Otomanguean historical linguistics: Exploring the subgroups

Eric W. Campbell

Abstract

This is the second of two articles that survey and assess progress and prospects in the historical linguistics of the Otomanguean language family, which is the most widely distributed and most diverse linguistic lineage of Mesoamerica. While considerable progress has been made in understanding the linguistic and cultural history of Otomanguean, in some ways it remains the Mesoamerican language family about which we know the least. It consists of eight major subgroups: Mè’phàà-Subtiaba, Chorotegan, Oto-Pamean, Chinantec, Mixtecan, Amuzgo, Zapotecan, and Popolocan. While the first article of this series addressed the historical linguistics of the higher levels of the Otomanguean family, this article focuses on the eight major subgroups, especially progress in their reconstruction, internal subgrouping, and areas in need of further research.

1 Introduction

Of the linguistic lineages of the Mesoamerican cultural and linguistic area (Kirchhoff 1967[1943]; Campbell et al. 1986; Carmack et al. 2007), the Otomanguean language family is the most widespread and most diversified. It reaches as far north as the Mexican state of San Luis Potosi (Pame), and it previously reached southeast to the Gulf of Nicoya in Costa Rica (Mangue). Kaufman (1988; 2006; 2015a) divides the Otomanguean family into Western Otomanguean and Eastern Otomanguean, each of which splits into two large subgroups, which each in turn split into two MAJOR SUBGROUPS. Of these eight major subgroups, several display considerable diversity of their own (Zapotecan, Mixtecan, Popolocan, Oto-Pamean), while others consist of a handful of languages and varieties (Chinantec), a few closely related
languages (Mè’phàà-Subtiaba, Chorotegan), or perhaps one language with a few varieties (Amuzgo). While Otomanguean is distributed over a much smaller area than the Indo-European language family, its overall diversity and the number and variable depth of its major subgroups are roughly comparable (Kaufman & Justeson 2010).

The eight major subgroups in turn consist of from one (Amuzgo) to six (Oto-Pamean) ETHNOLINGUISTIC GROUPINGS (INALI 2009), which themselves may consist of one or many languages (Figure 1). In most cases, communities within an ethnolinguistic grouping recognize some collective identity or history and their speech varieties usually have cognate autonyms with shared Spanish and English translations.

![Figure 1. Otomanguean classification (based on Kaufman 1988)](image-url)
The present article is the second of two articles whose main goals are to (i) provide a critical overview and state of the art of Otomanguean historical linguistics, (ii) highlight interesting or problematic aspects of the field, and (iii) identify areas in need of further attention. While the first article (Campbell 2017) addresses the higher levels of the family, the present one focuses on reconstruction and internal subgrouping within the eight major subgroups: Mè’phàà-Subtiaba (§2), Chorotegan (§3), Oto-Pamean (§4), Chinantec (§5), Mixtecan (§6), Amuzgo (§7), Zapotecan (§8), and Popolocan (§9). Typological characteristics of particular subgroups are mentioned in the sections, aside from features that are common to most of Otomanguean: VS and usually VAO constituent order, general head-initiality, head-marking, tone, laryngeal articulations, parataxis and juxtapositional structures, vigesimal numeral systems, body part terms as relational nouns (with few prepositions), prolific lexicalization through compounding, complex inflectional classes, lack of voice alternations but presence of layered unproductive valence-changing derivations, and main verbs incorporated into auxiliaries. Conclusions are given in §10.

2 Mè’phàà-Subtiaba

Mè’phàà is a cluster of speech varieties spoken in the eastern part of Guerrero state. It was previously referred to as Tlapanec, Yope, or Yopi, but community educators recently established the preferred autonym “Mè’phàà” (Carrasco Zúñiga 2006: 12; Navarro Solano 2012: 12). These languages display case marking, split ergativity with agentive marking in certain contexts, as well as indexation of 3rd person referents as given or new (Wichmann 2014); the morphology is synthetic and somewhat fusional.

Lehmann (1920) recognized that Mè’phàà was closely related to the geographically distant Subtiaba (also known as Nagranda or Maribio), a dormant and sparsely documented language that was spoken around León, Nicaragua (Campbell 1975). Mason (1940) referred to the pair as...
“Supanecan,” and Sapir (1925a: 403) considered them to be “dialects of a single language.” Radin (1935: 45) suggested that Azoyú Mé’phàà was more divergent from other Mé’phàà varieties than was Subtiaba, and Kaufman (1983: 118) concludes that Mé’phàà is as phonologically diverse as Mé’phàà and Subtiaba together.

Lehmann (1920: 973–978) suggested that Subtiaba might be related to the distant Washo language of California; Sapir (1925a, 1925b) followed up and argued that Subtiaba belonged to the Hokan-Coahuiltecan stock, whose own genetic unity remains controversial (Campbell 1997: 295). Weitlaner (1941: 252) suggested that Mé’phàà might be Otomanguean, and Swadesh (1960: 85) grouped it closely with Chiapanec-Mangue, ultimately related to his Macro-Mixtecan. But it was Rensch (1976, 1977a) who finally demonstrated that Mé’phàà—and by association Subtiaba—was Otomanguean. Oltrogge (1977) argued that Subtiaba was instead related to Tol (Jicaque) of Nicaragua and Chontal of Oaxaca (Tequistlateco), both generally considered isolates or possibly related to each other (Campbell 1997: 160).

Suárez (1979) endorsed but revised Rensch’s account of Mé’phàà’s phonological development from proto-Otomanguean and rejected the Hokan and Tol-Chontal proposals (Suárez 1983: 2). Bright (1978) was not convinced by Rensch, and Suárez (1986) responded by adding evidence from (i) idiosyncratic sound alternations shared between Mé’phàà and other Otomanguean languages, and (ii) cognate morphology and grammatical particles. Kaufman (1983: 45-47) lays out sound correspondences between Mé’phàà-Subtiaba and the other Otomanguean major-subgroup protolanguages and identifies a couple of morphological innovations shared with Chorotegan (Kaufman 1988, 2016a).

No reconstruction of proto-Mé’phàà-Subtiaba exists yet, so shared innovations cannot be identified for subgrouping. Wichmann (2014) considered short vocabulary lists to estimate relative lexical similarity between varieties. He places Subtiaba as the least similar, followed
by Azoyú, then Malinaltepec, and then Tlacoapa: the more southeasterly the variety, the more lexically divergent (Figure 2).

Figure 2. Mè’phàà-Subtiaba lexical similarity (Wichmann 2014)

3 Chorotegan: Mangue and Chiapanec

Chorotegan is the sister to Mè’phàà-Subtiaba. It consists of the now dormant Mangue and Chiapanec languages. Aside from a sketch Chiapanec grammar (de Albornoz 1875), little morpho-syntactic information about these languages exists. Mangue (also known as Chorotega or Diria) was spoken along the Pacific slope in what is now El Salvador, Honduras, Nicaragua, and Costa Rica. Brinton (1886) mentions that Mangue had already fallen out of use in much of its earlier range and that it was closely related to the Chiapanec language of Chiapas, Mexico. Brinton (1892) later reported that Chiapanec was related to Mazatec. Lehmann (1920) compiled and discussed much of the existing Chiapanec and Mangue data and classified them with Popolocan and Otomi-Mazahua in the group that Schmidt (1977 [1926]) later named “Otomí-Mangue,” the source of the name “Otomanguean” (Jiménez Moreno 1962).

Fernández de Miranda et al. (1959: 57) distanced Chorotegan from Otomanguean, based on lexicostatistical calculations. But then Fernández de Miranda & Weitlaner (1961) reconstructed 270 proto-Chiapanec forms and 159 proto-Chorotegan forms and compared them to Gudschinsky’s (1959) proto-Popolocan and Longacre’s (1957) proto-Mixtecan. They proposed
100 proto-Popolocan-Mangue reconstructions, placing Chorotegan back in Otomanguean. Echoing Brinton and Lehmann, Fernández de Miranda & Weitlaner (1961: 8) noted striking lexical similarity between Chorotegan and Popolocan, and Hamp (1964: 393) cautiously suggested a closer connection between those two groups than either with Mixtecan. Kaufman (1988) classifies Chorotegan and Popolocan as maximally distant within Otomanguean but finds grammatical evidence for contact between them before the (proto-)Chorotegans migrated south.

4 Oto-Pamean: Otomí, Mazahua, Matlatzinca, Tlahuica, Chichimeco Jonaz, and Pame

Oto-Pamean is the northernmost Otomanguean subgroup, spanning a large area of central Mexico. It is linguistically divergent, having restructured the verbal inflectional system (Kaufman 1988) and having undergone significant phonological reduction (Kaufman 1983: 61). It boasts some of the largest consonant inventories and most elaborate TAM systems found in the family. Oto-Pamean languages tend towards semantic alignment, have noun classes, and have a dual number category (Palancar, To appear).

Based on lexical and grammatical similarities, Pimentel (1903a: 434) proposed an “Othomí” language family consisting of Otomi, Mazahua, Pame, and Chichimeco Jonaz. He excluded Matlatzinca because of its polysynthetic structure, citing language contact as the source of shared similarities (Pimentel 1903b: 408). Soustelle (1937) wrote ethnographies and phonological and grammatical sketches of Otomí (San José del Sitio), Mazahua, Matlatzinca, Tlahuica, Pame, and Chichimeco Jonaz. He argued that (i) the phonological, lexical, and grammatical similarities among them were too deep to be due to chance or contact and they formed an “Otomí-Pame” family; (ii) they form pairs based on closest affiliation: Otomí-Mazahua, Matlatzinca-Tlahuica, and Pame-Chichimeco Jonaz (see also Pimentel 1903b: 426); and (iii) Otomí-Mazahua and Matlatzinca-Tlahuica together form a larger subgroup (Soustelle 1937: 439).
Bartholomew (1965: 1-10) identifies shared innovations and confirms Soustelle’s subgrouping, arguing against the competing lexicostatistical proposals of Manrique Castañeda (1958) and Swadesh (1960). Kaufman (1988) concurs with Soustelle and Bartholomew (Figure 3).

Bartholomew (1965) reconstructs 811 proto-Oto-Pamean forms, including tone, and inflectional morphology, but some of her reconstructions are based on data only within various Oto-Pamean subgroups. Kaufman (1983) replaces Bartholomew’s (1989) revised diphthongs with sequences involving glides and her “centralizing laryngeal” with the vowels *i and *ə.

4.1 Otomi and Mazahua

The first application of the Comparative Method to Oto-Pamean languages was Newman & Weitlaner’s (1950a) 333 proto-Otomí reconstructions, based on eleven Otomi varieties. They identified about a dozen lexical isoglosses and reconstruct some Spanish loans to proto-Otomí, suggesting only recent diversification. In their analysis, Otomí consists of four groupings: (i) Ixtenco, the most divergent; (ii) Northwestern (Mezquital, *ɔ merging with *a, and *mb merging with *b): Tepenené, Zimapan, Tecozautila (Hidalgo), and the variety documented in López Yépes (1826); (iii) Southwestern, the most diversified, including Tilapa, Magú, Octupán,
and Temoaya; and (iv) Northeastern (or Sierra Otomi), including San Pablito (Puebla) and San Pedro Tlachichilco (Hidalgo).

Newman & Weitlaner (1950b) deepened their work to 203 proto-Otomí-Mazahua reconstructions (their “Primitive Central Otomian”), adding data from the Mazahua varieties of San Miguel Tenoxtitlan and Ixtapa de Oro. According to them Otomi *õ merged with *ũ, and *r partially merged with *n, but otherwise Otomi is conservative and Mazahua innovative. Bartholomew (1960) argues for removing *r, *ʃ, *s, and *ɲ from Newman & Weitlaner’s proto-Otomí consonant inventory and revises their reconstructed consonant clusters based on insights from three more varieties: San Felipe (México state), and Tlacotlapilco and San Gregorio of Hidalgo.

Lastra (2006: 44–63) surveys 39 Otomi varieties and presents phonological and lexical isoglosses, but some involve innovations, others retentions, and others just correspondences without any indicated directionality. Lastra (2010: 872) reports that Otomi and Mazahua show significant grammatical differences but share about 70% of their lexicon and could perhaps be considered a single dialect chain. Palancar (2013) traces shifts and losses of the proto-Oto-Pamean dual number in multiple Otomi varieties.

4.2 Matlatzinca and Tlahuica

Matlatzinca is now spoken only in the village of San Francisco Oztotilpan near Toluca, but was formerly spoken in nearby Mexicalcingo and near Charo, Michoacán (Soustelle 1937: 302), where it was known as Pirinda (Pimentel 1903b). Tlahuica is currently spoken in San Juan Atzingo and neighboring towns in the municipality of Ocuilan, southeast of Toluca (Muntzel 1986: 3).

Weitlaner (1953) sketched out sound correspondences between Matlatzinca and proto-Otomian and found less than 50% lexical correspondence. Hasler (1961) offered 190 proto-Matlatzinca-Tlahuica reconstructions but did not discuss many of the irregularities in the data.
Pérez (2007) reconstructs the proto-Matlatzinca-Tlahuica segmental inventory and identifies sound changes in each. She provides 888 cognate sets, another 132 possible cognates, and 4,670(!) lexical sets illustrating the many sound correspondences, but she does not reconstruct full forms. Pascacio Montijo (2014) reconstructs the proto-Matlatzinca-Tlahuica independent pronouns.

4.3 Pamean: Pame and Chichimeco Jonaz

Pame was earlier considered to consist of two languages: Southern Pame and Northern Pame (Soustelle 1937). The former is now dormant and was spoken in Jiliapan, Hidalgo (Gibson & Bartholomew 1979: 309), and the latter is now considered to consist of two languages of southern San Luis Potosí: Northern Pame and Central Pame (Avelino 2006; Lastra 2010: 865). Excluding recent migration, Chichimeco is spoken today in one town: Misión de Chichimecas, San Luis de la Paz, Guanajuato. The same variety is reported to have been spoken in a few other towns (Angulo 1933: 153), and a more divergent variety was spoken in Fuencleta, San Luis Potosí (Lastra 2010: 865).

Bartholomew (1959) added Pame to Newman & Weitlaner’s (1950b) proto-Otomían and Weitlaner’s (1953) proto-Otomán-Matlatzinca studies, which she later expanded into her in-depth proto-Oto-Pamean reconstruction (Bartholomew 1965).

5 Chinantec

Chinantec is the sister group of Oto-Pamean; it is spoken in northern Oaxaca and a few towns in Veracruz. Chinantec languages are largely monosyllabic, with highly elaborate tonal inflection for person marking and aspect; animacy is highly activated in the grammar and the languages have been reported to display accusative (Castellanos Cruz 2014: 54) or perhaps agentive alignment (Castillo Ramírez 2011: 31).
Weitlaner & Smith (1962) reconstructed the proto-Chinantec segmental inventory and a dozen words. Rensch’s (1963; 1968) proto-Chinantec included 230 reconstructions and significantly differed from Weitlaner & Smith’s (1).

(1)  

<table>
<thead>
<tr>
<th></th>
<th>Weitlaner &amp; Smith (1963)</th>
<th>Rensch (1968)</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘basket’</td>
<td>*mïʔ</td>
<td>*miʔL</td>
</tr>
<tr>
<td>‘cacao’</td>
<td>*ciaʔ</td>
<td>*ziʔLH</td>
</tr>
<tr>
<td>‘possum’</td>
<td>*hái</td>
<td>*húH</td>
</tr>
<tr>
<td>‘fire’</td>
<td>*hïʔi̯</td>
<td>*hyiL</td>
</tr>
<tr>
<td>‘beans’</td>
<td>*hnai̯</td>
<td>*hniuL</td>
</tr>
</tbody>
</table>


Egland (1983) classified Chinantec varieties into 14 groupings based on inter-intelligibility surveys, but Rensch (1989) arrived at a much different subgrouping when he expanded his proto-Chinantec reconstruction to an impressive 968 forms, the largest of any Otomanguean major subgroup. He identifies 47 phonological and 36 lexical innovations. Although many isoglosses overlap, several pairs of varieties group together by sharing the same phonological innovations: Palantla with Tepetotutla (Group II); Valle Nacional with Ozumacin (Group III); Lalana with Lealao (Group IV); and Quiotepec with Comaltepec (Group V). Ojitlán, Usila, Tlacoatzintepex, Chiltepec, Sochiapan, and Quetzalapa together share 7 phonological and 14 lexical innovations (Group I). Group I shares 5 more innovations with Tepetotutla alone and
nine further innovations with both Tepetotutla and Palantla, joining Groups I and II into a larger subgroup. Two other innovations are shared by all groups except for the divergent Group V (Figure 4).

Figure 4. Chinantec subgrouping (Rensch 1989)

6 Mixtecan: Trique, Cuicatec and Mixtec

Longacre (1955, 1957) argued that Trique, Mixtec and Cuicatec form the Mixtecan family. He considered Amuzgo to be somehow related but largely omits it for lack of data. He reconstructs final syllables of 279 proto-Mixtecan forms with data mostly from (i) Chicahuaxtla Trique; (ii) Concepción Pápalo Cuicatec; and (iii) the Mixtec varieties of San Miguel el Grande, San Esteban Atatláhuca, Jicaltepec, and Metlatónoc. Longacre explored the complicated patterns in penultimate syllables, which required a sketch of the Mixtecan tense-aspect system. He posited stem-initial consonant alternations for proto-Mixtecan, which Kaufman (1983: 57) handles with additional sound correspondences and prefixes. Early studies probing subgrouping within Mixtec and Cuicatec (Holland 1959) and Mixtec and Trique (Arana Osnaya 1960) followed Longacre’s work, but these were based on lexicostatistics and not shared innovations.
6.1 Trique

Trique is spoken in the Western Mixteca Alta region and is generally considered to consist of three not greatly divergent varieties: Chicahuaxtla, Copala, and Itunyoso (Hollenbach 1977; Kaufman 2006a; DiCanio 2008; Matsukawa 2008). It is the most innovative Mixtecan subgroup. It underwent a series of sound changes that precipitated a restructuring of its verbal inflectional system (Longacre 1957: 17, 58), and it was earlier classified outside of Mixtecan (Weitlaner 1941), sometimes in Popolocan (Mechling 1912). In disagreement with Longacre, Swadesh (1960) placed Trique outside of Mixtecan again (lexicostatistically), and Longacre (1961) countered with detailed evidence from shared innovations, fixing Trique’s place in Mixtecan. Matsukawa (2005) presents about 580 preliminary proto-Trique reconstructions based mostly on Chicahuaxtla and Copala.

6.2 Cuicatec

Cuicatec is spoken in the northern part of Oaxaca around Cuicatlán and is the group most closely related to Mixtec (Belmar 1902: 4). INALI (2009) lists three varieties, which Kaufman (2006a: 120) considers a single language. Longacre (1957) deals with Cuicatec historical phonology in some detail, and though a lengthy bilingual dictionary (Anderson & Concepción Roque 1983) and a few descriptive works exist (Needham & Davis 1946; Davis & Walker 1955; Bradley 1991; Feist & Palancar 2016), Cuicatec remains one of the least documented and described Otomanguean subgroups.

6.3 Mixtec

Mixtec is spoken in much of western Oaxaca, parts of eastern Guerrero, and a few towns in Puebla. Jiménez Moreno (1962) considers Mixtec a language with seven diversified “dialect complexes.” Josserand (1983: 457–458) states that “Mixtec must include over a dozen mutually unintelligible varieties, or languages, each with many local dialects,” and Lewis et al. (2015) list 51 Mixtec “languages.” As these varied numbers reflect, Mixtec is highly diversified and
its internal classification remains poorly understood. While larger regions such as the Mixteca Alta, Mixteca Baja, and Mixteca de la Costa may have geographic utility, they are not linguistic groupings.

Mak & Longacre (1960) consider 28 Mixtec varieties and reconstruct final syllables of 91 proto-Mixtec forms, traced from Longacre’s proto-Mixtecan. Longacre (1957: 113) explains that he only reconstructs final syllables because they are the most likely to be fully cognate across Mixtecan, but it’s not clear why Mak & Longacre did not reconstruct full forms at least within Mixtec proper. They also leave tone aside, even though Longacre had reconstructed proto-Mixtecan tone.

From 1977 to 1982, Josserand (n.d.) directed an extensive lexical and syntactic survey in 120 Mixtec villages, which is now curated in the Archive of the Indigenous Languages of Latin America (AILLA). These data fueled significant advances in Mixtec historical linguistics. Bradley (1981) reconstructed proto-Mixtec kinship terminology. Bradley & Josserand (1982) investigated Mixtec dialectology and provided 45 proto-Mixtec reconstructions (without tone) and 16 sound changes with relative chronologies. They provide isogloss maps and discuss likely migrations and a diachronic stratigraphy of diffusion spheres, but offer no subgrouping. Josserand et al. (1984) examine colonial era Mixtec documents and determine that they represented local varieties and not some written standard or lingua franca Mixtec.

Josserand’s (1983) Mixtec dialect history focuses on vowels but also includes 188 proto-Mixtec reconstructions, with primary data from the 120 surveyed Mixtec villages. She posits 12 Mixtec dialect areas (Figure 5).
Higher subgrouping of Josserand’s Mixtec dialect areas is yet to be achieved, and she cautions that several divergent Mixtec groupings “look alike despite their checkered phonological histories… and only a thorough knowledge of the historical development of these varieties can result in an accurate genealogical classification” (Josserand 1983: 459). There is some evidence (*x > tʃ) that Coastal varieties emigrated from around San Juan Mixtepec (Bradley & Josserand 1982: 293, 297; Josserand et al 1984: 156), before Lord 8 Deer “Jaguar
Claw” ruled the Coastal kingdom of Tututepec (Smith 1963; Spores 2000; Joyce et al. 2004), but Josserand does not propose any Mixtepec-Coastal subgroup.

Kaufman (1983) reworks proto-Mixtec, and the differences between his, Mak & Longacre’s (1960), and Bradley & Josserand’s (1982) segmental inventories are as follows:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>*n d</td>
<td>*n d</td>
<td>*nt</td>
</tr>
<tr>
<td>*tn</td>
<td>*t &gt; tn / __ V</td>
<td>*t &gt; tn / __ V</td>
</tr>
<tr>
<td>*θ</td>
<td>*s</td>
<td>*s</td>
</tr>
<tr>
<td>*h</td>
<td>*x (velar?)</td>
<td>*x (velar)</td>
</tr>
<tr>
<td>*v</td>
<td>*w</td>
<td>*w</td>
</tr>
<tr>
<td>*m</td>
<td>*w &gt; m / __ V</td>
<td>*w &gt; m / __ V</td>
</tr>
<tr>
<td>*ñ</td>
<td>*j &gt; ñ / __ V</td>
<td>*j &gt; ñ / __ V</td>
</tr>
<tr>
<td>?</td>
<td>*V? (vocalic feature)</td>
<td>?</td>
</tr>
<tr>
<td>*a</td>
<td>*æ</td>
<td>*æ</td>
</tr>
<tr>
<td>*Vm#</td>
<td>*V</td>
<td>*V</td>
</tr>
</tbody>
</table>

Proto-Mixtecan stem-final glottal stop has been lost in apparently all Mixtec varieties except for Santa Maria Zacatepec (“Tacuate”, Western Coastal); Ayutla de los Libres, Guerrero (Southern Baja, Pankratz & Pike 1967); and perhaps Nuyoo (Western Alta, Harris 1995; Harris & Harris n.d.). Dürr (1987: 20) reconstructs proto-Mixtec tone and concludes that lost final glottal stop has left high floating tones (see e.g. Pike 1948; Macaulay 1995) or low floating tones (Daly & Hyman 2007: 202) in some varieties.

Many and varied records of earlier forms of Mixtec varieties exist, which are valuable resources for linguistic and cultural (pre)history (van Doesburg et al., Under review): (i) codices of pre-Columbian Mixtec pictographic writing (Caso 1965; Smith 1973); (ii) early colonial era documentation of Mixteca Alta varieties such as Teposcolula and Yanhuitlán carried out by Spanish friars (Jiménez Moreno 1962: 100–103), including vocabulary (Alvarado 1962[1593]), grammar (Reyes 1889[1593]), and translated Christian doctrine (Hernández 1567); and (iii) numerous alphabetic writings by Mixtec speakers covering a range of topics, which
began appearing shortly after the Spanish invasion (Josserand et al. 1984; Terraciano 2001; Restall et al. 2005).

7 Amuzgo

Amuzgo is spoken in the Costa Chica region in three municipalities of Guerrero and two of Oaxaca (Apóstol Polanco 2014). Kaufman (1988, 2006) counts one Amuzgo language, treating it as an Otomanguean major-subgroup isolate, but Buck (2015: 11) claims that there are three rather differentiated varieties: Guerrero Amuzgo, San Pedro Amuzgos (Oaxaca), and Ipalapa (Oaxaca), the last of which is undocumented. Amuzgo has undergone near complete monosyllabification of roots and has recently been analyzed as having semantic alignment (Apóstol Polanco 2014).

Orozco y Berra (1864: 237) and Belmar (1901: 5) reported that Amuzgo was sister to Mixtec. Longacre (1957; 1961: 21) placed it coordinate with Mixtecan (see also Longacre & Millon 1961), but Swadesh (1960: 84, 108) disagreed, first placing Amuzgo inside of Mixtecan with Trique outside, and then later including both (Swadesh 1964). Longacre (1964: 1018, 1966a, 1966b) changed his own position and considered Amuzgo to be a primary Otomanguean branch, no more closely related to Mixtecan than to Popolocan or Chorotegan, and Rensch (1966, 1976) adopted this position. Based on shared innovations, Kaufman (1988) returned Amuzgo to a position coordinate with Mixtecan in an Amuzgo-Mixtecan subgroup of Eastern Otomanguean. Kaufman (1983) analyzes the Amuzgo ballistic syllables (Bauernschmidt 1965) as sequences of $Vh(\bar{\theta})$.

8 Zapotecan: Chatino and Zapotec

Zapotecan consists of the Chatino and Zapotec languages and occupies much of the central and southern parts of Oaxaca and a few villages in Veracruz. Zapotec is highly diversified, like Mixtec (Smith Stark 1995), while Chatino is considerably less so but consists of several languages. The state of the art of Zapotecan classification is discussed in the following
paragraphs and is presented in Figure 6. While intermediate subgrouping is more elaborated for Zapotecan than for any other Otomanguean group, details on the Zapotec side are tentative since many of the nodes are defined by only one or two innovations.

Figure 6. Zapotecan internal classification

8.1 Chatino

Several sound changes define Chatino as a group apart from Zapotec. The Proto-Zapotecan single and geminate consonant series (Swadesh 1947) unconditionally merged in Chatino, and several other unconditioned changes occurred: *t > h, *s > t, *t' > t, and *f > s (Kaufman 2016b), in counterfeeding orders (Campbell 2013). Early Chatino also underwent several conditioned changes (Campbell, in press):

\[(2) \quad \text{pZn} \quad \text{pCh} \]

a. *l > *n / __ V
b. *? > Ø / before [-sonorant] or *l
c. *V₁ > V₂ / __ ?/h V₂ (translaryngeal vowel harmony)
Sullivant (2016) examines Belmar’s (1902b) data from the dormant and sparsely documented variety of Santo Domingo Teojomulco and determines that it reflected several Chatino changes: degemination, \( *t > *h \), \( *t' > *t \), and \( *l > *n \) / \( \_V \). It did not undergo the sibilant changes \( *(s)s > *t \) or \( *(ʃ)ʃ > *s \), reflecting sibilant distributions like Zapotec’s. This demonstrates that (i) Kaufman had the direction of these changes correct, even if counter to expected directions of sound change (cf. Suárez 1980: 52), (ii) Teojomulco was a most divergent Chatino variety, and therefore (iii) all other varieties form a primary subgroup: CORE CHATINO.

Boas (1913) reported that (Core) Chatino consists of three groups: (i) Zenzontepec, Tlapanalquiahuitl and Tlacotepec; (ii) Tataltepec; and (iii) Juquila, Yaitepec and all others. Upson & Longacre (1965) reconstruct 251 proto-Chatino lexemes (not tone) with data from Boas’ three groups but mistakenly include Elotepec Papabuco (Zapotec) (Rendón 1971). Like Boas, they did not suggest further subgrouping. Campbell (2013) presents 117 refined or new proto-Chatino reconstructions, with tone (Campbell & Woodbury 2010), and identifies four shared innovations that support Boas’ Juquila and Yaitepec group (Eastern Chatino) and seven additional innovations shared between Eastern Chatino and Tataltepec, which form a higher-level subgroup: Coastal Chatino. Campbell (In press) adds 68 additional reconstructions, and Campbell & Cruz (2010) reconstruct proto-Chatino numerals.

8.2 Zapotec

Swadesh (1947) published 94 proto-Zapotec reconstructions based on four varieties: Ixtlán (Sierra Juárez), Yatzachi el Bajo (Cajono), Tehuantepec (Central), and Miahuatlán (Miahuatecan). Fernández de Miranda (1995) reconstructed 501 proto-Zapotec forms, adding three more varieties to Swadesh’s four: Rincón, Mitla, and Santa María Coatlán. Suárez (1973) reconstructed 167 proto-Zapotec forms using mostly Fernández de Miranda’s data, and Benton (1988) has a manuscript of proto-Zapotec reconstructions and another on proto-Zapotec tone
Kaufman (2016b) compiles, analyzes, and expands all of this earlier work and provides about 850 reconstructions and verbal morphology in what is now the state of the art in proto-Zapotec(an) reconstruction. Beam de Azcona et al. (In press) reduce the proto-Zapotec vowel inventory from six to five vowels.

Only two proposed innovations are shared by all Zapotec languages to the exclusion of Chatino: (i) loss of vowel nasality, and (ii) shift from final syllable to penultimate syllable prominence. Kaufman (2006a: 122) attributes these changes to contact with non-Otomanguean prestige languages, probably Mixe-Zoquean, which were the source of lexical borrowings into Zapotec(an) (Kaufman & Justeson 2007; Kaufman 2016b: 83).

The first characterizations of Zapotec diversification were geographic and impressionistic (Pimentel 1875; León 1902; Belmar 1905; Radin 1925). Angulo & Freeland (1935) proposed three high-level groups, which Fernández de Miranda (1995) maintained: Southern Sierra, Valleys (Central), and Northern Sierra. Swadesh (1947: 230) and Suárez (1990: 42, 50) challenged the tripartite classification, concluding that there was no demonstrable subgrouping. Kaufman (2016b) presents a binary branching with Western Zapotec splitting off first, then Papabuco, then “Southern” Zapotec, and finally Northern and Central Zapotec (including Isthmus), but only one or few innovations define his groups (Kaufman 1987-1989).

Smith Stark’s (2007) Zapotec classification was until recently the state of the art, but new findings require new revisions. He determines that all Zapotec varieties excluding Western Zapotec and Soltec (Suárez 1972) form a primary subgroup, Core Zapotec, defined by two (perhaps related) innovations: *kʷ > b (unconditionally), and *kkʷ > p in post-tonic position. These remain the only evidence for Core Zapotec.

The geographically westernmost varieties of Elotepec, Zaniza, and Texmelucan form the Core Zapotec subgroup known as Papabuco, but the only reported isogloss unique to Papabuco is the first person singular independent pronoun (j)a (Smith Stark 2007: 105).
Recent advances have whittled away what was earlier referred to as “Southern Zapotec,” which previously included the Extended Coatecan, Miahuatecan, Cisyautepecan, and Tlacolulita groups (Smith Stark 2007). Inter-intelligibility testing had suggested considerable diversity in the Southern Sierra (Weathers 1975), which may be a diffusion zone into which multiple Zapotec groups independently migrated (Beam de Azcona 2014). Based on two shared innovations, *ts > z and *tt > tʃ, Beam de Azcona posits a new Macro-Coatecan group consisting of three subgroups: Coatecas Altas, Coatecan, and Amatecan, the last consisting of Amatec and Tlacolulita (*e > a / _a; *ã > e).

As Suárez (1990: 56) had indicated in isogloss maps, Operstein (2012) finds that a conditioned split of *tt/ occurred in all of Smith Stark’s Core Zapotec except for Papabuco, (Macro-)Coatecan, and a couple of isolated exceptions. This change reflects a possible subgroup within Core Zapotec, call it NARROW CORE ZAPOTEC, that would include only Central Zapotec, Northern Zapotec, Miahuatecan (*ss > t and *s > θ) and Cisyautepecan (mb > m, Beam de Azcona 2014).

Northern Zapotec is usually divided into four clusters (Fuente 1947, Smith Stark 2007): Sierra Juárez (serrano, *s > θ, *ss > θ), Cajono (xhon, *t > tʃ, except Betaza *t > dz [Campbell 2011]), Rincón (nexitzo), and Choapan (vijano), the last two lacking defining innovations. Northern Zapotec varieties retain traces of a proto-Zapotec stem-final *-k (Kaufman 2015b).

Smith Stark (2007) includes ten groups in Central Zapotec, which has only one proposed defining innovation: a Progressive Aspect prefix ka(f)j-, present also in Cisyautepecan (Méndez Espinosa 2004), which may have grammaticalized from a verb of posture (Smith Stark 2004; Broadwell 2015). However, Kaufman (1983: 60, 1988) reconstructs this prefix as proto-Zapotecan *kkaj-, a reflex of proto-Otomanguean *kai PROGRESSIVE/’to be.’

Zapotecan’s diversification center of gravity lies in the southwestern part of its distribution: Chatino, Soltec, Western Zapotec, Papabuco, and Macro-Coatecan. A “least moves” model of
language spread and diversification (Sapir 1949 [1916]; Dyen 1956) would suggest a Zapotecan homeland somewhere in this southwestern area. However, the stratigraphy of loans into Zapotec(an) from Mixe-Zoquean language(s) (Kaufman & Justeson 2007: 200; Kaufman 2016b: 83), which lie further to the east, must also be considered, as well as potential correlations with ancient Zapotec hieroglyphic writing at Monte Albán and other sites (Caso 1928; Whittaker 1980; Urcid 2001; Marcus 2003). The script is only partly deciphered, but the limited linguistic evidence suggests that the latest writings may have represented a proto-Northern-Central Zapotec stage (Kaufman & Justeson 2008: 233).

9 Popolocan: Popoloca, Chocho, Ixcatec, and Mazatec

Popolocan is sister to Zapotecan and is distributed around the Tehuacán Valley and the Cañada Region, mostly in northern Oaxaca and southern Puebla. Popolocan languages are known for their rich noun classifier systems (Veerman-Leichsenring 1995: 241).

Inspecting phonetic, lexical and morphological features, González Casanova (1925) proposed that (i) Popoloca and Chocho were very closely related, (ii) the two were fairly closely related to Ixcatec, and (iii) all three were possibly related to Mazatec. This is now the consensus view (Hamp 1958, 1960; Veerman-Leichsenring 2000: 336). The most in-depth historical linguistic treatment of Popolocan is Gudschinsky’s (1959) monograph, in which she reconstructs 356 proto-Popolocan forms, comparing proto-Mazatec (Gudschinsky 1956) with Popoloca, Chocho and Ixcatec, and cross-checking with Longacre’s (1957) proto-Mixtecan. Longacre (1962) revises Gudschinsky’s proto-Popolocan and presents additional Mixtecan comparisons. Veerman-Leichsenring (2000, 2004) reconstructs proto-Popolocan independent pronouns and noun classifiers, but see Costaouec & Swanton (2015) for revision of the latter.

9.1 Popoloc: Popoloca, Chocho, and Ixcatec

Ixcatec is a near-dormant language spoken by a few elders in Santa María Ixcatlán, Oaxaca. Ngigua (Xru Ngiwa, Chocho, Chocholtec) is a cluster of endangered speech varieties spoken
in the Coixtlahuaca and Tamazulapan Valleys of northern Oaxaca (Mock 1982; van Doesburg & Swanton 2011). Ngigua (Popoloca proper) consists of three clusters: Northern, Western, and Eastern—whose labels indicate their positions with respect to Tehuacán, Puebla.

With data from Ixcatec and Ngigua (Popoloca of San Felipe Otlaltepec and Chocho of Teotongo and Santa Catarina Ocotlán), Fernández de Miranda (1951) reconstructed 200 proto-Popoloc forms (without tone). She outlines the sound changes in each language and suggests a closest affiliation between Chocho and Ixcatec. Hamp (1958, 1960) examines her work and instead pairs Chocho with Popoloca. Kaufman (2006; 2015a) does not group Chocho apart from the three Popoloca clusters but places all four as coordinate.

9.2 Mazatec

Mazatec is spoken in northern Oaxaca and neighboring areas of Veracruz and Puebla. Gudschinsky (1956) reconstructed 93 proto-Mazatec forms based on seven varieties: Huautla de Jiménez, San Gerónimo Tecoatl, Mazatlán, Santa María Jiotes, San Miguel Huautla, San Pedro Ixcatlán, and Soyaltepec. She excluded tone because only Soyaltepec and Huautla had tonal analyses at that time, but she later decided that those two were sufficient for reconstructing proto-Mazatec tone (Gudschinsky 1959). She reports that Chiquihuitlán is the most divergent variety (Gudschinsky 1956: 2), an observation later supported in Kirk’s (1970) dialect intelligibility study.

In a first Mazatec subgrouping paper, Gudschinsky (1955) argues that Huautla de Jiménez separated early from the more lowland San Miguel Huautla and the first became a prestige variety that lexically and phonologically influenced the second. In another paper, focusing on innovations, Gudschinsky (1958) splits Mazatec into High Mazatec and Low Mazatec, each with internal subgroups based on single sound changes. She analyzes lexical isoglosses due to early lexical diffusion, retention of proto-Mazatec variation, and later innovations, discussing how the genetic and diffusional picture may correlate with Mazatec political history.
Kirk (1966) expands the reconstruction of proto-Mazatec, adding San Bartolomé Ayautla, Jalapa de Díaz, Santo Domingo del Río, and Cuaunecuitla to Gudschinsky’s (1956) eight varieties. He reconstructs 731 proto-Mazatec forms, with tone, and outlines the changes in the daughter languages, providing the primary data and sound correspondences. Kirk (1985) later adds a short discussion about proto-Mazatec numerals.

Leonard et al. (2012) propose a quite different subgrouping from Gudschinsky’s (1958), based on 31 phonological traits examined in a geolinguistic database created from Kirk’s (1966) reconstructions and cognate sets, but the linguistic argumentation for their classification is not provided. Swanton et al. (2013) have presented an intriguing study of the distributions of Mazatec innovations and their correlation with current and historical footpath networks that connect(ed) Mazatec communities.

10 Discussion and Conclusions

This article has surveyed the historical linguistics of Otomanguean’s eight major subgroups. Initial progress was made in the 1950s and 1960s. The 1980s saw significant advances in Mixtec, Chinantec and Kaufman’s proto-Otomanguean, and proto-Zapotecan has steadily been refined up to the present. Further documentation of Otomanguean languages is now crucial for expanding and refining the major-subgroup protolanguages and their internal subgrouping. Fortunately, significant documentary and descriptive work is underway, increasingly by young native-speaker linguists.

Unfortunately several Otomanguean language subgroups are highly endangered: Ixcatec, Chocho, Matlatzinca, and Tlahuica. The Pamean languages are also endangered and are some of the least described Otomanguean languages, and they are of special interest because of their location along and outside the northern limits of Mesoamerica. The most divergent varieties of Chatino and Zapotec are dormant or poorly documented, and with Chiapanec, Mangue, and Subtiaba now dormant, Mè’phàà is the only remaining group of half of Western Otomanguean.
The divergent Azoyú Mè’phàà variety is quite endangered (Lewis et al. 2015), making documentation of that and other Mè’phàà varieties (e.g. Marlett 2011) a high priority.

Rensch’s (1989) Chinantec etymological dictionary includes numerous reconstructions with tone, primary data from daughter languages, sound correspondences, comparative verb morphology, and subgrouping based on shared innovations, with consideration of overlapping isoglosses due to inter-group diffusion. Future work for other subgroups should follow such a model and attempt to correlate results with those from ethnohistory and archaeology. Only with expanded and refined subgroup reconstructions can we advance our understanding of the higher levels of Otomanguean (Campbell 2017), which can then be applied to studies of Mesoamerican linguistic and cultural prehistory.

ACKNOWLEDGEMENTS

Sincere thanks to Patience L. Epps and two reviewers for valuable commentary on earlier versions of this article. Parts of the Zapotecan section (§8) were presented at the Workshop on the State of the Art of Mesoamerican Linguistics at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, December 2014. Any remaining errors are mine alone.

References


Beam de Azcona, Rosemary; Francisco Arellanes Arellanes; Mario E. Chávez Peón; Mario Hernández Luna; Sofía Gabriela Morales Camacho; Carlos de Jesús Wagner Oviedo & Miriam Itzel Manzano Corona. In press. *Umlaut* (armonía vocálica) en el desarrollo histórico de las lenguas zapotecas. In: Lucero Meléndez Guadarrama, Cristina Buenrostro & Marcela San Giacomo Trinidad (eds.), *Discusiones recientes en torno a la lingüística histórica indomexicana*. Mexico, D.F.: UNAM.

Belmar, Francisco. 1901. *Investigación sobre el idioma amuzgo*. Oaxaca, Mexico.


Pimentel, D. Francisco. 1875. Cuadro descriptivo y comparativo de las lenguas indígenas de México, Tomo III. México: Tipografía de Isidro Epstein.


Suárez, Jorge A. 1983. La lengua tlapaneca de Malinaltepec. Mexico, D.F.: Universidad Autónoma de México.


van Doesburg, Sebastián; Michael W. Swanton; Alejandro de Ávila y Christian DiCanio. (Under review). Flores blancas, campos quemados y quetzales: la morfología histórica mixteca y la


