LANGUAGE DOCUMENTATION & CONSERVATION

Vol. 4 (2010), pp. 231-253
http://nflrc.hawaii.edu/ldc/
http://hdl.handle.net/10125/4481

Orthography Design for Chuxnabán Mixe

Carmen Jany
California State University, San Bernardino

Many discussions of orthography development center on the later stages of such endeavors and on the impact of newly developed orthographies over an extended period of time. This paper focuses on the early stages of orthography development for Chuxnabán Mixe, a previously undocumented language, and the establishment of a working orthography in collaboration with the community for the purpose of language documentation.

The orthography design follows many of the linguistic, pedagogical, sociopolitical, and practical principles observed in new orthographies, such as phonemic orientation, maximum ease of learning, local acceptability, and ease of use with computers and new media. While the community favors using conventions from the dominant language, Spanish, it also prefers dialectal particularity over multidialectal uniformity.

Chuxnabán Mixe is a Mixean language spoken by 900 people in one village in Oaxaca, Mexico. Limited documentation is available for some of the other Mixean varieties, but there is no widely used uniform orthography. Mixean languages and dialects differ primarily in their vowel systems, and each variety, if documented, has its own orthographic conventions established, often highlighting dialectal idiosyncrasies. This paper illustrates the orthography development and discusses some of the similarities to and differences from other orthographic conventions used for this language family.

1. INTRODUCTION.[1] Many discussions of orthography development center on the later stages of such endeavors and on their impact over an extended period of time. This paper focuses on the early stages of orthography development for Chuxnabán Mixe, an undocumented Mexican indigenous language, and discusses the establishment of a working orthography in collaboration with the community for the purpose of language documentation. As many studies have shown, these early stages of orthography development are essential and can determine the success of an orthography (Grenoble & Whaley 2006; Guérin 2008; Hinton 2001; Rehg 2004; Rice & Saxon 2002).

Chuxnabán Mixe is spoken in the eastern midlands of the Mexican southern state of Oaxaca (see Figure 1). The Mixean territory is composed of 290 communities (Torres Cisneros 1997). Each village speaks a different Mixean variety, many of which are mutually unintelligible. In many cases it has yet to be determined whether a particular variety represents a distinct language or dialect, since documentation of Mixean languages is limited. Some linguists divide the Mixean family into three main languages: Lowland Mixe, Midland Mixe, and Highland Mixe (INEA 1994, 1997a, 1997b, 1997c). More recently,

[1] This research was in part supported by a Pilot Project Grant (PPG0044) from the Hans Rausing Endangered Languages Project, for which I am very grateful. I would also like to thank the Chuxnabán Mixe community for their collaboration and support in the orthography development process. Furthermore, I am thankful to two anonymous reviewers for their useful comments on an earlier draft of this paper.

Licensed under Creative Commons
Attribution Non-Commercial No Derivatives License
E-ISSN 1934-5275

Ethnologue lists ten Mixean languages divided into three larger branches: Eastern Mixe with six languages, and Veracruz Mixe and Western Mixe with two languages each (Lewis 2009). Chuxnabán Mixe has been identified by its speakers as Midland Mixe and corresponds to Quetzaltepec Mixe in the *Ethnologue* entry. At present, there are a few published grammars and dictionaries of other Mixean languages (De la Grasserie 1898; Hoogshagen & Hoogshagen 1997; Ruiz de Bravo Ahuja 1980; Schoenhals & Schoenhals 1982; Van Haitsma & Van Haitsma 1976), but there is no widely used uniform orthography. Rather, each variety, if documented, has established its own orthographic conventions, often highlighting dialectal idiosyncrasies.

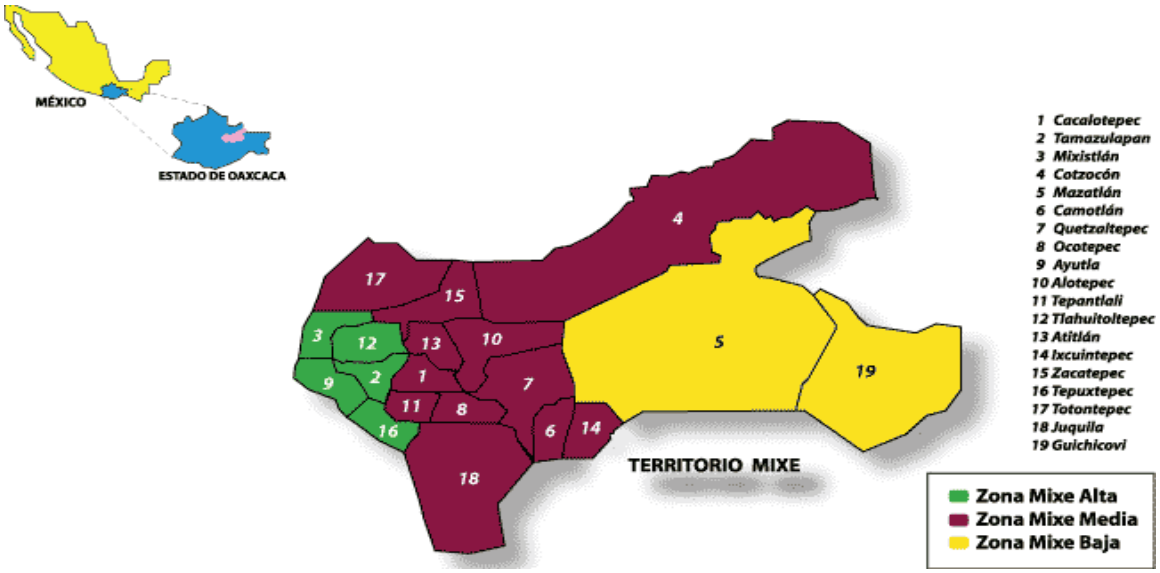

FIGURE 1: Map of the Mixean Region[2]

Previously undocumented and to some degree endangered, Chuxnabán Mixe is spoken by 900 people in one village. Following the framework established by UNESCO (2003) and the levels of endangerment discussed in Grenoble & Whaley (2006:16–19), Chuxnabán Mixe can be assessed as *unsafe*, *vulnerable*, or *at risk*. All community members speak the language, and there is intergenerational language transmission. However, education and literacy development occur exclusively in Spanish, and bilingualism is on the rise. Thus, while it continues to be learned as a first language by children, the proficiency of these speakers is limited. Language use is also rapidly declining as many young people migrate to the cities or the United States, where they often cease speaking the language. Currently,

[2] The map is available at the following website: http://www.redindigena.net/ser/pueblomixe/mapa.html.

Chuxnabán Mixe is used only within the community, mainly at home and for unofficial daily activities. At official gatherings, such as an informational village reunion, as well as at some religious events, both Chuxnabán Mixe and Spanish are spoken. Communication with members from other Mixean communities generally occurs in Spanish, especially if the variety is very distinct. Hence, there are restrictions in terms of use in language domains. Moreover, virtually no literacy materials existed, nor had an official orthography been adopted prior to the start of my documentation project.

My work on Chuxnabán Mixe began in 2006, when I began working with a speaker in the United States for several months before visiting the community for the first time to ask for permission, to establish contacts and search for collaborators, and to collect data. During this very early stage of the language documentation process I compiled several wordlists, established a list of phonemes, and shifted from a phonetic orthography based on the International Phonetic Alphabet (IPA) to a tentative phonemic orthography. This first orthography was based on my phonemic analyses. Existing materials on other Mixean languages (De la Grasserie 1898; Hoogshagen & Hoogshagen 1997; Ruiz de Bravo Ahuja 1980; Schoenhals & Schoenhals 1982; Van Haitsma & Van Haitsma 1976) were very helpful, given that the consonant systems are almost identical in the various Mixean languages. The vowels presented a greater challenge, as discussed in section 4.1. The main purpose of creating a written representation at that stage was language documentation, linguistic analysis, and the preparation of discussion points for a community-approved orthography. The result was a first tentative orthography designed to serve as a starting point for the development of a working orthography in collaboration with the community.

The data for this paper originated from my second and longer visit to the community in 2008. During the first week of that visit, I had three meetings with the local government to explain the language documentation project to the local officials and to find potential collaborators, consultants, and interpreters. Village officials identified community elders who could serve as storytellers and collaborators for orthography development and testing. They decided that orthography development should occur in consultation with younger speakers with a high school diploma, as the younger generation would benefit the most from having written records of their language and history. During the second week of my field trip, the orthography was elaborated and tested. Several changes were made to my first tentative orthography (see section 3.1). The new orthography was tested informally with several young community members and approved.

In general, the orthography design follows many of the linguistic, pedagogical, sociopolitical, and practical principles observed in new orthographies (Grimes & Gordon 1980) and discussed in the next section, such as phonemic orientation, maximum ease of learning, local acceptability, and ease of use with computers and new media. Furthermore, while the established orthography is very similar to those of other Mixean varieties, it also retains dialectal particularities, as seen in section 3.2. Overall, the orthography development was guided by four main factors treated in section 3.1. Section 2 focuses on general principles in orthography development for oral languages. Section 3 describes the established orthography and evaluates the principles employed. Section 4 treats specific challenges and is followed by summary and conclusions in section 5.

2. GENERAL PRINCIPLES IN ORTHOGRAPHY DEVELOPMENT FOR ORAL LANGUAGES. Introducing native literacy for an oral language can have many implications, positive and negative (Guérin 2008; Hinton 2001; Hornberger 1997). The first question one should ask when embarking on such an endeavor is: *Do the benefits outweigh the risks and disadvantages?* Some of the risks summarized in Hinton 2001 include the native speakers' loss of control over the language, the freezing and decontextualizing of the language, and the possible development of disagreements and divisions within a speech community. Once a language is written, a speaker may lose ownership and control over who has access to the language, and the natural course of the oral transmission—whereby narratives are constantly being shaped and adapted—is interrupted by having narratives fixed in time. Furthermore, there may be diverging opinions as to how a story should be recorded or how the language should be written down, which can create divisions within the community. While the pitfalls of having a written representation of a language seem significant, the benefits can be greater. Benefits of writing oral languages include written language documentation for language survival, possible expansion of language use to other domains, and the empowerment of the community due to the misconception that written languages are superior to spoken languages. Moreover, literacy is often linked to power in the society (Hornberger 1997:360). For the Chuxnabán Mixe community "to write or not to write their language" did not come up as a debate during my visit, because they acknowledged increasing language loss in the community; recognized the benefits of having their narratives, history, customs, and language recorded for the next generations; and felt a sense of pride in having their language written and, thus, being more similar to the dominant language, Spanish. In addition, several other Mixean varieties already had orthographies in place, there were Mixean street signs in other villages, and some of the younger speakers who attended a high school in another village had had lessons on how to read and write that particular Mixean variety. Once it was clear that the community was very eager to see their language in writing, many other factors needed to be contemplated.

In order to develop a new orthography for an oral language, a series of linguistic and non-linguistic factors need to be taken into consideration, the advantages and disadvantages for each option need to be assessed, and compromises need to be made. The following discussion of such factors in orthography design is based on Grenoble & Whaley 2006, Grimes & Gordon 1980, Sebba 2007, Seifart 2006, and Venezky 2004.

Most new orthographies that are created for the purpose of language documentation are phoneme-based; i.e., they show a one-to-one correspondence between phonemes and graphemes. Phonemic writing is generally preferred over phonetic writing, as phonemic systems change more slowly than phonetic realizations; phonemically-based orthographies are thus more stable (Grenoble & Whaley 2006:141). Creating a phonemic orthography implies at least a basic phonological analysis preceding its design. Grimes & Gordon (1980:94) note that when creating a phonemic orthography, "linguists know how to do a good job in a short period of time" and that "making sure that all phonemic distinctions get encoded is safe," while Rehg (2004:506) cautions that "faulty phonological analyses give rise to faulty orthographies." There are at least two consequences of phonemic orthographies: (1) possible inconsistency among different representations of a single word or morpheme, and (2) possible creation of homographs, i.e., words that are lexically different but pronounced the same (Grimes & Gordon 1980:93). Therefore, the principle of one

distinctive sound represented by one specific symbol cannot always be followed, as other linguistic factors also come into play. Such factors include orthographic depth, under- and overdifferentiation, and functional load.

Orthographic depth distinguishes shallow from deep orthographies and refers to the level of linguistic structure at which forms are represented (Seifart 2006:279). Shallow orthographies are those with a correspondence between the written representation and the surface realization of linguistic forms, i.e., the phonetic forms rather than the phonemic, and the phonemic form rather than the morphemic structure. Following Sebba 2007, fully phonemic orthographies are shallow. Deep orthographies are those with a correspondence between written representation and underlying form. Orthographic depth can also have a linguistic impact and result in under- or overdifferentiation, i.e., the representation of phonemic contrasts. Underdifferentiation or underrepresentation occurs when more than one phoneme is represented by the same symbol, possibly resulting in homographs. Overdifferentiation is present when different symbols are used for phonetic distinctions, such as allophonic variation (Grenoble & Whaley 2006:148). Underdifferentiation may be the result of simplicity, for example by limiting the number of symbols in the orthography, in order to omit contrasts with a low functional load. Functional load refers to the fact that some features are more frequent and more important than others in a language, and users of an orthography rely more on these features than on others for reading and writing. The above-discussed linguistic considerations are reflected in the Chuxnabán Mixe orthography design (see section 3).

While linguistic factors are basic in orthography design, they may conflict with non-linguistic factors, which ultimately override all others. The main non-linguistic factor in orthography development is social, political, and cultural acceptability. Written representations of a language serve as markers of identity (Grenoble & Whaley 2006; Sebba 2007). If a community does not approve the orthography, then it will not be used and will therefore fail. Depending on the sociopolitical situation, communities may choose to distinguish their language, dialect, and themselves from neighboring varieties through the use of a specific grapheme to represent their distinctive sound or sounds and, thus, their identity.[3] This is one of the main reasons why the uniform Mixean orthography developed by the Mexican National Institute for Adult Education has never been widely accepted and used (INEA 1994, 1997a, 1997b, 1997c; Suslak 2003). The idea of a unified Mixean writing system was first proposed at a regional summit in 1979 (Suslak 2003:558), but great dialectal variation and the desire of each community to keep its dialectal idiosyncrasies in the orthography have impeded its success. Similarly, the Chuxnabán Mixe community expressed that they do not wish to consider orthographies of neighboring or other Mixean varieties, but would rather create their own unique orthography.

A second important non-linguistic factor in the development of an orthography for an oral language is ease of use with computers and new media. With the world-wide web reaching even the remotest areas of the world and expanding in use, it becomes clear that a new orthography should be designed in a way so that its graphemes are readily available on standard keyboards. This will not only facilitate the language documentation process, it

[3] See Hinton 2001 for Havasupai and Hualapai, as well as Suslak 2003 for Totontepec Mixe.

will also encourage its use with new media and possibly in new domains. Given that one of the documentation projects for Chuxnabán Mixe involves an online dictionary resource (Jany 2009), technological usability was an important factor.

The task of an orthography developer is to evaluate the above-discussed factors, to balance the advantages and disadvantages of each option, and to find compromise (Seifart 2006). The next section describes the established orthography, summarizes the development process, and evaluates some of the principles employed.

3. CHUXNABÁN MIXE ORTHOGRAPHY DESIGN. The Chuxnabán Mixe orthography presented in this paper is the result of three design stages. When I first started documenting the language in 2006, I worked for several months with a female speaker who had recently left the community and was living in the United States. My first written representation was purely phonetic, using IPA symbols in order to begin documenting the language and to examine the sound system. Once I had defined the phonemes, I shifted to a tentative phonemic orthography. This first phonemic orthography was based on orthographies from other Mixean varieties and discussions with a native linguist who was working on a related language. This tentative orthography served to prepare discussion points for my next visit to the village and to continue the documentation process; it also became the basis for the working orthography. The working orthography presented in this paper was elaborated in collaboration with community members during my field trip in 2008. In this early period in the language documentation process, many linguistic features were yet to be uncovered and, therefore, this was not an established standardized orthography; rather, it was expected that revisions and additional rules would follow. For now, this working orthography has been informally tested with speakers and approved by the community.

Interestingly, during my field trip in 2008, two previously written representations of Chuxnabán Mixe surfaced in the village, each using its own writing conventions. One was a booklet containing narratives. A member from a different Mixean community had once collected the narratives and left the booklet at the local government agency. The other was several loose pages containing a wordlist. A community member had prepared the wordlist, but never continued his work (Gregorio Cirilo 2005). Neither source included explanations of the orthography. In fact, the community collaborators were unable to read much of it and resorted to the Spanish translations to identify the Mixean contents and words. Hence, they quickly decided not to consider these sources in the design process.

In what follows, I present the largely phonemic orthography elaborated in 2008, illustrate changes made to the previous tentative orthography, and discuss the principles that guided the design process. In addition, the Chuxnabán Mixe orthography is compared to those of other Mixean varieties.

3.1 FROM PHONEMES TO GRAPHEMES. Chuxnabán Mixe has twelve consonantal phonemes: four stops, two nasals, two fricatives, two affricates, and two glides, and an additional eight phonemes /b, d, g, f, s, ɾ, r, l/ occurring only in Spanish loans. The consonants are summarized in Table 1 with phonemes from Spanish loans in square brackets.

The vowel system includes a phonemic vowel length distinction and a phonemic phonation contrast between plain, aspirated, glottalized, and interrupted vowels, the same as in other Mixean varieties. While the consonant systems are almost identical across all Mixean

varieties, the vowel inventories vary greatly and range from five to nine phonemic vowels (Crawford 1963; Hoogshagen & Hoogshagen 1997; Schoenhals & Schoenhals 1982; Suslak 2003; Van Haitsma & Van Haitsma 1976; Wichmann 1995). Chuxnabán Mixe has six vowel phonemes and three marginal vowel phonemes discussed in section 4.1. Table 2 summarizes the vowel phonemes. Marginal vowel phonemes are in square brackets. Table 3 illustrates the phoneme-to-grapheme correspondences. Phonemes from Spanish loans are excluded from this table, since they conform to Spanish writing conventions (see also section 4.2).

	Bilabial	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosives	p/[b]	t/[d]			k/[g]	ʔ
Nasals	M	n				
Fricatives	[f]	[s]	ʃ			h
Affricates		ts	tʃ[4]			
[Rhotic]		[ɾ][r]				
[Lateral]		[l]				
Glides	w			j		

TABLE 1. Chuxnabán Mixe consonants

i	[ʏ]	ɨ	u
e	[ø]		o
	[æ]	a	

TABLE 2. Chuxnabán Mixe vowels

Phoneme	p	t	k	ʔ	m	n	ʃ	h	ts	tʃ	w	j
Grapheme	p	t	k	’	m	n	x	j	ts	ch	w	y
Phoneme	a	e	i	o	u	ɨ	ʏ	ø	æ			
Grapheme	a	e	i	o	u	ë	ü	ö	ä			

TABLE 3. Phoneme to grapheme correspondence

[4] This affricate generally results from morpheme-induced palatalization, a suprasegmental phenomenon, and is often not treated as a phoneme in descriptions of Mixean varieties. It is included here because the phonetic effects differ from those observed in other instances of palatalization, and in some instances it can not be traced back to palatalization (see also discussion in section 4.3)

The orthography shown in Table 3 avoids special IPA symbols from the International Phonetic Alphabet that are not readily available on standard keyboards, therefore satisfying the technological usability principle. Furthermore, it uses dieresis for all special vowel qualities and models several consonants on Spanish spelling conventions. Before discussing these and other general principles of orthography design at work, I will provide a brief overview of some allophonic variations for the consonants. Given the complexity of the vowel system, the orthographic conventions relating to the vowels are discussed in a separate section in 4.1

Some of the allophonic variations found in Chuxnabán Mixe, such as obstruent voicing and final sonorant devoicing, are common traits in Mesoamerica (Campbell et al. 1986). The consonants and their allophones are summarized in Table 4.

Phoneme	Allophones
/p/	[p, b, pʰ, p̚]
/t/	[t, d, tʰ, t̚]
/k/	[k, g, kʰ, k̚]
/ˀ/	[ˀ, ɣ̞]
/ʃ/	[ʃ, ʂ, ʒ, ʐ]
/h/	[h, ɦ]
/ts/	[ts, dz, s]
/tʃ/	[tʃ, dʒ]
/m/	[m, m̥]
/n/	[n, n̥, ŋ]
/w/	[w]
/j/	[j]

TABLE 4. Allophones of consonants

While most of the allophonic variations are not problematic for the orthography design, allophonic obstruent voicing yields three voiced stops [b, d, g], which are phonemes in Spanish.[5] Since Chuxnabán Mixe speakers are generally literate in Spanish, they are used to seeing the written representations of these voiced stops. As a result, two design principles, (1) phonemic writing and (2) Spanish spelling conventions, are in conflict here. After illustrating the allophonic variations with several examples, the community collaborators decided not to write the voiced stops [b, d, g] in Mixean words, because they felt it

[5] However, their phonetic realizations generally differ from the allophonic voiced stops found in Chuxnabán Mixe. For example, Spanish voiced stops are spirantized in intervocalic position.

was not necessary for comprehension; but they decided to retain them in Spanish loans. This is illustrated in the examples below.

(1) Mixean words with allophonic voiced stops

tëëpë	'the ones who'	→ [ˈtɨːbɨ]
aka'any	'it started'	→ [aˈgaˀaŋ]
ntëëjk	'my house'	→ [ndɨːʰk]

(2) Spanish loans with phonemic voiced stops

cebu	'type of bull'	from Spanish	*cebú*	'type of bull'
lugarë	'place'	from Spanish	*lugar*	'place'
Dios	'God'	from Spanish	*Dios*	'God'

In general, four main factors were considered in the orthography design: (1) only representing phonemes rather than allophones, (2) using some Spanish writing conventions, since community members are generally literate in Spanish, (3) having the simplest representation possible and using symbols found on basic keyboards rather than special IPA characters, and (4) consulting orthographies of other Mixean varieties and the general orthography established by the Mexican National Institute for Adult Education (INEA) applicable to all Mixean languages for comparison. In regards to the voiced stops, the first factor overrode the second one, and [b, d, g] are not represented in Mixean words. As with many other Mexican indigenous languages, Spanish orthographic conventions are used to represent various sounds, such as the palatalized affricate /tʃ/ and the glottal fricative /h/, which are represented with the graphemes *ch* and *j* respectively. While it may seem illogical to someone not familiar with Spanish spelling conventions to use the grapheme *j* for the glottal fricative /h/ rather than for the palatal glide /j/, informal testing with several speakers showed that they were only able to read the words when written with the Spanish spelling conventions. Speakers were confused and hesitated when the spelling was changed to *h* for /h/ and *j* for /j/. This has to do with two Spanish spelling rules: (1) in Spanish the grapheme *h* is never pronounced and (2) the grapheme *j* is pronounced as the glottal fricative /h/. It was clear that following these same spelling rules for the Chuxnabán Mixe orthography greatly facilitated reading and writing. Likewise, the grapheme *y* represents a palatal glide /j/ in Spanish and was, therefore, adopted to represent the palatal glide. The graphemes adopted from the Spanish spelling conventions are illustrated in the examples below.

(3) Grapheme *j* for the glottal fricative /h/

jot	'stomach'	→ [hotʰ]
poj	'wind'	→ [poh]
pojënë	'fast'	→ [ˈpoɦɨnɨ]

(4) Grapheme *y* for the palatal glide /j/

yaa'iix	'you can see it'	→ [jaːˈiːʂ]
koy	'pillow'	→ [koj]
piyë'kp	'to run' -	→ [piˈjɨʔkʰpʰ]

(5) Grapheme *ch* for the palatalized affricate /tʃ/

chëëjk	'his/her house'	→ [tʃɨːʰk]
tsu'uch	'meat'	→ [tsuʔutʃ]
kachyën	'in the basket'	→ [ˈkatʃɨn]

The orthography used in Spanish loans is discussed in section 4.2. Generally, Spanish loans are represented in their original spelling, with the exception of certain vowels and when occurring with Mixean morphology.

To sum up, the first three principles mentioned above—phonemic writing, simplicity for technological usability, and Spanish spelling conventions—were followed with only minor compromises, such as phonemic writing over Spanish spelling for the voiced stops and dieresis for special vowels over other types of graphemes more readily found on basic keyboards. The fourth principle, which is following already-established Mixean orthographies, was also largely met. Although the community expressed that they wished to develop their own unique orthography, the results are very akin to orthographies used for other Mixean languages and dialects, given that similar principles were followed in the design. For example, many Mixean varieties use dieresis to represent certain vowel qualities rather than IPA symbols, follow the same Spanish spelling conventions for /h/ and /j/, and use an apostrophe for the glottal stop. Furthermore, given that several community collaborators had attended high school in a different village and learned the respective Mixean orthography, they already had ideas about how a Mixean language should look in writing and favored similar outcomes. A comparison to other orthographies is presented in section 3.2.

The working orthography is almost identical to the tentative orthography I had established previous to my second field trip. Only three changes were made: (1) the use of the voiceless stops /p, t, k/ for the voiced allophones [b, d, g] respectively, (2) the use of *ë* instead of *ï* for the high central vowel [ɨ], and (3) suprasegmental palatalization is represented by the grapheme *y* following the palatalized consonant, except for *ch*, rather than having two *y* graphemes, one preceding and one following the palatalized consonant, to mirror the phonetic realization (see 4.3). The changes are summarized in Table 5.

Phoneme	p	t	k	ʔ	m	n	ʃ	h	ts	tʃ	w	j
Tentative orthography	p, b	t, d	k, g	’	m	n	x	j	ts	ch	w	y
Working orthography	p	t	k	’	m	n	x	j	ts	ch	w	y
Phoneme	a	e	i	o	u	ɨ	ʏ	ø	æ			
Tentative orthography	a	e	i	o	u	ï	ü	ö	ä			
Working orthography	a	e	i	o	u	ë	ü	ö	ä			

TABLE 5: Tentative versus working orthography

To conclude, the Chuxnabán Mixe orthography is primarily phonemic and shallow, showing phonemic surface realizations and, therefore, allomorphy (Seifart 2006:279). It does not follow the constancy principle (Venezky 2004), thus, some morphemes are represented differently depending on their phonemic surface realization.

(6) Allomorphy represented in orthography

mata’px	‘eighty’	from	*mataxk* ‘four’ + *i’px* ‘twenty’
tumpë	‘worker’	from	*tun* ‘to work’ + nominalizer *=pë*
nyöky	‘his/her paper’	from	*noky* ‘paper’ + 3rd person prefix *y-*

A major drawback of a shallow orthography is that it may create homographs (Grimes & Gordon 1980:93). So far, only a very few homographs have been identified in Chuxnabán Mixe. There are several advantages to having a shallow orthography. First, a shallow phonemic orthography is easy to learn and aids the language documentation process. Second, a deeper orthography representing morphemes rather than phonemes may approximate Chuxnabán Mixe to other Mixean dialects, and the community clearly wanted their orthography to represent their unique Mixean variety. This is better achieved with a shallow orthography. The next section compares the established orthography to those of other Mixean languages and dialects.

3.2 ORTHOGRAPHIES OF OTHER MIXEAN LANGUAGES AND DIALECTS. Efforts to create a unified writing system for all Mixean varieties have met with limited success (Suslak 2003:557). The first attempt to develop a Mixean orthography dates back to the colonial period when a Dominican friar wrote a grammar, dictionary, and translations of various religious texts (De Quintana 1733). According to Suslak (2003:557), the Dominican order was then forced to withdraw from the region and De Quintana’s “writing system ended up forgotten.” Much later, in the 1950s and 1960s, the Summer Institute of Linguistics (SIL) started to work on several Mixean languages. Interestingly, they often used the

same symbols for the same phonemes in distinct varieties but different orthographic conventions for each variety. SIL's strategies employed in orthography design are summarized in Benton 1999, which also addresses the issue of uniformity among related languages. SIL produced grammars, dictionaries, and bible translations. Acunzo (1991) and others claim that introducing different orthographies was a strategy aimed at linguistically and socially fragmenting the Mixean community. As a reaction to this belief, the idea of a unified Mixean orthography was proposed at a regional summit in 1979 to serve as a vehicle of Mixean unification (Suslak 2003:558). A local task force was created for this endeavor, but they were faced with the problem of great linguistic diversity, and the orthography was never a success. Tables 6, 7, and 8 compare eight different Mixean orthographies and include corresponding IPA symbols and Spanish graphemes:

(1) The established Chuxnabán Mixe orthography (Chux)
(2) The Chuxnabán Mixe orthography used by a community member (Chux2)

Four orthographies developed by SIL for four distinct Mixean varieties:
(3) San José El Paraíso Mixe (S.José)
(4) Coatlán Mixe (Coat)
(5) Tototepec Mixe (Toton)
(6) Tlahuitoltepec Mixe (Tlah)

(7) The uniform Mixean orthography developed by the National Institute for Adult Education (INEA)
(8) The Mixean orthography used by De Quintana in the colonial period (Quint)

Chux	Chux2	S.José	Coat	Toton	Tlah	INEA	Quint	IPAº	Span
p	p	p	p	p	p	p	p	[p, b]	p
*b**	b	b	b	*b*	*b*	*b*, p	b	[b]	b, v
t	t	t	t	t	t	t	t	[t, d]	t
d	d	d	d	d	*d*	*d*, t	*d*	[d]	d
k	k	k	c, qu	k, qu	k	k	c, k, qu	[k, g]	c, k, qu
g	g	g	g (gu)	g	*g*	*g*, k	g	[g]	g
‘	‘	ʔ, ˀ	‘	‘	ʔ	‘	-	[ʔ]	-
m	m	m	m	m	m	m	m	[m]	m
n	n	n, ŋ	n, ng	n	n	n	n	[n, ŋ]	n
x	x	š, ž	x	x	š	x	x	[ʃ,ʒ]	-
j	j	h	j	j	h	j	h, j	[h]	g, j
ts	ts, dz	c, ẓ	tz, dz	ts	c	ts	tz	[ts, dz]	tz
ch	ch	c̃	ch	ch	cy	tsy	ch	[tʃ, dʒ]	ch
y	y	y	y	y	y	y	y	[j]	ie, ll, y

Chux	Chux2	S.José	Coat	Toton	Tlah	INEA	Quint	IPA°	Span
w	w	w	hu, *v*	v	w	w	hu	[w]	-, (u)
f	*f*	*f*	*f*	*f*	*f*	*f*	*f*	[f]	f
l	*l*	*l*	*l*	*l*	*l*	*l*	*l*	[l]	l
r	*r*	*ř*	*r*	*r*	*r*	*r*	*r*	[r]	r, rr
s	*z*	*s, z*	*s*	*s, z*	*s*	*s*	*s*	[s, z]	c, s, z

* Graphemes in italics occur only in loans from Spanish and other languages
° Not all allophonic variations are included in the IPA column
ˇ If part of the syllable nucleus

Table 6. Comparison of graphemes for consonants

Chux	Chux2	S.José	Coat	Toton	Tlah	INEA	Quint	IPA	Span
py	py	p̃	py, yp	py	py	py	-	[pʲ]	-
ty	ty	t̃	ty, yt	ty	ty	ty	-	[tʲ]	-
ky	ky	k̃	ky, yk	ky	ky	ky	-	[kʲ]	-
xy	xy	s̃	xy, yx	xy	šy	xy	-	[ʃʲ]	-
ch	ch	c̃	ch	ch	cy	tsy	-	[tʃ]	-
my	my	m̃	my, ym	my	my	my	-	[mʲ]	-
ny	ñ	ñ	ñ	ñ	ny	ny	ñ	[ɲ]	ñ

Table 7. Comparison of graphemes for palatalized consonants

Chux	Chux2	S.José	Coat	Toton	Tlah	INEA	Quint	IPA	Span
a	a	a	a	a	a	a	a	[a]	a
ä	a̲	-	-	e̲	ɔ	ä	â	[æ, ɔ]	-
e	e	e	e	e	e	e	e	[e]	e
i	i	i	i	i	i	i	i, y	[i]	i, y
ë	ʉ	ɨ	ɨ	a̲	ʌ	ï, ë	ê, î	[ɨ, ʉ]	-
o	o	o	o	o	o	o	o	[o]	o
ö	-*	-	-	o	-	ö	ô	[ø]	-
u	u	u	u	u	u	u	u	[u]	u
ü	-*	-	-	u̲	-	ü	û	[ʏ]	-

* This vowel quality has not been found in the source due to the fact that it occurs in verb forms, and very few verbs were collected

Table 8: Comparison of graphemes for vowels

Tables 6, 7, and 8 clearly illustrate the variation among the distinct orthographies, in particular among the four writing systems developed by SIL. The main areas where the orthographies diverge are (1) the allophonic voiced stops [b, d, g] and more generally the adoption of certain allophonic variations, (2) the representations of /k/ and /h/ in regards to Spanish spelling conventions, (3) less commonly found vowel qualities, and (4) suprasegmental palatalization. Some orthographies, such as the ones for Chuxnabán Mixe (Chux2), San José El Paraíso Mixe (S.José), and Coatlán Mixe (Coat), and the one used by De Quintana during the colonial period (Quint), represent some allophonic variations in their spelling conventions, in particular the voiced stops [b, d, g], which are phonemic in Spanish. Similarly, some orthographies represent the voiceless velar stop /k/ and the glottal fricative /h/ following Spanish spelling conventions with the graphemes *c, k, qu* and *j* respectively. Other than that, the representations of the consonantal phonemes demonstrate few differences. The Mixean vowel system is a major challenge orthography developers are faced with for two reasons: it is the source for most of the differences among the Mixean varieties, and all varieties have one or more vowel qualities for which there is no symbol on a basic keyboard. In addition, each Mixean variety has at least one phonemic vowel that does not occur in Spanish. As with the consonants, the four orthographies elaborated by SIL show considerable variation. In general, dieresis, diacritics, or IPA symbols are used for the special vowel qualities. The fourth major variation among the orthographies is the spelling conventions used for suprasegmental palatalization. Again, while some varieties follow Spanish spelling for the palatalized /n/, a phonemic palatalized consonant found in Spanish, others choose to represent suprasegmental palatalization for all consonants in the same manner, a consonant followed by a palatal glide /j/ or a consonant with a superscript tilde (see section 4.3).

Overall, the comparison of Mixean orthographies illustrates the difficulties encountered by developers and users when attempting to unify pre-existing orthographies. Moreover, it shows why the uniform writing system established by the National Institute for Adult Education (INEA) has not been successful. Compared to the other orthographies, the Chuxnabán Mixe writing system is exclusively phonemic and resorts to Spanish spelling conventions only in cases where confusion would result otherwise, such as with the glottal fricative /h/ and the post-alveolar affricate /ch/. The next section treats specific challenges that merit more detailed discussion.

4. CHALLENGES AND SOLUTIONS. This section discusses some of the challenges we were faced with when developing the orthography in more detail. These challenges relate to the complex vowel system, the spelling of Spanish loans, suprasegmental palatalization, and word boundaries. Solutions to these challenges and the reasons behind the decisions made are presented and examined.

4.1 VOWEL QUALITIES, VOWEL LENGTH, AND LARYNGEAL FEATURES. One of the most interesting and complex features of the Chuxnabán Mixe phonemic system is the vowels. In addition to nine vowel qualities, there is a phonemic contrast between short and long vowels and between modal, breathy or aspirated, glottalized or creaky, and interrupted

vowels. Overall, the contrasts result in five types of syllable nuclei: V, VV, VV^{h}, V^{ʔ}, and V^{ʔ}V. Three out of the nine vowel qualities, <ä, ö, ü>,[6] are generally the result of suprasegmental palatalization, as shown in the following examples.

(7) Marginal vowel qualities <ä, ö, ü> in palatalized environments

tsu'uch	'meat'	→ *chü'üch*	'his/her meat'
juuy	'to buy'	→ *jyüüy*	'he/she bought it'
jo'kp	'to weave'	→ *jyö'kypy*	'he/she weaved it'
noky	'paper'	→ *nyöky*	'his/her paper'
		inä'äny	'he/she said'
		kääky	'tortilla'

As a result, palatalization could be viewed as a contextual cue and the vowel qualities as fronted allophones of /a, o, u/ respectively. However, the latter also occur in palatalized environments.

(8) Vowel qualities /a, o, u/ in palatalized environments

tëykyaapxp	'to tell the truth'
tekychu'uk	'one-legged'
toopy	'to burn'

To date, no minimal pairs have been identified for the o/ö and u/ü contrasts, but there are several minimal pairs for the a/ä contrast.

(9) Minimal pairs: /a ~ ä/

taak	'mother'	*maajtsk*	'the one who grabs'
tääk	'suddenly'	*määjtsk*	'two'

However, the two low vowels <a, ä> are also used in free variation in the same word and by the same speakers, most likely due to dialect borrowing.

(10) Free variation of <a, ä>

taapë/tääpë	'this one'
maajtsk/määjtsk	'two'
tsak/tsäk	'dull'

[6] Note that angled brackets < > mark spelling.

Overall, it was decided to include the marginal three vowels in the orthography using dieresis and to accept both spellings for the words with the low vowels <a, ä> in free variation. The central high vowel, also written with dieresis, equally posed a challenge. It could be represented either as <ïz> or as <ë>, as its allophones range in vowel height from an /e/ to an /i/. The community preferred to use <ë>, probably because some of the orthography collaborators had had Mixean reading and writing lessons in another village where this symbol was used for a similar sound. The central high vowel posed an additional challenge. It is devoiced word-finally due to a phonological process, but in certain words it retains the voicing inherent to a nominalizer. This results in having voiced and devoiced variants in the same environment and potentially leading to homographs. This is shown below.

(11) Voiceless [ɨ̥] word-finally:

xu'kpë	'to smell'	[ˈʂuˀkpɨ̥]
tsookpë	'to heal'	[ˈtsoːkpɨ̥]
pëëjkpë	'pain'	[pɨːʰkpɨ̥]

Voiced [ɨ] word-finally:

maa'tspë	'thief'	[ˈmaːˀtspɨ]
yaa'o'kpë	'murderer'	[jaːˀoˀkpɨ]
o'kpë	'cadaver'	[oˀkpɨ]

To date, only one such homograph has been identified. However, this issue will need further investigation as language documentation progresses.

(12) Homograph: ääjtspï → [ˈæːhtspɨ] 'dancer' and [ˈæːhtspɨ̥] 'to dance'

The next issue relating to the vowel system is vowel length. First, it had to be determined how many distinctive vowel lengths there are. A three-way phonemic vowel length distinction has been noted for other Mixean varieties (Hoogshagen 1959 ; Van Haitsma & Van Haitsma 1976). Jany (2006, 2007) demonstrates that there is no evidence for such a three-way contrast in Chuxnabán Mixe, but that there is a clear phonemic distinction between short and long vowels. It was decided to represent long vowels as two consecutive vowels, the same as in all other Mixean varieties.

The laryngeal features also represented a challenge. As noted above, Chuxnabán Mixe shows a phonemic contrast between modal, aspirated, glottalized, and interrupted vowels. Aspiration occurs in the last portion of the vowel and could be represented either with the same symbol as the glottal fricative /h/ or with a superscript <ʰ> following the vowel. Regardless of the fact that the phonological use and phonetic realization of aspiration is clearly different from the glottal fricative /h/, the community decided to use the same symbol <j> for both. Although this facilitates writing on basic keyboards, it could potentially lead to homographs, since the glottal fricative /h/ also occurs in coda position and word-finally. To date, no such homographs have been identified.

Non-modal phonation contrasts in Chuxnabán Mixe depend on laryngeal timing. While breathiness occurs only in the last portion of the vowel, glottalization or creakiness can be found in the last, the middle, or the first portion of the vowel. These timing differ-

ences are related to differences in function. The first two involve the phonemic contrast between plain, glottalized, and interrupted vowels. The third occurs in vowel-initial words where a glottal stop is inserted initially to function as an onset. It was quickly decided to represent the glottal stop or creakiness with an apostrophe. The only question that remained open was whether to orthographically represent the glottal stop word-initially, as it is fully predicable in this position and, therefore, not phonemic. The community chose not to represent glottal stops word-initially.

4.2 SPANISH LOANS AND SPANISH-BASED SPELLING CONVENTIONS. Using orthographic conventions from the dominant language can have advantages and disadvantages. Advantages include familiarity with the spelling conventions for literate users, which can aid literacy, technological usability, and the possibility for the non-dominant language to gain status. Disadvantages comprise the possibility of internal inconsistencies and the reduction in emblemacy (Seifart 2006; Sebba 2007; Grenoble & Whaley 2006). The main reason for choosing to adopt such spelling conventions is the sociopolitical relation to the dominant language. While indigenous communities in Latin America often favor Spanish spelling conventions, some have opted against Spanish orthography rules to distance themselves from the dominant language (Grenoble & Whaley 2006). Benton (1999:5) notes that in Mexico indigenous communities generally want their language to look like Spanish. This has also been the case for Chuxnabán Mixe when adopting the grapheme <j> for the glottal fricative /h/. However, the community chose not to represent the Chuxnabán Mixe allophones and Spanish phonemes /b, d, g/, except in Spanish loans.

(13) Spanish loans with /b, d, g/ in Chuxnabán Mixe

cebu	'type of bull'	from Spanish	*cebú*	'type of bull'
lugarë	'place'	from Spanish	*lugar*	'place'
Dios	'God'	from Spanish	*Dios*	'God'

Overall, eight phonemes /b, d, g, f, s, ɾ, r, l/ only occur in Spanish loans, and the spelling in loans is generally not adapted to Chuxnabán Mixe orthographic conventions, with the exception of loans that show a phonological adaptation, such as the final high central vowel <ë>, or Mixean bound morphology. Certain Spanish loans seem to have entered the language some time ago and have undergone several changes. These loans do not follow Spanish spelling conventions.

(14) Older Spanish loans in Chuxnabán Mixe

krukts	'cross'	from Spanish	*cruz*	'cross'
nanwelë	'grandmother'	from Spanish	*abuela*	'grandmother'

Spanish loans generally undergo the same phonological processes as words of Mixean origin. Nevertheless, given that Spanish has additional phonemes not found in Chuxnabán Mixe, such as the voiced stops /b, d, g/, the voicing rules do not always apply in the same

way as in Mixean. For instance, in Spanish loans word-initial and word-final voiced stops may occur,[7] as well as intervocalic voiceless stops.[8]

(15) Mixean phonological processes ignored in Spanish loans

Word-initial voiced and intervocalic voiceless stop:	*botë* 'boot' ['botɨ]
Word-initial voiced stop:	*burrë* 'donkey' ['burɾɨ]

Yet, phonological processes triggered by affixing, such as obstruent voicing and metathesis, apply equally to Spanish loans.

(16) Mixean phonological processes in Spanish loans

Obstruent voicing: *n-* 'my'+ *zapat* 'shoe'	→ *nzapat* 'my shoe' [nzaˈpaːt]
Obstruent voicing: *carton* 'box'+ *joch* 'in'	→*cartonjoch* 'in the box' [kaɾtonˈɦotʃ]
Metathesis: *y-* 'his, her' + *burrë* 'donkey'	→ *byurrë* 'his donkey' ['bjuɾɨ]

In compounding, however, the voicing rules do not always apply:

(17) Mixean phonological processes ignored in compounds

cerë 'wax' from Spanish *cera* + *pa'ak* 'sweet' → *cerëpa'ak* 'honey' [seːɾɨˈpaʔak]

To sum up, the Chuxnabán Mixe community is open to adopting Spanish spelling conventions where necessary. Spanish orthography is used in all Spanish loans, unless there are significant phonological adaptations or changes present.

4.3 SUPRASEGMENTAL PALATALIZATION. Palatalization in Chuxnabán Mixe, as in other Mixean languages (Hoogshagen & Hoogshagen 1997; Schoenhals & Schoenhals 1982; Van Haitsma & Van Haitsma 1976; Dieterman 2008), is a suprasegmental process affecting not only the palatalized consonant, but adjacent vowels as well. This palatalization is perceived as an onglide and offglide and is manifested by a change in the formant structure of adjacent vowels, lowering F1 and raising F2. The latter indicates fronting. One exception is palatalized /j/ (Dieterman 2008; Jany 2006). The palatal glide /j/ does not undergo any changes with morpheme-induced palatalization. Suprasegmental features, such as palatalization and tone, often represent challenges for orthography developers. In the case of palatalization, it has to be decided whether to write the phonetic effects and to what degree. One of the strongest phonetic effects is the perception of an onglide and

[7] The voiced allophones [b, d, g] do not occur in these positions in Chuxnabán Mixe, except in shortened forms.
[8] In Chuxnabán Mixe the voiced allophone would occur in intervocalic position.

an offglide with most consonants. The community decided to represent the offglide, but not the onglide. Furthermore, while suprasegmental palatalization in Mixean varieties is manifested as a secondary articulation of the consonants /p, m, h, ʔ, w/, it also changes the primary position of the alveolar and velar consonants /t, k, x, ts, n/, moving them toward the palatal region (Dieterman 2008). This is most apparent for the alveolar affricate /ts/, which changes to /tʃ/, a phoneme in Chuxnabán Mixe and in Spanish. Therefore, it was determined to also change its representation to the grapheme *ch*, in order to represent the post-alveolar affricate.

(18) Post-alveolar affricate /tʃ/ resulting from suprasegmental palatalization

y- '3rd person possessive' + *tëëjk* 'house' → *chëëjk* [tʃɨːʰk] 'his/her house'

Morpheme-induced suprasegmental palatalization occurs at word-edges, such as when the third-person possessive *y-* is added, as in the example above. A different palatalization process occurs in compounds word-medially and is not morpheme-induced. If a word ending in the palatal glide /j/ precedes another word in a compound, the first consonant of the second word is palatalized and voiced. It was decided to write the onglide and the offglide in these cases, because the onglide is lexically motivated and not the result of a suprasegmental process.

(19) Palatalization in compounding

tëy 'truth' + *kaapxp* ' to speak'→ *tëykyaapxp* [tɨjˈgjaːpʂp] 'to tell the truth'

Phonetically this process is different from morpheme-induced palatalization in that it also causes voicing of the following segment. To conclude, suprasegmental palatalization is represented with a palatal glide <*y*> following the palatalized consonant. If palatalization is not suprasegmental and results from a phonological process, as in compounding, the triggering palatal glide remains, and both the onglide and the offglide are written. This rule implies a certain grammatical and lexical knowledge of the language and may complicate the transcription of spoken language, but the community was confident that they knew where to write the onglide and where not to write it.

4.4 WORD BOUNDARIES. When writing a language one has to decide what constitutes a word in that language. This can be a highly complex issue, since there may be conflicting prosodic, morphosyntactic, and semantic criteria. Words are recognized as entire units and processed as such; they are also the basic units for language processing in reading (Seifart 2006:282). Generally, speakers of a language are able to identify word boundaries. However, when it comes to compound words, one could either break them up or write them as single words. Benton (1999:3) suggests two deciding factors for word breaks: (1) considering phonological factors pointing to a word break and (2) assessing whether native speakers think of a morpheme string as a unit or not. A list of compound words was not available when developing the orthography, and the decision whether to write compounds as one

word or not is made on a case by case basis. In general, speakers showed no preference in this regard, but three phonological processes occurring within words provide a clue. One is consonant cluster reduction, which occurs when the final consonant of the first word is the same as the initial consonant of the second, thereby creating underlying geminates.

(20) Consonant cluster reduction in compound words

tëëjk 'house' + *koojp* 'scatter seed' → *tëëjkoojp* 'build'
tutk 'turkey' + *kopk* 'mountain' → *tutkopk* 'plant type'

The second phonological process is obstruent voicing, which occurs only within words, but not across word boundaries.

(21) Obstruent voicing in compound words

naax 'earth' + *yuujk* 'animal'	→ *naaxyuujk* 'worm'		[na:jˈʒju:ʰk]
wiin 'eye' + *tu'uk* 'one'	→ *wiintu'uk* 'one-eyed'		[wi:nˈduuk]
kaajpn 'village' + *kopk* 'mountain'	→ *kaajpnkopk*	'capital'	[ka:ʰpnˈgopk]
jëën 'fire' + *pixk* 'flea'	→ *jëënpixk* 'spark'		[hɨ:nbiʂk]

The third phonological process is phoneme-induced palatalization, whereby the last phoneme in the first word is a palatal glide /j/ and triggers palatalization of the following consonant.

(22) Phoneme-induced palatalization in compound words (same as example 19)

tëy 'truth' + *kaapxp* 'to speak' → *tëykyaapxp* [tɨjˈgja:pʂp] 'to tell the truth'

Other compounds show no phonological processes that could provide clues about word boundaries, but they are equally written as one word.

(23) Compound words with no phonological cues

i'px 'twenty' + *teky* 'foot' → *i'pxteky* 'centipede'
yo'k 'neck' + *paajk* 'bone' → *yo'kpaajk* 'Adam's Apple'
wiin 'eye' + *waay* 'hair' → *wiinwaay* 'eyelash'
tsaajp 'heaven, sky' + *tëëjk* 'house' → *tsaajptëëjk* 'church'

The examples of compound words presented above do not pose a major challenge to readers, but as some orthography developers have noted (Guérin 2008:60, Benton 1999:3), long words are more difficult to read, especially for beginning readers. Given that Chuxnabán Mixe is a polysynthetic language, the verbal complex can include multiple prefixes and suffixes, in addition to root compounding, thereby creating very long strings of morphemes. Since there are no ideal breaks in such words, they are written as one word.

(24) Long words

ntuukjotwiijtpyën	'I saw it begin'
nkosnëktë'ënanë	'I feel like going to the bathroom'
jyä'äxpüxëtënä'ä	'He was cutting wood'
nyeekë'utsuktu'tënë	'He cut it in order to get out'

No word breaks can be inserted in the examples above, because there is bound morphology at both word edges, and speakers perceive them as single units.

To conclude, three phonological processes serve as cues for word boundaries in certain compounds. However, in the majority of cases such cues are not available. The challenge of word breaks has only minimally been discussed in the community, and community members generally do not feel strongly about writing compounds as one or more words. Rehg (2004:504) notes that issues such as word boundaries, punctuation, and capitalization only gain attention after an alphabet has been developed and implemented. It is expected that these issues will be revisited and further discussed during future community meetings.

5. SUMMARY AND CONCLUSIONS. The present paper describes the early stages of orthography design for Chuxnabán Mixe, a somewhat endangered Mexican indigenous language spoken in Oaxaca. These stages are very important and can have a lasting impact determining the success of the documentation and literacy development process. The Chuxnabán Mixe orthography was developed in collaboration with the community. My role was to provide guidance with linguistic issues and spelling options. All decisions were made by community members, and the resulting orthography was informally tested in the village and approved.

The orthography design follows the trends discussed by Grimes & Gordon (1980), the recommendations proposed by Grenoble & Whaley (2006), and the suggestions made by Rehg (2004). In general, one symbol represents one phoneme, and the orthography was built on existing practices in other Mixean varieties and in the dominant language, Spanish, although dialectal idiosyncrasies were retained. Most symbols are readily available on basic keyboards, and informal testing has shown that speakers are at least able to read the language. While most decisions were easily made, the complex vowel system, certain phonological processes—such as obstruent voicing and suprasegmental palatalization—and the presence of Spanish loans posed some challenges. Understanding the phonology, and to a degree the morphology, of the language was imperative to resolving these challenges. Not all issues have been addressed during these first orthography-design meetings in the community. Some, such as how to deal with casual speech or shortened forms, the unreleased /p/ that seems to occur in all words ending in the bilabial nasal /m/, punctuation and capitalization, and dialect borrowing, were identified and postponed for future meetings.

In conclusion, this work illustrates the orthography development process for a previously undocumented language and presents some of the issues that need to be addressed when undertaking such an endeavor. Some of the decisions made during this early stage may need to be revised after the documentation process has progressed and after extensive testing has been conducted.

References

Acunzo, Mario. 1991. *Educación e identidad étnica: El caso de la comunidad mixe de Santa María Tlahuitoltepec (México)*. Ecuador: Abya-Yala.

Benton, Joseph P. 1999. How the Summer Institute of Linguistics has developed orthographies for indigenous languages of Mexico. Summer Institute of Linguistics. http://www.sil.org/mexico/ilv/L001a-ILVOrto.htm. (1 July, 2010.)

Campbell, Lyle, Terrence Kaufman & Thomas C. Smith-Stark. 1986. Meso-America as a linguistic area. *Language* 62(3). 530–570.

Crawford, John Chapman. 1963. *Totontepec Mixe phonotagmemics*. Summer Institute of Linguistics, No. 8. University of Oklahoma.

De La Grasserie, Raoul. 1898 (reprint 1968). *Langue Zoque et Langue Mixe: Grammaire, dictionnaire*. Bibliothèque Linguistique Américaine, Tome XXII. Paris. (Reprint: Kraus. Nendeln, Liechtenstein).

De Quintana, Fray Augustín. 1733 (printed 1890). *Confesionario en Lengua Mixe*. Alençon.

Dieterman, Julia. 2008. Dallas, TX: Summer Institute of Linguistics. www.sil.org/silepubs/Pubs/.../50951_DietermanJ_Mixe_Palatalization.pdf.(1 July, 2010.)

Gregorio Cirilo, Cosme. 2005. *Vocabulario Mixe*. Unpublished ms.

Grenoble, Lenore & Lindsay Whaley. 2006. *Saving languages: An introduction to language revitalization*. Cambridge: Cambridge University Press.

Grimes, Joseph & Raymond Gordon, Jr. 1980. Design of new orthographies. In James Kavanagh & Richard Venezky (eds.), *Orthography, reading, and dyslexia*, 93–103. Baltimore, MD: University Park Press.

Guérin, Valérie. 2008. Writing an endangered language. *Language Documentation & Conservation* 2(1). 47–67.

Hinton, Leanne. 2001. New writing systems. In Leanne Hinton & Ken Hale (eds.), *The green book of language revitalization in practice*, 239–250. San Diego, CA: Academic Press.

Hoogshagen, Searle. 1959. Three contrastive vowel lengths in Mixe. *Zeitschrift für Phonetik und allgemeine Sprachwissenschaft* 12. 111–115.

Hoogshagen Noordsy, Searle & Hilda Halloran Hoogshagen. 1997. *Diccionario Mixe de Coatlán Oaxaca*, 2nd edn. (1st edn. 1993.) Mexico: Summer Institute of Linguistics.

Hornberger, Nancy H. (ed.). 1997. *Indigenous literacies in the Americas: Language planning from the bottom up*. Contributions to the Sociology of Language #75. Berlin: Mouton de Gruyter.

INEA (Instituto Nacional para la Educación del los Adultos). 1994. *Ayuk Ja'jy Y'ayuujk. Libro del Adulto, Población Mixe (Parte Baja), Oaxaca*. http://cdi.gob.mx/index.php?id_seccion=519. (1 July 1, 2010.)

INEA (Instituto Nacional para la Educación del los Adultos). 1997a. *Ayuk Ja'jy Y'ayuujk. Libro del Adulto, Población Mixe (Parte Alta), Oaxaca*. http://cdi.gob.mx/index.php?id_seccion=519. (1 July 1, 2010.)

INEA (Instituto Nacional para la Educación del los Adultos). 1997b. *Instructivo del Alfabetizador. Población Mixe, Oaxaca*. http://cdi.gob.mx/index.php?id_seccion=519. (1 July 1, 2010.)

INEA (Instituto Nacional para la Educación del los Adultos). 1997c. *Já'ájy m'ayuujk. Cuaderno de Ejercicios, Población Mixe, Oaxaca*. http://cdi.gob.mx/index.php?id_seccion=519. (1 July 1, 2010.)

Jany, Carmen. 2006. Vowel length and phonation contrasts in Chuxnabán Mixe. *Santa Barbara Papers in Linguistics 18: 9th Annual Workshop on Native American Languages*. University of California, Santa Barbara.

Jany, Carmen. 2007. Phonemic versus phonetic correlates of vowel length in Chuxnabán Mixe. *Berkeley Linguistic Society Proceedings 33S: Languages of Mexico and Central America*. University of California, Berkeley.

Jany, Carmen. 2009. Digital Fluency and Participatory Research: The Chuxnabán Mixe Online Dictionary. *Santa Barbara Papers in Linguistics: 12th Annual Workshop on Native American Languages*. University of California, Santa Barbara.

Lewis, M. Paul (ed.). 2009. *Ethnologue: Languages of the world*. 16th edn. Dallas, TX: SIL International. http://www.ethnologue.com/. (29 August, 2010.)

Rehg, Kenneth. 2004. Linguists, literacy, and the law of unintended consequences. *Oceanic Linguistics* 43(2). 498–518.

Rice, Keren & Leslie Saxon. 2002. Issues of standardization and community in Aboriginal language lexicography. In William Frawley, Kenneth Hill & Pamela Munro (eds.), *Making dictionaries: Preserving indigenous languages of the Americas*,125–154. Berkeley, CA: University of California Press.

Ruiz De Bravo Ahuja, Gloria. 1980. *Mixe Tlahuitoltepec, Oaxaca*. Archivo de Lenguas Indígenas de México. (Data provided by Don D. Lyon.)

Schoenhals, Alvin & Louise C. Schoenhals. 1982. *Vocabulario Mixe de Totontepec*, 2nd edn. (1st edn. 1965.) Hidalgo, Mexico: Summer Institute of Linguistics.

Sebba, Mark. 2007. *Spelling and society: The culture and politics of orthography around the world*. Cambridge: Cambridge University Press.

Seifart, Frank. 2006. Orthography development. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation*, 275–299. Berlin: Mouton de Gruyter.

Suslak, Dan. 2003. The Story of ö: Orthography and cultural politics in the Mixe Highlands. *Pragmatics* 13(4). 551–563.

Torres Cisneros, Gustavo. 1997. *Mixes*. Comisión Nacional para el Desarrollo de los Pueblos Indígenas. http://www.cdi.gob.mx. (1 July, 2010.)

UNESCO Ad Hoc Expert Group on Endangered Languages. 2003. *Language Vitality and Endangerment*. Paris. http://www.unesco.org/culture/ich/doc/src/00120-EN.pdf. (1 July, 2010.)

Van Haitsma, Julia Dieterman & Willard Van Haitsma. 1976. *A hierarchical sketch of Mixe as spoken in San José El Paraíso*. No. 44. Mexico: Summer Institute of Linguistics.

Venezky, Richard L. 2004. In search of the perfect orthography. *Written Language & Literacy* 7(2). 139–163.

Wichhmann, Søren. 1995. *The Relationship among Mixe-Zoquean languages of Mexico*. Salt Lake City, UT: University of Utah Press.

Carmen Jany
cjany@csusb.edu