BRILL

LANGUAGE DYNAMICS AND CHANGE (2023) 1–52

brill.com/ldc

Morphological diffusion and the internal subgrouping of Central Totonac

David Beck | ORCID: 0000-0001-8422-7728
University of Alberta, Edmonton, AB, Canada
dbeck@ualberta.ca

Received 4 February 2023 | Accepted 25 September 2023 |
Published online 28 November 2023

Abstract

The Totonac branch of the Totonacan (also known as Totonac-Tepehua) family is traditionally broken down into four divisions—Misantla, Northern, Sierra, and Lowland. Misantla is an obvious outlier, but the relationship among the remaining three, which comprise the Central Totonac division, is uncertain due to competing lines of evidence: lexical isoglosses group Sierra and Lowland against Northern while morphological changes appear to set Sierra off against the other two. The spatial distribution of the morphological innovations shows these not to be a coherent set of changes inherited from a common ancestor, but instead a series of successive innovations diffused in a wave-like pattern. This paper also demonstrates that the morphological innovations are more recent than the lexical changes, supporting the prior separation of Sierra-Lowland languages from Northern. The paper also explores the methodological issues associated with the classification of languages in close contact at shallow time depths.

Keywords

Totonacan – internal reconstruction – morphological diffusion – historical linguistics – language contact

Traditionally, models of the internal structure of language families favor an approach based on the phylogenetic tree, which is structured into branches

defined by inherited synapomorphic innovations. Although this model is widely accepted and gives good results, at least at greater time depths, there are some well-known limitations to it that have led to the proposal of alternate views of linguistic phylogeny (Schmidt, 1872; Bailey, 1973; Wolfram and Schilling-Estes, 2003). A particular problem with strict branching models surfaces in cases where diffusion or horizontal transfer of lexical or other material occurs, which can lead to a situation where different lines of evidence appear to support different ramifications of a family tree. The problem is particularly acute when dealing with families at shallow time depths whose members are in close geographic proximity to one another. Such cases often force us to consider the potential diachronic limits to traditional cladistic taxonomies and the methods that can be used to tease apart the relationships between closely related varieties within language families.

This paper considers these problems in the context of Central Totonac (Brown et al., 2011; Levy and Beck, 2012), a proposed division within the Totonac branch of the Totonacan language family. The internal reconstruction of Totonac has been hampered by a number of factors—shallow time depth, geographic proximity, and the borrowing of morphology and lexicon—that confound traditional cladistic taxonomies. Consideration of different lines of evidence in isolation has led to competing hypotheses as to the first split in the Central group. Lexical evidence (Davletshin, 2008, 2018; Brown et al., 2011; Moore, 2017) points to an initial division between Northern and Sierra-Lowland languages, whereas morphological patterns seem to point to a grouping of Northern with Lowland against Sierra (MacKay and Trechsel, 2011, 2014, 2015). Closer examination of the spatial distribution of the morphological innovations identified in Sierra languages by MacKay and Trechsel, however, shows that, rather than being innovations inherited from a putative Proto-Sierra language, these traits are distributed in a wave-like pattern such that the full set of innovations is found only in the core Sierra geographic region. Languages on the periphery of this area share only some innovations, the number of changes correlating inversely with their geographic distance from the center. This suggests horizontal transfer of features rather than inheritance of synapomorphic changes, and the fact that some of the changes appear to be incomplete in some languages suggests that these innovations are recent, or still in progress. The absence of the Sierra morphological changes from Cerro Xinolatépetl, a language that shares most of the Sierra-Lowland isoglosses but which is separated from the Sierra geographic area by Nahuatl, shows that the spread of morphological innovations must have postdated both the development of the Sierra-Lowland lexicon and the separation of Cerro Xinolatépetl from the rest of the Sierra languages. This paper thus argues that

it is the split between Northern and Sierra-Lowland Totonac seen in the lexical data that demonstrates the older, phylogenetically relevant division in the Central branch, whereas the morphological features found in Sierra languages are more recent horizontal diffusions. In the absence of lexical data cleanly separating Sierra and Lowland languages, this paper also suggests that this traditional division in the family may not be a valid cladistic grouping.

1 The Totonacan language family

Totonacan (or Totonac-Tepehua) languages are spoken by approximately 270,000 people (Instituto Nacional de Lenguas Indígenas, 2021) in the northern parts of the Mexican states of Puebla and Veracruz, and in southeastern Hidalgo. Figure 1 shows the locations of all the languages discussed in this paper as well as the wider geographic range of Totonacan varieties listed in the *Catálogo de las lenguas indígenas nacionales* (Instituto Nacional de Lenguas Indígenas, 2008). Most of the languages in the family are in the process of slow extinction (Instituto Nacional de los Pueblos Indígenas, 2021). Totonacan is widely agreed to show a primary split between Tepehua, which consists of three languages, and Totonac (Arana Osnaya 1953; MacKay and Trechsel, 2018), a larger and less well-understood branch that has traditionally been held to consist of four major subgroups—Misantla, Northern, Sierra, and Lowland (McQuown, 1940, 1990; Watters, 1988; MacKay, 1999, 2011; MacKay and Trechsel, 2006, 2011, 2014, 2015, 2018).[1] Northern languages are located in the area around Xicotepec de Juárez extending to the north and west, and eastward along the Necaxa River; Sierra languages occupy the highlands east of Zacatlán and the adjacent sierra (henceforth, the Sierra Sur); and Lowland languages are found in the coastal region to the east of that, extending northward to Tuxpan. Misantla is a geographic outlier, being spoken further south in the central region of Veracruz.

The four subbranches of Totonac have typically been treated as coordinate branches under a single Totonac node, as shown in Fig. 2. This model has been widely disseminated by scholars, although its propagation seems owed either to an abundance of caution or, perhaps, to a certain "scholarly inertia" (Childs 2003: 47) which has led authors to reproduce this tree for convenience in publi-

1 Lowland Totonac is also known as Papantla Totonac and Sierra is sometimes referred to as Highland Totonac. The *Catálogo de las lenguas indígenas nacionales* (Instituto Nacional de Lenguas Indígenas, 2008) refers to Lowland Totonac as "totonaco de la costa" (Coast Totonac) and Sierra as "totonaco central alto" (Central Highland Totonac).

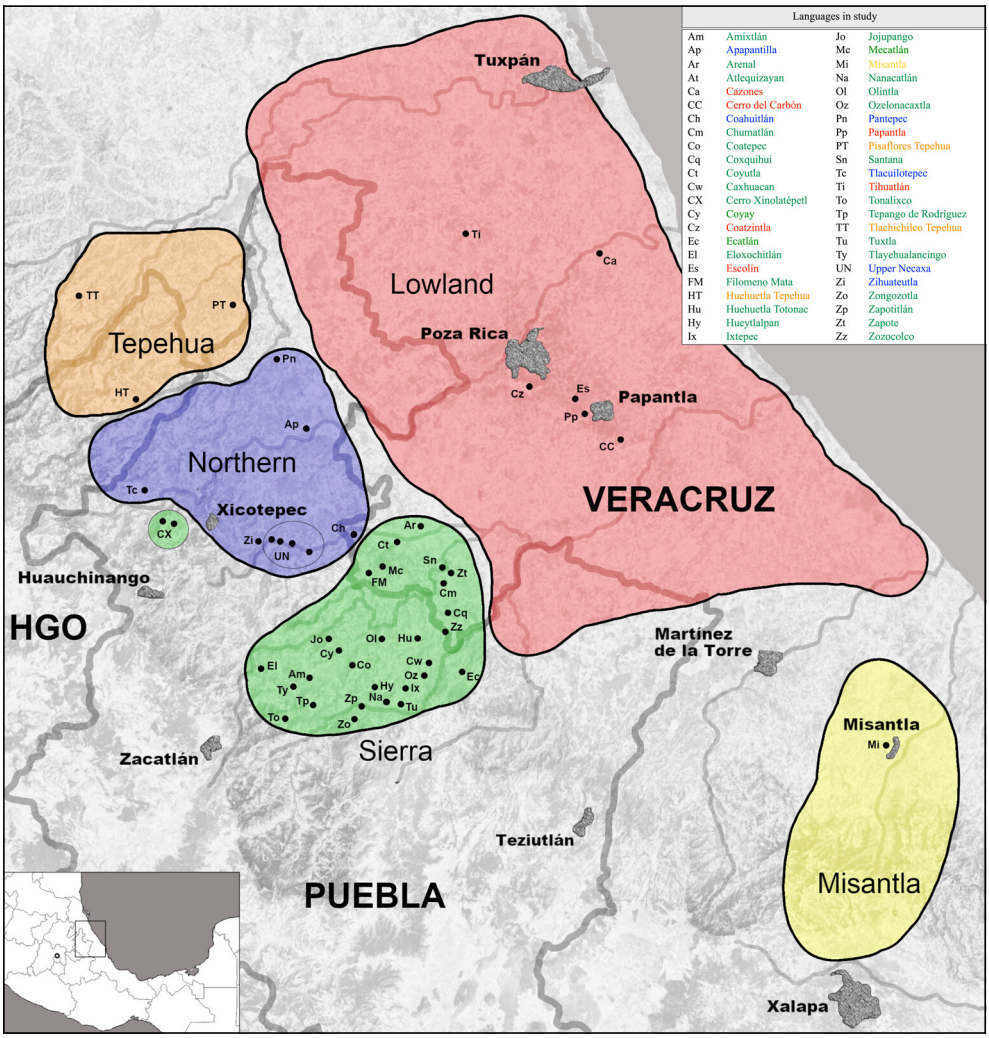

FIGURE 1 Totonacan languages and subgroupings

cations focusing on other aspects of the languages. Those researchers who take an explicit position on the internal relationship among the Totonac languages uniformly treat Misantla as a first branch in a more articulated tree (Ichon, 1969, citing H. Aschmann, personal communication; Hasler, 1993; Davletshin, 2008, 2018, 2019; Brown et al., 2011; Levy and Beck, 2012; Beck, to appear), as shown in Fig. 3, which groups together the remaining Totonac subgroups under the heading "Central Totonac" proposed in Brown et al. (2011) and Levy and Beck (2012). Most published claims advocating this configuration (with or without the term "Central Totonac") are merely informal statements reflecting researchers' intu-

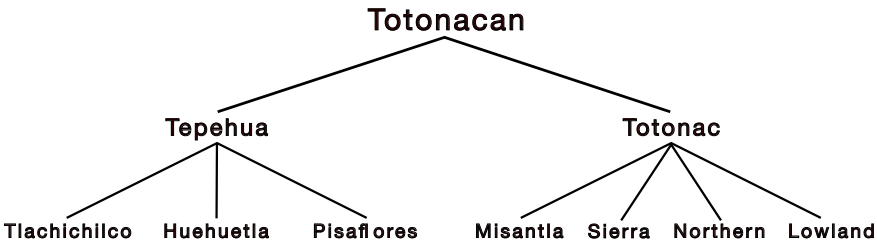

FIGURE 2 Traditional view of the Totonacan family

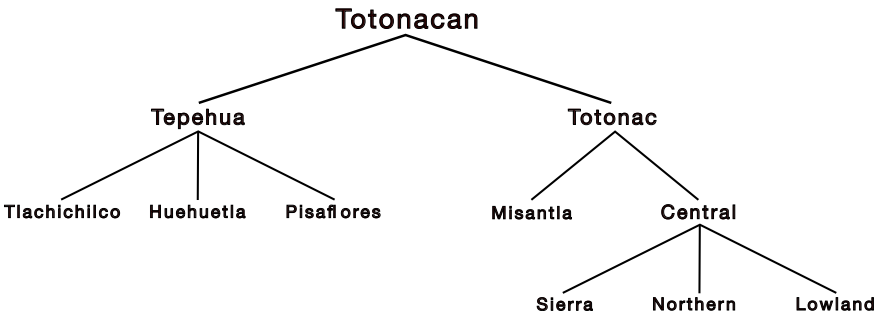

FIGURE 3 An alternative view of the Totonacan family

itions based on rather obvious morphological and lexical differences between Misantla and the Central group of Totonac languages. However, morphological patterns that set Misantla apart are documented in Beck (2012) and MacKay and Trechsel (2014, 2015), while Davletshin (2008, 2018, 2019, in press) uses lexicostatistical methods to show that the Misantla lexicon is the most distinct of the Totonac subgroups. Brown et al. (2011) offer cognate sets and reconstructions for 190 Proto-Totonacan forms and, based on these, propose a phylogenetic tree with a primary division within Totonac between Misantla and Central; this division is not explicitly argued for, but can be clearly seen by examining the cognate sets found in the paper.

Going beyond the tripartite division of Central Totonac seen in Fig. 3, there is a good deal of controversy about further ramification, and in the literature are found proposals for all three potential subgroupings—(i) Northern and Sierra against Lowland, (ii) Northern against Sierra and Lowland, and (iii) Sierra against Northern and Lowland. The first of these, treating Northern and Sierra together, was proposed in an Honors thesis by García Rojas (1978) based on a small dialectological survey, but has not received much subsequent support. The second alternative, grouping Sierra and Lowland Totonac against Northern, has more proponents (Ichon, 1969; Davletshin, 2008, 2018, 2019, in press; Brown et al., 2011; Levy and Beck, 2012; Moore, 2017) and is based on a

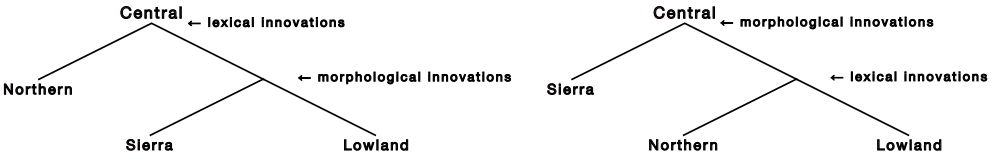

FIGURE 4 Competing hypotheses for the ramification of Central Totonac

well-known set of lexical isoglosses shared by Sierra and Lowland varieties but not found in Northern. The third option, placing Northern and Lowland languages together in opposition to Sierra, finds support in morphological patterns and changes to the person paradigms, identified by MacKay and Trechsel (2011, 2014, 2015), that are found in the highlands but not in the northern and coastal areas.

The lexical evidence for hypothesis (ii) and morphological evidence for (iii) both seem robust, leaving us with two conflicting views of the development of the Central branch of the Totonac group, shown in Fig. 4. Under the first hypothesis, the primary division between Northern and the other two branches would have been the result of lexical innovations, and the morphological traits distinguishing Sierra would have arisen later. The second hypothesis places the morphological innovations prior to the lexical changes, which would then have to be attributed to contact and borrowing between the now separate Lowland and Sierra languages. At first glance, the morphological changes might be presumed to be older, since morphology is often assumed to be more stable than the lexicon, and the kind of morphological innovation involved—the wholesale reorganization of person paradigms—seems like the type of fundamental change that might lead to "speciation" of linguistic varieties and the hard branching of phylogenetic trees. The sheer number of changes involved also gives the impression of being the cumulative product of innovation over a great deal of time. Nevertheless, closer examination of the agreement patterns found in a large sample of the languages spoken in the Sierra Sur shows that there is a great deal of variation in the way that these changes have played out. Rather than providing the set of clear and discrete morphological isoglosses that we might have hoped for, the spatial distribution of the changes in Sierra Totonac verbal paradigms reveals a wave-like diffusion of stepwise grammatical changes originating in the central Sierra Sur in or around Zapotitlán de Méndez and radiating outward. The absence of the morphological changes in one variety, Cerro Xinolatépetl, that shares many of the Sierra-Lowland lexical isoglosses favors the diachronic scenario where the lexical changes in Central Totonac precede the Sierra morphological innovations, suggesting that the correct subgrouping is that which opposes Northern against Sierra and Lowland.

TABLE 1 Principal published sources of data

Language	Data source
Apapantilla (Ap)	Reid et al. (1968); Reid and Bishop (1974); Reid (1991)
Cerro Xinolatépetl (CX)	Andersen (2012)
Coatepec (Co)	McQuown (1990)
Filomeno Mata (FM)	McFarland (2009)
Huehuetla Tepehua (HT)	Smythe Kung (2007)
Huehuetla Totonac (Hu)	Troiani (2004)
Misantla (Mi)	MacKay (1999)
Ozelonacaxtla (Oz)	Román Lobato (2008)
Papantla (Pp)	Aschmann (1973); Levy (1987, 1990)
Pisaflores Tepehua (PT)	MacKay and Trechsel (2010)
Tlachichilco Tepehua (TT)	Watters (1988)
Tlayehualancingo (Ty)	Espinoza Morales (1978)
Tuxtla (Tu)	A. Juárez Esteban (2016); T. Juárez Esteban (2020)
Upper Necaxa (UN)	Beck (2004, 2011a)
Zapotitlán (Zp)	Aschmann (1962)
Zihuateutla (Zi)	García-Vega (2022)

2 Lexical isoglosses

As noted above, most proposals for a first-level internal division within Central Totonac (CT) favor the grouping of Northern (N) against Sierra (S) and Lowland (L; Ichon, 1969; Davletshin, 2008, 2018, 2019; Brown et al., 2011; Levy and Beck, 2012). This view has generally been based on a small set of lexical isoglosses recognized by both speakers and linguists, the forms most frequently alluded to including words for WATER (N *ška:n, S-L *čʊ̨čut), HEART (N *li:stą́kną, S-L *nakúh), and MOON (N *ma:ɬkuyúʔ, S-L *papáʔ). Expanding on these, Moore (2017) identifies a set of 14 vocabulary items in seven languages that constitute isoglosses grouping Sierra and Lowland languages together against Northern varieties.

Since the publication of Moore (2017), a great deal of additional material has been collected or made public, allowing us to address the same issue based on a much larger set of data from a wider range of sources. Data for the current study has been drawn from the published sources listed in Table 1.[2]

2 In addition to the published sources listed in Table 1, Shoebox databases compiled for Zapotit-

TABLE 2 Languages covered by surveys

Survey	Languages covered
Kaufman et al. (2003–2005)	Amixtlán (Am), Arenal (Ar), Caxhuacan (Cw), Cazones (Ca), Chumatlán (Cm), Coahuitlán (Ch), Coatzintla (Cz), Coxquihui (Cq), Coyay (Cy), Coyutla (Ct), Eloxchitlán (El), Filomeno Mata (FM), Huehuetla Totonac (Hu), Hueytlalpan (Hy), Ixtepec (Ix), Jojupango (Jo), Mecatlán (Mc), Olintla (Ol), Ozelonacaxtla (Oz), Pantepec (Pn), Tepango de Rodríguez (Tp), Tihuatlán (Ti), Tlacuilotepec (Tc), Upper Necaxa (UN), Zapotitlán (Zp), Zihuateutla (Zi), Zongozotla (Zo), Zozocolco (Zz)
Additional applications of Kaufman et al.	Cerro Xinolatépetl (CX), collected by C. Snoek; Coahuitlán (Ch), collected by D. Moore; Zihuateutla (Zi), collected by M. García-Vega
Revision of Kaufman et al.	Atlequizayan (At), Ecatlán (Ec), Mecatlán (Mc), Nanacatlán (Na), Tepango de Rodríguez (Tp), Tonalixco (To), Tuxtla (Tu), Zongozotla (Zo), all collected by O. López Francisco; Santana (Sn), Zapote (Zt), both collected by F. Montes Castañeda

Data was also taken from a 683-item survey developed by Kaufman et al. (2003–2005) for the Project for the Documentation of the Languages of Mesoamerica (PDLMA). The original survey was applied in 29 communities, listed in the first row of Table 2.

The full PDLMA questionnaire was also re-implemented from scratch in three communities (second row of Table 2) in the mid-2010s by Devin Moore (Coahuitlán), Michelle García-Vega (Zihuateutla), and Conor Snoek (Cerro Xinolatépetl) as part of longer-term field documentation projects; see the second row of Table 3. In 2017, I created an abbreviated version of the questionnaire covering 405 lexical items (supplementary materials A) and some additional questions placing a more direct focus on key grammatical features. I

lán (Aschmann, n.d.a), Coyutla (Aschmann, n.d.b), Apapantilla (Reid et al., n.d.), and Tlachichilco Tepehua (Watters, n.d.), and lexical databases for Filomeno Mata (T. McFarland, personal communication) and Cerro Xinolatépetl (G. Andersen, personal communication) were also consulted. Emendations to the Shoebox database compiled by Aschmann for his Papantla dictionary (Aschmann, 1973) were also made available to me by Paulette Levy. "Papantla" in this context is a cover term for data gathered primarily in the communities of El Escolín and Cerro del Carbón in the Lowland language area.

piloted the abbreviated survey in Mecatlán, and then later hired Osbel López Francisco, a native speaker of Zongozotla, to implement the survey in other communities in the Sierra Sur. Two additional applications of the abbreviated survey were made by Faustino Montes Castañeda, a PhD student at the Centro de Investigaciones y Estudios Superiores en Antropología Social and a native speaker of Chumatlán, in El Zapote and Santana. Communities covered by the revised survey are given in the third row in Table 3. This expanded dataset puts us in a position to revisit the findings of Moore (2017) using forms for the 405 items on the abbreviated PDLMA survey from 45 Totonacan languages.

Examination of the expanded list of items confirms 11 of Moore's (2017) 14 lexical items distinguishing languages traditionally classified as Northern from Sierra-Lowland, and adds 25 more, giving us 36 items that can be used as lexical isoglosses.[3] In contrast, there are no items on the list of 405 that exclusively group Northern and Lowland against Sierra. The 36 isoglosses that unequivocally support the Sierra-Lowland vs. Northern division are shown in Table 3. For each item, Table 3 provides a tentative reconstruction of the contrasting Sierra-Lowland (second column) and Northern forms (third column) that constitute the isogloss, based on comparison of all the forms from each of the individual languages in each group.[4] The differences among the reflexes of the proposed reconstructions in each set are relatively minor and have to do for the most part with vowel quality and various types of regular syllabic reductions (e.g., Sierra-Lowland *sayi̱:n 'rain' is typically reduced to [se̱:n]). It is possible that an improved understanding of Proto-Central Totonac phonology would lead to some modifications of the reconstructions proposed in Table 3, but the differences between the Northern and Sierra-Lowland forms are stark enough that there can be no doubt as to which side of the isogloss a particular form belongs.

3 Three of the items proposed by Moore (2017)—LIVER, SAND, and GREEN—turn out not to be viable when the original set of seven languages is expanded to the current 45. In the case of LIVER, Moore proposes an isogloss based on the presence of the prefix *mak-* 'body' in Northern languages; however, several variants in the expanded set of Sierra languages also have a form with this prefix. The isogloss for SAND turns out to be confounded by the presence of two competing translations of the Spanish *arena* 'sand'—one for sand of the type found on a beach (*kukúh) and one for sand from a riverbed used for making concrete (*muntsáya). Many, perhaps most, languages have both, and it remains to be seen if the distribution of the two forms offered in single-word elicitation contexts is of interest for reconstruction. GREEN corresponds to at least three distinct etyma in Totonac, an issue complicated by the presence of more than one word that can be translated as Spanish *verde* in many Totonacan languages.

4 The transcription system used for the reconstructions in Table 3 is loosely phonemic, based on recordings of individual forms and what we know about common phonological processes in Totonac. Morphological segmentation of forms is not provided. The forms underlying the reconstructions in Table 3 are provided in supplementary materials B.

The final column in the table lists any exceptional distributions and cognates of either etymon found in Misantla Totonac (Mi) or in Tepehua (Tep). The absence of a Misantla or Tepehua word for a particular etymon in column 4 indicates that these data are not available.[5]

The data in Table 3 demonstrate the greater lexical similarity of Sierra and Lowland. An examination of the Misantla and Tepehua forms in the final column shows that the shared Sierra-Lowland vocabulary is likely to be innovative: in all but five cases where a cognate of one of the etyma in the table can be found in the data currently available for Misantla or Tepehua, the cognate corresponds to the Northern form (e.g., [34] WATER, S-L *čṵ́čut vs. N *škaːn, Tep *škaːn). Of the five cases where we see the other alignment, in three of them the Misantla and Sierra-Lowland forms are the same and the Northern form bears a greater resemblance to Tepehua ([9] DREAM, S-L *maːnišnan; N *laqawaːnán; Mi *maːnišnan*; HT *ʔalaqawan*, TT *aqlaqahun*; [17] ITCH, S-L *šḭn; N *pikšnḭ́; Mi *šin*; HT *taːpš*; [19] MOON, S-L *papáʔ; N *maːɬkuyúʔ; Mi *pap*; HT *maːɬkiyuʔ*, TT *maːɬkuyu*). In the other two, the Northern form is clearly an innovative departure from Proto-Totonac ([6] CLOTHES, S-L *ɬáqa̰ːt; N *lṵ́šṵ; Mi *ɬaqaːt*; HT *puːmpuʔ*, TT *láqčʼiːti*) or Proto-Totonacan ([18] MEAT, S-L *liːwá̰ːy; N *kiníːt; Mi *laːwa̰ː*; HT *ɬiːway*). There are four cases where Northern forms appear sporadically in Sierra-Lowland languages ([15] HEART, [24] RAINBOW, [27] SHIT_N, [28] SNOT), but only two cases where the opposite pattern holds, again suggesting that most of the Northern vocabulary is likely older and has been retained in individual languages in the Sierra or Lowland areas. The two exceptions to this pattern are found in the Northern language Tlacuilotepec and both involve slight shifts in meaning of pan-Totonacan lexicon ([25] RUN, S-L *qos, Tc *qos* 'run,' N **qos* 'fly'; [35] WEEP, S-L *ta̰sa, Tc *ta̰sa* 'weep,' N *ta̰sa 'vocalize').[6] Not all languages assigned to either group by this method have all the diagnostic forms—in a number of cases, individual languages or small sets of

5 It is worth noting that there are only 248 Misantla forms in my current database, as resources available for Misantla are rather limited. It may be that additional data will reveal more shared lexicon with Sierra-Lowland than we currently see in Table 3. Of the Misantla forms we do have, 52 (33%) either use roots not found in any other Totonac language or show extensive sound changes that differentiate them from the Central forms. This supports the findings of Davletshin (2008, 2018, 2019, in press) and Brown et al. (2011), as well as the intuitions of most Totonacanists, that the Misantla lexicon is the most divergent.

6 It is worth noting that the etymon of [35] WEEP recorded in the *Arte de la lengua totonaca* (1990)—an early colonial grammar of the Sierra variety Hueytlalpan, written in the sixteenth or early seventeenth century—is *kaɬwan*, the Northern form, rather than the *ta̰sa* form currently in use in the Sierra Sur today. Of the 13 concepts listed in Table 3 that can be found in the *Arte*, the other 12 all belong to the Sierra-Lowland lexical set.

TABLE 3 Lexical isoglosses grouping Sierra-Lowland against Northern

	Survey item	Sierra-Lowland	Northern	Notes
[1]	BASKET	*páqšu	*čq̣:štu̯	Mi *ta:nqe:*; HT *tanqalin*, TT *qáyč'i*
[2]	BELIEVE	*ka:nahla	*q̣qayi̯:	CX *yaqayi:*; Mi *qatq̣:yq̣:wa*; HT *laka?i:*
[3]	BIG†	*tɬánkq̣	*qq̣tɬq̣	Mi *papa:t*; Tep *qay
[4]	CHACHALACA	*ɬpáte:qa	*lákču	CX *lákča*
[5]	CHEST†	*kušmú:n	*ku̯šán	CX *ku̯šá:n*; Tep **tankiɬak-*
[6]	CLOTHES	*ɬáqq̣:t	*lu̯šu̯	Mi *ɬaqa:t*; HT *pu:mpu?*; TT *láqč'i:ti*
[7]	CLOUD	*púkɬni̯	*poqɬnú?	Mi *tapu:čuwanq?*; HT *?atapu:ts'i*, TT *taputs'i*
[8]	DANCE$_V$	*tantɬi:	*tɬi:	Mi *ti:*; Tep *t'i:
[9]	DREAM$_V$	*ma:nišnan	*laqawa:nán	Mi *ma:nišnan*; Tep *aqlaqawan
[10]	EAR†	*taqa:n	*q̣qašqǫ̯ɬ	Mi *qq̣qq̣šquɬ*; Tep **?aqašqoɬ
[11]	ELOTE‡	*tɬáqtɬq̣	*tsq̣́qtsq̣	FM *tsq̣́qtsq̣*; Mi *tsq̣:lakuš*; Tep *t'ákt'a
[12]	FIRE†	*ɬkuya:t	*makskút	Mi *mukskut*; HT *?alapnap*, TT *hikmi*
[13]	GIRL†	*tsu̯mq̣́:t	*tsu̯mq̣xá:t	FM *tsu̯muxq̣́:t*; HT *ti?*, TT *tsí?i*
[14]	GOOD†	*tɬa:n	*tsey	FM *tse:*; HT *qoš*, TT *o:ši*
[15]	HEART†	*nakúh	*li:stq̣́knq̣	Cq, Zo, Zt, Na *li:stq̣́knq̣*; Mi *la:kq̣tsi:n*; Tep *hałanuti
[16]	HOUSE POST	*tq̣layá:w	*tantú:n	CX *tantú:n*; HT *ta:ntu:n*, TT *ma:táw*
[17]	ITCH$_N$	*ši̯n	*pikšnį́	Mi *ši̯n*; HT *ta:pš*
[18]	MEAT	*li:wq̣́:y	*kiní:t	Mi *la:wq̣:*; HT *ɬi:way*
[19]	MOON	*papá?	*ma:ɬkuyú?	Mi *pap*; HT *ma:ɬkiyu?*, TT *ma:ɬkuyu*
[20]	NO	*ni:	*ła:	Mi *la*; Tep *ha:ntu
[21]	NOSE†	*kankan	*kínį́	CX *kínį́*; Mi *kį́?*; prefix CT *kinka-*, HT *kanqa-*
[22]	RAIN$_N$†	*sayį́:n	*ška:n	Mi *pušunu̯?*; Tep *ška:n; cf. [34]
[23]	RAIN$_V$†	*sayi̯:nan	*min ška:n	Northern lit. 'water comes'※
[24]	RAINBOW	*čq̣:makškúli̯t	*makčaškulinkq̣́t	Cz *makškúlit*; Ec *makškúli̯t*; Mi *ma:makšwali:ŋ*; HT *kiɬmakčata*, TT *škulumákča:ti*
[25]	RUN$_V$	*qos	*tu̯xnún	Tc *qos*; Mi *tsq̣:la*; HT *?ats'ala:*, TT *akyahuy*
[26]	SEE	*u̯kšiɬ	*laqtsín	FM *laqtsin*; Mi *lqqq̣n*; Tep *laqts'in
[27]	SHIT$_N$	*qq̣tsq̣́sq̣	*i̯ɬtį́n	Cz *i̯ɬti̯*; Mi *?i̯ɬti̯*; Tep *?iɬt'i
[28]	SNOT	*šánqat	*šilít	CX, FM, Ct, Mc *šili̯t*; Mi *šili̯t*; Tep *šiɬ

TABLE 3 Lexical isoglosses grouping Sierra-Lowland against Northern (*cont.*)

	Survey item	Sierra-Lowland	Northern	Notes
[29]	STAFF	*lí:štoqo	*pa:lák	Tep *pa:lik
[30]	TOMORROW†	*čą:lį́	*laqalí:	Mi *laqali:*; HT *ɬi?*, TT *ɬij*
[31]	URINE	*tsúlut	*sqolút	Mi *ʔiṣqɔt*; HT *tsulunti*
[32]	VILLAGE	*ką:čikį́:n	*ką:lakčíknį	CX *ką:lákčik*; Mi *la:tiwį́lan*; HT *laqačaqan*, TT *pu:táwlan*
[33]	WALL	*pá:tsaps	*tapá:k	HT *ta:ntu?*
[34]	WATER†	*čú̧čut	*ška:n	Mi *ška:n*; Tep *ška:n
[35]	WEEP	*tą̣sa	*kałwán	Tc *tą̣sa*; Mi *kałwan*; Tep *qałun
[36]	WHERE?	*ni:	*ła:	FM *ła:*; Mi *naɲčun*; HT *hunta:*, TT *tan*

† Identified in Moore (2017). ‡ Fresh corn on the cob. ⁜ Ca, Cz, Zt, FM use the same expression, combining [23] $RAIN_N$ *sayį́:n with *min* 'come.'

two or three languages make use of a third form to express the same concept (e.g., several languages use a Spanish borrowing, *anima* 'soul,' for [15] HEART). However, where the diagnostic form is not present in a particular language, the form used is neither the usual Sierra-Lowland nor the usual Northern form.

Based on this grouping method, the languages in this study can be classified as follows:

Northern: Apapantilla (Ap), Coahuitlán (Ch), Pantepec (Pn), Tlacuilotepec (Tc), Upper Necaxa (UN), Zihuateutla (Zi)

Sierra-Lowland: Amixtlán (Am), Arenal (Ar),[7] Atlequizayan (At), Caxhuacan (Cw), Cazones (Ca), Cerro Xinolatépetl (CX), Chumatlán (Cm), Coatepec (Co), Coatzintla (Cz), Coxquihui (Cq), Coyay (Cy), Coyutla (Ct), Ecatlán (Ec), Eloxchitlán (El), Filomeno Mata (FM), Huehuetla (Hu), Hueytlalpan (Hy), Ixtepec (Ix), Jojupango (Jo), Mecatlán (Mc), Nanacatlán (Na), Olintla (Ol), Ozelonacaxtla (Oz), Papantla (Pp), Santana (Sn),

7 Arenal is referred to as "Espinal" in Kaufman et al. (2003–2005), but this appears to be the name of the municipality rather than the community itself. Likewise, the survey conducted in the Upper Necaxa community of Patla is mislabeled "Jopala"—again, using the name of the municipality instead of the town.

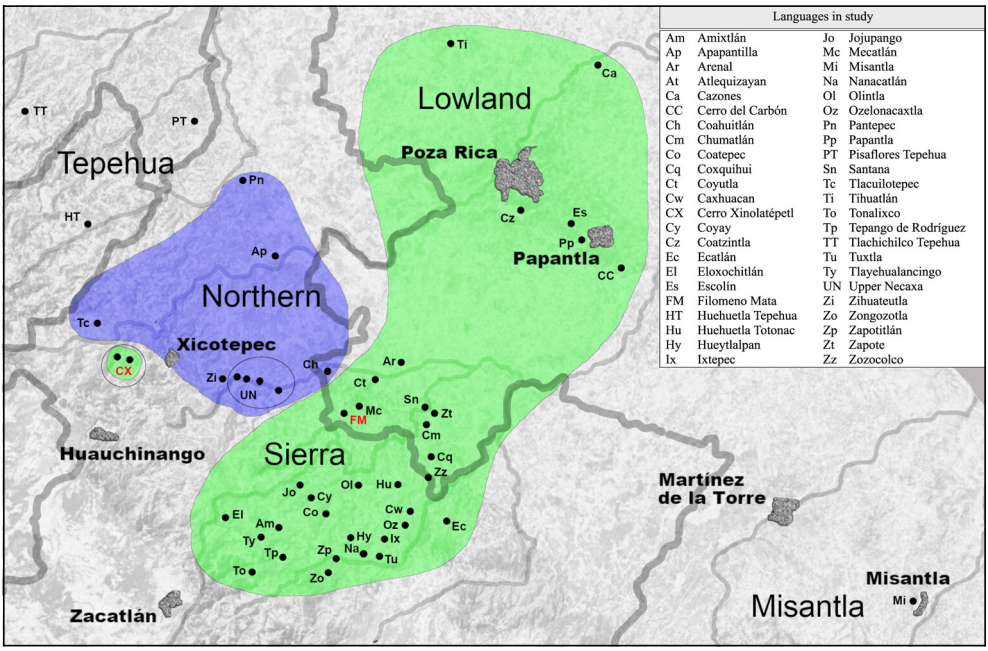

FIGURE 5 Northern vs. Sierra-Lowland isoglosses

Tepango de Rodríguez (Tp), Tihuatlán (Ti), Tlayehualancingo (Ty), Tonalixco (To), Tuxtla (Tu), Zapote (Zt), Zapotitlán (Zp), Zongozotla (Zo), Zozocolco (Zz)

The geographic distribution of these two groupings is shown in Fig. 5. This classification largely confirms earlier observations, but there are two things of note. The first is that none of the lexical sets associated with the concepts given in Table 3, or any other patterns found in the sets associated with the other 369 concepts compiled for this study, lend strong support for further differentiation between Sierra and Lowland languages, although there are some isoglosses that identify a "core" Sierra group centered on Zapotitlán de Méndez (Zp) that share forms not found in other Sierra or Lowland languages. It may be that further investigation will eventually turn up isoglosses that support a division between Sierra and Lowland, but at this stage the lexical evidence for this distinction is lacking.

The second finding from this comparison has to do with two languages in the dataset that resist easy classification—Filomeno Mata (FM) and Cerro Xinolatépetl (CX), shown in red in Fig. 5. These two languages have a substantial number of both Northern and Sierra-Lowland forms, as shown in Table 4.

TABLE 4 Diagnostic forms in Filomeno Mata and Cerro Xinolatépetl

	Sierra-Lowland form	Northern form	Other
Filomeno Mata	[1], [3], [4], [5], [6], [7], [8], [9], [10], [12], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [27], [29], [30], [31], [32], [33], [34], [35]	[11], [13], [14], [26], [27], [28], [36]	[2]
Cerro Xinolatépetl	[1], [3], [7], [9], [10], [11], [12], [14], [15], [17], [18], [19], [22], [23], [24], [26], [27], [29], [30], [31], [34], [35], [36]	[2], [4], [5], [16], [21], [28], [32]	[6], [8], [13], [20], [25], [33]

According to Table 4, Filomeno Mata has Sierra-Lowland forms for 29/36 concepts, Northern forms for 7/36, and uses other forms for 2/36 (FM has both Sierra-Lowland and Northern forms for [27] SHIT_{N}). These numbers are consistent with the conclusions of Moore (2017), who used a similar method with his smaller sample to tentatively classify Filomeno Mata as Sierra-Lowland. The preponderance of Sierra-Lowland over Northern forms in Filomeno Mata suggests that, on the balance of probabilities, what we see here is borrowing of lexicon from Northern into a Sierra-Lowland variety rather than the other way around. A similar case can be made for grouping Cerro Xinolatépetl with Sierra-Lowland: Table 4 shows Cerro Xinolatépetl to have 23/36 Sierra-Lowland forms as opposed to 7/36 Northern, and 6/36 independent forms.

It is also worth noting that of the 36 concepts listed in Table 3, a number belong to what is commonly referred to as "basic vocabulary": 12 of the concepts appear on the Swadesh 100 list (Swadesh, 1971) and nine on the Leipzig-Jakarta list (Tadmor et al., 2010). An examination of the diagnostic forms found in both Filomeno Mata and Cerro Xinolatépetl in Table 4 reveals that the preponderance of the basic vocabulary in both languages is drawn from Sierra-Lowland rather than Northern, as seen in Table 5.

Of the 12 concepts in Table 5 that appear on the Swadesh list, 11 of these are Sierra-Lowland forms in Filomeno Mata and only one is Northern. In Cerro Xinolatépetl, 10 are Sierra-Lowland forms, one is a Northern form, and one form is idiosyncratic. Making the same comparison with the more restricted Leipzig-Jakarta list, we see eight Sierra-Lowland forms vs. one Northern form in Filomeno Mata, and seven Sierra-Lowland vs. one Northern (plus one idiosyncratic form) in Cerro Xinolatépetl. If either set of basic vocabulary is, as is widely believed, more stable and resistant to borrowing, we would expect the larger group of basic vocabulary forms to be the inherited forms, which would add further support to including both Filomeno Mata and Cerro Xinolatépetl in Sierra-Lowland.

TABLE 5 Basic vocabulary in Filomeno Mata and Cerro Xinolatépetl

	Filomeno Mata		Cerro Xinolatépetl		
	S-L	**N**	**S-L**	**N**	**Other**
[3] BIG	✓		✓		
[7] CLOUD*	✓		✓		
[10] EAR	✓		✓		
[14] GOOD		✓	✓		
[15] HEART*	✓		✓		
[18] MEAT	✓		✓		
[19] MOON*	✓		✓		
[20] NO	✓				✓
[21] NOSE	✓			✓	
[23] RAIN	✓		✓		
[26] SEE	✓		✓		
[34] WATER	✓		✓		

*Not included in the Leipzig-Jakarta list (Tadmor et al., 2010)

3 Sierra morphological innovations

The most comprehensive analysis of morphological features that might be relevant to the classification of languages in the Totonacan family is found in MacKay and Trechsel (2014, 2015). These two papers systematize for the first time the differences in verbal inflectional patterns found across the family and show that these correlate to a large extent to the traditional Totonac family tree with a four-way split given in Fig. 2 above. Going a bit deeper, MacKay and Trechsel also write that the morphological evidence leads to the conclusion that "the languages comprising the Northern and Papantla branches of Totonac are less distinct than the traditionally 'flat' family tree might suggest" (MacKay and Trechsel, 2015: 154).[8] MacKay and Trechsel take the morphological evidence as cautious support for a Northern-Lowland grouping, although they acknowledge the conflicting lexical evidence:

8 MacKay and Trechsel use the term Papantla here for the group referred to as Lowland in this paper.

> Based on a preliminary look at a limited set of lexical items, we suggest that there are more significant lexical correspondences between the Papantla and Sierra languages than between the Papantla and Northern languages. Whether this is indicative of a deeper genetic affiliation between the Papantla and Sierra languages or simply the result of borrowing and diffusion has not yet been determined.
>
> MACKAY and TRECHSEL, 2015: 154

As we saw in the previous section, consideration of a much larger set of lexical items from a wider range of languages reinforces this observation.

Considering morphological features, MacKay and Trechsel (2015) state that the following six patterns are found uniquely among the Sierra languages:

a. Use of compositional forms in verbs with second person subjects and first person objects where one or both is plural (2(PL) > 1(PL)), as opposed to the use of a single syncretic form in Northern and Lowland (2015: Section 2.2.1)
b. Use of distinct forms in verbs with first person subjects and second person objects where one or both is plural (1(PL) > 2(PL)) in at least one Sierra language, as opposed to the use of a single syncretic form in Northern and Lowland (2015: Section 2.2.2)
c. Unambiguous expression of number of both subject and object in verb forms with third person subjects and first or second person objects (3 > 1/2; 2015: Section 2.2.3)
d. Unselective use of *-quu* 'PLURAL' to express the plurality of both subjects and objects in the third person (2015: Section 2.2.4)
e. Use of *ka:-* 'PLURAL OBJECT' with only first and second person objects in some Sierra languages, as opposed to its use with objects of all persons in Northern and Lowland (2015: Section 2.2.5)
f. The presence of the word-final *-y* 'IMPERFECTIVE' suffix in the first person singular and third person forms of vowel-final verb stems, as opposed to its absence in Northern and Lowland (2015: Section 2.2.6)

Of these, MacKay and Trechsel (2015: 138) assert that three can be considered diagnostic of the Sierra branch of Totonac languages: loss of the syncretic 2(PL) > 1(PL) form, (a) in the list above; the presence of *-quu* 'PLURAL' as a third person plural index, (d) in the list; and the presence of the final *-y* imperfective suffix, (f) in the list. The others appear exclusively in Sierra languages, but are not found in all of them.

In the remainder of this section, I will reexamine each of these features, beginning with (c), (d), and (e), which I will treat together in Section 3.1 as related changes stemming from the reanalysis of the Proto-Central Totonac

totalitive/terminative suffix *-qu̱ː as an index of argument plurality (Beck and Levy, 2019). Mapping out the geographic distribution of the reflexes of *-qu̱ː in the person paradigms of Sierra languages shows that its use with objects is much more limited than its use with subjects, meaning that the unselective use of *-qo̱ː* referenced in (d) is not a general feature of Sierra languages.[9] In Section 3.2, examination of the distribution of languages that have lost the syncretic 2(PL) > 1(PL) form in favor of compositional forms shows a different and significantly more limited distribution than we find for *-qo̱ː* as a plural subject index; the compositional forms turn out to be centered on the south-central region of the Sierra Sur, while languages on the periphery retain the syncretic form. Section 3.3 briefly illustrates the loss of the syncretic 1(PL) > 2(PL) form in Coatepec and, at least partially, in Zapotitlán de Méndez mentioned in MacKay and Trechsel (2015). Although the pattern is extremely limited, it does suggest a further progression in the overall regularization of grammatical paradigms represented by the reanalysis of *-qu̱ː and the loss of the syncretic 2(PL) > 1(PL) form. In Section 3.4, I discuss the imperfective *-y* suffix, and then in Section 3.5 I examine an additional feature found in most languages of the Sierrra Sur which is not mentioned in MacKay and Trechsel (2015)—the loss of the stative plural suffix *-nan*. Rather than defining a single, coherent set of languages sharing a common morphology, each individual feature considered here singles out a different set of languages in the Sierra Sur. Taken together, their spatial distribution suggests that the distinctive Sierra morphological patterns are the result of a wave-like diffusion of successive morphological innovations rather than a consistent and coherent set of patterns inherited from a common ancestor.

Grammatical forms discussed in this paper are drawn from the sources listed in Table 1 or extracted from example sentences found in the surveys listed in Table 2. The original version of the PDLMA survey (Kaufman et al., 2003–2005) was designed to elicit a broad range of Totonacan grammatical structures by asking consultants to translate Spanish words and sentences into their native language without context. This produced a rich dataset covering the key lexical items and morphological patterns needed for comparison, but the method of elicitation created certain obstacles. Because the interviewer was neither a

9 To avoid the circumlocution "reflexes of *-qu̱ː," I will refer to such elements simply as *-qo̱ː*, which is the form heard in the recordings available to me. MacKay and Trechsel write *-quu* because the [o] in most Sierra and Lowland languages is widely held to be an allophone of /u/ conditioned by proximity to the uvular stop (Levy, 1987). The form is consistently laryngealized in recordings from my surveys and from those collected by Kaufman et al. (2003–2005), as well as in all published sources in which laryngealization is included in transcriptions.

linguist nor a Totonac speaker, incorrect, anomalous, or off-task answers were not caught or queried. The result of this is that data for some forms in some languages is incomplete or uninterpretable because speakers either offered incorrect word-forms (e.g., a form with a 3PL object where a 2PL object was sought) or were inconsistent in the form used from one context to another (e.g., in the Caxhuacan survey the speaker offers forms with *-qo:* *-tit* as well as forms with *ka:-* in addition to *-qo:* *-tit* for 2PL > 3PL—see Section 3.1.2 below—at different points in the interview). There are fewer problems of this type with the data collected in later surveys, which were administered either by a trained Totonac speaker recording multiple repetitions and breaking between interview items, or by linguists studying the grammar over a period of years. Nevertheless, even in these surveys gaps remain. Data judged likely to be unreliable is excluded, which means that for some languages there is no information for certain morphological forms; however, there are not many gaps, and these are not in my view enough to seriously undermine the conclusions reached in this study.

3.1 *Reanalysis of* -qo: *as a 3PL argument index*

One of the morphological innovations identified by MacKay and Trechsel (2015) as being diagnostic of Sierra languages is the use of a reflex of the suffix *-qu: as an index of argument plurality for third persons, as seen in these examples from Coatepec:[10]

10 Examples are glossed following the Leipzig Glossing Rules. Abbreviations used: 1, 2, 3 first, second, third person; ALN alienative; ALTV allative applicative; ANT anterior; BEN benefactive applicative; CT Central Totonac; CLF numeral classifier; EXCL exclusive; FOC focus; FUT future tense; IDF indefinite voice; IMPF imperfective aspect; INCL inclusive; JUNCT phonological juncture; L Lowland Totonac; LOC locative; N Northern Totonac; NREL nonhuman relativizer; OBJ object; PAST past tense; PF perfect aspect; PFV perfective aspect; PL plural; PO possessive; PROG progressive aspect; QTV quotative; RCP reciprocal; S Sierra Totonac; SG singular; ST stative; SUB subject; TERM terminative; TOT totalitive. Abbreviations used for languages appear in alphabetical order in the legends for Figs. 1 and 5. When citing data from published sources, I have standardized transcriptions to an APA format using č, š, and y for IPA ʧ, ʃ, and j, respectively, to facilitate comparison of forms, though I have respected authors' other choices about phonemicization and orthographic representation, An acute accent represents word-level stress. Examples from McQuown (1971) are cited by text and line number, and were prepared by Paulette Levy based on McQuown's database, held at the University of Chicago. Uncited data are taken from my own field notes or the surveys listed in Table 2. Transcriptions of these data are my own. Transcriptions that I have made for languages that do not have phonological descriptions are loosely phonemic and rely on what is known in general about the phonology of the group to which that language belongs.

Coatepec

(1) *maqaːsaá niːquːníːt kintɬaːtkintsíʔ*
maqaːs≠aʔ *niː-**quː**-niːt* *kin-tɬaːt-kin-tsiːʔ*
long.ago≠now die-**3PL**-PF 1PO-father-1PO-mother
'My parents died long ago.' (McQuown, [1939–1968] 1971: 36.41)

(2) *wáː špaːstakqúː štɬaːtštsíʔ*
waʔ *š-paːstak-**quː**-∅* *š-tɬaːt-š-tsiːʔ*
that PAST-remember-**3PL**-IMPF 3PO-father-3PO-mother
'It was because he remembered his parents.' (McQuown, [1939–1968] 1971: 36.21)

In (1), *-qo̰ː* indexes the plurality of the subject, the parents, while (2) comes from a context where the subject, the person remembering, is a single individual and *-qo̰ː* indexes the plurality of the object.[11] Note the position of the suffix, immediately following the verb stem and preceding aspectual suffixes.

In both Northern and Lowland languages, subject and object plurality for third persons are expressed by different prefixes. Third person plural subjects are indexed by *ta-* '3PL.SUB':

Upper Necaxa

(3) *kin**ta**ma̰ʔwaʔó̰ːɬ i kinkṵ́šḭ škṵ́tḭ*
*kin-**ta**-ma̰ʔ-wa-ʔo̰ː-ɬ* *i* *kin-kṵ́šḭ* *škṵ́tḭ*
1OBJ-**3PL.SUB**-ALN-eat-TOT-PFV JUNCT 1PO-corn coatimundi
'The coatimundis ate all my corn.'

The prefix *ta-* expresses both the person and number of the subject. Number and person of objects are indexed separately, the prefix *kaː-* expressing the plurality of objects of all persons:

Upper Necaxa

(4) a. ***kinkaː**paškíː**n***
***kin-kaː**-paškiː-**n**-∅*
1OBJ-PL.OBJ-love-**2OBJ**-PFV
'S/he loved us.'

11 McQuown writes /-quː/ as he considers [o] an allophone of /u/. Coatepec as described by McQuown (1940, 1990) has lost laryngealization, although more recent research shows that the language today does have laryngealized vowels (Levy, in press). It is unclear whether this reflects a diachronic "regressive" change in Coatepec, some internal variation in the community, or an error on McQuown's part.

b. ***ka:****paški:****n***
ka:*-paški:-****n****-∅*
PL.OBJ-love-**2OBJ**-PFV
'S/he loved you guys.'

c. ***ka:****paški:****ɬ***
*∅-****ka:****-paški:-****ɬ***
3OBJ-PL.OBJ-love-PFV
'S/he loved them.' or 'They loved them.'

As we see here, *ka:-* expresses the plurality of objects in the first, second, and third person. The first person plural object form in (4a) combines the first and second person object affixes. The clusivity found with subjects is neutralized in this construction—that is, (4a) could mean 's/he loved me and some other(s), not including you.' Clusivity is also neutralized in possessive constructions. The ambiguity in the translation of (4c) stems from a restriction against the co-occurrence of the third person plural subject prefix *ta-* with *ka:-* when the latter pluralizes a third person object (see Section 3.1.3 below). In Upper Necaxa, speakers may choose either of the two affixes, depending on discourse context (so a possible form for 'they loved them' would be *tapaški:****ɬ***, which could also mean 'they loved him/her').

As MacKay and Trechsel (2015) observe, in many Sierra languages *ka:-* only pluralizes first and second person objects, while *-qo:* is used for third person objects:

Tuxtla
(5) a. ***kinka:****paški:****n***
kin-ka:*-paški:-****n****-∅*
1OBJ-PL.OBJ-love-**2OBJ**-PFV
'S/he loved us.'

b. ***ka:****paški:****n***
ka:*-paški:-****n****-∅*
PL.OBJ-love-**2OBJ**-PFV
'S/he loved you guys.'

c. *paški:****qó:ɬ***
*paški:-****qo:****-****ɬ***
love-**3PL**-PFV
'S/he/they loved him/her/them.' (T. Juárez Esteban, 2020: 30–31)

The first two forms in (5) are identical to the first two in (4). The third, however, bears *-qo̱:* (transcribed by Juárez Esteban without the laryngealized vowel, although laryngealization is audible in survey recordings from Tuxtla) instead of *ka:-*. This form is multiply ambiguous because of the "unselectivity" of *-qo̱:* noted by MacKay and Trechsel (2015: 140), meaning that *-qo̱:* can take scope over the subject, the object, or both. As we will see below, however, this unselectivity is not found in all the languages identified by MacKay and Trechsel as belonging to the Sierra group: while all their Sierra languages use *-qo̱:* to express the plurality of third person subjects, only a subset clustered in the south-central Sierra Sur use it to pluralize objects.

In Northern and Lowland, a cognate of *-qo̱:* is used as a totalitive or terminative marker. One frequent use is to express the total completion of an event or the exhaustivity of a process:[12]

Upper Necaxa

(6) *i̱kǫ̱ʔštę̱ʔwili:tǫ̱:kí:ɬ kimpu:kapéx, ǫ̱ʔpu:ti:ʔǫ́:ɬ*
i̱k-ǫ̱ʔštę̱ʔ-wili:-tǫ̱:ki:-ɬ kin-pu:kapéx ǫ̱ʔ-pu:ti:-ʔǫ̱:-ɬ
1SG.SUB-leave-put-flee-PFV 1PO-coffeepot head-evaporate-**TOT**-PFV
'I went off and left my coffeepot and [the coffee] all boiled off.'

The same affix can also be used to express affectedness of all the participants in an event:

Upper Necaxa

(7) *tatsukúɬ taní: ásta xǫ̱: tani:ʔǫ́:ɬ*
ta-tsuku-ɬ ta-ni: ásta xǫ̱: ta-ni:-ʔǫ̱:-ɬ
3PL.SUB-begin-PFV 3PL.SUB-die until where 3PL.SUB-die-**TOT**-PFV
'They started to die until they all died.'

(8) *ka:mǫ̱ʔni:ʔǫ́:ɬ tantu̱tún*
ka:-∅-mǫ̱ʔni:-ʔǫ̱:-ɬ tan-tu̱tún
PL.OBJ-3SG.SUB-kill-**TOT**-PFV CLF_{ANIMAL}-three
'He killed all three [pigs].'

As we see here, the totalitive can take scope over either the subject, as in (7), or the object, as in (8). Examples (7) and (8) come closest to the use of *-qo̱:* to index plurality that we saw in (1) and (2) above; however, in Northern and Lowland

12 In Upper Necaxa, the reflex of *-qu̱: is *-ʔǫ̱:*.

languages, verbs in sentences like (7) and (8) have genuine overt agreement affixes (*ta-* '3PL.SUB' and *ka:-* '3PL.OBJ'), and in neither case is the use of *-qo:* obligatory with plural arguments—it is freely chosen to express the meaning 'all.' The totalitive suffix occurs in exactly the same slot in the verbal template as does *-qo:* in Sierra. As argued in Beck and Levy (2019), it is a small shift in meaning from the use of *-qo:* to express 'all' to its use as an index of argument plurality.

Even though *-qo:* most frequently expresses argument plurality in Sierra languages, it is also found in contexts with a terminative meaning, the function seen in the Upper Necaxa example in (6) above, as shown by these examples from Coatepec:

Coatepec

(9) *lá: nakwaayin**qú:** naklaqán, mat nakiwanií tu: naktɬawá*
*la: na-k-wa:yin-**qu:**-ya: na-k-laq-ʔan-ya: mat*
how FUT-1SUB-eat-**TERM**-IMPF FUT-1SUB-ALTV-go-IMPF QTV
na-kin-wan-ni-ya: tu: na-k-tɬawa-ya:
FUT-1OBJ-say-BEN-IMPF NREL FUT-1SUB-do-IMPF
'As soon as I finish eating I'll go with him, he says he'll tell me what I'll do.' (McQuown, [1939–1968] 1971: 36.181)

(10) *nu:maá kintiwaníɬ kintɬá:t, škiwani**qu:**ni:taá ʔakšnií kčá:w kqu:yú:m*
nu:maá kin-ti-wan-ní-li kin-tɬá:t
thus 1OBJ-ANT-say-BEN-PFV 1PO-father
*š-kin-wan-ni-**qu:**-ni:t≠á ʔakxniʔ k-čá:n-w*
PAST-1OBJ-say-**TERM**-BEN-PF≠now when 1SUB-arrive-1PL.SUB
k-qu:yú:m
LOC-Huehuetla
'That's all my father told me, he'd finished telling me when we got to Huehuetla.' (McQuown, [1939–1968] 1971: 12.66)

In example (9), the subject of the verb *wa:yin* 'eat, dine' is singular and the verb itself is intransitive. Although it bears *-qo:*, there is no object to pluralize: *-qo:* can only mean 'finish' here. Likewise, in (10) both the subject and the object of the verb *waní* 'speak to' are singular, so again *-qo:* can only be the expression of the termination or completion of the event. The fact that *-qo:* in Northern and Lowland is a totalitive/terminative affix while in Sierra it serves both as a terminative and as a plural person index suggests that the latter use of *-qo:* is the more recent function, an extension of the totalitive use to the expression of argument plurality.

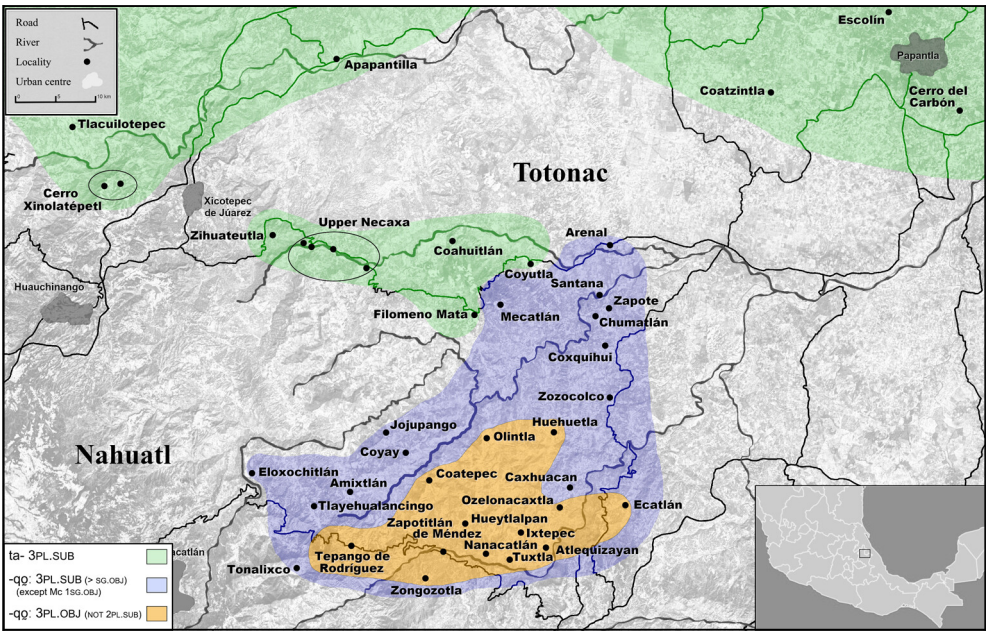

FIGURE 6 Distribution of *ta-* '3PL.SUB' and *-qǫ:* '3PL'

3.1.1 Distribution of *-qǫ:* as a 3PL index

Another indication of the recency of the use of *-qǫ:* as plural index for third persons can be seen in the spatial distribution of these uses across the Sierra Sur, shown in Fig. 6.[13] Figure 6 plots the distribution of the prefix *ta-* '3PL.SUB' typical of Northern and Lowland languages (and also found in Tepehua and Misantla) against the distribution of *-qǫ:* used as an index of plurality for (a) third person subjects of monovalent verbs and bivalent verbs with singular

13 The maps shown here and below are intended to give general ideas of the spatial distribution of features, which is represented in terms of contiguous geographical regions when the same feature is attested in communities that are located in close proximity. This is not to be taken as a claim that all undocumented communities within a region necessarily shown that feature, though it is considered probable that they would. Where there is no data but evidence of variation adjacent to a specific area, that area is left empty, as for instance in the region of lowland Veracruz between Arenal and Coatzintla, which would appear to be the transition zone between features of Sierra and Lowland languages. Three languages—Pantepec (Northern), Tihuatlán, and Cazones (both Lowland)—are not shown on most maps. These are all located farther to the north but share the features of the members of their respective groups as they appear on the maps. The mountainous area to the north of the Upper Necaxa communities and Coahuitlán is essentially unpopulated except for one or two small Spanish- or Nahuatl-speaking communities.

objects of any person, and (b) third person objects with subjects other than the second person plural.[14] The most striking feature of this map is that the set of languages that use *-qǫ:* for third person objects is a subset of the languages that use *-qǫ:* to index the plurality of third person subjects. Languages that use *-qǫ:* only for subjects occupy a large area extending at least as far north as Arenal and as far east as the line running north–south between Chumatlán and Ecatlán (shaded blue in Fig. 6);[15] however, those in which *-qǫ:* is used to index the plurality of objects as well are confined to a much smaller area in the south-central part of the region (shaded orange). Languages that use *ta-* '3PL.SUB' rather than *-qǫ:* to express the plurality of third person subjects and which retain the Central Totonac *ka:-* as an index of object plurality for third persons are shown in green.

The difference in agreement patterns between Sierra languages that use *-qǫ:* to pluralize third person objects (orange) and those that use *ka:-* (blue) can be seen in (11):

		Tuxtla (P. Levy, personal communication)	Eloxochitlán
(11)	a. 1SG>3PL	*k-pa:ški:-qǫ:-ɬ*	*k-**ka:**-ų̨kšiɬma:*
		'I loved **them**'	'I am seeing **them**'
	b. 2SG>3PL	*pa:ški:-qǫ:*	***ka:**-ų̨kšiɬpǫ̨:-t*
		'you$_{SG}$ loved **them**'	'you$_{SG}$ are seeing **them**'
	c. 3PL>1SG	*kim-pa:ški:-qǫ:-ɬ*	*ki-ų̨kšiɬma:-qǫ:*
		'**they** loved me'	'**they** are seeing me'
	d. 3PL>2SG	*pa:ški:-qǫ:-n*	*yų̨kšiɬma:-qǫ:-n*
		'**they** loved you$_{SG}$'	'**they** are seeing you$_{SG}$'
	e. 3PL>3SG	*pa:ški:-qǫ:*	*yų̨kšiɬma:-qǫ:*
		'**they** see him/her'	'**they** see him/her'
	f. 3PL>3PL	's/he/**they** see him/her/**them**'	***ka:**-yų̨kšiɬma:-qǫ:*
			'**they** see **them**'

The difference in object agreement is seen most clearly in (11a) and (11b), where Tuxtla uses *-qǫ:* and Eloxochitlán uses *ka:-* for third person plural objects. Both use *-qǫ:* consistently with third person plural subjects (11c–f). However, the

14 Third person objects with 2PL subjects show more varied patterns and are discussed in Section 3.1.2 below.

15 Although Mecatlán is included in this area, it uses *ta-* '3PL.SUB' rather than *-qǫ:* with 1SG objects: Mc *kin**ta**ų̨́kša* 'they see (*ų̨kš-*) me (*kin-*)'.

loss of *kaː-* for indexing the plurality of third person objects in Tuxtla means that Tuxtla has the same form for (11e) as for (11f), whereas Eloxochitlán has an unambiguous 3PL > 3PL form that uses *-qǫ̨*: for subject and *kaː-* for object plurality. This shows that the use of *-qǫ̨ː* with plural objects is a separate innovation from its use with subjects, and that the unselectivity of *-qǫ̨*: as a 3PL index remarked upon by MacKay and Trechsel (2015) is in fact limited to only a subset of languages—Huehuetla, Olintla, Coatepec, Ozelonacaxtla, Hueytlalpan, Zapotitlán, Ixtepec, Nanacatlán, and Tuxtla—in the central Sierra Sur. The more limited distribution of *-qǫ̨ː* with objects thus has the appearance of a more recent extension of its use with subjects which has yet to spread to other languages of the Sierra Sur.

3.1.2 Distribution of *-qǫ̨ː* in 2PL > 3PL forms

The idea that the use of *-qǫ̨ː* with objects is a more recent, and less advanced, change than its use with subjects is reinforced when we consider the use of *-qǫ̨ː* in forms with third person plural objects and second person plural subjects (2PL > 3PL). Here we find less consistency and more variation among the languages. Figure 7 shows the distribution of verb forms that use the following combinations of affixes:

(12)	a.	*kaː- -tit*	Upper Necaxa	***kaːlaʔtsinaːtít***	'you guys see them'
	b.	*-qǫ̨ː -tit*	Nanacatlán	***ǫ̨kšiɬqǫ̨ːyaːtít***	'you guys see them'
	c	*kaː- -qǫ̨ː -tit*	Zozocolco	***kaːǫ̨kšiɬqǫ̨ːyáːtit***	'you guys see them'
	d.	*∅- -tit*	Mecatlán	***ǫ̨kšiɬáːtit***	'you guys see them'

The pattern in (12a) is the Proto-Central pattern that uses *kaː-* as the plural object marker (green in Fig. 7). This pattern is maintained on the western periphery of Sierra (Tonalixco, Eloxochitlán, and Coyay) as well as in two languages on the northern edge (Huehuetla and Chumatlán). The pattern in (12b) replaces *kaː-* with *-qǫ̨ː* (blue) and, with the exception of Huehuetla, coincides with the languages that use *-qǫ̨ː* for plural objects with first and third person subjects that we saw in the previous section in Fig. 6.

There are also three communities—Caxhuacan, Zozocolco, and Coxquihui (in red)—that use the pattern in (12c) that combines both *kaː-* and *-qǫ̨ː* in a single verb-form. None of these three use *-qǫ̨ː* to express the plurality of third person objects with other persons of subject. Although Caxhuacan appears to be separated from the others on the map by Huehuetla, Huehuetla is in fact located on the other side of a mountain ridge; travel between Caxhuacan and the other two is more direct, suggesting greater potential for regular contact and mutual influence. These three communities, located to the east of the cen-

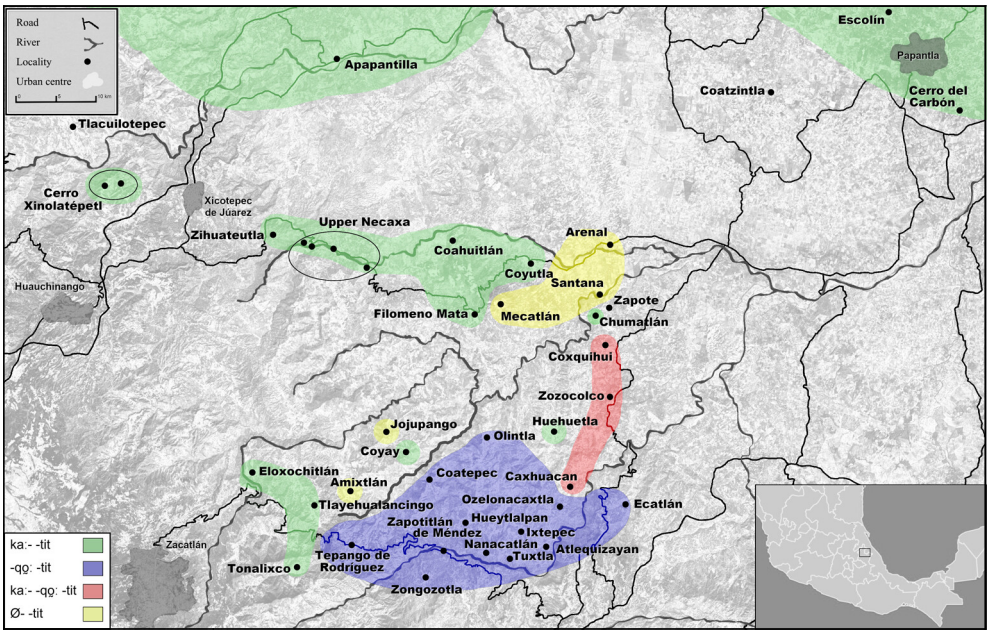

FIGURE 7 Distribution of 2PL > 3PL forms

tral Sierra Sur, would also be in more immediate contact with Lowland varieties that exclusively use *kaː-* to index object plurality.

The fourth form in (12d), which lacks object agreement altogether (so has an additional reading of 'you guys see him/her/it'), appears in the data from five communities (Amixtlán, Jojupango, Mecatlán, Santana, and Arenal—shown in yellow), none of which use *-qǫː* for the indexation of third person objects with other persons of subject. It is entirely possible, however, that in these languages the acontextual nature of the elicitation task led to the consultants simply not producing a plural object index of any kind.[16] In this case, these languages could be grouped with either those showing the pattern in (12a) or those that show the pattern in (12b).

3.1.3 Distribution of *-qǫː* in 3PL > 1/2PL forms

The variation we see above in 2PL > 3PL verb forms is suggestive of a change in progress in the Sierra Sur with languages running the full spectrum between

16 It should be noted that the consultant that I worked with in Mecatlán resisted suggestions that we might be able to add either *kaː-* or *-qǫː* to the form given in (12d) and retain the desired meaning. Whether the same results would have been obtained with a different speaker, or with the same speaker on a different day, remains to be seen.

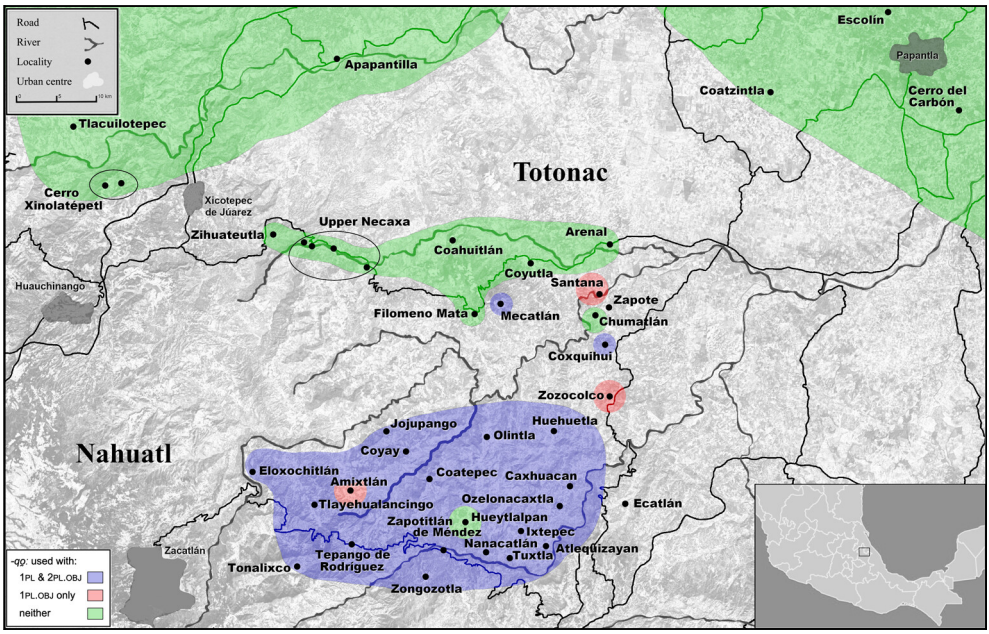

FIGURE 8 Distribution of *-qǫ:* '3PL' with 1/2PL.OBJ

those that fully conserve the Proto-Central forms with *ka:-* to those that eliminate *ka:-* entirely in favor of *-qǫ:* as an index of object plurality. We see a similar type of variation with respect to whether or not *-qǫ:* is used as a plural subject index with first or second person objects. As shown in Fig. 8, there are languages in the central part of the Sierra Sur that allow *-qǫ:* in all 3PL > 1/2PL forms (shown in blue), those that allow it only in 3PL > 1PL forms (shown in red), and those that disallow *-qǫ:* with both first and second person objects—these are shaded green in the map, as are the languages that retain the Proto-Central pattern and use *ka:-* to index object plurality in all persons.

The largest contiguous area among the Sierra languages is the blue region in which languages have compositional 3PL > 1/2PL forms like those in (13):

Nanacatlán

(13) a. ***kinka:ɻukšiɬqǫ:yá:n***
kin-ka:-ɻukšiɬ-qǫ:-ya:-n
1OBJ-PL.OBJ-see-**3PL.SUB**-IMPF-**2OBJ**
'They saw us.'

b. ***ka:ɻukšiɬqo:yá:n***
ka:-ɻukšiɬ-qo:-ya:-n
PL.OBJ-see-**3PL.SUB**-IMPF-**2OBJ**
'They saw you guys.'

These forms have the expected combination of affixes for first or second person plural objects and use *-qo:* to index the plurality of a third person subject, as do the examples from Coatepec in (1) and (2) above. The process of shifting *-qo:* to an all-purpose third person subject suffix would thus seem to be complete in these languages, which occupy a contiguous area only as far north as Huehuetla, with outliers in Mecatlán and Coxquihui.

In Hueytlalpan and Chumatlán (the isolated green areas on the map), on the other hand, no forms with *-qo:* '3PL.SUB' and first or second person objects are attested, though the two languages differ in their strategies for expressing 3PL > 1/2PL. Chumatlán simply leaves the plurality of the subject unexpressed, whereas the speaker in Hueytlalpan made use of the indefinite agent suffix *-kan*, which suppresses the expression of the subject altogether. The latter could be an artifact of bilingual elicitation, since *-kan* is often used to translate Spanish sentences with elided 3PL subjects such as those used as the prompts in the PDLMA survey. In Amixtlán, Santana, and Zozocolco (the isolated red areas), *-qo:* is used only with 1PL objects. However, the data from Amixtlán is inconsistent: one out of two forms with a 3PL subject and a 1PL object in the prompt has *-qo:*, while the other form shows no indexing of subject plurality at all. In Santana, three out of four 3PL > 1PL forms have *-qo:* in them and no 3PL > 2PL forms do. There are also two languages left ungrouped in Fig. 8—Ecatlán and Zapote. Both have forms with 1PL objects with *-qo:*, but the 2PL object forms obtained in the surveys were considered unreliable and were excluded, making it unclear if these languages pattern with Amixtlán, Santana, and Zozocolco, or with the group that uses *-qo:* with both 1PL and 2PL objects.

It is possible, of course, that the variability in the data here means that consultants supplied the incorrect form at least some of the time, or that the indexation of subject plurality is dependent in part on discourse context. This could also mean that some of the languages marked on the map as not having *-qo:* with first and second person objects do in fact have the option of using it, though we know this not to be true either of Northern and Lowland languages, or of Filomeno Mata and Cerro Xinolatépetl (all languages that do not use *-qo:* for third person subjects at all). Clearly, there is room for more careful study of these patterns. Nevertheless, the data in Fig. 8 concur with what is shown on the other maps, that the role of *-qo:* in various parts of

the person paradigm is not the same for all of the Sierra languages, and that distribution of *-qǫ:* forms shows a pattern of decreasing integration into the person-marking system as we move further away from the south-central Sierra Sur.

The incomplete distribution of compositional 3PL > 1/2PL forms is also discussed by MacKay and Trechsel (2015: 139–140) as part of their more general examination of 3SUB > 1/2PL forms. These authors use this part of the Sierra paradigm to draw some interesting comparisons between Sierra languages like Ozelonacaxtla that have nonambiguous 3SUB > 1/2PL forms and languages from outside the Sierra Sur where the corresponding forms are ambiguous—specifically, Papantla and Filomeno Mata. Consider the examples in (14). The first three are drawn from MacKay and Trechsel (2015), to which I have added an ambiguous Cerro Xinolatépetl form and a nonambiguous Upper Necaxa form:

(14)	Ozelonacaxtla (Sierra)	***ká:-lkapá:sták-qǫ́:-yá:-n***
		'They remember you guys.' (p. 139)
	Papantla (Lowland)	***ká:-akšił-ni***
		'S/he/they saw you guys.' (p. 146)
	Filomeno Mata	***ká:-láqtsį-n***
		'S/he/they saw you guys.' (p. 146)
	Cerro Xinolatépetl	***ka:-akšił-n***
		'S/he/they saw you guys.'
	Upper Necaxa (Northern)	***ka:-ta-lą?tsí-n***
		'They saw you guys.'

All four forms use the combination of *ka:-* 'PL.OBJ' and *-n* '2OBJ' to express the 2PL object; the unambiguous Ozelonacaxtla form uses *-qǫ:* to index the 3PL subject; the Upper Necaxa form is also unambiguous, though it achieves this by using *ta-* '3PL.SUB.' In Filomeno Mata, Papantla, and Cerro Xinolatépetl, which use *ta-* to index 3PL subjects in other forms, there is a prohibition against the co-occurrence of *ta-* and *ka:-*, so the former is excluded in (14), resulting in an ambiguous verb-form expressing the number of the object but not of the subject.

Because MacKay and Trechsel (2015: 146, n. 15) classify Filomeno Mata as a Northern language and treat Upper Necaxa (and Apapantilla, which also has *ka:- ta-* in this form) as innovative members of the Northern group, they posit that unambiguous 3SUB > 2PL forms are an exclusive feature of Sierra, whereas ambiguous forms are a shared feature of Northern and Lowland. However, examination of the data from additional Northern languages in the sample—

Coahuitlán, Pantepec, Tlacuilotepec, and Zihuateutla—show that unambiguous *kaː- ta- -n* forms are found in these languages as well. This either means that Filomeno Mata is an outlier in Northern, or it supports the conclusion reached in Section 2, based on lexical isoglosses, that Filomeno Mata belongs with Sierra-Lowland. The data in (14) show that Cerro Xinolatépetl also shares the Lowland restriction against the co-occurrence of *ta-* and *kaː-*. As with Filomeno Mata, excluding Cerro Xinolatépetl from Northern would meant that the presence of nonambiguous 3SUB > 1/2PL forms is typical of both Northern and Sierra, although this is true for different reasons. In Northern, unambiguous forms occur because these languages tolerate the combination of *ta-* and *kaː-* with first and second person plural objects. The emergence of nonambiguous forms in Sierra, on the other hand, is a natural consequence of the use of *-qo̱ː* as an index of 3PL subjects, given that adding *-qo̱ː* to an ambiguous 3SUB > 1/2PL form (Pp ***kaːakšiɬni*** 's/he/they saw you guys') creates the nonambiguous 3PL > 1/2PL forms seen in the south-central part of the Sierra Sur (Na ***kaːukšiɬqoːyaːn*** 'they see you guys').

3.2 *2(PL) > 1(PL) forms*

Another pattern that MacKay and Trechsel (2015) identify as being unique to Sierra languages is the loss of a syncretic form found in all other branches of the Totonacan family (including Tepehua and Misantla). Outside the Sierra Sur, Totonacan languages express all situations in which a second person acts on a first person and one or both is plural (2(PL) > 1(PL)) using a single verb-form. The pattern is illustrated by the following data from the Northern language Upper Necaxa, where we see a single form with three interpretations:

Upper Necaxa

(15) ***kilaːpaːsto̱ká·w***
kin-laː-paːsto̱k-yaː-w
1OBJ-RCP-remember-IMPF-**1PL.SUB**
(i) 'You$_{SG}$ remember us.' 2SG > 1PL
(ii) 'You$_{PL}$ remember us.' 2PL > 1PL
(iii) 'You$_{PL}$ remember me.' 2PL > 1SG

The form in (15) makes use of the first person object prefix *kin-*, the reciprocal prefix *laː-*, and the first person plural subject suffix *-w* to express each of the three separate translations. In contrast, languages of the Sierra group like Huehuetla Totonac use compositional forms for each of the three meanings:

Huehuetla

(16) a. ***kinka****tsúka?*
kin-ka*-tsuk-a-?*
1OBJ-PL.OBJ-embrace-**IMPF-2SG.SUB**
'You$_{SG}$ embrace us.'

b. ***kin****tsukátit*
kin*-tsuk-a-****tit***
1OBJ-embrace-**IMPF-2PL.SUB**
'You$_{PL}$ embrace me.'

c. ***kinka****tsukátit*
kin-ka*-tsuk-a-****tit***
1OBJ-PL.OBJ-embrace-**IMPF-2PL.SUB**
'You$_{PL}$ embrace us.' (Troiani, 2004: 70–71)

Each of the Huehuetla forms uses regular expressions of the second person subject (*-tit* in the plural and—for this language with a verb of this conjugation—a final glottal stop in the singular) and the first person object (*kin-*); the plurality of the object in (16a) and (16c) is expressed by *ka:-*.[17]

As with the other features we have examined, the geographic distribution of 2(PL) > 1(PL) forms (Fig. 9) proves to be more complex than originally thought. On the map we can see that languages that preserve the Proto-Totonacan *kin- la:- -w* pattern in (15) above (in green) form a periphery around the group that has replaced it with compositional forms (in blue), this group being somewhat coterminous with, but smaller than, the group that uses *-qǫ:* as a 3PL.SUB suffix shown in Fig. 6. The conservative languages that use *-qǫ:* and maintain *kin- la:- -w* include the languages north of Chumatlán (Mecatlán, Santana, and Arenal) and those on the western edge of the group (Eloxochitlán, Amixtlán, Tonalixco, and Jojupango); also in this western area are two communities (in yellow), Tlayehualancingo and Coyay, where the syncretic form is not used for all three of the meanings shown in (14).

Also outside the group that uses compositional forms are Tepango de Rodríguez and Zongozotla (in red), which share the novel 2(PL) > 1(PL) form in (17):[18]

17 Troiani (2004) does not indicate vowel duration in her transcriptions.

18 The form corresponding to translation (17i) was accidentally left out of the 2017 survey for the verb *ɻukšił* 'see' but the same inflectional pattern was recorded for *waní* 'speak to.' I later confirmed the 'see' form with Osbel López Francisco and his father, and recorded it in my field notes as given here.

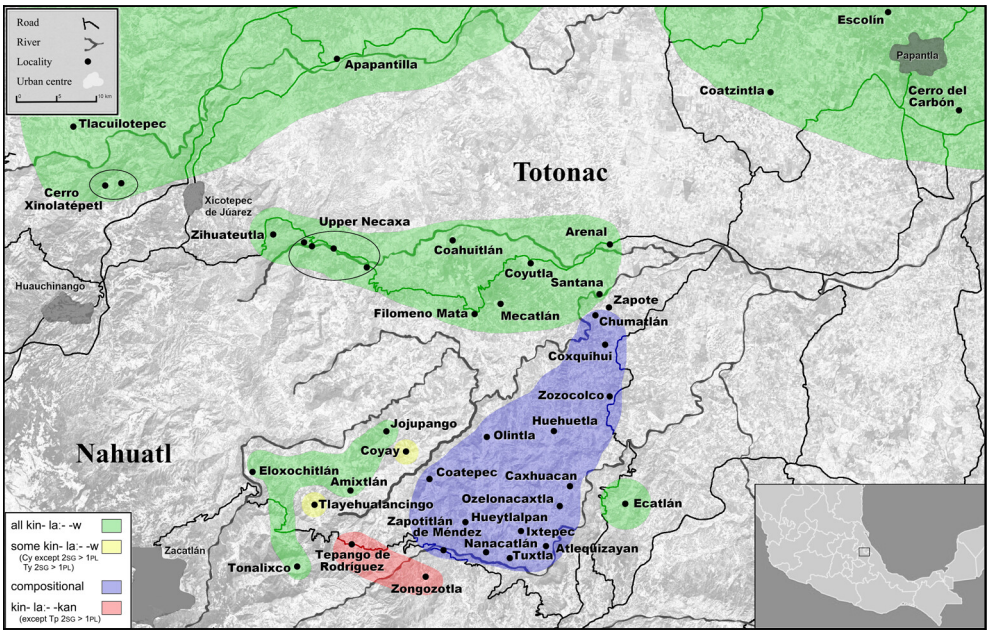

FIGURE 9 Distribution of 2(PL) > 1(PL) forms

Zongozotla

(17) ***kiła:ɻukšiɬma:kᵩ***
kin-ła:-ɻukšiɬ-ma:-kᵩ
1OBJ-RCP-see-PROG-**IDF:PFV**

(i) 'You$_{SG}$ are seeing us.'
(ii) 'You$_{PL}$ are seeing us.'
(iii) 'You$_{PL}$ are seeing me.'

Example (17) closely resembles the conservative *kin- la:- -w* form, but replaces the semantically anomalous first person plural suffix *-w* with the indefinite agent suffix *-kᵩ* (the perfective form of *-kan*). The *-kan* form is used in Zongozotla with all three readings given in (17), but in Tepango de Rodríguez it is attested only with the readings in (ii) and (iii), while the 2SG > 1PL form of the verb *wanɻ́* 'say something to' is *kinka:wánᵩ*, which is analogous to the compositional Huehuetla form in (16a) above. It is worth noting that the *-kan* form is not attested at all in the data collected in Tepango de Rodríguez by Kaufman et al. (2003–2005), who found the compositional patterns for all three of the readings in (17). Their Zongozotla survey produced inconsistent results, with instances of both *-kan* and compositional forms for 2SG > 1PL, compositional forms for 2PL > 1SG, and a *-kan* form for 2PL > 1PL. It is possible that both patterns are present in these communities.

3.3 *1(PL) > 2(PL) forms*

Another common syncretism found in Totonacan languages occurs in forms with first person subjects acting on second person objects, where one or both is plural (1(PL) > 2(PL)). In Tepehua, Misantla, and most of the Central Totonac languages, a single verb-form is used for all situations of this type, although the details of the form in the various branches of the family are slightly different. The Central pattern is illustrated by the Upper Necaxa forms in (18):

Upper Necaxa

(18) ***i̱ka:****pa:ški:yá:****n***
i̱k-ka:*-pa:ški:-ya:-****n***
1SG.SUB-PL.OBJ-remember-IMPF-**2OBJ**
(i) 'I love you$_{PL}$.' 1SG > 2PL
(ii) 'We love you$_{SG}$.' 1PL > 2SG
(iii) 'We love you$_{PL}$.' 1PL > 2PL

The form in (18) is expected to have only the translation in (18i), but it is also expresses (18ii) and (18iii). The expected forms for translations (18ii) and (18iii) would normally include *-w* '1PL.SUB,' which seems to be blocked by the presence of the second person object suffix *-n*.

In most Sierra languages, the form used for 1(PL) > 2(PL) also follows the syncretic pattern seen in (18). However, MacKay and Trechsel (2015) observe that the Sierra language Coatepec has replaced the single verb-form with three separate, unambiguous forms:

Coatepec

(19) a. ***ka:****paški:yá:****n***
k-ka:*-paški:-ya:-****n***
1SG.SUB-PL.OBJ-love-IMPF-**2OBJ**
'I love you guys.' (McQuown, 1990: 166)

b. ***ka:****paški:yá:****w***
k-ka:*-paški:-ya:-****w***
1SG.SUB-PL.OBJ-love-IMPF-**1PL.SUB**
'We love you$_{SG}$.' (McQuown, 1990: 164)

c. ***ka:****paški:****qu:****yá:****w***
k-ka:*-paški:-****qu:****-ya:-****w***
1SG.SUB-PL.OBJ-love-3PL-IMPF-**1PL.SUB**
'We love you$_{PL}$.' (McQuown, 1990: 166)

Of these forms, only the expression in (19a), which is formally identical to (18), is compositional. The forms in both (19b) and (19c) are missing the second person object suffix *-n*, which appears to be in competition for the final slot in the verb template with the 1PL subject suffix *-w*. We saw the same competition in (18) for readings (18ii) and (18iii), though there it is resolved differently. In (19b), the prefix *kaː-* seems to stand in for the second person object suffix rather than pluralizing the object (which is singular here). In (19c) the suffix *-quː* expresses object plurality, albeit (uniquely to this form) for a second person. This very limited and unique use of *-qo̱ː* represents another step in the expansion of the suffix into the object-person paradigms, and it seems consistent with the overall trend we see in the changes in Sierra person paradigms away from ambiguous and syncretic forms to more transparent inflectional patterns.

Although the loss of the syncretic 1(PL) > 2(PL) form is not well attested in any other language, the innovative form in (19b) is found in Aschmann's (n.d.a) lexical database for the nearby town of Zapotitlán de Méndez, compiled in the 1950s and 1960s:

Zapotitlán de Méndez

(20) *šmaːn wa̱ wi̱š miniːni̱y na**kkaː**ka̱ːkniːnaniyaː**w** …*
šmaːn wa̱ wi̱š miniːni-y
just FOC you deserve-IMPF
*na-**k-kaː**-ka̱ːkniːnan-ni-yaː-**w***
FUT-**1SG.SUB-PL.OBJ**-respect-BEN-IMPF-**1PL.SUB**
'Only you deserve that we worship you …' (Aschmann, n.d.a: line 7984, entry for *šliːƛanka̱*)

Likely because this combination of persons is textually infrequent, there are only two examples in the database (the other is in the entry for *a̱ːmaqtum*, line 130). On the other hand, the form in (19c) is not attested in Aschmann's database—though neither are sentences with the 1PL.SUB > 2PL.OBJ combination in the translations that would correspond to it. However, the syncretic form corresponding to (18) is given in the database entry *kkaːwaniyaːn* (line 1400) with all three possible 1(PL) > 2(PL) translations—'I say it to you guys/we say it to you$_{SG}$/we say it to you guys.' In the PDLMA survey for Zapotitlán, conducted some 60 years later, only the syncretic form is seen. This mixed data suggests that both the older syncretic forms and at least one of the innovative Coatepec-type forms may coexist (or may have coexisted) in Zapotitlán. It also suggests that the regularization of the object-person paradigm seen in the Coatepec data in (19) might, at one point, have been in the process of

spreading to Zapotitlán. If so, this would be consistent with the other examples of wave-like diffusion of morphological innovations creating more compositional (or at least more transparent) forms. The fact that it is these two communities that may have been in the process of further disambiguating their person paradigms is consistent with their location in the central part of the Sierra Sur, the apparent epicenter of the morphological innovations in the region.

3.4 *Imperfective* -y

Another morphological feature singled out by MacKay and Trechsel (2015) as diagnostic of Sierra is the presence of the imperfective suffix *-y* in word-final position of the first person singular and third person subject forms of vowel-final verb stems. This suffix is found in Tepehua but is absent in Northern and Lowland (and Misantla):

	Upper Necaxa		Huehuetla (Troiani, 2004: 88)	
(21)	*i̱kɬawá*	'I do it'	*ktɬawáy*	'I do it'
	ɬawáya̱	'you do it'	*tɬawáya?*	'you do it'
	ɬawá	's/he does it'	*tɬawáy*	's/he does it'
	i̱kɬawayá:w	'we$_{\text{EXCL}}$ do it'	*ktɬawayáw*	'we$_{\text{EXCL}}$ do it'
	ɬawayá:w	'we$_{\text{INCL}}$ do it'	*tɬawayáw*	'we$_{\text{INCL}}$ do it'
	ɬawaya:tít	'you guys do it'	*tɬawayátit*	'you guys do it'
	taɬawá	'they do it'	*tɬawaqúy*	'they do it'

The underlying form of the imperfective is *-ya:*, which only surfaces when the affix is protected from final syncope by the presence of some other suffix to its right (e.g., UN *ɬawa-ya:-tít*, Hu *tɬawa-yá-tit* 'you guys do it'). In Sierra languages, word-final syncope removes only the vowel, whereas in Northern and Lowland the entire suffix is lost. The retention of *-y* is reported to be very consistent across the Sierra group and is heard in the isolation forms of words in recordings from all the languages in the region other than Tonalixco and Eloxochitlán; however, it is not attested in any of the Northern or Lowland languages in the sample, in Cerro Xinolatépetl, or in Filomeno Mata (Fig. 10).

Unlike the other features we have seen, the presence of *-y* is more likely to be conservative than innovative, as the suffix is also found in Tepehua languages. Loss of *-y* in most of the Central group seems to be part of a more general tendency to elide or lenite word-final segments. This word-final syncope has had a particular impact on the form of aspectual suffixes. Example (22) illustrates this for the perfect suffix *-ni̱:tan*. As in (21), we see syncope when the affix is

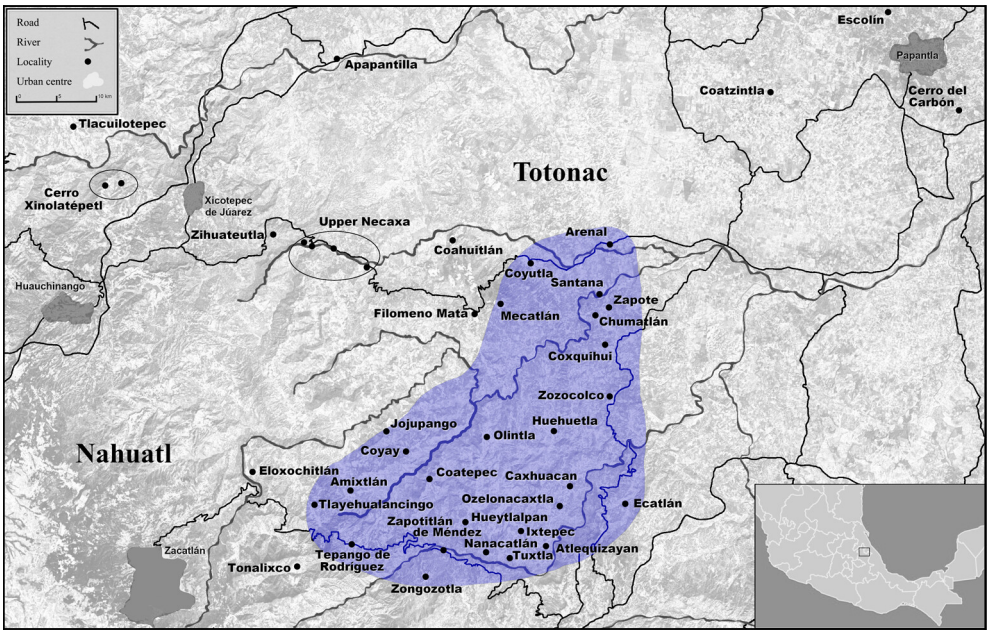

FIGURE 10 Distribution of imperfective *-y*

word-final and a full form when it is followed by another affix (or in the second person singular, where the subject is expressed by laryngealization of the final vowel, historically a suffix, *-ʔ):

	Upper Necaxa		Huehuetla (Troiani, 2004: 89)	
(22)	*i̱kɬawn̰i̱:*	'I have done it'	*iktsuní:t*	'I have done it'
	ɬawan̰i̱:ta̱	'you have done it'	*tsukní:ta*	'you have done it'
	ɬawan̰i̱:	's/he has done it'	*tsukní:t*	's/he has done it'
	i̱kɬawan̰i̱:táw	'we$_{\text{EXCL}}$ have done it'	*ktsukni:táw*	'we$_{\text{EXCL}}$ have done it'
	ɬawan̰i̱:táw	'we$_{\text{INCL}}$ have done it'	*tsukni:táw*	'we$_{\text{INCL}}$ have done it'
	ɬawan̰i̱:tantít	'you guys have done it'	*tsukni:tántit*	'you guys have done it'
	taɬawan̰i̱:	'they have done it'	*tsukqun̰í:t*	'they have done it'

The loss of the imperfective *-y* in (21) thus appears to be the result of the same process illustrated in (22). The loss of the final *-y* suffix in separate branches of the Totonac group may thus be an example of Sapir's (1921) notion of "drift," in this case the tendency toward the syncope of word-final aspectual suffixes, although why this process was arrested in Sierra remains an open question.

3.5 *Stative plural suffix* -nan

An additional inflectional feature of interest, not discussed in MacKay and Trechsel (2015), is the use of the stative plural suffix *-nan*. This suffix is found in Northern and Lowland languages both with verbs in the progressive aspect and with a small class of stative posture verbs (*ma: 'be lying,' *ya: 'be standing,' *wi: 'be sitting,' and *wakaꞌ 'be high') when either have plural subjects, as seen in (23):

Upper Necaxa

(23) a. *i̱kmaꞌ:**náw***
*i̱k-maꞌ:-**nan**-w*
1SG.SUB-be.lying-ST.PL-1PL.SUB
'We$_{EXCL}$ are lying down.'

b. *paꞌ:**nantít***
*paꞌ:-**nan**-tit*
be.lying:2SUB-ST.PL-2PL.SUB
'You guys are lying down.'

c. *tamaꞌ:**naɬ***
*ta-maꞌ:-**nan**-li̱*
3PL.SUB-be.lying-ST.PL-PFV
'They are lying down.'

d. *i̱kli:tsi̱:maꞌ:**naw***
*i̱k-li:tsi̱:n-maꞌ:-**nan**-w*
1SG.SUB-laugh-PROG-ST.PL-1PL.SUB
'We$_{EXCL}$ are laughing.'

e. *li:tsi̱mpaꞌ:**nantít***
*li:tsi̱:n-paꞌ:-**nan**-tit*
laugh-PROG:2SUB-ST.PL-2PL.SUB
'You guys are laughing.'

f. *tali:tsi̱maꞌ:**náɬ***
*ta-li:tsi̱:n-maꞌ:-**nan**-li̱*
3PL.SUB-laugh-PROG-ST.PL-PFV
'They are laughing.'

Diachronically speaking, the Central Totonac progressive is based on a compound combining the stem of a dynamic verb with a reflex of the stative posture verb *ma: 'be lying' (Beck, 2011b). Stative posture verbs all take *-nan* when their subjects are plural, as shown in (23a–c) for the reflex of *ma: 'be lying' in Upper Necaxa, *ma:ɬ*. This use of *-nan* is maintained the progressive aspect in Northern and Lowland languages (23d–f). Note that *ma: has a suppletive second person subject form, *paꞌ:, seen in (23b), and this is maintained in the progressive inflection in (23e). Many languages like Upper Necaxa also preserve an "empty" use of the perfective suffix in these contexts, which can be seen in (23c) and (23f). In Misantla and Tepehua the progressive aspect is realized differently, but *-nan* does occur on stative posture verbs with plural subjects.

Most languages in the Sierra Sur, on the other hand, follow the pattern shown in (24), where *-nan* is lost in both with stative posture verbs and in the progressive aspect:

Zongozotla

(24) a. *kmaːw*
k-maː-w
1SG.SUB-be.lying-1PL.SUB
'We$_{EXCL}$ are lying down.'

b. *pa̲ːtit*
pa̲ː-tit
be.lying:2SUB-2PL.SUB
'You guys are lying down.'

c. *maːqo̲ːɬ*
maː-qo̲ː-li̲
be.lying-3PL-PFV
'They are lying down.'

d. *kliːtsi̲ːmáːw*
k-liːtsi̲ːn-maː-w
1SG.SUB-laugh-PROG-1PL.SUB
'We$_{EXCL}$ are laughing.'

e. *liːtsi̲ːmpa̲ːtit*
liːtsi̲ːn-pa̲ː-tit
laugh-PROG:2SUB-2PL.SUB
'You guys are laughing.'

f. *liːtsi̲ːmaːqo̲ːɬ*
liːtsi̲ːn-maː-qo̲ː-li̲
laugh-PROG-3PL-PFV
'They are laughing.'

With the exception of the third-person plurals, which show the contrast between the Northern *ta-* and the Sierra *-qo̲ː* for 3PL.SUB, most of the forms in (24) are identical to those in (23)—except for the fact that *-nan* is absent. The paradigms in (24) still show suppletion in the second person forms and include an empty perfective suffix.

The distribution of *-nan* is illustrated in the map in Fig. 11. As can be seen, the suffix is found consistently in Northern, Lowland, and Cerro Xinolatépetl (green), but it is not found in any of the languages in the region running from Santana and Zapote south (blue). There is, however, a small band of languages, all grouped with Sierra-Lowland by lexical isoglosses, on the northern edge of this area where *-nan* is attested only in specific persons. In Filomeno Mata (red), McFarland (2009: 54) reports that *-nan* appears with 3PL subjects, while survey data for Mecatlán and Arenal (orange) indicate that in these varieties *-nan* is found only with 1PL subjects. We should also note that there is at least some variability in Coyutla which, according to the data found in Kaufman et al. (2003–2005), does not have progressive or stative forms with *-nan* in the second person plural; however, textual examples with 2PL subjects drawn from Aschmann (n.d.b) do use *-nan*. Given that Aschmann's data were collected several decades earlier, it may mean that there has been a very recent loss of *-nan* with 2PL subjects in Coyutla. Thus, as with the distribution of 2(PL) > 1(PL) forms, the spatial distribution of *-nan* gives the impression of a center bounded by a frontier area where a morphological change is in process.

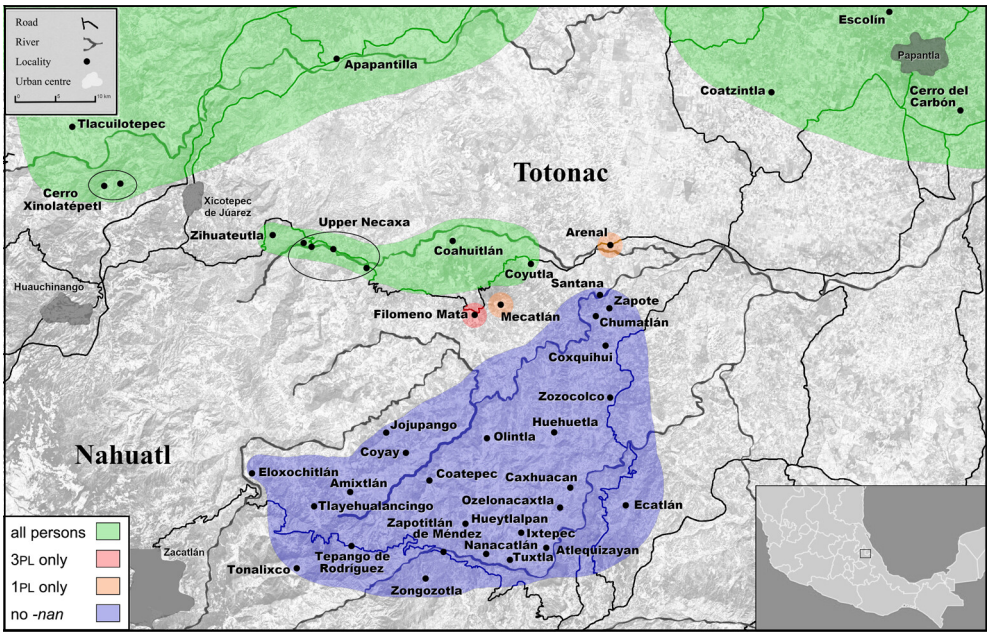

FIGURE 11 Distribution of the stative plural suffix *-nan*

4 Conclusions

Determining the internal classification of the Central Totonac branch of the Totonacan family is a challenging task, given the shallow time depth, the close geographic proximity of the languages, and the potential for borrowing and horizontal transfer of grammatical features between the members of the group. This paper presents lexical evidence, based on the comparison of items on a 405-word list for 45 languages, for the primary division in Central Totonac being a division between Northern Totonac and groups traditionally designated "Sierra" and "Lowland." This distinction is supported by the presence of 36 isoglosses, 31 of which show the Sierra-Lowland group to be lexically innovative with respect to Proto-Totonac and/or Proto-Totonacan. The same set of isoglosses also supports Moore's (2017) conclusion that Filomeno Mata is more closely affiliated with Sierra-Lowland, and suggests that Cerro Xinolatépetl should also be included in that group, although both languages show lexical influence from adjacent Northern languages. Lexical data do not currently support the traditional division between Sierra and Lowland Totonac.

The distinction between Sierra and Lowland seems somewhat easier to maintain on the basis of morphological features (MacKay and Trechsel, 2011, 2014, 2015), although even that is not unproblematic. MacKay and Trechsel

(2015) identify six morphological traits as being unique to, but not present in all, Sierra languages and they argue that three of these are diagnostic of this branch of the family—loss of the syncretic 2(PL) > 1(PL) form; the presence of the *-y* imperfective suffix; and the presence of *-qo:* 'PL' (MacKay and Trechsel, 2015: 138). Later in the same paper, MacKay and Trechsel key on the presence of *-qo:* 'PL' as the crucial feature "both necessary and sufficient to identify a language as Sierra Totonac" (MacKay and Trechsel, 2015: 140). However, it is not entirely clear in the paper whether they are referring to the "unselective" use of *-qo:* as a 3PL index (for subject, object, or both), or whether any use of *-qo:* to index 3PL at all would be necessary and sufficient to classify a language as belonging to the Sierra branch. As we saw in Section 3.1.1 above, the two conditions define different sets of languages. The unselective use of *-qo:* occurs only in Huehuetla, Olintla, Coatepec, Ozelonacaxtla, Hueytlalpan, Zapotitlán, Ixtepec, Nanacatlán, and Tuxtla, all located in the central area of the Sierra Sur. A much larger group of languages would be singled out if we also included those that use *-qo:* only to pluralize subjects, and this latter group would coincide more closely with the traditional conception of "Sierra." We thus might want to reformulate MacKay and Trechsel's diagnostic to specify that any and all uses of *-qo:* as a 3PL.SUB index would suffice to classify a language into the Sierra branch. This would differentiate Sierra from other Totonacan languages that use *ta-* '3PL.SUB'—that is, it would define Sierra languages as those lying in the blue and orange areas of Fig. 6.

Even if we were to take the presence of *-qo:* '3PL.SUB' as necessary and sufficient for a language to be classified as Sierra, this still does not resolve the problem of accounting for all of the concomitant morphological innovations seen in the proposed Sierra branch in terms of the traditional cladistic model of the family tree—that is, describing the developments in the family in terms of synapomorphic innovations inherited by descendant groups of languages from common ancestors. It might be possible to claim that an initial reanalysis of *-qu: from a totalitive/terminative affix to a 3PL subject index was the first such innovation; however, accounting for the micro-variation in the uses of *-qo:* and its idiosyncratic distribution in the verbal paradigms of individual languages that we see, for instance, in Figs. 7 and 8, becomes extremely difficult: to do so, we would need to posit identical, sporadic changes to the person paradigms as having occurred independently in different subregions of the Sierra Sur. Once we add to the mix the other Sierra morphological traits—the loss of syncretic 2(PL) > 1(PL) forms (Fig. 9), the retention of *-y* (Fig. 10), and the loss of *-nan* (Fig. 11)—drawing clean lines of descent that maintain the geographic integrity of the descendant languages so defined becomes nearly impossible. A cladistic phylogenetic model would require us to hypothesize many parallel indepen-

dent innovations in noncontiguous languages, and would have no real account for languages like Mecatlán that "mix" Sierra and Proto-Central person indexing systems, or for languages like Filomeno Mata, Mecatlán, and Arenal that seem to have only partially undergone paradigmatic changes (in these languages, the incomplete loss of *-nan*).

On the other hand, the spatial distribution of the morphological patterns typical of the Sierra Sur becomes less problematic once we recognize that they follow a wave-like pattern resembling the diffusion of a chain of morphological innovations originating in the southern part of the region and moving outwards. This pattern is most obvious in Fig. 6, where we see a first wave of *-qo:* '3PL.SUB' in an area centered around Zapotitlán de Méndez, followed by a subsequent wave as the suffix makes a further change to become a plural object index. Similar patterns can be seen in the maps for all the Sierra morphological innovations considered in this paper (with the exception of *-y* 'IMPERFECTIVE'). The most natural interpretation of these distributions is that what we are seeing is a cascade of changes to the person paradigm in the process of spreading horizontally between languages by diffusion rather than inheritance, with the epicenter of innovation residing in the south-central Sierra Sur and radiating outwards to the northeast and to the west. Some of the changes are incomplete in individual languages, while languages closer to the center show greater consistency in the adoption of diffused patterns and a much stronger tendency to have all of the unique Sierra morphological features.

In a wave-like model of diffusion, the innovations with the greatest geographic extension are interpreted to be the earliest, which in this case points to the following temporal sequence of innovations:

1. *-qo:* used for 3PL > 3OBJ forms (Fig. 6)
2. loss of *-nan* (Fig. 11)
3. *-qo:* used for 3PL > 1/2PL (Fig. 8)
4. *-qo:* used for all 3PL objects (other than with 2PL subjects; Fig. 6)
5. regularization of the syncretic 2(PL) > 1(PL) form (Fig. 9)

The first step in the chain is the reanalysis of the Central Totonac totalitive/terminative *-qo:* as a 3PL subject index, a change from a free use of the affix indicating that an entire group performed an action to its use as an obligatory index of subject plurality for third persons. The second step, the loss of *-nan*, seems somewhat independent of the integration of *-qo:* into the person-marking system, but the fact that *-nan* has been all but erased from Sierra suggests that it was an early change, contemporaneous, or nearly so, with the spread of the use of *-qo:* with subjects. It is possible that the elimination of *-nan* was driven by the redundancy of having two expressions of subject plurality adjacent to one another in the verbal template. Steps 3 and 4 represent further encroachment

of *-qo:* into the person paradigms. The final change in the sequence, the loss of the syncretic *kin- la:- -w* 2(PL) > 1(PL) form, like step 2, is independent of the adoption of *-qo:* as a 3PL index, but represents a further simplification of the inflectional system in that it replaces a single multiply ambiguous form having three possible readings with separate compositional forms. As we saw in Section 3.3, Coatepec has taken a further step down this road and has replaced the syncretic 1(PL) > 2(PL) form found in other Totonac languages with three separate, nonambiguous verb forms.

The sequence of changes posited here is consistent with what we see in the *Arte de la lengua totonaca* (1990), a grammar written in the sixteenth or early seventeenth century in Hueytlalpan, which is located at the apparent epicenter of innovation in the central Sierra Sur. The *Arte* shows *-qo:* being used for the plurality of third person subjects and no use of *-nan*, indicating that steps 1 and 2 in the sequence were complete in Hueytlalpan at that point in time. There are no examples of 3PL > 1/2PL forms, so the status of step 3 is unknown; however, the *Arte* reports that both *ka:-* and *-qo:* are used interchangeably for the indexation of third plural objects (p. 63), indicating that step 4 in the sequence was in progress at the time the grammar was written. The PDLMA survey shows that in modern Hueytlalpan *-qo:* has fully replaced *ka:-* for 3PL.OBJ. The *Arte* also shows that the conservative *kin- la:- -w* form was in use at the time of writing, at least for 2SG > 1PL (p. 66) and 2PL > 1SG (p. 68) forms (there are no examples of 2PL > 1PL), meaning that step 5 had yet to take place, although it has since taken hold in modern Hueytlalpan. As noted in footnote 6 above, to the extent that the *Arte* provides lexical information about sixteenth-century Hueytlalpan, the Sierra-Lowland lexical isoglosses seem to have been in place at this point in time, except for the replacement of *kałwan 'weep' with the metonymic use of *tąsa 'vocalize' that we see in the modern Sierra languages. This shows that the isoglosses that distinguish Northern and Sierra-Lowland predate at least some of the morphological innovations that define the Sierra group.

More evidence of the relative timing of the lexical vs. morphological innovations comes from consideration of Cerro Xinolatépetl and some of the languages located on the western edge of the Sierra Sur—Amixtlán, Coyay, Eloxochitlán, Jojupango, Tonalixco, and Tlayehualancingo. These languages share a very salient phonological characteristic, the presence of a palatal glide /y/ in the onset of words that in other Totonac languages have an initial vowel or glottal stop. This is illustrated in Table 6 in the form of the word for FLEA. As we see in Fig. 12, Cerro Xinolatépetl is currently isolated from the other *y*-languages by Nahuatl-speaking (or formerly Nahuatl-speaking) communities; however, the fact that they share the *y*-onset suggests that at one point they must have been in closer proximity, either because this feature was inherited from a com-

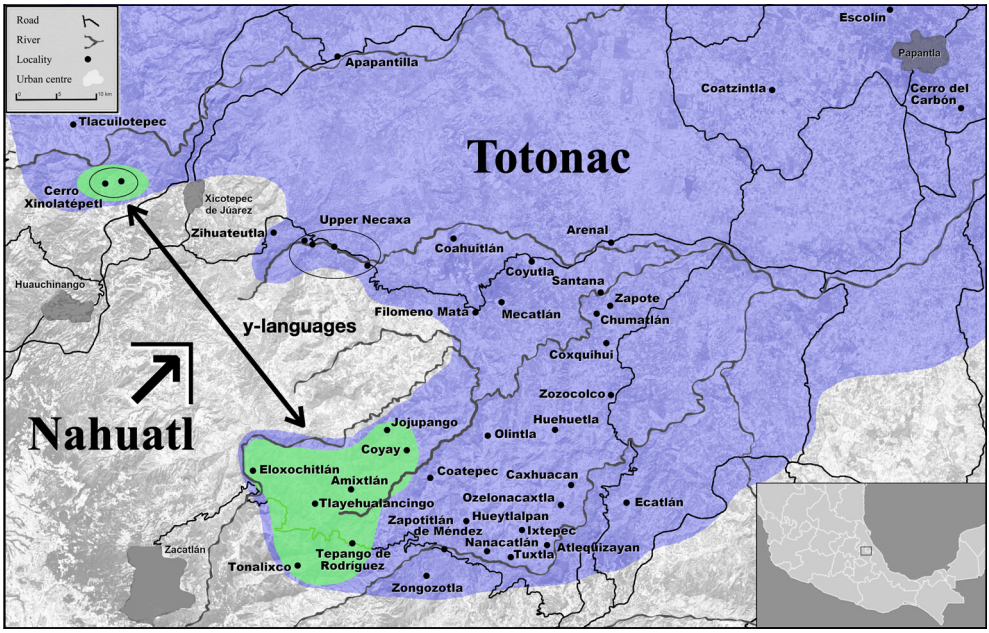

FIGURE 12 Separation of *y*-languages by Nahuatl migration

TABLE 6 Central Totonac words for FLEA

	Northern	*y*-languages		Sierra-Lowland
FLEA	Ap *ɋqtsɻ́ːs* Pn *ɋqtsíːs* UN *ɋʔtsíːs* Zi *ɋqtsíːs*	CX *yɋ́qtsiːs*	Am *yáqtsis* El *yáqtsiːs* Tp *yɋqtsɻ́ːs* To *yɋ́qtsiːs*	At *ɋ́qtsiːs* Mc *ɋ́qtsɻ́ːs* Pp *aqtsíːs* Zp *ɋ́qtsɻ́ːs*

mon ancestor language or because it diffused across a set of languages in close contact. The *y*-languages in the Sierra Sur share in the lexical isoglosses differentiating the Sierra-Lowland group from Northern Totonac—as does, for the most part, Cerro Xinolatépetl, which has 23 of the 36 Sierra-Lowland forms given in Table 3 above. However, Cerro Xinolatépetl shares none of the Sierra morphological innovations, whereas the earliest of these (use of *-qɋː* for 3PL subjects with all persons of object and the loss of *-nan*) is found in the other *y*-languages.

The presence of the lexical isoglosses in Cerro Xinolatépetl but the absence of any of the grammatical innovations indicates that the separation of Cerro

Xinolatépetl from the other *y*-languages postdates the lexical changes in Sierra-Lowland but predates the morphological changes. Dating the separation is difficult, but historical accounts place the arrival of the Nahuatl in the area around the mid-fifteenth century, when the Aztecs under Moctezuma conquered the area (Davies, 1973; Acuña, 1985).[19] The Nahuatl intrusion into the region might have isolated Cerro Xinolatépetl somewhere around this time, which would have been before the Sierra morphological innovations began affecting the western part of the Sierra Sur. Another possibility is that the Cerro Xinolatépetl speakers moved to their current location from an area closer to that occupied by the other *y*-languages after the arrival of Nahuatl, which would place the spread of the morphological features found in the Sierra Sur at an even later date. In either case, the innovations in the Sierra-Lowland lexicon would necessarily have taken place before the isolation of Cerro Xinolatépetl, which prevented its acquisition of the Sierra morphological innovations.

Based on the evidence available to us, then, it is clear that the lexical innovations that characterize Lowland and Sierra predate the morphological innovations that characterize languages in the Sierra Sur: the latter give the appearance of being relatively recent and, perhaps, still in the process of spreading. This scenario, or one very much like it, puts us in a position to propose a more articulated view of the internal structure of Totonacan, shown in Fig. 13. Figure 14 shows how the languages included in this study would be classified according to this hypothesis.

The tree in Fig. 13 takes a traditional cladistic view of the family, treating the innovations or changes that distinguish each branch in the tree as "speciating" events that define new phylogenetic divisions in the family. As we have seen in this paper, however, some of the innovations that define the different groups are, at least potentially, features that have appeared in individual languages due to horizontal transfer rather than inheritance. The farther down the tree we go the less certain we can be that what we see are traditional phylogenetic groups. At the top of the tree, Totonac and Tepehua are well-established and uncontroversial phylogenetic units. The next division within Totonac distinguishing Central and Misantla would almost certainly be an actual phylogenetic division as well, given the marked morphological and lexical differences that set Misantla apart. The lexical isoglosses presented in Table 3 above seem to fur-

19 Another event found in historical accounts that could point to a much earlier date for the arrival of Nahuatl would be the advent of groups of "Chichimecas" in the Sierra Sur near Zacatlán sometime in the late twelfth century (Acuña, 1985: 170; Hasler, 1993). This would push the timing of the separation of Cerro Xinolatépetl even further back, but would have no effect on the relative timing of the lexical and morphological innovations.

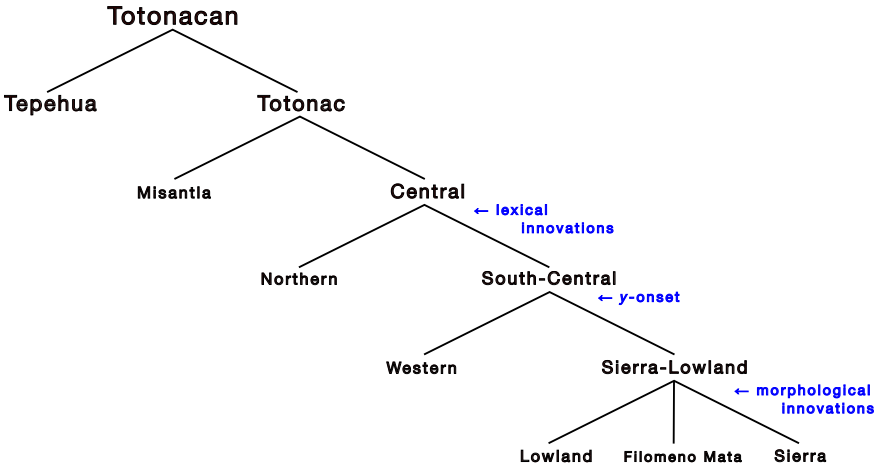

FIGURE 13 Cladistic view of the Totonacan family

TEPEHUA
Tlachichilco (TT), Pisaflores (PT), Huehuetla (HT)
TOTONAC
Misantla (Mi)
CENTRAL TOTONAC
NORTHERN
Apapantilla (Ap), Coahuitlán (Ch), Pantepec (Pn), Tlacuilotepec (Tc), Upper Necaxa (UN), Zihuateutla (Zi)
SOUTH-CENTRAL
WESTERN SIERRA
Amixtlán (Am), Cerro Xinolatépetl (CX), Coyay (Cy), Eloxochitlán (El), Jojupango (Jo), Tepango de Rodríguez (Tp), Tlayehualancingo (Ty), Tonalixco (To)
SIERRA-LOWLAND

LOWLAND
Cazones (Ca), Cerro del Carbón (CC), Coatzintla (Co), Coyutla (Ct), Escolín (Es), Tihuatlán (Ti)
Filomeno Mata (FM)
SIERRA
Arenal (Ar), Atlequizayan (At), Caxhuacan (Cw), Chumatlán (Cm), Coatepec (Co), Coxquihui (Cq), Ecatlán (Ec), Huehuetla (Hu), Hueytlalpan (Hy), Ixtepec (Ix), Mecatlán (Mc), Nanacatlán (Na), Olintla (Ol), Ozelonacaxtla (Oz), Santana (Sn), Tuxtla (Tu), Zapote (Zt), Zapotitlán (Zp), Zongozotla (Zo), Zozocolco (Zz)

FIGURE 14 Internal divisions within the Totonacan family

ther ramify the Central group into Northern and what could be referred to as the South-Central branch. On the balance of probabilities, I feel this is most likely also a genuine cladistic phylogenetic division, although it is certainly possible that the lexical items that currently define the South-Central branch could have been spread throughout the area by borrowing. However, if this were the case, we would expect a much spottier distribution of the lexicon in question. This is not what we see, nor do we see (as we did with the Sierra morphological changes) any sign of a potential center of diffusion, distance from which predicts the likelihood of South-Central vocabulary being present. Instead, the distribution of the lexical items, at least in the data we have currently, is very homogeneous across the region, which is consistent with the dispersion of lexical items through migration and dispersal of populations. Nevertheless, it would be premature to rule out the borrowing scenario entirely.

Within South-Central, the *y*-languages, labeled "Western Sierra" in Fig. 13, are an easily identifiable subgroup, although if the *y*-onset turns out to have been diffused among, rather than inherited by, these languages, Western Sierra would not necessarily be a traditional cladistic grouping. Further divisions within the newer, more restricted version of Sierra-Lowland are also difficult to establish with any precision, given that, as we have seen, the morphological features usually used to define the Sierra division do not paint a consistent picture: depending on which specific morphological pattern we choose, the divisions between Sierra and Lowland will be made differently. Even if we were to settle on the most widespread feature, the use of *-qo:* as a 3PL subject index, it is not clear that the division we would be making in the family is a genuine cladistic division, defined by a synapomorphic innovation made in one language and inherited by its descendants. The wave-like distribution of the successive morphological changes found in the Sierra Sur fits much better with a model of diffusion or horizontal transfer of morphological innovations through contact, spread through what might be best viewed as dialect continuum, with easily identifiable—perhaps mutually unintelligible—variants at either end and a series of intermediate varieties in between. Thus, based on the data presented in this paper, "Lowland" and "Sierra" (and perhaps "Western Sierra") appear to be no more than loose descriptive or geographical terms, though further investigation may in the end give us other means to support the distinction.

Looking beyond the implications of this study for the internal reconstruction of Central Totonac and the larger Totonacan family tree, comparison of the lexical and morphological data available to us has revealed some implicit biases in our usual treatment of this kind of data. The first of these, noted above, is a tendency to assume morphology to be more conservative and slower to change or diffuse than lexicon, which is widely accepted to be both more borrowable

and more easily subject to horizontal transfer within a family tree. This is the opposite of the case documented here, where it is the lexicon that has remained relatively stable within the South-Central group and the morphology that has undergone a rapid series of innovations, some which have the appearance of still being in progress. The importance of considering the spatial distribution of grammatical features is also brought to the fore, as in this case it reveals a well-defined sequence of smaller innovations radiating outward from a fairly obvious geographic center in a wave-like pattern. The changes diffused in this way are extensive and salient enough to have created the appearance of a phylogenetic distinction widely accepted by specialists.

Part of the problem with coming to firm conclusions about a tree-like structure for Central Totonac is almost certainly the shallow time depth being considered, and the fact that the languages in question are still geographically proximate, the populations having remained in close contact throughout the period in question. Much like the tree model for biological species, the tree model for historical linguistics is at its most effective when examining older relationships and relationships between populations separated by geographic or long-term sociopolitical barriers. In cases like Central Totonac, where all of the complicating factors—shallow time depth, geographic proximity, borrowing of morphology and lexicon, and wave-like diffusion of key patterns—come into play, we are forced to consider what the lower diachronic limits to traditional cladistic taxonomies might be and what methods we can use to tease apart the relationships between closely related varieties within language families.

Supplementary materials

- Supplementary materials A: Abbreviated PDLMA questionnaire with bilingual index
- Supplementary materials B: Lexical isoglosses and reconstructions
- Supplementary materials C: Beck and Levy (2019) handout

These supplementary materials can be found here: https://doi.org/10.6084/m9.figshare.24566005.

Acknowledgments

The author would like to thank Albert Davletshin, Paulette Levy, Osbel López Francisco, Carolyn MacKay, Teresa McFarland, Faustino Montes Castañeda,

Devin Moore, Susan Smythe Kung, Conor Snoek, Michelle García-Vega, and Jim Watters for their contributions of data, ideas, and encouragement. I would also like to thank Jeff Good and three anonymous reviewers for *Language Dynamics and Change* for helping to make this a much better paper. Any remaining deficiencies in the work are due to my own limitations. Research for this project was funded by the Social Sciences and Humanities Research Council of Canada and by a grant from the Endangered Languages Documentation Programme.

References

Acuña, René (ed.) 1985. *Relaciones geográficas del siglo XVI: Tlaxcala, vol. 2*. Mexico City: Universidad Nacional Autónoma de México.

Andersen, Gerry. 2012. Totonaco del Cerro Xinolatépetl. In Paulette Levy and David Beck (eds.), *Las lenguas totonacas y tepehuas: Textos y otros materiales para su estudio*, 181–193. Mexico City: Universidad Nacional Autónoma de México.

Arana Osnaya, Evangelina. 1953. Reconstrucción del protototonaco. *Revista Mexicana de Estudios Antropológicos* 23: 123–130.

Arte de la lengua totonaca. 1990. Edited by Norman A. McQuown. Mexico City: Universidad Nacional Autónoma de México.

Aschmann, Herman P. 1962. *Vocabulario totonaco de la sierra*. Mexico City: Summer Institute of Linguistics.

Aschmann, Herman P. 1973. *Diccionario totonaco de Papantla*. Mexico City: Summer Institute of Linguistics.

Aschmann, Herman P. n.d.a. Shoebox lexical database for Zapotitlán de Méndez Totonac. In possession of the author.

Aschmann, Herman P. n.d.b. Shoebox lexical database for Coyutla Totonac. In possession of the author.

Bailey, Charles-James N. 1973. *Variation and Linguistic Theory*. Washington, DC: Center for Applied Linguistics.

Beck, David. 2004. *Upper Necaxa Totonac*. Munich: LINCOM Europa.

Beck, David. 2011a. *Upper Necaxa Totonac Dictionary*. Berlin: Mouton de Gruyter.

Beck, David. 2011b. Lexical, quasi-inflectional, and inflectional compounding in Upper Necaxa Totonac. In Alexandra Aikhenvald and Pieter Muysken (eds.), *Multi-verb Constructions: A View from the Americas*, 63–106. Leiden: Brill.

Beck, David. 2012. Apéndice: Tablas de morfología comparativa. In Paulette Levy and David Beck (eds.), *Las lenguas totonacas y tepehuas: Textos y otros materiales para su estudio*, 587–596. Mexico City: Universidad Nacional Autónoma de México.

Beck, David. to appear. Totonacan. In Søren Wichmann (ed.), *The Languages and Lin-*

guistics of Middle and Central America: A Comprehensive Guide. Berlin: Mouton de Gruyter.

Beck, David and Paulette Levy. 2019. From 'done' to 'all' to 'everyone' to 'plural': The evolution of a number agreement suffix in Central Totonac. Coloquio Pluralidad Nominal y Verbal de las Lenguas de las Américas, Universidad Nacional de México, October 9. (See handout in supplementary materials C.)

Brown, Cecil H., David Beck, Grzegorz Kondrak, James K. Watters and Søren Wichmann. 2011. Totozoquean. *International Journal of American Linguistics* 77: 323–372.

Childs, G. Tucker. 2003. *An Introduction to African Linguistics*. Amsterdam: John Benjamins.

Davies, Nigel. 1973. *The Aztecs*. London: Folio Society.

Davletshin, Albert. 2008. Classification of the Totonacan languages. Problemy izučenij dal'nego rodstva jazykov (k 55-leti C.A. Starostina), Russian State University for the Humanities, Moscow, March 25–28.

Davletshin, Albert. 2018. Las vocales finales, los procesos fonéticos finales y mediales en el prototototonaco-tepehua: Un primer acercamiento. In Elisa Cristina Buenrostro Díaz, Lucero Meléndez Guadarrama and Marcela San Giacomo Trinidad (eds.), *Lingüística histórica de lenguas indomexicanas: Hallazgos y discusiones recientes*, 139–186. Mexico City: Universidad Nacional Autónoma de México.

Davletshin, Albert. 2019. Laryngealized vowels and laryngealized consonants in the history of the Totonacan languages of Mexico. *Journal of Language Relationships* 17(3): 177–196.

Davletshin, Albert. In press. La clasificación léxico-estadística y la glotocronología de las lenguas tepehua-totonacas. In Y. de la Cruz Lira López (ed.), *Estudios antropológicos en Veracruz, el Instituto de Antropología a 65 años de su fundación*. Xalapa: Universidad Veracruzana.

Espinoza Morales, Antonio. 1978. *Gramática del totonaco de San Andrés Tlayehualancingo*, MA thesis, Escuela Nacional de Antropología e Historia, Mexico City.

García Rojas, Blanca. 1978. *Dialectología de la zona totonaco-tepehua*. Honors thesis, Escuela Nacional de Antropología e Historia, Mexico City.

García-Vega, Michelle. 2022. *Morphology of Zihuateutla Totonac*. PhD dissertation, University of Alberta.

Hasler, Juan A. 1993. La formación de los grupos totonaco. *La Palabra y el Hombre* 86: 5–21.

Ichon, Alan. 1969. *La religion de los totonacos de la sierra*. Mexico: Instituto Nacional Indigenista.

Instituto Nacional de Lenguas Indígenas. 2008. *Catálogo de las lenguas indígenas nacionales: Variantes lingüísticas de México con sus autodenominaciones y referencias geoestadísticas*. Mexico City: Imprenta del Gobierno Federal, SEGOB.

Instituto Nacional de Lenguas Indígenas. 2021. Población de 5 años y más hablante de alguna lengua indígena por familia lingüística, comparativo 2000–2015. https://site.inali.gob.mx/Micrositios/estadistica_basica/estadisticas2015/pdf/general/general1A.pdf (accessed October 22, 2021).

Instituto Nacional de los Pueblos Indígenas. 2021. Códice México. http://www.inpi.gob.mx/codicemexico/ (accessed October 22, 2021).

Juárez Esteban, Adela. 2016. *Los predicados complejos en el totonaco de Tuxtla, Puebla*. MA thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social, San Cristobal de las Casas.

Juárez Esteban, Teodoro. 2020. *Orden de constituyentes de la oración simple en el totonaco de Tuxtla*. MA thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social, San Cristobal de las Casas.

Kaufman, Terrence, Carolyn MacKay and Frank R. Trechsel. 2003–2005. Totonac-Tepehua Dialect Survey. Archive of the Indigenous Languages of Latin America. https://ailla.utexas.org/islandora/object/ailla%3A124508 (accessed September 30, 2023).

Levy, Paulette. 1987. *Fonología del totonaco de Papantla, Veracruz*. Mexico City: Universidad Nacional Autónoma de México.

Levy, Paulette. 1990. *Totonaco de Papantla, Veracruz*. Mexico City: El Colegio de México.

Levy, Paulette. In press. *Los textos totonacos de Coatepec de N.A. McQuown y M. Oropeza Castro: La jerarquía prosódica en una lengua polisintética*. Mexico City: Universidad Nacional Autónoma de México.

Levy, Paulette and David Beck (eds.). 2012. *Las lenguas totonacas y tepehuas: Textos y otros materiales para su estudio*. Mexico City: Universidad Nacional Autónoma de México.

MacKay, Carolyn J. 1999. *A Grammar of Misantla Totonac*. Salt Lake City: University of Utah Press.

MacKay, Carolyn J. 2011. Una reconstrucción del acento primario en el proto-totonaco-tepehua. In Ana Lidia Munguía Duarte (ed.), *Colección de estudios lingüísticos 1: Fonética, morfología y tipología semántico-sintáctica*, 93–125. Hermosillo: Universidad de Sonora.

MacKay, Carolyn J. and Frank R. Trechsel. 2006. Totonacan languages. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, 2nd edn, 3–8. Oxford: Elsevier.

MacKay, Carolyn J. and Frank R. Trechsel. 2010. *Tepehua de Pisaflores, Veracruz*. Mexico City: El Colegio de México.

MacKay, Carolyn J. and Frank R. Trechsel. 2011. Relaciones internas de las lenguas totonaco-tepehuas. *Proceedings of the Conference on Indigenous Languages of Latin America-V*. Archive of the Indigenous Languages of Latin America. https://www.ailla.utexas.org/node/97 (accessed October 19, 2023).

MacKay, Carolyn J. and Frank R. Trechsel. 2014. Diagnósticos morfológicos para la clasificación de las lenguas totonaco-tepehua. In Rebeca Barriga Villanueva and Esther Herrera Zendejas (eds.), *Lenguas, estructuras y hablantes: Estudios en homenaje a Thomas C. Smith-Stark*, vol. 2, 843–870. Mexico City: El Colegio de México.

MacKay, Carolyn J. and Frank R. Trechsel. 2015. Totonac-Tepehua genetic relationships. *Amerindia* 37: 121–158.

MacKay, Carolyn J. and Frank R. Trechsel. 2018. An alternate reconstruction of Proto-Totonac-Tepehua. *International Journal of American Linguistics* 84: 51–92.

McFarland, Teresa Ann. 2009. *The phonology and morphology of Filomeno Mata Totonac*. PhD dissertation, University of California, Berkeley.

McQuown, Norman A. 1940. *A Totonac grammar*. PhD dissertation, Yale University.

McQuown, Norman A. (with Manuel Oropeza Castro). (1939–1968) 1971. Textos totonacos. Microfilm Collection of Manuscripts on Cultural Anthropology 100. Joseph Regenstein Library, University of Chicago.

McQuown, Norman A. 1990. *Gramática de la lengua totonaca, Coatepec, Sierra Norte de Puebla*. Mexico City: Universidad Nacional Autónoma de México.

Moore, Devin. 2017. Subgrouping of Coahuitlán Totonac. *Canadian Journal of Linguistics* 62: 84–117.

Reid, Aileen A. 1991. *Gramática totonaca de Xicotepec de Juárez*. Mexico City: Summer Institute of Linguistics.

Reid, Aileen A. and Ruth G. Bishop. 1974. *Diccionario totonaco de Xicotepec de Juárez*. Mexico City: Summer Institute of Linguistics.

Reid, Aileen A., Ruth G. Bishop and Ella M. Button. n.d. Shoebox lexical database for Apapantilla Totonac. In possession of the author.

Reid, Aileen A., Ruth G. Bishop, Ella M. Button, and Robert Longacre. 1968. *Totonac: From clause to discourse*. Norman, OK: Summer Institute of Linguistics.

Román Lobato, Gabriela. 2008. *La juntura fonológica en el totonaco de Ozelonacaxtla, Huehuetla, Puebla*. MA thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social, San Cristobal de las Casaa.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and World.

Schmidt, Johannes 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*. Weimar: Böhlau.

Smythe Kung, Susan. 2007. *A Descriptive Grammar of Huehuetla Tepehua*. PhD dissertation, University of Texas at Austin.

Swadesh, Morris. 1971. *The Origin and Diversification of Language*. Chicago: Aldine.

Tadmor, Uri, Martin Haspelmath and Bradley Taylor. 2010. Borrowability and the notion of basic vocabulary. *Diachronica* 27(2): 226–246. https://doi.org/10.1075/dia.27.2.04tad.

Troiani, Duna. 2004. *Aperçu grammatical du totonaque de Huehuetla, Puebla, Mexique*. Berlin: LINCOM Europa.

Watters, James K. 1988. *Topics in Tepehua grammar*. PhD dissertation, University of California, Berkeley.

Watters, James K. n.d. Lexical database for Tlachichilco Tepehua. In possession of the author.

Wolfram, Walt and Natalie Schilling-Estes. 2003. Dialectology and linguistic diffusion. In Brian D. Joseph and Richard Janda (eds.), *The Handbook of Historical Linguistics*, 713–735. Oxford: Blackwell.