

## Introducing San Agustín Mixtepec Zapotec

Rosemary G. Beam de Azcona

University of California, Berkeley

### 1. Introduction & Endangerment status

With this paper I introduce a new language, San Agustín Mixtepec Zapotec (SAMZ). This paper contains background information on SAMZ and its speech community. I highlight SAMZ data for historical and comparative implications within Southern Zapotec (SZ), including both segmental and suprasegmental correspondences. As this is a working paper, I also feature two not-yet fully analyzed but intriguing problems in the synchronic phonology of SAMZ having to do with nasal-obstruent clusters and tonal morphology respectively. The information given in this paper is also especially valuable because of the lack of previous documentation and because of SAMZ's impending death.

SAMZ is spoken in the single town of San Agustín Mixtepec, in the ex-district of Miahuatlán in the Southern part of the state of Oaxaca, Mexico. SAMZ is today only spoken fluently (when spoken at all) by two elderly residents of the town for which the language is named. There are reportedly also two semi-speakers. The information in this paper comes from the youngest fluent speaker, who is now 72 years old. The findings presented here are very preliminary, the result of only three weeks' total fieldwork in 2002 and 2003<sup>1</sup>. More long-term fieldwork is planned for the next three years.

This community apparently agreed to stop speaking Zapotec in 1965. In addition to the undeserved stigma formally assigned to the language at that time, immigration out of SAM has no doubt been a key factor in this language's demise. Modern land disputes have contributed to the town's loss of resources and further loss of population. Yet the exodus out of SAM began much earlier with a dispute that divided the townspeople into two groups, with one group going South to found San Agustín Loxicha (SAL) in the ex-district of Pochutla (see <http://www.laneta.apc.org/rio/loxicha/historia.htm>). Both of these San Agustín towns have stories about the statue of the patron saint of each town being switched. Each version has the opposite town coming out the winner. In the SAM version this switch is the cause of the SAL's population increase and SAM's population decrease.

### 2. History, Geography & Classification

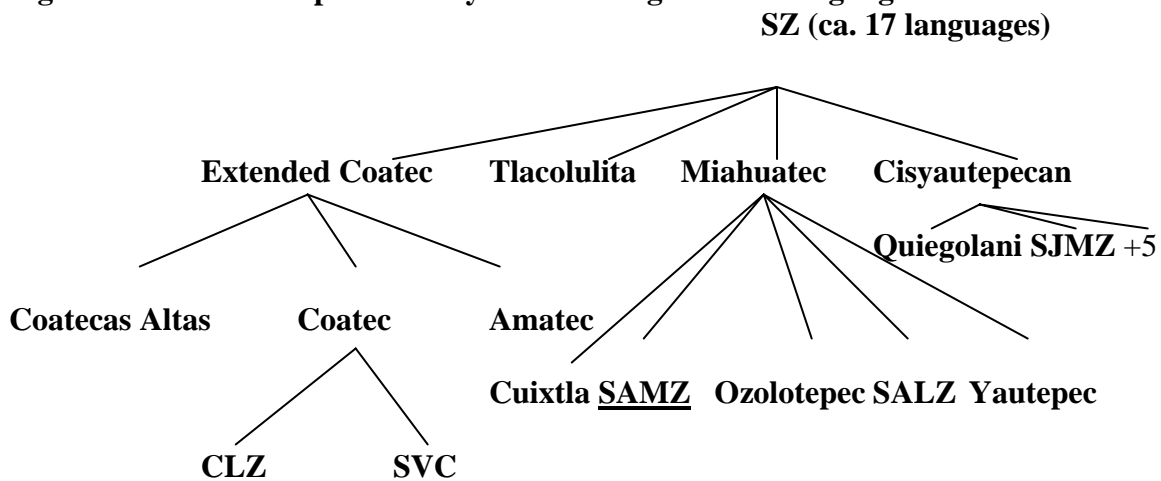
SAMZ belongs to the Miahuatec subgroup of Southern Zapotec (<Zapotecan < Otomanguan). According to Smith Stark (2003), SZ languages are defined by the addition of an /m/ at the beginning of words for animals and other words which take *b*- in other Zapotec languages. This renders the prefix *mb*- in Miahuatec languages such as SAMZ. Miahuatec languages differ from other SZ languages by having *t* and *d* reflexes of the Proto-Zapotec (PZ) phonemes reconstructed as \**ss* and \**s* (Kaufman, 2003).

According to Rojas (1950), SAM was founded by people from Miahuatlán, and the linguistic evidence supports this, although the town's enemies say that it was founded by people from Sola de Vega, an excuse for taking land away. Miahuatlán was supposedly founded by a group from Coatlán (Rojas, 1950), an assertion which, if accurate, would suggest that Miahuatec is most closely related to the Coatec group within SZ.

The Miahuatec group of languages gained significant ground with the southward expansion led by emigrants from SAM. The language of SAL, which must have developed from SAMZ, is today spoken vigorously in several towns near the Pacific coast, such as Cozoaltepec, Santo Domingo de Morelos, and many of the towns named Loxicha. Linguistic evidence of a more northerly origin for that southernmost Miahuatec language comes in the form of a borrowed placename. The SALZ name for the town of Santa María Colotepec, is clearly a borrowing from Coatlán-Loxicha Zapotec (CLZ), and specifically from a northern dialect of that language (Beam de Azcona, 2004), one which lies much closer to SAM than to SAL. Thus, SAMZ is a language which historically links the languages of the northern and southern parts of the Miahuatec region.

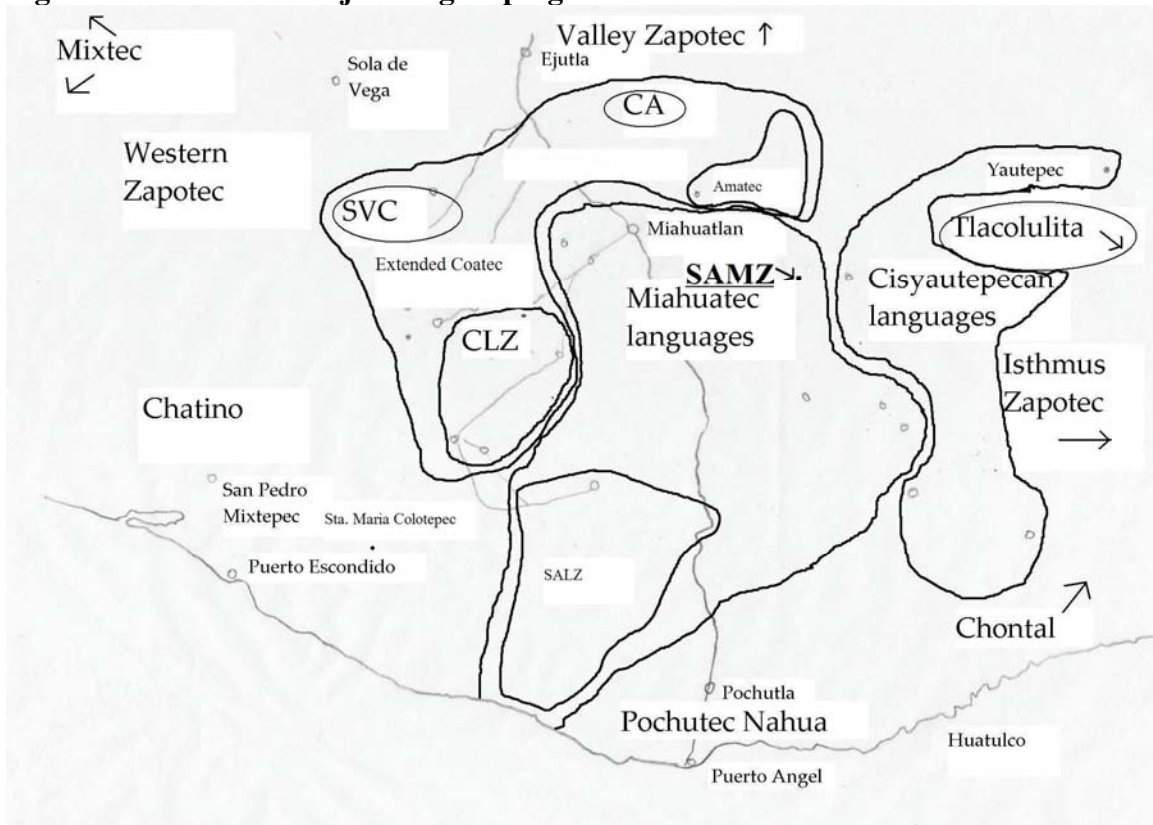
SAM is surrounded by and politically linked to towns where the use of other Zapotec languages is vigorous. SAM is subject to the municipality of San Cristobal Amatlán, where some 86.5% of the residents speak Amatec, an “extended Coatec” language also spoken in several other towns (Smith Stark, 2003). The similarly named language of San Juan Mixtepec Zapotec (SJMZ) also shares a geographic border with SAMZ and belongs to a third subgroup of SZ, Cisyau-tepecan. It is more closely related to Quiegolani, Xanica, Xanaguía, Quierí, Lapaguía and the very endangered Xadani Zapotec. Cisyau-tepecan languages are defined by Smith Stark as those which have lost the *b* and have only an *m*- prefix for most animal words. Many of the languages mentioned so far are shown in relation to SAMZ in Figure 1 (based on Smith Stark, 2003).

**Figure 1: Southern Zapotec family tree showing selected languages**



Each of the main groups shown in Figure 1 is located geographically, along with SAMZ, and some other individual languages and towns referenced here, in the map in Figure 2. The location of towns in this map is based on a map found in Rojas (1950) with supplementary information from the modern INEGI maps. The linguistic boundaries shown here are based on Smith Stark (2003) and on my own fieldwork with these languages.

**Figure 2: SAMZ and major subgroupings of SZ**



### 3. Sounds

Here I give a brief account of the phoneme inventory. Phonology is covered in 4.

#### 3.1 Segments

SAMZ has a six vowel system, as shown in Figure 3.

**Figure 3: SAMZ vowels**

i		u
e		o
ë [ɛ ~ æ]		
	a	

SAMZ has three types of obstruents: voiceless plosives<sup>2</sup>, voiceless spirants and voiced spirants. Voiced plosives do occur phonetically as part of the prenasalized plosive series which I tentatively posit here. My analysis of this series is as yet incomplete and I cover the details of the problem in 4.2.2. The full consonant inventory of SAMZ is shown in the practical orthography in Figure 4, with marginal and loan phonemes in parentheses

**Figure 4: SAMZ consonants**

	Labial	Dental	Alveolar	Palatal	Retroflex	Velar	Labiovelar	Glottal
<b>Voiceless plosives</b>	<b>p</b>	<b>t</b>	<b>tz</b>	<b>(ky)</b>	<b>ch</b>	<b>k</b>	<b>kw</b>	
<b>Voiceless spirants</b>		<b>th</b>	<b>s</b>		<b>x</b>			<b>(h)</b>
<b>Voiced spirants</b>	<b>b</b>	<b>d</b>				<b>g</b>		
<b>Prenasal -plosives</b>	<b>mb</b>		<b>ndz</b>		<b>ndzh</b>	<b>ng</b>	<b>ngw</b>	
<b>Nasals</b>	<b>m</b>	<b>n</b>				<b>(nh)</b>		
<b>Liquids</b>			<b>r, (rr)</b>	<b>(ly)</b>	<b>l</b>			
<b>Glides</b>	<b>w</b>			<b>y</b>				

### 3.2 Suprasegmentals

SAMZ has three tones: high, low, and rising. The SAMZ low tone is phonetically mid-to-low falling, the same as in the northern dialects of CLZ.

In addition to tone there is a three-way contrast involving glottalization: plain V, rearticulated V<sup>7</sup>V, and checked V<sup>7</sup>. These terms are fairly self-explanatory. Plain vowels lack glottalization. Rearticulated vowels are interrupted by a glottal stop which is followed by an echo of the same vowel (this is phonologically one segment, I do not mean to say that there is a vowel plus a second echo vowel in terms of the number of segments represented). Checked vowels are vowels that end in a glottal stop.

All three tones can occur on plain and checked vowels but only level tones can occur on rearticulated vowels. The pre-glottal stop portion of a checked vowel is longer in duration than the pre-glottal stop portion of a rearticulated vowel, and this is likely the reason why the language's single contour tone does not occur on rearticulated vowels. The possible combinations of tone and vowel type are shown with examples in Figure 5.

**Figure 5: Combinations of tone and glottalization in SAMZ**

	Plain vowels	Checked vowels	Rearticulated vowels
<b>Low tone</b>	<i>lâg</i> 'hoja; leaf'	<i>mbë<sup>7</sup>l</i> 'culebra; snake'	<i>yë<sup>7</sup>ël</i> 'hondura; swimming hole'
<b>High tone</b>	<i>lâr</i> 'ropa; clothing'	<i>xí<sup>7</sup>l</i> 'ala; wing'	<i>xí<sup>7</sup>il</i> 'algodón; cotton'
<b>Rising tone</b>	<i>ngǒ</i> 'huevo; egg'	<i>ngǒ<sup>7</sup>n</i> 'buey; ox'	Not attested

## 4. Phonology & Morphology

In this section I provide selected details of SAMZ phonology and morphology, including some interesting but unresolved morphophonological issues.

### 4.1 Possession

There is some interesting tonal morphology involved in marking a first person singular possessor on a possessed noun in this language, described in 4.1.2 below. Details of the morphology of possession marking are also potentially enlightening in a problem

of phonological analysis discussed in 4.2.3. In order to better understand the implications discussed below, I give a basic description of how possession works in this language.

#### 4.1.1 Segmental marking of possession

The morphology and syntax of possessed noun phrases found in SAMZ is fairly typical of Zapotec languages as a whole, though individual languages have variations on these patterns (as in CLZ, see Beam de Azcona, 2004). Possessors follow possessed nouns in these left-headed languages. Zapotec languages typically display a difference between alienable and inalienable possession. As in many of these languages (see, for example, Pickett et al., 1998) in SAMZ alienable possession is marked with an *x*-initial prefix. This pattern affects loanwords as well as native words, as shown in (1). Inalienable possession in SAMZ only differs by the lack of this prefix, as shown in (2), and does not make use of a separate noun or prefix as some languages do (e.g. CLZ).

- |     |              |                                       |                      |                                   |
|-----|--------------|---------------------------------------|----------------------|-----------------------------------|
| (1) | <i>kũch</i>  | <b>‘cochino; pig’</b>                 | <i>xkũch lè</i>      | <b>‘tu cochino; your pig’</b>     |
|     | <i>kwě̀l</i> | <b>‘totomoxtle; cornhusk’</b>         | <i>xkwě̀l mě́ yá</i> | <b>‘su totomoxtle; his husk’</b>  |
| (2) | <i>làb</i>   | <b>‘huarache; sandal’<sup>3</sup></b> | <i>làb lè</i>        | <b>‘tu huarache; your sandal’</b> |
|     | <i>bě̀l</i>  | <b>‘hermana; woman’s sister’</b>      | <i>bě̀l mě́ yá</i>   | <b>‘su hermana; her sister’</b>   |

As in other, for example Valley, Zapotec languages (see Suárez, 1973), nouns beginning in lenis consonants in the citation form have root variants with the corresponding fortis consonant after *x*-, as shown in (3).

- |     |             |                       |                 |                             |
|-----|-------------|-----------------------|-----------------|-----------------------------|
| (3) | <i>bě̀l</i> | <b>‘llama; flame’</b> | <i>xpě̀l nà</i> | <b>‘mi llama; my flame’</b> |
|     | <i>yèt</i>  | <b>‘olla; pot’</b>    | <i>xkèt lè</i>  | <b>‘tu olla; your pot’</b>  |

The nasal portion of an animacy prefix is often lost following *x*-, as shown in (4).

- |     |               |                            |                     |   |
|-----|---------------|----------------------------|---------------------|---|
| (4) | <i>mbèr</i>   | <b>‘guajolote; turkey’</b> | <i>xpèr mě́ yá</i>  | <b>‘su guajolote; her turkey’</b>         |
|     | <i>ngṍn</i>  | <b>‘buey; ox’</b>          | <i>xkṍn mě́ yá</i> | <b>‘su buey; his/her ox’</b>              |
|     | <i>ngwě̀n</i> | <b>‘dedo; finger’</b>      | <i>xkwě̀n lè</i>    | <b>‘tu dedo; your finger’<sup>4</sup></b> |

#### 4.1.2 Tonal marking of possession by the first person singular

In most Zapotec languages a floating high tone is added to verb roots with 1s subjects and noun roots with 1s possessors (see, for example, Beam de Azcona, 2004; Bickmore & Broadwell, 1998). In CLZ this floating H is added as a suffix:

- |     |               |                           |  |      |     |
|-----|---------------|---------------------------|--|------|-----|
| (5) | <i>líd</i>    | <b>‘casa; home’</b>       |  |      |     |
|     | <i>lĩd nâ</i> | <b>‘mi casa; my home’</b> |  |      |     |
|     |               |                           |  |      | CLZ |
|     |               |                           |  | home | 1s  |

In SAMZ the floating high tone affects both both types of possessed nouns equally.

- |     |                     |                        |                        |                        |
|-----|---------------------|------------------------|------------------------|------------------------|
| (6) | <i>kwě̀l</i>        | <i>xkwě̀l mě́ yá</i>   | <i>xkwě̀l nà</i>       | Alienable possession   |
|     | <b>‘totomoxtle’</b> | <b>‘su totomoxtle’</b> | <b>‘mi totomoxtle’</b> |                        |
|     | <b>‘cornhusk’</b>   | <b>‘his cornhusk’</b>  | <b>‘my cornhusk’</b>   |                        |
| (7) | <i>lě̀</i>          | <i>lě̀ lè</i>          | <i>lě́ nà</i>          | Inalienable possession |
|     | <b>‘nombre’</b>     | <b>‘tu nombre’</b>     | <b>‘mi nombre’</b>     |                        |
|     | <b>‘name’</b>       | <b>‘your name’</b>     | <b>‘my name’</b>       |                        |

In CLZ when the targeted root’s tone ends high (i.e. has a high or rising tone), the floating high tone morpheme has no effect on the surface form, as shown in (8).

- |     |             |                          |                |                                |            |
|-----|-------------|--------------------------|----------------|--------------------------------|------------|
| (8) | <i>bě́l</i> | <b>‘hermana; sister’</b> | <i>bě́l nà</i> | <b>‘mi hermana; my sister’</b> | <b>CLZ</b> |
|-----|-------------|--------------------------|----------------|--------------------------------|------------|

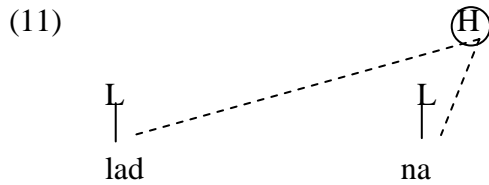
In SAMZ a root whose tone ends high is likewise unaffected by the floating high tone. However, in this case the pronoun itself undergoes a change in tone. Note that this tonal effect on the pronoun must be the result of the same floating high tone involved in first person singular marking and cannot be the result of some kind of tone sandhi caused by the concatenation of a certain type of root and a low-toned CV pronoun because the second person familiar pronoun is unaffected in the same context. In (9) for comparison I give the rising-toned bare root, and the second person and first person singular possessed forms. This behavior of the floating high tone in SAMZ makes it seem more like a clitic than an affix (which is how I’ve analyzed the cognate tonal marker in CLZ) because although tones are necessarily phonologically bound, the domain of this floating high tone seems to be the noun phrase rather than the head noun itself.

- |     |                         |                                 |                               |
|-----|-------------------------|---------------------------------|-------------------------------|
| (9) | <i>lǒs</i>              | <i>lǒs lè</i>                   | <i>lǒs nǎ</i>                 |
|     | <b>‘lengua; tongue’</b> | <b>‘tu lengua; your tongue’</b> | <b>‘mi lengua; my tongue’</b> |

The docking of the floating high tone on the first person singular pronoun is consistent, predictable, and easy to explain in the environment just described, an environment in which it would be phonologically impossible for this tonal morpheme to mark the possessed noun. The behavior of the floating high tone with low-toned possessed noun roots, as I described it above seems equally unsuspecting and natural but in reality in that environment it is more complicated. Sometimes when a low-toned root is possessed, the floating H docks on both the root and the pronoun, as shown in (10).

- |      |                       |                               |                             |
|------|-----------------------|-------------------------------|-----------------------------|
| (10) | <i>làd</i>            | <i>làd lè</i>                 | <i>lǎd nà</i>               |
|      | <b>‘cuerpo; body’</b> | <b>‘tu cuerpo; your body’</b> | <b>‘mi cuerpo; my body’</b> |

To account for this one could posit a floating high tone that targets all the word-final low tones in the possessed noun phrase, as modeled in (11).



The problem is predicting when both root and pronoun will be affected vs. the root alone. I have considered historical, phonological, and semantic explanations. For example, some SAMZ low tones correspond to CLZ low tone and others to CLZ falling tone, but there are both kinds of correspondences in each group of SAMZ words.

(12)	<b>CLZ</b>	<b>SAMZ</b>	<b>1s possessed</b>	
‘sal; salt’	<i>zèd</i>	<i>tèd</i>	<i>těd nà</i>	<b>Head noun only</b>
‘nombre; name’	<i>lě^</i>	<i>lě̀</i>	<i>lě̃ nà</i>	
‘diente; tooth’	<i>ndě̀y</i>	<i>lě̀y</i>	<i>lě̃y nă</i>	<b>Pronoun also</b>
‘cuerpo; body’	<i>lâz</i>	<i>lâd</i>	<i>lăd nă</i>	

The difference cannot be attributed to a difference in alienability of possession. Both types of possessed noun phrases show the same diversity, with the first person singular pronoun taking the floating high tone along with some low-toned possessed nouns but not with others, as shown in (13).

(13)	<b>Gloss</b>	<b>Noun in isolation</b>	<b>1s possessed</b>	<b>Type of possession</b>
	‘huarache; sandal’	<i>lâb</i>	<i>lăb nă</i>	inalienable
	‘nombre; name’	<i>lě̀</i>	<i>lě̃ nà</i>	inalienable
	‘guajolote; turkey’	<i>mbèr</i>	<i>xpěr nă</i>	alienable
	‘comal; griddle’	<i>sìl</i>	<i>xsìl nà</i>	alienable

Of the documented SAMZ low-toned nouns, those that remain glottalized with 1s possession never allow a tonal change on the accompanying 1s pronoun. Deglottalization with first person singular marking is also found on a small number of verbs in CLZ (Beam de Azcona, 2004), and on alienably possessed nouns in Lachixío Zapotec (Sicoli, 1998). Thus far I have seen deglottalization with first person singular marking on too few roots in SAMZ to tell whether it is predictable or not. The fact that the pronoun is unaffected tonally when the root stays glottalized suggests that the floating high tone first docks onto the possessed noun root and only spreads to the following pronoun if there is no glottal stop following the root vowel to block the spreading. This generalization is based on relatively few examples, however. In (14) I show two nouns with checked vowels, exhibiting the generalization that when there is deglottalization the high tone can spread and when there is not the glottalization blocks any spreading. Note though, that my notion of spreading here is not linear because it is not as though the high tone is simply copied and spreads one pace rightward. The high tone cannot concatenate in front

of a low tone in the same word, here the pronoun, and surface that way because there are no phonologically falling tones in this language (though the low tone does fall phonetically). Since the sequence HL is impossible in the language, the kind of spreading that takes place when the high tone docks on both noun and pronoun may involve tonal metathesis on the pronoun.

- (14) **‘barriga; belly’**    *lè<sup>7</sup>n*    →    *lě<sup>7</sup>nă*    (deglossalized)  
**‘papel; paper’**        *yè<sup>7</sup>tz*    →    *xkè<sup>7</sup>tz.nà*    (not deglossalized)

There are a small number of, perhaps irregular, low-toned nouns which either show no tonal alternation on the root or have a surface high rather than rising tone when possessed by 1s, and which do not allow tonal alternation on the accompanying pronoun. I show two of these in (15).

- (15) **‘yerno; son-in-law’**    *xùs*    →    *xús.nà*  
**‘llama; flame’**            *bè<sup>7</sup>l*    →    *xpè<sup>7</sup>l.nà*

For unglottalized low-toned nouns with surface rising tone when possessed by the first person singular, I have not found a phonological predictor for when the pronoun will be affected and when it will not.

- |      |   |                          |                            |
|------|---|--------------------------|----------------------------|
| (16) | Floating H affects root only                          |                          | Also affects pronoun       |
|      | <i>blàn</i> <b>‘carbón; coal’</b>                     | <i>dìtz</i>              | <b>‘espalda; back’</b>     |
|      | <i>kwè<sup>7</sup>l</i> <b>‘totomoxtle; cornhusk’</b> | <i>gòn</i>               | <b>‘limosna; alms’</b>     |
|      | <i>lè<sup>7</sup></i> <b>‘nombre; name’</b>           | <i>kwàt</i>              | <b>‘cachete; cheek’</b>    |
|      | <i>mbyìn</i> <b>‘pájaro; bird’</b>                    | <i>làb</i>               | <b>‘huarache; sandal’</b>  |
|      | <i>sìl</i> <b>‘comal; griddle’</b>                    | <i>làd</i>               | <b>‘cuerpo; body’</b>      |
|      | <i>tèd</i> <b>‘sal; salt’</b>                         | <i>lèy</i>               | <b>‘diente; tooth’</b>     |
|      | <i>xlè<sup>7</sup></i> <b>‘fruta; fruit’</b>          | <i>lò</i>                | <b>‘cara; face’</b>        |
|      | <i>xnì</i> <b>‘luz; light’</b>                        | <i>mbèr</i>              | <b>‘guajolote; turkey’</b> |
|      | <i>yù</i> <b>‘tierra; soil’</b>                       | <i>mbè<sup>7</sup>kw</i> | <b>‘perro; dog’</b>        |
|      | <i>yùx</i> <b>‘arena; sand’</b>                       | <i>mbè<sup>7</sup>l</i>  | <b>‘pescado; fish’</b>     |

It is unclear at this time whether a phonological solution will present itself or whether these lexical items fall into classes based on their morphological behavior when marked for a first person singular possessor. Nevertheless I include the above information for those who are interested in seeing fresh data from this Zapotec language, which I expect will differ in this respect from other Zapotec languages which some readers will be familiar with.

## 4.2 Segmental phonology

In this section I give two clear rules of segmental allophony and I explore the evidence for positing a separate series of prenasalized plosive phonemes.

#### 4.2.1 Devoicing

SAMZ has word-final devoicing, as seen in (17).

(17) *làg* [làx] ‘**hoja; leaf**’      *lár* [láɾ] ‘**ropa; clothing**’

There is also anticipatory voicing assimilation when /β/ in a prefix precedes a voiceless segment in a root, as shown in (18).

(18) *bdō* [βðō] ‘**plátano; banana**’    vs.    *bthē*’ [βθèʔ] ‘**epazote**’  
*bchú’ux* [βčúʔuʂ] ‘**jitomate; tomato**’

The *b* in the *mb*- animacy prefix also assimilates to the voicelessness of following consonants.

(19) *mblàn* ‘**liebre; jackrabbit**’      vs.    *mpsĩn* ‘**venado; deer**’  
*mptìn* ‘**ratón; mouse**’

The same thing happens with the homophonous class A<sup>5</sup> completive prefix. This prefix is shown before roots with initial voiced and voiceless consonants in (20).

(20) Completive	+	Verb stem	=	Surface form
<i>mb-</i>		<i>-dĩ’b</i> ‘ <b>costurearlo; sew (it)</b> ’		<i>mbdĩ’b</i>
<i>mb</i>		<i>-rò’o</i> ‘ <b>salir; go out</b> ’		<i>mbrò’o</i>
<i>mb-</i>		<i>-tè’ed</i> ‘ <b>aprender; learn</b> ’		<i>mptè’ed</i>
<i>mb-</i>		<i>-thà’as</i> ‘ <b>aplastarse; get smashed</b> ’		<i>mpthà’as</i>

#### 4.2.2 Prenasalized stops and affricates

As defined by Smith Stark (2003), SZ languages have nasal-initial animal words. In all branches but Cisyautepacan these words in fact begin in nasal-obstruent sequences. Besides the animacy prefix, SAMZ (a Miahuatec language) and CLZ (a Coatec language) both have homorganic nasal-obstruent sequences in the habitual and completive aspect prefixes. One might ask whether these sounds are truly clusters or whether they are an independent series of phonemes. In SAMZ these would be /<sup>m</sup>b, <sup>n</sup>dz, <sup>ɲ</sup>j, <sup>ŋ</sup>g, <sup>ŋ</sup>gw/.

There are no voiced stop or affricate phonemes in SAMZ. Under either analysis some synchronic or historical rules must be posited. These include homorganic nasal assimilation, occlusion of voiced fricatives and or voicing of voiceless plosives. In the case of *mb* and *ng* a phonological rule could cause occlusion of /β/ and /ɣ/. This is certainly what happened historically since words which here have *mb* and *ng* are

reconstructed with lenis *\*p* and *\*k* (Kaufman, 2003). *Ngw* results historically from the earlier marker /ko/ cited by Córdova (1578) and reconstructed by Kaufman (2003). The labialization is a reduction of the vowel. Synchronically, the cluster analysis is more problematic for *ndz* and *ndzh* [<sup>n</sup>dz, <sup>n</sup>ʒ] than for *mb*, *ng*, *ngw*. While there are voiced fricatives in SAMZ that correspond to *mb* and *ng(w)*, there are no voiced sibilants. A synchronic analysis would have to involve either voicing of the voiceless affricates *tz* and *ch* or the voicing of the sibilants *s* and *x* with additional stop epenthesis.

If these sounds are underlying clusters we might expect to find evidence that they can result synchronically from the concatenation of the component phonemes and that there can also be synchronic reduction or deletion of one of the component phonemes. There are nouns sharing a derivational relationship which differ only by the initial nasal, e.g. *bě`l* ‘llama; flame’ and *mbě`l* ‘estrella; star.’ On the surface these would seem to suggest that *m* is added to form the more animate word, but in fact the animacy prefix is *mb-* and in all the cases I’ve found the less animate word is *b-*initial so it could be that the *b* of the prefix causes the initial /β/ of the root to delete since [mbβ] does not occur. As seen in (20) above, the *b* of the *mb* prefix is devoiced while the *m* is not, preceding voiceless sounds. Better evidence is the fact that the *b* of the animacy prefix can delete altogether. In SAMZ I have only seen this happen before *k*, as shown in (21).

(21)	Classifier	+	Noun root	=	Surface form
	<i>mb</i>		<i>kĩ</i> ‘liendre; nit’		<i>mkĩ</i>

In SAMZ, the habitual aspect marker [<sup>n</sup>ʒ] also reduces before consonant-initial verb roots but instead of reducing to a plain nasal *n* it reduces to a prenasalized stop *nd*.. The sequence [nd] without any following sibilant, only occurs in this reduced form of the completive aspect marker and not in any other word in the language that I have seen.

(22)	Habitual	+	Verb root	=	Surface form	<b>SAMZ</b>
	[ <sup>n</sup> ʒ]		- <i>àth</i> ‘morir; die’		[ <sup>n</sup> ʒ àθ]	
	[ <sup>n</sup> ʒ]		- <i>bó’o</i> ‘sacar; take out’		[ <sup>n</sup> dβóʔo]	

Since the cluster analysis is stronger for *mb*, *ng*, *ngw* than for *ndz* and *ndzh*, I look now to the historical origins of the latter two sounds. Instances of [<sup>n</sup>ʒ] in SAMZ are cognate with *nd* in CLZ. There are two historical sources for this consonant (I take PZ forms from Kaufman, 2003), *\*(nV+)l* and *\*(nV)+ty*. Lenis *\*ty* otherwise usually reflects as /t/ in SAMZ, as in the verb ‘salir; go out’ in (20) above, except when it preceded *\*i*, as in ‘squirrel.’ This split took place in all SZ languages except in Coatec proper. In other Mihuatec languages for which I have collected data, (SAL, Santa Cruz Xitla, San Sebastián Río Hondo), the reflex of *\*ty* before *\*i* is /z/, e.g. ‘7’ *\*katyi* > *yaz* and ‘hogar; home’ *\*lityi* > *liz*. In SAMZ there is no /z/ and instead the reflex is /s/. In some Cisyauatepecan languages, including SAMZ’s closest neighbor, San Juan Mixtepec, the reflex of *\*ty* before an earlier front vowel is /dz/, the same as in the SAMZ word for

‘squirrel.’ Instances of [ʰdz] in SAMZ usually correspond to CLZ *nzh* [ʰj̃] and are either reconstructed with \*y or retroflex \*ɣ.

**Figure 6: PZ origins of and CLZ correspondences with SAMZ [ʰj̃] and [ʰdz]**

PZ	gloss	CLZ	SAMZ
*ni+lana	<b>‘tizne; soot’</b>	<i>ndān</i>	[ʰj̃ān]
*lawo	<b>‘cara; face’</b>	<i>ndô</i>	<i>lõ</i>
*nV+tyeʔ	<b>‘aquí; here’</b>	<i>ndë̃^ ~ ndë̃ʔ</i>	[ʰj̃ éʔe]
*kwe+tyi:ʔ(s)sa	<b>‘ardilla; squirrel’</b>	<i>ndiʔz</i>	[ʰdzíʔit]
*ti+yaka(ʔ)	<b>‘oreja; ear’</b>	<i>nzhâ</i>	[ʰdzàx]
*ɣ-okwaʔ	<b>‘maíz en grano; dried corn kernels’</b>	<i>nzhõp</i>	[ʰdzõʔβ]

The sounds *mb*, *ng*, and *ngw* seem to be more recently formed or transparent clusters, while the sounds *ndz* and *ndzh* seem to be less analyzable and more like single segments.

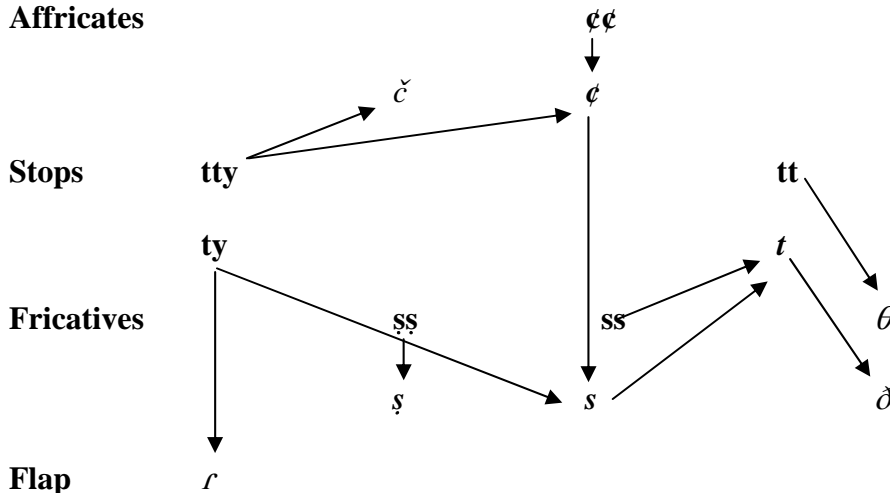
### 5. Historical and comparative data of interest

The segment found initially in the word ‘mouth’ is reconstructed by Swadesh (1947) and Fernández de Miranda (1965) as \**r*, by Suárez (1973) as \**t* and by Benton (1988) and Kaufman (2003) as \**ty*. Most modern Zapotec languages, including SAMZ, have /*r*/ here. In (Beam de Azcona, 2001) I argue that the change from \**ty* to /*r*/ is post-contact and in fact related to contact with Spanish which itself has an alveopalatal flap. Evidence that SAMZ probably did not have /*r*/ at the time of contact with Spanish is found in the loanword *lěnkw* ‘jorobado; hunchback’ < Spanish *renco* ‘limping, gimp.’

Changes in coronal obstruents in Southern Zapotec are interesting and often involve chain shifts. Sounds that have been reconstructed by Kaufman (2003) are given in bold and sounds that exist in modern SAMZ are italicized:

**Figure 7: Coronal obstruents from Proto-Zapotec to SAMZ**

#### Affricates



Most of the changes in Figure 7 took place in other Miahuatec languages too. The only changes from Figure 7 that also took place in Coatec languages are \*ʂ > ʂ, \*tt > θ, and \*t > ð. The chain shift in which ty > s > t > ð is thus probably a drag chain, beginning before the breakup of Miahuatec and Coatec.

Few attempts have been made to reconstruct tone in Proto-Zapotec or its daughters, mostly due to lack of reliable data. The first attempt was made by Swadesh (1947) and the most recent by Benton (2002). Having studied the tone system of Coatlán-Loxicha Zapotec intensely since 1996, and now embarking on this study of SAMZ, I hope that by comparing the suprasegmental systems of both languages I will be able to make my own contribution to the diachronic study of tone in this family. The system of suprasegmental contrasts found in SAMZ appears more conservative than that of CLZ. There are fewer tones in SAMZ and more phonation type contrasts, something more like the system of languages like Isthmus Zapotec (e.g. see Pickett, 1959). In Figure 8 are the results of a survey of 220 cognates between these two languages. Tonal correspondences are represented in rows and vowel type correspondences are represented in columns. Symbols on the left of a tilde refer to SAMZ and on the right of the tilde to CLZ. Thus, the third box in the second row indicates that there are two words which have low tone in SAMZ corresponding to falling tone in CLZ which also have a checked vowel in SAMZ and a plain vowel in CLZ. Since glottalization acts as a tone in CLZ but can be cognate with glottalization in SAMZ (which is not tonal), it is indicated along both axes.

**Figure 7: Suprasegmental combinations in correspondence---SAMZ ~ CLZ**

	V ~ V	V ~ V <sup>ʔ</sup>	V <sup>ʔ</sup> ~ V	V <sup>ʔ</sup> ~ V <sup>ʔ</sup>	V <sup>ʔ</sup> V ~ V	V <sup>ʔ</sup> V ~ V <sup>ʔ</sup>	Total
L ~ L	31		5		8		44
L ~ F	46		2		3		51
L ~ R	1		2				3
L ~ ?				8		10	18
R ~ R	14		11				25
R ~ F	11		1				12
R ~ L	5		2				7
R ~ H	2						2
R ~ ?		6		2		1	9
H ~ F	10		1		3		14
H ~ H	1				2		3
H ~ ?		3		4		22	29
H ~ R			1		1		2
H ~ L					1		1
<b>Total</b>	<b>121</b>	<b>9</b>	<b>25</b>	<b>14</b>	<b>18</b>	<b>33</b>	

I have not made a reconstruction of tone for the ancestor of these two languages, but certain significant gaps and tendencies can be found in the table above from which we can make some generalizations.

More than 80% percent of CLZ words with ? tone have some type of glottalization in SAMZ. Unglottalized SAMZ words that correspond to glottalized CLZ words have SAMZ H or R tone but not L. Only about 60% of SAMZ checked and rearticulated words have ? tone in CLZ. Regarding SAMZ as conservative, it seems that glottalization was lost in CLZ about 40% of the time, but that the CLZ glottal tone does indeed reflect historical glottalization. Another 20% of words with CLZ glottal tone seem to have developed it not from glottalization but from the high pitch which is found with high and rising tones and also from glottalization.

More than 80% of CLZ low-toned words correspond to SAMZ low-toned words. About 60% of CLZ falling-toned words have low tone in SAMZ. More than 80% of CLZ rising-toned words have rising tone in SAMZ. CLZ high-toned words are rare but never correspond to SAMZ L tone. Low and rising tones are common and are found in both languages, not having changed in 80% of the words surveyed. While just over half of CLZ falling tones seem related to the SAMZ low tone, the other 40% may have resulted from the addition of a low tone following a high tone, for example with post-tonic vowel deletion, in CLZ and Coatec. Of the 60% of CLZ falling tones that correspond to SAMZ low tone, while it is possible that there was a merger in SAMZ, it is equally possible that there was a split in CLZ. The SAMZ low tone is not level at all but falls phonetically, the same as in the most Northern dialects of CLZ. Upstep, floating tones, or other details about the tonal environment in these words may have contributed to the development of a higher falling tone that exists contrastively in CLZ today.

Looking at the same results but with the focus on SAMZ, we find that SAMZ high is most likely to correspond to CLZ high (30%) and glottal (60%) tones. SAMZ rising tone is most likely to correspond to CLZ rising tone (almost 50%) but can correspond with any CLZ tone. SAMZ low tone is most likely to correspond to CLZ falling and low tones (about 40% each).

About 25% of CLZ words in this sample have glottal tone. About 25% of SAMZ words have checked vowels. About 25% of SAMZ words have rearticulated vowels. About 50% of SAMZ words are not glottalized. So, while about half of SAMZ words have some type of glottalization, this number has been reduced by half in CLZ.

Looking at the frequency with which tones occur we find that low is the most common SAMZ tone, making up a little over 50% of this sample. High and rising tones are equally common in SAMZ. This differs from CLZ in which high tone is rare.

## **6. Conclusion**

The description of SAMZ will have much to offer toward the comparison with other Zapotec languages and the documentation of the history of the Southern Zapotec region itself. This language has made important contributions to this region and is unfortunately now at the end of its days. The complexity, e.g. of tonal morphology, found in this language will surprise those accustomed to seeing more morphological simplification in dying languages. I see two possible explanations for this. It is possible that forms which were once predictable based on particular generalizations come to be used more haphazardly. However, if this were the case one would expect more free variation than I

have seen so far with the first person possessed forms I have recorded. More likely, the complexity preserved by the last two speakers of this language, survives because this language virtually ceased to be spoken with one fell blow in 1965. If it had ceased to be used in a more gradual way the expected simplifications could have spread through the dwindling speech community. Instead, the only two people on earth who can speak this language fluently acquired it as children when the language was not so endangered and for political reasons stopped speaking it to each other or passing it along to younger generations. The language is almost mummified, resisting change because of lack of use. Yet, the language is not dead yet and the younger generation today is more interested than the generation of 1965. While no children are learning this language, young adults frequently visit the speakers and ask them how to say things in Zapotec. I hope that whatever documentation I am able to produce will also be of value to future generations of people in San Agustín Mixtepec, Miahuatlán, Oaxaca, Mexico.

### Notes

<sup>1</sup> My fieldwork on this language in the summers of 2002 & 2003 was generously funded by the Endangered Language Fund, The Survey of California and Other Indigenous Languages, and the California Indian Language Center. I am also especially grateful to Lázaro Díaz Pacheco & my anonymous SAMZ consultant.

<sup>2</sup> I use the term “plosive” to include both stops and affricates. This may not be a standard usage for this term but it is the one I originally acquired in my acquisition of linguistic jargon and I find it convenient to use this definition when talking about series of Zapotec phonemes that include both stops and affricates.

<sup>3</sup> This noun is obligatorily possessed. To simply refer to any old sandal one would use the compound *yě`l yìd* and by contrast that noun is alienably possessed, e.g. *xkě`lyìd lè* ‘tu huarache; your sandal.’

<sup>4</sup> Most body parts are inalienably possessed but some extremities are alienably possessed, perhaps because these body parts can come off in accidents or violent incidents during one’s lifetime, while the loss of more core body parts cannot be lost before the end of one’s life. A similar example is the loanword *kól* ‘cola; tail’ in *xkól mǎ*<sup>7</sup> ‘su cola del animal; it’s tail.’

<sup>5</sup> For discussion of class A see Kaufman (1993, 2003) or Beam de Azcona (2004). For different analyses see Bartholomew (1983) or López & Newberg (1990).

### References

- Bartholomew, Doris A. 1983. *Gramática Zapoteca* in Nellis, Neil and Jane G. Nellis, comps. *Diccionario zapoteco de Juárez: zapoteco-español, español-zapoteco (Ca titsa' qui' ri'u)*. Vocabularios (y Diccionarios) Indígenas ``Mariano Silva y Aceves" 27. Mexico: Instituto Lingüístico de Verano.
- Beam de Azcona, Rosemary G. 2001. Proto-Zapotec coronal stops in Coatlán-Loxicha Zapotec and other Southern Zapotec languages. ms.

- , 2004. *A Coatlán-Loxicha Zapotec Grammar*. Ph.D. dissertation. University of California, Berkeley.
- Benton 1988. Proto-Zapotec Phonology. ms.
- , 2002. A reconstruction of the tone system of Proto Zapotec. ms.
- Bickmore, Lee and George Aaron Broadwell. 1998. High Tone Docking in Sierra Juárez Zapotec. *International Journal of American Linguistics*, vol. 64, no. 1, pp. 37-67.
- Black, Cheryl A. 1994. *Quiegolani Zapotec Syntax*. Ph.D. dissertation. University of California, Santa Cruz.
- Córdova, P. Fray Juan de. 1987 [1578]. *Arte del Idioma Zapoteco*. México: INAH.
- Fernández de Miranda, María Teresa. 1995 [1965]. *El Protozapoteco*. México: El Colegio de México y el Instituto Nacional de Antropología e Historia.
- Kaufman, Terrence S. 2003. Proto-Zapotec reconstructions. ms.
- López L., Filemón and Ronaldo Newberg Y. 1990. *La Conjugación del Verbo Zapoteco: Zapoteco de Yalálag*. México, D. F.: Instituto Lingüístico de Verano.
- Pickett, Velma B. y colaboradores. 1959. *Castellano-zapoteco, zapoteco-castellano [Vocabulario zapoteco del Istmo]. Vocabularios (y Diccionarios) Indígenas "Mariano Silva y Aceves" 3*. Mexico: Instituto Lingüístico de Verano. ii, 246 pp. [2nd ed., rev. and enlarged (1965), xii, 163 pp.; repub. (1968, 1971)].
- Pickett, Velma B., Cheryl Black, and Vicente Marcial Cerqueda. 1998. *Gramática Popular del Zapoteco del Istmo*. Juchitán: Centro de Investigación y Desarrollo Binnizá A. C. e Instituto Lingüístico de Verano A.C.
- Rojas, Basilio. 1958. *Miahuatlán: Un Pueblo de México*. Monografía del distrito de Miahuatlán, estado de Oaxaca.
- Sicoli, Mark. 1998. "Glottalization and Categories of Possession in Lachixío Zapotec." Paper presented at the annual meeting of the Society for the Study of the Indigenous Languages of the Americas, New York.
- Smith Stark, Thomas C. 2003. "Algunas isoglosas zapotecas." Trabajo preparado para las actas del III Coloquio Internacional de Lingüística "Mauricio Swadesh." México, D.F.: UNAM.
- Suárez, Jorge A. 1973. On proto-Zapotec phonology. *IJAL* 39: 236-49.
- Swadesh, Morris. 1947. The Phonemic Structure of Proto-Zapotec. *International Journal of American Linguistics* 13: 220-30.