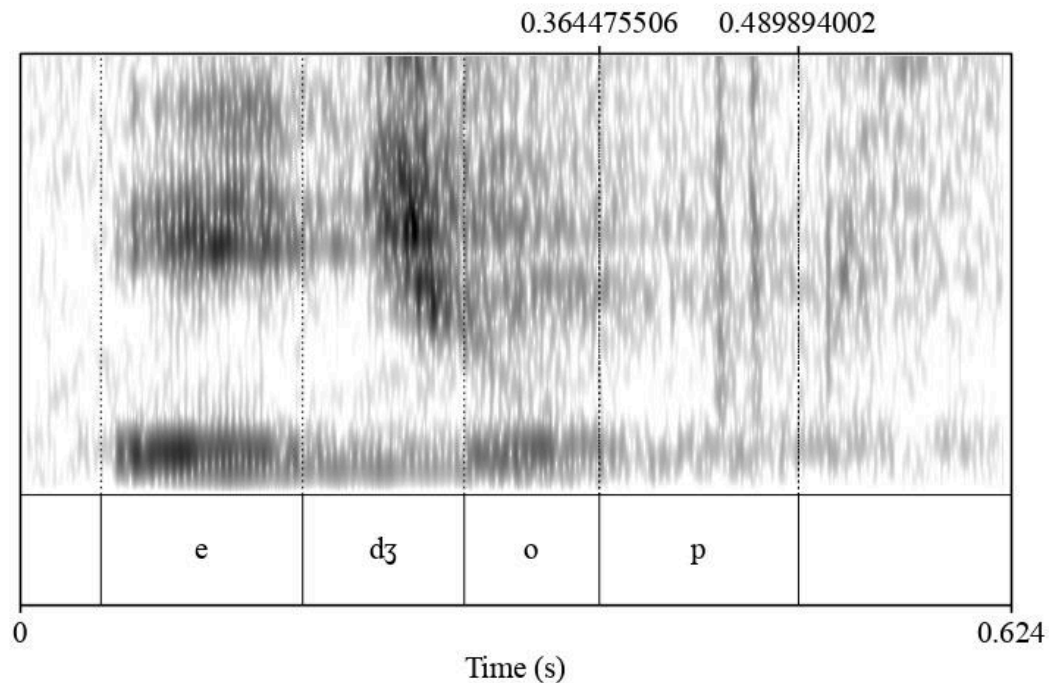
In this thesis, Spanish translations and SMdVZ orthography are provided in italics while English translations are written in standard font. English translations come first followed by Spanish translations and if applicable, Zapotec orthography. This is because Spanish is the

second language (or first language) for many of the speakers of SMdVZ and narratives and word lists were designed in a Spanish-Zapotec style, meaning that the Spanish translation is the closest equivalent to the elicited SMdVZ word or narrative. English is not a language spoken by an overwhelming majority of speakers in the SMdV community, so Spanish was necessary to communicate and collaborate with speakers. English translations are simply provided because this thesis is written in English and for ease of understanding for those who do not speak Spanish.

3.1.3 Palatalization

Palatalization is a common feature in my database, but for the purposes of this thesis, palatalization is defined here as a consonant acquiring a secondary palatal articulation or a consonant followed by a palatal consonant /j/ (Hume 1994). Sequences of a consonant followed by a palatal consonant /j/ are common in SMdVZ. It seems to occur on consonants in coda position both syllable-finally and word-finally. Based on the data used in this thesis, I analyze this phenomenon as a palatal consonant /j/ following a non-palatalized consonant (/C-j/) and not as a palatalized consonant (C^j), as there doesn't seem to be an environment that triggers palatalization to occur. There is no consistent environment where an underlying consonant would become palatalized, so I analyze these as sequences of consonant-glides rather than as palatalized consonants. As of right now, there are no minimal pairs between a given consonant and consonant followed by a palatal consonant /j/ that would lead me to believe that palatalization creates a difference in meaning. The figure below illustrates palatalization occurring word finally:

Figure 5. Spectrogram of a palatalized consonant in /e.'dʒopj/ ‘to suck, *chupar*’ by female speaker Francisca García Hernández



In the spectrogram above, we can see the palatalization of the final consonant /p/ through the light formants shown after the release of the final consonant, narrowly transcribed as [e.'dʒopj]. Glides and palatalization often show up in this way on spectrograms due to the fact that they have vowel-like qualities, which is why we see light formants appearing at the end of the word.

Palatalization in this language is not limited to just obstruents or lenis consonants like the /p/ shown in the figure above; it appears to occur with any possible consonant. However, not all consonants in word-final position are followed by /j/ and there is no consistent predictor of when a word-final consonant should be followed by /j/.

3.2 Vowels

SMdVZ has seven phonemic vowels, as shown below in Table 5. In Table 6, minimal and near minimal pairs are provided to illustrate the phonemic status of each vowel. It should be

noted that the under tilde /ŷ/ represents laryngealization and the glottal superscript /Vʔ/ represents glottalization. These phonation types will be discussed in more detail in later sections.

Table 5. Vowel phonemic inventory

	Front	Mid	Back
High	/i/	/i̠/	/u/
Mid	/e/		
	/æ/		/o/
Low		/a/	

A phonetic [ɛ] occurs as well, often appearing as an allophone of /æ/ in free variation.

The presence of a mid central vowel has been posited, but further research is necessary to determine the phonemic status of this vowel. Below is a list of minimal and near minimal pairs to illustrate the phonemic status of each vowel, with adjacent vowel pairs listed first:

/i/ vs /i̠/: /'zɪtj/ 'onion, *cebolla*', /'dzɪtj/ 'bone, *hueso*'

/i/ vs /e/: /'βdi/ 'ant, *hormiga*', /'ndě/ 'that, *ese*'

/i̠/ vs /u/: /'βñ:j/ 'people, *gente*', /'βʒû:n:/ 'run, *correr*'

/u/ vs /o/: /'dɯ/ 'rope, *cuerda/mecate*', /'do/ 'spike, *espiga*'

/e/ vs /æ/: /'βě/ 'broom, *escoba*', /'βæn/ 'do! (IMP), *haz*'

/æ/ vs /a/: /'βæn/ 'do! (IMP), *haz*', /'βan:j/ 'mud, *lodo*'

/a/ vs /o/: /'dǎnj/ 'hill, *cerro*', /'do/ 'spike, *espiga*'

/i/ vs /u/: /'βdi/ 'ant, *hormiga*', /'βduf/ 'barks, *ladra*'

/i/ vs /æ/: /'βin:j/ 'grain (to sow), *grano (para sembrar)*', /'βæn/ 'do! (IMP), *haz*'

/i/ vs /a/: /'diβj/ 'wash, *lava*', /'tap/ 'four, *cuatro*'

/i/ vs /o/: /'βdi/ 'ant, *hormiga*', /'do/ 'spike, *espiga*'

/i/ vs /e/: /'dzit/ 'bone, *hueso*', /'yět/ 'below, *debajo*'

/i/ vs /æ/: /'βin/ 'people, *gente*', /'βæn/ 'do! (IMP), *haz*'

/i/ vs /a/: /'dzit/ 'bone, *hueso*', /'natj/ 'slow, *lento*'

/i/ vs /o/: /'βinj/ 'people, *gente*', /'βō/ 'coal, *carbón*'

/u/ vs /e/: /'du/ 'rope, *cuerda/mecate*', /'dět/ 'rainbow, *arcoíris*'

/u/ vs /æ/: /'βzûn:j/ 'run, *corre*', /'βæn/ 'do! (IMP), *haz*'

/u/ vs /a/: /'du/ 'rope, *cuerda/mecate*', /'dănj/ 'hill, *cerro*'

/e/ vs /a/: /'dět/ 'rainbow, *arcoíris*', /'dănj/ 'hill, *cerro*'

/e/ vs /o/: /'dět/ 'rainbow, *arcoíris*', /'do/ 'spike, *espiga*'

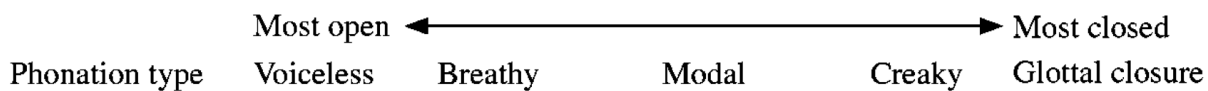
/æ/ vs /o/: /'βzǎβj/ 'orphan, *huérfano*', /'do/ 'spike, *espiga*'

4. Suprasegmental phonology

4.1 Vowel quality (phonation)

Ladefoged (1971) suggests that there is a continuum in phonation types, which is illustrated below in Figure 6. This continuum “is defined in terms of the aperture between the arytenoid cartilages, ranging from voiceless (furthest apart), through breathy voiced, to regular, modal voicing, and then on through creaky voice to glottal closure (closest together)” (Gordon and Ladefoged 2001). Zapotec is a language in which phonological contrasts are created through the use of different phonation types.

Figure 6. Continuum of phonation types (after Ladefoged, 1971)



Like other Zapotec varieties, SMdVZ has multiple phonation types. SMdVZ has three distinct voice qualities that create a contrast in meaning: modal, laryngealized, and glottalized. In other literature on Zapotec languages, laryngealized phonation is also known as “creaky” and glottalized phonation is known as “checked”. For the purposes of this thesis, I will use the terms “laryngealized” and “glottalized” when referring to these phonation types. These contrasts are illustrated in the list below through minimal and near minimal pairs. They are also illustrated visually through Praat analysis with laryngealized versus modal in Figures 7 and 8, and laryngealized versus glottalized in Figures 9 and 10.

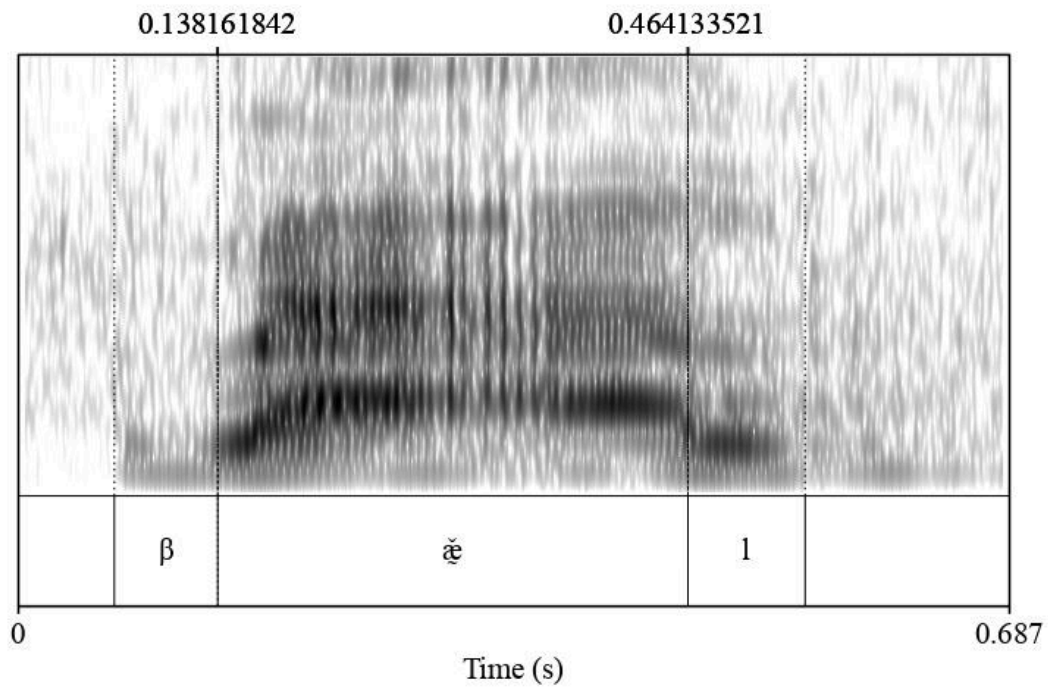
According to Figure 6, laryngealized phonation is a subtype of the glottalized phonation type. The laryngealized phonation type is characterized by multiple consecutive glottal closures, resulting in a “creaky” voice. In contrast, the glottalized phonation type is characterized by one single glottal closure, resulting in a “checked” voice. The ultimate difference between the two phonation types is the amount of glottal closures, making the laryngealized phonation type a subtype of the glottalized phonation type, though not allo-phonation types of the same phonation type. If we following according to Ladefoged’s continuum provided in Figure 6, the far right of the continuum would be the complete, singular glottal closure present in the glottalized phonation type, while the multiple, repeated, smaller glottal closures present in the laryngealized phonation type would be toward the right of the continuum past “modal”, but not quite as far as “glottal closure”. However, in SMdVZ and many other Zapotec varieties, these two phonation types are distinct both phonetically and phonologically, illustrated below in Figures 7, 8, 9, and

10.

Table 6. Vowel quality (phonation) contrasts

Phonation type	SMdVZ		
Modal	/ˈdo/ ‘spike, <i>espiga</i> ’		/ˈβæɫ/ ‘fire, <i>fuego</i> ’
Laryngealized	/ˈdu/ ‘rope, <i> cuerda/mecate</i> ’	/ˈβjǎ/ ‘nopal, <i>nopal</i> ’	/ˈβǎɫ/ ‘meat, <i>carne</i> ’
Glottalized	/ˈoroʔ/ ‘grow/to grow, <i> crece/crecer</i> ’	/ˈβjǎʔ/ ‘dance, <i>baila</i> ’	

Figure 7. Spectrogram of a laryngealized vowel in /ˈβǎɫ/ ‘meat, *carne*’ by female speaker Carmen Santiago



We can see during the vowel portion of the word in Figure 7 that there are multiple

vertical striations in the spectrogram, indicating repeated glottal closures, meaning that the vowel is laryngealized. Compare this with the following Figure 8, in which the vowel does not have any vertical striations (no glottal closures), meaning that the vowel is modal:

Figure 8. Spectrogram of a modal vowel in /'βæɫ/ 'fire, *fuego*' by female speaker Carmen Santiago

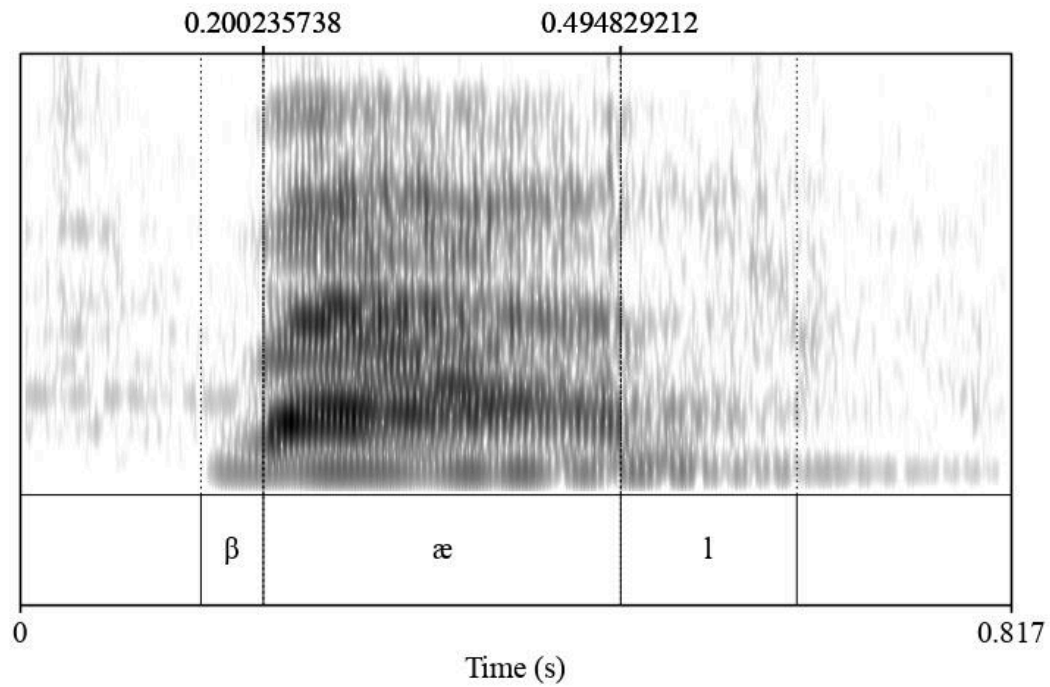
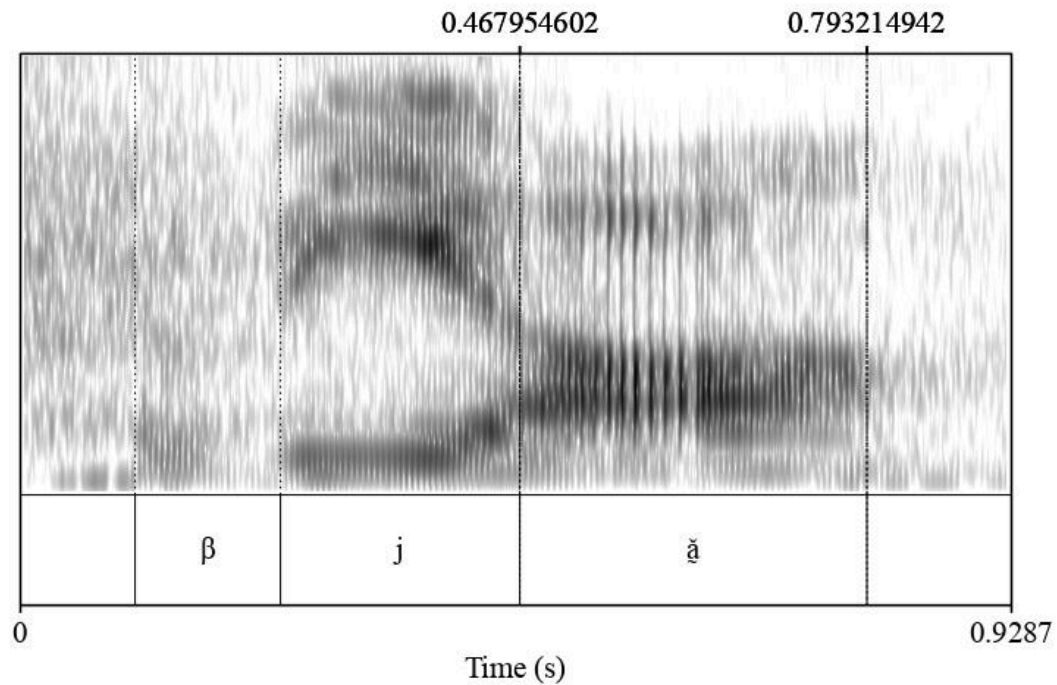
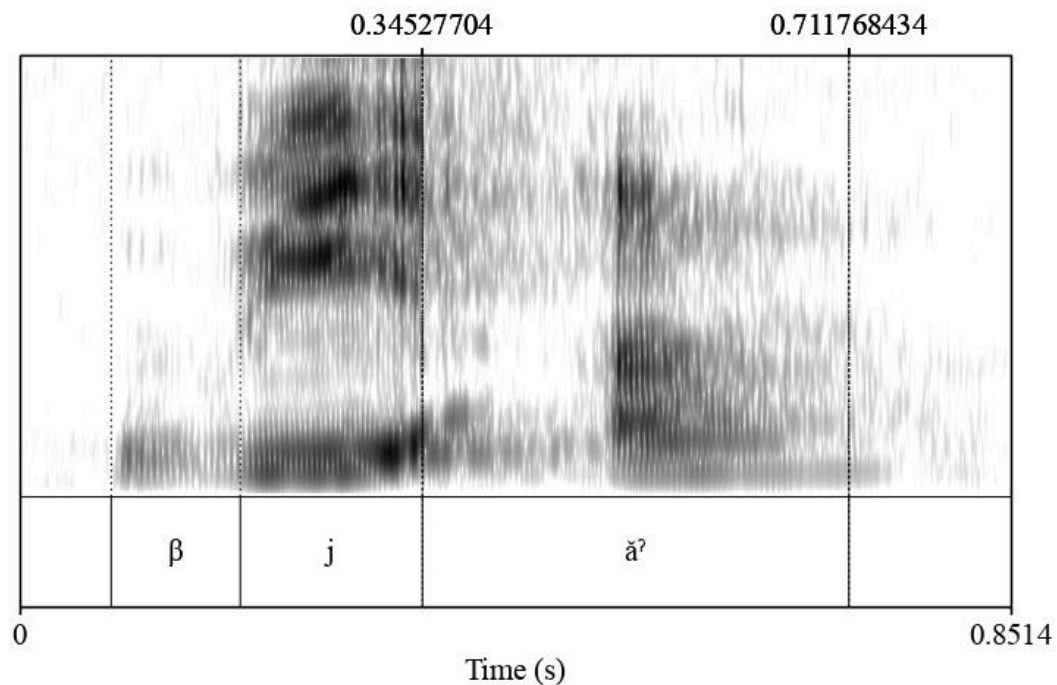


Figure 9. Spectrogram of a laryngealized vowel in /'βjǎ/ 'nopal, *nopal*' by female speaker Carmen Santiago



Again, here we see multiple repeated vertical striations over the course of the vowel in Figure 9, indicating laryngealization. On the other hand, we can contrast this with the following Praat spectrogram in Figure 10, which has one singular glottal closure, indicating glottalized phonation:

Figure 10. Spectrogram of a glottalized vowel in /'βjǎʔ/ 'dance, *baila*' by female speaker Francisca García Hernández



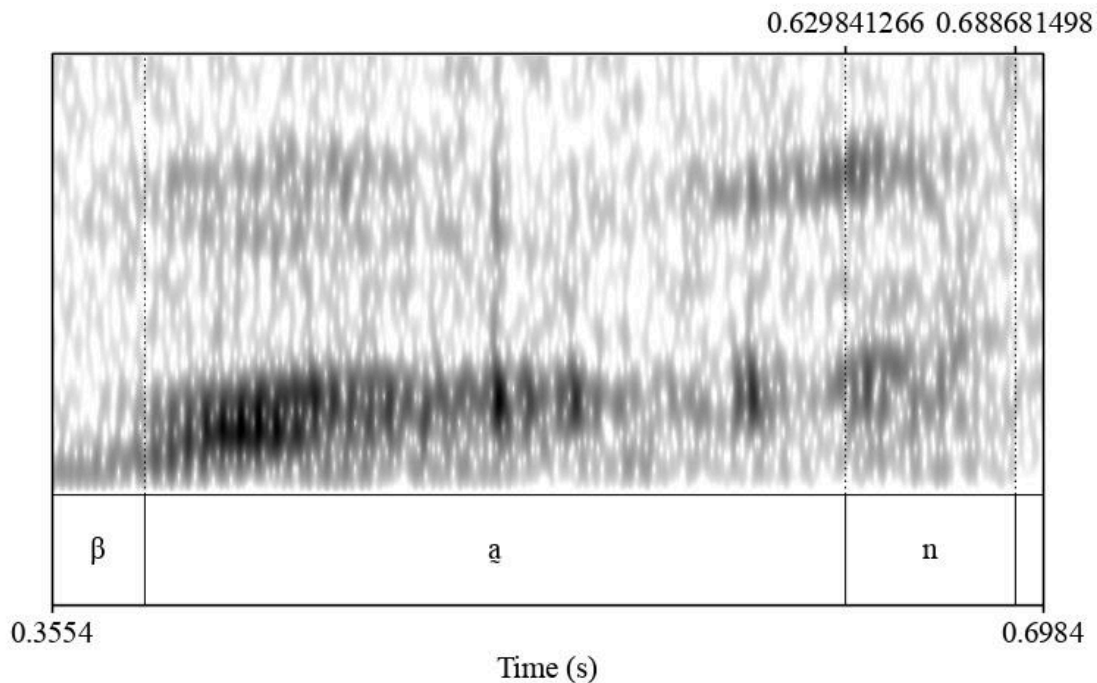
The distribution of glottalized vowels and laryngealized vowels seems to be complementary in certain contexts; namely, in closed syllables. Glottalized vowels precede fortis consonants in closed syllables (as in San Pablo Guilá Zapotec), while laryngealized vowels precede lenis consonants in closed syllables. In these contexts, the distinction between the two phonation types is lost. Modal vowels can occur in both of these contexts too. Therefore, glottalized and laryngealized vowels are only contrastive in open syllables. These contrasts are illustrated below in Figures 11 and 12.

Environments to describe the environments where laryngealized versus glottalized vowels occur:

$$V \rightarrow \underset{?}{V} / C_C_{[+lenis]}$$

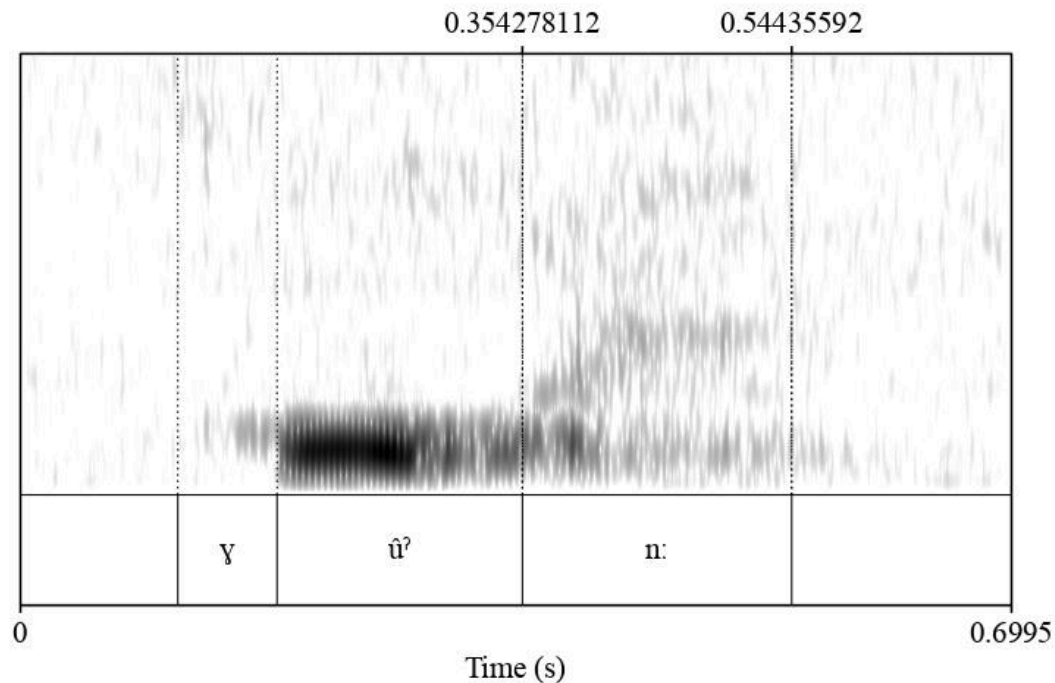
$$V \rightarrow V^? / C_C_{[+fortis]}$$

Figure 11. Spectrogram of a laryngealized vowel followed by a lenis sonorant in /do. 'βan/ 'leather strap, *mecapal*' by female speaker Carmen Santiago



In this figure, we can see that the laryngealized vowel is followed by a lenis consonant /n/. This illustrates that in closed syllables, laryngealized vowels are always followed by lenis consonants. The fortis and lenis distinction between sonorant consonants in Central Valley Zapotec varieties is mainly duration, with fortis consonants being long and lenis consonants being shorter in duration. The duration of this lenis /n/ consonant is 0.05884 seconds long, contrasting with the following fortis /n:/ in the following figure:

Figure 12. Spectrogram of a glottalized vowel followed by a fortis consonant in /'yû²n:/ 'bull, *toro*' by female speaker Francisca García Hernández



In this figure, we can see that the glottalized vowel is followed by a fortis consonant /n:/. This demonstrates that in closed syllables, glottalized vowels are always followed by fortis consonants. The duration of this fortis /n:/ consonant is 0.19008 seconds long, indicating that it is a fortis consonant.

4.1.1 Lack of a phonemic breathy phonation type

Unlike some other varieties of Zapotec, such as San Lucas Quiavini Zapotec (SLQZ), SMdVZ does not have a phonemic breathy phonation type (Pérez Báez and Uchihara 2016). This is illustrated through words in SLQZ with breathy voice quality that do not have breathy voice quality in SMdVZ. In Figure 13 below, we can see that the Praat spectrogram shows no signs of breathy phonation.

SLQZ /təp/ vs. SMdVZ /'tap/ 'four, *cuatro*'

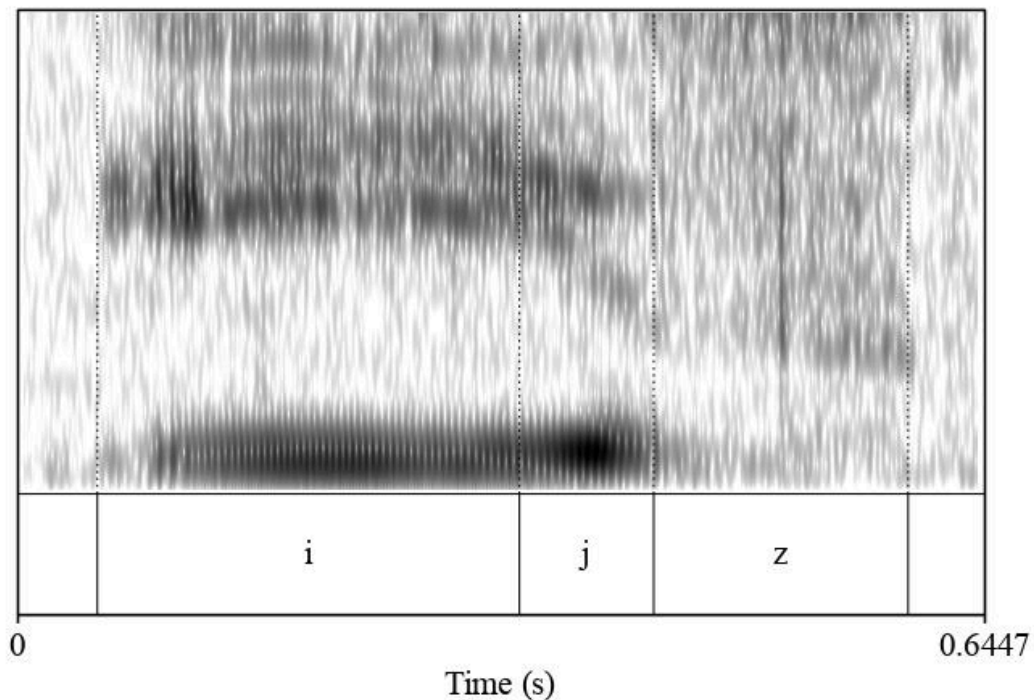
SLQZ /nəs/ vs. SMdVZ /'n:əs/ 'day before yesterday, *anteayer*'

SLQZ /nâf/ vs. SMdVZ /'naʃ/ ‘chocolate/sweet, *chocolate/dulce*’

SLQZ /nâ:/ vs. SMdVZ /'nâ:/ ‘now, *ahora*’

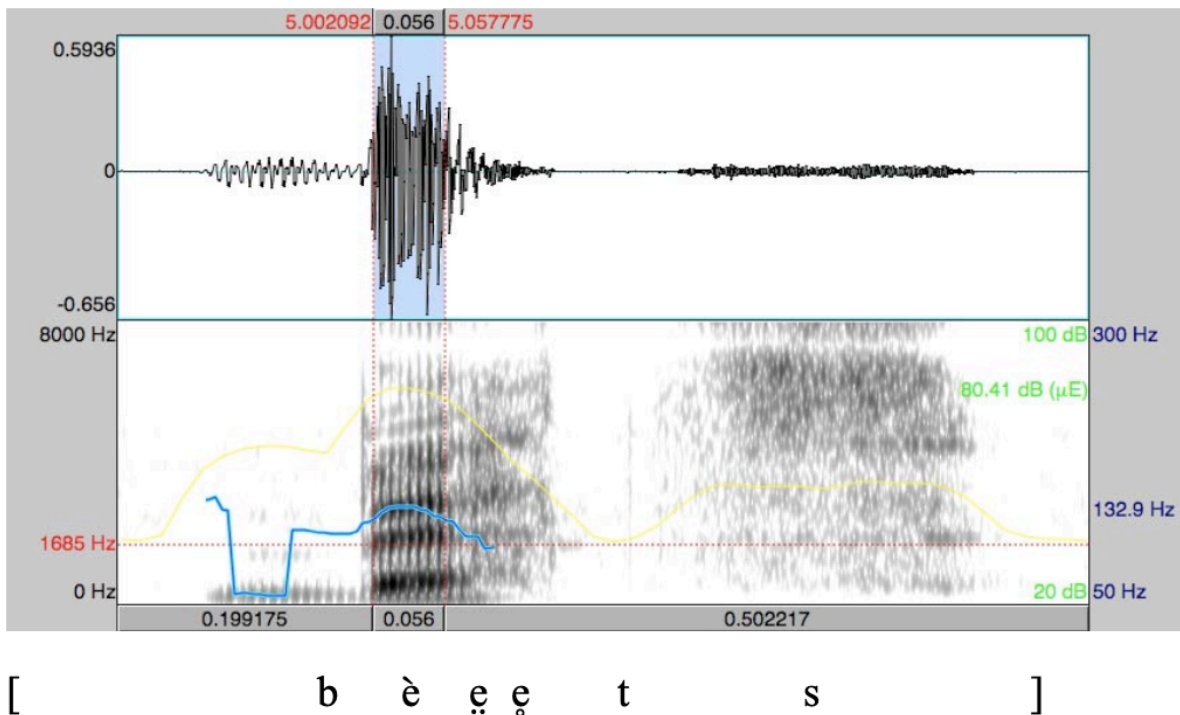
It appears that words with the breathy phonation type in other varieties of Zapotec are realized with a modal phonation type in SMdVZ. In SLQZ, the word ‘year, *año*’ is transcribed broadly as /iãz/, with the second part of the diphthong being breathy (Pérez Báez and Uchihara 2016). If the same word in SMdVZ were to be breathy, we would expect to see the voicing bar (which is indicative of voicing) at the bottom of the spectrogram for either both part of the vowel-glide sequence or at least half of the vowel-glide sequence (most likely on the glide portion of the sequence to be consistent with the SLQZ transcription of the word). However, we do not see this demonstrated in the following figure:

Figure 13. Spectrogram of a modal vowel in /'ijz/ ‘year, *año*’ by female speaker Francisca García Hernández in SMdVZ



In the figure above, we can see that the voicing is maintained throughout the entire duration of the vowel, indicated by the voicing bar at the bottom of the spectrogram. Meanwhile, in the following spectrogram in Figure 14, we can see a breathy vowel for a portion of the duration of the vowel, indicated by the absence of the voicing bar:

Figure 14. Spectrogram and waveform of a breathy vowel in /'bɛts/ ‘(man’s) brother, *hermano (de un hombre)*’ by female speaker in SLQZ (Pérez Báez and Uchihara 2016)

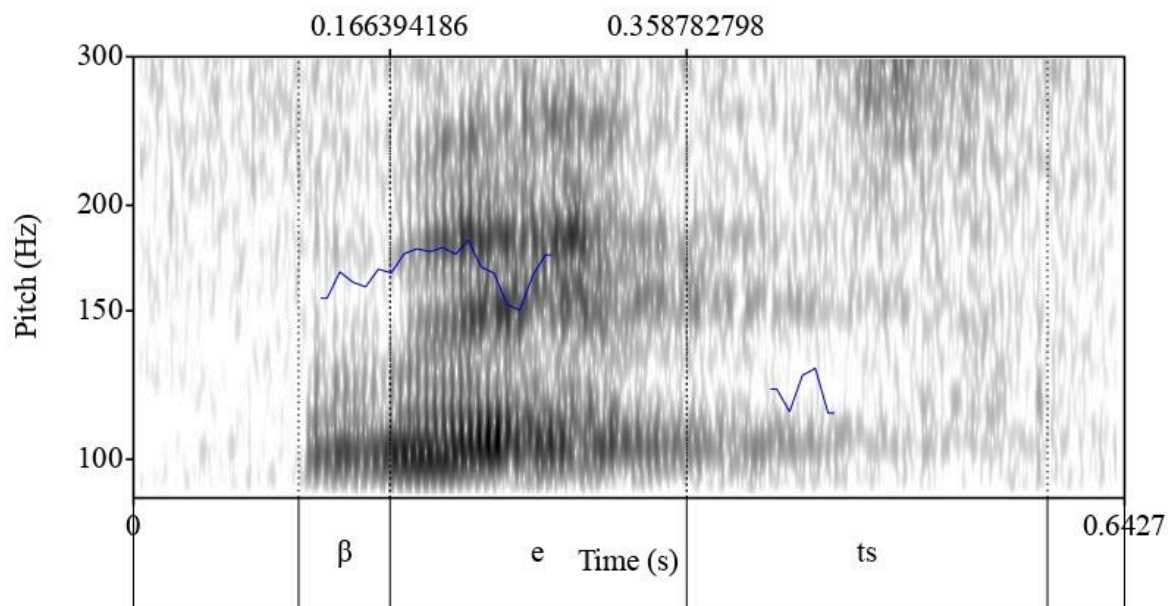


In comparing Figures 13 and 14 above, we can see that the former shows consistent, strong vertical striations throughout the duration of the vowel, while the latter shows a strong vowel at first, followed by lighter formants, indicating breathy phonation. This demonstrates that while the breathy phonation type occurs in SLQZ, it does not occur in SMdVZ in many words.

There is evidence of a phonetic breathy phonation type shown below in Figure 15. While this breathiness occurs, it is not distinct from the modal phonation type, making it a purely phonetic phenomenon in this language. Speakers produce the phonetic breathiness in some

words, while they consistently use a modal phonation type instead of a breathy phonation type in other words, as seen previously in Figure 13. This phonetic breathiness used in some words is also produced interchangeably with the modal phonation type, making the breathy phonation type not contrastive in this language.

Figure 15. Spectrogram of a phonetic breathy vowel in /'βets/ 'brother, *hermano*' by female speaker Francisca García Hernández in SMdVZ



In the figure above, we can see that the vowel looks similar to the vowel shown previously in Figure 14, demonstrating that this is a phonetic breathy vowel ['βets]. The formants and voicing bar both start out strong at the onset of the vowel, but eventually fade before the beginning of the following affricate. Like in SLQZ, this word is produced with breathy voice, but unlike SLQZ, this breathy voice is not contrastive and thus is not a phonemic phonation type in SMdVZ.

4.2 Diphthongs

Diphthongs are not phonemic in SMdVZ. Phonetically, they occur, but I analyze them rather as a glide-vowel or vowel-glide sequence. For the purposes of this thesis and assessing the status of diphthongs, I analyze these glide-vowel and vowel-glide sequences the same.

Sequences of two different vowels or vowel + glide (/j/ or /w/) appear to occur in front of lenis consonants in closed syllables, in front of fortis consonants, and in coda position:

(1) a. /'kojn:j/ 'finger, *dedo*' b. /'njes/ 'water, *agua*' c. /'ɣaj/ 'five, *cinco*'

(2) a. /'ɣwôz/ 'hunt, *caza*' b. /'ɣwê/ 'guava, *guayaba*'

In other varieties of Zapotec such as TdVZ diphthongs do not occur and are analyzed as a glide-vowel sequence through the assignment of tone to the glide-vowel sequence (Gutiérrez Lorenzo 2021). This will be discussed in SMdVZ later in this section. I am fairly certain that diphthongs are not present in this language, though more analysis from above the word level is required in order to determine the status of diphthongs in SMdVZ. It appears that these glide-vowel and vowel-glide sequences can occur with either glide /j/ or /w/ and any vowel besides /w/ and /u/. As far as I can tell, there are no /wu/ or /uw/ sequences in this language. This is illustrated by the blank cell in the table below. Examples of each pairing are shown below in Table 7.

Table 7. Possible glide-vowel and vowel-glide sequences

	/j/	/w/
/u/	/'juʒ/ 'sand, arena'	
/o/	/'joʔ/ 'house, casa'	/'ɣwots/ 'say something to him, di algo a él'
/æ/	/'βsjæ/ 'eagle, águila'	/'ɣi.'twæz/ 'figleaf gourd, chilacayote'
/e/	/'jeʃ/ 'avocado, aguacate'	/'ɣwe/ 'go! (IMP), anda'
/i/	/'ji.'βa/ 'sky, cielo'	/'ɣwitʃ/ 'shell explodes, cascarón estalla'
/i/	/'stɪj/ 'another, otro'	/'βwidj/ 'domestic bird, ave doméstica'
/a/	/'jaŋ/ 'cob, olote'	/'lwâ/ 'my face, mi cara'

These vowel-glide/glide-vowel sequences are shown below in Figures 16 and 17 with formants highlighted in the spectrograms to show the transition from vowel to glide or glide to vowel:

Figure 16. Spectrogram of a glide-vowel sequence in /'zjûl/ 'long, *largo*' by female speaker Francisca García Hernández

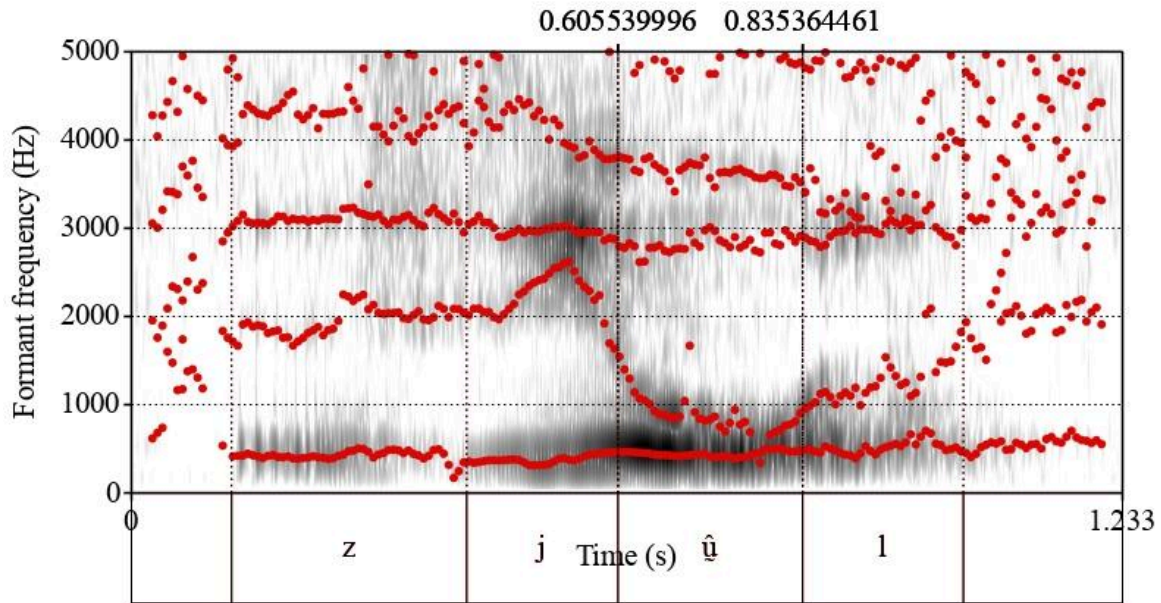
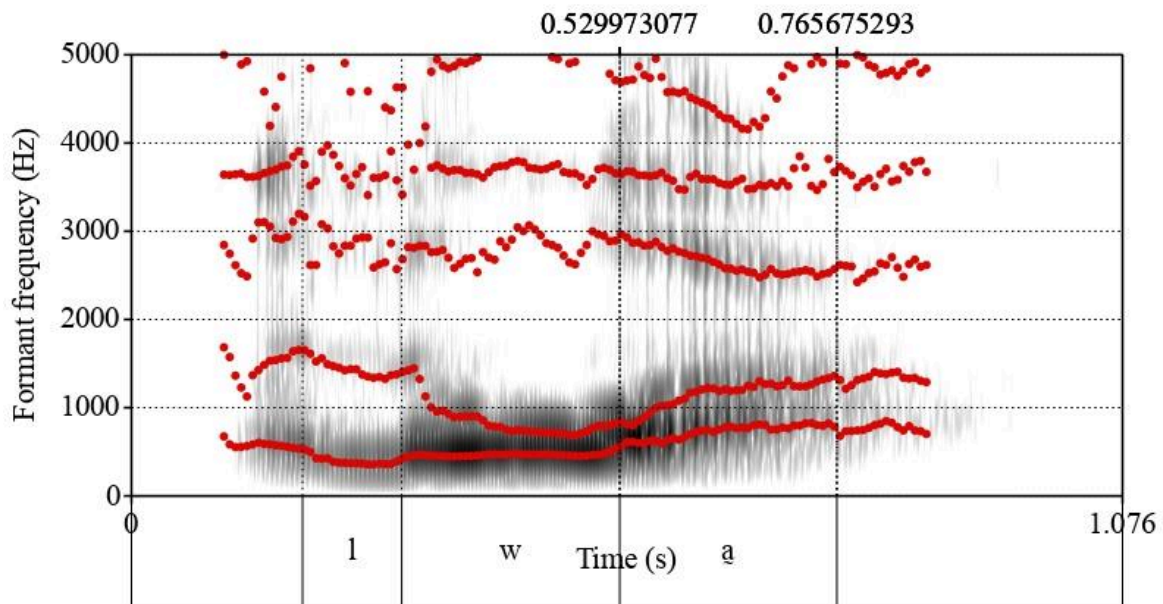


Figure 17. Spectrogram of a glide-vowel sequence in /'lwa/ 'my face, *mi cara*' by female speaker Carmen Santiago

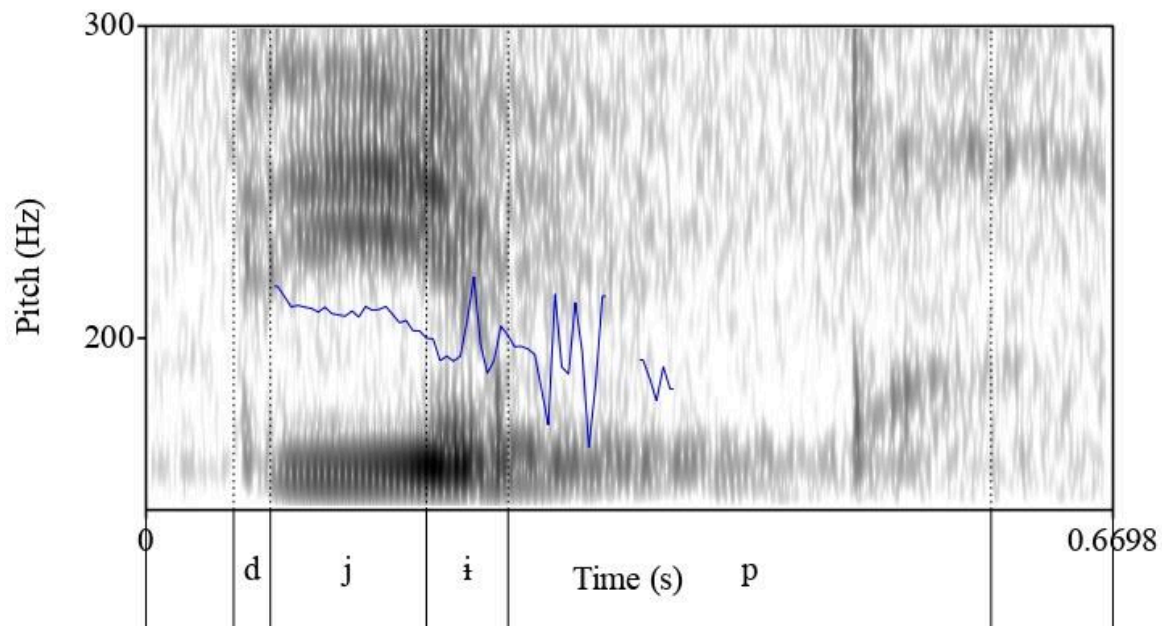


From the figures above, we can see that the formants shift throughout the duration of the glide-vowel sequence, indicating a potential change in vowel throughout the duration.

Interestingly, it should be noted that the laryngealization only affects the vowel portion of the glide-vowel sequence, which can be seen through the repeated vertical striations in the /a/ portion of the spectrogram and the lack thereof during the /w/ portion of the spectrogram. This seems to be evidence for these sequences to be interpreted as glide-vowel or vowel-glide sequences, as if it were to be interpreted as a vowel-vowel sequence, the phonation (in this case, laryngealization) should carry on throughout the entirety of the vowel-vowel sequence, not just one part.

It should be noted that another key reason why diphthongs are not likely to exist in this language is because of the interaction of tone and the glide-vowel sequences. If diphthongs were able to exist, it should be possible for two different tones to occur on the two vowels that make up the diphthong. However, this never occurs and the same tone is carried on throughout the vowel-glide or glide-vowel sequence. This is demonstrated in the following figure:

Figure 18. Spectrogram of a low tone in a glide-vowel sequence in /'djip/ 'strong, *fuerte*' by female speaker Francisca García Hernández



Here, we can see that the pitch is consistent throughout the entire duration of both the glide and the following vowel, meaning that the vowel carries the tone that the glide shares. A glide-vowel sequence carrying two different tones has not been attested for. The glide is not capable of carrying its own tone in addition to the tone carried by the following vowel, meaning that the glide should not be analyzed as being part of a diphthong and instead should be analyzed as a separate glide.

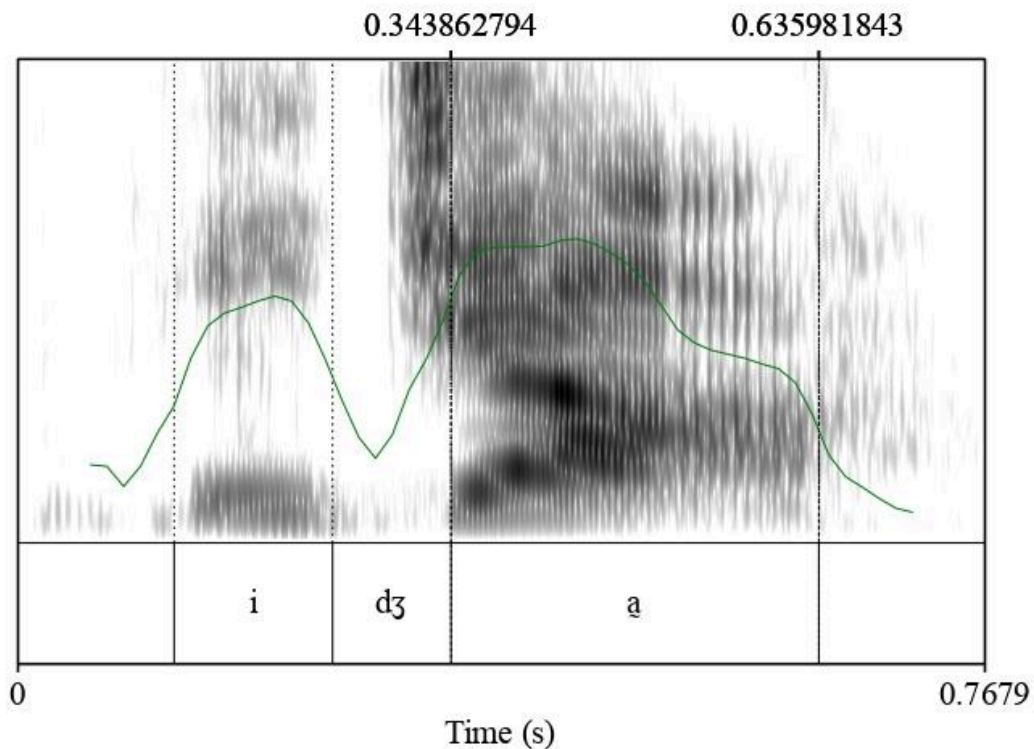
4.3 Stress

Stress is known to be a suprasegmental feature of many Zapotec languages (Chávez Peón 2010; Gutiérrez Lorenzo 2021; Arellanes 2009; Beam de Azcona 2023). Stress is defined here as prominence of a syllable through phonetic means, such as vowel duration and intensity. Of

course, pitch would not be an indicator of stress in this language, as pitch is already meaningfully involved in the production of tone, which will be covered in the following section.

In Zapotec languages, stress usually falls on the root of a verb or noun. Because of this, stress usually falls on the second syllable of a word, as the root is generally preceded by a prefix in a multisyllabic word. Phonetically, stress is indicated through lengthened vowels and higher intensity in the stressed syllable. For example, in Figure 19 below, a disyllabic word is given to illustrate the role of both vowel duration and intensity of the ultimate syllable, indicating that it is stressed:

Figure 19. Spectrogram of the word /i.'dʒa/ ‘animal, *animal*’, demonstrating stress that falls on the ultimate syllable, by speaker Francisca García Hernández



In the figure above, intensity is indicated by the green line. We can see that the maximum intensity value in the second syllable is 81.38 dB, while the maximum intensity in the first

syllable is 75.45 dB, indicating that the second syllable reaches a higher intensity. In this particular noun, the second syllable is the root of the word while the first syllable is a prefix attached to the front of the root. Because of this, the second and final syllable of the word carries prominence.

We can also see that the length of the vowel in the second syllable is longer in duration compared to the vowel in the first syllable, with the former being 0.2921 seconds long and the latter being 0.1259 seconds long. This again indicates that the stress falls on the second syllable of the word.

4.4 Tone

Tone is considered here to be the production of pitch to distinguish lexical or grammatical meaning. According to Yip, tones occur in 60-70% of the world's languages (2002, p. 1). A tonal language can be defined as “a language with tone is one in which an indication of pitch enters into the lexical realisation of at least some morphemes” (Hyman 2006).

I propose that SMdVZ has three phonemic tones, including low (L), rising (ʌ), and falling (V). There seems to be a phonetic mid tone, which I analyze to be an allotone of the rising tone, which will be discussed in further detail in a later on in the section. See Figure 21 for a phonemically rising tone reanalyzed as a mid tone. There is also a phonetic high tone which has only been attested for in unstressed syllables in multisyllabic words and is considered to be an allotone of the falling tone. See Figure 23 for a phonemically falling tone reanalyzed as a high tone. This will be discussed in further detail later on in the section.

For the purposes of this thesis, the low tone is represented orthographically and in IPA transcription without a diacritic *a* or /a/. The mid tone is represented orthographically and in IPA transcription with the diacritic *ā* or /ā/. The high tone is represented orthographically and in IPA

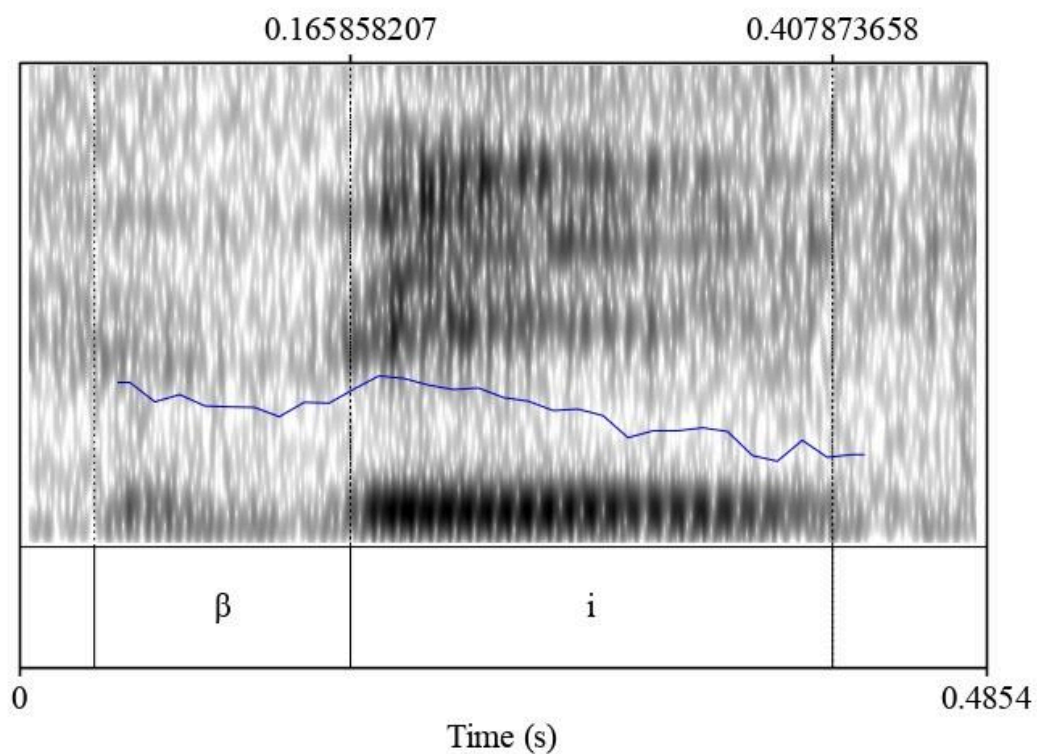
transcription with the diacritic *á* or /*á*/. The rising tone is represented orthographically and in IPA transcription with the diacritic *ǎ* or /*ǎ*/. The falling tone is represented orthographically and in IPA transcription with the diacritic *â* or /*â*/.

A set of example tones were used to determine the tones of unknown words. These example tone words were selected through comparison with other varieties and analyzed through Praat to determine their pitch ranges. Once the pitch range had been determined for both male and female speakers, these example tone words were assigned tones and compared to other words with unknown tones. If another word by the same speaker happened to follow the same tone pattern as an example word, the same tone pattern would be assigned to the word in question. In this way, I assigned tones to all of the words that my collaborators provided. I also used Pike's method of categorizing tones by first phonetically grouping tones and then further analyzing them (Pike 1949).

Each tone falls within a certain range that varies depending on the gender of the speaker. According to my data, for female speakers, the low tone falls within the range of 221 to 152 Hz while for male speakers, the range is 109 to 92 Hz. I only had two male collaborators for this dataset, so this range could be due to the fact that I had a limited number of participants. In the future, I would like to work with more male speakers. For female speakers, the falling tone falls within the range of 239 to 153 Hz and 164 to 84 Hz for male speakers. Interestingly, the falling and low tones are generally within the same range in terms of pitch, but what differentiates the two is the range of pitch covered throughout the tone and the length of the vowel associated with the tone. Figure 20 below illustrates the length of an open-syllable low tone word and Figure 25 later on in the section illustrates the length of an open-syllable falling tone word to compare the length of their respective vowels. Generally speaking, the low tone does not cover as much of a

range throughout the vowel while the falling tone starts from a higher pitch and falls to a lower pitch compared to the low tone. So while the low tone is phonetically falling, it is phonemically distinct from the falling tone due to the range of pitch covered throughout the tone. Additionally, the low tone is generally shorter in duration compared to the falling tone.

Figure 20. Spectrogram of a low tone in an open-syllable in the word /'βi/ 'air, *aire*' by speaker Humerto Hernández García



In comparing the figure above, we can see that the and low tone is phonetically falling, indicated by the pitch tracker, but the duration of the vowel differs, with the low tone in Figure 20 being notably shorter at 0.2420 seconds and the falling tone in Figure 24 being longer at 0.4498 seconds.

The high and mid tones are phonetically level tones, meaning that they do not cover a change in pitch over the course of the vowel; they are phonetically stable. However, they cover

different pitch ranges: the high tone falls between 225 to 257 Hz for female speakers and between 124 to 127 Hz for male speakers.

The high tone has only been attested for in unstressed syllables in multisyllabic words. It is uncommon compared to other tones in terms of frequency. I hypothesize that it is an allotone of the falling tone, as they have a complementary distribution: the high tone only occurs in unstressed syllables, while the falling tone cannot occur in unstressed syllables. For this reason, the high tone is not phonemically distinct from the falling tone, making it an allotone that only occurs in specific conditions, namely in unstressed syllables, usually occurring word-initially. It should be noted that this high tone cannot be an allotone of the rising tone because the rising tone is attested for in stressed syllables, putting them in contrastive distribution. The falling tone is realized as a high tone in unstressed syllables, which is demonstrated in the rule below. I would like to explore the distribution of the falling and high tones in further research, as stress seems to be the only factor in determining the distributions of the allotones currently.

Rule to describe the distribution of the falling tone:

$\hat{V} \rightarrow \acute{V} / \text{unstressed syllables}$

The rising tone phonetically rises from a lower pitch to a higher pitch, the only tone in SMdVZ to do so. This tone rises from around 230 to 297 Hz for female speakers and from around 109 to 115 Hz for male speakers. It should be noted that the rising tone does not cover as much of a range compared to the other contour tone (falling).

As of right now, there is no definitive tonal phonological process (such as tone Sandhi) that leads us to believe that there is a phonological contrast between a mid tone and a low tone, but more research is needed to determine the phonological status of this mid tone.

Historically, the mid tone in TdVZ split off from the rising tone in a phenomenon called

mid tone splitting. We see the same thing occurring in SMdVZ, as the mid tone currently alternates with the rising tone, and is sometimes used interchangeably with the rising tone within the same word by the same speaker. While the two tones are completely distinct in TdVZ, they seem to be in the process of divergence in SMdVZ. In Figure 21, we see an example of this phenomenon in the word /'zā/ 'cloud, *nube*'. Phonologically, this word should have a rising tone. However, it is reanalyzed as a mid tone in this context, demonstrating that the rising and mid tones are used interchangeably by speakers in certain words. Because of this, the mid tone is thought to be an allotone of the rising tone, as it is not phonemically distinct from the rising tone. Over time, we might expect to see this mid tone eventually break off from being an allotone of the rising tone and become its own phonemic tone, as we've seen in other varieties of Zapotec. However, because the mid tone is used interchangeably with the rising tone in certain situations, it is analyzed here as an allotone of the underlying rising phonemic tone. It should be noted that there are still some words in which the rising tone is always analyzed as a rising tone and others where the rising tone might be reanalyzed as a mid tone.

Because I have only looked at single words and not at the interaction between words, it is not yet possible to know if this mid tone triggers tone Sandhi, as it does in other varieties of Zapotec (Gutiérrez Lorenzo 2021). In SMdVZ, it appears that some words are always produced with a rising tone and never a mid tone, while others tend to be produced with a mid tone and occasionally a rising tone. This indicates that this mid tone reanalysis occurs in specific environments and that the mid tone is indeed splitting off from the rising tone. The specific environments in which this mid tone reanalysis occurs is not yet known; further exploration is needed to determine which environments trigger this mid tone reanalysis.

In Figures 21, 22, 23, 24, and 25, a pitch tracker (the blue line) is used in Praat analysis to

illustrate the different phonological tones of SMdVZ. In Table 8, multiple minimal and near minimal pairs are provided to illustrate the phonemic status of each tone.

Figure 21. Spectrogram of the phonetic mid tone represented in the word /'zǎ/ ['zā] 'cloud, *nube*' by speaker Carmen Santiago, a female speaker

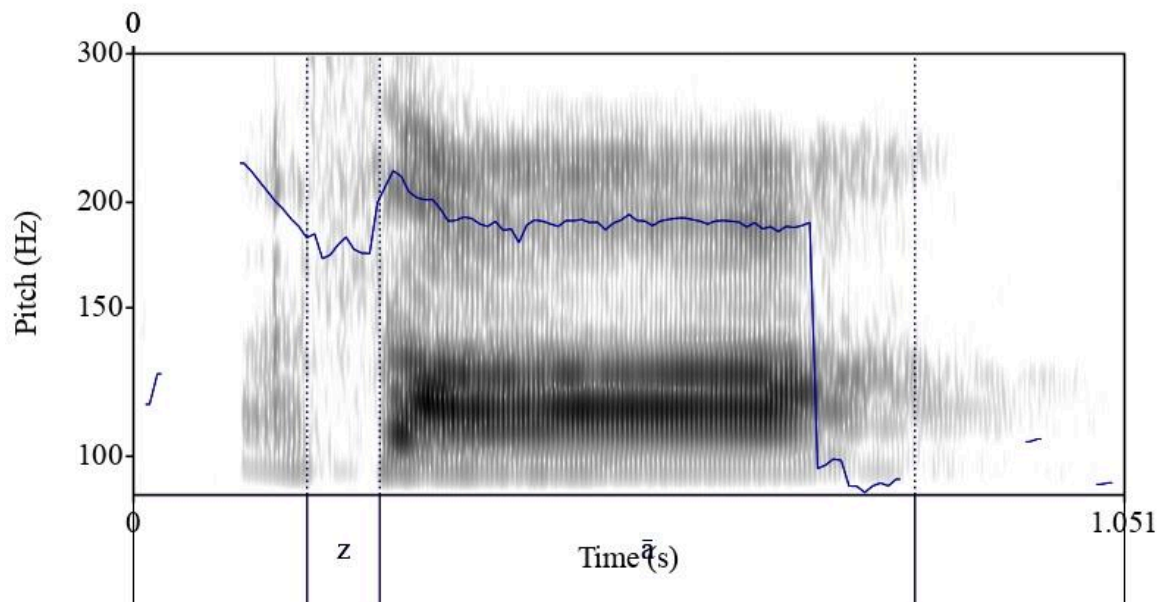
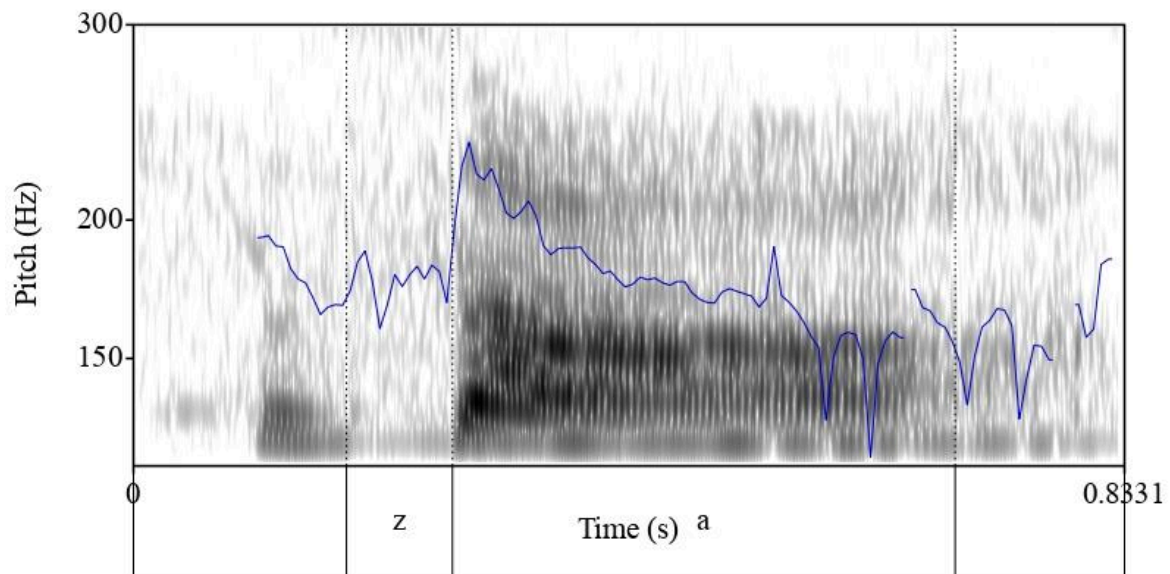


Figure 22. Spectrogram of the phonemic low tone represented in the word /'za/ 'fat, *grasa*' by speaker Carmen Santiago, a female speaker



It should be noted that the low tone is relatively short compared to other tones and is phonetically falling, but is realized as a low tone compared to the true falling tone.

Figure 23. Spectrogram of the phonetic high tone represented in the first syllable of the word /tú.ˈnē/ ‘who, *quién*’ by speaker Carmen Santiago, a female speaker

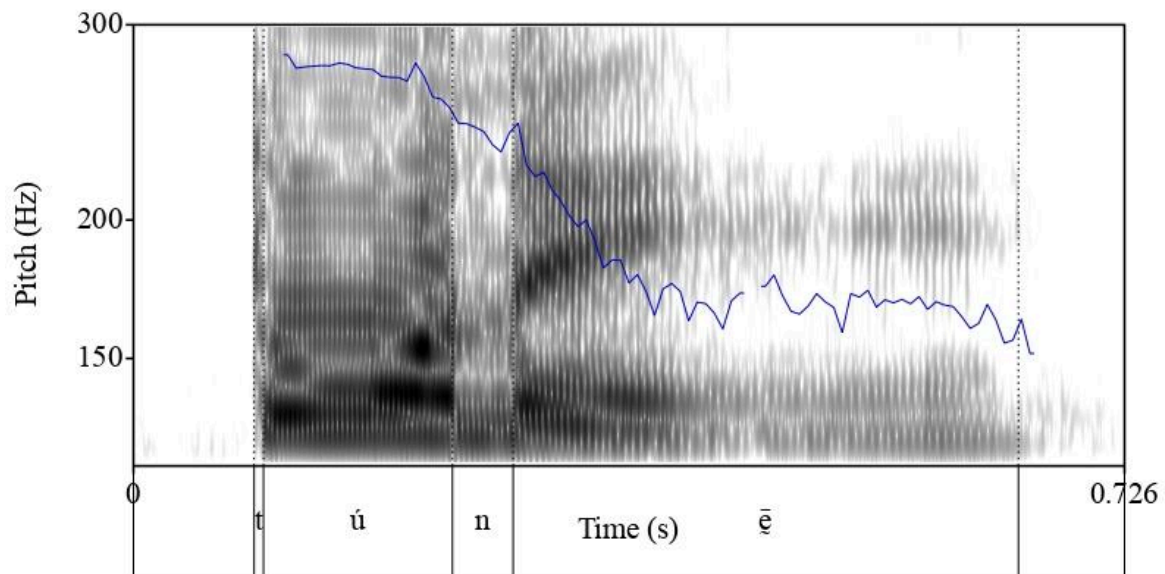


Figure 24. Spectrogram of the phonemic rising tone represented in the word /ˈdǎnj/ ‘hill, *cerro*’ by speaker Francisca García Hernández, a female speaker

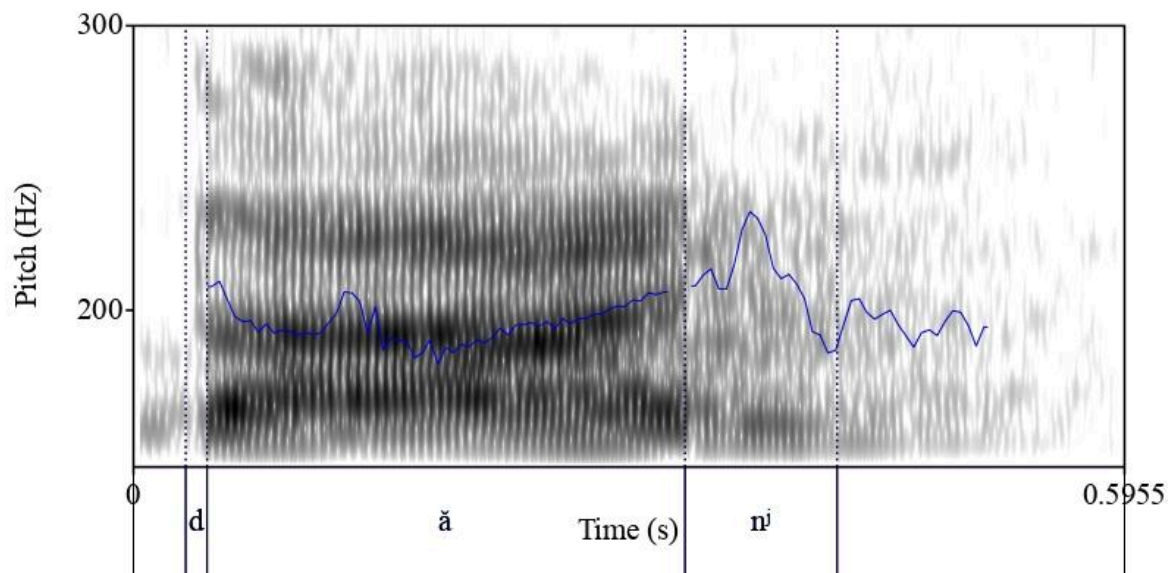
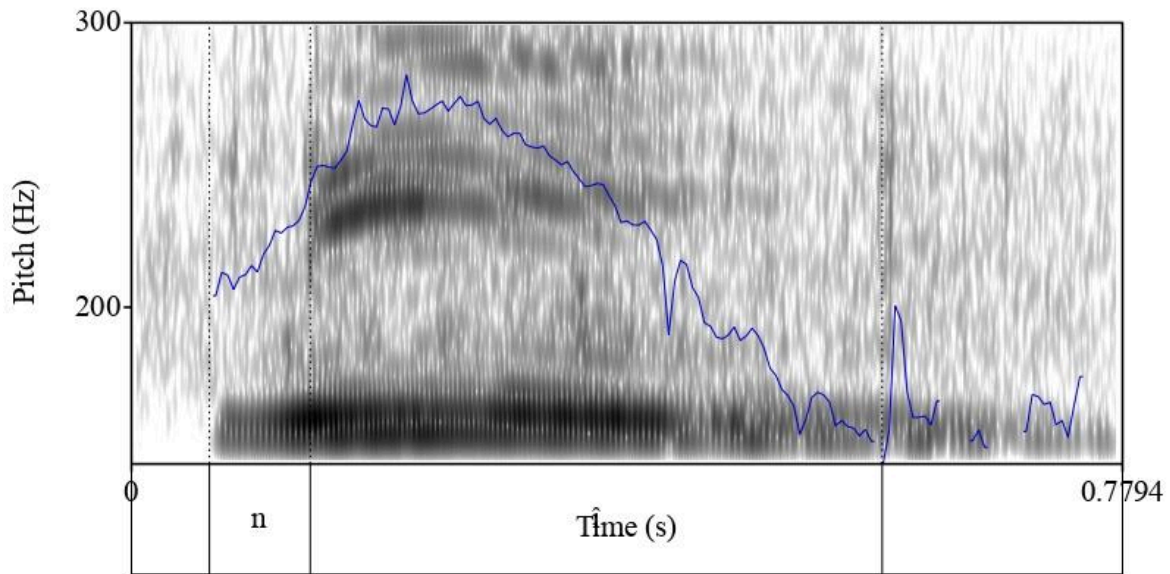


Figure 25. Spectrogram of the phonemic falling tone represented in the word /'nî/ 'sour, *agrio*' by speaker Carmen Santiago, a female speaker



The falling tone starts from a higher pitch and falls to a lower pitch compared to the low tone.

Table 8. Tone contrasts

Tone	SMdVZ			
Low	/'yɛt/ 'tortilla, <i>tortilla</i> '	/'zɑ/ 'fat, <i>grasa</i> '	/'kuβj/ 'mass, <i>masa</i> '	/'βæɫ/ 'fire, <i>fuego</i> '
Rising	/'yɛt̃/ 'below, <i>abajo</i> '	/'zã/ 'cloud, <i>nube</i> '		/'βæɫ̃/ 'meat, <i>carne</i> '
Falling			/'kũβj/ 'new, <i>nuevo</i> '	

5. Conclusions

I found that SMdVZ is similar to other Central Valley Zapotec languages in its phonemic inventory, especially to Teotitlán del Valle Zapotec (Gutiérrez 2021). Similar to other Central Valley Zapotec varieties, SMdVZ has the fortis and lenis distinction, twenty-two phonemic consonants, and seven phonemic vowels. It also has three distinct tones, including low, rising, and falling, with two allotones: the high tone as an allotone of the falling tone and the mid tone as an allotone of the rising tone. It also makes a distinction between fortis and lenis consonants, with vowels lengthening before lenis consonants and not lengthening before fortis consonants (though vowel length is not contrastive). Additionally, there is a three-way distinction between phonation types; modal vowels contrast with laryngealized vowels and glottalized vowels.

6. Further research

There are three main areas of research that require further investigation; the status of the mid and high tones, palatalization, and stress. Though a phonetic mid tone exists, there is no tone phonological process that currently leads us to believe that this mid tone is phonemic. This phonetic mid tone alternates with both the rising tone, with SMdVZ speakers using them interchangeably, often within repeated productions of the same word. Because of this, I analyze the mid tone as an allotone of the rising tone. However, because I have only looked into single word utterances, it is not possible to know if this mid tone triggers tone Sandhi across word boundaries, as it does in other varieties of Zapotec (Gutiérrez Lorenzo 2021). In the future, I would like to explore the relationships between tones across word boundaries in the narratives that collaborators have provided me with, as it would be reasonable to expect some sort of tone Sandhi to exist in this language. This would ultimately help me to determine the phonemic status of the mid tone. Because only certain words allow for the rising tone to be reanalyzed as a mid

tone, more research is required to determine the specific environments in which this reanalysis occurs. More data at both the word level and across word boundaries would be helpful in determining the environments that trigger this tone reanalysis.

In an earlier section, I posited that the high tone is an allotone of the falling tone, as they occur in complementary distribution, with the high tone occurring only in unstressed syllables and the falling tone unable to occur in unstressed syllables. I would like to further explore the distribution of the falling tone and its proposed high allotone, as so far, stress is the only indicator of their distributions.

I am also unsure on the status of consonant-glide sequences or if they should be analyzed instead as palatalized consonants. So far, there are no minimal pairs of a given consonant and a palatalized consonant that would lead me to believe that palatalization creates a difference in meaning. However, more research is required to ultimately determine the status of these possible palatalized consonants.

Another prosodic feature that requires further exploration is stress. It is known that Zapotec languages have stress, but analyzing these patterns went beyond the scope of this description. In the future, more research on the interaction of tone, stress, and voice quality is required to explore the prosodic nature of SMdVZ.

As noted in the Methodology section, the majority of my collaborators were female speakers; I only had two male speakers participate. In the future, I would like to work with a larger group of collaborators, namely more male speakers, to gain a bit more clarity into the pitch range for each tone.

The glottalized phonation type often sees vowels either having glottalization in the middle of the vowel or at the end of the vowel. Because of this, vowels that have the single

glottal closure in the middle of the vowel often have a rearticulated part of the vowel occur, while the vowels that have the glottal closure at the end of the vowel do not have this rearticulation. Though this rearticulation of the vowel does not create a contrast and result in a different meaning of a given word, I am unsure what causes it and would like to explore the contexts that trigger this rearticulation in further research.

Finally, I would like to explore the relationship between phonation type and tone in SMdVZ. It is possible for all three different phonation types to carry the different tones, but more research into this area is required to understand how the two influence each other. I predict that non-modal phonation types probably have an effect on the fundamental frequency (f_0) in some way, most likely lowering the measurement and making the perceived pitch lower.

References

- Arellanes, Francisco Arellanes. *El sistema fonológico y las propiedades fonéticas del zapoteco de San Pablo Güilá: Descripción y análisis formal*. El Colegio de México, 2009.
- Beam de Azcona, Rosemary G. *The historical dialectology of stative morphology in Zapotecan*. "Journal of Historical Linguistics 13, no. 1 (2023): 115-172.
- Beck, David. "Language loss and linguistic suicide: A case study from the Sierra Norte de Puebla, Mexico." (2008).
- Chávez Peón, Mario E. *The interaction of metrical structure, tone, and phonation types in Quiavini Zapotec*. Diss. University of British Columbia, 2010.
- Esposito, Christina. "WPP, No. 103: Santa Ana del Valle Zapotec Phonation." (2004).
- Esposito, Christina M. "Variation in contrastive phonation in Santa Ana del Valle Zapotec." *Journal of the International Phonetic Association* 40.2 (2010): 181-198.
- Esposito, Christina Marie. *Santa Ana del Valle Zapotec phonation*. Diss. University of California, Los Angeles, 2003.
- GARCIA, EDER SANTIAGO. "INTERACCIONES EN ZAPOTECO EN HOGARES BILINGÜES DE MITLA, OAXACA: UN ESTUDIO DE LOS NICHOS DE SOCIALIZACIÓN EN ABUELOS Y NIETOS." *Publicación impresa* (2015).
- Gordon, Matthew, and Peter Ladefoged. "Phonation types: a cross-linguistic overview." *Journal of phonetics* 29.4 (2001): 383-406.
- Gutiérrez Lorenzo, Ambrocio. *A description and analysis of the syntax and functions of*

subordinate clauses in Teotitlán del Valle Zapotec (with an introductory overview of TdVZ phonology and morphosyntax). Diss. 2021.

Hamel, Rainer Enrique. "Indigenous language policy and education in Mexico." *Encyclopedia of language and education* 1 (2008): 301-313.

Hume, Elizabeth V. *Front vowels, coronal consonants and their interaction in nonlinear phonology*. Routledge, 2018.

Hyman, Larry M. "Word-prosodic typology." *Phonology* 23.2 (2006): 225-257.

Jaeger, Jeri J., and Robert D. Van Valin Jr. "Initial consonant clusters in Yateé Zapotec." *International Journal of American Linguistics* 48.2 (1982): 125-138.

Munro, Pamela. "Zapotec grammar without tears (except perhaps for the grammarian)." (2004).

Nellis, Donald G. & Barbara E. Hollenbach. 1980. *Fortis versus lenis in Cajonos Zapotecphonology*. *International Journal of American Linguistics* 46. 92–105.

<https://doi.org/10.1086/465639>

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

Silverman, Daniel. 1997. Laryngeal complexity in Otomanguean vowels, *Phonology* 14, 235-262.

Uchihara, Hiroto, and Gabriela Pérez Báez. "Fortis/lenis, glides and vowels in Quiavini Zapotec." *Glossa: a journal of general linguistics* 1.1 (2016).

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