Supplementary Material: Standardization of tone notations

Rates of change and phylogenetic signal in Mixtec tone

Sandra Auderset[1, 2]

[1]Max Planck Institute for Evolutionary Anthropology
[2]University of California Santa Barbara

August 15, 2022

This document gives a brief overview of how were standardized across varieties and provides additional information about the sources. I also provide an overview of the tone melodies attested for each variety in the data and point to further issues and how they were resolved. Varieties are sorted by subgroups and within that alphabetically.

Standardization of tone notations

The IPA offers two principal ways of displaying tone: diacritics that are placed above the tone bearing unit (TBU) or tone bars, with the latter method suggested as the preferred one. Tone diacritics are useful in practical orthography but not well suited to alignments, for which it is more useful to represent the tone after the TBU as its own character (even if that is not the phonetic reality). Tone bars are difficult to type and read and thus not well suited for the task at hand. I use Chao's tone numbers (Chao 1930) instead, since they are widely known and easy to type and read.

In this system, each distinctive pitch level is assigned a number from one to five, with one being the lowest and five the highest, cf. Table 1. The interval between the lowest and highest pitch is assumed to correspond roughly to an augmented fifth. Contour tones are represented as combinations of these levels. A high to low falling tone, for example, would be noted as 51, while a low to mid rising tone would be noted as 13. This is not meant as a statement regarding their status as tonemes, but rather as a practical decision facilitating comparison between languages. The reader should keep this in mind when consulting the profiles, since this means that the profile does not necessarily tell you how many tonemes a language has.

In the source materials, we find a multitude of tone notations (if tones are noted at all). Most descriptions of varieties with three tonemes use diacritics, while those with more tonemes are usually represented with numbers. Below I summarize the most common systems.

Diacritics: Notations with diacritics are most common in languages with three level tonemes and five tonemes of which two are contour tones. But they also occur with four level tonemes.

Table 1: Overview of tone notation based on Chao (1930)

Chao's number	Musical comparison	Label
5	G#	high
4	F	half-high
3	E	mid
2	D	half-low
1	C	low

High tone is invariable marked with an acute accent, rising tone with a caron, and falling tone with a circumflex. The marking of mid and low tones varies across sources and usually one of them is left unmarked. If the mid tone is marked, this is usually done with a macron. If the low tone is marked a grave accent, underbar, or more rarely macron is used. Sources that mark tones with diacritics but are not accompanied by a an explanation of the representation are often ambiguous, because of the overlapping uses of macron and no mark for mid and low tones, respectively.

Numbers: Notations with numbers are common in languages with more than three levels and/or multiple contour tones. Sources vary widly as to what system of numbering they use. The two most common ones are the Chao notation or a similar system in which the highest tone corresponds to the highest number and an inverse Chao notation, in which the lowest number correspond to the highest tone. Further differences pertain to what exact levels the numbers represent. Countours are represented by joining the respective numbers, e.g. 15 for a low-high contour in a Chao notation.

The tone notations found for level tones in the sources and their standardization used in this study are summarized in Table 2.

Group 1

San Juan Coatzospam Mixtec

The data for this variety comes from two sources: Dürr (1987) and Small (1990). Dürr's (1987) data and analysis are based on Pike & Small 1974, but the tone notation is standardized (as with all the varieties in Dürr 1987). The main difference between the two sources is that Small (1990:268) also posits toneless syllables, noting that "Unaccented syllables do not carry contrastive tone; they assimilate to the tone of the following syllable. Tone is not marked on such syllables in this sketch". In both sources, there are two level tones (low and high) and a contour tone, which is restricted to the second mora. In Dürr's data, the only contour we find in our data set is rising, while in Small's data it is falling. Both sources describe this variety as having extensive tone sandhi with surface tones written, which could explain the differing interpretation of the contour tone.

The tones given for cognate items often do not match across these sources. The apparent lack of consistency across tone groups was already mentioned by Dürr (1987:27, 33), leading the author to state that Coatzospam simply does not fit any of the generalizations. While the issue cannot

Table 2: Tone notations in source materials

Label	Description	5	4	3	2	1
AM	all three marked	acute (á)		macron (ā)		grave (à)
MULL	mid-low grave, low underbar	acute (á)		unm. (a)	grave (à)	underbar (a̱)
MMLU	low unm., mid macron	acute (á)		macron (ā)		unm. (a)
LU	low unm., no mid	acute (á)				unm. (a)
MULM	mid unm., low macron	acute (á)		unm. (a)		macron (ā)
LM	no mid, low macron	acute (á)				macron (ā)
MULG	mid unm., low grave	acute (á)		unm. (a)		grave (à)
LG	no mid, low grave	acute (á)				grave (à)
MULB	mid unm., low underbar	acute (á)		unm. (a)		underbar (a̱)
C	Chao	˥	˦	˧	˨	˩
C4	Chao with 4	˦		˧	˨	˩
C3	Chao with 3	˧		˨		˩
IC	inverse Chao	˩	˨	˧	˦	˥
IC4	inverse Chao with 4	˩	˨	˧		˦
IC3	inverse Chao with 3	˩		˨		˧

Table 3: Contrastive tone melodies in Coatzospam Mixtec (reflexes of basic tones highlighted)

Small			Dürr		
level tones		contour	level tones		contour
1.1	5.1	1.51	1.1		1.15
1.5	5.5		1.5	5.5	5.15

Table 4: Contrastive tone melodies in Agustín Chayuco Mixtec (reflexes of basic tones highlighted)

1.1	3.1	
1.3	3.3	5.3
	3.5	5.5

Table 5: Contrastive melodies in Colorado Mixtec (reflexes of basic tones highlighted)

1.1	3.1	5.1
1.3	3.3	5.3
	3.5	5.5

be resolved without gathering and analyzing more data, the two accounts can be reconciled and summarized with respect to the already reconstructed basic tones. The final glottal stop has no influence and thus this variety merges the *HH with *HH? sets and the LL with the LL? sets. *H has a low tone reflex, while *L seems to have a high tone (although the interpretation is difficult). There also seems to be (at least) a merger of *LH and *HL sets.

Group 2

San Agustín Chayuco Mixtec

The data for this variety comes from three sources Pensinger et al. (1974), Josserand (1983), and Dürr (1987). Only a few entries from Josserand (1983) and Pensinger et al. (1974) are marked for tone. In the few cases where we had multiple cognates marked for tones, the tones were always identical. In all three sources, tones are marked by diacritics (MMLG). This variety has three level tones (low, mid, high). In our data set, the low and high tone show a restricted distribution never co-occurring with each other. This results in 7 contrastive tone melodies.

The final glottal stop had no effect on tone in this variety, producing a merger of *HH with *HH? and *LL with *LL? sets. Reflexes are very regular: *H has a mid tone reflex, *L a low tone.

San Juan Colorado Mixtec

The data for this variety comes from Stark et al. (1986). Tones are marked throughout (MULG). This variety has three level tones (low, mid, high), but the combination of low-high is not attested, resulting in 8 contrastive tone melodies.

The final glottal stop does not affect tones in this variety, producing the familiar mergers. The reflexes of *HL and *LH sets are very interesting. *LH has mid-mid reflex, merging these sets with the *HH and *HH? ones. The numerals show different reflexes from other *LL sets, namely mid-mid. This could be analogical to ‘one’, where the mid-mid tones are expected. *HL appears to have split reflexes: once the expected mid-low, but also high-mid. It seems that vowel-initial forms have mid-low, while consonant-initial forms have high-mid.

Table 6: Contrastive melodies in Jicaltepec Mixtec (reflexes of basic tones highlighted)

1.1		
1.3	3.3	5.3
	3.5	5.5

Table 7: Contrastive melodies in Zacatepec Mixtec (reflexes of basic tones highlighted)

level tones		final contour
1.1	3.1	
1.3	3.3	3.31
1.5	3.5	

Santa María Jicaltepec Mixtec

The data for this variety come from Dürr (1987). This variety has three level tones (low, mid, high), but not all combinations are attested. Low and high tone do not co-occur and the mid-low combination is not attested either, resulting in 6 contrastive tone melodies.

The final glottal stop does not affect tones. *H has low tone reflex generally, *L a high tone, but there is missing data with respect to *LL sets apart from numerals. Again, *HL and *LH sets show somewhat unexpected reflexes. *HL has a mid-mid reflex, while *LH merges with the other low tone sets into high-high.

Santa María Zacatepec Mixtec

The data for this variety comes from Swanton & Mendoza Ruíz (forthcoming). This variety is very important for reconstruction because it preserves final glottal stop. The data from the recent sketch grammar by Towne (2011) had to be excluded, because tones are only marked on a few entries and the final glottal is missing from many entries. The data from Swanton & Mendoza Ruíz (forthcoming:fn.10) are preliminary, i.e. the documentation and analysis are ongoing. The tone system of this variety is not further elaborated on, but the correspondences are very systematic. Tones are marked with numbers in this source (C3). This variety has three level tones (low, mid, high) and one contour tone (falling). The high tone does not appear on the first mora and the contour only appears in the second mora combining with a mid tone. This results in 7 contrastive tone melodies. The glottal stop does not affect tones. All the *H-initial sets are merged into mid-mid reflexes and all the *L-initial sets into mid-mid.low reflexes.

Santiago Jamiltepec Mixtec

The data for this variety come from Johnson (1988). Tones are marked with diacritics (MMLU). This variety has three level tones (low, mid, high), but not all combinations are attested. Low and high tone do not co-occur and the low-mid combination is not attested either, resulting in 6 contrastive tone melodies.

The glottal stop has no influence on the tone reflexes. *H generally is mid and *L is low, but

Table 8: Contrastive melodies in Jamiltepec Mixtec (reflexes of basic tones highlighted)

1.1	3.1	
	3.3	5.3
	3.5	5.5

Table 9: Constrastive tone melodies in Peñasco Mixtec (reflexes of basic tones highlighted)

level tones			final contour
	3.1	5.1	
1.3	3.3	5.3	
1.5	3.5	5.5	5.15

there is also a further merger of *HL rendering low-low as well. *LH shows the inverse reflex of what might be expected, namely mid-low.

Group 3

Magdalena Peñasco Mixtec

The data for this variety comes from an extensive, recently published dictionary by Hollenbach (2017) using a latin-based system. Tones are not marked on the entry but given in square brackets after the entry, represented with letters; for better comparison, I have changed this so that the entries are now directly marked for tone (with MULG). There are three level tones (low, mid, high) and one contour (rising). The level tones all combine with each other, except there is no low-low combination. The contour tone only appears on the second mora and in our data set only combines with the high tone. This results in 9 contrastive tone melodies. The glottal stop has no influence on tone reflexes. Generally, *H has mid tone reflex, and *L has a low tone reflex.

Santo Domingo Huendío Mixtec

The data for this variety were shared with me by Braulio Becerra Roldán and are an updated and corrected version of Becerra Roldán (2015). Tones are marked throughout (MMLG). There are three level tones (low, mid, high) and one contour (rising). The level tones all combine with each other, except there is no low-low combination. The contour tone only appears on the first mora and in our data set only combines with the high tone. This results in 9 contrastive tone melodies. The glottal stop has no influence on tone reflexes. Generally, *H has mid tone reflex, and *L has a low tone reflex.

Table 10: Constrastive tone melodies in Huendío Mixtec (reflexes of basic tones highlighted)

level tones			initial contour
	3.1	5.1	
1.3	3.3	5.3	
1.5	3.5	5.5	15.5

Table 11: Constrastive tone melodies in Peñoles Mixtec (reflexes of basic tones highlighted)

level tones		initial contour
1.1	5.1	51.1
1.5	5.5	51.5

Group 4

Santa María Peñoles Mixtec

The data for this variety come from Josserand (1983) and Dürr (1987). Tones are marked with diacritics (LG). There are two level tones (low and high) and one contour tone (falling). The level tones combine with each other freely, but the the contour is restricted to the first mora, resulting in 6 contrastive tone melodies. The glottal stop has no effect on the tone reflexes. The reflexes could be summarized as tone inversion, with *H rendering low tone and *L rendering high tone.

Linkage 5

San Juan Diuxi Mixtec

There is data for this variety from two sources: Dürr (1987) and Kuiper & Oram (1991). They differ quite drastically in their tone marking, because there are two competing analyses of the tone system of this language. Pike & Oram (1976) analyze Diuxi Mixtec has having two tonemes (high and low) and constrastive stress on the final syllable, while in Daly's (1978) analysis there are four tonemes (high, low, high modified, low modified) based on two features (high and modified) but no contrastive stress. This variety exhibits various processes of tone sandhi, which further complicates matters. The forms in Daly (1978) are all underlying and produce the surface realizations via a set of rules. Neither the materials in Dürr (1987) nor those in Kuiper & Oram (1991) have been standardized, resulting in a mixture of notations for both sources. Furthermore, the sections written by Kuiper in Kuiper & Oram (1991) use a modified version of Daly's system writing surface tones. We are thus left with three different tone notations/system for this variety, making the analysis in terms of historical processes difficult to impossible. This variety thus had to be excluded from the current study.

Table 12: Contrastive tone melodies in Ocotepec Mixtec (reflexes of basic tones highlighted)

1.1	3.1	5.1
1.3	3.3	5.3
1.5	3.5	5.5

Table 13: Contrastive tone melodies in Yucuhiti Mixtec (reflexes of basic tones highlighted)

level tones		
1.1	3.1	
1.3	3.3	
1.5	3.5	5.5

Group 6

Santo Tomas Ocotepec Mixtec

The data for this variety comes from three sources: Josserand (1983), Dürr (1987) and Alexander (1988). Josserand (1983) uses the same tone notation system as Dürr (1987). Alexander (1988:170) also notes three tones, but with a different notation (MMLU). This variety has three level tones (high, mid, and low) which freely combine with each other, resulting in 9 contrastive tone melodies.

Where there are multiple entries from these sources, the tones often coincide, but not always. The most systematic difference is found in *LL and *LL? sets, where Josserand often exhibits mid-low where Alexander has low-low. As the latter is more frequent over all in these sets, we regard this as the default reflex. The final glottal stop has no effect on tone realization. Generally, *H has a mid tone reflex and *L a low tone.

Santa María Yucuhiti Mixtec

The data for this variety is from Josserand (1983). Tones are marked with diacritics throughout (MULG). This variety has three level tones (high, mid, and low) but not all combinations are attested. The high tone in the first mora shows the most restricted distribution, only co-occuring with another high tone. This results in 7 contrastive tone melodies. The final glottal stop has no effect on tone realization. Generally, *H has a mid tone reflex and *L a low tone.

Santiago Nuyoo Mixtec

The data for this variety is from Josserand (1983). Tones are marked with diacritics throughout (MULG). There is relatively little data. There are three level tones (low, mid, high) which combine freely. There is one entry with a falling contour on the second mora, resulting in 10 contrastive tone melodies. Generally, *H has a mid tone reflex and *L a low tone. There is a merger of the *HL set with the other high tone sets.

Table 14: Constrastive tone melodies in Nuyoo Mixtec (reflexes of basic tones highlighted)

level tones			contour
1.1	3.1	5.1	
1.3	3.3	5.3	
1.5	3.5	5.5	5.51

Table 15: Constrastive tone melodies in Chalcatongo Mixtec (reflexes of basic tones highlighted)

Macaulay			Swanton & Mendoza Ruíz		
1.1	3.1	5.1	1.1	3.1	
1.3	3.3	5.3	1.3	3.3	5.3
1.5	3.5	5.5			5.5

Chalcatongo de Hidalgo Mixtec

The data for this variety comes from two sources: Macaulay (1996) and Swanton & Mendoza Ruíz (forthcoming). Macaulay (1996:19) uses an Americanist notation with tones marked by diacritics (MULG). Swanton & Mendoza Ruíz (forthcoming) write in a form of IPA and use numbers for tone notation (C3). In both sources, there are three level tone (low, mid, high). The distribution is, however, not the same. In Macaulay, all combinations are possible and we thus find 9 contrastive tone melodies. In the newer source, only 6 melodies are attested, but there is also less data.

Where we have entries from both sources, the tone markings sometimes coincide but other times do not. While it is difficult to draw generalizations on such a limited set, we observe that usually when there are divergences, the reflexes in Macaulay (1996) are one level higher than those in Swanton & Mendoza Ruíz (forthcoming). That is, often the former will have a high or mid tone, where the latter has a mid or low tone, respectively. In general, the data from Macaulay (1996) do not result in consistent tone sets, even though there are many more entries. Most likely, this is due to floating high tones which appear in connected speech (Macaulay 1996:34-35) being marked on some items in isolation. We have resolved incoherent sets on the basis of the systematic correspondences from Swanton & Mendoza Ruíz (forthcoming). The final glottal stop has no effect on tone realizations. There is a further merger of all the sets with a final low tone, cf. Table **??**.

San Esteban Atatlahuca Mixtec

The data for this variety comes from three sources Alexander (1980) (MULM), Josserand (1983), and Dürr (1987). Tones are marked in all sources, but not on every entry in Josserand (1983). We base our analysis predominantly on Alexander (1980), because the correspondences are most regular in that source. There are three level tones (low, media, high) and all but one combination (low-low) are attested, resulting in 8 contrastive tone melodies.

Table 16: Constrastive tone melodies in Atatlahuca Mixtec (reflexes of basic tones highlighted)

level tones		
	3.1	5.1
1.3	3.3	5.3
1.5	3.5	5.5

Table 17: Contrastive tone melodies in El Grande Mixtec (reflexes of basic tones highlighted)

level tones		
1.1	3.1	5.1
1.3	3.3	5.3
1.5	3.5	5.5

San Miguel el Grande Mixtec

The data for this variety comes from two sources: Josserand (1983) and Dürr (1987). This variety has three level tones (high, mid, and low) which freely combine with each other, resulting in 9 contrastive tone melodies. The final glottal stop has no effect on tone realization. Generally, *H has a mid tone reflex and *L a low tone. There is a merger of all sets that have a final low tone into mid-low. This could be explained by low tone dissimilation for the low-low sets.

San Pedro Molinos Mixtec

The data for this variety comes from two sources: Josserand (1983) and Dürr (1987). This variety has three level tones (high, mid, and low) which freely combine with each other, resulting in 9 contrastive tone melodies. The final glottal stop has no effect on tone realization. Generally, *H has a mid tone reflex and *L a low tone. Due to low coverage, this variety had to be excluded from further analysis.

Santiago Yosondua Mixtec

The data for this variety comes from Farris (1992). Tones are marked with diacritics (MMLU). There is extensive tone sandhi in this variety, but the tones are written as underlying, i.e. before the application of any sandhi rules. There are three level tones (low, mid, high) and all but one

Table 18: Contrastive tone melodies in Molinos Mixtec (reflexes of basic tones highlighted)

level tones		
1.1	3.1	5.1
1.3	3.3	5.3
1.5	3.5	5.5

Table 19: Contrastive tone melodies in Yosondua Mixtec (reflexes of basic tones highlighted)

1.1	3.1	5.1
1.3	3.3	5.3
1.5		5.5

Table 20: Contrastive tone melodies in Abasolo Mixtec

level tones		
1.1	3.1	5.1
1.3	3.3	5.3
1.5	3.5	5.5

combination (mid-high) are attested, resulting in 8 contrastive tone melodies. The correspondence sets are not as regular as with other varieties, especially in the sets involving *L tones. The final glottal stop has no effect on tone realization. Generally, *H has a mid tone reflex and *L a low tone. There is a merger of the *HL set with the other high tone or low tone sets.

Group 7

Group 71

Abasolo del Valle Mixtec

This variety is spoken in a diaspora comunity in the state of Veracruz. The data come from Galindo Sánchez (2009). The author claims that the variety is identical to that spoken in San Juan Mixtepec. There are three tones: high = acute, mid = unmarked, low = grave (MULG). The tones, however, do not correspond well with those of Mixtepec reflexes and there seem to have been extensive merging processes, collapsing almost all basic tone sets. This variety will thus be excluded from further analysis, because it would be difficult to impossible to tell which sets certain entries belong to.

Santa Maria Yucunicoco Mixtec

The data for this variety is from Josserand (1983). Tones are marked with diacritics throughout (MMGL). This variety has three level tones (low, mid, high), which combine to all 9 possible tone melodies. There is one form given with a mid-low contour on the last mora ('feather'), but given this is an isolated case and tone notation in the source is not standardized or strictly phonological, we do not add this contour to the inventory. Generally, *H has a mid tone reflex and *L a low tone. The lost glottal stop has a raising effect, i.e. it renders a high tone.

Table 21: Contrastive tone melodies in Yucunicoco Mixtec (reflexes of basic tones highlighted)

1.1	3.1	5.1
1.3	3.3	5.3
1.5	3.5	5.5

Table 22: Contrastive tone melodies in Mixtepec Mixtec (reflexes of basic tones highlighted)

level tones			final contour	initial contour
1.1	3.1	5.1	‡1.15	15.3
1.3	3.3	5.3	(‡3.15)	51.3
1.5	3.5	5.5		51.5

San Juan Mixtepec Mixtec

The data for this variety comes from two sources: Josserand (1983) and Dürr (1987). There are three levels tones (low, mid, high) and two contours (rising and falling) combining into 14 distinct tone melodies in our data. The level tones can appear on either mora and combine with each other freely. The rising contour tone can appear on either mora as well. The rare falling tone seems to be restricted to the first mora. We have relatively few entries compared to other languages in this subgroup, so some of the absent tone melodies might reflect a data gap. We assume this is the case for the absent mid-rising melody.

Generally, *H has a mid tone reflex and *L a low tone. The lost glottal stop has a raising effect, i.e. it renders a high or high-low tone. There is a further split in the sets with final glottal stop, such that monosyllables (CVV and CV?V) exhibit a final high tone, but bisyllables (CVCV) exhibit a final falling tone (although there is not enough data to confirm this for HH?).

La Batea Mixtec

The data for this variety was shared with us by collaborator and native speaker Yésica Ramírez (MICOP). The data were collected between 2016-2020 in Oxnard and Santa Barbara, California as part of the collaboration between UCSB and MICOP. Tones are marked with diacritics (MULG). There are three levels tones (low, mid, high) and two contours (rising and falling) combining into 16 distinct tone melodies in our data. The level tones can appear on either mora and combine with each other freely. The rising contour tone can appear on either mora as well, although it is rarely found on the first mora. The much rarer falling tone seems to have a very restricted distribution, but this might be due to a gap in our data.

Generally, *H has a mid tone reflex and *L a low tone. The lost glottal stop has a raising effect, i.e. it renders a high or high-low tone. There is a further split in the sets with final glottal stop, such that monosyllables (CVV and CV?V) exhibit a final high tone, but bisyllables (CVCV) exhibit a final falling tone.

Table 23: Contrastive tone melodies in La Batea Mixtec (reflexes of basic tones highlighted)

level tones			final contour		initial contour
1.1	3.1	5.1	ǂ1.15		15.1
1.3	3.3	5.3	ǂ3.15	3.51	15.3
1.5	3.5	5.5	5.15		51.5

Table 24: Contrastive tone melodies in Yucunani Mixtec (reflexes of basic tones highlighted)

level tones			final contour		initial contour
1.1	3.1	5.1	ǂ1.15		15.1
1.3	3.3	5.3	ǂ3.15	3.51	15.3
1.5	3.5	5.5	5.15		

Yucunani Mixtec

The data for this variety was shared by collaborator and native speaker Jeremías Salazar. The orthography used is latin-based **ʔ** and tones are marked throughout with diacritics (MULG). There are three levels tones (low, mid, high) and two contours (rising and falling) combining into 15 distinct tone melodies in our data. The level tones can appear on either mora and combine with each other freely. The rising contour tone can appear on either mora as well, although it is rarely found on the first mora. The much rarer falling tone seems to have a very restricted distribution, but this might be due to a gap in our data. The same is probably true for the missing rise-high melody.

Generally, *H has a mid tone reflex and *L a low tone. The lost glottal stop has a raising effect, i.e. it renders a high or high-low tone. There is a further split in the sets with final glottal stop, such that monosyllables (CVV and CVʔV) exhibit a final high tone, but bisyllables (CVCV) exhibit a final falling tone.

Linkage 72

Ixpantepec Nieves Mixtec

The data for this variety is predominantly from (Carroll 2015), but some entries are from Josserand (1983). In both, tones are marked with diacritics (MMLG). This variety has three level tones (high, mid, and low) which freely combine with each other, resulting in 9 contrastive tone melodies.

Table 25: Contrastive tone melodies in Ixpantepec Mixtec

1.1	3.1	5.1
1.3	3.3	5.3
1.5	3.5	5.5

Table 26: Contrastive tone melodies in Duraznos Mixtec

1.1	3.1	5.1
1.3	3.3	5.3
1.5	3.5	5.5

Table 27: Contrastive tone melodies in Tecomaxtlahuaca Mixtec

1.1	3.1	5.1
	3.3	5.3
1.5	3.5	5.5

San Martín Duraznos Mixtec

The data for this variety comes from two sources: Josserand (1983) and the ongoing documentation project of Sandra Auderset and native speaker Carmen Hernández Martínez. Tones are only marked on a few entries in Josserand (1983) (with MMLG), but throughout in the newer source (with MULG). The tones almost alwasy match between the two sources. This variety has three level tones (high, mid, and low) which freely combine with each other, resulting in 9 contrastive tone melodies.

The lost final glottal stop has a lowering effect, resulting in a merger of the low tone sets. *H generally has a mid tone reflex and *L a low tone. There is a further merger of HL with the HH set.

San Sebastian Tecomaxtlahuaca [teco]

The data for this variety is from Josserand (1983). Tones are marked with diacritics throughout (MMLG). There is a limited amount of data, making generalizations difficult at this stage. This variety has three level tones (high, mid, and low) and one combination is not attested in the data set (low-mid), resulting in 8 contrastive tone melodies.

It seems that the lost final glottal stop has a lowering effect. *H generally seems to have a high tone reflex and *L a low tone. But the correspondence sets are not very regular and thus difficult to interpret.

Group 73

Alacatlatzala Mixtec [alac11]

The data for this variety comes from five sources: Josserand (1983), Dürr (1987), Zylstra (1991), Anderson (2006) and Zylstra (2012). Tones are marked in all of these sources, but not for all entries in Josserand (1983). All of them mark tones with diacritics (Anderson (2006:viii) and Zylstra (2012:13) with MULB, Zylstra (1991:8) with MMLU). In the majority of cases, all the sources have the same tone marking across cognates. Where they diverge, the entries from Zylstra (1991) usually show different tones that do not always fit the default reflexes of the set. Zylstra (2012)

Table 28: Contrastive tone melodies in Alacatlatzala Mixtec

level tones		
1.1	3.1	5.1
1.3	3.3	5.3
1.5	3.5	5.5

Table 29: Contrastive tone melodies in Alcozauca Mixtec

level tones			contour	
1.1				
	2.2	3.2	5.2	
1.3		3.3	5.3	
1.5		3.5	5.5	51.5

is not only the most recent work, but also has the most consistent reflexes. We thus in general place higher confidence in the tone marking of this source than the others.

This variety has three level tones (high, mid, and low) which freely combine with each other, resulting in 9 contrastive tone melodies.

The lost final glottal stop has a lowering effect, resulting in a merger of the low tone sets. *H generally has a mid tone reflex and *L a low tone. There is a further merger of HL with the HH set.

Alcozauca de Guerrero Mixtec [alco11]

The data for this variety come from two sources: Josserand (1983) and Swanton & Mendoza Ruíz (forthcoming). Tone is marked for a few entries only in Josserand (1983) with diacritics. In Swanton & Mendoza Ruíz (forthcoming), tone is marked throughout with numbers (C4). This variety has four level tones (low, mid-low, mid, high) and two contours (rising and falling) and a complex tone (high-low-high) (Swanton & Mendoza Ruíz forthcoming), but the latter does not appear in our data set. Tone 1 and 2 have an almost complementary distribution, since tones 2, 3 and 5 do not combine with tone 1, which is thus restricted to the first mora. All other tones can appear on either mora. The rising contour is absent from our data set, so we cannot say anything about its distribution. The falling contour only appears in the first mora before a high tone.

The lost final glottal stop has a raising effect, resulting in high tone reflex. *H generally has a 3 tone reflex, while *L has a 2 reflex after H, but 1 otherwise. There are no mergers in the basic tones of this variety.

Cahuatache Mixtec [cahu11]

The data for this variety originally comes from Schultze-Jena (1938), a collection of early travel accounts. It is reproduced in Dürr (1987). Although a valuable resource, the interpretation of the material is not easy, especially regarding the tone notation. Therefore, I only use the data as reproduced in the later collections, not from the original. Even so, the sets are less regular

Table 30: Contrastive tone melodies in Cahuatache Mixtec

level tones		
1.1	3.1	5.1
1.3	3.3	5.3
1.5	3.5	5.5

Table 31: Constrastive tone melodies in Coicoyán Mixtec

level tones			contour
1.1	3.1	5.1	
1.3	3.3	5.3	
1.5	3.5	5.5	31.5?

than in other varieties and there seem to be more mergers than in other varieties, suggesting that some tone differences might have gone unnoticed. This variety will thus be excluded from further analysis. This variety has three level tones (high, mid, and low) which freely combine with each other, resulting in 9 contrastive tone melodies.

Coicoyán de las Flores Mixtec

The data for this variety is from Josserand (1983). Tones are marked with diacritics throughout (MMLG). There is a limited amount of data, making generalizations difficult at this stage. There are three level tones (low, mid, high) which combine freely. There is one entry with a falling contour on the first mora. It seems that the lost final glottal stop has a raising effect.

El Jicaral Mixtec

The data for this variety was provided by JN Martin. The orthography used is Latin-based and tones are marked throughout with diacritics (MULB). There are three level tones (low, mid, high) and two contours (rising and falling). While level tones can appear on either mora and combine with each other freely, the contours are only found on the first mora and combining with one level tone each in our data set. This results in 11 contrastive tone melodies.

The lost final glottal stop has a raising effect, resulting in a high tone. Generally *H has mid reflex and *L a low one. There is a merger of the HL with the HH set.

Table 32: Contrastive tone melodies in Jicaral Mixtec

level tones			contour
1.1	3.1	5.1	
1.3	3.3	5.3	15.3
1.5	3.5	5.5	51.5

Table 33: Contrastive tone melodies in Metlatonoc Mixtec

1.1	3.1	5.1
1.3	3.3	5.3
1.5	3.5	5.5

Table 34: Contrastive tone melodies in Progreso Mixtec

level tones		
1.1	(3.1)	5.1
1.3	3.3	5.3
1.5	3.5	5.5

Metlatonoc Mixtec

The data for this variety comes from Josserand (1983) and Dürr (1987). This variety has three level tones (high, mid, and low) which freely combine with each other, resulting in 9 contrastive tone melodies. We only have few entries for this variety and the correspondence sets show a lot of variation. This variety thus had to be excluded from further analysis.

San Jerónimo Progreso Mixtec

The data for this variety come from two sources: Dürr (1987) and Shields (1988). Note that the sources refer to this variety as Silacayoapan Mixtec, using the name of the municipality. They all specify, though, that the data was gathered in the town of San Jerónimo Progreso. Shields (1988:312) marks tone with diacritics as well (MMLU). This variety exhibits tone sandhi and surface tones are marked. In many cases, if we have entries from both sources, the tones coincide. Where they do not, most often Shields (1988) has mid-mid where Dürr (1987) has low-low. This variety has three level tones (high, mid, and low) which freely combine with each other, resulting in 9 contrastive tone melodies.

It seems that the lost glottal stop has a lowering effect. Generally *L has a low tone reflex, which *H it is more difficult to tell. *HL shows mixed reflexes but has merged either with *HH or *LL.

San Marcos de la Flor Mixtec

The data for this variety was shared with us by collaborator and native speaker Moisés (MICOP). The data were collected between 2019-2020 in Oxnard, California. This variety is closely related to Piedra Azul Mixtec.

This variety has three level tones (low, mid, high) and two contour tones (falling and rising), combining into a total of 18 contrastive tone melodies. The level tones have no restrictions in their distribution, the but the contour tones do. Only the rising contour can appear on the first mora, while both falling and rising contours can combine with level tones when they appear on the second mora. There is also a rare melody composed of two contours, a rising and falling one.

Table 35: Tone melody contrasts in San Marcos la Flor Mixtec

level tones			final contour		initial contour
1.1	3.1	5.1	1.15	1.51	15.1
1.3	3.3	5.3	3.15	3.51	15.3
1.5	3.5	5.5	5.15	5.51	15.51

Table 36: Tone melody contrasts in Piedra Azul Mixtec

level tones			final contour		initial contour
1.1	3.1	5.1	1.15	1.51	15.1
1.3	3.3	5.3	3.15	3.51	15.3
1.5	3.5	5.5	5.15	5.51	15.51

The lost glottal stop has a high-low reflex. Generally *H has a mid tone reflex and *L has a low tone reflex. There are mixed reflexes for the *HL group.

Piedra Azul Mixtec

The data for this variety was provided by linguist and native speaker Gabriel Mendoza (p.c.). The orthography used is latin-based and tones are marked throughout with diacritics (MULG notation). This variety has three level tones (low, mid, high) and two contour tones (falling and rising), combining into a total of 18 contrastive tone melodies, one of which is possible but not attested in our dataset (cf. Peters 2018). The level tones have no restrictions in their distribution, the but the contour tones do. Only the rising contour can appear on the first mora, while both falling and rising contours can combine with level tones when they appear on the second mora. There is also a rare melody composed of two contours, a rising and falling one (Peters 2018:26-36) While all 18 tone melodies are attested on bisyllabic words, monosyllables have a more restricted inventory and never show an initial contour tone.

The lost glottal stop has a high-low reflex. Generally *H has a mid tone reflex and *L has a low tone reflex. There are mixed reflexes for the *HL group.

Tepango Mixtec

The data for this variety come from two sources, Dürr (1987) and Hills (1990). Tones are marked throughout in both sources (with MMLU in the latter). Both sources refer to the variety as 'Ayutla Mixtec', referencing the municipality that Tepango is part of. They all specify that the data was gathered with speakers in the village of Tepango. This is one of only a few Mixtec varieties that retains final glottal stop and is thus particularly important for reconstruction.

Both sources describe the variety as having three level tones (low, mid, high), but the attested tone melodies are quite different, resulting in mismatches between tone markings within cognates. In Hills (1990), the high tone is rare and has a very restrictive distribution (it seems to appear mostly on inflected verbal forms, which we excluded from this study). It appears only once in our data set, combining with a mid tone. In Dürr (1987), the high tone freely combines

Table 37: Contrastive tone melodies in Tepango Mixtec by source

Hills		Dürr		
1.1	3.1	1.1	3.1	5.1
1.3	3.3		3.3	
	3.5	1.5	3.5	5.5

Table 38: Contrastive tone melodies in Tlahuapa Mixtec

level tones			final contour	initial contour
1.1	3.1	5.1		
1.2				
1.3	3.3	5.3	3.15	15.3
1.5	3.5	5.5	5.15	51.3

with all others, but the mid tone in the second mora only appears with a mid tone in the first mora. Generally, where Dürr has a high tone, Hills has a mid tone. Due to the presence of the final gottal stop, this variety has a tone system quite unlike the others in this group.

Tlahuapa Mixtec

The data for this variety come from collaborator and native speaker Griselda Reyes Basurto (MICOP, Oxnard CA). The collaboration with Griselda Reyes Basurto is part of the collaboration between UCSB Linguistics and MICOP. Reyes Basurto (2020) uses a latin-based systems and tones are marked throughout as high = acute, mid = unmarked, mid-low = macron, low = grave, high-low = circumflex, low-high = haček (MULG). The mid-low tone is rare outside of verbs. The lost final glottal stop has a raising effect, resulting in a high tone. Generally *H has mid reflex and *L a low one. There is a merger of the HL with the HH set.

Xochapa Mixtec

The data for this variety come from Stark et al. (2013), who use a latin-based orthography. Tones are marked with diacritics (MULL). This variety has four level tones (low, mid-low, mid, high) and two contours (rising and falling). The latter are rare though and only appear on the first mora in our data set. Tone 2 (mid-low) also has a restricted distribution and does not co-occur with either tone 3 or tone 1. Due to these restriction, there are 12 contrastive melodies. The *HL set gives three different reflexes (1.1, 2.2, 3.3) but otherwise the sets are very regular. The lost final glottal stop has a raising effect.

Yoloxóchitl Mixtec

The data for this variety come from Amith & Castillo García (n.d.). Tones are marked with numbers (C4). There are four level tones (low, low-mid, mid, high) and five contour tones, but the distribution especially of the latter is very restricted, leading to a total of 16 contrastive tone

Table 39: Contrastive tone melodies in Xochapa Mixtec

level tones				contour
1.1			5.1	
1.3	2.2	3.3	5.3	15.3
1.5	2.5	3.5	5.5	51.5

Table 40: Contrastive tone melodies in Yoloxóchitl Mixtec

level tones			final contour	initial contour
1.1		5.1	1.32	
	3.2	5.2	3.52	13.2
1.3	3.3	5.3		15.3
1.5	3.5	5.5	5.25	15.5

melodies in our data. Tone 2 can only appear on the second mora and does not combine with tone 1. We also do not find entries with a mid-low tone melody. Falling contour tones are rare and can only appear on the second mora. They are noted as 32 and 52, respectively, but given that they both only combine with one level tone, they need not be distinguished to maintain all the contrasts. Raising contours are mostly found on the initial mora, but there is also one melody that has a rise in the second mora. As with falling contours, these could all be labeled rising and the contrasts would still be maintained.

The lost final glottal stop has a raising effect, resulting in high tone reflex. *H generally has a 3 tone reflex, while *L has a 2 reflex after H, but 1 otherwise. There are no mergers in the basic tones of this variety.

Group 74

San Jerónimo Xayacatlan Mixtec

The data for this variety come from Dürr (1987). There are three level tones (low, mid, high), but only five contrastive melodies attested in our data. Given the rarity/absence of high tones on the first mora in the closely related variety of Xaycatlan de Bravo, this gap might not simply reflect missing data.

The lost final glottal stop has a lowering effect, resulting in a merger of the low tone sets. *H generally has a mid tone reflex and *L a low tone. There is a further merger of HL with the HH set.

Table 41: Contrastive tone melodies in SJ Xayacatlan Mixtec

1.1	3.1
1.3	3.3
1.5	

Table 42: Contrastive tone melodies in Xayacatlan de Bravo Mixtec

1.1	3.1	5.1
1.3	3.3	
1.5	3.5	

Table 43: Contrastive tone melodies in Yutatío Mixtec

	3.1	5.1
1.3	3.3	5.3
1.5	3.5	5.5

Xayactlan de Bravo Mixtec

The data for this variety come from Josserand (1983). There are three level tones (low, mid, high), but only one entry with a high tone on the first mora, resulting in 7 contrastive tone melodies.

The lost final glottal stop has a lowering effect, resulting in a merger of the low tone sets. *H generally has a mid tone reflex and *L a low tone. There is a further merger of HL with the HH set.

Linkage 75

San Andrés Yutatío Mixtec

The data for this variety come from an extensive recent dictionary Williams et al. (2017). The orthography used in this source is latin-based and tones are marked throughout (MULB). There are three level tones, but one combination (low-low) is not attested, resulting in 8 contrastive tone melodies.

The lost final glottal stop has a lowering effect. *H generally has a mid tone reflex and *L a low tone. However, there also seems to be low tone dissimilation in the LL sets and a further merger of the HL with the HH set. The high tone in HH? could be the result of avoiding a merger with the LL sets.

Santiago Cacaloxtepec Mixtec

The data for this variety come from Dürr (1987). There are three level tones (low, mid, high), but the high and mid tone never combine, resulting in 7 contrastive tone melodies.

The lost final glottal stop has a lowering effect, resulting in a merger of the low tone sets. *H generally has a mid tone reflex and *L a low tone. There is a further merger of HL with the HH?

Table 44: Contrastive tone melodies in Cacaloxtepec Mixtec

1.1	3.1	5.1
1.3	3.3	
1.5		5.5

Table 45: Contrastive tone melodies in Yucunuti Mixtec

1.1	3.1	5.1
1.3	3.3	

Table 46: Contrastive tone melodies in Yucuquimi Mixtec

1.1		5.1
1.3	3.3	5.3

set.

Yucunuti de Benito Juarez

The data for this variety come from Josserand (1983), but few entries are marked for tones. This variety will thus be excluded from analysis for now.

Yucuquimi de Ocampo Mixtec

The data for this variety comes from Swanton & Mendoza Ruíz (forthcoming). Tones are marked with numbers (C3). There are three level tones (low, mid, high). All of them can appear in the first mora, but the high tone never appears on the second mora. There is also no mid-low combination, resulting in only 5 contrastive tone melodies in our data.

Group 76

San Sebastián del Monte Mixtec

The data for this variety come from collaborator and native speaker Juvenal Solano (MICOP, Oxnard CA). The collaboration with Juvenal Solano is part of the collaboration between UCSB Linguistics and MICOP. In this source (Solano 2020) tones are marked throughout: high = acute, mid = unmarked, low = grave (MULG).

This variety has three level tones (high, mid, and low) which freely combine with each other, resulting in 9 contrastive tone melodies.

The lost final glottal stop has a lowering effect, resulting in a merger of the low tone sets. *H generally has a mid tone reflex and *L a low tone. There is a further merger of HL with the HH set.

Table 47: Contrastive tone melodies in Monte Mixtec

1.1	3.1	5.1
1.3	3.3	5.3
1.5	3.5	5.5

References

Alexander, Ruth Mary. 1980. *Gramática mixteca de Atatlahuca.* Instituto Lingüístico de Verano, A.C.

Alexander, Ruth Mary. 1988. A syntactic sketch of Ocotepec Mixtec. In C Henry Bradley & Barbara E Hollenbach (eds.), *Studies in the syntax of Mixtecan languages*, vol. 1, 151–304. Dallas & Texas: Summer Institute of Linguistics and the University of Texas at Arlington.

Amith, Jonathan D. & Rey Castillo García. n.d. Recursos lexicosemánticos para el mixteco de Yoloxóchitl, municipio de San Luis Acatlán, Guerrero (Glottocode yolo1241; ISO 639-3 xty). unpublished.

Anderson, Lynn. 2006. *Vocabulario de palabras que se relacionan con el maíz en mixteco de Alacatlatzala, Guerrero.* México, D.F.: Instituto Lingüístico de Verano, A.C. 2nd edn.

Becerra Roldán, Braulio. 2015. *Un estudio fonológico del mixteco de Santo Domingo Huendio, Oaxaca.* México, D.F. ENAH MA thesis.

Carroll, Lucien Serapio. 2015. *Ixpantepec Nieves Mixtec word prosody*: University of California San Diego PhD dissertation.

Chao, Yuen-Ren. 1930. ə sistim əv "toun-letəz" [A system of tone letters]. *Le maître phonétique* 8(30). 24–27.

Daly, John P. 1978. Notes on Diuxi Mixtec tone. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session* 22(1).

Dürr, Michael. 1987. A preliminary reconstruction of the Proto-Mixtec tonal system. *Indiana* 11. 19–61.

Farris, Edwin R. 1992. A syntactic sketch of Yosondúa Mixtec. In C Henry Bradley & Barbara E Hollenbach (eds.), *Studies in the syntax of Mixtecan languages*, vol. 4, 1–171. Summer Institute of Linguistics and the University of Texas at Arlington.

Galindo Sánchez, Bernardo. 2009. *Vocabulario Básico Tu'un Savi - Castellano.* Xalapa, Veracruz: Academia Veracruzana de las Lenguas Indígenas 1st edn.

Hills, Robert A. 1990. A syntactic sketch of Ayutla Mixtec. In C. Henry Bradley & Barbara E. Hollenbach (eds.), *Studies in the syntax of Mixtecan languages*, vol. 2, Summer Institute of Linguistics and the University of Texas at Arlington.

Hollenbach, Barbarba Elena Erickson. 2017. *Diccionario mixteco de Magdalena Peñasco.* Instituto Lingüístico de Verano, A.C.

Johnson, Audrey F. 1988. A syntactic sketch of Jamiltepec Mixtec. In C Henry Bradley & Barbara E Hollenbach (eds.), *Studies in the Syntax of Mixtecan languages*, vol. 1, 11–150. Summer Institute of Linguistics and the University of Texas at Arlington.

Josserand, Judy Kathryn. 1983. *Mixtec dialect history*: University of Michigan Ann Arbor PhD dissertation.

Kuiper, Albertha & Joy Oram. 1991. A syntactic sketch of Diuxi-Tilantongo Mixtec. In C Henry Bradley & Barbara E Hollenbach (eds.), *Studies in the syntax of Mixtecan languages*, vol. 3, 185–408. Summer Institute of Linguistics and the University of Texas at Arlington.

Macaulay, Monica Ann. 1996. *A grammar of Chalcatongo Mixtec.* University of California Press.

Pensinger, Brenda J et al. 1974. *Diccionario Mixteco: mixteco del este de Jamiltepec, pueblo de Chayuco.* México, D.F.: Instituto Lingüístico de Verano, A.C.

Peters, Simon L. 2018. *The inventory and distribution of tone in Tù'un Ndá'vi, the Mixtec of Piedra Azul (San Martín Peras), Oaxaca.* University of California Santa Barbara MA thesis.

Pike, Eunice V & Joy Oram. 1976. Stress and tone in the phonology of Diuxi Mixtec. *Phonetica* 33(5). 321–333.

Pike, Eunice V & Priscilla Small. 1974. Downstepping terrace tone in Coatzospan Mixtec. In Ruth M. Brend (ed.), *Advances in Tagmemics*, 105–134. Amsterdam: North-Holland.

Reyes Basurto, Griselda. 2020. Diccionario del Mixteco de Tlahuapa, Guerrero, Mexico. unpublished.

Schultze-Jena, Leonhard. 1938. *Bei den Azteken, Mixteken und Tlapaneken der Sierra Madre del Sur von*

Mexico, vol. 3 Indiana. Jena: Gustav Fischer.

Shields, Jana. 1988. A syntactic sketch of Silacayoapan Mixtec. In C. Henry Bradley & Barbara E. Hollenbach (eds.), *Studies in the Syntax of Mixtecan Languages*, vol. 1, 305–449. Summer Institute of Linguistics and the University of Texas at Arlington.

Small, Priscilla C. 1990. A syntactic sketch of Coatzospan Mixtec. In *Studies in the Syntax of Mixtecan languages*, vol. 2, 261–479. Summer Institute of Linguistics and the University of Texas at Arlington.

Solano, Juvenal. 2020. Vocabulario del Mixteco de San Sebastián del Monte. unpublished.

Stark, Sharon C., Audrey P. Johnson & Benita González de Guzmán. 2013. *Diccionario básico del mixteco de Xochapa, Guerrero*. Instituto Lingüístico de Verano, A.C. 3rd edn.

Stark, Sharon C., Andrea Johnson Peterson & Filiberto Lorenzo Cruz. 1986. *Diccionario Mixteco de San Juan Colorado*. México, D.F.: Instituto Lingüístico de Verano, A.C.

Swanton, Michael & Juana Mendoza Ruíz. forthcoming. Observaciones sobre la diacronía del tono en el Tu'un Savi (mixteco) de Alcozauca de Guerrero. *SLI* .

Towne, Douglas. 2011. *Gramática popular del tacuate (mixteco) de Santa María Zacatepec, Oaxaca*. Instituto Lingüístico de Verano, A.C.

Williams, Judith F., Gerardo Ojeda Morales & Liborio Torres Benavides. 2017. *Diccionario mixteco de San Andrés Yutatío, Tezoatlán, Oaxaca*. Ciudad de México: Instituto Lingüístico de Verano, A.C.

Zylstra, Carol F. 1991. A syntactic sketch of Alacatlatzala Mixtec. In C Henry Bradley & Barbara E Hollenbach (eds.), *Studies in the Syntax of Mixtecan languages*, vol. 3, 1–177. Summer Institute of Linguistics and the University of Texas at Arlington.

Zylstra, Carol F. 2012. *Gramática del Tu'un Savi (la lengua mixteca) de Alacatlatzala, Guerrero*. Instituto Lingüístico de Verano, A.C.