ASSESSING THE CHITIMACHA-TOTOZOQUEAN HYPOTHESIS[1]

DANIEL W. HIEBER

UNIVERSITY OF CALIFORNIA, SANTA BARBARA

1. Introduction[2]

Scholars have attempted to genetically classify the Chitimacha language of Louisiana ever since the first vocabulary of the language was collected by Martin Duralde in 1802. Since then, there have been numerous attempts to relate Chitimacha to other isolates of the region (Swanton 1919; Swadesh 1946a; Gursky 1969), Muskogean as part of a broader Proto-Gulf hypothesis (Haas 1951; Haas 1952), and even languages as far afield as Yuki in California (Munro 1994). The most recent attempt at classification, however, looks in a new direction, and links Chitimacha with the recently-advanced Totozoquean language family of Mesoamerica (Brown, Wichmann & Beck 2014; Brown et al. 2011), providing 90 cognate sets and a number of

---

[1] [Acknowledgements]

[2] Abbreviations used in this paper are as follows:

*	reconstructed form
**	hypothetical form
intr.	intransitive
post.	postposition
tr.	transitive
AZR	adjectivizer
CAUS	causative
NZR	nominalizer
PLACT	pluractional
TRZR	transitivizer
VZR	verbalizer

morphological parallels as evidence. Now, recent internal reconstructions in Chitimacha made available in Hieber (2013), as well as a growing understanding of Chitimacha grammar (e.g. Hieber forthcoming), make it possible to assess the Chitimacha-Totozoquean hypothesis in light of more robust data. This paper shows that a more detailed understanding of Chitimacha grammar and lexicon casts doubt on the possibility of a genetic connection between Chitimacha and Mesoamerica. Systematic sound correspondences prove to be unattainable for the data provided in Brown, Wichmann & Beck (2014). However, groups of correspondences do appear in the data, suggestive of diffusion through contact rather than genetic inheritance. I argue that regional trade networks between the Lower Mississippi Valley and Mesoamerica, and the coastal position of the Chitimacha and Totonacan peoples, would have made such diffusion possible.

This paper proceeds as follows: Section 2 provides some background on the relevant languages, their geneological tree, and their locations. Section 3 describes the source of the data for this study and provides some details regarding methdology, including details on how the data from Brow, Wichmann & Beck (2014) has been revised in light of data from Chitimacha. Section 4 then presents the new set of sound correspondences based on these revised word sets. Section 5 concludes by arguing for a situation of language contact or linguistic diffusion rather than genetic relatedness, and suggests a possible mechanism by which linguistic diffusion could have taken place.

2. Background

2.1. Chitimacha

The first written record of people we presume to be speakers of Chitimacha is from Hernando de Soto's meandering expedition out of Florida and along the Mississippi, when his men were accosted by spear-throwing (i.e. atlatl) inhabitants at the mouth of the Mississippi in 1543 (Swanton 1938). Upon the later arrival of the French in the early 1700s, this same territory was inhabited by the Washa and Chawasha people, said to be relatives of the Chitimacha further west that spoke the same language (Swanton 1917). At the time, the range where Chitimacha must have been spoken extended from Grand Lake in the west to the mouth of the Mississippi in the east (see Figure 1). The eastern bands were quickly killed off in the first twelve years of French presence in the area, before anything could be recorded of their language, so that our only knowledge of the language comes from the dialect spoken around Bayou Teche, near present-day Charenton, Louisiana.

Figure 1. Historic range of the Chitimacha Tribe

The earliest vocabulary of Chitimacha was a 385-item wordlist recorded in 1802 (Duralde 1802) and published in Vater (1821). Based on this and other vocabularies, Gallatin (1836:118) says of the languages of this region that "each of those tribes speaks a distinct language and different from any other known to us." Prior to this, the languages of the U.S. Southeast, then little-known, were presumed to be part of a 'Floridian' or 'Southern' stock (Duponceau 1819).

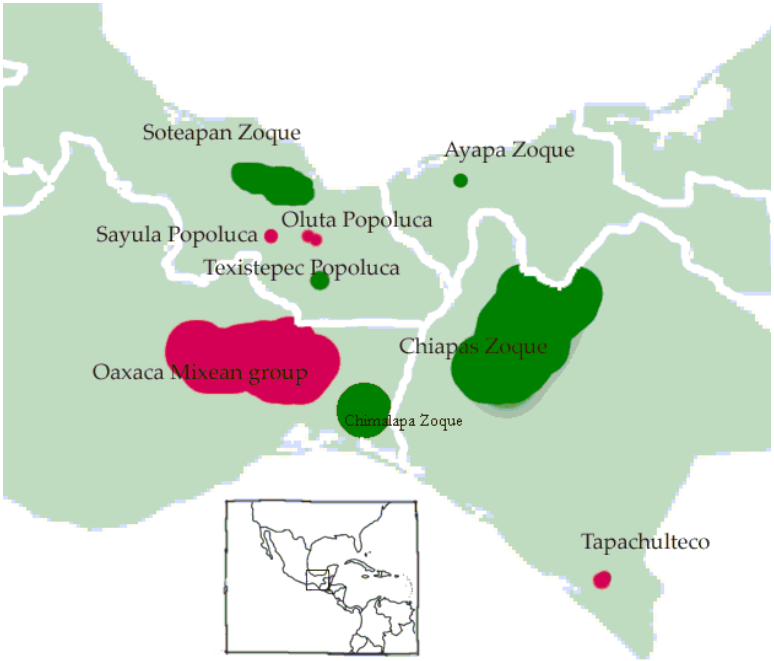
The most extensive and phonetically accurate documentation of the language comes from fieldwork conducted by Morris Swadesh from 1930–1934 in Charenton. He worked with Chief Benjamin Paul and his niece Delphine Ducloux, then the last two native speakers of the language. Chief Paul died in 1934, followed by Mrs. Ducloux in 1940. Swadesh's documentation produced over 200 pages of typewritten texts, a dictionary of approximately 3,500 entries, and a 250-page grammar, but these materials were never published, and now reside at the American Philosophical Society Library in

Philadelphia, PA (Swadesh 1953). Though Swadesh did publish several shorter works on the language (Swadesh 1933; Swadesh 1934; Swadesh 1946a) and a brief grammar sketch (Swadesh 1946b), the inaccessibility of his more extensive documentation meant that little was known of the language by other linguists for many decades. This state of affairs changed when the Chitimacha Tribe began a language revitalization program in the 1990s, and procured digital copies of most of the extant archival materials. Since then, the tribe has started language lessons in the tribal school, created a Chitimacha version of the well-known Rosetta Stone language-learning software, started a preschool immersion program, and many other initiatives. The recent availability of the archival materials has also fostered a small explosion of research on the language, including Iannucci (2009), Hieber (2013), Brown, Wichmann & Beck (2014), Mithun (to appear), and Hieber (forthcoming).

2.2. Mixe-Zoquean

The Mixe-Zoque languages are spoken in the Oaxaca, Chiapas, Veracruz, and Tabasco districts of Mexico (see Figure 2). Roughly speaking, the languages labeled Mixe are spoken in Oaxaca, the Popoluca languages are spoken in Veracruz, Ayapa Zoque is spoken in Tabasco, and Chiapas Zoque is spoken in Chiapas.

Figure 2. The Mixe-Zoque languages



The first attempt at classifying these languages is Wonderly (1949), and the main division between Mixe and Zoque comes from Nordell (1962) and Kaufman (1962; 1963). The history of classification for these languages is summarized in Thomas (1974) and Wichmann (1994). Wichmann's (1995) detailed classification of the Mixe-Zoquean languages remains the authoritative source on the subgroupings of this language family, and contains 2,218 cognate sets. Wichmann proposes the family tree shown in Figure 3.

Figure 3. The Mixe-Zoquean family tree

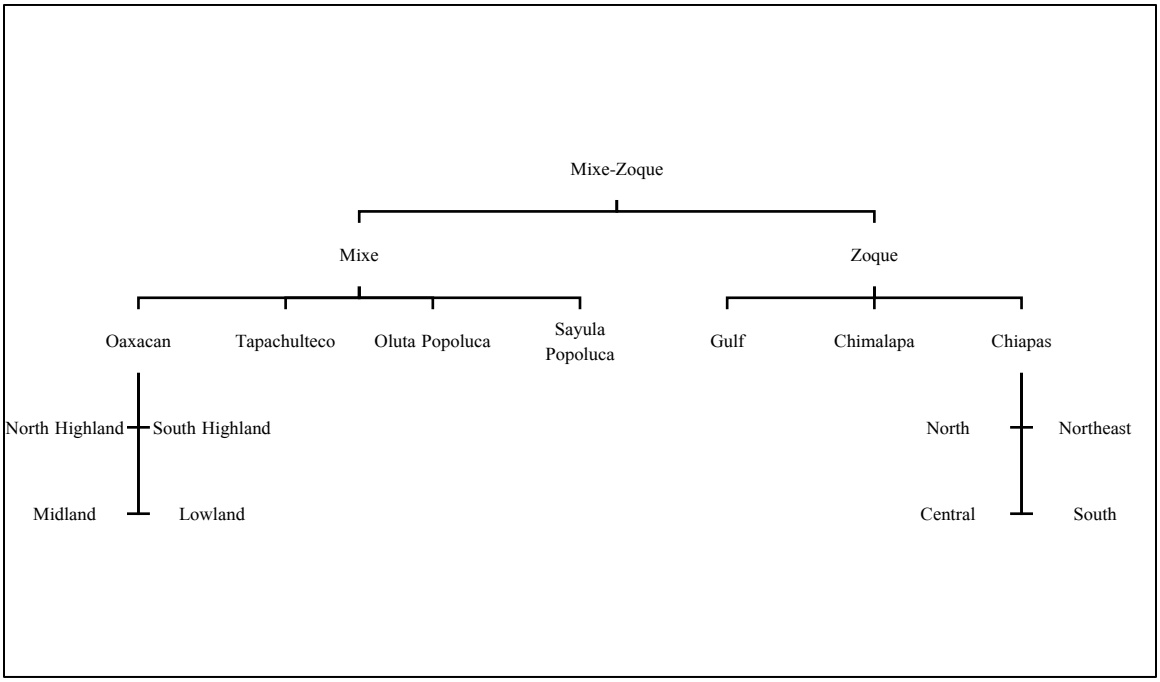


FIG. 1.—Family tree for Mixe-Zoquean, following Wichmann (1995)

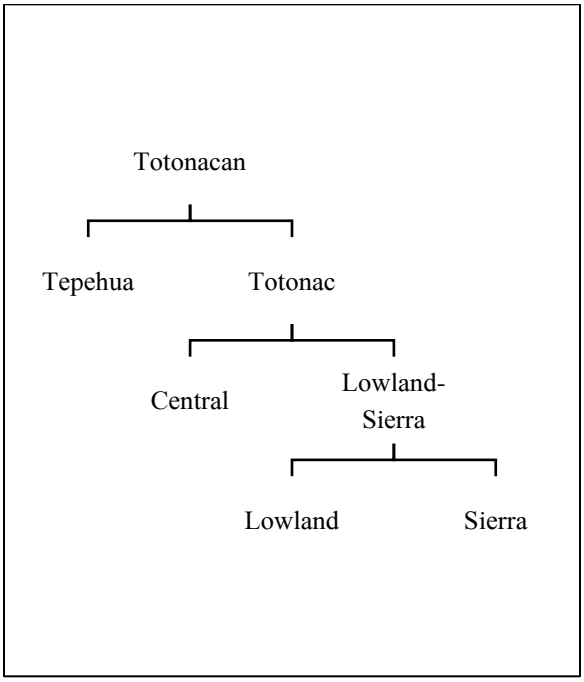
2.3. Totonacan

The Totonacan language family consists of two subgroups – Totonac and Tepehua, spoken by approximately 300,000 people mostly in the Veracruz and Puebla districts of Mexico near the Gulf Coast, some 200–400 miles north of the Mixe-Zoquean languages along the coast (see Figure 4). While Totonac and Tepehua have recently been shown to be genetically related (Kondrak, Beck & Dilts 2007), Brown et al. (2011) is the first attempt at reconstructing Proto-Totozoquean – the proposed protolanguage subsuming all the Tepehua-Totonac and Mixe-Zoquean languages – though the authors note that their Proto-Totonacan reconstructions should be regarded as both preliminary and provisional because the available data on Tepehua-Totonac languages is still scant. Still,

the authors reconstruct 190 words for Proto-Totonacan, and propose the family tree shown in Figure 5.

Figure 4. Locations of the Totonacan and Mixe-Zoquean language families

Figure 5. Proposed family tree for Totonacan (Totonac-Tepehua)



Finally, using the reconstructions for Proto-Totonacan, they reconstruct 188 terms from Proto-Totozoquean, and suggest the family tree provided in Figure 6.

Figure 6. Proposed family tree for Totozoquean

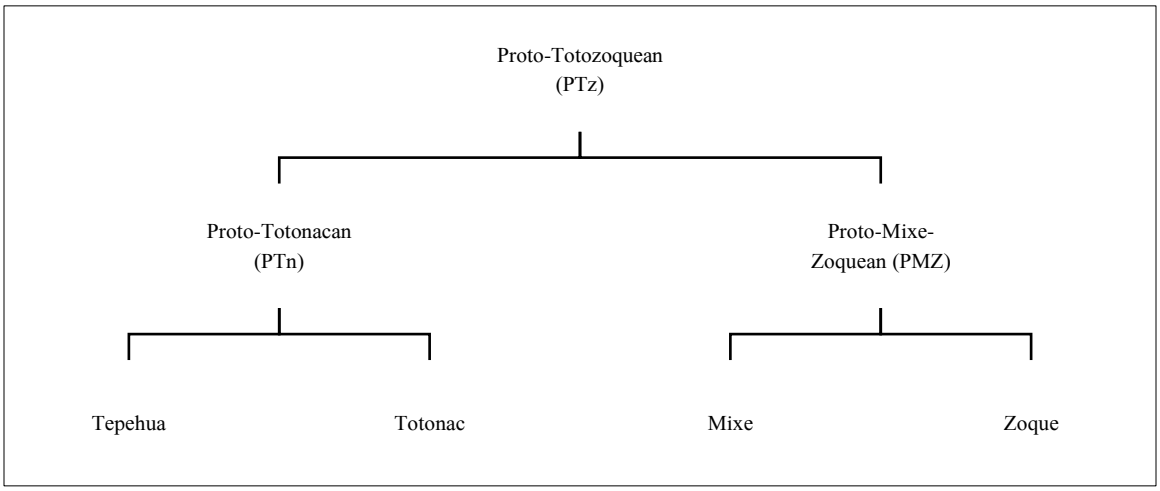


Figure 7 shows the locations of all languages included in this study.

Figure 7. The Chitimacha, Totonacan, and Mixe-Zoquean language families

3. Data & Methodology

The methodology for this study proceeds as follows: First, each of the cognate sets presented in Brown, Wichmann & Beck (2014) (hereafter BWB) are examined in light

of internal evidence from Chitimacha and revised accordingly. Second, I determine the new sound correspondences that arise from these revised sets of words. Once the sound correspondences have been determined, words or sets that exhibit irregular or one-off correspondences and cannot be explained via an appropriate conditioning context are discarded. On the basis of this new data, I determine whether Chitimacha and Totozoquean should be considered genetically related.

The complete set of revisions are provided in **APPENDIX I**. In representing the forms, I use an Americanist orthography slightly different from that of BWB. Correspondences to IPA notation are provided in the footnote below.[3] All other glossing abbreviations are provided in a second footnote below.[4]

The evidence from Chitimacha comes in many forms, including: insight into original or additional meanings of a word; knowledge of internal morpheme boundaries; internal

---

[3]

Grapheme	IPA value
<c>	/t͡s/
<c'>	/t͡sˀ/
<č>	/t͡ʃ/
<č'>	/t͡ʃˀ/
<k'>	/kˀ/
<p'>	/pˀ/
<š>	/ʃ/
<t'>	/tˀ/
<y>	/j/

[4]

Abbreviation	Meaning
AZR	adjectivizer
DTRZR	detransitivizer
NZR	nominalizer
TZR	transitivizer
VZR	verbalizer

evidence for roots that are no longer productive; and internal reconstruction of the original phonetic shape of a morpheme. The primary source of these pieces of evidence is Hieber's (2013) dictionary of Chitimacha, which provides morphemic analyses of each word in the lexicon, numerous forms reconstructed from internal evidence, and many accompanying notes with relevant linguistic details. One other general piece of recurring evidence is that Chitimacha verb roots and most noun roots are canonically CV(:)C, CV(:)hy, CVyk, or CV(:){k/p}(')S, where S = sibilant. Whenever a root is found that contains more phonetic material than this, it is usually possible to analyze the form into distinct morphemes, often followed by a nominalizer or verbalizer suffix. An example of this is *pa:nt'in* 'wing'. Swadesh doesn't list a root for this word, but it's clear that the form derives from a root *pa:n-* + *-te* VZR + *-(ʔ)in* P.NZR – a common derivational pattern in the language.

A last important methodological note is that the internal reconstructions in Hieber (2013) for Chitimacha were arrived at independently of any evidence from Totozoquean. Data from Totozoquean may of course shed more light on the internal history of Chitimacha, but for the present paper only that evidence which can be independently motivated is used to argue for a Chitimacha-Totozoquean connection.

These pieces of evidence from Chitimacha allow for various kinds of revisions to BWB's cognate sets. First, a cognate set may be simply rejected, as for Ch *k'a:ct-* 'cut around' : PMZ **ko:ʔc* 'break'. The Ch word here derives from a root **k'aʔ-* 'circularly, around' that also appears in words like *k'a:ye-* 'be positioned around, coiled'. The

meaning of ‘cut’ is only derivative, stemming from the addition of a *-c* suffix meaning ‘affect by touching’. The Ch and PMZ forms are therefore not likely to be cognate.

Sometimes the revisions consist solely of changing which phonetic material is considered relevantly cognate. For example, BWB’s cognate set Ch *č’ak’u(mt)* ‘chew’ : PTn **cą́qą́* ‘chew’ treats the /u/ of Ch and the **á* of PTn as cognate material, but internal evidence from Chitimacha suggests that *-umt* ‘having to do with the head’ is a separate morpheme (visible in words such as *he:humti* ‘topknot’ and *katumti* ‘coiffure’). Moreover, Ch ejectives occur only before vowels or after a long vowel, so that the historical root of *č’ak’umt-* must have been *č’ak’V*.

The consequence of these revisions for BWB’s sound correspondences is that certain cognate sets can no longer be taken as evidence for particular correspondences. For example, rejecting BWB’s cognate set Ch *k’a:ct-* ‘cut around’ : PMZ **ko:ʔc* ‘break’ means that there is one less set supporting the Ch /k/ : Tn *k : PMZ *k correspondence (which still has 24 other supporting sets), and also one less set supporting their Ch /c/ : *c : *c correspondence, which is left with only 4 supporting cognate sets. This change by itself is relatively minor, but in combination with the 54 revisions proposed in APPENDIX I to the 90 cognate sets which BWB put forward, the overall effect on the sound correspondences is quite drastic.

Therefore, this paper takes the *revised* set of cognates and recalculates the sound correspondenes that derive from them. The question then becomes, of course, does the systematicity of sound correspondences found in BWB still hold? What differences do we see in the correspondenes, and do the same trends still occur? Or must the

possibility of genetic relatedness be abandoned entirely? It will be shown in the following sections that the correspondences proposed by BWB lose a great deal of their systematicity when the cognate sets are revised to take additional data from Chitimacha into account, but that some of the same trends that BWB find do still occur in the data. The implication of this result for the genetic relatedness of Chitimacha and Totozoquean is discussed in §5.

4. Sound Correspondences

Taking the revisions to the cognate sets shown in **APPENDIX I** into account, this section now examines the new sound correspondences that result, comparing them to BWB's correspondences as it proceeds.

4.1. Consonants

4.1.1. Voiceless Stops

For voiceless stops, BWB posit the correspondences listed in Table 1, as well as a rule whereby PCh-Tz *C → Ch /C'/ / __V̰, where C = any obstruent but /p/.

Table 1. BWB's voiceless stop correspondences

Ch	PTn	PMZ	# Sets
p	*p	*p	15
t	*t	*t	11
č	*t	*t	4
c	*t	*t	3
n	*t	*t	4
k	*k	*k	32
k	*q	*k	25
n	*q	*k	5

The sound correspondences that result from my revisions to the Chitimacha are somewhat different, as shown in Table 2. A cell with an m-dash (—) indicates that there is no data for that language family, while a null value (∅) indicates a correspondence to zero (i.e. deletion or epenthesis). While BWB reconstruct four separate series for their PCh-Tz *t, the three least common correspondences have been eliminated from the revised correspondences in Table 2 because the sets supporting them are irregular in their correspondences, or evidence from Chitimacha suggests that they are in fact not cognate. BWB also reconstruct a single segment *k for all three of the final sets in Table 1, whereas I divide the data into two groups: those sets showing a *k in the PTn and those showing a *q.

Table 2. Revised voiceless stop correspondences

Environment	Ch	PTn	PMZ	# Sets
	p	*p	*p	14
_V̲ / Vʔ	p’	—	*p	1
	t	*t	*t	9
_V̲ / _Vʔ	t’	*t	*t	3
	k	*k	*k	20
_V̲ / _Vʔ	k’	*k	*k	6

_#	k	*k	*k / *ʔ	7
	k	*q	*k	5
_V̲ / _Vʔ	k’	*q	—	2
???	n	*q	*k	3

There are four other correspondences which would likely reconstruct to /t/, but are supported by only one or two cognate sets each, and are also irregular in some of their other correspondences as well. They are not included in Table 2.

Finally, there are three correspondences of Ch /n/ : PTn *q : PMZ *k, shown in Table 3, that may not seem phonetically similar at first glance, but whose segments each participate in several regular correspondences. On the other hand, it is not clear what the conditioning environment for this correspondence should be. I continue to include these cognates in the correspondences for the time being.

Table 3. The Ch /n/ : PTn *q : PMZ *k correspondence

Gloss	Ch	PTn	PMZ
	n	*q	*k
‘field, crop’	huwo	*qa̲wa:	—
‘wing’	pa:n(t’in)	*paqa	*pak
‘mat’	t’a:na	*(š)ta̲qa(t)	*ta:kʔ

Tables 4, 5, and 6 show the cognate sets supporting the voiceless stop series in more detail. Outlined cells indicate irregularities / exceptions.

Table 4. Sets underlying the /p/ correspondences

	Environment / Gloss	Ch	PTn	PMZ
(14 sets)		p	*p	*p
	‘person’	**p**an(š)	*(la:)**p**ana	*p~~i~~n
	‘stem of plant’	ka:**p**(ti)	—	*ka**p**(e)
	‘stab, shoot with arrow’	ti**p**(te-)	*ti**p**	*ti**p**
(1 set)	_V̲ / _Vʔ	p’	—	*p
	‘split’	**p’**a:p(te-)	—	***p**ap(s)

Table 5. Sets underlying the /t/ correspondences

	Environment / Gloss	Ch	PTn	PMZ
(8 sets)		t	*t	*t
	‘swing’	**t**ahy(te-)	*(S)**t**iwí	*t**i**ːʔy
	‘sharp point, blade’	si**t**(’ik’i)	*si**t**	—
(3 sets)	_V̲ / _Vʔ	t’	*t	*t
	DEM	**t’**a	*t**a̲**	—
	‘mat’	**t’**aːna	*(š)t**a̲**qa(t)	*t**a**ːkʔ
	‘become wet, drip’	**t’**eyk(te-)	*(S)t**a̲**x	*t**a**ʔks

Table 6. Sets underlying the /k/ correspondences

	Environment / Gloss	Ch	PTn	PMZ
(18 sets)		k	*k	*k
	‘grind’	**k**ih(ci-)	*(S)**k**ití	*t**k**ːʔt
	‘basket’	kaː**k**(t)	—	*kaʔ**k**
(6 sets)	_V̲ / _Vʔ	k’	*k	*k
	‘eat’	**k’**uš(t-)	*t**k**ːʔs	—
	‘corn’	**k’**as(ma)	*t**k̲**s(pa̲)	*(ʔ**t**)**k**s(i)
	‘ant’	(c’iːsc’ik’ah)c’i**k’**ut	—	*cu**k**ut
(7 sets)	_#	k	*k	*k / *ʔ
	‘seize’	po**k**(št-)	—	*poːʔ(t)
	‘write’	haː**k**(šte-)	—	*haːyʔ
	‘young cane ‘reed’	waːs(imiš)	—	*waːsu**k**
	‘head’	ku**t**	*ku**k**	*ko~koʔ
	‘be wet’	t’ey**k**(te-)	*(S)t**a̲**x	*taʔ**k**s
	‘break’	to**h**-	*tu**k**ša	—

Several trends arise out of the above data. By far the most robust of these is a correspondence of Ch /C’V/ : PTn *CV̲ : PMZ *Vː or *Vʔ. Wherever /p’, t’, k’/ occur in Ch, a *V̲ or *Vʔ sequence occurs in PTn and a *Vː or *Vʔ sequence in PMZ. The

reverse is also true for Ch /t', k'/ but not /p'/. That is, on occasion a laryngealized vowel or glottal occurs in the Totozoquean forms with no corresponding ejective consonant in the Chitimacha. This correspondence and its exception in Ch /p'/ was also identified by BWB, and so in this respect the revised sound correspondences still strongly support this part of their findings.

The segment /k/ appears as /k'/ before a laryngealized vowel, shows some variation word-finally, and otherwise remains /k/. The lack of systematicity in word-final position casts some doubt on the cognate sets showing word-final variation in *k, but for these sets it is difficult to tell which segments correspond precisely to which, so it may be that these correspondences could be cleaned up with additional evidence from Totozoquean. And the PMZ forms are consistent in displaying either *ʔ, *k, or both word-finally. So while the correspondences are not systematic, nor are they problematic enough to be discarded.

4.1.2. Sibilants & Affricates

The sibilants and affricates reconstructed by BWB are given in Table 7.

Table 7. BWB's sibilant and affricate correspondences

Ch	PTn	PMZ	# Sets
č	*c	*c	5
c	*c	*c	5
s	*s	*s	5
s	*š	*s	4
š	*š	*s	8
č	*š	*s	5

Table 8 shows the revised series of affricates and sibilants. The main difference between the two is the conditioning environments posited. While BWB simply reconstruct different segments for each of the different correspondences, resulting in an inventory that includes *s, *s^{y}, *š, and *š^{y}, I instead suggest relevant conditioning environments for the different correspondences, while eliminating some of sets showing irregular correspondences.

Table 8. Revised sibilant and affricate series

Environment	Ch	PTn	PMZ	# Sets
_#	c	—	*c	2
_V̲ / _Vʔ	c’ / č’	*c	*c	5
X_X	s	*s	*s	6
#_	č’	*s	*s	5
	š	*š	*s	11
_#	∅	*s / š	*s	3

The cognate sets supporting these correspondences are exemplified in Table 9.

Table 9. Sets supporting the revised sibilant and affricate series

	Environment / Gloss	Ch	PTn	PMZ
(2 sets)	_#	c	—	*c
	‘herb, plant’	muhc	—	*ʔuhc
	‘stick, adhere’	ʔuc(te-)	—	*ʔoːʔc
(5 sets)	_V̲ / _Vʔ	c’ / č’	*c	*c
	‘bug’	č’iʔiš ~ čiːš	*cḭːš ~ cḭːs	*cis(ik)
	‘talk, say’	č’aːn-	—	*caːm | cam
	‘ant’	(c’iːsc’ik’ah)c’ik’ut	—	*cukut
(6 sets)	X_X	s	*s	*s
	‘maize’	k’as(ma)	*kḭ́s(pa̰)	*(ʔɨ)ks(i)
	‘type of cane’	(piya) waːsi(miš)	—	*waːsuk
	‘back of, under’	his-	—	*hɨs
(4 sets)	#_	č’	*s	*s

	‘sour’	č’am-	*sṷ́ːn	*šun | suːn
	‘young woman’	kiča, kici, kič	—	*kisa(y)
	‘oak’	čuh(ču)	—	*soho
(9 sets)		š	*š	*s
	‘shell corn’	kuːš(pa)	*kúšḭ	*(ʔɬ)ks
	‘fish’	makš	—	*ʔaks(a)
	2.PRO	was	*wḭš	—
(3 sets)	_#	∅	*s / š	*s
	‘night, dark’	č’i(ma)	*cḭːs	*cis
	‘break’	toh-	*tukša	—
	‘spread apart’	weːk(te-)	—	*waʔks
(1 set)		h	*s	*s
	‘boil’	šuh(t-) / šuš(t-)	—	*soːs

The correspondences for PTz-Ch *c and *s once again show a glottalization of the Ch consonant in the environment of /_V̰/ or /_Vʔ/. PTz-Ch *s appears to have stayed /s/ intervocalically, while PTz-Ch *š remained /š/ in Ch and PTn and became *s in PMZ.

4.1.3. Glottal Consonants

BWB propose only one glottal series, shown in Table 10, although they also posit a PCh-Tz *h (unattested in any of the daughter languages) in order to explain differences in vowel length. By comparison, Table 11 shows the revised sound correspondences, where I include a series for both *h and *ʔ. Table 12 shows examples of the cognate sets underlying them.

Table 10. Glottal series proposed by BWB

Ch	PTn	PMZ	Sets
∅	∅	*ʔ	(18 sets)

Table 11. Revised series of glottal consonants

Environment	Ch	PTn	PMZ	Sets
V_C	∅	∅	*ʔ	(14 sets)
#_	ʔ	—	*ʔ	(6 sets)
V_V	ʔ	∅	∅	(1 set)
	h	—	*h	(4 sets)
???	h	—	*ʔ	(1 set)
	h	*x	*h	(1 set)

Table 12. Sets underlying the revised series of glottal consonants

	Environment / Gloss	Ch	PTn	PMZ
(14 sets)	V_C	∅	∅	*ʔ
	‘grind’	kih(ci-)	*(S)kití	*kɨːʔt
	‘squirrel’	kuː(mit)	*(š)kú(tɨ̲)	*kuʔ(y)
	‘become wet, drip’	t’eyk(te-)	*(S)tɑ̲x	*taʔks
(6 sets)	#_	ʔ	—	*ʔ
	‘see, look at’	ʔam-	—	*ʔaʔm
	‘see, seek’	ʔiš(i-)	—	*ʔis
	‘go and return, arrive’	ʔuy-	—	*ʔoy
(1 set)	V_V	ʔ	∅	∅
	‘bug’	č’iʔiš ~ čiːš	*cɨ̲ːš ~ cɨ̲ːs	*cis(ik)
(4 sets)		h	—	*h
	‘herb, plant’	muhc	—	*ʔuhc
	‘oak’	čuh(ču)	—	*soho
	‘draw, write’	haːkš(te-)	—	*haːyʔ
(1 set)	???	h	—	*ʔ
	‘good, well’	huy-	—	*ʔoyV
(1 set)		h	*x	*h
	‘inside’	huh-	*xuː	*hoh

The segments /h/ and /ʔ/ appear consistently as such in all three families, although although most of the languages in these families (including Chitimacha) have a

phonotactic constraint that words must begin with a consonant, so it is not clear whether this initial segment in each of these languages is necessarily connected to that of other languages. The loss of any initial segment might have triggered the insertion of an epenthetic /ʔ/ in its place, so this may be a parallel development due to shared phonotactics. Intervocalically, glottals in all three families have a tendency to delete, while the vowels on either side merge and undergo compensatory lengthening. There are still synchronic reflexes of this process in various of the modern languages, including Chitimacha, which shows regular alternations like *čiʔiš ~ čiːš* 'leaf' and *poʔ ~ poː* 'herb, grass'.

When in the environment of /V_C/, the *ʔ appears to have been retained in PMZ, triggered laryngealization in PTn (though the PTn data is sparse for this sets of correspondences, as Table 12 makes evident), and had one of several different effects in Chitimacha: glottalization of the consonant preceding the vowel (or following in the case of some sonorants and long vowels, cf. the set Ch *paˀn* : PTn *pa(pą́)* : PMZ *poʔo*), compensatory lengthening (sometimes followed by the insertion of an offglide; cf. the set for 'become wet'), or in some cases the glottal stop simply disappears. So once again we see some general trends surrounding the effect of glottals, but not the systematicity we would hope for in proving a genetic relationship.

4.1.4. Sonorants

The sonorants reconstructed by BWB are given in Table 13. They also suggest a rule whereby PCh-Tz *m → ∅ / #__, which I retain in the revised correspondences as shown in Table 14.

Table 13. BWB's sonorant correspondences

Ch	PTn	PMZ	# Sets
m	*m	*m	10
n	*n	*n	11
w	*w	*w	7
y	*t	*y	5

Table 14. Revised sonorant correspondences

Environment	Ch	PTn	PMZ	# Sets
	w	*l	*ʔ	1
	m	—	*m	3
#_	m	—	*ʔ	3
	n	*n	*n	4
_#	m	*n	*n	2
#_	∅	*n	*n	2
	w	*w	*w	7
	y	—	*y	5
V_V	y	∅	—	2

The revised data show that the sonorant correspondences are significantly more complex than proposed by BWB, though it is possible to find likely conditioning environments for each of the different correspondences.

Representative cognate sets for the above correspondenes are provided in Table 15.

Table 15. Sets underlying the sonorant correspondences

Environment / Gloss	Ch	PTn	PMZ
	m	—	*m
'sprout, stem'	kaː(mu)	—	*ka(ma)
'bend down'	kam(te-)	—	*kɨʔmɨ
'see, look at'	ʔam-	—	*ʔaʔm
#_	m	—	*ʔ
'fish'	makš	—	*ʔaks(a)

‘bottom’	mak(ta)	—	*ʔoːk
‘herb, plant’	muhc	—	*ʔuhc
	n	*n	*n
‘whip, beat (in conflict)’	nakš	*naq ~ nik	*naks
‘person’	pan(š)	*(laː)pana	*pɨn
‘red’	piːh(ne), pin	*(S)pi̱n	*(nɨʔ)pin
_#	n	—	*m
‘talk, say’	č’aːn-	—	*caːm | cam
_#	m	*n	*n
‘flower’	šaːmu	*ša̱ná	—
‘sour’	č’am-	*sú̱ːn	*šun | suːn
_#	∅	*n	—
‘rain, water’	kaya	*(š)kaːn	—
‘belly, side’	pay(e)	*paːn	—
	w	*w	*w
‘speak, say’	weʔn	*wan	*wan
2.PRO	was	*wi̱š	—
‘adze, hew’	wač(unkšt-)	*wa̱tá̱	*wɨt
	y	—	*y
‘cry’	yeh(t-)	—	*yaʔs
‘swing’	tahy(te-)	—	*tɨːʔy
‘go and return, arrive’	ʔuy-	—	*ʔoy
V_V	y	∅	—
‘rain, water’	kaya	*(š)kaːn	—
‘belly, side’	pay(e)	*paːn	—

The segment *m appears to have been lost in the Totozoquean languages word-initially, and otherwise remained /m/. A small piece of synchronic evidence from Chitimacha supports this pattern as well: cf. *ʔokun* ‘shoulder’ and *mokun* ‘knee’.

The segment *n remains /n/ in most environments, but shows a great deal of variation word-finally, with an /m/ appearing inconsistently in one language family or

the other or simply deleting, similar to the trend with *k seen earlier. In both cases, the correspondences for these consonants are otherwise quite robust, and the variation occurs just word-finally.

Both *w and *y appear to have deleted intervocalically in Chitimacha and Proto-Totonacan, triggering a merger of the surrounding vowels with compensatory lengthening. This is the same process as seen earlier with *ʔ, and again there is ample synchronic evidence in Chitimacha for this pattern, seen in allomorph pairs like *now-* ~ *no:-* 'be ripe' and *k'ay-* ~ *k'a-* 'be not' (only *w triggers compensatory lengthening). Thus we seem to have a more general process whereby glides and glottals deleted intervocalically and triggered compensatory lengthening in Ch, while triggering just compensatory lengthening in PTn, and deleting entirely in PMZ.

The two sets reconstructing to *y are technically compatible and could possibly be combined into a single correspondence, but this seems more due to coincidental gaps in the data more than anything else. It is also unlikely considering that both *w and *ʔ have separate correspondences for intervocalic position and *y seems to pattern with them. Thus I leave the two *y correspondences separate.

4.2. Vowels

The vowel correspondences in the data are considerably less clean than the consonant correspondences, though some trends emerge. BWB themselves do not attempt to include details of stress or length in their reconstruction of PCh-Tz due to lack of systematicity. My revisions to the cognate sets based on improved data from Chitimacha do not appear to have improved this situation. Table 16 shows the full set of

vowel correspondences posited by BWB, and Table 17 provides the revised vowel correspondences.

Table 16. Vowel correspondences proposed by BWB

Ch	PTn	PMZ	Sets
a	i	i	2
a	i̯	i	2
a	a	ɨ	3
a	a̯	ɨ	2
a	i	ɨ	3
a	i̯	ɨ	3
a	a	a	19
a	a̯	a	13
e	a	a	2
e	a̯	a	6
i	*i	*i	9
i	*i̯	*i	7
i	i	ɨ	4
i	i̯	ɨ	3
u	a	a	4
u	a̯	a	4
u	u	u	5
u	u̯	u	2
a	u	u	1
a	u̯	u	2
u	u	o	8
u	u̯	o	4
o	u	o	3
o	u̯	o	2
a	a	o	4
a	a̯	o	4

Table 17. Revised vowel correspondences

Environment / Gloss	Ch	PTn	PMZ
	a	*a	*a
_ʔ	aː	*a	*a
+ glottal	aː	*a̰	*aː
+ glottal	a	*a	*aː
_[+ glottal/glide]	a	*aː	*a
_N	e	*a	*a
_ʔC	e	*a̰	*a
_ʔ#	e	*a̰ː	—
	i	*i	*i
_[+ glottal/glide]	i	*ḭ(ː)	*i
	o	*u	*oː
q_	o	*a̰	—
	u	*u	*u
_ʔ	uː	*u	*u
_[+ glottal/glide]	u	*uː	*u
C_[+ glottal/glide]	u	*u	*o(ː)

What Table 17 illustrates is that, while the data do cluster around a variety of correspondences, their conditioning environments are not at all clear, and overlap greatly. While BWB again posit different reconstructed segments for each correspondence and I instead suggest possible conditioning environments, the overall situation is approximately the same for each table – that is, messy and complicated.

Table 18 begins to look at these correspondences in more detail, starting with the most well-supported correspondences in the vowels, where either same segment appears

in all three languages, or a long vowel appears in Ch and a short vowel in PTn and PMZ.

Table 18. Vowel correspondences with the same vowel in all three families

Environment / Gloss	Ch	PTn	PMZ
	a	*a	*a
‘whip, beat (in conflict)’	nakš	*naq	*naks
‘see, look at’	ʔam-	—	*ʔaʔm
‘young woman’	kiča, kici, kič	—	*kiša(y)
	i	*i	*i
‘sharp point, blade’	sit(’ik’i)	*sit	—
‘see, seek’	ʔiš(i-)	—	*ʔis
‘young woman’	kiča, kici, kič	—	*kiša(y)
	u	*u	*u
‘coals, embers’	ku(ps)	*(š)qu	—
‘herb, plant’	muhc	—	*ʔuhc
‘ants’	(c’iːsc’ik’ah)c’ik’ut	—	*cukut
_ʔ	aː	*a	*a
‘wing, arm, bone’	paːn(t’in)	*paqa	*pak
‘split’	p’aːp(te-)	—	*pap(s)
‘stem of plant’	kaːp(ti)	—	*kap(e)
_ʔ	uː	*u	*u
‘shell corn’	kuːš(pa)	*kúšn̰	*(ʔɨ)ks
‘squirrel’	kuː(mit)	*(š)kú(tn̰)	*kuʔ(y)

The two *a correspondences in the table, where Ch shows an /a/ and /aː/ respectively, are both rather robust with 5 and 7 supporting cognates, and at first glance do not seem to have any clear distinguishing environment. However, this may again be due to a simple gap in the data. As is evidenced by the table, there are few attested PTn forms (due simply to the lack of extensive documentation for many of the Totonacan langauges). Since glottals, as seen in 4.1.3, tend to trigger compensatory vowel

lengthening in Ch and often have no reflex in PMZ, it is likely that this Ch /Vː/ : PTn *V : PMZ *V correspondence appears in the environment of /_ʔ/, which we would expect to see as laryngealization in the PTn. However, no such laryngealization is present in the available PTn forms. Thus the different vowel lengths in the Ch forms do not have a clear conditioning environment.

Allowing for greater deviation, however, a familiar trend emerges where a long vowel occurs in one family in the environment of a glottal stop, ejective consonant, or laryngeal vowel in one of the other families. Again, there is little consistency as to which family exhibits which trait, but these patterns are at least somewhat consistent in that these features co-occur regularly. Table 19 shows the correspondences where the same segment occurs across all three languages in the environment of a glottal feature.

Table 19. Vowel correspondences where same segment occurs in all three families, in the environment of a glottal

Environment / Gloss	Ch	PTn	PMZ
+ glottal	aː	*a̱	*aː
‘mat’	tʼaːna	*(š)ta̱qa(t)	*taːkʔ
‘flower’	šaː(mu) | šaːmu	*ša̱ná	—
‘chew’	čʼakʼ(umt-)	*ca̱qá	—
+ glottal	a	*a	*aː
‘mat’	tʼaːna	*(š)ta̱qa(t)	*taːkʔ
‘see, look at’	ʔam-	—	*ʔaʔm

A much clearer conditioning environment can be seen in the correspondences in Table 20, where the PTn forms show compensatory lengthening in response to the deletion of a glottal or glide consonant. In the set for ‘rain’, for example, the

intervocalic *y has deleted in PTn and triggered compensatory lengthening in the surrounding vowel. Likewise the deletion of the word-final *h in the PTn form for ‘inside’ has triggered lengthening of the preceding vowel.

Table 20. Vowel correspondences showing compensatory lengthening

Environment / Gloss	Ch	PTn	PMZ
_[+glottal/glide]	a	*aː	*a
‘rain, water’	kaya	*(š)kaːn	—
‘belly, side’	pay(e)	*paːn	—
_[+glottal/glide]	i	*i̲(ː)	*i
‘night, dark’	č’i(ma)	*ci̲ːs	*cis
‘bug’	č’iʔiš ~ čiːš	*ci̲ːš ~ ci̲ːs	*cis(ik)
‘red’	piːh(ne), pin	*(S)pi̲n	*(niʔ)pin
_[+glottal/glide]	u	*uː	*u
‘inside’	huh-	*xuː	*hoh

There are also several correspondences containing /e/ or /o/ shown in Table 21. Note that these correspondences also follow the trends just discussed, whereby in the set for ‘become wet’ the Ch form shows an off-glide in response to delete of a glottal stop while PTn form shows laryngealization, and in the set for ‘be lying’ the PTn form shows compensatory lengthening in the same environment but word-finally. Table 21 also shows a correspondence whereby the PMZ shows *o(ː) while the other forms show *u. A possible conditioning environment explaining why for why *o(ː) appears instead of *u in the PMZ forms is that these words are all CVC(V) ending in a glottal consonant or glide.

Table 21. Vowel correspondences with /e/ or /o/

Environment / Gloss	Ch	PTn	PMZ
_N	e	*a	*a
‘speak, say’	we^{ʔ}n	*wan	*wan
_ʔC	e	*a̱	*a
‘fold’	k’ep(tki-)	*qa̱p(s)	—
‘become wet, drip’	t’eyk(te-)	*(S)ta̱x	*taʔks
_ʔ#	e	*a̱ː	—
‘be lying’	pe- ~ peʔ ~ pe(h)	*pa̱ː	—
	o	*u	*oː
‘cloud’	pok(u)	*puq(ɬni)	—
‘cut roughly’	pok(št-)	—	*poːʔ(t)
‘break’	toh-	*tukša	—
q_	o	*a̱	—
‘leached corn’	now | huwo	*qa̱waː	—
C_[+glottal/glide]	u	*u	*o(ː)
‘head’	kut	*kuk	*ko ~ koʔ
‘cook by boiling’	šuh(t-) | šuš(t-)	—	*soːs
‘oak’	čuh(ču)	—	*soho
‘go and return, arrive’	ʔuy-	—	*ʔoy
‘stick, adhere’	ʔuc(te-) | ʔu(cte-)	—	*ʔoːʔc
‘good, well’	huy-	—	*ʔoyV

The above trends exhaust any potential regularity that exists for the vowels. Beyond these, it is difficult to tease out any regularity of correspondences. Table 22 shows the remaining ‘residue’ in the vowel correspondences. Again, many of the forms seem to undergo processes like compensatory lengthening or glottalization of a consonant in the expected environment of /Vʔ/, but the real issue with these correspondence is the segments themselves, which occur in a variety of combinations with no clear

conditioning context to distinguish them. Even deciding what correspondence should reconstruct to what vowel is not at all clear.

Table 22. Miscellaneous vowel correspondences

Environment / Gloss	Ch	PTn	PMZ
???	a	—	*eː
‘grow, grown’	yaʔ-	—	*yeːʔk
???	a	*a	*oː
‘bottom’	mak(ta)	—	*ʔoːk
‘reed’	ʔak(t)	—	*ʔoːk(win)
???	a	*ṷː	*u
‘cold’	č’ak(i)	—	*suk
‘sour’	č’am-	*sṷːn	*šun | suːn
???	eː	*ṷ	*a
‘pretty, handsome’	k’eːs(ik’i)	*kṷsi	—
‘spread apart’	weːk(te-)	—	*waʔks
m_ (?)	u	*a	*a
‘sprout, stem’	kaː(mu)	—	*ka(ma)
‘flower’	šaː(mu) | šaːmu	*šạ́ná	—
???	uː	*ạ	*a
‘break’	puːk(te-)	*pạq(S)	*pak
???	i	—	*u
‘ants’	(c’iːsc’ik’ah)c’ik’ut	—	*cukut
???	∅	—	*u
‘type of cane’	(piya) waːsi(miš)	—	*waːsuk | *waːšuk

While these vowel correspondences certainly lack any regularity when compared to the ones earlier in this section, none of these sets should necessarily be discarded for that reason alone, for with the exception of the particular vowel that participates in these exceptional cognates, the remainder of the segments in each of these words participates

in correspondences that are much better supported. Therefore each of the words in Table 22 should be tentatively retained.

5. Discussion

The discussion of sound correspondences in the preceding section makes clear that certain robust trends do appear in the data that suggest some type of connection between Chitimacha and the Totozoquean languages, if more so for the consonants than the vowels. For consonants, the most general trend that emerges across a number of different sets is glottalization of Chitimacha voiceless stops in the historical environment of a glottal stop. For vowels, the corollary to this trend is compensatory lengthening in response to deletion of the historic glottal stop. Sonorant and glottal segments also have a tendency to delete intervocalically, leaving a long vowel in their stead. These broader generalizations, as well as some of the more well-supported correspondences, have made it possible to reject a number of cognate sets proposed by BWB when those sets participate only marginally in these regularities.

Despite these regularities, a great deal of variation exists in the data. For any given set, this variation is not enough in itself to rule out the set as unrelated. Typically, an irregular set will participate in a number of well-established correspondences, but contain one exceptional segment that defies explaining. This was especially the case with the sets in Table 22, where a single vowel segment in each cognate was problematic, but the words were otherwise very well supported. But in some cases even the irregularity was somewhat rule-governed, such as the way *k was unpredictable in its reflex, but only in word-final position.

What to do with these seemingly contradictory types of data? If one takes regular and clearly conditioned sound correspondences as the *sine qua non* of historical reconstruction, the Chitimacha-Totozoquean data clearly fail the test. Moreover, having rejected a total of 23 of 90 cognates proposed by BWB, the remaining 67 sets of similar words do not seem like enough evidence in themselves to argue for genetic relatedness. Yet the similarities and trends in the data are also too extensive to simply ignore. Taken together, these data suggest that Chitimacha and Totozoquean are connected not by a shared genetic inheritance, but rather language contact or diffusion through intervening langauges at some point in their history.

A number of anthropological and archaeological facts give credence to this hypothesis. The Chitimacha were known for their canoeing technology and ability to navigate the waterways of the bayou region, sometimes building huge canoes that could seat up to 40 people (Swanton 1911). So it is certainly possible that the Chitimacha might have navigated along the Gulf coast and made connections with peoples farther south. More likely, as many of the words under consideration related to maize, these words may have been borrowed as part of a broader dispersion of maize agriculture out of Mesoamerica, perhaps accompanied by some population movement that would have brought maize technology with it. The sets that show a possible connection to maize are listed in Table 23.

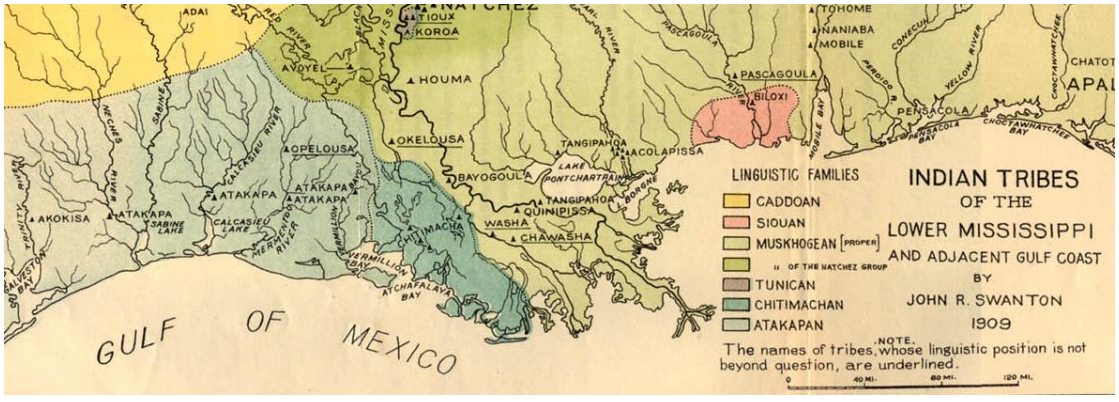
Table 23. Sets potentially related to maize

Gloss	Chitimacha	Proto-Totonacan	Proto-Mixe-Zoquean
'sprout, stem'	kaːmu	—	kama
'stem, stalk'	kapti	—	kape

‘eat, bite’	k’ušt	—	kɬːʔs
‘maize’	k’asma	kʃ̲spa̲	ʔɬksi
‘grind’	kihci	(S)kití	kɬːʔt
‘shell corn’	kuːšpa	kúši̲	ʔɬks
‘leached corn’	nowa	qa̲waː	—

Trade connections are another possibile source of diffusion. It is known that the Chitimacha once lived as far west as Grand Lake, in the territory of the Ishak (known as Atakapa in the literature; see Table 24), who themselves brought goods from Texas and Mesoamerica and traded them with the rest of the Southeast. Moreover, the Chitimacha possessed a number of cultural practices generally associated with Mesoamerica. For example, they are the only Southeastern group described in Swanton’s (1911) survey as making nixtamalized maize (hominy) using a lime solution, a practice employed in Mesoamerica but not in the U.S. Southeast, and they may have used Mesoamerican-style atlatls as well (Swanton 1938).

Table 24. Map of the historic ranges of the Atakapa (light blue) and Chitimacha (dark blue) peoples



While none of this evidence is sufficient in itself to prove a linguistic connection to Mesoamerican, the confluence of evidence between the sound correspondences,

anthropology and archaeology all suggest that *some* connection exists, but that this connection is not one of genetic relationship as Brown, Wichmann & Beck (2014) propose.

APPENDIX I

REVISIONS TO THE COGNATE SETS IN BROWN, WICHMANN & BECK (2014)

Parentheses around phonetic material indicate that the material is not considered cognate. The columns labeled 'Original' and 'Revised' show the Chitimacha (Ch), Proto-Totonacan (PTn), and Proto-Mixe-Zoquean (PMZ) forms between slashes respectively. An m-dash (—) indicates that no cognate word was found for that language. A vertical bar (|) indicates uncertainty as to which form is correct, while a tilde (~) indicates allomorphy between two forms. Chitimacha forms preceded by an asterisk (*) are internal reconstructions, and may be considered Pre-Chitimacha (PCh). Forms preceded by a double asterisk (**) are hypothetical for the purpose of discussion (i.e. forms that are predicted but not attested). Since different orthographic conventions are used in each of the datasets, I have standardized the data using an Americanist orthography.

The revisions below consist of two types: First are revisions to the Chitimacha form based on new or more in-depth data from Hieber (2013). Second are those revisions which arise from examining the sound correspondences in §4 of this paper. Those words or sets which show irregular sound correspondences have been rejected.

Gloss	Original Ch / PTn / PMZ	Revised Ch / PTn / PMZ	Supporting Evidence
'sew'	č'u(ši) / *cą̲(pą́) / —	(Rejected)	The Ch root is **č'uš* and not further analyzable.
'chew'	č'ak'u(mt) / *cą̲qą́ / —	č'ak'(umt) / *cą̲qą́ / —	*č'ak'umt* is analyzable as *č'ak+(ʔ)umt*.

‘say, speak’	č’aːm / — / cam	č’aːm / — / cam | caːm	Only the Proto-Zoque is reconstructable, to *cam*, but based on the Ch BWB posit that the PMZ would be ***caːm*.
‘night, dark’	č’i(ma) / c̲iː(s) / ci(s)	č’i(ma) / c̲iːs / cis	There is no reason to exclude /s/ from the cognate material before examining the sound correspondences.
‘bug, worm’	č’iːš / c̲iːš ~ c̲iːs / cis(ik)	č’iʔiš ~ č’iːš / c̲iːš ~ c̲iːs / cis(ik)	The Ch has an allomorph *č’iʔiš*.
‘maize’	kaːmu / — / kama	kaː(mu) / — / ka(ma)	Ch *kaːmu* appears to derive from a historic root *kaʔ* ‘sprout, stem, go up’ + *ma* PL. It is however possible that the /ma/ of PMZ *kama* is cognate to the Ch plural.
‘lime’	kupšeš / qaš(tax) / (ʔa)kas	(Rejected)	Ch kupšeš is from a root *kup ‘dark, coal’, possibly in addition to šiʔ ‘dust’ (the word ‘lime’ is often formed of words meaning ‘dust’ or ‘ash’).
‘bend down’	kam(te) / — / kɨʔm(ɨ)	kam(te) / — / kɨʔmɨ	There is no evidence for excluding the final vowel of the PMZ form from the cognate

			set without first examining the sound correspondences.
‘young woman’	kiča / — / kiša(y)	kiča, kič, kici / — / kiša(y) | kisa(y)	*kiča* has alternate forms *kič* and *kici*. Only the Proto-Mixe is reconstructable, but based on the Ch BWB posit that the PMZ is ***kisay*.
‘shoulder’	(mo)ku(n) / (cu)qu(s) / ko(sok)	(Rejected)	*mokun* relates to *ʔokun* ‘knee’ plus an unknown segment /m/ (possibly from *ni* NZR), suggesting internal morphology of *ʔoka* + *n* AZR. The cognate sets underlying the PTn and PMZ suggest that the /s/ is also relevantly cognate. Thus the comparison should be between *ʔoka* and *qus* / *kos*, which do not appear to be cognate.
‘cut’	k’aːc(t) / — / koːʔc	(Rejected)	The Ch /c/ is not cognate material, nor do the semantics of the root match the PMZ. *k’aːct* < *kaʔ*

			‘move circularly’ + *c* ‘touch’ + *t* TRZR
‘rain’	kay(a) / (š)kaːn / —	kaya / (š)kaːn / —	No internal evidence suggests an internal morpheme boundary or the addition of a final vowel for *kaya*.
‘be cold’	k’as(te) / kąs(ní) / —	k’as(te) / kas(ní) | kąs(ní) / —	Only the Proto-Tepehua is reconstructable. BWB posit ***kąsní* for PTn based on the Ch.
‘grind’	kihci / (S)kití / kɨʔt	kih(ci) / (S)kití / kɨʔt	/c/ is not cognate material. *kihci* < *kih* ‘pull’ + *c* ‘touch’ + *i* VZR
‘shell corn’	kuːs(pa) / kúšį / (ʔɨ)ks	kuːš(pa) / kúšį / (ʔɨ)ks	The Ch root is documented as *kuːš* rather than *kuːs*.
‘bone’	kaci / (lu)kuti / —	(Rejected)	This set does not participate in the sound correspondences in §4 and must be rejected.
‘pretty, beautiful’	k’eːsi(k’i) / kų̨si / —	k’eːs(ik’i) / kų̨si / —	The historic root of *k’eːsik’i* is likely **k’eːs*.
‘louse’	cat / (s)káːta / (ʔa)wat	cat / (s)káːta / ʔawat	BWB posit *k^{y}w > /c:k:w/, but I find *k > /c:k:ʔ/ more plausible. The /áː/ of *skáːta* is then explainable as compensatory vowel

			lengthening following the loss of /w/ in *ʔawat*. However, this set does not participate in regular sound correspondences, so should be rejected.
‘crop, field’	nowa / qa̱waː / —	now | huwo / qa̱waː / —	The Ch is either from *now* ‘be ripe, cooked’ or related to *huwo* ‘field’.
‘outside’	nuk / (š)kuk / (š)quq	(Rejected)	The semantics of *nuk* ‘outside’ do not seem sufficiently close to *(š)quq* ‘row (of plants), furrow’ or *kuk* ‘middle’.
DEM	man(ki) / a̱n / —	(Rejected)	The /n/ of Ch *manki* belongs to the enclitic *=ki* with its allomorph *=nki*. The form of the demonstrative is *ma*.
‘bush’	muhc / — / ʔuc	muhc / — / ʔuhc	Only the Proto-Mixe is reconstructable, to *ʔuhc*, but BWB posit ***ʔuc* for the PMZ. However, every form in the PMZ cognate set shows an /h/.
‘write’	naːk(št) / — / noki	(Rejected)	*naːkšt* derives from *ni* DTRZR + *haːkš* ‘write’ + *t* TRZR.

‘break up’	puːk(te) / pa̱q(S) / pak	puːk(te) / pa̱q(S) / pak | paːk	Only the Proto-Zoque is reconstructable, to *pak*, but BWB posit ***paːk* for PMZ.
‘wing, bone’	paːn(t’in) / paq(a) / pak	paːn(t’in) / paqa / pak	All the Totonacan forms show a final vowel (/a/ or /e/) for *paqa*. I know of no evidence that this final vowel is morphologically or historically separable.
‘pile’	pac(t) / pi̱nk(š) / piw	(Rejected)	pact < *pa* ‘pile, gather’ + *c* ‘touch’ + *t* TRZR.
‘red’	pin / (S)pi̱n / (nɨʔ)pin	piːh(ne) | pin / (S)pi̱n / (nɨʔ)pin	Ch words for ‘red’ all derive from *piːhne*, which is itself composed of *piːh* + *ne*. *pinun* < *piniwa* + *n* < *piihne* + *wa*; *pinika* < *piihne* + *ka*. It is still possible that the Ch *pin-* forms are cognate to the Totozoquean though.
‘light’	(kap)pič / (s)pi̱t / pit	(Rejected)	This form is actually *kap ʔič*, and *ʔič* is further composed of *ʔi* + *č*. Swanton misheard a sequence of /p#ʔ/ as a geminate /pp/.

‘cane’	piya / — / pɨːyV(n)	piya / — / pɨːyV(n) | pɨːya	Only the Proto-Mixe is reconstructable, but BWB posit ***pɨːya* for PMZ.
‘cut’	po(kšt) / — / poːʔ(t)	pok(št) / — / poːʔ(t)	The Ch root is **pokš* and most likely unanalyzable, though there is a small but unlikely possibility that it is related to a historic root **pok* ‘seize’.
‘moon, month’	pa(nʔ) / pa(pą́) / po(yʔa)	(Rejected)	Ch *panʔ* is unanalyzable, and I believe a better reconstruction for PMZ is **poʔo* (the Zoque forms appear to be composed of the Mixe **poʔo* + another morpheme **ya*). The Ch /ʔ/ is actually preglottalization. This set does not participate in the correspondences laid out in §4 and should be rejected.
‘sharp’	siťi(k’i) / sit / —	sit(’ik’i) / sit / —	The Ch derives from a historic root **sit* ‘sharp’.
‘flower’	šaːmu / šą́ná / —	šaː(mu) / šą́ná / —	*šaːmu* derives from *šaʔa* + *ma* (cf. *šaʔap*

			‘May pop’). It is possible that the bimorphemic ša:mu is still cognate to the Totonacan forms though.
‘boil’	šuš(t) / — / so:s	šuh(t) | šuš(t) / — / so:s	The most fluent Ch speaker in the available documentation used *šuht*, while the less fluent one used *šušt*.
‘sun’	č’a(ʔa) / — / sɨw	(Rejected)	*č’aʔa* is not further analyzable. This set does not participate in the correspondences outlined in §4 and should be rejected.
‘sour’	č’am(i) / šu̱:n / šun	č’am / šu̱:n / šun | su:n	The Ch form is actually *č’am*. Only the Proto-Mixe is reconstructable, but BWB posit ***su:n* for PMZ.
‘be wet’	t’e(ykte) / (S)ta̱(x) / taʔk(s)	t’eyk(te) / (S)ta̱x / taʔks	Ch **t’eyk* cannot be analyzed into separate morphemes. The final sibilants of the Totozoquean forms are likely cognate.
‘sibling’	ta(t’in) / (s)tá(nku̱) / tɨ(wɨ)	(Rejected)	I can see no evidence in the Totozoquean forms underlying these

			cognates that suggests internal morpheme boundaries for these words. The Ch root is most likely **tat*, though **ta* is possible. In addition, this set does not participate in the regular sound correspondences outlined in §4, and so must be rejected.
‘swing’	tah(yte) / (S)ti(wí) / tɨːʔ(y)	tahy(te) / (S)tiwí / tɨːʔy	The Ch form cannot be divided into separate morphemes, nor is there evidence for internal boundaries within the Totozoquean. Additionally, /y/ : /w/ : /y/ is a very plausible correspondence.
‘break’	to(h) / tu(kša) / —	toh / tukša / —	There is no evidence that any of the phonetic material in these forms should not be considered cognate. The words are all unanalyzable.
‘touch, tickle’	čikin(e) / — / tikin	(Rejected)	The Ch breaks down as čiki + ne. This set does not participate in the regular

			correspondences in §4 and must be rejected.
‘all’	nak / — / tɨk	(Rejected)	This set does not participate in the sound correspondences in §4 and must be rejected.
‘fruit’	nan(u) / tin̲ / tɨm	(Rejected)	There is no basis for excluding /u/ as relevant to the cognate set without first examining the potential sound correspondences. However, this set does not participate in the regular correspondences outlined in §4 and must be rejected.
‘land, earth’	ne(yʔ) / ti̲(ya̲t) / —	(Rejected)	Neither Ch neyʔ nor PTn ti̲ya̲t show evidence of a morpheme boundary before /y/. The Ch /ʔ/ is actually a preglottalization. This set does not participate in the sound correspondences laid out in §4 and must be rejected.

‘league with an animal spirit’	nek(ma) / tiːk(ú̲ː) / —	(Rejected)	The PTn /ú̲ː/ should not be excluded until the sound correspondences are examined. However, upon examination, this set turns out not to participate in the regular correspondences outlined in §4, and so should be rejected.
‘small’	(na)c’i(k’i) / ci̲(nk) / —	(Rejected)	The root of the Ch form is a monomorphemic **nac*, and the PTn shows no signs of being internally divisible.
‘young cane reed’	waːs(imiš) / — / waːš(uk)	waːs(imiš) / — / waːšuk | waːsuk	The PMZ shows no evidence that the final /uk/ is morphologically separable. ***waːšuk* is a reconstruction of the PMZ based on the Ch, but only the Proto-Mixe can actually be reconstructed, to *waːsuk*.
‘wrap, bundle’	waːc(t) / — / wat	waː(ct) | waːt / — / wat	The Ch most likely derives from *waʔa* ‘another’ + *c* ‘touch’ + *t* TRZR, although

			the root might also be **wa:t*, with the final /t/ obscured by the *-c* suffix.
‘tongue’	wen(ʔ) / wan / wan	weʔn / wan / wan	The final /ʔ/ of the Ch cannot be separated from the root, and is in fact a preglottalization.
‘adze, hew’	waču(ŋkšt) / wą̱tą̱ / wɨt	wač(unkšt) / wą̱tá / wɨt	Based on analogous verbs with similar phonotactics, the morpheme boundaries for this word are probably *wača-nk-š-t*, though the meaning of the root is unknown.
‘write’	haː(kšte) / — / haː(yʔ)	haːkš(te) / — / haːyʔ	The Ch root *haːkš* is unanalyzable, and it is quite plausible that the Ch /k/ might be cognate to the PMZ /ʔ/.
‘cry out’	ye(ht) / tą̱(sá) / yaʔ(s)	(Rejected)	There is no evidence suggesting that the final /h, s/ of these roots should be treated separately from the rest of the root. This set does not participate in the regular correspondences in §4

			and should be rejected.
‘grow’	yaːkʾi / (s)tak / yeːʔk	(Rejected)	Ch yaːkʾi is a synchronically analyzable word consisting of yaʔ ‘grow’ + kʾi AZR. This set does not participate in the correspondences presented in §4 and so must be rejected.
‘stick something’	(ʔ)uc(te) / — / ʔoːʔc	ʔu(cte) / — / ʔoːʔc	Omitting the initial glottal of the Ch from this cognate set appears to be a formatting error in BWB. The final /c/ is most likely the *-c* ‘touch’ suffix, given the meaning of the verb, but could plausibly be cognate to the PMZ /c/ regardless.
‘cheek’	(w)aːku / (l)aka / (ʔ)aːka	(Rejected)	There is no evidence that the initial segments in this set should not be taken into account when considering cognancy. As such, this is a highly irregular

			correspondence, and the set should be rejected.
‘house’	(h)ana / (č)aqa / (t)ɨk	(Rejected)	There is no basis for excluding the initial consonants from this cognate set. Additionally, this set does not participate in the regular correspondences outlined in §4, and must be rejected.
‘good’	(h)uyi / — / (ʔ)oyV	huy / — / ʔoyV	The root of the Ch form is *huy*, and there is no evidence that suggests the initial consonants of these words could not be cognate.

REFERENCES

Brown, Cecil H., David Beck, Grzegorz Kondrak, James K. Watters & Søren Wichmann. 2011. Totozoquean. *International Journal of American Linguistics* 77(3). 323–372.

Brown, Cecil H., Søren Wichmann & David Beck. 2014. Chitimacha: A Mesoamerican language in the Lower Mississippi Valley. *International Journal of American Linguistics* 80(4).

Duponceau, Peter S. 1819. Report of the Corresponding Secretary to the Committee, of his Progress in the Investigation committed to him of the General Character and Forms of the Languages of the American Indians. *Transactions of the Historical & Literary Committee of the American Philosophical Society*, vol. 1, xvii–l. Philadelphia: American Philosophical Society.

Duralde, Martin. 1802. Vocabulaire de la langue des Chetimachas.

Gallatin, Albert. 1836. *A synopsis of the Indian tribes within the United States East of the Rocky Mountains, and in the British and Russian Possessions in North America*. (Transactions & Collections of the North American Antiquarian Society 2).

Gursky, Karl-Heinz. 1969. A lexical comparison of the Atakapa, Chitimacha, and Tunica languages. *International Journal of American Linguistics* 35(2). 83–107.

Haas, Mary R. 1951. The Proto-Gulf word for "water" (with notes on Siouan-Yuchi). *International Journal of American Linguistics* 17(2). 71–79.

Haas, Mary R. 1952. The Proto-Gulf word for "land" (with a note on Proto-Siouan). *International Journal of American Linguistics* 18(4). 238–240.

Hieber, Daniel W. Semantic alignment in Chitimacha.

Hieber, Daniel W. 2013. *A dictionary of Chitimacha (Scholar's Edition)*. Charenton, LA: Chitimacha Tribe of Louisiana.

Iannucci, David J. 2009. Aspects of Chitimacha phonology. University of Utah.

Kaufman, Terrence S. 1962. Mixe-Zoque subgroups and the position of Tapachulteco. *Actas y memorias del XXV Congreso Internacional de Americanistas, 1962, Vol. 2*, 403–411. México, D.F.

Kaufman, Terrence S. 1963. Mixe-Zoque diachronic studies.

Kondrak, Grzegorz, David Beck & Philip Dilts. 2007. Creating a comparative dictionary of Totonac-Tepehua. In John Nerbonne, T. Mark Ellison & Grzegorz Kondrak (eds.), *Computing and historical phonology: Proceedings of the Ninth Annual Meeting of the ACL Special Internet Group in Computational Morphology and Phonology*. Prague: Association for Computational Linguistics.

Mithun, Marianne. Linguistic areas of Native North America. In Raymond Hickey (ed.), *Handbook of areal linguistics*. Cambridge: Cambridge University Press.

Munro, Pamela. 1994. Gulf and Yuki-Gulf. *Anthropological Linguistics* 36(2). 125–222.

Nordell, Norman. 1962. On the status of Popoluca in Zoque-Mixe. *International Journal of American Linguistics* 28(2). 146–149.

Swadesh, Morris. 1933. Chitimacha verbs of derogatory or abusive connotation with parallels from European languages. *Language* 9(2). 192–201.

Swadesh, Morris. 1934. The phonetics of Chitimacha. *Language* 10(4). 345–362.

Swadesh, Morris. 1946a. Phonologic formulas for Atakapa-Chitimacha. *International Journal of American Linguistics* 12(3).

Swadesh, Morris. 1946b. Chitimacha. In Harry Hoijer (ed.), *Linguistic structures of Native America*, 312–336. ((Viking Fund Publications in Anthropology 6) 6). New York: Viking.

Swadesh, Morris. 1953. *Chitimacha grammar, texts and vocabulary*. Philadelphia, PA: American Philosophical Society.

Swanton, John R. 1917. Unclassified languages of the southeast. *International Journal of American Linguistics* 1(1). 47–49.

Swanton, John R. 1938. Historic use of the spear-thrower in southeastern North America. *American Antiquity* 3(4). 356–358.

Swanton, John Reed. 1911. *Indian tribes of the Lower Mississippi Valley and adjacent coast of the Gulf of Mexico*. (Bureau of American Ethnology Bulletins 43). Washington, D.C.: Smithsonian Instiutiion.

Swanton, John Reed. 1919. *A structural and lexical comparison of the Tunica, Chitimacha, and Atakapa languages*. (Bureau of American Ethnology Bulletins 68). Washington, D.C.: Smithsonian Institution.

Thomas, Norman D. 1974. *The linguistic, geographic and demographic position of the Zoque of Southern Mexico. Papers of the New World Archaeological Foundation*. (Papers of the New World Archaeological Foundation 36). Provo: Brigham Young University.

Vater, Johann Severin. 1821. *Analekten der Sprachenkunde*. . Vol. 2. Leipzig: Dykschen Buchhandlung.

Wichmann, Søren. 1994. Mixe-Zoquean linguistics: A status report. In Davis Bartholomew, Yolanda Lastra & Leonardo Manrique (eds.), *Panorama de los estudios de las lenguas indígenas de México*, vol. 1, 193–267. (Biblioteca Abya-Yala 16). Quito: Abya-Yala.

Wichmann, Søren. 1995. *The relationship among the Mixe-Zoquean languages of Mexico*. (Studies in Indigenous Languages of the Americas). Salt Lake City: University of Utah Press.

Wonderly, William L. 1949. Some Zoquean phonemic and morphophonemic correspondences. *International Journal of American Linguistics* 15(1). 1–11.