Further Analysis of Lexical and Morphological Comparanda Pertaining to the Proto-Mayan-Mijesokean Hypothesis

David F. Mora-Marín
davidmm@unc.edu
UNC-Chapel Hill

December 1, 2018
December 18, 2018
January 1, 2019
April 20, 2019
June 19, 2021
May 7, 2023

Abstract
This paper follows up on several issues in Mora-Marín (2016) that were suggested for further research in connection with the Proto-Mayan-Mijesokean Hypothesis. The paper focuses on the nature of comparanda that are phonologically and semantically similar but which were excluded from the lexical datasets in Mora-Marín (2016) due to a number of potential confounders—only one recurring correspondence, possible onomatopoeia, seemingly irregular correspondences, unmatched segments. A few are argued to be viable as potential cognates, others more likely to be loanwords, and still others must remain excluded. Additional grammatical morphemes are compared, and the question of cognacy involving not only morphemes but also paradigmatic patterns is discussed. The most important result is the expansion of the number of occurrences of previously identified sound correspondences, as well as a few new ones.

1. Introduction

This paper revisits the test of the Proto-Mayan-Mijesokean (pMaMiso) hypothesis advanced in Mora-Marín (2016), which proposes a remote genealogical relationship between the Mayan (Ma) and Mije-Sokean (MiSo) language families.[1] Mora-Marín (2016) established a set of regular sound correspondences that provide evidence for cognacy; these were carefully screened against the correspondences evident among the diffused etyma previously recognized by several authors (e.g. Campbell and Kaufman 1976; Justeson et al. 1985; Wichmann 1995, 1999), some of which differ significantly from the correspondences that were argued to be indicative of cognacy. In total, that author presented 78 comparanda (cf. his Addenda 1-3), among them 12 grammatical morphemes, as support for a relationship of cognacy. Based on these, he argued for 21 sound correspondences that provide strong support given their patterns of recurrence, and an additional set of 24 sound correspondences that are less supportive due to insufficient recurrence or somewhat extreme phonological distance (e.g. pMa *q' compared to pMiSo *¢). In addition, he put forth a list of 39 items exhibiting similarity but also problematic traits that were worth exploring further (cf. his Addendum 4), and a list of 5 items (cf. his Addendum 5) whose

[1] I term the hypothesis Proto-Mayan-Mijesokean to indicate, by means of the hyphens, the following hierarchical relationship: [[Mayan] [Mije-Sokean]], in which the Mijean and Sokean languages form a single Mije-Sokean unit first, which is coordinate with Mayan. In contrast, a term Proto-Mayan-Mije-Sokean would indicate a different hierarchical relationship: [[Mayan] [Mijean] [Sokean]]. Such term would thus give the impression that Mayan, Mijean, and Sokean are coordinate units, which would not be accurate.

phonological traits, wide distribution in both language families, and especially semantic domains (e.g. cultigens, technology) preclude a straightforward differentiation between cognates and archaic loans. Finally, he also compiled a list of 93 items (cf. his Addendum 6) that are very likely loans between the two language families; this list included 16 etyma that had not been previously identified in the literature cited above. These were used to extract sound correspondences typical of loanwords, with the goal of serving as a control against which to assess sound correspondences of possible cognates.

Overall, Mora-Marín concluded that the hypothesis was viable but must confront two major challenges before it can be validated: the difference in the sizes of the phonemic inventories of the two languages (Ma with 26 consonants and 5 vowels, for a total of 31 segments; MiSo with 11 consonants and 6 vowels, for a total of 17 segments), and the difficulty in distinguishing between possible cognates and archaic loans. Although the identification of new cognates and additional correspondences will no doubt improve the chances of understanding how the two languages (pMa, pMiSo) diverged so drastically, phonologically, after separation, this paper addresses the second challenge more directly—that of distinguishing between cognates and archaic loans. An important result of Mora-Marín (2016:141-143, 145, 170) was the existence of etymological doublets, instances of cognates and loans that can be traced back to the same etymon: in other words, cognates that, subsequent to the differentiation of pMaMiso, and their retention in the two respective families, experienced diffusion from one family to the other. Such etyma allow for relative assessment of phonetic distance of sound correspondences, and thus can be useful in attempting to distinguish between cognates and archaic loans.

The present paper revisits the pMaMiso hypothesis, specifically by addressing the problematic comparanda excluded by Mora-Marín (2016) from his initial assessment. It also offers further discussion of the grammatical morphemes introduced previously. Section 2 reviews prior work, as well as the methods and findings of Mora-Marín (2016). Section 3 examines the problematic lexical comparanda, and assesses their validity. Section 4 does the same for the grammatical comparanda. And section 5 offers conclusions and pointers for future research.

2. Methods and Framework

2.1. Mayan and Mije-Sokean.

Please refer to Mora-Marín (2016) and the sources cited therein for basic background to the Mayan and Mije-Sokean language families, the phonological structures of proto-Mayan and proto-Mije-Sokean, and prior attempts to compare the two. Below I only refer to details of direct relevance to the argumentation.

2.2. Constraints on Comparanda

As in Mora-Marín (2016), I apply the following theoretical and methodological constraints:

1) For Mayan, only etyma reconstructible to pMa, LpMa, and CMa are utilized; for Mije-Sokean, only etyma reconstructible to pMiSo, pMi, or pSo are utilized;
2) Strict semantic isomorphism criterion;
3) Regular sound correspondences;

4) Comparanda with $C_1VC_2(V)$… shapes must exhibit biconsonantal correspondences;
5) Allow significant phonological distance for only one of the (minimum) two consonants per etymon, with the other consonantal correspondence(s) exhibiting little or no phonological distance;
6) 3+ occurrences are preferred, but if only two occurrences for one of the consonants in the $C_1VC_2(V)$… comparison are found, the other consonant must exhibit more than two occurrences;
7) Unmatched sequences of segments are not allowed unless morphological analysis supports a morphemic breakup that makes the comparison valid following the criteria above, or unless the semantic and phonological similarity is otherwise very high;
8) *Favor* sound correspondences that do not occur in the likely loanword set; this criterion does not preclude sound correspondences present in the likely loanword set.

A major challenge to this long-range comparison, as already noted, is the drastic difference between the sound inventories of the languages. For one, pMa can be reconstructed with 26 consonants, while pMiSo only 11, fewer than half of the pMa inventory. This means that, if the two languages are related by descent from a common ancestor, pMiSo has undergone a major reduction via process of loss or merger, which inevitably would result in a situation wherein multiple pMa sounds could correspond to a single pMiSo sound; of course, one could argue that pMa may have also elaborated its sound inventory compared to a putative pMaMiso ancestor. Either way, criterion 5 above is designed specifically as a counterweight to this problem, to avoid, as much as possible, a situation where "anything goes."

The situation with the consonants, however, is not as overwhelming as one might think—not when compared to vowels. Indeed, even though pMa has 5 basic vowels and pMiSo 6, matters are complicated dramatically when one considers vowel complexity: V, VV, Vʔ, Vh for pMa; V, VV, Vʔ, VVʔ, Vh for pMiSo. This means that for pMa there are potentially 20 distinct vowel nuclei, and 30 for pMiSo. Thus, any one vowel can, in principle, have many more possible vocalic correspondences than your average consonant. This is why Mora-Marín (2016) focused on the consonants, and offered only a few observations regarding vowels comparisons. The present paper is no exception, although increasing the viable dataset of cognates, one of the objectives of this paper, will facilitate, eventually, a more detailed analysis of the vocalic correspondences. It may also be worthwhile to analyze the /ʔ/ and /h/ of so-called complex vowels (e.g. pMiSo Vʔ, VVʔ, Vh), as consonants (i.e. compare them as consonantal correspondences) instead of as part of the vowel nuclei. Doing so would yield only two types of vowels in both pMiSo and pMa, /V/ and /VV/.

Another challenge pertaining to the aforementioned asymmetry between these protolanguages bears mention. If related phylogenetically, the historical depth of the split between the Ma and MiSo families would be rather significant, at least comparable to that of Uto-Aztecan (ca. 3000 BCE) or Chibchan (3600 BCE) or even Otomangean (4000 BCE). Previously, Mora-Marín (2011, 2014) was able to obtain only 13-15 items from the Swadesh 100-word list to calculate a basic estimate for the time of separation.[2] If so, and assuming the

[2] Utilizing the standard glottochronology formula and .805 retention level per millennium (Lees 1953:117), and isolating 14 core vocabulary items from Swadesh's 100-word list, Mora-Marín (2011) obtained an estimate of 6,083 BP for the divergence of a putative pMaMiso language. No attempt to calculate a margin of error was made at the

validity of glottochronology for the sake of argument, 85-87% of the shared core vocabulary between these languages has already been lost (or the meanings shifted so drastically as to render cognates unrecognizable), and likely a greater percentage of non-core vocabulary; the chances of being able to document many examples of recurrence of individual sound correspondences will be low for some sounds, particularly sounds that are rare to begin with. This is likely to remain the case for correspondences involving marked sounds, such as Mayan *tʸ' and *q'.[3]

Grammatical morphemes are also discussed in this paper. For some, they constitute strong evidence of a historical relationship, including phylogenetic relationship, especially when two or more languages under comparison share unusual patterns or aberrancies (Meillet [1925]1958), or when they form a part of a broader pattern involving multiple cognate grammatical morphemes—what Nichols (1996) has called multidimensional paradigmaticity. Kaufman (1990:25) alludes in this fashion to "distribution" when he states that "Grammatical correspondences involving agreements in sound, function, and distribution should be found." And yet, caution must be exercised when studying grammatical morphemes, for as Campbell (1997:219-222) and Campbell and Poser (2008:189) observe, such forms exhibit two tendencies that can make proposals based heavily on such data unreliable: they tend to be rather short in length (cf. problems with comparisons of short forms); and they tend to utilize a rather small subset of the consonants and vowels present in a language—the "less marked segments." These two facts combine to make chance resemblances more common than some scholars have thought.

Consequently, I assume that if any one type of evidence is to be used on its own, such evidence must consist of regular sound correspondences. Consequently, the validity of grammatical comparanda must be ascertained on the basis of consistency with regular sound correspondences established on the basis of lexical comparanda (Kaufman 1990:20).[4] It is further necessary to acknowledge that grammatical roots are prone to the influence of analogical change (Kuryłowicz 1947), and phonological erosion, aided in great part by the process of cliticization or affixation characteristic of grammaticalization (Meillet 1958[1912]).

Finally, for now, a few words about the final *Cs clusters of pMiSo are in order. Wichmann (1995:66-67) describes the following morpheme-final consonant clusters for pMiSo: *ps, *ks, *pʔ, *tʔ, *cʔ, *kʔ, *sʔ, *hʔ, *mʔ, *nʔ, *wʔ, *yʔ. He adds that of these clusters, *sʔ, *hʔ, and *yʔ are absent after short vowels. He also discusses the case of the *ps and *ks clusters,

time. Following Lees (1953:124), and using a total number of 57 possible lexical cognates from the Primary Lexical Dataset in Mora-Marín (2016) from which the 14 core vocabulary items were selected, a standard error (7/10 confidence level) of ±438 (i.e. 6,083 ± 438) can be calculated as time-depth estimates for the separation of pMaMiso.

[3] I do not expect that a satisfactory reconstruction of the sound system of pMaMiSo will be forthcoming on the basis of cognates alone; instead, it may be necessary to utilize typological patterns rather extensively to accomplish such a task.

[4] In other words, if only one kind of evidence is to be adduced to propose genetic inheritance between two or more languages, the most reliable evidence is regular sound correspondences among lexical and grammatical morphemes. Similarities among grammatical morphemes, on their own, may not be reliable enough, but instead should be checked again previously established regular sound correspondences.

specifically. These occur with two root canons: *CVʔCC and *CVCC (Wichmann 1995:68).[5] Lastly, he hypothesizes that *ps and *ks, "the only permitted clusters [not also involving *ʔ] in pMZ morphemes," may have originated "in the same sequence, namely, -kʷVs-," with the roundedness value of the vowel "determin[ing] whether the labiovelar fell out as *p* or *k*." However, his motivation for such a hypothesis seems to lie in a single, possible parallel between MiSo and Uto-Aztecan (UA), in the form of the pMi root *hiʔkʃ 'to suffocate' and pUA *hikwɨ(sɨ) 'to breathe'. This solution is not only speculative, relying on the unproven hypothesis of a relationship between MiSo and UA, but it is also based on a likely onomatopoeic term ('to suffocate'), and, furthermore, would not account for the likelihood that there existed *ts clusters that simply became indistinguishable from *¢ (i.e. *t͡s) in root-final position.

Mora-Marín (2016:170-171) analyzes the *ks and *ps clusters of MiSo as likely indicative of frozen derivational morphology largely restricted to verbs, and raises the slight possibility of cognacy with the pMa suffix *-i-sa ~ *-e-sa 'causativizer'. This is a matter to which I return in §3.1 and §4.

2.3. Prior Results

Mora-Marín (2016) established four datasets. The first was a dataset of likely loanwords, including over 90 cases compiled primarily from Campbell and Kaufman (1976), Wichmann (1995, 1998), and his own research. He isolated the sound correspondences present among loanwords to serve as a kind of control, i.e. to determine to what extent possible cognates differed from loanwords phonologically. **Table 1** lists the sound correspondences founds among loanwords.

Table 1. Sound correspondences among loanwords.

#	Ma	MiSo	loan #
1.	p	p	4, 6, 16, 19, 24, 36, 47, 52, 55, 61, 82, 92
2.	b'	p	87, 93
3.	-p	-Ø	39
4.	t	t	5, 34, 44, 54, 84
5.	k	k	1, 9, 12, 14, 15, 16, 26, 29, 35, 36, 38, 45, 46, 48, 50, 57, 65, 69, 82, 86
6.	k	¢	85
7.	k'	k	5, 15, 62, 64, 78, 91
8.	q	k	51, 87
9.	q'	k	70, 71, 80
10.	s	s	8, 12, 31, 39, 48, 85
11.	s	¢	43
12.	s	š	4
13.	š	š	2, 33
14.	š	s	20
15.	¢	¢	7, 10, 27, 28, 35, 44, 55, 56, 61, 73

[5] Wichmann argues that the former canon typically conveys a meaning related to 'the meeting of two uniform objects', while the latter canon typically conveys a meaning related to 'pointed intrusion upon an object' (Wichmann 1995:70)

16.	¢	š	56
17.	¢’	¢	6
18.	č	s	76
19.	č	¢	3, 38, 40, 47, 65, 69
20.	č’	¢	23
21.	m	m	7, 15, 18, 20, 24, 25, 27, 31, 32, 40, 43, 45, 46, 49, 58, 70, 73, 74, 83
22.	n	n	5, 8, 53, 63, 71
23.	m	n	51
24.	(m)	ø	16
25.	l	n	62
26.	n	(m)	72
27.	n	m	64
28.	ŋ	[ŋ] /w/	83
29.	n	[ŋ] /w/	74
30.	ŋ	n	91
31.	ŋ	h	80
32.	-n	-ʔ	13
33.	-n	-ø	3, 28, 37
34.	w	w	3, 13, 17, 22, 26, 28, 30, 33, 34, 63, 86
35.	w	[-ŋ]	14
36.	-ø	-w	30
37.	y	y	9, 14, 16, 18, 25, 29, 58, 75, 76, 92
38.	x	h	23, 37, 54, 75, 78, 84
39.	x	k	85
40.	x	n	11
41.	h	h	11, 17, 63, 92
42.	h	y	19, 22
43.	-h/-ø	-y	15
44.	ʔ	ʔ	10, 13, 30, 32, 49, 53, 58, 73, 86
45.	ʔ	y	37
46.	ʔ	ø	16, 29
47.	ʔ	n	8

Mora-Marín (2016) also established a Primary Lexical Dataset for cognates abiding by the strictest constraints listed above, as well as a Grammatical Dataset that in general abides by such criteria, except, in a few cases, the biconsonantal correspondence criterion. **Table 2** lists the sound correspondences isolated from these two datasets. Two of these correspondences exhibit minimal recurrence, two cases each; however, their incorporation rests on the fact that the two correspondences are parallel, *t^{y} : *t and *t^{y}’ : t. Another correspondence, *b’ : *ʔ, was argued in Mora-Marín (2016) to exhibit three examples, but is now reduced to two, as a result of a mistake in the original dataset, explained in §4.

Table 2. Sound correspondences from Primary Lexical Dataset and Grammatical Dataset.

	pMa	pMiSo	Recurrence	Reconstruction	Examples
1	p	p	7x	*p	14, 32, 45, 56, 63, 72
2	x	h	4x	*x	22, 39, 56, 83
3	l	n	4x	*l	17, 70, 71, 152
4	m	m	3x	*m	26, 58, 101
5	k	k	3x	*k	33, 41, 46
6	k’	k	4x	*k’	9, 42, 60, 89
7	q	k	9x	*q	2, 25, 26, 45, 57, 71, 77, 103, 124
8	q’	k	6x	*q’	4, 11, 13, 20, 63, 101
9	r	t	4x	*r	13, 16, 18, 57
10	t	t	4x	*t	37, 110, 120, 124
11	t^{y}	t	2x	*t^{y}	19, 24
12	t^{y}’	t	2x	*t^{y}’	32, 84
13	w	w	3x	*w	51, 122, 151
14	y	y	3x	*y	25, 33, 84
15	w	y	4x	*w	85, 122, 152
16	h	h	5x	*h	87, 125, 149, 152, 152
17	ŋ	w	3x	*ŋ	8, 59, 103
18	b’	ʔ	2x	*b’	4, 23
19	s	s	3x	*s	11, 104, 107
20	š	s	3x	*š	3, 9, 89
21	ʔ	ʔ	12x	*ʔ	16, 17, 27, 54, 59, 70, 85, 87, 103, 120, 122, 125

And last, Mora-Marín (2016) established a Secondary Lexical Dataset that abides by most of the constraints listed above, including the biconsonantal correspondence criterion, but contains a few etyma that could potentially be onomatopoeic (e.g. ‘to drink’, ‘to smash’), which nonetheless exhibit recurring correspondences consistent with those that characterize etyma in the Primary Lexical Dataset. **Table 3** lists the sound correspondences isolated from such dataset.

Table 3. Sound correspondences from Secondary Lexical Dataset.

	pMa	pMiSo	Recurrence	Tentative Reconstructions	Examples
22	b’	p-…(ʔ)	1x	*b’	2
23	ŋ-	n-	1x	*ŋ	23
24	n	n	1x	*n	7
26	-l	-h	1x	*l	149
27	t	¢	1x	*t	20
28	t^{y}’	¢	1x	*t^{y}’	41
30	č	¢	1x	*č	36
31	č	s	1x	*č	58
32	č’	s	1x	*č’	39
33	¢	¢	1x	*¢	77
34	¢’	¢	1x	*¢’	46

35	¢’	s	1x	*¢’	22
39	q	s	1x	*q^y	72
41	q’	¢	1x	*q^{y}’	54
43	x	y	1x	*x	6
44	x	w	1x	*x	19
45	h	y	1x	*h	18
25	-l	-y	2x	*l	42, 51
29	t’	t	2x	*t’	75, 83
37	k’	¢	2x	*k^{y}’	6, 104
38	q	¢	2x	*q^y	24, 60
40	q’	s	2x	*q^{y}’	8, 36
42	r	¢	2x	*r	3, 14
36	k	¢	3x	*k^y	7, 27, 37
46	-ʔ	-y	3x	*y	1, 10, 151

At this point it is worth noting that Brown (2017a) has described a procedure, dubbed the Beck-Wichmann-Brown (BWB) system, updated from Brown et al. (2014), that aims to “provide much-needed uniformity to the comparative method” (Brown 2017a:252). The BWB system weighs several criteria for assessing the validity of a comparative set, by assigning each criterion a certain number of points, and it does so separately and differently. Instead of applying strict semantic isomorphism, the BWB system uses “translation equivalence,” which allows for the comparison of etyma that are not semantically isomorphic but which exhibit feasible shifts. Brown offers the following breakdown of their criteria (2017a) for evaluating each comparison set: 2 points “For each sound correspondence”; 2 points “For translation equivalence”; 1 point “For lack of potential onomatopoeia”; and 1 point “For lack of unexplained, unmatched phonological segments.” The author argues for the application of the following procedure: “A sound correspondence is deemed observed (and thus earns 2 points) only if it is supported by at least *three* comparative sets (each of which shows at least 7 BWB points)” (Brown 2017a). Brown (2017a) then applied this procedure to several proposed genetic groupings in the literature, including some that are well accepted (e.g. Indo-European) and others that are controversial (e.g. Na-Dene/Yeniseian). He also applied it to Mora-Marín’s (2016) Proto-Mayan/Proto-Mije-Sokean proposal, using three items deemed to correspond to basic vocabulary (‘bone’, ‘to give’, ‘white’), and the results conferred support to the proposal, with two sets (‘bone’ and ‘to give’) receiving 8 points each, and the third receiving 7 points (Brown 2017b); these sets exhibit three recurrences each for the correspondences of pMa *q to pMiSo *k, and pMa *aa/*a to pMiso *a/*aa. Brown concludes that the proposal exhibits “At least moderate support for genealogical relationship, but one of very great chronological depth” (2017b:300).

Altogether, Mora-Marín (2016) defined 46 sound correspondences present among potential cognates, 21 of which showed sufficient recurrence to support cognacy. Next, I apply the same methods in order to test additional comparanda for cognacy, diffusion, or accidental similarity. This task will make reference, whenever it is deemed useful, to Brown, Holman, and Wichmann (2013a, 2013b), who have utilized “standardized word lists for over half of the world’s languages” to isolate 692 sound correspondences among related languages by means of a computer program. They also provide “frequency data for each observed correspondence.” The approach by Brown et al. (2013a, 2013b) can offer historical linguists a useful heuristic.

Assessing the "global frequency" of a particular sound correspondence can be used to weigh the evidence. In fact, referring to the shared innovation of pMa *r > y among the Huastecan and Greater Tzeltalan subgroups of Mayan, Brown et al. (2013a:15) state the following:

> If the change is common, it can be assumed to have arisen independently in two groups of languages; if it is unusual, it provides stronger evidence that a shared innovation occurred and thus supports a scenario where the two language groups—in this case Huastec and Greater Tzeltalan—form a subgroup.

Brown et al.'s (2013b) global frequency measures for the r : y correspondence (and thus for the pMa *r > y change), as those authors note, suggests that the change is relatively common. It occurs among 12 language genera (NG), out of 276 language genera among which "the two corresponding sounds are frequent enough for a correspondence between them to be identified if it is actually present" (AG), and thus it exhibits a correspondence percentage (CP) of 4.35. To Brown et al. (2013a), these measures are sufficient to support independent innovation.[6]

Lastly, the resulting sound correspondences—both previously established ones from Mora-Marín (2016) as well as new ones—will be listed and discussed.

3. Analysis

3.1. Lexical Comparanda

I begin with lexemes from Mora-Marín (2016) that were excluded from the Primary Lexical Dataset (Addendum 1), especially those with no recurrence or unusual correspondences, and also with items from Mora-Marín's (2016) Addendum 4, those exhibiting intriguing similarities but problematic traits. In the following comparisons, cognates reviewed for the first time will be presented with the set number between parentheses.

Mora-Marín (2016:152-153) provides a single etymon, item #23 in **Table 4**, as evidence for a tentative *ŋ : *n correspondence.[7] One of the items in Addendum 4 could support this correspondence, #64. However, the item in question does not abide by strict semantic isomorphism, and requires, instead, that at least two obstacles be overcome: comparing a noun in Mayan to an adjectival participle in Mije-Sokean, and assuming that 'tail' can be derived metaphorically from the meaning 'bent' or 'twisted'. This is a feasible shift, e.g. 'twisted/bent' > 'tail', but the comparison should be considered tentative.

Table 4. Evidence for *ŋ : *n correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
23	ŋ-	n-	23	pCMa *ŋab'/nɨɨʔ	rain/water
			(64)	ŋeeh/new-e	Tail/bent, twisted

[6] It is important to note that this heuristic should be applied with the proviso: "all other things being equal." If all other things are not equal, i.e. if evidence for loanwords and relative chronologies of sound changes based loanwords is available, then one must consider such evidence first. Indeed, Kaufman and Justeson (2008) have provided evidence in the form of several Huastecan loanwords among several Otomanguean and Totonacan languages that must have occurred prior to the *r > y change (as well as other sound changes that characterize Huastecan), but subsequent to the arrival of Huastecan speakers to the Huasteca region.

[7] This correspondence is not too rare, according to Brown et al. (2013:s): NG 30, CP 14.02, AG 214. In fact, its CP is comparable to that for the q : k correspondence in their data: NG 12, CP 15.79, AG 76.

Table 5 provides the single etymon used as evidence for the *n : *n correspondence in Mora-Marín (2016:153), item #7 'snake'. I have added an additional item from Addendum 4, item #158, pMa *hoonon 'wasp' and pMiSo *weeʔnV(k) 'wasp', and a newly identified example, item #182, pCMa *naq 'to fight' and pMiSo *naks 'to beat, to whip'. There may exist several problems with item #158. First, the Mayan etymon, at least, could very well be onomatopoeic. Second, there are no *entirely unproblematic* cases of a correspondence of pMa *h to pMiSo *w; the only other example would be item #64, **Table 4** above, lacking semantic isomorphism (but exhibiting a plausible semantic shift 'bent, twisted' > 'tail'). And third, the final consonant in pMa, *n, does not plausibly match the likely final consonant of the pMiSo form, *k, although this consonant may not have been present, given the parentheses in Wichmann's (1995) reconstruction. Item #182 is more straightforward, showing a correspondence of pCMa *n- and pMiSo *n-, as well as pCMa *-q and pMiSo *-k (correspondence 7), the latter well attested (10 prior occurrences). Although the *n : *n correspondence is attested in the loanword dataset, items #7 and #182, together, would support its viability, and so would #158, albeit more tentatively due to its likely onomatopoeic nature and unmatched/mismatched segments.

Table 5. Evidence for *n : *n correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
24	-n	-n	7	kaan/pSo *¢ahin	snake/snake
	-n-	-n-	158	hoonon/weeʔnV(k)	Wasp/wasp
	n-	n-	182	pCMa *naq/pMiSo *nak(-s)	To fight/to beat, whip

The following sets involving nasals seem more promising. **Table 6** provides two etyma from Addendum 4 supporting a minimally recurring correspondence of pMa *ŋ to pMiSo *m.[8] The forms are as semantically isomorphic as one could expect. The MiSo form is assumed, as discussed above in connection with the *ks clusters, to be traced to a pre-pMiSo form **miʔk-s, bearing a **-s 'verbalizer' suffix. The remaining consonantal correspondences would involve, for #108, *h : *ʔ, and *k : *k, and for #179, *ʔ and *ʔ. The problem is the situation of unmatched segments, contained within square brackets. The pMiSo form would contain two segments, *ut, that do not correspond to any segments in the Mayan form. However, despite the violation of this criterion, the forms are otherwise too similar, phonologically and semantically, to warrant exclusion. They are likely to be either cognates or loans.[9] Since there is no ŋ : m correspondence among the loanwords, I will consider them potential cognates, and defer the question of the unmatched segments as a matter for future research.

[8] This correspondence is somewhat rare, according to Brown et al. (2013:s): NG 5, CP 2.33, AG 215.

[9] Given the morphological structure of both Mayan and Mije-Sokean languages, it is more likely that the bulk or entirety of a lexical root or stem consist of the first three CVC sequence (pMa *ŋiiʔ, pMiSo *mɨʔ…), and that the remaining, unmatched segments (pMiSo *…ut) could potentially constitute a previously unidentified suffix, whether inflectional or derivational. Wichmann (1995:317) appears to reconstruct a suffix of the right shape, *-ut, in pMiSo *hukʔ-ut 'fire, light', a noun, based on pMiSo *hukʔ 'to smoke', a verb. Perhaps pMiSo *mɨʔut contains such a suffix too; so far, though, there is no evidence for a pMiSo verb of the shape mɨʔ* that could serve as the basis for a derived noun with such a hypothetical suffix.

Table 6. Evidence for *ŋ : *m correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
47	ŋ	m	(108)	ŋihk/miʔks	Temblar/temblar
			(179)	ŋiiʔ/mɨʔ[ut]	parent-in-law; son-in-law/son-in-law (also parent-in-law across both branches)

Table 7 presents evidence for a correspondence of pMa *b' to pMiSo *m, a correspondence not discussed in Mora-Marín (2016). In this correspondence, pMa *b' appears finally (#78, #80, #92) or intervocalically (#79), while pMiSo *m appears intervocalically. There exist four instances. There is one case that is semantically questionable, item #79, for it requires comparing a verb form 'to hear', involving a derivational suffix **-i*, to a nominal form 'deaf' that does not obviously involve a suffix, and they seem to be antonymic. Item #92 is not semantically isomorphic, but the meanings are relatable. The remaining examples show semantic isomorphism (#78) or at least overlap (#80) among available meanings. This correspondence thus shows some merit, and it is plausible. It further raises the question of two previously unidentified correspondences: *h : *ʔ, *l : *h.

Table 7. Evidence for pMa *b' : pMiSo *m correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
48	-b'	-m-	(78)	haʔb'/ʔaame	year/year
			(79)	ʔab'.i/ʔuuma(ʔ)	hear/deaf
			(80)	laab'/pSo *hama	witch, nagual/day, sun, nagual
			(92)	kab' ~ kaab'/kama	earth, land/cornfield

Next, **Table 8** provides the evidence for the correspondence of pMa *l to pMiSo *m in initial position. This correspondence is rare, according to Brown et al. (2013:s17): NG 1, CP 0.34, AG 291. Item #105 is problematic semantically: it could very well be sound symbolic in both Mayan and Mije-Sokean. In fact, the Mayan form, **leq'*, even bears a similarity to the English form, *lick*. But even after excluding this item, three comparanda remain viable based on the criteria that were laid out. One of these three items could prove problematic, nonetheless, for culture-historical reasons: item #93 refers to objects made of clay in both pMa and pMiSo. This brings up an interesting question: what type of clay objects do these terms refer to? In today's Mayan languages, reflexes of pMa **laq* typically refer to *fired* pottery cups/bowls. But the invention/arrival of pottery in/to Mesoamerica very likely postdates the breakup of any putative pMaMiso language, potentially by a millennium or more. So if this term is in fact cognate, it is possible that its meaning may have pertained to unfired clay (e.g. sun-baked), but not pottery per se. The three viable items are not only sufficient to propose a recurring sound correspondence, but one of the other consonantal correspondences attested in these three forms have already been shown to exhibit recurrence: *q : *k (correspondence 7, 9x). Two additional sound correspondences among these three comparanda were previously documented: *q' : *(ʔ)¢ (correspondence 41, 1x), *¢' : *¢ (correspondence 34, 1x). These examples therefore raise correspondences 31 and 34 to 2x each.

Table 8. Evidence for pMa *l : pMiSo *m correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
49	l-	m-	(29)	looq’/muuʔ¢	adobe, mud, wall/toast, make adobes
			(34)	lah¢’/ma¢	squeezed, narrow/to grasp, rub, squeeze
			(93)	laq/mɨk-i	clay plate, cup/object made of clay
			(105)	leq’/muuʔk	to lick/to suck

Table 9 illustrates a correspondence already supported with three comparanda in Mora-Marín (2016:142, correspondence 4). Three additional examples exist. The first, item #81, exhibits this correspondence: semantically, the pMa and pMiSo forms are comparable. The main issue lies in justifying the correspondence of pMa *č’ to pMiSo *ks or pMiSo *ʔks. If I follow my assumption that the *s of the *ks cluster is a pre-pMiSo innovation (**-s ‘verbalizer’), then pMiSo *k or *ʔk would be the segment (or sequence of segments) comparable to pMa *č’. For now, I cannot provide a satisfactory solution to this word-final correspondence, nor is there another example to support it. That said, the correspondence of pMa *C’ to pMiSo *ʔC or *Cʔ, argued in Mora-Marín (2016:159-162) to be present in a majority of correspondences involving pMa *C’, is observed in this case. Another new instance of the *m : *m correspondence is available in the form of #160, which I discuss in more detail below in connection to the *t : *¢ correspondence. And a third new instance is found in item #91, which I review in more detail with regard to the *k’ : *¢ correspondence below. Thus, there exist now six items to support the *m : *m correspondence.

Table 9. Further evidence for *m : *m correspondence

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
4	m	m	(81)	mič’ ~ moč’/mɨɨʔks	To clench, squeeze in hand/To squeeze, extract
			(160)	LL+WMa *kitaam/pMi *ʔii¢ɨmɨ	Peccary/peccary
			(91)	k’am/¢ɨm	to receive, take, carry/to carry on the back

A last set including a nasal sound requires review, shown in **Table 10**. The first set, #86, seems viable semantically, assuming semantic shifts through metonymy (e.g. Amate tree > paper made from the Amate tree > paper used for writing > to write). The second set, #40 requires the assumption of several semantic shifts too, but roasting in ashes and coals is a well-known cooking technique, and thus a metonymic change could result in such fashion. For now, the correspondence of -(ʔ)ŋ : -y(ʔ) must be regarded as tentative, and will be listed last in case it requires removal at a later time.

Table 10. Further evidence for *ŋ : *y correspondence

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
61	-(ʔ)ŋ	-y(ʔ)	40	t^yaʔŋ/¢aayʔ	ashes, quicklime/to roast
			86	huʔŋ/haayʔ	amate (Ficus), paper/to write

The newly identified correspondence of *h : *ʔ, mentioned above with respect to #78 (**Table 8**) is not isolated; see **Table 11**. There exists another set from **Addendum 4** that exhibits it, item #62, pMa *huul 'to arrive here' and pMiSo *ʔoy 'to leave and come back; to come back and leave'. They are both intransitive verbs of motion with a partial overlap in meaning, but I am not confident that one meaning is easily derivable from the other, and consequently, their semantics could be problematic. For now, this correspondence deserves only tentative support.

Table 11. Evidence for pMa *h : pMiSo *ʔ correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
50	h-	ʔ-	62	huul/ʔoy	to arrive here/to leave and come back; to come back and leave
			78	haʔb'/ʔaame	year/year

The last etymon discussed, #62, does raise the issue of another sound correspondence, *l : *y, for which two potential cognates had previously been identified, items #42 and #51, as seen in **Table 12**. This correspondence has therefore acquired additional support from #62, and also from an additional item that had previously been noticed by Kaufman and Norman (1984:124), namely, an intransitive root reconstructed as pMa *laaj 'to (be) finish(ed); to end', and pSo *yah 'to be finished'. More recently, Kaufman with Justeson (2003:1271) have revised the reconstruction to proto-Central Mayan (pCMa) *laj or *laaj, while Wichmann (1995:515) reconstructs it to pSo as an intransitive root *yah 'to finish'.

Table 12. Evidence for pMa *l : pMiSo *y correspondence

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
25	-l	-y	42	k'uul/kuy	brush, woods, forest/tree
			51	wol/woy	spherical/to roll up
			62	huul/ʔoy	to arrive here/to leave and come back; to come back and leave
	l-	y-	183	pCMa *lax or *laax/pSo *yah	to end; to finish

Next, **Table 13** provides the evidence for the *l : *h correspondence in Mora-Marín (2016:153-154). Item #80, mentioned above in connection with the *b' : *m correspondence, exhibits the *l : *h correspondence too and thus provides additional support.[10]

Table 13. Evidence for *l : *h correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
26	-l	-h	149	hil ~ hihl/pSo *heh	to rest/to rest
	l-	h-	80	laab'/pSo *hama	witch, nagual/day, sun, nagual

Table 14 presents the *t : *¢ correspondence in Mora-Marín (2016:154), which he described as consistent with the apparent conditions for the $*t^{y}$' : *¢ correspondence attested in

[10] This correspondence is somewhat rare, according to Brown et al. (2013:s): NG 5, CP 2.26, AG 221.

two other etyma. One item from Addendum 4 exhibits this correspondence and could provide additional support. However, it is problematic on two counts. First, the Mayan etymon may be internally diffused; but if it's not, its distribution could suggest a LpMa stage. Second, the initial consonant of the Mayan etymon, *k, would not be the most likely correspondence to the initial consonant of the MiSo etymon, *ʔ.[11] It is feasible that **kitaam* could be a reanalysis of **k-(ʔ)itaam* (i.e. /k+ʔitaam/) 'their peccary', in which case a Mayan root **ʔitaam* would be more easily compared to the Mijean root *ʔii¢ɨmɨ; if correctly analyzed, then these comparanda are more likely to constitute a case of cognacy, rather than borrowing, for Mayans would have likely heard a MiSo *¢ as /¢/, [¢], and Mije-Sokeans would have likely heard a Mayan *t as /t/, [t].[12] Also, a correspondence of *t : *¢ does not appear among the correspondences for the proposed loans between Mayan and MiSo. For now, I will tentatively suggest that the Mayan etymon is indeed LpMa *kitaam (and possibly *k-itaam), and that it is a possible cognate with pMiSo *ʔii¢ɨmɨ.

Table 14. Evidence for *t : *¢ correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
27	-t	-¢	20	q'ut/kaʔ¢	to grind in grinding stone/to stone, grind
			160	LL+WMa *kitaam/pMi *ʔii¢ɨmɨ	Peccary/peccary

The correspondence between *t^{y}' and *¢ identified by Mora-Marín (2016:154) for item #41 was also noted for item #74, seen in **Table 15**. The problematic issue lies with the possibility that #74 could be onomatopoeic in either pMa or pMiSo (pSo). That said, it shows the same correspondence as item #41, which is unlikely to be onomatopoeic, and the *t^{y}' : *¢ correspondence does not occur among the loanword correspondences. Also, the word-initial *p : *p correspondence in #74 is unproblematic and well attested. Brown et al. (2013) do not document a t^{y} : ¢ correspondence. Within Mayan alone, however, similar developments can be documented: Greater Mamean exhibits *t^{y} > ¢, and *t^{y} > č'; since MiSo languages typically do not have *t^{y}, a correspondence between pMa *t^{y}' and pMiSo *¢ seems reasonable. And since the model proposed in Mora-Marín (2016) calls for deglottalization or ejective fission of ejectives in pre-pMiSo times, an initial change could have been pMaMiso *t^{y}' > pre-pMiSo *t^{y}, followed by a change similar to that seen in Greater Mamean, pre-pMiSo *t^{y} > pMiSo *¢. So, tentatively, I support this correspondence as exhibiting minimal recurrence and worthy of further exploration.

Table 15. Evidence for *t^{y}' : *¢ correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
28	-t^{y}'	-¢	41	t^{y}'eken/¢uku	driver ant/ant
			74	pat^{y}'/pSo *pa¢	to smash, crack/to squash

There is a similar correspondence, of pMa *t^{y} to pMiSo *¢, attested in two sets, items #138 and #40, seen in **Table 16**. The first refers to a bodily function or secretion, and could

[11] Brown et al. (2013) list this correspondence, of k : 7, as not that uncommon worldwide: NG 17 CP 9.19, AG 185. However, it would be the only such example in the datasets I am using, and in my experience, it is a rare correspondence, internally, within Mayan and Mije-Sokean.

[12] This correspondence (¢ : t) is not very rare, according to Brown et al. (2013:s): NG 7, CP 4.83, AG 145.

potentially be also onomatopoeic. I discuss it further, together with other semantically similar etyma, in section 3.2. Set #40 is also problematic, but the semantic shifts required are plausible; for now, the correspondence will receive tentative support, and listed toward the bottom of the comprehensive list of sound correspondences in **Table 38**, just in case its removal is deemed necessary at a later time.

Table 16. Evidence for *t^y : *¢ correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
60	t^y-	¢-	138	t^yuh[b’] ~ ¢uh[b’]/¢uh, ¢uh-i(C)	spittle/to spit, spittle
			40	t^yaʔŋ/¢aayʔ	ashes, quicklime/to roast

Mora-Marín (2016:154) examined the correspondence of pMa *t’ and pMiSo *t seen in **Table 17**. The first etymon, #75, could be argued to not exhibit semantic isomorphism; however, the superficial similarity between the mollusk inside a snail and an earthworm (both slick and elongated) is, in my opinion, likely to be semantic basis for this etymon.[13] The other, #83, could be argued to be onomatopoeic, at least the Mayan version, given the meaning, ‘to drip’. At the present, there exist no additional examples of this correspondence, and thus it must remain somewhat tentative.

Table 17. Evidence for *t’ : *t correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
29	t’-	t-	75	t’oot’/pSo *toʔt	snail/earthworm
			83	t’ux/tuuh	to drip; drop/rain

A similar correspondence, *t’ : *¢, parallel to that of *t : *¢, discussed above, is attested in etymon #136, seen in **Table 18**. The problem here lies primarily in the possible onomatopoeic nature of this etymon, not unlike the case of #83. However, the correspondence of *l : *n is already documented (Mora-Marín 2016:142, correspondence #3). I discuss both etyma, #83 and #136, below, in connection to analogical reanalysis of sound symbolism as a confounding factor.

Table 18. Evidence for *t’ : *¢ correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
56	t’-	¢-	136	t’ul/¢un	drop/drip

Mora-Marín (2016:154-155) describes two correspondences of pMa *č to pMiSo *¢/*s, and pMa *č’ to pMiSo *s, as shown in **Tables 19** and **20**, respectively. As already noted, one case of a similar correspondence, of pMa *č to pMiSo *¢, occurs in the secondary lexical dataset with #96, which could be onomatopoeic: pMa *kaač ‘horsefly’ : pMiSo *¢a¢i ‘horsefly’. Also, a correspondence of pMa *č’ to pMiSo *¢ is evident in items #76, pMa č’ax ‘to wash’ and pSo *¢eʔ ‘to wash’, and #171, GQ’+GMa (also Q’eqchi’) *č’aʔb’iʔn ‘Crotalaria spp.’ and pMiSo *¢ɨpʔV ‘quelite’. The first set, #76, is problematic because of the correspondence between the

[13] I consider the case of Mayan *t’oot’ to be derived, originally, from a form of the shape **t’oot or **toot’, before undergoing a process of assimilation between homorganic stops, as argued in Mora-Marín (n.d.2). Thus, a better comparison would be between pre-pMa *t’oot (or *toot’) and pSo *toʔt.

second consonants, *x and *ʔ, which is not otherwise attested. Below, in Section 3.2, I suggest that this irregular correspondence could be due to reanalysis involved in synesthetic sound symbolism within Mayan. Loanword correspondence set #19 (**Table 1**) attests to several instances of the č : ¢ correspondence among likely loans, and loanword correspondence set #20 attests to one instance of č : s. There are no loanword correspondences of č' : s or č' : ¢. I consider both correspondences to be further supported, if only tentatively. More tentatively, it may be noted that the *č : *s and *č' : *s correspondences appear to be in complementary distribution with the *č : *¢ and *č' : *¢ correspondences: within Mayan, the former are associated with high vowels and the latter with low vowels; within Mije-Sokean, the former are associated with back vowels, the latter with non-back.

Table 19. Evidence for *č : *¢/*s correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
30	č-	¢-	36	čaaq'/¢eʔs	bed[14]/maize granary, bed
			96	kaač/¢a¢i	Horsefly/horsefly
31	č-	s-	58	čiim/sum	bag, to bag/to gather, net

Table 20. Evidence for *č' : *¢/*s correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
52	č'-	¢-	76	č'ax/pSo *¢eʔ	To wash/to wash
			171	GQ'+GMa (+Q'eq) *č'aʔb'iʔn/¢ɨpʔV	Chipilín (Crotalaria spp.)/Quelite
32	-č'	-s	39	xuč'/hooʔs	to grind corn, dough/nixtamal (cooked corn)

Mora-Marín (2016) proposed two sets of correspondences of pMa *¢ : pMiSo *¢, supported by etymon #77, and pMa *¢' : pMiSo *¢/*s, supported by etyma #46 and #22. Additional etyma exhibiting these correspondences may be seen in **Tables 21** and **22**, respectively. The first, #88, corresponds to pCMa *¢ak, reconstructed as WM+ by Kaufman with Justeson (2003:903), and glossed variously as 'to chase after', 'to grasp', 'to hunt'; its presence in Q'eqchi' could potentially point to a CMa reconstruction. If so, this term could be compared to pSo *¢ɨk 'to grasp', and the two exhibit an initial *¢ : *¢ identity correspondence, as well as a final *k : *k correspondence. Addendum 4 listed two additional items supporting the second correspondence of pMa *¢' : pMiSo *¢, shown in Table 22: pMa *lah¢' 'squeezed, narrow' and pMiSo *ma¢ 'to grasp, rub, squeeze', as well as pMa *¢'u¢' 'to kiss, suckle' and pMiSo *¢uʔ¢ 'to suckle, chew'. Mora-Marín (2016) had previously disqualified this item for being potentially onomatopoeic and exhibiting irregular reconstruction within Mayan, i.e. pCMa *¢'u¢' ~ *¢'ub'. However, Mora-Marín (n.d.1) has since shown this apparent irregularity yields to internal reconstruction and the discovery and reconstruction of a pre-pMa *-Vb' suffix utilized with stems that refer to actions of the mouth (kiss, yawn, whistle, cough, spit); see Section 3.2 for a summary of this analysis. Although it would appear that there are two instances of this correspondence within item #137, one of them is likely the result of assimilation to a homorganic ejective obstruent in Mayan (i.e. pre-pMa *¢'u¢ or *¢u¢' > pMa *¢'u¢'); consequently, #137

[14] See footnote on #36 in Addendum 1 for justification.

exhibits a correspondence of *¢ : *¢ as well. An additional example with the ¢’ : ¢ correspondence was included in Mora-Marín’s (2016) Addendum 5 (‘Possible Archaic Loans or Cognates’): pMa *¢’utux ‘corn tassel’ and pSo *¢utu ‘ear of corn’. This item also shows a correspondence, in word-final position, of *x : Ø, not otherwise attested. No further examples have been adduced exhibiting the *¢’ : *s correspondence (#35) identified by Mora-Marín (2016). There is a hint of complementary distribution between the *¢(’) : *¢ and *¢(’) : *s sets, but with so few examples a strong case cannot be made. It would seem that the *¢ : *s correspondence correlates with a high back vowel in Mayan, instead of a low central vowel; and it would seem that the *¢’ : *s correspondence correlates with a mid-back vowel in Mayan, instead of the low central or high back vowel.

Table 21. Evidence for *¢ : *¢ correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
33	¢	¢	88	pCMa *¢ak/pSo *¢ɨk	to chase after, to grasp, grab/to grasp
			77	qa¢/pMi *kɨɨʔ¢	to rip/to break
53	¢-	s-	21	¢uk/so¢	to seize; to fish, to tie up/to tie

Table 22. Evidence for *¢’ : *¢/*s correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
34	¢’	¢	46	pCMa *¢’ak/¢ookʔ	medicine/to heal
			137	¢’u¢’ (pre-pMa *¢u¢’ or *¢’u¢)/¢uʔ¢	kiss, suckle/suckle, chew
			48	¢’utux/pSo *¢utu	corn tassel/ear of corn
			34	lah¢’/ma¢	squeezed, narrow/to grasp, rub, squeeze
35	¢’	s	22	xo¢’/huʔs	to scratch, shave/to shave

Previously, Mora-Marín (2016:156) had identified a correspondence (36) of pMa *k to pMiSo *¢ evident in three items (#7, ‘snake’; #27, ‘quern/stone’; #37, ‘palm tree’). He proposed pMaMiso $*k^y$ as the source of such correspondence. He had also noted the presence of this correspondence in Addendum 4, specifically in #96 from the secondary lexical dataset: pMa *kaač ‘horsefly’ : pMiSo *¢a¢i ‘horsefly’. And he had also observed the presence of this correspondence among the loanwords. Besides item #96, two other etyma exhibit this correspondence, as seen in **Table 23**: #73, pMa *kiis ‘fart’ and pMiSo *¢iisʔ ‘fart’, and #21, pMa *¢uk ‘to seize; to fish, to tie up’ and pMiSo *so¢ ‘to tie’. Etymon #96 is problematic for two reasons: first, it has been proposed to be *potentially* onomatopoeic; and second, its initial consonant in pMa exhibits variation, given the multiple reconstructions as *kiis ~ *¢iis ~ *tiis (Kaufman with Justeson 2003:319). The evidence from most subgroups (Yu, EMa, WMa) points to *kiis; and there is reason to believe that the irregular forms could be the result of internal diffusion within Mayan.[15] Two additional items were previously proposed to illustrate this

[15] Within Western Mayan, three languages (Tzotzil, Tzeltal, Popti’) attest to the form *¢is*, but given their close proximity to Sokean speakers, it is possible that they could have borrowed it from Sokean. It is also attested in Huastecan as *tiθ*, which Kaufman with Justeson (2003:328) point out could be traced to either *tiis or *¢iis. Both *kis ‘fishy-smelling’ and *tis ‘fart’ can be reconstructed to proto-Ch’olan (Kaufman and Norman 1984:123). Given

correspondence by Fox (1978) and Brown and Witkowski (1979): pMa *kaq ‘red’ and pMiSo *¢apa¢ ‘red’ (#49) and pMa *kaʔŋ ‘sky’ and pMiSo *¢ap(-hɨ) ‘sky’. Fox (1978) and Brown and Witkowski (1979) both assumed their cognacy. I did too earlier in my research. However, these two items are highly problematic: they both exhibit unmatched segments that cannot be explained; and one of them (#49) would require two correspondences (*k : *¢ and *q : *¢) that involve significant phonological distance, a violation of my methodological assumption that only one correspondence per etymon be allowed to do so. This correspondence can be argued to gain support from items #96, #73, and #21, because they otherwise observe regular correspondences (*č : *¢ for item #96; *s : *s for item #73), although this additional support must remain tentative due to the possible onomatopoeic nature of perhaps two of these items (#73, #96) suggested by Campbell and Kaufman (1980:853).[16]

Table 23. Evidence for *k : *¢ correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
36	k	¢	96	kaač/¢a¢i	horsefly/horsefly
			73	kiis/¢iisʔ	fart/to fart
			(49	kaq/¢apa¢	red/red)
			(109	kaʔŋ/¢ap(-hɨ)	sky/sky)
			21	¢uk/so¢	to seize; to fish, to tie up/to tie
			7	kaan/pSo *¢ahin, pMi *¢anay (or pMi *¢anʔay)[17]	snake/snake
			27	kaaʔ/¢aaʔ	quern/rock, stone
			37	tuuk/tuu¢[-kuy]	coyol (type of palm tree)/palm tree

Mora-Marín (2016:156) proposed a correspondence of *k’ : *¢ based on two etyma (#6, #104) parallel to the foregoing correspondence of *k : *¢, based on three etyma (#7, #27, #37). However, one of the two etyma proposed in support of the *k’ : *¢ correspondence must be removed from consideration: #104. It was incorrectly glossed as ‘to shrink’ (Spanish ‘encoger’), instead of the correct gloss as ‘to choose, select, pick out’ (Spanish ‘escoger’); the item must be removed from the *s : *s correspondence, correspondence 19 (Mora-Marín 2016:148-149), which I address below. The correct item that should have been compared, that was in my original dataset notes, was item #91, pMa *k’am ‘to receive, take, carry’ and pMiSo *¢ɨm ‘to carry on the back’. With this correction the pattern remains the same: a correspondence that shows minimal recurrence (items #6 and #91).

---

their narrow distributions, it is possible that the form *tiis and *¢iis could have been innovated much later, long after the breakup of pMa, and then diffused internally, the second one possibly from a Mije-Sokean language.

[16] Campbell and Kaufman (1980:853) state that “The involvement of onomatopoeia in some of these forms is clear, since they illustrate irregular correspondences within the Mayan langauges [sic].” In the case of ‘horsefly’, though, the correspondences within Mayan appear to be thoroughly regular, but given the referent, a sound-imitative quality cannot be discounted, though it is by no means obvious that this should be so, as Witkowski and Brown (1981:907) imply. As far as ‘fart’ is concerned, there exist irregular correspondences within Mayan, but I have discussed these in the preceding footnote and treated such irregular correspondences as the result of possible internal diffusion due to the areal restriction of the irregular reflexes.

[17] Kaufman (personal communication, 2012) reconstructs pMi *¢anʔay ‘snake’.

Table 24. Additional vidence for *k' : *¢ correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
37	k'	¢	91	k'am/¢ɨm	to receive, take, carry/to carry on the back
	k'	¢	6	k'ax[aaŋ] ~ k'ax[ŋ]/¢ay	rope, vine/rope, cord (also, by extension, 'hammock'; in some forms of So, 'vine')

Previously, Mora-Marín (2016:148-149) proposed a correspondence of *s : *s. As already noted, item #104 requires removal due to an error. That said, at least two other etyma from the primary lexical set in Mora-Marín (2016) attest to the correspondence, as seen in **Table 25**. Two additional etyma at the highest level of reconstructions (pMa, pMiSo) attest to it as well (**Table 25**): #73 and #166. Both items have been suggested in the past to be potentially onomatopoeic, though. A third item, #162, is available, but problematic due to the final correspondence of pCMa *p to pSo *w [ŋ], which is otherwise not attested elsewhere; however, this correspondence—between pCMa *p and pSo */w/—is not at all implausible and will be listed toward the bottom of the comprehensive list in **Table 38**, due to the fact that so far it is unique.

Table 25. Evidence for *s : *s correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
19	s	s	11	pCMa *siiq'/suuʔk	to smell/to smell
			107	saq/saaka	white/white
			73	kiis/¢iisʔ	fart/to fart
			166	ʔus/ʔuusu	gnat/mosquito
			162	pCMa *siip/pSo *siŋ (*/siw/)	to swell/to swell

Mora-Marín (2016) also described the correspondences of *q : *¢/*s and *q' : *¢/*s as unusual, in reference to phonological distance involved (and the number of presumed sound changes necessary to account for such a correspondence). In fact, no such set of correspondences occurs in the Brown et al. (2013) dataset. Mora-Marín (2016) described three etyma in support of each correspondence: #24, #60, #72 for *q : *¢/*s; #8, #36, #54 for *q' : *¢/*s. Two additional examples of the *q : *¢ correspondence are evaluated here, seen in **Table 26**. The first is #44, involving a pMa root *ʔihq- corresponding to pMi *wih¢ ~ *wi¢. The pMa root in question is present in pCMa *ʔihq.a¢ 'carga' and EMa+GQ' *ʔihq.a 'cargar' (Kaufman with Justeson 2003:893). This etymon would exhibit an initial correspondence of pMa *ʔ to pMi *w, and a final correspondence of pMa *q to pMi *¢. This *ʔ : *w correspondence is attested elsewhere, as discussed below. A second item, #49, would also support a *q : *¢ correspondence, but it was already discarded due to the phonological distance and unmatched segments criteria. Thus additional support for the *q : *¢ correspondence is found only in item #44, raising the number of occurrences to four. A hint of complementary distribution between *q : *¢ and *q : *s is evident: the Mije-Sokean correspondences show *s after a high back vowel, and *¢ elsewhere.

Table 26. Evidence for *q : *¢/*s correspondence

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses

38	q	¢	44	ʔihq-/pMi *wih¢ ~ *wi¢	load, to carry on back or head/carry, carry on the head
			(49	kaq/¢apa¢	red/red)
			24	t^{y}aq/tɨɨ¢~ tɨɨ¢ʔ ~ tɨʔɨ¢	dry/to dry out, become thin; dry, thin; tooth
			60	k’oq/kooʔ¢	to cut, tear/to tan, break
39	q	s	72	pooq/puʔus	dust/dust

There are two additional items, seen in **Table 27**, that could provide support for the *q’ : *¢/*s correspondence for which three etyma had already provided support in Mora-Marín (2016). Item #29, discussed above in connection with the *l : *m correspondence also bears an instance of the *q’ : *¢ correspondence, or perhaps more precisely the *q’ : *ʔ¢ correspondence (i.e. *C’ : *ʔC). Also, item #55 provides support for the same correspondence, but this item also exhibits a correspondence of *n : *m that is not otherwise attested; it will also be added toward the bottom of the comprehensive list of sound correspondences in **Table 38**. Making this allowance for now, there exist now four examples in support of the *q’ : *¢/*s correspondence. While the q(’): ¢(’) correspondence does not occur in the Brown et al. (2013) dataset, Mora-Marín (2016) does not consider it likely that a change of *q(’) > ¢(’) took place, but instead, a sequence of changes such as *q^y(’) > k^y(’) > ¢ on one branch (e.g. MiSo) and *q^y(’) > q(’) on the other branch (e.g. Ma).[18] I have once again collapsed two correspondences due to the possibility of complementary distribution: the correspondence of *q’ : *s correlates with non-back and non-low vowels (*e, *ɨ) in Mije-Sokean, while *q’ : *¢ correlates with low central (*aa) and back vowels (*uu) in Mije-Sokean.

Table 27. Evidence for *q’ : *¢/*s correspondence

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
41	q’	¢	29	looq’/muuʔ¢	adobe, mud, wall/toast, make adobes
			55	q’an/¢aamʔ	yellow; ripe/to ripen
			54	ʔaʔq’/pMi *ʔaaʔ¢ ‘to grow (vine)’, pMi *ʔaaʔ¢-a ‘vine’	vine/to grow (vine), vine
40	q’	s	8	q’iiŋ/sɨw	sun, day/sun, day, fiesta, name
			36	čaaq’/¢eʔs	bed/maize granary, bed

Next, I address the *ʔ : *w correspondence that was discussed before in connection with item #44, which attests to the *q : *¢ correspondence. Two etyma, #15 and #44 support this correspondence, as seen in **Table 28**. Etymon #15 also provides further support for the *k’ : *k correspondence. This *ʔ : *w correspondence parallels the *ʔ : *y correspondence (46) described

[18] Brown et al. (2013) do catalog the ¢ : ky correspondence: NG 2, CP 8.00, AG 25. Their entry for the ¢ : k correspondence shows the following figures: NG 1, CP 0.68, AG 146. For c” : k”, they provide the following figures: NG 1, CP 6.67, AG 15.

in Mora-Marín (2016), supported by three sets of comparanda, repeated below in **Table 29**. A comparison of the two correspondences is suggestive of contrastive distribution.

Table 28. Evidence for *ʔ : *w correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
54	ʔ-	w-	15	LpMa *ʔak’/wɨɨkV	turkey hen/crested guan[19]
			44	ʔihq-/pMi *wih¢ ~ *wi¢	load; to carry on back or head/to carry (on the head)

Table 29. Evidence for *ʔ : *y correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
46	ʔ	y	1	ʔehq’ (or ʔexq’)/yɨk(ɨk)	black/black
			10	ʔaq’/yak	to give, place, put/to give (SaP also has ‘to place, put’)
			151	wiʔ/wa(ʔ)y	hair/hair

There is a set of correspondences identified in Mora-Marín (2015:158-159) that exhibit similarities (**Table 30**): pMa *x : pMiSo *y, pMa *x : pMiSo *w, and pMa *h : pMiSo *y. These occur word-finally.[20] Item #6 is given multiple reconstructions by Kaufman with Justeson (2003:960) and Kaufman (2017:100); the latter reconstructs it as *k’aj/haanh/m, i.e. *k’ajaaŋ or *k’ahaaŋ or *k’ajaam or *k’ahaam, and glosses it as ‘string, cord, rope; vine’. I analyze it as *k’axaaŋ or *k’ahaaŋ, and also as bimorphemic, i.e. *k’ax.aaŋ, in which case *.aaŋ would be a derivational suffix. There is evidence for a pMa *.aaŋ ‘nominalizer’ morpheme: pMa *b’ohr.aaŋ, with glosses such as ‘manojo (handful, bunch)’ and ‘rollo (roll)’, is likely derived from pMa *b’or ‘enrollar (to roll up)’ (Kaufman with Justeson 2003:914), in part by the addition of *.aaŋ. However, if pMa *k’ax.aaŋ/*k’ah.aaŋ is derived in such a manner, it is unclear what the root *k’ax-/*k’ah- would be. This would not be the only such case: pMa *t^yaq’ and *t^yaq’.aanh bear the same gloss, ‘ripe’ or ‘cooked’, making the function of the *.aaŋ suffix unclear. Item #12 was not included in the original comparison. It involves LpMa *waax, glossed here as ‘corn food’, instead of ‘tortilla’, the gloss in the contemporary Mayan languages, given that tortillas were innovated in Mesoamerica during the Late Classic period (Kaufman with Justeson 2003:1192). If item #6 is analyzed as *k’ax-, then it and #12 support a root-final correspondence of *x : *y. If item #6 is analyzed as *k’ah-, then it and #18 support a root-final correspondence of *h : *y. Finally, there is a similar word-final correspondence of pMa *x and pMiSo *w seen in item #19. This last correspondence differs from the other two contextually in the Mayan data: the *x : *y and *h : *y correspondences follow a low central vowel in Mayan, either *a or *aa/*aʔ, while the *x : *w correspondence follows a mid back vowel in Mayan, *o. In Mije-Sokean there is also a difference, but not as marked: the *x : *y and *h : *y correspondences follow a low central

[19] The crested guan is *Penelope purpurascens*. They are members of the Cracidae family; some of its members, including the crested guan, resemble turkeys, and in fact, were originally classified as related to turkeys.

[20] A word-medial context for the pMa *x : pMiSo *y correspondence may be seen in pMa *ʔaaxaaw : pMiSo *hayaw ‘man’ (a gloss of ‘big man’ might be more appropriate for pMa *ʔaaxaaw). This comparison would also call for a *ʔ : *h correspondence, which is otherwise not attested: instead, a *h : *ʔ correspondence is attested twice. The comparison also calls for a *w : *w correspondence, which is otherwise attested four times. The vowels would agree in quality, but not length. Despite the problematic *ʔ : *h correspondence the overall similarity makes this comparison one worth revisiting.

vowel, *a, or mid back vowel, *o, while the *x : *w correspondence follows a high central vowel, *ɨ. Previously, item #64 was said to be tentative due to a lack of semantic isomorphism, although a semantic shift from 'bent' or 'twisted' to 'tail' is conceivable. This item bears a correspondence of *h : *w, following a mid-front vowel, *ee in pMa, *e in pMiSo. If included, a pattern emerges in Mayan: low central vowel is associated with the *x/*h : *y correspondence, while a mid-vowel is associated with the *x/*h : *w correspondence. In pMiSo the contexts remain distinct but more complex: *y after *a and *o, *w after *e and *ɨ. An item previously argued to be somewhat problematic due to potential onomatopoeia, #158, exhibits a root-initial *h : *w correspondence. In any case, the evidence remains scant, insufficient to solidly establish either set of correspondences, but worth pursuing further.

Table 30. Evidence for *x : *y, *h : *y, and *x : *w correspondences.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
43	-x-	-y	6	k'ax[aaŋ] ~ k'axŋ/¢ay	rope, vine/rope, cord ('vine' in some so)
	-x	-y	12	LpMa *waax/way-e	corn food (tortilla)/to grind corn; dough
44	-x	-w	19	tʸox/tɨw	straight/straight
45	-h	-y	18	ra(a/ʔ)h/toy-a	pain/affliction, pain
55	h	w	64	ŋeeh/new-e	tail/bent, twisted
			158	hoonon/weeʔnV(k)	wasp/wasp

Next, prior evidence for a *x : *h correspondence was presented in Mora-Marín (2016), with at least four comparanda with biconsonantal correspondences, enough to provide support. One additional item has been found, providing further support (**Table 31**), namely, #183, which consists of pCMa *lax (or *laax) 'to finish/end' and pSo *yah 'to finish', already cited above in connection with *l : *y correspondence. Below, this correspondence is discussed in connection with additional comparanda that do not meet the strictest criteria but which nonetheless merit discussion.

Table 31. Evidence for *x : *h correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
15	x-, -x	h-, -h-, -h	22	xo¢'/huʔs	to scratch, shave/to shave
			39	xuč'/hooʔs	to grind corn, dough/nixtamal (cooked corn)
			53	b'ax/paahʔ	to nail/to chisel
			83	t'ux/tuuh	to drip; drop/rain
			183	pCMa *lax or *laax/pSo *yah	to end; to finish

Finally, I revisit a correspondence of interest, that between pMa *č' : pMiSo *(ʔ)ks, identified in item #81 above, and another similar example, item #172, exhibiting a correspondence of pMa *š : pMiSo *ks, seen in **Table 32**. Overall, I have argued for the need to analyze the *s of the pMiSo *ks and *ps clusters as a frozen suffix, pre-pMiSo *-s 'verbalizer', and to leave it out of comparisons. However, in these cases, it seems that the only possible motivation for a pMa *č' or *š is affrication/spirantization of a *k adjacent to a *s. If so, perhaps their irregular pattern, showing up as an affricate in one instance, and as a fricative in the other, could point to borrowing as the more likely explanation. The affricate in question is not just an affricate, but an ejective. This could suggest that a pre-pMa *k' may have been involved originally for this form, as opposed to simply a *k. Perhaps the putative pre-pMiSo *-s 'verbalizer' suffix was present already in pMaMiSo, in some form: if so, a pre-pMa form *mik'-s (or *mok'-s), for example, could have given rise to *mič' (or *moč'), and a pre-pMa form *ʔik-s- could have given rise to *ʔiš-. A viable cognate to this pre-pMiSo *-s 'verbalizer' could be the pMa *-esa ~ *-isa 'causativizer'. Both items in question here, #81 and #172, are transitive verbs in Mayan, and thus, it is plausible that they could bear a frozen causativizer. This is highly speculative at this point, but worth entertaining by searching for additional correspondences of pMa *č(') and *š to pMiSo *ks, which would be predicted to occur with transitive verbs in Mayan. For now I will *not* list these correspondences in the comprehensive list, since there is a chance of diffusion, especially with item #172, which I will consider in a paper in preparation on additional loans between Mayan and Mije-Sokean (Mora-Marín n.d.3).

Table 32. Problematic correspondences of pMa *č' : pMiSo *ʔks and pMa *š : pMiSo *ks.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
	č'	ʔks	81	mič' ~ moč'/mɨɨʔks	To clench, squeeze in hand/To squeeze, extract
	š	ks	172	ʔiš.i, ʔiš.iʔm/ʔɨks, ʔɨks-i	To shell corn, shelled corn/to shell corn, shelled corn

3.2. Onomatopoeia and Internal Reconstruction

Mora-Marín (2016) excluded several potentially onomatopoeic etyma that exhibit irregular correspondences among the Mayan cognates from the initial comparison. For example, Kaufman with Justeson (2003:1180, 1182) reconstruct pCMa *¢'u¢' ~ *¢'ub' as 'to kiss, suckle'. Similarly, they reconstruct pMa *šušub' ~ *šuub' with glosses for the cognates centering around 'to whistle' (2003:754-755). Based on such evidence, as well as a detailed analysis of the Mayan cognates, including variation within individual languages, Mora-Marín (n.d.1) has argued that such irregular correspondences in Mayan yield to internal reconstruction, allowing for the establishment of invariant pre-pMa forms that are more comparable to the pMiSo counterparts, as seen in **Table 33**. Such invariant pre-pMa forms point to the existence of two derivational suffixes of the shape **-Vb'* , one used to derive 'resultative' stems from verbal or verbal noun roots, and the other corresponding to the 'instrumentalizer' derivational suffix that applies to verbs.

Table 33.

#	Mayan	Mije-Sokean	Glosses Ma/MiSo	Class
137	pre-pMa **¢u¢'- (or **¢'u¢-)	*¢uʔ¢	kiss, suckle/suckle, chew	vt/vt
69	pre-pMa **šuuš	*suus(ʔ)	to whistle/to whistle	vi/vi
135	pre-pMa **ʔox-	pMi *ʔoh(V), pMi *ʔohoʔ (vi)	cough/cough; to cough	n/n, vi
138	pre-pMa **t^yuh-	*¢uh, *¢uh-i(C)	spittle/to spit, spittle	n/vi,

				n
153	pre-pMa **wax-	*weh	noise(?)/to shout	n/vi
	pre-pMa **(ty/ch)uq'-	pSo *tuʔk	Hiccup/hiccup	

The resulting pre-pMa forms are readily comparable to their semantically isomorphic pMiSo forms, and they conform to the regular correspondences defined in Mora-Marín (2016). The *¢' : *¢ correspondence has already been discussed in this regard, cf. **Table 22**. In addition, the *š : *s correspondence (correspondence 20) supported with three examples of recurrence, can now be substantiated with pre-pMa *šuuš.ub' and pMiSo *suus(ʔ), as seen in **Table 34**. Similarly, pre-pMa **ʔox.ob' 'cough' and pMi *ʔohoʔ 'cough; to cough' exhibit a one-to-one correspondence: *ʔ : *ʔ, *o : *o, *x : *h, *o : *o, *b' : *ʔ. More importantly, each of the consonantal correspondences has been already established with recurrence; the *x : *h in this etymon is included in **Table 35** below. Next, pre-pMa **t^yuh.ub' 'spittle' and pMiSo *¢uh 'to spit' and *¢uh-i(C) 'to spittle' also show regular correspondences previously established, including the case of *t^y : *¢, cf. **Table 16**. The case of pre-pMa **wax.ab' 'musical instrument' and pMiSo *weh 'to shout' is a bit problematic: the implied pre-pMa root *wax- cannot be glossed at the moment, but it undoubtedly existed, and had a meaning related to 'noise' or even 'loud noise'.

Table 34. *š : *s correspondence

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
20	š	s	89	šiik'/sak	wing, feather/underarm
			3	raʔš/¢us	green/green
			9	k'uš/kɨɨʔs	to bite, gnaw/to pull, bite into something tough
			69	šuuš-/suus(ʔ)	To whistle/to whistle

Another etymon that was included in **Addendum 4** and excluded due to likely onomatopoeia is the case of pMa *¢eʔ 'laughter' and pMiSo *siikʔ, *siikʔ-V 'laugh, smile, laughter'. While it exhibits a previously established correspondence of *¢ : *s, the correspondence of *ʔ : *kʔ is not otherwise attested, and thus, for now, must await confirmation. A similar case is that of pMa *xiq' ~ *xaq' 'to choke', compared to pMi *hiʔkš 'to choke'. The key correspondences are well established in other comparanda: *x : *h, *q' : *ʔk. The final *ʔkš cluster in pMi is assumed to be comparable to pMa *q' only, for pMi *š (< pMiSo *s) has been analyzed above as a frozen pre-pMiSo suffix (*-s). The lingering problem with this form is the variable vowel in the pMa reconstruction. For now, given its agreement with previously established recurring correspondences, I would argue that this etymon should not be excluded. The basic question was whether these etyma should be excluded as possibly onomatopoeic. The answer seems to be no: their patterns of correspondences are too close and consistent with regular correspondences established on the basis of non-onomatopoeic etyma to warrant their exclusion.

3.3. Analogical Reanalysis, Sound Symbolism, Onomatopoeia

Systematic sound symbolic forms, also referred to as synesthetic onomatopoeia, are assumed to be phonologically conventionalized segments (or segment sequences) associated with semantic domains not directly associated with natural sounds, but instead, with size, shape, manner, intensity, etc.

Given the analogical productivity evident in sound symbolism, it is possible that a set of cognates, in this case between pMa and pMiSo, could become more differentiated under the analogical influence of forms with similar meanings. I argue that this is what has happened with #66, pMa *xul ~ *xool 'hole' : pMiSo *hut 'to perforate': they exhibit a regular correspondence in their initial consonant, *x : *h, attested in at least 5 other instances (**Table 31**), but not in the final consonant. I propose this is possibly due to the analogical influence from pMiSo *sut 'to perforate'. Thus, analogical reanalysis (essentially a type of paradigmatic levelling) seems to account for the irregular correspondence. It is possible to tentatively add this item to **Table 35**. If so, then it is expected that pMaMiso would have had *xul for #66. Table 35 also includes two likely onomatopoeic items that exhibit regular correspondences, #135 and #150, as well as an item that exhibits a biconsonantal correspondence but also an unmatched sequence of two segments in MiSo, #82, and a second item, #56, with a biconsonantal correspondence but also two unmatched segments.

Table 35. Evidence for *x : *h correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
15	x-, -x	h-, -h-, -h	22	xo¢'/huʔs	to scratch, shave/to shave
			39	xuč'/hooʔs	to grind corn, dough/nixtamal (cooked corn)
			53	b'ax/paahʔ	to nail/to chisel
			82	kox/kaah[aw][21]	cougar/jaguar
			83	t'ux/tuuh	to drip; drop/rain
			56	LpMa *paax/pah[ak] (or pa[ʔ]ah[k])	sour/sweet
			135	ʔox.ob'/ʔohoʔ	cough/cough; to cough
			150	xiq' ~ xaq' /pMi *hiʔkš	to choke/to choke
			66	xul ~ xool/hu[t]	hole/to perforate
			183	pCMa *lax or *laax/pSo *yah	to end; to finish

Another instance of possible analogical reanalysis is #76, pMa *č'ax 'to wash' and pSo *¢eʔ 'to wash', which exhibit a regular correspondence in their initial consonant, but not the final consonant. This could be due to the analogical influence, this time on the Mayan side, from other forms that exhibit a similar meaning and a final *x: pMa *tʸ'ax 'dripping, splattered' and pMa *t'ux 'drip, drop'. Taking this into account, #76 was included in **Table 20** above. If so, then it is expected that pMaMiso would have had *č'Vʔ for #76.

Another similar case could be #167, pCMa *tʸur 'to cut fruit' and pMiSo *tuk 'to cut fruit', where the pMiSo form could be the result of analogical reshaping due to the influence of

[21] This set, #82, was suggested to me by Kaufman (personal communication 2011). Unfortunately, it contains an unmatched sequence of segments that could disqualify this item.

other MiSo terms related to ‘cutting’ that exhibit a final *k, such as pMi *¢uk ‘to cut with a knife’, pSo *wɨk ‘to slice’, and pMiSo *¢ik ‘to peal’. Given this, the expected pre-pMiSo form would have been *tut, prior to analogical reanalysis, given the correspondence between pMa *r : pMiSo *t. This means that #167 could belong among other cases with the *t^y : *t correspondence, as in **Table 36**.

Table 36. Evidence for *t^y : *t correspondence.

=	pMa	pMiSo	#	pMa/pMiSo lexemes	pMa/pMiSo glosses
11	t^y-	t-	19	t^yox/tɨw	straight/straight
			24	t^yaq/tɨɨ¢ ~ tɨɨ¢ʔ ~ tɨʔɨt¢	dry/to dry out, become thin; dry, thin; tooth
			167	pCMa *t^yu[r]/*tu[k]	to cut fruit/to cut fruit

Other comparanda exhibiting irregular correspondences may eventually yield to this line of analysis. However, for now, the etyma in **Table 37** must remain excluded.

Table 37.

#	pMa			pMiSo		
163	ʔil	to see	v	ʔis	to see	iv, tv
165	peč	nest	n	peʔn; pMi *peʔn-i 'nest'	to make a nest; nest	iv; n

3.3. Irreconcilable Unmatched Segments and Other Problematic Etyma

Table 38 lists a few etyma from Addendum 4 in Mora-Marín (2016) that exhibit unmatched segments or sequences of segments that cannot at this time be reconciled or accounted for and must therefore remain excluded. Comparanda #99 and #100 are excluded because they are nursery forms, for one, and also may have diffused both internally and externally.

Table 38.

43	ERG-tuu[k-eel]	emphatic (reflexive) pronoun	pron	tu[m]	one[22]	num
133	[lu]qum	worm	n	kumu	worm	n
164	[ʔaʔ]¢’aam	salt	n	pMi *[taa]ma¢	salt	n
109	kaʔ[ŋ]	sky	n	¢a[p(-hɨ)]	sky	n
49	ka[q]	red	adj	¢a[pa¢]	red	adj
50	pMa ¢aaʔ ~ LL #taaʔ	excrement	n	tɨɨ[n]ʔ/ti[n]	to defecate/excrement	n
99	taat	father	n	tata	father	n
100	naan	mother	n	nana	mother	n

[22] The numeral one is reconstructed as pMiSo *tum by Kaufman (2007). There is an unmatched final segment in these forms. Brown and Witkowski (1979) first suggested this comparandum (as a cognate).

Two items with unmatched segments, #82 and #56, have been tentatively included among likely cognates illustrating the *x : *h correspondence in **Table 35** because they include biconsonantal correspondences that show regularity, even when the unmatched segments are considered.

3.4. Summary of Sound Correspondences

Table 39 lists 61 sound correspondences, 46 of which were identified in Mora-Marín (2016) and further supported in this paper, and 15 of which constitute additional sound correspondences identified in this paper. **Table 40** lists the sound correspondences arranged in ascending order based on the Mayan sounds. **Table 41** lists the 20 sound correspondences that occur at least four times; the etyma listed in this table thus constitute the most solid foundation for further testing of the pMaMiSo hypothesis and the set to which the BWB method can be expected to yield most promising results.

Table 39. Comprehensive list of sound correspondences sorted for recurrence.

#	pMa	pMiSo	Recurrence	Examples
22	b’	p-…(ʔ)	1x	2
31	č	s	1x	58
32	č’	s	1x	39
35	¢’	s	1x	22
39	q	s	1x	72
44	x	w	1x	19
45	h	y	1x	18
53	¢	s	1x	21
56	Ø	k	1x	67
56	t’	¢	1x	136
57	x	ø	1x	48
58	p	w	1x	162
59	n	m	1x	55
43	x	y	2x	6, 12
12	t^{y}’	t	2x	32, 84
18	b’	ʔ	2x	4, 23
23	ŋ	n	2x	23, 64
26	l	h	2x	149, 80
27	t	¢	2x	20, 160
28	t^{y}’	¢	2x	41, 74
29	t’	t	2x	75, 83
30	č	¢	2x	36, 96
40	q’	s	2x	8, 36
42	r	¢	2x	3, 14
47	ŋ	m	2x	108, 179
50	h	ʔ	2x	62, 78
52	č’	¢	2x	76, 171
54	ʔ	w	2x	15, 44
55	h	w	2x	64, 158
60	t^{y}	¢	2x	40, 138

61	ŋ	y	2x	40, 86
11	t^y	t	3x	19, 24, 167
14	y	y	3x	25, 33, 84
17	ŋ	w	3x	8, 59, 103
24	n	n	3x	7, 158, 182
25	l	y	4x	42, 51, 62, 183
33	¢	¢	3x	77, 88, 137
37	k’	¢	3x	6, 91, 104
38	q	¢	3x	24, 44, 60
41	q’	¢	3x	29, 54, 55
46	ʔ	y	3x	1, 10, 151
6	k’	k	4x	9, 42, 60, 89
9	r	t	4x	13, 16, 18, 57
15	w	y	4x	85, 122, 152, 177
20	š	s	4x	3, 9, 69, 89
34	¢’	¢	4x	46, 34, 38, 137
48	b’	m	4x	78, 79, 80, 92
49	l	m	4x	29, 34, 93, 105
3	l	n	5x	17, 70, 71, 136, 152
13	w	w	5x	12, 51, 122, 151, 153
10	t	t	5x	37, 48, 110, 120, 124
19	s	s	5x	11, 73, 107, 166, 162
4	m	m	6x	26, 58, 81, 91, 101, 160
2	x	h	7x	22, 39, 56, 83, 135, 153, 183
5	k	k	6x	33, 41, 46, 88, 92, 108
36	k	¢	6x	7, 21, 27, 37, 73, 96
16	h	h	7x	86, 87, 125, 138, 149, 152, 152
8	q’	k	7x	4, 11, 13, 20, 63, 101, 105
1	p	p	7x	14, 32, 45, 56, 63, 72, 74
7	q	k	11x	2, 25, 26, 45, 57, 71, 77, 93, 103, 124, 182
21	ʔ	ʔ	18x	16, 17, 27, 40. 54, 59, 70, 79, 85, 86, 87, 103, 120, 122, 125, 135, 166, 179

Table 40. Comprehensive list of sound correspondences organized alphabetically (by pMa).

#	pMa	pMiSo	Recurrence	Examples
33	¢	¢	3x	77, 88, 137
53	¢	s	1x	21
34	¢’	¢	4x	46, 34, 38, 137
35	¢’	s	1x	22
18	b’	ʔ	2x	4, 23, ~~148~~
22	b’	p-…(ʔ)	1x	2
48	b’	m	4x	78, 79, 80, 92
30	č	¢	2x	36, 96

31	č	s	1x	58
32	č’	s	1x	39
52	č’	¢	2x	76, 171
16	h	h	6x	86, 87, 125, 138, 149, 152, 152
45	h	y	1x	18
50	h	ʔ	2x	62, 78
55	h	w	2x	64, 158
5	k	k	6x	33, 41, 46, 88, 92, 108
36	k	¢	6x	7, 21, 27, 37, 73, 96
6	k’	k	4x	9, 42, 60, 89
37	k’	¢	3x	6, 91, 104
3	l	n	5x	17, 70, 71, 136, 152
25	l	y	4x	42, 51, 62, 183
26	l	h	2x	149, 80
49	l	m	4x	29, 34, 93, 105
4	m	m	5x	26, 58, 81, 91, 101, 160
24	n	n	3x	7, 158, 182
59	n	m	1x	55
17	ŋ	w	3x	8, 59, 103
23	ŋ	n	2x	23, 64
47	ŋ	m	2x	108, 179
61	ŋ	y	2x	40, 86
56	Ø	k	1x	67
1	p	p	7x	14, 32, 45, 56, 63, 72, 74
58	p	w	1x	162
7	q	k	11x	2, 25, 26, 45, 57, 71, 77, 93, 103, 124, 182
38	q	¢	3x	24, 44, 60
39	q	s	1x	72
8	q’	k	7x	4, 11, 13, 20, 63, 101, 105
40	q’	s	2x	8, 36
41	q’	¢	3x	29, 54, 55
9	r	t	4x	13, 16, 18, 57
42	r	¢	2x	3, 14
19	s	s	5x	11, 73, 107, 166, 162
20	š	s	4x	3, 9, 69, 89
10	t	t	5x	37, 48, 110, 120, 124
27	t	¢	2x	20, 160
29	t’	t	2x	75, 83
56	t’	¢	1x	136
11	t^{y}	t	2x	19, 24, 167
60	t^{y}	¢	2x	40, 138
12	$t^{y’}$	t	2x	32, 84
28	$t^{y’}$	¢	2x	41, 74
13	w	w	4x	12, 51, 122, 151, 153

15	w	y	4x	85, 122, 152, 177
2	x	h	7x	22, 39, 56, 83, 135, 153, 183
43	x	y	1x	6, 12
44	x	w	1x	19
57	x	ø	1x	48
14	y	y	3x	25, 33, 84
21	ʔ	ʔ	18x	16, 17, 27, 40. 54, 59, 70, 79, 85, 86, 87, 103, 120, 122, 125, 135, 166, 179
46	ʔ	y	3x	1, 10, 151
54	ʔ	w	2x	15, 44

Table 41. 20 most common sound correspondences (at least 4x).

#	pMa	pMiSo	Ocurrences	Examples
6	k’	k	4x	9, 42, 60, 89
9	r	t	4x	13, 16, 18, 57
15	w	y	4x	85, 122, 152, 177
20	š	s	4x	3, 9, 69, 89
34	¢’	¢	4x	46, 34, 38, 137
48	b’	m	4x	78, 79, 80, 92
49	l	m	4x	29, 34, 93, 105
25	l	y	4x	42, 51, 62, 183
3	l	n	5x	17, 70, 71, 136, 152
10	t	t	5x	37, 48, 110, 120, 124
13	w	w	5x	12, 51, 122, 151, 153
19	s	s	5x	11, 73, 107, 166, 162
2	x	h	7x	22, 39, 56, 83, 135, 153, 183
5	k	k	6x	33, 41, 46, 88, 92, 108
36	k	¢	6x	7, 21, 27, 37, 73, 96
4	m	m	6x	26, 58, 81, 91, 101, 160
16	h	h	7x	86, 87, 125, 138, 149, 152, 152
16	h	h	7x	86, 87, 125, 138, 149, 152, 152
1	p	p	7x	14, 32, 45, 56, 63, 72, 74
8	q’	k	8x	4, 11, 13, 20, 61, 63, 101, 105
7	q	k	11x	2, 25, 26, 45, 57, 71, 77, 93, 103, 124, 182
21	ʔ	ʔ	19x	16, 17, 27, 40, 54, 59, 61, 70, 79, 85, 86, 87, 103, 120, 122, 125, 135, 166, 179

Lastly, it is worth revisiting the issue of glottalized obstruents in Mayan raised by Mora-Marín (2016), who argued for a significant correlation between pMa *C’ and pMiSo instances of *Cʔ or *ʔC. To test this idea, **Table 42** compiles the following information: number of instances of pMa *C’, number of instances of pMa *C’- and pMa *-C’, number of instances of pMa *C’ correlated to pMiSo *Cʔ or *ʔC, number of instances of pMa *C’- correlated to pMiSo *Cʔ or *ʔC, and number of instances of pMa *-C’ correlated to pMiSo *Cʔ or *ʔC. I have excluded two etyma, #75 and #137, which exhibit a pMa shape *C’$_1$V(V)C’$_1$, which Mora-Marín (n.d.2) argues to be traceable to pre-pMa *C’$_1$V(V)C$_2$ or *C$_1$V(V)C’$_2$, where *C(’)$_1$ and *C(’)$_2$ are homorganic obstruents. Because of the uncertainty of which of the homorganic obstruents was glottalized in pre-pMa—the root-initial or root-final one—these etyma have been excluded. Given that pMa allows *C’ both root-initially and root-finally, but pMiSo only allows *Cʔ or *ʔC root-finally, I would expect there to be a bias: more correlations between pMa *-C’ and pMiSo *Cʔ or *ʔC (or *C…ʔ…). It would seem that this is supported: 62.5% of pMa *-C’ correlate to pMiSo *Cʔ or *ʔC (or *C…ʔ…), as opposed to 44.4% for cases of pMa *C’-.

Table 42

	Out of total pMa *C’- (18)	Out of total pMa *-C’ (16)	Total pMa *C’ : pMiSo *ʔC
pMa *C’- : pMiSo *ʔC	(8) 44.4%		8
pMa *C’- : pMiSo *C	(10) 55.6%		
pMa *-C’ : pMiSo *ʔC		(10) 62.5%	10
pMa *-C’ : pMiSo *C		(6) 37.5%	
	100.0%	100.0%	52.9% out of total cases of pMa *C’

In total, 52.9% of ejectives (18/34 instances) in pMa correlate with pMiSo *Cʔ or *ʔC (or *C…ʔ). This may not seem significant—it may seem random. However, for one, the distribution is uneven based on position of pMa *C’ (root-initial or root-final), being tipped in favor of pMa *-C’, as expected given the fact that pMiSo does not allow initial *ʔC or *Cʔ clusters. But more importantly, if this were random, one would expect to find a similar proportion of cases of pMiSo *ʔC or *Cʔ correlating to pMa *C (non-ejectives). This is most certainly *not* the case. When one considers the cases of pMa *C_1V(V)(h)C_2 (i.e. not pMa *C_1V(V)ʔC_2) and pMiSo *ʔC or *Cʔ in which *neither* pMa *C_1 nor pMa *C_2 could be ejective (e.g. pMa *s, *ʃ, *m, *n, *w, *y, etc), then only 7 instances out of the total of 119 etyma in the dataset appear: 7/119 = 6%. This is drastically different from the total cases of pMa *C’ (ejectives) that correlate with pMiSo final *ʔC or *Cʔ, 38/119, or 32%. (I have reintroduced the two etyma with pMa *C’ that I had excluded earlier, #75 and #137, because the uncertainty over the position of the pre-pMa *C’ is irrelevant for this count.

A Chi-Square Test of Independence bears out and supports this, as seen in **Table 43**. With all the cells greater than 5, meeting the conditions of the test, it shows a p-value of <.00001 that points to a statistically significant relationship of dependence between pMa *C’ and pMiSo *7C/C7.

Table 43. Chi-Square Test of Independence. Contingency table displays information as follows: observed frequencies for each cell, (expected frequencies for each cell), [chi-square statistic for each cell].

a) Contingency Table

	pMZ *7C	**pMZ *C**	**Row Totals**
pM *C'-	8 (3.60) [5.38]	10 (14.40) [1.34]	18
pM *-C'	10 (3.20) [14.45]	6 (12.80) [3.61]	16
pM *C	7 (18.20) [6.89]	84 (72.80) [1.72]	91
Column Totals	25	100	125 (Grand Total)

b) Chi-Square Statistics

Chi^2	33.4001
p	<.00001

A second test, this time collapsing pMa *C'- and *-C' into a single category, also shows statistically significant results: X^2 (1, N = 114) = 19.25, p < .001.

Table 44. Chi-Square Test of Independence: Presence of pMa ejectives (*C') versus presence of pMiSo *7C or *C7.

a) Observed and expected frequencies

		pMiSo *7C/C7		
		No	Yes	Total
pMa *C' (ejective)	No	60	7	67
		49.96	*17.04*	
	Yes	25	**22**	47
		35.04	*11.96*	
	Total	85	29	114

b) Chi-Square Statistics

Chi^2	19.25
df	1
p	<.001
Cramér's V	0.41 (medium effect size)

Lastly, a third test was carried out to check whether there is a significant relationship between likelihood of pMa *C' correlating with pMiSo *7C/C7, on the one hand, and place of articulation of pMa *C'. The reason for this test was the basic descriptive statistics seen in **Table 45**. Note that cases involving pMa *q' constitute 34% of the total cases of pMa ejectives, and 25.53% of cases of pMa *C' correlating with pMiSo *7C/C7. Hence, the third test, with results seen in **Table 46**, was carried out to compare pMa *q' against other types of pMa *C'. According to the test, cases of pMa *q' are significantly overrepresented with cases of pMiSo *7C/C7, compared to other types of pMa *C'.

Table 45.

		pMiSo *7C/C7					
		No		Yes		Total	
		n	%	n	%	n	%
pMa *C'	¢'	3	6.38%	2	4.26%	5	10.64%
	b'	6	12.77%	3	6.38%	9	19.15%
	č'	1	2.13%	1	2.13%	2	4.26%
	k'	5	10.64%	3	6.38%	8	17.02%
	q'	4	8.51%	12	25.53%	16	34.04%
	t'	2	4.26%	1	2.13%	3	6.38%
	ty'	4	8.51%	0	0%	4	8.51%
	Total	25	53.19%	22	46.81%	47	100%

Table 46. Chi-Square Test of Independence: pMa Ejectives according to place of articulation (Uvular/Non-Uvular) versus presence of pMiSo *7C or *C7.

a) Observed and expected (italics) frequencies

		pMiSo *7C/C7		
		Yes	No	Total
pMa *C’ Articulation	Uvular	**12**	<u>4</u>	16
		7.49	*8.51*	
	Non-uvular	<u>10</u>	**21**	31
		14.51	*16.49*	
	Total	22	25	47

b) Chi-Square Statistics

Chi^2	7.74
df	1
p	.005
Cramér’s V	0.41 (medium effect size)

I suspect a combination of factors was at work. First, pMaMiSo had *C’ (ejectives) in both root-initial and root-final positions. At some point, pre-pMiSo began a process of ejective fission, *C’ > *Cʔ or *ʔC. And then, pre-pMiSo further underwent a process or reduction of initial consonant clusters, disallowing cases of *Cʔ (or *ʔC) as onsets. Under such a scenario, the following steps may have taken place: 1) pMaMiSo *C’ > pre-pMiSo *Cʔ or *ʔC (ejective fission) in both root-initial and root-final positions; 2) pre-pMiSo *Cʔ- clusters > pMiSo *C-, resulting in deletion of many cases of *ʔ, but others were relocated to coda position, i.e. pre-pMiSo *CʔVC > pMiSo *CVC or *CVCʔ or *CVʔC. During both steps of the process, some instances of *ʔ resulting from ejective fission were deleted, but more so in cases where the ejective was root-initial than root-final, and less so in cases where pMaMiSo had *C’ with pMa *q’ as a reflex.

So far, I have ignored the cases of pMa *b’. In Mora-Marín (n.d.2) it is argued that pMa *b’ exhibits a different distribution from all pMa *C’ that were ejectives, and therefore, that pMa *b’ was likely already an implosive, /b’/, rather than an ejective, /p’/. For example, in pMa the only roots of the shape *C’$_1$VC’$_2$ are those where *C’$_1$ and *C’$_2$ are homorganic ejectives. There is only one instance of a pMa root of the shape *C$_1$VC’$_2$, pMa *kik’ ‘blood’. Mora-Marín (n.d.2) argues that starting in pMa, a process of ejective assimilation between homorganic obstruents began to take hold. However, pMa *b’ does not participate in these restrictions. As long as either *C’$_1$ or *C’$_2$ is *b’, then pMa can have roots of the shape *C’$_1$VC’$_2$ without the homorganic restriction, i.e. *C’$_1$Vb’ and *b’VC’$_2$ in which *C’$_1$ and *C’$_2$ can correspond to any ejective obstruent present in the pMa inventory. **Table 47** provides the instances of pMa *b’ in the proposed cognates with pMiSo. Two of them are root-initial: one of these, #53, corresponds to pMiSo *p…(C)ʔ, while the other, #2, simply to pMiSo *p. These two cases could be collapsed into a single correpondence, obviously, thus showing minimal recurrence: pMa *b’ : pMiSo *p. Four intervocalic instances of pMa *b’ correlate with pMiSo *m. Thus, out of the total of 8 instances of pMa *b’ in the dataset, one involves a likely process of ejective fission, #53. The other two can be explained as a result of a process encountered across Mayan languages, wherein final *b’ is often realized phonetically as [ʔ], undergoing a process of debuccalization. The fact that there are no obvious cases of pMa *C’ ejectives corresponding to pMiSo *ʔ, whether in initial or final position, supports a different behavior of pMa *b’ with respect to pMa *C’ ejectives.

Table 47.

pMa	pMiSo	Occurrences	Examples
b'-	p…ʔ	1x	53
b'-	p	1x	2
-b'-	m	4x	78, 79, 80, 92
-b'	ʔ	2x	4, 23

4. Morphological Comparanda

Previously, Mora-Marín (2016:167-168, Addendum C) presented 12 grammatical comparanda as evidence for either cognacy or archaic loans. These correspond to the first 12 items listed in **Table 48**. He explained that 6 exhibit only a single consonant to compare, while the other 6 do exhibit biconsonantal correspondences. Furthermore, he noted that some exhibit consonantal correspondences also attested in the loanwords dataset, while a few exhibit correspondences attested in the primary and secondary lexical datasets but not in the loanword dataset, and thus constituting stronger evidence for cognacy. For example, the correspondence of pMa *w to pMiSo *y is present in items #122 and #177. Mora-Marín (2016) also noted the agreement of pMa *-Ø and pMiSo *-Ø, both as 'third person singular ergative', but noted that this is common in morphologically ergative-absolutive languages, and does not constitute strong evidence for a relationship.

Table 43

#	pMa	pMiSo	Gloss	Class
16	ʔar	ʔit	there is/to be, exist	part/vi
120	teʔ	teʔ	demonstrative, locative/demonstrative	part/part
121	nu+	na+	1sERG/1sERG	procl/procl
122	waʔ	yɨʔ	demonstrative/demonstrative	part/part
123	-o(h/w) ~ -V	-wɨ	plain (indicative) status of TV/completive of independent TV/IV	suff/suff
124	-taq ~ -tiq	-tɨk(ay)	pluralizer/pluralizer	suff/suff
125	haʔ	he(ʔ)	demonstrative base/demonstrative base	part/part
126	ʔu+	ʔi+	3sERG/3sERG	procl/procl
129	-e-Vn	-ʔin	imperative of root IV/optative	suff/suff
177	-ow ~ -aw	pSo *-ʔoy	antipassivizer/antipassivizer	suff/suff
178	-ik	-ik	nominalizer/deverbalizer	suff/suff
118	-i(-k) ~ -i-h	-hi	plain (indicative) completive IV/completive dependent	suff/suff
127	-a(h/w) (R) ~ -Vŋ (D)	-wɨ	Imperative of TV/vetative	suff/suff
109	-o-ex ~ -ex	^ay[23]	perfect participle/perfective	suff/suff
180	-a-ʔ	-ɨʔ (> -aʔ)	dependent (of root transitives)/imperative	suff/suff

[23] Wichmann (1995:545) reconstructs pMiSo *^ay 'perfective' (where the circumflex, <^>, indicates that "glottal stop and/or length [have been] disregarded in the rendering of the canical shape"), but does not describe it in detail. Kaufman and Justeson (2004:1087) do not reconstruct such form, but they do reconstruct a pMiSo *-nay7 'perfect/progressive' suffix, presumably different form the enclitic reconstructed by Wichmann.

181	-i-sa (~ -e-sa)	**-s	causativizer/verbalizer	suff/suff

First of all, I must correct a mistake that I made in Mora-Marín (2016): item #148, comparing pMa *-a-b' ~ *-b' ~ *-i-b' to pMiSo *-ʔ ~ *-Vʔ is invalid. I misread a source alluding to an instrumentalizing suffix of the shape *-ʔ* for San Miguel Chimalapa Soke as saying that it was a Sokean reconstruction. This is not the case, and therefore this item must be removed for now; it is possible that careful examination of other Mije-Sokean languages will show this suffix to be represented more widely, but at the moment there is no evidence for this.

Five additional items also shown in Table 41 are assessed next. The first is pMa *-i-k ~ *-i-h 'plain completive of intransitives' and pMiSo *-hi 'completive dependent'. Clearly, the pMa variant *-i-h resembles the pMiSo form *-hi, but several assumptions are required to support their cognacy: a significant shift in their meaning/function; and regularity in their sound correspondences. The former is problematic, and at first seemingly disparate. The latter, if *-i-h and *-hi are compared, seems consistent with previously established correspondence 16, of pMa *h : pMiSo *h. But the issue becomes problematic if pMa *-i-k is brought into the picture. I return to this set below, before arriving at a conclusion.

The second set for comparison, pMa *-a(w) 'imperative of root transitives' and pMiSo *-wɨ 'vetative (negative imperative of dependent verbs)', is more straightforward: the meaning/function of the suffixes is closer, and the correspondence of pMa *-a(w) to pMiSo *-wɨ is more likely, since pMa *w : pMiSo *w is already attested as correspondence 13. For now, this set seems promising and will be incorporated as a potential cognate.

The third set, pMa *-o-ex 'perfect participle' and pMiSo *^ay 'perfective' may exhibit a closer match in meaning than the first two, as well as a correspondence between pMa *x and pMiSo *y that was documented above with two occurrences (#6, #12). Like the previous set, this one seems promising too.

The fourth set involves pMa *-a-ʔ 'dependent of root transitives' and either pMiSo *-ɨʔ (> *-aʔ) 'imperative'. The resemblance in form is strong; and a shift from imperative to dependent could result from the use of imperatives in quotative constructions. Both correspondences of pMa *a : pMiSo *ɨ and pMa *a and pMiSo *a are well attested. But the risk of a chance resemblance cannot be discounted.

More speculatively, Mora-Marín (2016:170-171) proposed that the MiSo *ks and *ps clusters could point to a frozen derivational suffix of the shape **-s. He argued that overwhelmingly, pMiSo roots with final *Cs clusters are verbs, whether transitive or intransitive, and proposes, through internal reconstruction, a **-s 'verbalizer' of pre-pMiSo. Wichmann (1995:73) does not even include final CC clusters, including *ps and *ks clusters, among the major non-verb canons of pMiSo. Indeed, there exist very few exceptions—four in total—for pMiSo, pMi, and pSo etyma: for *ps, pMiSo *ʔi:ʔps '20' and pMi *hɨn-tɨpʃ 'gusano de mariposa (caterpillar)'; for *ks, pMiSo *piʔks (adj?) 'punteado (pointed)', and pSo *neʔks (adj) 'pegajoso (sticky)'. Otherwise, all other cases seem to be verbs. It is possible that this **-s 'verbalizer' could be cognate with the pMa suffix *-i-sa ~ *-e-sa 'causativizer', which derives transitives from non-transitive roots. However, it is much too early to support this proposal.

At this point it is worth considering the problem of multidimensional paradigmaticity, or the possibility of correspondences not only in form/function but also distribution (Kaufman 1990; Nichols 1996). Several of the proposed cognates belong to the respective paradigms for verb status inflection in Mayan and Mije-Sokean. Table 44 lists the paradigm for pMiSo by Kaufman and Justeson (2004:1083), while Table 45 lists the paradigm for pMa by Kaufman (2015:278-279, 416). The forms in bold letters are suggestive of correspondence between the two paradigms, and the superindexes point to the proposed correspondences between morphemes. Kaufman (2015:195) explains that in pMa the plain status, when unmarked for aspect, represented the completive.

Table 44. pMiSo verb status paradigm.

<table>
<tr><td></td><td>Independent</td><td>Dependent</td></tr>
<tr><td>Incompletive</td><td>*-pa</td><td>*-e (> *-i after V_{high})</td></tr>
<tr><td>Completive</td><td>*-wʉ[1]</td><td>*-hi[2]</td></tr>
<tr><td>Imperative</td><td>*-ɨʔ[3] (> *-aʔ[3] after V_{mid})</td><td></td></tr>
<tr><td>Vetative</td><td></td><td>*-wɨ[4] (neg. dep. imp)</td></tr>
<tr><td>Optative</td><td>*-ʔin[5] (pSo only)</td><td></td></tr>
<tr><td>Subjunctive</td><td></td><td>*-ɨ (> *-a after V_{mid}; Proto-Soke)</td></tr>
</table>

Table 45. pMa verb status paradigm.

<table>
<tr><td></td><td colspan="2">Root and derived transitives</td><td>Root intransitives</td></tr>
<tr><td>Incompletive participle/gerund</td><td colspan="2">*-o-al ~ *-al</td><td>*-e-al</td></tr>
<tr><td>Plain/indicative (completive)</td><td colspan="2">*-o(h/w)[1] ~ *-V/h</td><td>*-i(-k)[2] ~ *-i-h[2]</td></tr>
<tr><td>Imperative</td><td>*-a(h/w)[4] ~</td><td rowspan="2">*-Vŋ</td><td>*-e-Vn[5]</td></tr>
<tr><td>Dependent</td><td>*-a-ʔ[3] ~</td><td>*-oq</td></tr>
<tr><td>Perfect</td><td colspan="2">*-o-ʔm ~ *-ʔm</td><td>*-i-naq</td></tr>
<tr><td>Perfect participle/gerund</td><td colspan="2">*-o-ex (active) ~ *-ex, *-b’il (passive) ~ *-b’il</td><td>*-e-ʔm</td></tr>
</table>

Mayan and Mije-Sokean differ in one crucial way: Mayan splits the paradigm systematically according to transitivity (and then further by distinguishing root transitives from derived transitives), whereas Mije-Sokean does so according to the status of the clause as dependent or independent; while Mayan does exhibit dependent status suffixes, Mije-Sokean does not have transitivity markers of any kind. Given these provisos, one can explore the similarities and attempt to detect a consistent pattern of correspondence in the distribution of the morphemes with form/function similarities, i.e. multidimensional paradigmaticity.

A more detailed and comprehensive overview is thus needed. For example, the completive status markers for independent and dependent verbs in pMiSo, *-wɨ and *-hi, respectively, bear a resemblance to the plain status markers for (root) transitive and intransitive verbs in Mayan (*-o(h/w), *-i(-k) ~ *-i-h), which code the completive when the verb is otherwise explicitly unmarked with aspect-marking prefixes (Kaufman 2015:194-195). Thus, there is a dual correspondence: two pMiSo suffixes that mark the completive status are phonologically comparable with two pMa suffixes that mark the completive when no explicit aspect-marking prefix or particle is present. What is not apparent is why the marker of independent verbs in pMiSo would correspond to the marker of root transitives in pMa, or why the marker of dependent verbs in pMiSo would correspond to the marker of root intransitives in pMa.

Similarly, the imperative of root transitives of pMa, *-a(h/w), bears a similarity to the vetative (negative imperative of dependents), *-wɨ, of pMiSo. Also, the (Mije-)Sokean 'optative', *-ʔin, resembles the pMa 'imperative of root intransitives', *-e-Vn. And lastly, the dependent of root transitives in pMa, *-a-ʔ, resembles the pMiSo imperative, *-ɨʔ ~ *-aʔ.

Of course, there are problems, already noted earlier, with the formal correspondences: the unmarked nature of the consonants involved, and the fact that these markers only bear one consonantal correspondence. Such factors increase the risk of chance resemblances. To validate these proposals, a systematic account of how an ancestral paradigm of verb status suffixes could give rise to the respective pMa and pMiSo paradigms would be necessary. I am nowhere near such a stage in my research. Nevertheless, when the whole paradigm and subparts of the paradigm are considered, the following correspondences seem to be involved, in terms of general grammatical categories: pMa 'plain/indicative' : pMiSo 'completive'; pMa 'imperative' : pMiSo 'vetative'; pMa 'dependent' : pMiSo 'imperative', etc. Addtionally, formal similarities between individual grammatical morphemes, and coincidence in sets of formal similarities are suggestive of the multidimensional paradigmaticity principle: not only are the phonological resemblances close on a case-by-case basis, e.g. pMa *-i(-k) ~ *-i-h : pMiSo *-hi, but also, similar forms within pMa, such as *-o(h/w) 'plain/indicative completive of transitives' and *-a(h/w) 'imperative of transitives', correlate with similar forms among the corresponding grammatical categories within pMiSo, such as *-wɨ 'completive independent' and *-wɨ 'vetative'. These seem to be interrelated correspondences of the nature highlighted by both Kaufman (1990) and Nichols (1996), though perhaps not showing the level of arbitrariness of Meillet's "shared aberrancies" ([1925]1958).

Thus, there are form/function similarities among suffixes, consistent with the sound correspondences proposed, and even coincidences in their patterns of internal distribution. What is currently lacking is a broader pattern that would allow for an account of the development of the respective systems that would explain the implied shifts within the paradigms: Why would the imperative of root intransitives in pMa correspond to the optative of independent verbs in pMiSo; why would an imperative of root transitives in pMa correspond to the vetative of independent verbs in pMiSo; why would the plain (completive) status of root intransitives in pMa correspond to the completive dependent status in pMiSo, while the plain (completive) status of root transitives in pMa corresponds to completive independent status in pMiSo?

5. Conclusions

The approach outlined by Mora-Marín (2014, 2016) has been applied to additional lexical and grammatical comparanda, allowing for the expansion of the dataset of possible cognates between Mayan and Mije-Sokean. It has also increased the number of sound correspondences, 61 in total, 20 of which show four or more occurrences. Some examples that were previously excluded due to the presence of non-corresponding segments can be explained as cases that underwent analogical reanalysis, possibly due to synesthetic sound symbolism, in either Mayan or Mije-

Sokean. Other examples must remain excluded despite overall similarities due to unexplained sequences of unmatched segments, among other violations of the criteria established for allowed comparisons. Lastly, additional grammatical morphemes were discussed, and a comparison of distributional patterns between the two verbal morphology paradigms, specifically those involving aspect and mood marking, was carried out. The results point, preliminarily, to a potential agreement in terms of multidimensional paradigmaticity (Nichols 1996); this line of research thus seems promising, and further grammatical correspondences should be sought.

References

Brown, Cecil H. 2017a. Evaluating Proposals of Language Genealogical Relationship: The Beck-Wichmann-Brown (BWB) System. *Language Dynamics and Change* 7:252-285.

-----. 2017b. Online Appendix: Evaluating Proposals of Language Genealogical Relationship: The Beck-Wichmann-Brown (BWB) System. *Language Dynamics and Change* 7.

Brown, Cecil, Eric W. Holman, and Søren Wichmann. 2013a. Sound correspondences in the world's languages. *Language* 89:4-29.

-----. 2013b. Sound correspondences in the world's languages: online supplementary materials. *Language* 89.

Brown, Cecil H., Søren Wichmann, and David Beck. 2014. Chitimacha: A Mesoamerican language in the Lower Mississippi Valley. *International Journal of American Linguistics* 80:425-474.

Brown, Cecil H., and Stanley R. Witkowski. 1979. Aspects of the Phonological History of Mayan-Zoquean. *International Journal of American Linguistics* 45:34-47.

Campbell, Lyle, and Terrence Kaufman. 1976. A Linguistic Look at the Olmecs. *American Antiquity* 41:80-89.

_____. 1980. On Mesoamerican Linguistics. *American Anthropologist* 82:850-857.

_____. 1983. Mesoamerican Historical Linguistics and Distant Genetic Relationship: Getting It Straight. *American Anthropologist* 85:362-372.

Fox, James. 1978. Proto-Mayan Accent, Morpheme Structure Conditions, and Velar Innovations. Unpublished Ph.D. Thesis, University of Chicago.

Hock, Hans Heinrich. 1991. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.

Kaufman, Terrence. 1990. Language History in South America: What We Know and How to Know More. In *Amazonian Linguistics: Studies in Lowland South American Languages*, edited by Doris L. Payne, pp. 13-67. Austin: University of Texas Press.

-----. 2015. Mayan Comparative Studies. URL: https://www.albany.edu/ims/pdlma/2015%20Publications/Kaufman-Mayan%20Comparative%20Studies.pdf. Accessed January of 2017.

-----. 2017. Aspects of the Lexicon of Proto-Mayan and Its Earliest Descendants. In *The Mayan Languages*, edited by Judith Aissen, Nora C. England, and Roberto Zavala Maldonado, pp. 62-111. London and New York: Routledge/Taylor and Francis Group.

Kaufman, Terrence, and John Justeson. 2004. Epi-Olmec. In *The Cambridge Encyclopedia of the World's Ancient Languages*, edited by Robert D. Woodard, pp. 1071-1108. Cambridge University Press.

Kaufman, Terrence, with John Justeson. 2003. *Preliminary Mayan Etymological Dictionary*. http://www.famsi.org/reports/01051/index.html.

Lees, Robert B. 1953. The Basis of Glottochronology. *Language* 29:113-127.

Mora-Marín, David. 2011. The Proto-Maya-Mijesokean Hypothesis: Change and Transformation in Approaches to An Old Problem. Paper presented at the 44th Annual Chacmool Archaeological Conference, Calgary, Alberta, Canada, Saturday, November 12th, 2011.
-----. 2014. The Proto-Maya-Mijesokean Hypothesis: Change and Transformation in Approaches to an Old Problem. In *Climates of Change: The Shifting Environment of Archaeology, Proceedings of the 44th Annual Chacmool Conference*, edited by Sheila Kulyk, Cara G. Tremain and Madeleine Sawyer, pp. 213-225. Chacmool Archaeological Association, University of Calgary.
-----. 2016. Testing the Proto-Mayan-Mije-Sokean Hypothesis. *International Journal of American Linguistics* 82:125-180.
-----. n.d.1. Actions of the Mouth: Internal Reconstruction of Certain Mayan Sound Symbolic Forms. Unpublished manuscript.
-----. n.d.2. Distributional Analysis and the Internal Reconstruction of Pre-Proto-Mayan Glottalized Consonants. Unpublished manuscript.
-----. n.d.3. Additional Evidence of Loanwords Between Mayan and Mije-Sokean Languages. Manuscript under preparation.
Nichols, Johanna. 1996. The comparative method as heuristic. *The comparative method reviewed: regularity and irregularity in language change*, edited by Mark Durie and Malcolm Ross, pp. 39-89. Oxford: Oxford University Press.
Wichmann, Søren. 1994. Mixe-zoquean linguistics: a status report. Panorama de los estudios de Las lenguas indígenas de méxico, ed. Doris Bartholomew, Yolanda Lastra, and Luis Manrique, Vol. 1, pp. 193–267. Quito: Abya-Yala.
-----. 1995. *The Relationship Among the Mixe-Zoquean Languages of Mexico*. Provo: University of Utah Press.
-----. 1999. A conservative look at diffusion involving mixe-zoquean languages. Archaeology And Language II: Archaeological Data and Linguistic Hypotheses, ed. Roger blench and Matthew Spriggs, pp. 297–323. London: Toutledge.
Witkowski, Stanley R., and Cecil H. Brown. 1978. Mesoamerican: A Proposed Language Phylum. *American Anthropologist* 80:942-944.
_____. 1981. Mesoamerican Historical Linguistics and Distant Genetic Relationship. *American Anthropologist* 83:905-911.