

STATEMENT OF CURRENT ORTHOGRAPHY

LANGUAGE: Mixtec VARIANT: Magdalena Peñasco STATE: Oaxaca
TRANSLATORS: Bruce and Barbara Hollenbach DATE: September, 2012

(A copy of this statement must be on file with the orthography consultant in order for you to obtain permission to publish. If it is tentative, indicate below with an asterisk all symbols which are considered tentative and are being tested in one way or another. On a separate sheet describe the orthographical alternatives you are considering. Your orthography consultant should send a copy to the permanent SFP file and notify the PubCoord, Literacy, Linguistics and FTS of approvals.)

Tentative orthography approved by Chuck Speck, February, 1997

Working orthography approved by Chuck Speck, November 1997

Updated working orthography form submitted in August 1999

(NOTE: All of the above were submitted by e-mail, and were not on the regular orthography form.)

A further updated working orthography statement was submitted in March 2009. In October, 2009, one further change was made in the tone orthography at the request of our Mixtec co-workers (the accent that marks present tense and subjunctive mood will always be marked on the first syllable, even if the high tone is on the second syllable).

In January 2010, I submitted a new version of the statement incorporating this change in order to request an upgrade to established orthography status. And this revision, submitted in September 2012, adds a new phoneme to the inventory that occurs in only one word, which recently came to our attention.

We are using the same basic symbols we used in February 1997, which follow very closely the system proposed by the Mixtec Academy, and it is hard to imagine that anything else would be feasible. Our Mixtec co-workers can read it very fluently, and we never hear anyone suggest a change in these symbols.

Since 1999, I have changed the way I analyze syllable breaks in words with long and interrupted vowels, and I have also discarded the idea that ñ [ɲ] is an allophone of y [j]. Neither of these analytical changes affects the orthography; we continue to write VV, V'V, and ñ.

Over the intervening years, I have completed an analysis of the tone system, and we are now writing tone in some words, as we see that our co-workers need help in reading. See the description in section 3.2 below.

I have copied in the text I submitted back in 1997 at the bottom of this form, and the only changes I needed to make in it were to fix the spelling of two words that we originally heard wrong (this is, note, not an orthographic change).

The area that is perhaps most difficult for Mixtec is deciding where to write word breaks, and we are still debating some of these issues for compounds and other words with more than one morpheme. Some are written as one word, some are written as two words, and still others are written with a hyphen, as explained in section 4.2 below. Our co-workers seem able to read any of these quite well.

1. VOWELS

List all of the vowel phonemes of the language in phonetic chart order.

There are five vowels, the same ones that occur in Spanish.

	<i>Front</i>	<i>back unrounded</i>	<i>Back rounded</i>
<i>close</i>	i		u
<i>close mid</i>	e		o
<i>open</i>		a	

List below the phonemes from the chart above with other information as indicated. For any phoneme which has more than one phonetic variant, use a separate line for each variant.

SYMBOLS			EXAMPLE	GLOSS	EXPLANATIONS / COMMENTS
<i>Phnm</i>	<i>Phnt</i>	<i>Orth</i>			
/i/	[i]	i	titi	‘flower in form of a ball’	
/e/	[e]	e	sete	‘to cut hair’	
/a/	[a]	a	tata	‘seed corn’	
/o/	[o]	o	toto	‘rock’	
/u/	[u]	u	tutu	‘paper’	
/u/	[u̠]	u	tutu	‘paper’	optional variant for some speakers

See section 3. Prosodic Elements for a description of long, interrupted, and checked vowels, and also of nasalization.

2. CONSONANTS

List all of the consonant phonemes of the language in phonetic chart order.

	<i>Labial</i>	<i>Labio-dental</i>	<i>Dental</i>	<i>Alveo-palatal</i>	<i>Retro-flex</i>	<i>Palatal</i>	<i>Velar</i>	<i>Labio-velar</i>	<i>Glottal</i>
vl. stop	p		t				k (g)		
vd. stop	b		d (ð)				g (ɣ)		
affricate				tʃ					
vl. fricative		f	s	ʃ	ʂ		x (h)		
vd. fricative	β								
nasal	m		n (ɲ)			ɲ			
trill			r						

flap			r					
lateral			l					
glide						j (ʃ, ʒ)	w	
laryngeal								ʔ

List below the phonemes from the chart above with other information as indicated.

For any phoneme which has more than one phonetic variant, use a separate line for each variant. Use another sheet if needed.

SYMBOLS			EXAMPLE	GLOSS	EXPLANATIONS / COMMENTS
Phom	Phon	Orth			
			Stops		
/p/	[p]	p	tipii	‘warbler’	This phoneme occurs only in a few onomatopoeic words, and in many loans from Spanish.
/t/	[t]	t	tutu	‘paper’	
/k/	[k]	k	koo	‘snake’	
/k/	[g]	k	minki jin kata	‘brains’ ‘sing (fut. pl.)’	This allophone of /k/ occurs within roots following the nasal consonant /n/. It also occurs at the beginning of a root following a nasal proclitic or prefix. The voicing is especially strong following the future plural jin and the subjunctive plural jín, but we write k to preserve the form of the root, and people read it with no trouble.
/k/	[kŋ]	k	ki'in	‘to go (fut.)’	This allophone of /k/ occurs optionally in nasal words preceding a nasalized vowel. We have not written this, but it could be written kn (kni'i ‘to go’). Speakers seem to be fairly aware of it, but it is not as stable as tn (see clusters below).
/b/	[b]	b	lambombon	‘sound of bell’	This phoneme occurs only in a few onomatopoeic words, and in many loans from Spanish.
/d/	[d]	d	dañu	‘harm’	This allophone occurs only in loans.
/d/	[ð]	d	de lado	‘he’ ‘side’	This allophone occurs in two enclitic pronouns, and in loans; in loans it occurs in intervocalic position, as in Spanish.
/g/	[g]	g gu	gastu guitarra	‘expense’ ‘guitar’	This phoneme occurs only in loans.
/g/	[ɣ]	g gu	amigo águila	‘friend’ ‘hawk’	This allophone occurs only in loans; it occurs in intervocalic position, as in Spanish.
			Affricates		

/tʃ/	[tʃ]	ch	chuku	‘louse’	
			Fricatives		
/β/	/β/	v	vii	‘clean’	
/f/	[f]	f	foto	‘photograph’	This phoneme occurs only in loans.
/s/	[s]	s	sutu	‘priest’	
/ʃ/	[ʃ]	x	uxi	‘ten’	This alveopalatal sibilant occurs in only one word, in root-medial position, and it is written with the same symbol as the retroflex sibilant. Note that the [ʃ] allophone of /y/ occurs only in root-initial position.
/ʂ/	[ʂ]	x	xixi	‘aunt’	This retroflex sibilant corresponds to the alveopalatal sibilant in other towns.
/x/	[x]	j	jito	‘bed’	The friction is very light.
/x/	[h]	j	jito	‘bed’	The [h] allophone varies freely with [x].
			Nasals		
/m/	[m]	m	ma'a mesa	‘raccoon’ ‘table’	This phoneme occurs in nasal morphemes and in loans.
/n/	[n]	n	nani Nuyu	‘to be called’ ‘Cenobio’	This phoneme occurs in nasal morphemes and in loans.
/n/	[ŋ]	n	ñenku	‘desnutrido’	This allophone occurs in cluster preceding a velar stop in a limited number of words, many of which are loans.
/ɲ/	[ɲ]	ñ	ñuu albañil	‘town’ ‘mason’	This phoneme occurs in nasal morphemes and in loans.
/ɲ/	[ỹ]	ñ	ñuu albañil	‘town’ ‘mason’	This allophone varies freely with the above before high vowels.
			Trill		
/r/	[r]	r rr	rosa sarruun karru	‘rose’ ‘sound of wings’ ‘vehicle’	This phoneme occurs in a few onomatopoeic words and in loans; it is spelled r or rr according to the rules of Spanish.
			Flap		
/ɾ/	[ɾ]	r	ri ro'o triu	‘yo’ ‘tú’ ‘trigo’	This phoneme occurs in four pronouns (two enclitic and two independent) and in loans
			Lateral		

/l/	[l]	l	lulu lii kalu	‘baby’ ‘smooth’ ‘broth’	This phoneme is somewhat marginal in native words, and they usually have a diminutive or similar meaning; it occurs freely in loans.
			Glides		
/j/	[ʃ]	y	yoso	‘metate’	This phoneme is a voiceless sibilant in root-initial position, but the contrast with /s/ is maintained because the latter is always retroflex. Everyone is happy to write this [ʃ] sound with the letter y.
/j/	[ʒ]	y	kuiya yo ya	‘year’ ‘habitual’ ‘He (deity)’	This phoneme is a voiced sibilant in intervocalic position, and at the beginning of a proclitic or enclitic. Again, everyone is happy to write this sound with the letter y.
/j/	[j]	i y	tioko siu'u yelo	‘ant’ ‘to frighten’ ‘ice’	This phoneme is a glide when it follows another consonant, and it is written with the vowel i. In the speech of bilinguals, loans with initial /j/ are pronounced with a glide, whereas monolinguals use [ʃ], as they would in native words. We write y for both.
/w/	[w]	w u	uwa wa'a kuiya anua puente	‘bitter’ ‘gives’ ‘year’ ‘heart’ ‘bridge’	This phoneme is infrequent in intervocalic or root-initial position. It is very common following /k/, where it is written u. It also follows other consonants in loans.

			<i>Laryngeals</i>		
/ʔ/	[ʔ]	'	ve'e u'vi	'house' 'painful'	<p>Glottal stop is a frequent sound in the language, but a problematic one. It occurs in interrupted and checked vowels, where it is written with a vertical stroke (saltillo).</p> <p>It is also written with this stroke in extrasystemic interjections.</p> <p>Another use of glottal stop is at the end of yes/no questions (also content questions for some speakers), and here we do not write the stroke, but let the question marks on the sentence carry it. Speakers put the glottal stop in automatically.</p> <p>Glottal stop also occurs at the beginning of a root that would otherwise be vowel-initial. Speakers seem to be unaware of it in this position, and we simply write a word space.</p> <p>In compounds in which the second element is vowel-initial, we use a hyphen.</p>
		¿?	jm' ¿Ki'in ro?	'I don't know' 'Are you going?'	
		space	uu ita	'two flowers'	
		hyphen	la-ina	'dog'	
<i>Important clusters (consonantes compuestas)</i>					
/kw/	[k ^w]	ku	kuee	'slow'	This sequence is very common in native words.
/tn/	[tn]	tn	tnuu	'black'	This sequence occurs in nasal morphemes.
/nd/	[nd]	nd	ndaa ndijan kuenda	'straight' 'sandal' 'account'	This sequence occurs in oral morphemes, initially in some two-syllable nasal roots, and in loans.

NOTES ON FORMS USED IN OTHER TOWNS

The information given above for Magdalena Peñasco applies fairly well to the surrounding towns. I am aware of the following differences.

San Agustín Tlacotepec

This town has developed a sibilant offglide following /t/ ([t̪s]) and /nd/ ([nd̪z]) before the vowel /i/; these are treated as clusters and written ts and nts.

kitsi	animal	(cp. Magdalena kiti)
ntsi'i	to finish	(cp. Magdalena ndi'i)

The enclitic pronoun ri is pronounced with a strong trill [r] rather than a flap [ɾ], and it is written rri.

Ñatuu ní ja'an rri.	No fui.	(cp. Magdalena Ñatuu ní ja'an ri.)
---------------------	---------	------------------------------------

San Mateo Peñasco

The four enclitic and independent pronouns that have a flap [r] in Magdalena have a strongly velarized lateral [lʷ] in San Mateo. Even though it contrasts with the clear l [l] found in loans and in a few other words, both are written with the letter l.

¡Ma sa'a lo! Don't do it! (cp. Magdalena ¡Má sa'a ro!)

San Pedro El Alto

This town has six vowels, the five of Magdalena plus the high central vowel, /ɨ/ [ɨ], which is written ɨ.

ñii salt (cp. Magdalena ñii)

Most surrounding towns

The towns around Magdalena (all of the above plus San Cristóbal Amoltepec) do not have devoicing of root-initial /j/; it is pronounced [ʒ], rather than /ʃ/. In these towns, there is an /ʃ/ phoneme that is not retroflex, rather than the retroflex /ʃ̣/ found in Magdalena (except in the word uxi 'ten'). Note that the contrast between the fricative /ʃ/ (or /ʃ̣/) and the glide /j/ is clearly maintained in each town.

		Magdalena	San Cristóbal and others
yoo	moon	[ʃo:]	[ʒo:]
xito	uncle	[ʃito]	[ʒito]

(Note that the situation in Teita is different: the /j/ has been devoiced in root-initial position, and this has resulted in the loss of the contrast between /j/ and /ʃ/ in that position. The orthography uses x for both.)

NOTE ON NASAL VS. ORAL WORDS

When loans are taken into consideration (and Magdalena incorporates Spanish words very freely), it simply does not work to set up [m] and [ɲ] as allophones of /w/ and /j/ that occur in nasal words, even though this conditioning factor seems to be beyond question as the historical source of elements like m, ñ, tn, and nd. Also, the heavy pressure exerted by the prestige given to the Spanish alphabet works against any attempt to set up m and ñ as allophones of w and y. I have never seen a Mixtec spontaneously propose writing w or v instead of m, or of writing y instead of ñ, or t instead of tn. In addition, the Mixtec Academy seems to be quite unaware of these historical relations. The relation between n and nd is even more problematic. I have never seen a Mixtec propose that we write nd without the d; instead they debate about writing nd versus nt.

Note the difference between tniñi 'mouse' and the loan tina 'basin' (not *tnina). Note the difference among nduyu 'stake' and ñuñu 'bee' vs. the loan Nuyu 'Cenobio'. The name has not been adapted to either *Nuñu or *Nduyu.

There are also some exceptions to the generalization that nd occurs only in oral roots. We get ndijan 'sandal' and a number of other nasal roots that have initial nd.

There is also in San Cristóbal Amoltepec a reduced form of yukan 'that', which is yuun. It seems to be on its way to becoming a definite article. This word has not been adapted to the rule about using [ɲ] in nasal words and changed to *[ɲũũ]; it is pronounced [ʒũũ].

3. PROSODIC ELEMENTS

List all contrastive features of this language that were not treated under the consonant or vowel section. For example, tone, stress, length, glottalization, aspiration, palatalization, labialization, nasalization. Include all symbols used for each feature and list all of the sounds with which that feature occurs. (Example: length: ii, ee, aa, oo, uu.) List below the features with other information as indicated. For any feature which has more than one phonetic variant, use a separate line for each variant.

FEATURE	SYMBOLS			EXAMPLE	GLOSS	EXPLANATIONS / COMMENTS
	Phom	Phon	Orth			
Long vowels (VV)	/aa/ /ee/ /ii/ /oo/ /uu/	[a:] [e:] [i:] [o:] [u:]	aa ee ii oo uu	lasaa nee kandii koo nuu	'bird' 'dark' 'sun' 'snake' 'face'	Long vowels are treated as sequences of two like vowels. Such sequences occur only at the end of a word, and the sequence forms part of two syllables, each with its own tone.
Interrupted vowels (V?V)	/a?a/ /e?e/ /i?i/ /o?o/ /u?u/	[a?a] [e?e] [i?i] [o?o] [u?u]	a'a e'e i'i o'o u'u	sa'a ve'e yi'i ko'o yu'u	'to do' 'house' 'weeds' 'plate' 'mouth'	A sequence of two like vowels interrupted by a glottal stop is treated as a sequence of three phonemes. Such sequences occur only at the end of a word, and the sequence forms part of two syllables, each with its own tone.
Checked vowels (V?)	/a?/ /e?/ /i?/ /o?/ /u?/	[a?] [a?a] [e?] [e?e] [i?] [i?i] [o?] [o?o] [u?] [u?u]	a' e' i' o' u'	ka'nu ja'un ndiva'u nde'ya ne'un li'li ko'lo su'ma	'big' 'fifteen' 'coyote' 'peach' 'among' 'rooster' 'to snore' 'tail'	Vowels checked by /?/ are treated as sequences of a vowel and the glottal stop phoneme. They occur mainly in the penultimate syllable of a word. They usually precede a sonorant consonant, but in a few words, they precede an unlike vowel. When checked vowels precede a consonant, they very often have rearticulation [ka?anu]. Even though native speakers are very aware of this, we have chosen not to write it. The rearticulated phone never carries a contrastive tone.

Nasalization	Vn	[ã] [ẽ] [ĩ] [õ] [ũ]	an en in on un NV	kuaan ndijan ke'en kuijin kokon xu'un maa kani ñu'u tnu'u	'yellow' 'sandal' 'to take' 'white' 'thick' 'money' 'self' 'long' 'fire' 'word'	Nasalization is a feature of the morpheme. It is always heard on the final vowel, and it carries back through /ʔ/ and /j/ until it hits a consonantal barrier. We write it with a word-final n if the morpheme does not contain any nasal consonants, or if it contains nd. If a native morpheme has a nasal consonant (m, n, ñ, tn), it is always nasal, and the final n is not written. See note below for spread of nasalization to enclitics.
Palatalization	CjV	[C ^j V]	CiV	tiaka sia'a palio	'fish' 'to pass' 'canopy'	Palatalization of a consonant is marked by a very strong offglide, and palatalized consonants are treated as clusters containing /j/. The /tj/ cluster is usually a contraction of a ti- prefix meaning animal, and the /sj/ cluster is found in causative verbs. Other clusters with /j/ are found in loans.
Labialization	CwV	[C ^w V]	CuV	kuee fuersa sueter	'slow' 'strength' 'sweater'	Labialization of a consonant is marked by a very strong offglide, and labialized consonants are treated as clusters containing /w/. In native words, /kw/ is very common; other clusters are found in loans.

NOTE ON LONG AND INTERRUPTED VOWELS

I have opted for treating long and interrupted vowels as sequences because each vowel mora carries a tone. Words that end in long and interrupted vowels (CVV and CV'V) do, however, behave somewhat differently from words with CVCV or CV'CV in terms of certain tone-spreading rules, and it would be equally rational to treat long and interrupted vowels as single complex vowels that occur only at the end of words, create a heavy syllable in that position, and attract the word stress. In terms of orthography, however, it does not matter which analysis one adopts because they will be written VV and V'V in either case.

NOTE ON NASALIZATION

Nasalization of the root spreads to the right and is heard on vowel-initial enclitics, but we write the underlying form of the enclitic. See treatment of vowel clusters in section 4.6 below.

nuu o	our faces	[nũ: õ]
nuu o a	our faces	[nũ: õ ã]
tinana i	the child's tomato	[tinãñã ĩ]

Ka'an i.	the child will speak	[kãʔi]
Ndene'e o i.	we will take it out	[ndenẽʔõ i]

The nasalization of a nasal enclitic also spreads to the right, and is heard on a second enclitic.

ve'e ni a	your house	[veʔe nĩ ã]
-----------	------------	-------------

Nasal proclitics, especially those that do not have a nasal consonant, like the future plural jin, voice a following k and sometimes ch. We write the basic form with k and ch.

jin kata o	we will sing	[jĩ gata o]
jin chi'i o	we will sow	[jĩ d͡ʒiʔo]

For many speakers, ni 'past tense', even though nasal, does not have this strongly voicing effect.

List below features that are not symbolized or that are symbolized only sporadically, and give a brief statement of why they are not symbolized.

3.1 Stress

The canonical form of a root contains two syllables, and stress normally occurs on the first of these.

[ʔi.te]	ite	grass
[ʔe.ʔe]	ve'e	house
[ʔiʔ.nde]	vi'nde	nopal cactus
[ʔi.i]	vii	clean

When the tone of the second syllable is high and the first is mid or low, stress is sometimes heard on the high-tone syllable, especially if this syllable is followed by a word with a lower tone.

[ki.ʔi]	kixí	is sleeping
[ʔũ.ʔũ]	xu'un	money

In compounds with more than two syllables, the first syllable of a three-syllable word often has secondary stress, especially in deliberate speech.

[,ti.'nã.nã]	tomato
--------------	--------

Stress is not symbolized in the orthography, since it can be predicted from the location of word space. One reason for not writing enclitics together with the preceding root or other element is that this would make it harder for a reader to find the root and locate the stressed syllable.

3.2 Tone

There are three levels of tone, high, mid, and low. When low occurs on a syllable with either primary or secondary stress, it is realized as a mid-low glide. In the first part of compound words, two syllables are often reduced to a single syllable. If the original tone pattern was either mid low or low mid, the resulting syllable will have low. In very precise speech, both tones are sometimes retained and form a glide. In addition to the three levels, there are three floating tones, high, low-high, and low, each of which affects the following word in its own way. These floating tones effect a change in maybe 30 percent of the words in running text. In

addition, there is a rule that spreads the final low tone of a root to a following unstressed syllable under certain conditions, and another rule that moves a final high tone to the right under certain conditions.

Tone carries a fairly heavy functional load. Not only does it differentiate many lexical sets, but a floating high tone preceding a verb also marks the grammatical categories present tense and subjunctive mood. In addition, many words vary between two different tone patterns, or they vary as to whether or not they have a floating tone.

We mark tone on a very limited number of words. Our Mixtec co-workers do not want to fill up the page with tone marks, and they opted for marking tone only on the words in which they found themselves making false starts as they read because the context was not adequate to distinguish them from some other word. If we were to write tone on all words, the biggest question would have been whether to mark underlying or surface tone. A case can be made for either alternative.

One of the ways in which we mark tone is to write an acute accent for high tone in the following environments, where there is a significant contrast with mid or low.

1. Subjunctive mode is marked by a floating high tone that affects the next word. In the singular, the only difference between future indicative and subjunctive is the presence of this high tone on the root, which is arbitrarily written by an accent on the first syllable. This high tone often occurs on the first syllable, but it may occur on both or only on the second.

With high tone on the first syllable:

SUBJUNCTIVE		FUTURE	
ndíko ña	let her grind	ndiko ña	she will grind

With high tone on both syllables:

ndúku de i	let him look for the child	nduku de i	he will look for the child
------------	----------------------------	------------	----------------------------

With high tone on the second syllable:

táva i	let him take it out	tava i	he will take it out
--------	---------------------	--------	---------------------

In the plural, the high tone occurs on the plural marker jin, and an accent is written there (jín).

2. Present tense is also marked by a floating high tone; it affects the next word in the same ways as subjunctive. In many words, however, there is something else in the verb conjugation that signals present tense, and our Mixtec co-workers decided not to write an accent on these forms, only on the forms in which tone is the only marker of the tense.

Singular forms written with accent (In these verbs, present singular is homophonous with subjunctive):

PRESENT		FUTURE	
ndíko ña	she is grinding	ndiko ña	she will grind
ndúku de i	he looks for the child	nduku de i	he will look for the child
táva i	he takes it out	tava i	he will take it out

Singular forms written without accent:

jinu de	he is running	kunu de	he will run
---------	---------------	---------	-------------

jani de	he stands it up	kani de	he will stand it up
kixi i	the child is sleeping	kusu i	the child will sleep
jini de	he sees	kuni de	he will see

In plural forms high tone occurs on the ka plural particle preceding the root, but it is not written because a different plural particle jin is used in the future.

ka ndiko ña	they are grinding	jin ndijo ña	they will grind
ka nduku de i	they are looking for her	jin nduku de i	they will look for him
ka kixi de	they are sleeping	jin kusu de	they will sleep

In causative verbs, the root already has a high tone, and does not change in subjunctive or present, and so no accent is written.

skaka de	he is driving	skaka de	he will drive
----------	---------------	----------	---------------

In the dictionary, the present singular form of each verb is given as part of the entry in order to help speakers remember when to write an accent.

3. Selected functors: We write an acute accent on certain function words with high tone to distinguish them from function words with other tones:

má	imperative negative	vs.	ma	future negative
máko	imperative negative	vs.	ma ko	future negative + habitual
jín	subjunctive plural	vs.	jin	future plural
nú (~ ñúu)	if, when (present and future)	vs.	nuu	when (past), place where

4. Lexical pairs. We write an acute accent on the final syllable of certain lexical roots that show a contrast in tone with another word of the same lexical class.

ñu'ú	earth	vs.	ñu'u	fire (also ñu'u 'earth spirit')
ñíí	skin	vs.	ñii	salt
ka'án	think	vs.	ka'an	speak
neé	crazy, foolish	vs.	nee	dark

5. Other. We write an accent on the second syllable of the imperative of the verb 'to go'.

kua'án	imperative of to go	vs.	kua'an	incompletive tense of to go
--------	---------------------	-----	--------	-----------------------------

In some cases we want to differentiate a pair of morphemes in which no high tone occurs. Even though the Mixtec Academy recommends underline to mark low tone, we have decided not to use this. Instead of writing a tone mark, we write a circumflex on the less frequent member of the pair.

jâ	already	vs.	ja	thing, that which
ûni	really	vs.	uni	three

vâa	apparently	vs.	vaa	because
îyo	frightening, fearful	vs.	iyo	exists
kûu	to die	vs.	kuu	be, be able
kêe	to descend	vs.	kee	leave
ndêe	to descend (repetitive)	vs.	ndee	leave (repetitive)
kâku	be born	vs.	kaku	escape, be rescued

One of the pairs our co-workers feel the greatest need to differentiate is ja ‘already’, which has mid tone, vs. ja ‘thing, that, that which’, which has a mid-low glide plus floating low-high. On the grounds of phonology, we would be inclined to mark the word with the more complex tone pattern, but our choice was based on frequency. ‘Already’ is much less common than ‘thing, that, that which’.

In a few cases we mark one member of a pair in some other arbitrary way.

We write the number ‘one’ in (with a single vowel), and the number ‘nine’ iin, even though both have long vowels and the difference is in the tone: ‘one’ is mid mid, and ‘nine’ is mid low. This convention was established before we decided to use the circumflex, and it seems to work very well.

We write the word tna'a ‘also’, which has a mid low tone pattern, using the optional contracted form tna, to differentiate it from tna'a ‘each other, together’, which has a high high or low high tone pattern. This second word occurs in a similar environment to the first, but it is never contracted.

One other pair which we differentiate by writing a contracted variant (in addition to an accent) is nú ‘if, when’, which has high tone, versus nuu ‘when’, which has mid low plus floating low-high. Nú also has a longer variant núu.

I have appended to this orthography statement a document entitled “Guía ortográfica: Mixteco de Magdalena Peñasco”, which was written to explain these conventions to native speakers.

4. OTHER FEATURES

Answer the following questions adding pages as necessary.

4.1 How do you decide where to place sentence breaks?

This is a difficult area because Mixtecs do not think in terms of sentences, and they often link clauses with te ‘and’ and just keep going until they run out of things to say. We have looked for some phonological or grammatical signal to mark the end of a sentence, but we have not found any except the final glottal stop at the end of questions. The domain of sandhi does not serve as an indicator of a sentence break because any minor pause or syntax break within a sentence will break a sandhi chain. Sandhi never carries over to chi ‘pues’, nor to te ‘and’, nor even (usually) from the first part of a difrasismo (semantic doublet) to the second part.

We have often inserted sentence breaks where they felt right to us, often where there is a change of subject, time, or theme. When, however, our Mixtec co-workers read material with sentence breaks where we have put them, they read very well and don’t seem to notice the breaks. They are certainly not disturbed by them.

4.2 How do you decide where to place word breaks?

This area could well be considered the most difficult part of the entire language. A two-syllable root is the nucleus of the word, and roots combine with prefixes, proclitics, and enclitics, most of which are single syllables and lack primary stress. For various reasons, however, it does not seem best to write all of these elements together with the root as single words.

One problem occurs when the first element is nasal, and needs to be written with a final n. We do not want to put this n in a word-internal position because the n could be interpreted as a segmental consonant, rather than as a mark of nasalization. If we want to join a nasal element with another one, we usually use a hyphen.

The second problem area occurs when the second element is vowel initial, and writing it together with the first element would create a garden path situation; ku- + asi, can't be written kuasi, which would be misread as kwasi. We use a hyphen before a vowel-initial root that serves as the second part of a compound: ku-asi.

Roots also combine with other roots to create compounds. Some compounds are four-syllable, and consist of two two-syllable roots. Others are three-syllable, and consist of a contracted root followed by a full root. Often there is variation between a contracted and noncontracted form. Roots with the syllable patterns CVV and CVV are the ones most likely to undergo contraction: the second vowel and the glottal stop are lost. The tone pattern sometimes remains in the form of a glide, however, and nasalization is usually retained. In general, we write compounds with a contracted root together with the following root. If the contracted part is nasal, we write an n at the end, and join it to the second part with a hyphen. We also use a hyphen when the second part starts with a vowel.

Compounds that are not contracted are written as a single word if they are well established and common.

Here is a summary of what we are doing.

NOUNS

We write the following proclitics together with the root.

ti-	animal classifier	tisu'ma	scorpion
la-	animal classifier	lasaa	bird
ndi-	animal classifier	ndikolo	kind of lizard
nde-	liquid classifier	ndevisi	soda pop

We write the following proclitic as a separate word.

tnu	plant classifier	tnu yuja	pine tree
-----	------------------	----------	-----------

We write the following proclitic both ways in different combinations

ja- / ja	thing, that which	javixi	fruit
		ja taxi	witchcraft
		ja nani	name of

Some compounds written together

With nda'a 'hand':

ndatniñu	thing
ndakaa	key

With ve'e 'house':

veñu'u	church
vekaa	jail

Some compounds written separate:

lasaa kua'a	vermilion flycatcher
koo nda'a	centipede

ADJECTIVES

We write the following proclitics together with the root.

inu-	plural	chiki inukuii	green tunas
ti-	having characteristics of	tichuku	having lice
ndi-	various meanings	ndiñu'ma	grayish

We also write the following derivational prefixes together with the root.

ku-	inchoative	kunee	to get dark
ndu-	inchoative repetitive	nduva'a	to get better (from a sickness)
sa-	causative	sakuechi	to chop up

VERBS

We write the following prefixes together with the root.

ko-	future	konde'ya	will look
nda-	repetitive	ndataa	to give back
s-	causative	skee	to harvest corn
x-	causative	xta'vi	to raise (chickens)
xs-	causative	xsjaa	to inflate
ja-	inchoative	jakinee	to get in
nu-	inchoative repetitive	nukinee	to get back in

A few verbs have jin- instead of ja- for inchoative, and a few have sun- for causative, and we also write these together in spite of the final n:

jinkuiñi	to stand up
sunkani	to stand (something) up

All negatives, which range from one to three syllables and precede the verb proclitics, are written as separate words

Ñatuu kixí i. He is not sleeping.
 Ma ki'in de. He will not go.

We write the following proclitics as separate words.

jâ already
 ni past
 ni past CTF
 ka plural
 jin future plural
 yo past habitual
 ko future habitual
 jín subjunctive plural

Ni ka yo jika sa. We used to walk.
 Jin ko kaka sa. We will walk all the time.

One inflectional element that follows the verb is the motion verb plural *koo*. It is never contracted, and it always receives primary stress. It is written as a separate word, even when the verb that precedes it is contracted.

Kivi koo i. They will enter.
 Ni ka ka koo de. They went up. (cp. Ni kaa de. 'He went up.')

Motion auxiliary verbs show contraction, but we write them as separate words.

Ve ndi'i nuni. The corn is getting used up.
 Kuan ki'in ña ndute. She went to fetch water.

Some verb compounds show contraction of the first part and are written together:

satniñu to work (sa'a 'to do')

keja'a to begin (kee 'to leave')

Some verb compounds are written together that do not have contraction, but the meaning is idiomatic.

kataja'a dance
 kakan-ta'vi to ask a favor
 ka'an-so'o to exhort

Some verb compounds are written separate.

One element that follows the verb and is always written separate is *ini*, which forms emotion verbs. This element is never reduced and it always receives primary stress.

Ni ndakiti ini de. He became angry.

Another element that is always written separate is *tna'a*, 'each other' or 'together'. This is never contracted.

Ni ka jikan-tnu'u tna'a de. They asked each other.

There is a set of twelve enclitic pronouns. They commonly follow nouns to indicate possessor, verbs or adverbs to indicate subject, and prepositions to indicate the complement. Because they are enclitics, they are stressless and the root retains its penultimate stress. We write enclitic pronouns as separate words, which is the preference of our Mixtec co-workers, even though the Mixtec Academy writes them together with the previous word. By writing them separate, it is easier for a reader to recognize the preceding root, and to stress it correctly. Ten of the enclitics have the form CV.

nda'a ri	my hand
Ká'an de.	He speaks.
Ká'an xeen ña.	She speaks a lot
ji'in ti	with it (the animal)

Two of them, however, are simply V; even though they fuse with the stem, they are written as separate words.

ja'a i	his (the child's) foot
Ka jaa o.	We (inclusive) are eating.

See 4.6 under vowel clusters for further discussion of these two pronouns.

The document entitled “Guía ortográfica: Mixteco de Magdalena Peñasco”, which is appended to this orthography statement, describes these conventions in order to help native speakers write their language.

4.3 Dividers

Do you use any dividers other than word breaks, such as a hyphen?

We use a hyphen to separate parts of compounds when the first part is nasal, or when the second part is vowel-initial.

sukan-va'a	in order that
kun-kava	to fall
ku-asi	to become tasty
yu-itu	edge of a field

We also use a hyphen to separate off the possessed marker xi ~ x

tee xi-itu	owner of the field
------------	--------------------

We use a hyphen in ordinal numbers

yuva ku-uu	stepfather (lit.: 2nd father)
------------	-------------------------------

We also use a hyphen before the enclitic limiter particles -ni, -ka, and -nka.

kuee-ni	just slowly
va'a-ka	better (lit.: more good)
in-nka	just one

4.4 Loan Words

How do you treat loan words? What categories of loans do you recognize, and what principles do you apply to each category?

Spanish words are freely incorporated into Mixtec sentences in the speech of bilinguals at the present time, and Spanish words incorporated into Mixtec range from those that show no adaptation to Mixtec to those that show quite a lot.

We write words that show considerable adaptation as if they were Mixtec.

kalu	caldo
Su'si	God (from Jesús)
lunku	trompo

Sometimes there is some adaptation. People like to change a final o to u to “Mixtec-ize” a word. We go ahead and write the u.

burru	burro
frasku	jar

We are not consistent in how we write the consonants. Using a v instead of a b in donkey looked too jarring, and so we keep the b, but the k in frasku seems less jarring.

We spell casual, unadapted loans as in Spanish, and, in the dictionary and grammar, I enclose such code-switching words in quotes to show that they are not subject to the phonological rules of Mixtec.

ita “bugambilia”	bougainvillea
“programa civil”	civil program

When in doubt, we write the word as it is spelled in Spanish.

llanta	wheel (llanta); we could also write yanta
águila	hawk

We are still debating whether or not to keep the accent in more assimilated loans.

4.5 Syllable structure

What kinds of syllable structure does the language have? (Give examples.)

There are open syllables and syllables checked by glottal stop. The latter occur almost exclusively in the penultimate syllable of a root. The syllable onset can have from zero to three consonants.

Open syllables — final syllable of word

V	ja'.un	fifteen
	ko.o	snake
CV	tu.tu	paper
	ve.'e	house

CCV	ka.nda	to endure
	ya.kua	crooked
	ka'.ndi	to explode

CCCV	xtna	as
------	------	----

Open syllables — penultimate syllable of word

V	a.si	tasty
CV	tu.tu	paper
CCV	sku.su	to put to sleep
	ndo.'o	palm basket
	nda.xta.a	to sweep
CCCV	xsji.nu	to finish
	xtna.'a	to disturb
	ka.xtnu.'u	to let know
	skui.i	to peel
	skui.so	to promise

Open syllables — antepenultimate syllable of word

V	a.ndi.vi	heaven (V is infrequent, and andivi varies with jandivi)
CV	ku.va.'a	to be made
	nu.ko.o	to be seated
CCV	ndu.tu.tu	to be gathered together
	kua.tni.ñu	to use
CCCV	xtnakue'e	to wound

Closed syllables — penultimate syllable of word

V?	i'.ni	hot
CV?	su'.nu	shirt
	ja'.un	fifteen
CCV?	xta'.vi	to raise (animals)
	nda'.vi	poor
CCCV?	xnda'.vi	to deceive
	skua'.nu	to raise (children)

4.6 Clusters

What kinds of clusters occur (vocalic/consonantal)?

VOWEL CLUSTERS WITHIN A ROOT

There are no vowel clusters within a syllable. Consonant clusters with /j/ and /w/ are written with i and u giving a sequence of two orthographic vowels, but they do not constitute a vowel sequence.

Across syllable boundaries there are like vowels; see the treatment of lengthened vowels in section 3.above.

VOWEL CLUSTERS ACROSS MORPHEME BOUNDARIES

Vowels that come together at the boundaries of compound words are separated by a phonetic glottal stop, and this is written with a hyphen. See discussion above.

Two enclitic pronouns and one enclitic definite article consist of a vowel: i (third person child or thing), o (we inclusive), and a (proximal definite article). These enclitics, unlike roots, do not have a phonetic glottal stop at the beginning, and they show some fusion with the preceding element.

When these enclitics follow a CV.'V word, they take the place of the final vowel, but we write the underlying form.

CV.'V	chi'i o	we will sow	[tʃiʔo]
	ká'an i	the child speaks	[kãʔi]

(It would also be a rational choice to write the fused forms (chi'o, ka'in), but our LAs prefer the underlying forms.)

When these vowel-initial enclitics follow a word that has a different syllable structure, the final vowel of the root and the vowel of the enclitic form a cluster or a lengthened vowel, but we write a word space.

CV.V	koo o	we (I) will exist	[ko:]
	koo i	the child will exist	[ko:i]
CV'CV	sa'ya i	her child	[saʔɜai]
CV'.V	ja'un i	15 children	[xãʔũ]
CV.CV	kiti i ma	the child's animal	[kiti:mã]

A cluster is also formed when a vowel-initial enclitic follows another enclitic.

V.CCV	anua sa a	my heart	[an ^w ãsa:]
-------	-----------	----------	------------------------

The article can follow a pronoun, and it is possible to get a sequence of four phonetic vowels.

	anua o a	our heart	[an ^w ãõã]
--	----------	-----------	-----------------------

(Nasalization of the root carries over to vowel-initial enclitics; see the discussion in section 3. above.)

CONSONANT CLUSTERS WITHIN A ROOT

Some common combinations of two consonants that occur within a root are treated as consonantes compuestas.

CC	tn	tniñi	mouse
	nd	nduchi	beans
	kw	kuii	grass
	nw	anua	heart

A few other clusters occur in roots.

CC	nk	ndoko chinko	white zapote fruit
	ns	ndonso	to fall

The cluster nk also occurs at the beginning of some verbs, where the n is a reduced form of a proclitic element.

nk	nkuiñi	var. of jinkuiñi 'to stand'
----	--------	-----------------------------

These clusters also occur in loans.

nk	lunku	top (toy)
ns	gansu	goose

CONSONANT CLUSTERS ACROSS MORPHEME BOUNDARIES

There are two “consonantes compuestas” that occur in stems that are historically more than a single root, but which cannot synchronically be segmented.

CC	tj	tiukun	housefly
	sj	siu'u	to frighten (someone) (cp. yu'u 'to be frightened')

In addition, there are a large number of CC and CCC clusters that occur in causative verbs

CC	sk	skusu	to put to sleep
	sx	sjaa	to inflate (variant of xsjaa)
	ʃt	xtivi	to spoil, to waste
	ʃtʃ	xchi'yo	to cook a meal
	ʃs	xsa	to heat
	ʃn	xnani	to name
CCC	skw	skuiso	to promise
	ʃnd	xndi'i	to use up
	ʃtn	xtnakue'e	to hurt, to wound
	ʃsx	xsjaa	to inflate (variant of sjaa)

Some of these clusters occur in other compound words as well.

The following clusters occur in compound words other than causative verbs.

CCC	ʃkw	xkueyi	staircase, ladder
	nʃ	nchuxi	greetings (probably contains a loan adios)

Other clusters, like nt, are found in loanwords.

CC	nt	kapisantu	cemetery
		soo manta	length of cloth to tie baby on back

4.7

Do you use one orthographic symbol to represent more than one phoneme or phonemic feature?

The letter n represents a consonant phoneme when it occurs in a syllable onset, and it represents nasalization when it occurs at the end of a word. This is what the Mixtec Academy uses in their alphabet, and it causes problems only in writing compounds, where we need to use hyphens or word spaces, instead of joining the parts into a single word.

The letter i represents a vowel phoneme most of the time, but it represents /j/ when it occurs between a consonant and an unlike vowel. Likewise, the letter u represents a vowel phoneme most of the time, but it represents /w/ when it occurs between a consonant and an unlike vowel. This is most frequent in the complex consonant /kw/, written ku.

4.8

Are there some phonemes or phonemic features symbolized in one position in the word or phrase, but not elsewhere?

Nasalization is written with a final n in words that do not have a nasal consonant in the onset, but it is not written in words that have such a consonant.

Glottal stop is written in roots where it forms interrupted and checked vowels, and at the end of certain interjections, but it is not written at the end of questions or at the beginning of a root that is otherwise vowel-initial.

4.9

Are there some phonemes or phonemic features symbolized only sporadically? Examples: stress – written only where there is ambiguity, or tone – written only on certain words, or only certain tones written, etc.

No___ Yes___ (explain why you chose to do this)

Yes, tone. See explanation in section 3.2 above.

4.10 Text

Include a one-page glossed text using your orthography so your consultant can see orthographical choices in context. (Your orthography should follow Spanish punctuation conventions as much as possible.)

SAMPLE TEXT 1: Las niñas y la bandolera

Ni yo in ña'a,
había una mujer

te ni yo neva'a ña uu sa'ya ña,
y tenía ella dos hijas de.ella

te ni tetniñu ña i
y ordenó ella a.ellas

ja jin koo i jin nduku i ji'i.
que irán ellas buscarán ellas hongos

Había una mujer que tenía dos hijas, y les mando a que fueran a buscar hongos.

Te ni jan koo i,
y fueron ellas

te ni ka ni'i i ji'i ma.
y encontraron ellas hongos los

Kovaa sani te ni kenda in ña'a,
pero entonces llegó una mujer

te vaa ñuyuku ma kúu ña.
y pues bandolera la es ella

Fueron y encontraron los hongos. Pero en aquel momento salió una mujer, y en realidad era una bandolera.

Te jiña'a ña:
y le.dice ella

—¡Ne koo ya'a!
vengan acá

te taa ri uu xu'un
y daré yo dos monedas

—kúu ña jiña'a ña.
suena ella le.dice ella

Y ella dijo: --Vengan ustedes acá, y voy a darles algunas monedas --les dijo.

Te vaa lista ka ndaa maa i,
y pues listas están mismas ellas

te ni ka jito i ja'a ña ma,
y vieron ellas pies de.ella los

te ni ka jini i
y supieron ellas

ja ja'a ñuyuku kúu.
que pie bandolera es

Pero las niñas eran muy listas, y le miraron los pies, y supieron que era una bandolera.

Te ka jiña'a i ja:
y le.dicen ellas que

—Vaa ni ja xu'un a ni masu jan koo sa
pues ni por dinero el ni no llegaremos nosotros

chi ndijin chi ñuyuku kúu ni
pues usted pues bandolera es usted

—ka kuu i ka jiña'a i.
suenan ellas le.dicen ellas

Y le dijeron: --Ni siquiera por el dinero vamos, porque usted es una bandolera --dijeron ellas.

Sani te ka jinu i kuanu koo i.
entonces corren ellas regresaron ellas

Entonces regresaron corriendo a su casa.

HERE IS THE SAME TEXT WRITTEN IN PARAGRAPH FORMAT

Ni yo in ña'a, te ni yo neva'a ña uu sa'ya ña, te ni tetniñu ña i ja jin koo i jin nduku i ji'i.

Te ni jan koo i, te ni ka ni'i i ji'i ma. Kovaa sani te ni kenda in ña'a, te vaa ñuyuku ma kúu ña.

Te jiña'a ña:

—¡Ne koo ya'a! te taa ri uu xu'un —kúu ña jiña'a ña.

Te vaa lista ka ndaa maa i, te ni ka jito i ja'a ña ma, te ni ka jini i ja ja'a ñuyuku kúu.

Te ka jiña'a i ja:

—Vaa ni ja xu'un a ni masu jan koo sa chi ndijin chi ñuyuku kúu ni —ka kuu i ka jiña'a i.

Sani te ka jinu i kuanu koo i.

SAMPLE TEXT 2: The careless boys

(from orthography statement submitted in 1997)

In jichi, te uu suchi kuan koo yuku.
one time and two boy went plural woods

One time two boys went to the woods.

Te ni ka jakondee i ka kasiki i,
and past plural begin they pl play they

te ñatuu ni ka jini i.
and not past plural see they

And they began to play, and they didn't see [weren't aware of what was happening].

Te kuan koo ndixi'yu i ma itu xito i ma.
and went plural goats their the field uncle their the

And their goats went into the cornfield of their uncle.

Te ja ni kunuu ka yinee ti
and already past lasted plural are.inside animals

ka jaa ti viyu ma.
plural eat animals cornplants the

And already it got to be a long time that they were there eating the corn plants.

Te sana ni ka jini i, te chukan ma,
and then past plural see they, and that the

te ni kenda-ni xito i ma,
and past arrive-just uncle their the

te ni ndonda de nuu ndunduu i ma.
and past scold he face both them the

And then they saw it, and at that point, their uncle showed up, and he scolded the two of them.

HERE IS THE SAME TEXT WRITTEN IN PARAGRAPH FORMAT

In jichi, te uu suchi kuan koo yuku. Te ni ka jakondee i ka kasiki i, te ñatuu ni ka jini i. Te kuan koo ndixi'yu i ma itu xito i ma.

Te ja ni kunuu ka yinee ti ka jaa ti viyu ma. Te sana ni ka jini i, te chukan ma, te ni kenda-ni xito i ma, te ni ndonda de nuu ndunduu i ma.

BIBLIOGRAPHY

Further information about the sound system of Magdalena Peñasco Mixtec is found in:

Hollenbach, Barbara E. 2003a. "The historical source of an irregular Mixtec tone-sandhi pattern." Págs. 535–552 en Mary Ruth Wise, Thomas Headland y Ruth Brend, editores, *Language and life: essays in memory of Kenneth L. Pike*. Dallas: SIL International y University of Texas at Arlington. Online: <http://www.barbaraelenahollenbach.com/PDFs/mxgenSANDHI.pdf>.

Erickson de Hollenbach, Elena. 2004. "Los tonos del mixteco de Magdalena Peñasco" <http://www01.sil.org/mexico/mixteca/magdalena-penyasco/G030-TonosMixtecoMP-xtm.htm>.

- Erickson de Hollenbach, Elena. 2006a. “Paradigmas de los pronombres enclíticos del mixteco de Magdalena Peñasco” (<http://www.sil.org/mexico/mixteca/magdalena-penyasco/G037-ParadigmasPronombresMP-xtm.htm>).
- Erickson de Hollenbach, Elena. 2006b. “Paradigma del verbo ‘correr’ en el mixteco de Magdalena Peñasco” (<http://www.sil.org/mexico/mixteca/magdalena-penyasco/G036-correrMPenyasco-xtm.htm>).
- Erickson de Hollenbach, Elena. 2006c. “Paradigma del verbo ‘correr’ en el mixteco de San Cristóbal Amoltepec” (<http://www.sil.org/mexico/mixteca/magdalena-penyasco/G035-correrSCAmoltepec-xtm.htm>).
- Erickson de Hollenbach, Elena. s.f. “Gramática del Mixteco de Magdalena Peñasco (Sa'an N̄uu Savi)”. Ms.
- Erickson de Hollenbach, Elena. s.f. “Guía ortográfica: Mixteco de Magdalena Peñasco”. Ms., marzo de 2009.