Otomanguean historical linguistics: past, present and prospects for the future

Abstract

Among the linguistic lineages of Mesoamerica, the Otomanguean family is the most diverse and most widely spread. Long occupying a central position in one of the cradles of human civilization, speakers of Otomanguean languages have played important roles in the region, about which their languages have much to tell. However, Otomanguean is perhaps the least understood of the major Mesoamerican language families, due to its great diversity, the remarkable structural complexity of Otomanguean languages, and the history of the field of Otomanguean historical linguistics, which has seen great achievement alternating with periodic controversy and doubt. With a focus on the higher levels and more ancient time depths of the family, this article surveys Otomanguean historical linguistic work and presents a state of the art perspective on Otomanguean classification, reconstruction, linguistic prehistory, remaining challenges, and prospects for the future.

1 Introduction

Otomanguean is an expansive language family that has been centered around the core of the Mesoamerican cultural (Kirchhoff 1967[1943]) and linguistic area (Campbell et al. 1986) for as long as we can detect. It extends a little beyond the northern limits of Mesoamerica into the state of San Luis Potosí, Mexico (Pame), and it previously reached as far south as the Gulf of Nicoya along the Pacific slope of Costa Rica (Mangue). While the family still occupies a wide range, it now falls entirely within Mexico, not counting recent emigration. Its ancient and central, yet widespread, position in one of the cradles of civilization, where agriculture, complex states, elaborate monuments, and even writing were developed, make Otomanguean linguistic history important for understanding not only Mesoamerican (pre)history but also human history and cultural development more generally.

No other accepted Mesoamerican language family is as diverse as Otomanguean, and it remains the Mesoamerican family about which we know the least (Kaufman & Justeson 2010:
226–227), at least relative to its diversity. While this diversity itself may be one reason for this, another reason is that Otomanguean languages present remarkable complexity in their sound systems and word structure, which has made them challenging to analyze, describe, and compare, but which also makes them of general synchronic, diachronic and aesthetic interest.

The name “Otomanguean” is attributed to Schmidt (1977 [1926]), who listed an “Otomi-Mangue” group on his map of North and Middle American languages, and the name was later reduced to “Otomangue” by Jiménez Moreno (1936; 1962: 62).¹ These are just two of the many names that have been used to refer to Otomanguean, and these authors are only two of the many voices that have played roles in the long, contentious, and challenging process of determining that Otomanguean is a genetic unit, which languages belong to it, and how those languages are more or less closely related to each other.

The most reliable means for answering these questions is the comparative method of historical linguistics. Its basis in the regularity of sound change enables us to reconstruct the sounds and words of unattested earlier stages of languages, which in turn provide a foundation for handling less regular morphosyntactic and semantic change, and which all together may be leveraged for identifying changes due to language contact and for exploring human intellectual, social, and cultural prehistory in coordination with other fields such as archaeology, ethnohistory, epigraphy, and genetics.

This article is a critical account of past and present historical linguistic research on Otomanguean languages and has three primary goals: (i) to provide an informative and reference-rich resource for students and scholars of the many fields for which historical

¹ Otomanguean is sometimes written with a hyphen: “Oto-Manguean.” The hyphenless spelling reflects that the name does not indicate a coordination of two genetic groupings, like Sino-Tibetan, Mixe-Zoquean, or Oto-Pamean, or even two geographic regions, as in Afro-Asiatic or Indo-European.
linguistics is relevant, (ii) to shed light on interesting or problematic aspects of Otomanguean historical linguistics, and (iii) to offer suggestions for future work. The scope of this article is largely limited to topics involving the higher-level phylogeny of Otomanguean; a companion article dealing with Otomanguean’s major subgroups will follow. In §2 the current classification of Otomanguean languages is presented, and the history of scholarship and supporting evidence is surveyed. A few points of interest in the structure of Otomanguean languages are outlined in §3. Topics in language and prehistory are discussed in §4, and proposals for long distance relationship in §5. Conclusions are provided in §6.

2 Classification and reconstruction

This section presents the currently most recognized classification of Otomanguean languages (§2.1), references to earlier classification proposals (§2.2), a summary of research on Otomanguean comparative reconstruction and subgrouping (§2.3), and a recent challenge to Otomanguean as an established language family (§2.4).

2.1 Current Otomanguean language classification

The most widely recognized Otomanguean language classification is Kaufman’s (1988, 2006a, 2015a, 2016b). He first divides the family into Western Otomanguean and Eastern Otomanguean, which each split into two subgroups. Those four subgroups each bifurcate, yielding eight MAJOR SUBGROUPS, which are roughly on the order of Indo-European subgroups such as Romance, Germanic, Indo-Iranian, Balto-Slavic, etc. Kaufman’s high-level binary branching is shown in Figure 1. The rightmost part of the figure reflects the internal structure of the eight major subgroups, which is detailed in the sequel article to this, and is shown here.

2 Line length in the diagram does not necessarily reflect relative time depth.
only down to the level of the ETHNOLINGUISTIC GROUPING (INALI’s 2009 *agrupaciòn lingüistica*).^3^ Some names and spellings used here differ from those found in other sources. Kaufman’s dehispanicized spellings reflect the Mayan indigenous academic tradition (Mateo Toledo 2003), but no such agreement across Otomanguean groups yet exists. Some groups have recently established autonymic standards, rejecting names of external or derogatory origin: e.g. Tlapanec → Mè’phàà (Carrasco 2006). Other groups are increasingly using endonyms, e.g. Mixtec → Tu’un Savi (Guadalupe Joaquina 2014), but these may vary depending on the variety: *tnu’u³² dauc³²* (Ramírez Pérez 2014). Names and spellings are important, sensitive, dynamic, and political. An attempt is made here to reflect community desires in clearer cases, but when in doubt, or when there is risk of privileging one variety over others, more standard names and spellings are used.

Figure 1. Otomanguean classification (based on Kaufman 1988)

---

^3^ Some names and spellings used here differ from those found in other sources. Kaufman’s dehispanicized spellings reflect the Mayan indigenous academic tradition (Mateo Toledo 2003), but no such agreement across Otomanguean groups yet exists. Some groups have recently established autonymic standards, rejecting names of external or derogatory origin: e.g. Tlapanec → Mè’phàà (Carrasco 2006). Other groups are increasingly using endonyms, e.g. Mixtec → Tu’un Savi (Guadalupe Joaquina 2014), but these may vary depending on the variety: *tnu’u³² dauc³²* (Ramírez Pérez 2014). Names and spellings are important, sensitive, dynamic, and political. An attempt is made here to reflect community desires in clearer cases, but when in doubt, or when there is risk of privileging one variety over others, more standard names and spellings are used.
Some Otomanguean ethnolinguistic groupings are single languages with minor variation (e.g. Tlahuica, Chichimeco Jonaz, Ixcatec). Others consist of multiple languages, which may have their own internal variation (Otomí, Trique, Chatino, Mazatec), and still others are diversified themselves almost enough to be considered families in their own right (Zapotec, Mixtec, Chinantec). In striking contrast to the Mayan language family, which has seen over 30 years’ relative consensus that there are 31 Mayan languages (Campbell & Kaufman 1985: 188; Law 2013: 143), there is nothing near a consensus for the number of Otomanguean languages. In some areas variation may be so great but too gradual to ever count “languages.” In most cases, communities within ethnolinguistic groupings share some broader identity or recognize some degree of collective history, and refer to themselves or their languages with cognate autonyms (see e.g. Merrifield 1966: 581 for Chinantec).

2.2 Early Otomanguean classification proposals

Otomanguean has long been considered a genealogical linguistic grouping, but its name and membership have frequently changed over time. The many proposed Otomanguean classifications preceding Rensch’s (1966) proto-Otomanguean reconstruction will not be detailed here, but the primary references are provided in the next paragraph. Summaries of the early work can be consulted in Fernández de Miranda & Weitlaner (1961: 6–9), Rensch (1976: 1–8), and especially Jiménez Moreno’s (1962: 54–85) detailed account.

Early Otomanguean classificatory work can be roughly split into four periods according to the methodologies used. Works during the earliest period (1864-1911) were largely impressionistic, though nonetheless foundational (Orozco y Berra 1864: 25-29; Pimentel 1875; 4

---

4 Martínez Ortega (2012: 43-45) argues for using the name “Tlahuica” or the autonym pijyekajo instead of “Ocuiltec(o).”

5 Kaufman (1988) lists 48 virtual or emergent languages and “language areas” for Otomanguean, while Lewis et al. (2015) list 57 “languages” for Zapotec alone, and 52 for Mixtec.
Brinton 1891, 1892; León 1902; Belmar 1905; Thomas and Swanton 1911). Most proposals of
the second period (1912-1937) had some basis in closer examination of lexical and grammatical
data, with a tendency toward splitting languages into separate families (Mechling 1912;
Lehmann 1920; Angulo 1926a, 1926b; Angulo & Freeland 1935; Mendizábal & Jiménez
Moreno 1937; Soustelle 1937: 441; cf. Jiménez Moreno 1936). Works of the third period (1939-
1944) drew on comparison of typological features, which led to the more inclusive proposals
of Ecker (1939) and Weitlaner (1941)6 (see also Mason 1940; Radin 1944; cf. McQuown 1955).
During the fourth period (1959-1964), most works were based on lexicostatistics and
glottochronology (Swadesh 1950, 1952; Lees 1953) and included claims of longer-range
relations (see §5) (Fernández de Miranda et al. 1959; Swadesh 1959, 1960, 1964b). These
works provoked strong but respectful criticism for their lack of methodological rigor (Longacre
1960, 1961; Olmstead 1961; Callaghan & Miller 1962; Kaufman 1988), and lexicostatistics
and especially glottochronology are widely considered problematic methods for language
classification (Hojier 1956; Bergsland & Vogt 1962; Matisoff 2000).
If one thing were shared by all of the works just mentioned, it would be their lack, either
entire or relative, of argumentation based on careful application of the comparative method.

2.3 Comparative reconstruction(s)

While new approaches for establishing or refining language relationships continue to be
explored, the gold standard is still the comparative method, which requires systematic
comparison of large lexical data sets, identification of regular sound correspondences,

6 Weitlaner’s (1941) “Macro-Otomangue” group included an Olmec family that consisted of the Popolocan and Mixtecan families. The term “Olmec” refers to the earliest monumental Mesoamerican civilization (Jiménez Moreno 1942; Stirling 1968; Coe 1968), situated along the Gulf Coast of modern Veracruz and Tabasco states, whose population likely spoke a Mixe-Zoquean, not Otomanguean, language (Campbell & Kaufman 1976).
reconstruction of protolanguages, and tracing of changes and shared innovations in and among
daughter languages. The strongest proposals are those that are reinforced by evidence from
multiple or idiosyncratic grammatical correspondences, and all details of an analysis should be
carefully examined with consideration of likely or common pathways of change and
similarities due to language contact (Hock 1991: 556–580; Rankin 2003; Campbell & Poser
2008: 162–223; Crowley & Bowern 2010: 108–117). The remainder of this sub-section surveys
comparative reconstruction of Otomanguean major subgroups before Rensch (1966) (§2.3.1),
Rensch’s own contributions (§2.3.2), and Kaufman’s comparative Otomanguean (§2.3.3).

2.3.1 Subgroup reconstruction before Rensch (1966)

Comparative reconstruction of Otomanguean subgroups began with Swadesh’s (1947)
proto-Zapotec. In Oto-Pamean, Newman and Weitlaner (1950a) reconstructed part of proto-
Otomí and then added Mazahua (Newman and Weitlaner 1950b). Weitlaner (1953) added
Matlatzinca, and Bartholomew (1959) later added Pame. Fernández de Miranda (1951)
reconstructed proto-Popoloc (Popoloca, Chocho and Ixcatec), and Gudschinsky (1956) proto-
Mazatec. After Longacre (1955, 1957) reconstructed proto-Mixtecan, things took off: proto-
Popolocan (Gudschinsky 1959; Longacre 1962), proto-Mixtec (Mak & Longacre 1960), proto-
Choroteegan (Fernández de Miranda & Weitlaner 1961), proto-Chinantec (Rensch 1963, 1968),
proto-Chatino (Upson & Longacre 1965), and extensive reconstructions of proto-Oto-Pamean
(Bartholomew 1965) and proto-Mazatec (Kirk 1966).

Some of these works included the first systematic comparisons and preliminary
reconstructions across major subgroups: Gudschinsky’s (1959) 112 proto-Popotecan
reconstructions (proto-Popolocan and proto-Mixtecan); Fernández de Miranda & Weitlaner’s
(1961) 100 proto-Popoloca-Mangue reconstructions with proto-Mixtecan comparisons; and
Bartholomew’s (1965) proto-Oto-Pamean and proto-Popotecan sound correspondences and
cognate sets. Swadesh (1964a) provides about 200 reconstructions of his “proto-Oaxacan” (i.e.
Eastern Otomanguean plus Huave, the latter of which is no longer widely considered to be related to Otomanguean. Other works contributed to the reconstruction of proto-Otomanguean kinship terminology (Harvey 1963; Merrifield 1981) and the kinship system in the abstract (Casasa García 1979).

2.3.2 Rensch’s comparative Otomanguean phonology

Rensch (1966, 1976) modified the major subgroup reconstructions listed above, excluding Swadesh’s proto-Zapotec (1947), and reconstructed final syllables of 427 proto-Otomanguean forms. He did not include Subtiaba or Mè’phàà, which had been excluded from most Otomanguean discussion following Sapir’s (1925a, 1925b) influential classification of Subtiaba as Hokan—a relation now only possible if much more remote (see §5). Rensch reconstructs proto-Otomanguean phonology and devotes chapters to developments within each of the seven included major subgroups. His appendix includes his 427 cognate sets, which consist mostly of his reworking of the major subgroup reconstructions listed in §2.3.1 and the primary data that they are based on. While many of his sets or reconstructions (with superscript numerals representing tone), such as **(h)kʷen ‘mountain’ (Set 167) and **(n)(h)kʷe(h)(n)³ ‘straw mat’ (Set 181), might appear phonologically and/or semantically loose, his work received considerable praise from other Otomangueanists (Longacre 1977: 122; Suárez 1980; Kaufman 1983: 38). Rensch (1977a) subsequently added Mè’phàà-Subtiaba as an eighth major subgroup and traced its historical phonology from proto-Otomanguean, which was revised but enthusiastically endorsed by Suárez (1979) and later expanded by Kaufman (2016a).

Both Rensch (1976; 1977a) and Longacre (1977: 101) placed the seven, later eight, major subgroups all coordinate within Otomanguean. In another study, Rensch (1977b) highlighted phonological isoglosses and shared innovations across major subgroups that implied multiple population movements and frequently emerging and dissolving contact zones. The picture remained cloudy, and Rensch put forth no proposal for higher-level subgrouping. The other
major gap in Rensch’s work, and in Otomanguean historical linguistics of the time, was the lack of evidence from comparative morphology.

While Suárez (1979) was convinced by Rensch that Mè’phàà was Otomanguean, there were skeptical (e.g. Bright 1978). In response, Suárez (1986) brought to light evidence from idiosyncratic morphophonological alternations shared between Mè’phàà and other Otomanguean languages, and he identified a significant amount of cognate morphology and grammatical particles across major Otomanguean subgroups.

2.3.3 Kaufman’s comparative Otomanguean

Kaufman’s two main Otomanguean monographs deal with comparative phonology (Kaufman 1983) and comparative morphology (Kaufman 1988); they remain unpublished. Two recently web-published monographs contain some of their details: one on Otomanguean linguistic prehistory (Kaufman 2015a) and the other on the Otomanguean affiliation of Mè’phàà-Subtiaba and potential affiliation of Otomanguean with Hokan (2016a).

For comparative Otomanguean phonology, Kaufman (1983), like Rensch, worked through the major subgroup reconstructions listed in §2.3.1, adding additional work on Zapotec (Swadesh 1947; Suárez 1973) and Mixtec (Bradley & Josserand 1982). His revised analyses of the major subgroup phonologies and their discrepancies with earlier work are summarized in his appendices. He expands the cross-subgroup sound correspondences and significantly revises Rensch’s proto-Otomanguean phonology and the historical developments in the major subgroups. He advances two main arguments: (i) the consonant alternations proposed for proto-Mixtecan (Longacre 1957; Rensch 1976), proto-Popolocan (Gudschinsky 1959) and proto-Oto-Pamean (Bartholomew 1965) need not be analyzed as such nor reconstructed for proto-Otomanguean, and (ii) Longacre’s (1962: 35) theory of coda nasals having affected vowel quality was incorrect, and proto-Otomanguean instead had 9 vowels, not 4 as Rensch proposed.
Kaufman (1983) mentions that his working file contained 480 Otomanguean cognate sets, but his monograph and other publications do not include all of them or indicate exactly how they relate to his sound correspondences. Instead, he refers to Rensch’s numbered cognate sets, which from there can be traced back to the primary data that went into the major subgroup reconstructions or worked forward through Kaufman’s revised historical phonologies of the subgroups and Rensch’s proto-Otomanguean phonology. What this means is that one will not find all of the primary data, intermediate reconstructions, and proto-Otomanguean reconstructions laid out together in Kaufman’s monographs, and one will not find reference to all of Rensch’s reconstructions that Kaufman considered to be flawed or too semantically lenient. However, Kaufman (1983) does include his revisions of 83 of Rensch’s monosyllabic Otomanguean phonology. Kaufman also includes 18 bisyllabic proto-Otomanguean forms and cognate sets. In his web-published work on Otomanguean prehistory, Kaufman (2015a) includes 173 proto-Otomanguean reconstructions, 14 more for Western Otomanguean, 26 for Eastern Otomanguean, 15 for Oto-Pamean-Chinantec, 3 for Popolocan-Zapotecan, 14 for Amuzgo-Mixtecan, and 1 more particular to Tlapanec-Chorotegan, but without references to Rensch’s sets.

For Kaufman’s (1988) second major Otomanguean monograph he sifted through all grammatical descriptions of Otomanguean languages available at the time. Using his cross-subgroup sound correspondences, he presents his reconstruction of about 27 proto-Otomanguean tense, aspect, mood, voice, and nominalization markers and their positions of occurrence in the verbal templates of each major subgroup protolanguage. From there he reconstructs proto-Otomanguean verbal morphology, tracing function shifts and other

---

7 Kaufman’s cognate sets file has been deposited for ingestion into the Archive of the Indigenous Languages of Latin America.
innovations in order to propose his higher-level phylogeny of the family. The innovations defining Kaufman’s (1988, 2016a) high-level subgrouping are listed below:\(^8\)

**Western Otomanguean:** *ai* merged with *a*, *mu* HYPOTHETICAL, *kwa* ‘go/come to’ auxiliary, *kwi* ‘get up to’, *ci* NEGATIVE,

**Tlapanec-Chorotegan:** *kkwa ~ *kkwau* CAUSATIVE, *wai* IMPERSONAL, *ha* IMPERATIVE, *tau* OPTATIVE, and maybe *ni* ‘being’

**Oto-Pamean–Chinantec:** *ia* > *u*, monosyllabification of all roots, *mi* IMPERFECT, *rV* STATIVE/PERFECTIVE

**Eastern Otomanguean:** *ia* merged with *i*, *ea* merged with *a*, *kwe* animal classifier, maybe *se* CAUSATIVE

**Amuzgo-Mixtecan:** *ts* and *s* merged, *i= INDEFINITE > DURATIVE; *kwe* COMPLETIVE > POTENTIAL, causative and optative constructions reformed as auxiliaries followed by main verb in Potential Mood

**Popolocan-Zapoteco:** *au* merges with *u*, pOM *r* or *θ* > *t*, *(Y)tí PERFECT > PROGRESSIVE

### 2.4 Brown’s challenge to Otomanguean

Brown (2015a) has reviewed some of the published Otomanguean major subgroup reconstructions and Kaufman’s (1983) monograph and concludes that Otomanguean is not convincingly demonstrated as a genetic unit because the cross-subgroup sound correspondences are not cross-referenced to cognate sets there or in Rensch’s work, and otherwise the evidence is not sufficient to rule out chance or extensive areal diffusion as reasons for cross-subgroup lexical similarities (Brown 2015b). Since Otomanguean is a deep and highly diversified family, and several Otomanguean subgroups have undergone significant phonological change (§3), cognates across major subgroups are indeed difficult to recognize.

---

\(^8\) The reader may consult Kaufman (2016a) for proto-Otomanguean morphemes lost at each node and also for the innovations that define the major subgroups after the immediately higher level subgroups, e.g. from proto-Tlapanec-Chorotegan to proto-Mé’phàà-Subtiaba and from proto-Tlapanec-Chorotegan to proto-Chorotegan.
Furthermore, since most proto-Otomanguean lexical reconstructions are only single syllables of *(C)CV shape, the potential for chance resemblance is relatively high (Ringe 1999), especially if the semantics of compared forms are not exact, as in many of Rensch’s cognate sets. Kaufman’s (1988) reconstructed morphemes and verbal template through the various levels of the family are crucial evidence in support of Otomanguean unity, but Brown’s challenge will likely remain unresolved until somebody publishes further comparative phonological evidence illustrating regular sound correspondences across major groups that are cross-referenced to semantically tight cognate sets that yield more solid and numerous proto-
Otomanguean reconstructions.

3 Some typical Otomanguean features

All Otomanguean languages appear to be tonal, and it is probable that proto-Otomanguean was tonal. Although Fernández de Miranda & Weitlaner (1961) maintained some doubt about Chiapanec and Mangue, Brinton (1886: 244) wrote that the Mangue “words for bird, snake and flower are the same; but Albornoz gives this very example to illustrate the importance of accent, *nolô, a snake, nolô, a flower.”

Tones seem to change faster than segments do, and they are more challenging to confidently reconstruct. Josserand (1983: 243) states that “tone is among the first features to vary between towns speaking similar varieties of Mixtec,” and Bartholomew (1994: 351) ran into “problems because the two contrastive tones of Matlatzinca seemed to correspond to any and all of the three or four contrastive tones of the other” Oto-Pamean languages. Despite such challenges, preliminary tonal reconstructions have been put forth for proto-Mixtecan (Longacre 1957), proto-Mixtec (Dürr 1987), proto-Chinantec (Rensch 1968), proto-Oto-Pamean (Bartholomew 1965), proto-Mazatec (Gudschinsky 1959; Kirk 1966), proto-Popolocan (Gudschinsky 1959), proto-Zapotec (Benton 2001), proto-Chatino (Campbell & Woodbury 2010), and to some extent proto-Otomanguean (Rensch 1976).
The partly segmental and partly suprasegmental patterning of laryngeals or laryngealization poses challenges in the synchronic analysis of some Otomanguean languages (Macaulay & Salmons 1995; Golston & Kehrein 1998; Campbell 2014), especially when laryngeals interact with tone (Bradley & Josserand 1982: 283; Silverman 1997; DiCanio 2012). These challenges are only compounded in comparative studies (Longacre 1957: 75; Kirk 1966: 48; Fernández de Miranda 1951: 72), and such complexities make Otomanguean an important and intriguing case for understanding the diachrony of laryngeals, tone, and their interaction.

Contrastive vowel nasality is widespread in Otomanguean but absent in Zapotec, and Matlatzinca and Tlahuica both lost it, though independently (Pérez 2007: 235). Kaufman (2006a: 122) attributes these losses to ancient contact with non-Otomanguean prestige languages. Vowel nasality is also reportedly mostly lost in Mazahua of San Miguel Tenoxtitlan (Newman & Weitlaner 1950b), and Fernández de Miranda & Weitlaner (1961: 18) doubted that there were nasal vowels in Chiapanec. While nasal vowel correspondences are robust across Chatino languages, they are weak in Mixtecan (Longacre 1957: 30) and Mè’phàà (Suárez 1979: 372). Cross-family vowel nasality correspondences are likewise weak, and this has inspired proposals of post-vocalic nasal consonants, and not nasal vowels, in proto-Otomanguean (Rensch 1976: 38; Kaufman 1983).

Most Otomanguean languages lack rhythmic stress, but prosodic prominence falls on either final or penultimate syllables of stems, with most roots being historically bisyllabic. Kaufman (1983: 61) says that Zapotec, Mixtec and Cuicatec “shifted stress to the first syllable of stems.” In Mixtec, the greatest degree of phonological contrast occurs in final syllables (Longacre 1957: 113), which probably reflects the earlier position of prominence. In Trique and Amuzgo, prominence falls on final syllables (Josserand 1983: 140), suggesting that that was the proto-Mixtecan pattern. Proto-Mazatec is reconstructed with final-syllable prominence (Gudschinsky 1956: 7; Kirk 1966: 9, 167). In Mè’phàà, long vowels, nasal vowels, and multiple tones occur
only in final syllables (Suárez 1983: 6; Carrasco Zúñiga 2006: 68), and while the Chiapanec
records remain ambiguous for tone or accent, any traces of these are always on final syllables
(Fernández de Miranda & Weitlaner 1961: 18). While Muntzel (1986: 45) reports that in
Tlahuica polysyllabic words always have stress on initial syllables, Kaufman (1983: 61)
cautions that Oto-Pamean, Chinantec and Amuzgo preserve only one syllable of what in proto-
Otomanguean were perhaps up to four-syllable words or stems with clitic(s). The weight of
evidence thus suggests that proto-Otomanguean had final-syllable prominence: fewer changes
are implied than in the reverse direction, and nobody has suggested any explanation for
prominence shifts onto final syllables as Kaufman has proposed for the reverse.

Otomanguean languages have head-initial syntax and predominantly head-marking
morphology. Verbs may consist of multiple phonological words but tend to have a fairly fixed
templatic structure and consist minimally of a root with some aspect or mood inflection.
Kaufman’s (1988) reconstructed proto-Otomanguean (pre-)verbal template is shown in Figure
2. The preverbal slots may be prefixes, proclitics or particles, depending on the language.

<table>
<thead>
<tr>
<th>NEG</th>
<th>POS 4</th>
<th>POS 3</th>
<th>POS 2</th>
<th>POS 1</th>
<th>DERV</th>
<th>root</th>
</tr>
</thead>
<tbody>
<tr>
<td>negation</td>
<td>tense and time adverbs</td>
<td>aspect and mood</td>
<td>plural subject</td>
<td>auxiliary or higher predicate</td>
<td>deriv-ation</td>
<td>verb root</td>
</tr>
</tbody>
</table>

Figure 2. proto-Otomanguean verbal template (Kaufman 1988)

While the verbal template has been reduced in some Omanguean subgroups (e.g. Amuzgo,
see Apostol Polanco 2014) and restructured in others (Oto-Pamean, Kaufman 2015a: 10), it has
remained relatively stable across the family, even if the prosodic status of some of the positions
may vary across subgroups or languages. The Zoochina Zapotec example in (1) shows an initial
aspect prefix (position 3), a causative prefix in the auxiliary slot (position 1), a verb root, and subject and object enclitics. The Zoochina Zapotec (López Nicolás 2009: 42)

(1)  sh-w-âw=â=bâ?

ICPL-CAUS-eat=S.1SG=O.3INF

‘I feed him.’

The Mè’phàà example in (2) shows aspect in position 3, a causative marker in the derivational slot, the verb root, and an adverbial enclitic among the post-verbal person markers. Such adverbial enclitics are found widely throughout Otomanguean: e.g. Mazatec (Pike 1948: 124), Chinantec (Anderson 1990: 109), Zapotec (López Cruz 1997: 82), and Chatino (Rasch 2002: 139).

(2)  ni-tsi2-kh-a2=maʔ3=laʔ1=ne

CPL-CAUS-burn-A.3SG.INAN=already=E.2PL=that

‘You (pl.) already burned that.’

The Chalcatongo Mixtec example in (3) shows a temporal adverbial prefix (position 4), an aspect prefix (position 3) that is cognate to the Mè’phàà one in (2), a marker of plural subject (position 2), and verb root.

(3)  a-ni-ka-xáʔnã=Ø

already-CPL-PL-cut=3

‘They already cut (it).’

Otomanguean languages are noted for having impressively complex verbal inflectional classes, with rich allomorphy in person or aspect/mood inflection (usually position 3 and/or

---

9 A = absolutive, CAUS = causative, CPL = Completive Aspect, E = ergative, ICPL = Incompletive Aspect, INAN = inanimate, INF = informal, O = object, PL = plural, S = subject, SG = singular
tone change on the stem) (Smith Stark 2002; Wichmann 2006; Campbell 2011a; Palancar 2011).

4 Language and prehistory

From Otomanguean’s great time depth, spread, and central position in Mesoamerica we can imagine that Otomanguean groups played important roles in the development of Mesoamerican civilization. A key component of that development was agriculture, particularly the nutritious and productive triad of maize, beans, and squash (Kirchhoff 1967[1943]; Gasco et al. 2007). After “squash seed, and cobs of wild or incipiently domesticated maize” were excavated from a cave in the Tehuacán Valley of southern Puebla, carbon dated at $5560 \pm 250$, and then older remains were found there (Crane & Griffin 1962: 200), the Tehuacán Archaeological-Botanical Project was formed (Byers 1967; MacNeish 1967). Besides maize and squash, the Tehuacán Project found remains of beans, gourds, domesticated and wild avocado, black and white sapote, guava, hog plum, cotton, chili pepper, amaranth, maguey, and more.

Since Rensch (1966) had reconstructed proto-Otomanguean words for maize, beans, squash, chili, avocado, and cotton, and the Tehuacán Valley is located centrally in Otomanguean territory, Tehuacán seems a likely early Otomanguean center (Amador Hernández & Casasa García 1979). Glottochronological calculations, though now largely considered unreliable, had placed proto-Otomanguean contemporaneous with the Coxcatlán phase of the Tehuacán sequence (Hopkins 1984), and the Coxcatlán phase showed evidence for cultivation and greater sedentism (MacNeish 1967: 23). These findings led to the hypothesis that the location of the Otomanguean homeland, or Urheimat, was in or around the Tehuacán Valley. Kaufman (2015a: 53) finds that reconstructed proto-Otomanguean vocabulary is “compatible with a somewhat dry highland habitat after the domestication of some plants and before the rise of full-blown agriculture and village life,” which fits the Tehuacán profile. Since most of the same vocabulary is reconstructed for proto-Mayan (Kaufman 1976) and proto-Mixe-Zoquean (Campbell &
Kaufman 1976), the geographic component is essential to the hypothesis of the Tehuacán Valley as Otomanguean homeland. Not only is Tehuacán in Otomanguean territory, but it is occupied or surrounded by languages of several major subgroups of the family (Popolocan, Chinantec, and Mixtecan). Thus from the perspectives of linguistic “centre of gravity” (Sapir 1916: 79–82,) and a fewest moves model of language spread (Dyen 1956), Tehuacán remains a viable candidate for the Otomanguean homeland.

Winter et al. (1984: 68) point out that “the Tehuacán Valley is not the only center of early agriculture nor was it necessarily more important than other centers.” Evidence for domesticated maize and squash that predates Coxcatlán has since been found in the Guilá Naquitz Cave near Mitla, Oaxaca (Piperno & Flannery 2000) and near Iguala, Guerrero in the Central Balsas River Valley (Ranere et al. 2009; Piperno et al. 2009), where the wild ancestors of maize (teosinte) and squash occur. Therefore, while Tehuacán is still an important site, it is no longer considered to be the place where maize and squash domestication began.

Since Otomanguean is more diversified and presumably temporally deeper than Mayan, Mixe-Zoquean, and Totonacan, one may wonder if Otomanguean language speakers were the innovators of Mesoamerican agriculture. With Otomanguean’s great spread, in Mesoamerican terms at least, this would accord with the farming/language dispersal hypothesis (Bellwood 1991; Bellwood & Renfrew 2002; Diamond & Bellwood 2003). However, speakers of widely spread language groups are not always the innovators of the agricultural practices that may enable their spread (Nichols 1997: 375; Comrie 2002), and it is difficult, but worth trying, to correlate linguistic evidence with archaeological evidence for incipient agriculture. Recent “paleobiolinguistic” research has suggested that Otomanguean terms for maize, beans, squash, and chili pepper are some of the oldest reconstructable words for those crops (Brown et al. 2013a, 2013b, 2014a, 2014b), but Brown (2015a) later cautions that inferences based on those findings should be viewed as tentative because Otomanguean has not yet been demonstrated to
be a language family in any published thorough application of the comparative method (Brown 2015b).

Kaufman (2015a: 11-12) suggests possible homelands for the Otomanguean major subgroups, all of which fall within the wider reach of the Tehuacán tradition. His hypotheses are based on evidence from linguistic geography, his Otomanguean high-level subgrouping (Kaufman 1988), and language contact, either among Otomanguean groups or between Otomangueans and non-Otomangueans. In one case, a toponym provides some evidence: since the Mangues “were also known as Chorotegas (Nawa /cholol-te:ka-h/ ‘people from /cholol-la:n/’; /cholol-la:n/ is Cholula) it seems feasible to locate the Mang[u]ean homeland in the valley of Puebla, whose main center was Cholula” (Kaufman 2015a: 11).

While ethnonyms and toponyms may hold such clues for identifying the language of earlier inhabitants of an area (Vennemann 2003), toponymical evidence for Mesoamerican prehistory is complicated by a widespread and enduring practice of calquing, or loan translation. For example, the Spanish name of the town Tututepec, near the coast of Oaxaca, is from Nahua to:to:tl tepe:-k ‘bird hill-on’. In colonial Mixtec it was yucu-dzaa ‘mountain-(of)bird’ (Jiménez Moreno 1962: 98), and in Zenzontepec Chatino it is kē kinī ‘mountain (of)bird.’ Since the Nahuas were likely a late arrival to Mesoamerica (Fowler 1983: 245; Kaufman & Justeson 2010)—though that is still disputed (Hill 2001, 2012)—we might tentatively rule out Nahua as the original source of the name. Linguistic evidence suggests that Mixtecs expanded towards the coast from near San Juan Mixtepec (Bradley & Josserand 1982: 293, 297; Josserand et al 1984: 156), and the Mixtec Lord 8 Deer “Jaguar Claw” ruled the Coastal kingdom of Tututepec around 1000-1100 C.E. (Smith 1963; Spores 1993; Joyce et al. 2004). We might therefore tentatively rule out Mixtec and infer that the name originated in Chatino; some epigraphic (Urcid 1993) and physical anthropological (Christensen 1998) evidence suggests earlier occupation of the region by Chatinos. However, the inability to linguistically determine the
directionality of the calquing between Mixtec and Chatino, and the possibility that one or both of those groups might have calqued the name from yet another language group, leave some doubt about the name’s ultimate origin. Nevertheless, this type of converging evidence from different fields enables us to posit Chatino as the earliest detectable source. While probing ancient prehistory in toponyms presents some challenges, toponyms may provide more transparent evidence that bears on historical linguistic or ethnohistorical questions of the more recent past (Merrifield 1966; Doesburg & Swanton 2011).

Language contact offers a window onto prehistory, but relatively little such work exists for Otomanguean languages. Kaufman & Justeson (2010: 222) note that in “Mesoamerica, lexical borrowing among languages occurs at fairly low levels, so its occurrence reflects a serious amount of interaction,” which is what they claim for cases of borrowing from Mixe-Zoquean into Zapotecan and other Mesoamerican languages under Olmec influence (Kaufman & Justeson 2007: 200). A couple of other reported cases involve proto-Oto-Pamean-Chinantec words borrowed into proto-Uto-Aztecan (Hill 2012) and Mixtec borrowings in Chatino (Campbell 2013: 414).

If lexical borrowing is rare in Mesoamerica, perhaps grammatical influence is a more common outcome of language contact, as in the Vaupés region of the northwest Amazon (Sorensen 1967; Aikhenvald 2002; Epps 2006). Quite a few cases have been reported, both between Otomanguean groups and between Otomangueans and others. Kaufman (1988) proposes a Central Otomanguean contact area in which Popolocan borrowed several grammatical markers from Chiapanec-Mangue, presumably before the latter migrated south: *wi IMPERSONAL, *o plural subject, *tau OPTATIVE, *ha IMPERATIVE. Other reported cases include a Mixtec numeral classifier borrowed into Chatino (Sullivant 2012), Chatino phonological patterns transferred into Pochutec Nahua (Bartholomew 1980), noun classifiers borrowed from Chiapanec into Mayan languages (Hopkins 2012), and Mayan (Huastec)
influence on Otomian syntax (Kaufman 2015a). In an archaeological survey, Balkansky et al. (2000) conclude that “Monte Albán [Zapotec] and the Mixtec states arose together from an interacting nexus of pre-urban societies.” Such a proposal should impel linguists to search for parallel evidence from language contact. So far, not much evidence for contact between Zapotec and Mixtec has been found, though Kaufman (1988) identifies a pluralizer *kka that is found only in those two groups (see example (3)).

Another important resource for Otomanguean historical linguistics and prehistory are ancient Mesoamerican writing systems and iconography, some of which were used by Otomanguean groups. Of these, the ancient Zapotec (Caso 1928; Urcid 2001; Marcus 2003) and Mixtec writing systems (Caso 1965; Smith 1973) are the best known, but the Ñuiñe script (Moser 1977) of the Mixteca Baja is another.

Though not from the prehistoric era, another valuable source for Otomanguean historical linguistics is the documentary and descriptive work carried out by Spanish friars in the colonial period. These works, often of good analytical quality, give us text translations and snapshots of the grammar and lexicon of earlier forms of Otomanguean languages from as early as the 1550s. A few notable examples are Castro’s (1557) Matlatzinca vocabulary written in the margins of a copy of Molina’s Nahua dictionary; Córdova’s (1578a, 1578b) Antequera (Oaxaca City) Zapotec grammar and dictionary; and Reyes’ (1593) Mixtec grammar and Alvarado’s 1962 [1593] Mixtec vocabulary. Perhaps even more valuable, and certainly more culturally rich, are the numerous ethnohistorical records of Otomanguean languages written by their speakers, which exist from all periods after the Conquest (Terraciano 2001; Restall et al 2005; Oudijk 2008; Doesburg & Swanton 2011). For good examples of incorporation of early post-Conquest data into reconstructions, see Josserand et al. (1984) for Mixtec and Pérez (2007) for Matlatzinca-Tlahuica.
5 Proposals of external and long-distance relationship

Using lexicostatistics and glottochronology, Swadesh (1959; 1960) relates Otomanguean not only to the isolate Huave but also to other neighboring language groups, particularly Purépecha (Tarascan), Miskito (Misumalpan) and Chibchan. He proposed a particularly close relationship with Huave (Swadesh 1960, 1964a, 1964b). Rensch (1977b: 164) accepted the Huave idea, Longacre (1977: 122) was cautiously receptive, Kaufman (1988) was skeptical, and not much has been heard about it since.

Witkowski & Brown (1978) state that Otomanguean belongs to a “Mesoamerican phylum” with Mayan, Mixe-Zoquean, Huave, Totonacan, Lencan, and Tol (Jicaque), but the only linguistic evidence they refer to is their proposed connection between Mayan and Mixe-Zoquean (Brown & Witkowski 1979), which was hotly debated for a time (Campbell & Kaufman 1980, 1983; Witkowski & Brown 1981). Mora-Marín (2016: 128) recently took a deeper look into the possible Mayan and Mixe-Zoquean connection using the comparative method, concluding that it “can be supported, tentatively,” but he does not consider any possible Otomanguean link.

Greenberg (1987: 123) placed Otomanguean in a “Central Amerind” group with Uto-Aztecan and Kiowa Tanoan. This was in turn part of his proposed Amerind macro-phylum, along with all of the indigenous languages of the Americas except for the Na Dené and Eskimo-Aleut families. Although a few Americanist linguists were receptive to Greenberg’s results (Golla 1987; Hymes 1987), Greenberg’s methodology of mass comparison was flawed (Ringe 1992: 71–76), and his American languages classification has been heavily criticized (see e.g. Chafe 1987; Campbell 1988; 1997; Matisoff 1990; Rankin 1992). Crucially, Greenberg’s flawed classification of American languages should not be correlated with evidence about human prehistory from other disciplines, such as genetics (Bolnick et al. 2004), archaeology, or ethnohistory.
Since Sapir (1925a, 1925b) classified Subtiaba as Hokan, and Rensch later showed Subtiaba to be Otomanguean, it is worth considering whether Otomanguean and Hokan are ultimately related. While Hokan itself is not universally accepted (Campbell & Oltrogge 1980: 222; Poser 1995; Campbell 1997: 295; Mithun 1999: 304), Kaufman (2006b: 366, 2015b: 1, 2016a) does believe Hokan is a family and that Otomanguean is related to it. More work is needed to explore this possible connection, but its time depth will approach the limits of the comparative method.

6 Current trends and future directions

From the late 1940s through the 1960s, significant historical linguistic work on Otomanguean languages was carried out, leading to reconstructions of parts of six of Otomanguean’s eight major subgroups. This trajectory led Longacre (1964: 1016) to proclaim that the then forthcoming proto-Otomanguean reconstruction “will not be incomparable with the accomplishment of Indo-European scholarship.” While the diversity of Otomanguean may rival that of Indo-European, and Rensch’s proto-Otomanguean work was indeed a great accomplishment, the cavalry did not come, and few scholars aside from Kaufman have been working on historical linguistics at the higher levels of Otomanguean.

Several Otomanguean major subgroups have their internal subgrouping sketched out (Campbell, to appear), and we have seen additional important contributions to the reconstruction of some of the families and their subgroups, especially Mixtec (Josserand 1983), Zapotec (Fernández de Miranda 1995; Benton 1988; Kaufman 2016b), Chinantec (Rensch 1989), Trique (Matsukawa 2005), and Chatino (Campbell & Cruz 2010; Campbell 2013, In press). These and future works should provide the basis for advancing the historical linguistics of the higher levels of the family. What is currently needed is a coming together of experts on the various lower-level families to further support or revise the higher-level classification and reconstructions. Further interdisciplinary work involving linguists, ethnohistorians,
archaeologists, and geneticists is also needed to gain a clearer picture of Otomanguean and Mesoamerican prehistory.

Documentation and description of Otomanguean languages is rapidly expanding, as evidenced by the biannual *Coloquio de Lenguas Otomangues y Vecinas* and the *Taller de Tonos* and *Taller de Gramáticas Pedagógicas* at the Biblioteca Juan de Córdova in Oaxaca, supported by Mexico’s National Institute for Indigenous Languages (INALI) and the Harp Helú Foundation. But this new era of Otomanguean linguistics has an important and transformative difference from earlier ones: speakers of Otomanguean languages are now training as linguists in greater numbers (several are cited in this article), especially in the graduate program at CIESAS in Mexico. A new collaborative model with greater community participation and leadership is propelling the current growth of Otomanguean linguistics, which has the potential to significantly advance Otomanguean historical linguistics. Unfortunately, some Otomanguean language groupings, such as Ixcatec, Chocho, Cuicatec, Matlatzinca, Tlahuica, Chichimeco Jonaz, and Northern Pame remain sparsely documented or are facing serious endangerment, and the vitality of other Otomanguean languages is declining as a result of 500 years of colonialism, marginalization and now neo-liberalism and its continued exploitation. On the other hand, technological advances and linguistic expertise among community members and educators in communities and in academia offers some potential to stabilize languages, or at least better document them.

**References**


Ramírez Pérez, Elodia. 2014. La predicación no verbal y las construcciones copulares en el *tmu’u*.


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


