Syllable nuclei of proto-Mayan disyllabic stems

Søren Wichmann
Max Planck Institute for Evolutionary Anthropology & Leiden University
Cecil H. Brown
Northern Illinois University

1. Introduction.[1] This study complements the investigation of proto-Mayan (pM) syllable nuclei of Brown and Wichmann (2004).[2] In the latter work, the syllable nuclei of pM monosyllabic stems are reconstructed. Syllable nuclei of pM disyllabic stems constitute the focus of the present study.[3]

Brown and Wichmann (2004) reconstruct ten syllable nuclei types for proto-Mayan: *V, *V_S, *VV, *V*h*, *VV*h*, *V', *VV', *V'*h*, *VV'*h*, and *V*j* (where V = short vowel, V_S = special vowel, VV = long vowel, ' = glottal stop, *h* = glottal fricative, and *j* = velar fricative). [4] While Brown and Wichmann (2004) present cognate sets demonstrating the reflexes of each of these nuclei types, the cognates are all monosyllabic forms, i.e. forms of the shape CXC (where C = consonant, and X = syllable nucleus). The rationale behind this selective procedure was to work out the syllable nucleus inventory from reflexes in identical contexts before engaging with issues of differential behavior due to factors such as stress, vowel harmony, or consonants being affected by a stem-medial position.

[1] We are grateful to Lloyd Anderson for comments on an earlier version of this paper.

[2] Brown and Wichmann (2004) should be consulted for background materials on Mayan languages such as their names, locations, and classification within the Mayan language family.

[3] See Fox (1978) for some earlier generalizations about pM disyllabic stems.

[4] The orthography used here for Mayan words is the same as that employed in Brown and Wichmann (2004).

This study extends our previous work by examining how pM syllable nuclei behave in disyllabic stems showing the canon CXCXC.[5] The present analysis extends only to CXCXC forms that are either non-analyzable roots or roots having non-productive, fossilized derivational affixes. As an example of the latter, pM **xangab'* 'sandals' and **xuulub'* 'horn' are included since -V*b*' is not productive in any Mayan language anymore even though it is recognizably a nominalizing instrumental suffix ('walk'-INSTR for 'sandals' and 'poke'-INSTR for 'horn'). An example of an excluded item is **yaab'-iil* 'illness' ('sick'-STATE OF), whose suffix *-iil* is still productive in some Mayan languages. In contrast, a lexical item such as **aatz'aam* 'salt' is included since it is non-analyzable, i.e., neither of its two syllables on its own has a recognizable meaning or function. We use 'disyllabic stem' as a convenient way of collectively referring to a non-analyzable disyllabic root or a root containing non-productive affixes

2. The proposed system. Table 1 presents contemporary Mayan syllable-nucleus reflexes of pM nuclei in each of the two syllabic positions of CXCXC forms. A nucleus immediately followed by a hyphen, e.g., V-, indicates first syllable position. A nucleus immediately preceded by a hyphen, e.g., -V, indicates second syllable position. V is to be understood as a cover symbol for the vowels *e*, *i*, *o*, and *u*. For Itzaj, Mopan, Chol, Chontal, and Tzotzil the vowel *a* is treated separately because its reflexes have varying vowel qualities in these languages depending on their position in the CXCXC stem and the type of nucleus to which they pertain. A pM syllable nucleus heads each column of Table 1. Below each pM nucleus are pertinent nucleus reflexes in the individual languages. For example, the Yucatec (Yuc[6]) reflex of pM -*V (see third column of Table 1) is V. Data supporting generalizations of Table 1 are found in the Appendix which lists

[5] Disyllable stems of the canon CXCCXC also occur in Mayan languages, but are vastly less common than those of the CXCXC pattern.

[6] Mayan language abbreviations are (in the order of Table 1): Yuc (Yucatecan), Itz (Itzaj), Mop (Mopán), Hua (Huastec), Chl (Chol), Chn (Chontal), Ch'r (Ch'orti'), Tzo (Tzotzil), Tze (Tzeltal), Tek (Teko), Mam (Mam), Ixh (Ixhil), Awa (Awakateko), Pqch (Poqomchi'), Kaq (Kaqchikel), Tz'ut (Tz'utujiil), K'ich (K'iche'), Q'eq (Q'eqchi'), Moch (Mocho), Q'an (Q'anjob'al), Aka (Akateko), Jak (Jakalteko), Toj (Tojolab'al), Chu (Chuj).

disyllabic stems reconstructed for proto-Mayan in alphabetical order, given with their proposed meanings and their reflexes in Mayan languages.[7]

Table 1. Contemporary Mayan syllable-nucleus reflexes of pM nuclei in each of the two syllabic positions of CXCXC forms.

Proto-Form	*V-	-*V	*VV-	-*VV	*V*h*-	-*V*h*	*V'-	-*V'
Yuc	V	V	V [$V^{L}V$]	V [$V^{L}V$]	$V^{H}V$	V [$V^{L}V$]	V^{H}'V	V^{H}'V
Itz	V/*ä*	V/*a* [V/*ä*]	V/*a*	V/*a*	V/*a* [VV/*aa*]	V/*a* [VV/*aa*]	V/? [V'V/a'a]	V'V/*a'a*
Mop	V/*ä*	V/*a* [V/*ä*]	V/*a*	V/*a*	V/*a* [VV/*aa*]	V/*a* [VV/*aa*]	V/? [V'V/a'a]	V'V/*a'a*
Hua	V	V	V [VV]	VV	V	V		V
Chl	V/*ä*	V/*ä*	V/*a*	V/*a*	V*h*/*ah*	V/*a* [V*h*/*ah*]	V/? [V/ä]	V/*a* [V/*ä*]
Chn	V/*ä*	V/*ä*	V/*a*	V/*a*	V/*a*	V/*a*		V/*a* [V/*ä*]
Ch'r	V	V	V	V	V*h*	V [V*h*]		V
Tzo	V/*a* [V/*o*]	V/*a* [V/*o*]	V/*a*	V/*a*	V/*a* [V/*o*]	V/*o*	V/? [V/o]	V/*a* [V/*o*]
Tze	V	V	V	V	V*h*	V [V*h*]	V	V
Tek	V	V	VV	VV	VV	VV	V [V']	V'
Mam	V	V	VV	VV	VV	VV	V [V']	V'
Ixh	V	V	V	V	V	V	V [V']	V'
Awa	V	V	VV	V [VV]	VV	VV		V'
Pqch	V	V	V [VV]	VV	V*h*	VV [V*h*]		V'
Kaq	V	V	V [VV]	VV	V [VV]	VV		V'
Tz'ut	V	V	VV	VV	V*h*	VV [V*h*]	V [V']	V'

[7] Sources for words from individual Mayan languages are the same as those used in Brown and Wichmann (2004:163-164

K'ich	V	V	V [VV]	VV	V [VV]	VV	V [V']	V'
Q'eq	V	V	V	V	VV	V [VV]	V	V
Moch	V	V	VV	VV	VV	VV	V [V']	V'
Q'an	V	V	V	V	V	V	V	V
Aka	V	V	V	V	V	V	V	V
Jak	V	V	V	V	V	V	V	V
Toj	V	V	V	V	V*h*	V [V*h*]	V [V'V]	V [V'V]
Chu	V	V	V	V	V	V	V [V'V]	V'V

For the most part, reflexes of pM nuclei are the same in disyllabic stems as in monosyllabic stems. However, there are 23 instances in which a reflex of a pM syllable nucleus occurring in the first position of a disyllabic stem is different from that found in a monosyllabic stem. Also, there are altogether 19 instances in which a reflex occurring in the second position of a disyllabic stem is different from that found in a monosyllabic stem. Different reflexes in CXC stems are indicated in square brackets in Table 1.

CXCXC stems provide little or no evidence for the nuclei reconstructed as *Vs, *VVh, VV', *V'h, and VV'h in Brown and Wichmann (2004), and *V*j* also does not occur. Thus, the theoretically possible combinations of nuclei in first and second position, which would be 100 (10 x 10), is much reduced. Table 2 indicates the number of tokens of each attested combination of nuclei. It shows that in 82 cases (80%) either one or both of the two slots is occupied by a simple (*V) nucleus, and that in the remaining 18 cases (20%) both slots are occupied by a complex nucleus (*VV, *Vh, *V'). It also shows that there are differences in overall frequences: *V > *VV > *Vh > V' (read: "*V is more frequent than *VV", and so on). Surprisingly, perhaps, the data do not indicate any clear preference for complex nuclei to appear in one of the two slots. Overall there are 47 tokens of complex nuclei in the first syllable and 54 such tokens in the second syllable, and individual types also do not show any clear preferences.

Table 2. Attested tokens of combinations of syllable nuclei in proto-Mayan CXCXC stems

	*V-	*VV-	*Vh-	*V'-
-*V	23	11	12	3
-*VV	17	10	3	0
-*Vh	12	1	3	0
-*V'	4	3	1	0

In daughter languages the tendency to reduce complexity in CXCXC forms continues in different ways. In 11 individual-language reflexes of proto-nuclei containing glottal stops these glottals are lost, although CXC forms preserve them. In 7 cases a reflex is short in comparison to the corresponding reflex of the same nucleus type in a CXC context. Such shortening takes place both in initial and final syllables in various languages. Only in a scattered minority of cases does a reflex in the CXCXC context come out as what looks like a long vowel as opposed to a short one in the corresponding CXC context reflex. But all these cases involve vowel qualities—e.g., the Tzo V/*a* reflex of *V- in CXCXC stems as opposed to V/*o* in CXC stems. It is likely that such differences are not really to be interpreted as involving vowel quantity, but rather as differences in the treatment of vowel qualities having to do with relatively recent processes of assimilation or dissimilation due to the presence of the vowel in the other, adjacent syllable.

Below, we have extracted from the Appendix all the pM disyllabic stems that relate to each of correspondence series of Table 1. These stems are organized into sets defined by nuclei of second syllables. Within sets, stems are organized by canons.

pM -*V/-a***

CVCVC: **ab'at* 'messenger', **ajan* 'elote', **aqan* 'foot', **awal* 'milpa, sown field', **b'aqal* 'corncob', **b'eleng* 'nine', **ch'upaq* 'soap plant', **ekaj* 'axe', **elaq'* 'theft', **jalab'* 'loom', **jayum* 'to yawn', **kakaw* 'cocoa', **k'olol* 'oak', **k'utub'* 'finger span', **majan* 'lad', **pahay* 'skunk', **pojow* 'pus', **poqoq* 'dust', **sanik* 'ant', **sib'aq* 'soot', **tahab'* 'tumpline, twenty', **xangab'* 'sandals' **CVVCVC:** **aab'ang* 'stone, plum',

aaq'in* 'field work', **chiiwoh* 'tarantula', **hoonon* 'bumblebee', **jaawan* 'woman's son-in-law', **keenaq'* 'beans', **k'aak'as* 'bad, evil', **looqoq* 'mud', **muuxan* 'a kind of plant', **xuulub'* 'horn' **CV*h*CVC: **ahk'al* 'watery area, fertile ground, flat land', **ahq'ab'* 'night', **ahq'ol* 'above', **b'ahlam* 'jaguar', **ihtz'in* 'younger sibling', **kahlam* 'jaw', **kahwoq* 'thunder', **kohtom* 'coati', **kuhkay* 'firefly', **pehtaq* 'prickly pear', **sahb'in* 'weasel', **yahlang* 'below, under' **CV'CVC:** **k'e'wex* 'soursop', **k'u'k'um* 'feather', **pu'huy* 'bird sp.'

pM -*VV/-*_aa_

CVCVVC[8]**:** **al(')iib'* 'daughter-in-law', **b'ehoom* 'rich', **chaluun* 'kind of tree', **ikaan* 'uncle', **ityaaj* 'greens', **jukuub'* 'canoa', **keleem* 'young', **majaan* '(a) loan', **pataan* 'tribute', **pixaan* 'soul', **q'anaal* 'fatness' **saqiil* 'squash seed', **taq'aang* 'savanna', **tyaqiing* 'dry' **CVVCVVC**[9]**:** **aak'aach* 'female turkey', **aalaaq'* 'pet, domesticated animal', **aamaaq'* 'patio', **aatyooty* 'house', **aatz'aam* 'salt'**,** **koob'aan* 'a kind of chili', **puumuuy* 'a kind of dove', **siijoom* 'soap' **CV*h*CVVC**[10]**:** **ahyiin* 'crocodile', **ehtaal* 'mark, image', **ihchaam* 'woman's brother-in-law'

pM -*V*h*/-*_ah_

CVCV*h*C: **ajahw* 'lord, master', **atihn* 'bath, to bathe', **ikahq'* 'nephew, cousin', **inuhp* 'ceiba', **kaqahj(l)* 'palo mulato', **kyitahm* 'peccary', **k'exoohl* 'namesake', **ky'ajahng* 'rope', **lajuhng* 'ten', **peteht* 'spindle', **toq'ohr* 'willow', **tyaq'ahng* 'ripe, cooked', **CVVCV*h*C:** **aanahm* 'earth, dirt', **CV*h*CV*h*C:** **hahlahw* 'tepescuintle', **mahtahn* 'gift', **mahtzahb'* 'eyebrow, eyelash'

pM -*V'/-*_a_'

[8] Several of these, i.e., **b'ehoom*, **chaluun*, **jukuub'*, **keleem*, **majaan*, **saqiil*, **taq'aang*, and **tyaqiing*, could alternatively be reconstructed as *CVCV*h*C.

[9] Some of these, i.e., **aak'aach*, **koob'aan*, **puumuuy* and **siijoom* could alternatively be reconstructed as CVVCV*h*C.

[10] A theoretically possible alternative reconstruction is *CV*h*CV*h*C.

CVCV'C: **ab'a'q* 'soot', **ati't* 'grandmother', **ixi'm* 'corn', **same't* 'comal', **CVVCV'C:** **aala's* 'toy, game', **aaq'e'n* 'tray, mat, trough', **tz'uunu'n* 'hummingbird', **CVhCV'C:** **sihna'ng* 'scorpion'

3. Context-conditioned changes. The following are descriptions of most changes pertaining to the nuclei of pM disyllabic stems that have reflexes in contemporary Mayan languages different from those reported in Table 1. These differences relate to conditioning environments, which are identified.

3.1. Final laryngeals. Reflexes of pM **a* and **aa* behave differently from those charted in Table 1 when these two pM vowel types are in the second syllable of a disyllabic stem and are immediately followed by a stem-final laryngeal, either -*'# or -*h#. Table 3 presents reflexes of pM -**a*'#, -**aa*'#, -**ah*#, and -**aah*#. Also given in brackets are expected reflexes in monosyllabic stems whenever these differ from those of the disyllables (see Table 10 in Brown and Wichmann 2004).

Table 3. Reflexes of pM **a* and **aa* immediately followed by a stem-final laryngeal.

Proto-Form	*-a'#*	*-aa'#*	*-ah#*	*-aah#*
Yuc	*-a'#*			
Itz				
Mop	*-a'#*	*-a'#*		
Hua				*-a'#* [*aa*C#]
Chl	*-a'#*	*-a'#*	*-a#* [*ah#*]	*-a#* [*ah#*]
Chn	*-a'#*	*-a'#*	*-a#* [*ah#*]	*-a#* [*ah#*]
Ch'r			*-ah#*	*-ah#*
Tzo	*-o'#*	*-a'#*	*-oh#/-ow#* [*a#*]	*-a#*
Tze	*-a'#*	*-a'#*	*-a#*	*-a#*
Tek		*-a'#* [*aa'#*]		*-aa'#* [*aa'#*]
Mam		*-a#* [*aa'#*]		*-a#* [*aa'#*]
Ixh		*-a'#*		*-a#*

				[*aj*#]
Awa	*-a'*#	*-a'*#		*-aa'*#
		[*aa'*#]		
Pqch	*-aa'*#	*-aa'*#		
	[*a'*#]			
Kaq	*-a'*#	*aa'*#		*-aay*#
		[*a'*#]		
Tz'ut	*aa'*#	*-aa'*#		*-aay*#
	[*a'*#]			
K'ich	*aa'*#	*-aa'*#		*-ah*#
	[*a'*#]			[*aa'*#]
Q'eq	*-a'*#		*-a*#	
Moch	*-aa'*#	*-aa'*#	*-ah*#	Ø#
	[*a'*#]			[*aah*#]
Q'an	*-a'*#	*-a'*#		*-a*#
Aka	*-a'*#	*-a'*#	*-a'*#	*-a*#
			[*aa*#]	[*aa*#]
Jak		*-a'*#	*-a'*#	*-a*#
			[*ah*#]	[*ah*#]
Toj	*-a*#	*-a'*#	*-ah*#	*-a*#
	[*a'*#]			[*ah*#]
Chu		*-a'*#	*-a*#	*-a*#

Presented below are the pM disyllabic stems which relate to each of the four series of Table 3 (see the Appendix for reflexes).

pM -***a'**# **huula'* 'visitor'

pM -***aa'**# **umaa'* 'mute', **kooyaa'* 'tomato', *meeb'aa'* 'orphan, widow'

pM -***ah**# **patah* 'guayaba'

pM -***aah**# **tzimaah* 'guacal'

Since the generalizations of Table 3 are not based on abundant examples of cognate sets, they should be regarded as tentative and possibly subject to future revision.

3.2. Yucatecan languages. CXCXC proto-forms showing **j*, **w*, or **h* as a medial consonant and having an **a* or **ah* as the nucleus of their first syllables undergo a special development in Yucatecan languages (Yuc, Itz, and Mop). This involves loss of the medial **j*, **w*, or **h*, deletion of the nucleus of the second syllable, and lengthening of the

vowel of the first syllable. In Yucatec, the acquired long vowel shows high tone.[11] In formulaic summary,

1. pM *CX_1jXC or *CX_1wXC or *CX_1*h*XC

> Yuc C*áa*C, Itz and Mop C*aa*C, where X_1 = **a* or **ah*

Proto-forms illustrating shift 1 include **kahwoq*, **pahay*, **tahab'*, **ky'ajahng*, and **majaan*. Yucatecan language reflexes of the latter are respectively Yuc *cháak*, Itz and Mop *chaak*, Yuc *páay*, Itz and Mop *paay*, Yuc *táab'*, Itz *taab'*, Yuc *k'áan*, Itz and Mop *k'aan*, and Mop *maan*-.

In disyllables exhibiting a short *a* in the first syllable and an *i*-colored vowel in the second syllable the *a* goes to *i* by vowel harmony. In summary,

2. pM **a*- > Yuc, Itz, Mop *i*-/*__CXC, where X = *i*, *ii* or *ih*

Proto-forms illustrating shift 2 include **atihn*, **al(')iib*', **sanik*, **tyaqiing*, and **sakiil*. Reflexes in Yucatec are *ichíin*, *ilib'*, *síinik*, *tikin*, and *sikil*. Both Itz and Mopan have the irregular reflex *äli'* of **al(')iib*', while Itz attested other reflexes are *sinik* and *sikil* and Mop attested other reflexes are *ichn*-, *sinik*, *tikin*, and *sikil*.

3.3 Huastec. Disyllable proto-forms having **j* and **w* as a medial consonant and having an **a* or **ah* as the nucleus of their first syllables and that have a word-initial consonant other than a glottal stop (or no word-initial consonant at all) undergo a special change in Huastec (Huastecan subgroup). This entails loss of the medial **j* or **w*, deletion of the nucleus of the first syllable, and lengthening of the vowel of the second syllable (if not already long) while maintaining the quality of the vowel found in the nucleus of the second syllable. In summary,

[11] Yucatec has three tones, high (V^H), low (V^L), and neutral (V).

3. pM *$C_1X_1jX_2C$ or *$S_1X_1wX_2C$ > Hua CX_3C, where $C_1 \neq$ *', where X_1 = **a* or **ah*, where X_3 = VV, and where the vowel quality of X_3 = the vowel quality of X_2

Proto-forms illustrating shift 3 include **kahwoq* and **ky'ajahng*. Huastec reflexes of the latter are respectively, *tzook* and *tz'aah*.

3.4. Chol and Chontal. In two Cholan languages, Chol and Chontal, pM **a*- regularly shifts to *ä*-. However, when the medial consonant is **j*, **a*- develops as *a* in the two languages. In summary,

4. pM *C*a*C_1XC > Chl and Chn C*a*CXC, where C_1 = **j*

Proto-forms illustrating shift 4 include **ajahw*, **ky'ajahng*, and **majaan*. Chl and Chn reflexes of the latter are respectively, *ahaw*, *ch'ahan*, and *mahan*. Chl *halb'* from pM **jalab'* suggests that pM **a*- also develops as Chl *a* when **j* is stem-initial.[12]

In Chol, pM **a*- also shifts to *a* when the medial consonant is **h*. In summary,

5. pM *C*a*C_1XC > Chl C*a*CXC, where C_1 = **h*

Proto-forms illustrating shift 4 include **pahay* and **tahab'*. Chl reflexes of the latter are respectively, *pahäy* and *tahb'*.

In Chontal, pM **a*- regularly shifts to *ä*- (see Table 1). However, when **a*- is immediately preceded by **p*, it develops as Chn *a*-. In summary,

6. pM *C_1*a*CXC > Chn C*a*CXC, where C_1 = **p*

[12] In monosyllabic stems, pM **a* regularly shifts to Chl and Chn *ä*. However if a **j* also occurs in such stems, in certain circumstances **a* develops as *a* in the two languages. For details, see Brown and Wichmann (2004: 152).

Proto-forms illustrating shift 6 include **patah* and **pataan*. Chn reflexes of the latter are respectively, *pata* and *patan*.

3.5. Ch'orti'. Disyllabic proto-forms showing **j* as a medial consonant and having a short vowel as a first syllable nucleus undergo a special change in Ch'orti' (Cholan). This entails replacing **j* with ' (glottal stop) while deleting the nucleus of the second syllable. In summary,

7. pM *CV*j*XC > Ch'r CV'C

Proto-forms illustrating shift 7 include **ajan*, **pojow*, **majan*, **pahay*, **tahab'*, and **ky'ajahng*. Ch'orti' reflexes of the latter are respectively, *a'n*, *po'w*, *ma'n*, *pa'y*, *ta'b'*, and *ch'a'n*.

3.6. Tzotzil. Proto-Mayan **a-* and **ah-* yield *a-* in Tzotzil (Tzeltalan) (See Table 1). However, when the nucleus of the second syllable is -**a*, -**ah*, or -**a'*, pM **a-* and **ah-* yield *o-* in the language. In summary,

8. pM **a-* and **ah-* > Tzo *o-*/C___CX_1C, where X_1 = -**a*, -**ah*, or -**a*'

Proto-Mayan -**a*, -**ah*, and -**a'* yield -*a* in Tzotzil (see Table 1). However, when the nucleus of the first syllable is **a-* or **ah-*, pM -**a*, -**ah*, and -**a*' yield -*o* in the language. In summary,

9. pM -**a*, -**ah*, and -**a'* > Tzo -*o*/CX_1C___C, where X_1 = **a-* or **ah-*

Proto-forms illustrating interrelated shifts 8 and 9 include **ab'a'q*, **ahq'ab'*, **awal*, **jalab'*, **kahlam*, **kakaw*, **kyaq'ahng*, **patah*, **yahlang*, **xangab'*, *tyaq'ahng*, **ajahw*, **mahtzahb'*, **hahlahw*, **b'ahlam*, **ahq'ab'*, and **mahtahn*. Tzo reflexes of the latter are respectively, *ob'ok*, *ok'ob'*, *owol*, *holob'*, *kolom-*, *kokow*, *ch'ohon*, *potoh*, *olon*, *xonob'*, *tok'on*, *ohow*, *motzob'*, *holow*, *b'olom*, *ok'ob'*, and *moton*.

Exceptions to shifts 8 and 9 are pM **aqan* and **b'aqal* which yield respectively Tzo *akan* and *b'akal.* Tzo *okon* and *b'okol* are expected. Also, pM **ajan* yields Tzo *ahan* rather than expected *ohon*. Apparently, shifts 5 and 6 are blocked when first and second nuclei are both **a* and either **q* or **j* is the stem-medial consonant.

Disyllabic proto-forms of the canon *C*aha*C undergo a special change in Tzotzil. This entails loss of the medial **h* and deletion of the nucleus of the second syllable. The **a* which remains shifts to *o*. In summary,

10. pM *C*aha*C > Tzo C*o*C

Proto-forms illustrating shift 10 include **pahay* and **tahab'*. Tzotzil reflexes of the latter are respectively, *poy* and *-tob'*.

3.7. **Teko**. Proto-Mayan -*V*h* and -*VV yield -VV in Teko (Mamean) (see Table 1). However, when found adjacent to **j*, -*V*h* and -*VV yield -V in Teko. In summary,

11. pM -*V*h* and -*VV > Tek -V/*CXC$_1$___C$_2$, where C$_1$ = **j* or C$_2$ = **j*

Proto-forms illustrating shift 11 include **ityaaj*, and **lajuhng*. Teko reflexes of the latter are respectively, *itzaj* and *laajuj*.

Second-position pM **a* is deleted in Teko when the stem-final consonant is **q*'. In summary,

12. pM *CVC*aq*' > Tek CVC*q*'

Proto-forms illustrating shift 11 include **keenaq'* and **elaq'*. Teko reflexes of the latter are respectively, *kiinq'* and *elq'*-.

3.8. Mam. The second position nucleus of a disyllabic proto-form is deleted in Mam (Mamean) when the following conditions hold: (1) the pM stem begins with a vowel,[13] (2) the vowel quality of vowels of both first position and second position is low central, (3) at least one of the two vowels is not a short vowel, i.e., not **a*, and (4) the second position nucleus is not **a*'. In summary,

13. pM *#$X_1C_1X_2C_2$ > #$X_1C_1C_2$, where both X_1 and X_2 are low central in vowel quality, where X_1 and X_2 are not both **a*, and where X_2 is not **a*'.

Proto-forms illustrating shift 13 include **ajahw*, **ahq'ab'*, **aab'ang*, and **aamaaq'*. Mam reflexes of the latter are respectively, *-ajw*, *aq'b'-*, *aab'j*, and *aamq'*.

Mam deletes the nucleus of a first syllable of a disyllabic proto-form when the nucleus is **a*. In summary,

14. pM *C*a*CXC > Mam CCXC

Proto-forms illustrating shift 14 include **aqan*, **sanik*, **tyaq'ahng*, **tyaqiing*, and **chaluun*. Mam reflexes of the latter are respectively, *qan*, *sniky*, *tzq'aaj*, *tzqiij*, and *txluun*.

3.9. Ixhil. Disyllabic proto-forms showing either **j* or **ng* as a medial consonant undergo a special development in Ixhil (Mamean).[14] This involves loss of the medial **j* or **ng* and deletion of the nucleus of the first syllable accompanied by the change of the nucleus of the second syllable to a long vowel (if not already a long vowel) while maintaining the quality of the vowel found in the nucleus of the second syllable. In summary,

[13] More precisely, if initial vowels are assumed to be preceded by phonemic glottal stops, the Mam deletion regularity entails disyllabic words beginning with glottal stops.

[14] Proto–Mayan **j* and **ng* both are realized as **j* in Proto–Mamean.

15. pM *CX_1jX_2C or *CX_1ngX_2C > Ixh CX_3C, where X_3 = VV and where the vowel quality of X_3 = the vowel quality of X_2

Proto-forms illustrating shift 15 include **xangab'*, **pojow*, and **ky'ajahng*. Ixh reflexes of the latter are respectively, *xaab'*, *poow*, and *k'aa*.

3.10. Awakatek. Proto-Mayan -*V*h* yields -V in Awakatek (Mamean) (see Table 1). However, when -*V*h* is -**uh*, the Awa reflex is -*u*. In summary,

16. pM -**uh* > Awa -*u*

Proto-forms illustrating shift 15 include **inuhp* and **lajuhng*. Awa reflexes of the latter are respectively, *unup* and *lajuj*.

3.11. Poqomchi'.[15] Proto-Mayan **ah*- yields *ah*- in Poqomchi' (K'iche'an) (see Table 1). However when **ah*- is the onset of the first syllable of a disyllabic proto-form, **ah*- is realized as *a* in the language.[16] In summary,

17. pM **ah*- > Pqch *a*-/*#___CXC

Proto-forms illustrating shift 17 include **ahk'al* and *ahq'ab'. Poqomchi' reflexes of the latter are respectively, *ak'al* and *aq'ab'*.

In disyllables exhibiting a short *a* in the first syllable and an *i*-colored vowel in the second syllable the *a* goes to *i* by vowel harmony. In summary,

18. pM *a- > Pqc i-/*__CXC, where X = *i* or *ii*

[15] Possibly, the Poqomchi' V from *VVh, *VV, and *VV' in first syllables of disyllabic words is not a short vowel, but rather a vowel that has been neutralized for length.

[16] If initial vowels are assumed to be preceded by phonemic glottal stops this shift involves disyllabic words beginning with glottal stops rather than with vowels.

Proto-forms illustrating shift 18 include **sanik* and **tyaqiing*, which respectively yield *sinik* and *chiqiij* in Pqch. Since a similar shift is seen in Yucatecan (see 2 above) this process could have diffused from the latter.

3.12. Kaqchikel. Disyllabic proto-forms showing the canon *CV_1jV_1C lose the nucleus of the second syllable and also the final consonant of the second syllable in Kaqchikel (K'iche'an). In summary,

19. pM *CV_1jV_1C > Kaq CV_1j

Proto-forms illustrating shift 19 include **ajan* and **pojow*. Kaqchikel reflexes of the latter are respectively *aj* and *puj*.

3.13. K'iche'. Disyllabic proto-forms showing the canon *CV_1jV_1C lose the nucleus of the second syllable and also the final consonant of the second syllable in K'iche' (K'iche'an). In summary,

20. pM *CV_1jV_1C > K'ich CV_1j

Proto-forms illustrating shift 20 include **ajan* and **pojow*. K'iche' reflexes of the latter are respectively *aj* and *puj*.

Proto-Mayan *VV- yields V- in K'iche'. However, when a disyllabic stem's onset consonant is any consonant except *' (glottal stop) and when the stem-medial consonant is **b*', *VV- yields VV- in the language. In summary,

21. pM *VV- > K'ich V-/*C_1___b'XC, where $C_1 \neq$ *'

Proto-forms illustrating shift 21 include **koob'aan* and **meeb'aa'*. K'iche' reflexes of the latter are respectively, *koob'aan* and *meeb'aa'*.

3.14. Tz'utujil. Disyllabic proto-forms showing the canon *CV_1jV_1C lose the nucleus of the second syllable and also the final consonant of the second syllable in Tz'utujil (K'iche'an). In summary,

22. pM *CV_1jV_1C > Tz'ut CV_1j

Proto-forms illustrating shift 22 include **ajan* and **pojow*. Tz'utujl reflexes of the latter are respectively *aj* and *puj*.

Proto-Mayan **ah*- yields *ah*- in Tz'utujil (see Table 1). However, when **ah*- is the onset of the first syllable of a disyllabic proto-form, **ah*- is realized as *aa* in the language.[17] In summary,

23. pM **ah*- > Tz'ut *aa*-/*#___CXC

Proto-forms illustrating shift 23 include **ahk'al*, **ahyiin* and **ahq'ab'*. Tz'utujil reflexes of the latter are respectively *aak'al*, *aayiin*, and *aaq'ab'*.

Proto-Mayan *VV- yields VV- in Tz'utujil (see Table 1). However, when the disyllabic stem's medial consonant is **n*, *VV- is realized as V in the language. In summary,

24. pM *VV- > Tz'ut V-/*C___*n*XC

Proto-forms illustrating shift 24 include **hoonon*, **keenaq'*, and **tz'uunu'n*. Tz'utujil reflexes of the latter are respectively *wonon*, *kinaq'*, *tz'unun*.

[17] If vowels are assumed to be preceded by phonemic glottal stops this shift involves disyllabic words beginning with glottal stops rather than with vowels.

3.15. Q'eqchi'. Disyllabic proto-forms showing **j*, **ng*, **w*, or **h* as a medial consonant and having an **a-* or **ah-* as the nucleus of their first syllable undergo a special development in Q'eqchi' (K'iche'an). This involves loss of the medial consonant, deletion of the nucleus of the second syllable, and changing **a-* or **ah-* to *aa-*. In summary,

25. pM *CX_1jX_2C or *CX_1wX_2C or *CX_1hX_2C > Q'eq C*aa*C, where X_1 = **a* or **ah*

Proto-forms illustrating shift 25 include **xangab'*, **pahay*, **tahab'*, **ky'ajahng*, and **kahwoq*. Q'eqchi' reflexes of the latter are respectively, *xaab'*, *paar*, *taab'*, *k'aam*, and *kaaq*.

Disyllabic proto-forms showing the canon *CV_1jV_1C lose the nucleus of the second syllable and also the final consonant of the second syllable in Q'eqchi'. In summary,

26. pM *CV_1jV_1C > Q'eq CV_1j

Proto-forms illustrating shift 26 include **ajan* and **pojow*. Q'eqchi' reflexes of the latter are respectively *aj* and *poj*.

Disyllabic proto-forms with second syllable -**a*, which is adjacent to **q* or **q'*, lose the nucleus of the second syllable in Q'eqchi'. In summary,

27. pM *$C_1XC_2aC_3$ > Q'eq $C_1XC_2C_3$, where one of the consonants C_1 or C_2 = **q* or **q'*

Proto-forms illustrating shift 27 include **aqan*, **elaq*, and *keenaq'*. Q'eqchi' reflexes of the latter are respectively *aqn*, *elq'*, and *kenq'*.

3.16. Mocho. Proto-Mayan *V*h-* yields VV in Mocho (Q'anjob'alan) (see Table 1). However, when **ah-* is immediately followed by stem-medial **q'*, **ah-* develops as Mocho *a*. In summary,

26. pM **ah-* > Moch *a-*/*C___*q*'XC

Proto-forms illustrating shift 28 include **ahq'ab'* and **ahq'ol*. Mocho reflexes of the latter are respectively, *aq'ab'* and *aq'ol*.

3.17. Akatek. Disyllabic proto-forms showing **j* as a medial consonant undergo a special change in Akatek (Q'anjob'alan). This involves loss of the medial **j* and deletion of the nucleus of the second syllable accompanied by the change of the nucleus of the first syllable to a long vowel (if not already a long vowel) while maintaining the quality of the vowel found in the nucleus of the first syllable. In summary,

29. pM *CX_1jX_2C > Aka CX_3C, where X_3 = VV and where the vowel quality of X_3 = the vowel quality of X_1

Proto-forms illustrating shift 29 include **ajahw*, **majaan*, **majan*, **ajan*, **pojow*, and **lajuhng*. Akatek reflexes of the latter are respectively, *aaw*, *maan*, *maan*, *aan*, *poow*, and *laan-*.

4. Discussion. In this paper, as well as in Brown and Wichmann (2004) we have followed four principles of reconstruction:

1. each series of unconditioned correspondences should be matched by a distinct reconstructed element;
2. when multiple reflexes of a given reconstructed element are found in a given language, these should be accounted for by special phonological or morphological conditioning environments;
3. the number of exceptions to conditioned and unconditioned correspondences should be minimal.
4. the reconstructed elements should be as natural as possible, i.e., found in the given group of languages (preferably), in the given area, or in other languages of the world (minimally).

We doubt that any practitioner of the comparative method will disagree with these four principles. Different scholars, however, may weight the relative importance of each principle differently. In Brown and Wichmann (2004) we were able to satisfy principle (1) by a relatively large set of nucleus types which, together with an extensive set of conditioned changes (principle 2) minimized exceptions (principle 3) to the extent that 93% of all individual language reflexes were accounted for. Among the 10 nucleus types, however, several are not found in extant Mayan languages although nearly all are found in Mesoamerica, and all of them are found in some language in the world (cf. discussion in Brown and Wichmann 2004: 158-161). An item for future research continues to be a search for conditioning environments (principle 2) which would reduce the set of nucleus types such as to make for more natural reconstructions (principle 4). Thus far, however, no proposals to this effect have been published.

In the present work we also adhere to the four principles and reconstruct four different nucleus types which head the major correspondence series (principle 1) and which, together with conditioned reflexes (principle 2), minimize irregulaties (principles 3) such that of the among the 2232 reflexes of syllables (including lost ones and syllables in variant forms within one and the same language) in the sets of cognates in the Appendix only 142 or 6% behave in a way not accounted for. Finally, in contrast to our previous work on nuclei occurring in monosyllabic roots, we were able to reconstruct a very simple inventory of syllable nuclei (*V, *VV, V*h*, *VV) (principle 4) without introducing an unacceptable amount of irregularity in reflexes.

Inasmuch as several principles are involved, reconstruction is to a certain extent a matter of choice. It would be possible to reduce the amount of irregularity by around one percent, by introducing the nucleus types *V_s, *VV*h* or *VV', but the cost of introducing more complexity in the inventory of nuclei appearing in pM CXCXC stems seems much greater than the cost of having to qualify a small number of reflexes as being irregular. Moreover, it stands to reason that there should be more differentiation among possible nuclei in CXC forms than in CXCXC forms given that the latter have more phonological material—a whole additional syllable—available for differentiating between different lexemes. Finally, it is quite possible that the complexity of monosyllables arose

historically as a compensation for loss of segments, whereas disyllables, which are less likely to have suffered phonological reduction, may reflect an inventory of nucleus types more similar to that of some pre-proto-stage of Mayan than do pM monosyllables.

It should be mentioned that while the set of nuclei occurring in disyllabic stems is phonologically simple and natural there is one case, namely that of -*V*h*, where the reconstruction is actually fairly abstract. The reason that -*V*h* is best seen as an abstract reconstruction is that a segment *h* does not appear among the regular[18] reflexes of -*V*h* in any Mayan language. The reason why we nevertheless reconstruct -*V*h* is that unless one wants to introduce types such as *V_s or *VV*h*, there is no other viable candidate, and it is not possible to do without some reconstruction different from either -*VV or -*V given the abundance of evidence for the existence of the three different correspondence series which we have chosen to summarize by the reconstructions -*VV, *-V, *-V*h*.

5. Conclusion. Above we describe most generalizations relating to syllable nuclei in disyllabic words of the canon CXCXC in Mayan languages. There are others that might be described from existing data, and, surely, more to emerge as new data from Mayan languages and dialects become available.

[18] The Chol and Ch'orti' reflexes of the **peteht* 'spindle' exhibit an *h*, which we are forced to treat as irregular. Possibly, however, the reconstruction is not a valid pM etymon but rather a form which diffused from the one of the Lowland languages to neighboring ones. Moreover, the form is semi-analyzable as consisting of a verb root **pet* plus a rare *-*Vht* suffix.

APPENDIX

In the following cognate sets supporting disyllabic forms of the canon CXCXC reconstrued for Proto-Mayan are presented. Unexpected syllable nuclei are given in **bold** type. No attempt has been made to indicate semantic shifts undergone by some reflexes of the various proto-forms have undergone.

ab'a'q* 'soot': Chl ***a**b'****ä****k*, Tzo *ob'ok*/***o****b'****a****k*, Tze *ab'ak*, Ixh ***a'****b'****a****q*/***aa****b'a'j*/***aa****b'****a****q*, Pqch *ab'****aa****q'*, Tz'ut *ab'a'q*, K'ich *ab'a'q*, Q'eq *ab'aq*, Q'an *ab'aq*, Aka *ab'aj*, Jak *ab'aj*, Toj *ab'ak*.

ab'at* 'messenger': Hua *abaat*, Tze *ab'at*, Moch (Tuz) *ab'**a****t*.

ahk'al* 'watery area, fertile ground, flat land': Yuc *áak'al*, Itz *ak'al*, Chl *ok'ol*, Tzo ***a**ch'el*, Tze *ahch'al*, Ixh *ak'al*, Pqch *ak'al*, Tz'ut *aak'al*, K'ich *ak'al*, Q'an *ak'al*, Aka *ak'al*, Jak *ak'al*.

ahq'ab'* 'night': Yuc *áak'ab'*, Itz *ak'**ä****'*, Mop *ak'****ä****'*, Chn *ak'äb'*, Tzo *ok'ob'*, Tze *ahk'ab'*-, Mam ***a****q'b'*-, Ixh *aq'b'*-, Awa ***a****q'b'*-, Pqch *aq'ab*', Kaq *aq'a'*, Tz'ut *aaq'ab'*, K'ich *aq'ab'*-, Q'eq *aaq'****a****b'*, Moch *aq'ab'*, Q'an *aq'b'*-, Aka *aq'ab'*-, Jak *aq'b'*-.

**ahq'ol* 'above': Yuc *óok'ol*, Itz -*ok'ol*, Mop *ok'ol*, Tzo *ak'ol*, Tze *ahk'ol*, Moch *aq'ol*.

ahyiin* 'crocodile': Yuc *áayin*, Itz *ayim*, Mop *ayin*, Hua *ah**i****n*, Chl *ahin*, Chn ***ä****hin*, Ch'r *ahyin*, Tzo *ain*, Tze *ahyin*, Ixh *ayin*, Kaq *ayiin*, Tz'ut *aayiin*, K'ich *ayiin*, Q'eq ***a****hin*/***a****yin*, Moch (Tuz) *ahiin*, Q'an *ayin*, Toj ***a****yin*, Chu *ayin*.

ajahw* 'lord, master': Yuc -*a****haw*, Hua *ah****aa***-, Chl *ahaw*, Chn *ahaw*, Tzo *ohow*, Tze *ajaw*, Mam -***a****jw*-, Awa *ajaaw*, Pqch -*ajaaw*, Kaq *ajaaw*, K'ich *ajaaw*, Q'eq *ajaw*, Moch *ajaaw*, Q'an *ajaw*, Aka *aaw*, Jak *ahaw*-, Toj *ahaw*-.

**ajan* 'elote': Hua *ahan*, Ch'r *a'n*, Tzo *ahan*, Tze *ajan*, Kaq *aj*, Tz'ut *aj*, K'ich *aj*, Q'eq *aj*, Moch *ahan*, Q'an *ahan*, Aka *aan*, Jak *ahan*, Toj *ahan*, Chu *ajan*.

al(')iib'* 'daughter-in-law': Yuc *ilib'*, Itz ***ä**li'*, Mop ***ä****li'*, Hua *al'iib*, Chl *ä'lib'*, Chn *älib'*, Ch'r *arib'*, Tzo *alib'*-, Tze *alib'*-, Tek *aliib'*, Mam ***i****liib'*, Ixh *alib'*, Awa ***i****liib'*-,

Tz'ut -*alii'*, Q'eq *alib'*, Moch *aliib'*, Q'an *alib'*-, Aka *alib'*-, Jak *alib'*-, Toj *alb'*-, Chu *alib'*.

**aqan* 'foot': Hua *akan*, Tzo *akan*, Tze *akan*, Mam *qan*, Ixh *aqan*, Awa *aqan*, Pqch *aqan*, Kaq *aqan*-, Tz'ut *aqan*, K'ich *aqan*, Q'eq *aqn*, Q'an *aqan*-, Aka *ajan*-, Jak *ajan*, Toj *akan*.

atihn* 'bath, to bathe': Yuc *ichíi**n*-, Mop *ichn*-, Hua *achin*, Tzo *atin*-, Tze *atin*-, Tek *achiin*-, Ixh *achin*, Awa ***ii**ch**i'**n*, Kaq *at**i**n*-, Tz'ut *atiin*-, K'ich -*atiin*, Q'eq *atin*, Moch *achiin*, Q'an *achin*-, Toj *atin*, Chu *ating*-.

awal* 'milpa, sown field': Tzo *owol*, Tze *awal*, Tek -*awal*, Mam ***awal*, Ixh *awal*, Awa *aw**aa**l*, Moch *awal*, Q'an *awal*, Aka *awal*, Jak *awal*, Chu *awal*.

aab'ang* 'stone, plum': Hua *aba*, Chn *ab'än*, Tek *aab'aj*, Mam *aab'j*/*a**b'j*, Pqch *ab'aj*, Kaq *ab'aj*, Tz'ut *aab'aj*, K'ich *ab'aj*.

aak'aach* 'female turkey': Chl *ak'ach*, Ch'r *ak'ach*, Ixh *ak'atx*, Pqch *ak'a**ch*, Q'eq *ak'ach*, Q'an *ak'atx*, Aka *ak'atx*, Chu *ak'ach*.

aalaaq'* 'pet, domesticated animal': Yuc *àalak'*, Itz *alak'*, Mop *alak'*, Chl *alä**k'*-, Chn ***ä**l**ä**k'*-, Ch'r *arak'*, Tzo *alak'*, Tze *alak'*, Toj *alak'*. [If the Toj form is a loan, then this set does not attest to a pM form.]

aala's* 'toy, game': Chl *alas*/*ä**las*, Chn alas/***ä**l**ä**s*, Ch'r *aras*, Mam *aala's*, Ixh *ala's*, Kaq *ala's*, Tz'ut ***a**la's*, K'ich *ala's*.

aamaaq'* 'patio': Tzo *amak'*, Tze *amak'*, Tek *aamaaq'*, Mam *aamq'*, Kaq *amaaq'*, K'ich *amaaq'*, Q'eq *amaq'*, Moch *aama**q'*, Q'an *amaq'*, Aka *amaq'*, Jak *amaq'*, Chu *amak'*.

**aanahm* 'earth, dirt': Hua *anam*, Chl *anam*, Ch'r *anam*.

aaq'e'n* 'tray, mat, trough': Tze *ek'en*, Mam ***aq'**e**n*, Ixh *aq'e'm*, Kaq *aq'**ee**n*, Tz'ut ***a**q'**ee**m*, K'ich *aq'**ee**n*, Moch *aaq'e'n*.

**aaq'in* 'field work': Chl *ak'in*, Ch'r *ak'in*, Tzo *ak'in*, Ixh *aq'in*, Q'an *aq'in*, Aka *aq'in*, Jak *aq'in*, Toj *ak'in*. [If the Ixh form is a loan, then this set does not attest to a pM form.]

aatyooty* 'house': Yuc *otoch*, Itz *otoch*, Hua *ataa*, Chl *otot*, Chn *otot*, Ch'r *otot*, Ixh *otzotz*, Kaq -*achooch*, Tz'ut -*oochooch*, K'ich -*o'**ch*, Q'eq *ochoch*, Q'an -*atut*, Aka *atut*, Chu *atut*.

aatz'aam* 'salt': Hua *at'e**m*, Chl *atz'am*, Chn *atz'am*, Ch'r *atz'am*, Tzo *atz'am*, Tze *atz'am*, Tek *aatz'aam*, Ixh *atz'am*, Awa *aatz'**u**m*, Pqch *atz'aam*, Kaq *atz'aam*, Tz'ut *aatz'aam*, Q'eq *atz'am*, Q'an *atz'am*, Aka *atz'am*, Jak *atz'am*, Toj *atz'am*, Chu *atz'am*.

ati't* 'grandmother': Hua ***aach*, Pqch *ati't*, Kaq *ati't*, Tz'ut *ati't*, K'ich *ati't*. [The Hua form is from Proto-Huastecan (pH) *aatyit. pM *t > pH *ty/___*i or *u.]

b'ahlam* 'jaguar': Yuc *b'áalam*, Itz *b'alum*, Mop *b'aa**lum*, Chl *b'ahlum*, Chn *b'aläm*, Ch'r *b'ahrum*, Tzo *b'olom*, Tze *b'ahlam*, Mam *b'aalan*, Ixh *b'alam*, Awa *b'aalum*, Pqch *b'ahlam*, Tz'ut *b'ahlom*, Q'eq *b'aalam*, Moch *b'aalam*, Q'an *b'alam*, Aka *b'alam*, Jak *b'alam*, Toj *b'ahlam*, Chu *b'alam*.

b'aqal* 'corncob': Yuc *b'akal*, Itz *b'äkä**l*, Mop *b'äk**ä**l*, Chl *b'äkäl*, Chn *b'äkä*, Ch'r *b'akar*, Tzo *b'akal*, Tze *b'akal*, Moch *b'**aa**qal*, Q'an *b'aqal*, Aka *b'ajal*, Jak *b'ajal*, Toj *b'akal*, Chu *b'akal*.

b'eleng* 'nine': Hua *belee**w*, Mam *b'elaj*, Ixh *b'ele*, Awa *b'eluj*, Pqch *b'elej*-, Kaq *b'elej*-, Tz'ut *b'elej*-, K'ich *b'elej*-, Moch (Tuz) *b'elej*-.

b'ehoom* 'rich': Tzo *b'eom*, Pqch *b'ehoom*, Kaq *b'eyoom*, Tz'ut *b'eyoom*, K'ich *b'eyoom*, Q'eq *b'eho**m*, Q'an *b'eyom*, Aka *b'eyom*, Jak *b'eum*.

chaluun* 'kind of tree': Tzo *chalon*, Mam *txluun*, Tz'ut *chalu**m*, Q'eq *chalum*.

**chiiwoh* 'tarantula': Yuc *chìiwoh*/*chìiwol*, Itz *chiwoh*, Mop *chiwoh*/*chiwo*, Chn *-chiwo'*, Ch'r *chiwiw*, Toj *chiwoh*. [If the Toj form is a loan, then this set does not attest to a pM form.]

ch'upaq* 'soap plant': Chl *ch'upuh**k*, Tzo *ch'upak*, Tze *ch'upak*, Tek *tx'upaq*, Mam *ch'upeq*, Pqch *ch'ipaq*, Kaq *ch'upaq*, Tz'ut *ch'apaq*, K'ich *ch'ipaq*, Q'eq *ch'upaq*, Moch *ch'upaq*, Toj *ch'upaq*.

**ehtaal* 'mark, image': Chl *-ehtal*, Ixh *echl*-, Kaq *etaal*, Tz'ut *ehtaal*, Aka *echel*-, Jak *echel*-, Toj *ehtal*, Chu *echel*.

ekaj* 'axe': Tze *echej*, Pqch *ikej*, Kaq *ikaj*/*ikej*/*ikee**j*, Tz'ut *ikaj*, Aka *eche*, Jak *eche*, Toj *echej*.

**elaq'* 'theft': Tzo *elek'*, Tek *elq'*-, Ixh *elaq'*, Awa *alaq'*, Kaq *eleq'*, Tz'ut *elaq'*, Kich *elaq'*, Q'eq *elq'*, Moch *elq'*-, Jak *eleq'*, Toj *elk'*-, Chu *elk'*-.

hahlahw* 'tepescuintle': Chl *ha**law*, Chn *hal**ä**w*, Tzo *holow*, Pqch *hahlaaw*, Q'eq *h**a**law*, Moch (Tuz) *halaaw*[19], Q'an *alaw*, Aka *halaw*, Jak *halaw*, Toj *h**a**law*, Chu *halaw*.

hoonon* 'bumblebee': Ch'r *honon*, Tzo *honon*, Tze *honon*, Mam *oonan*, Ixh *onon*, Awa *oonum*, Kaq *wonoo**n*, Tz'ut *wonon*, K'ich *wonon*, Q'eq *honon*, Moch *hoonon*, Q'an *onon*, Aka *honon*, Jak *honon*, Toj *honon*-, Chu *honon*.

huula'* 'visitor': Yuc ***ú'ula'*, Mop *uda'*, Chl *hula'*, Chn *hula'*, Tzo *hulo'*, Tze *ula'*, Awa *uula'*, Pqch *ulaa'*, Kaq *ula'*, Tz'ut *uulaa'*, K'ich *ulaa'*-, Q'eq *ula'*, Moch *uulaa'*, Q'an *ula'*, Aka *hula'*, Toj *ula*.

**ihchaam* 'woman's brother-in-law': Yuc *íicham*, Itz *icham*, Mop *icham*, Tek *iichaam*, Moch *iichaam*, Q'an *icham*, Aka *icham*, Jak *icham*, Chu *icham*.

**ihtz'in* 'younger sibling': Yuc *íitz'in*, Itz *itz'in*, Mop *itz'in*, Chl *ihtz'in*, Chn *itz'in*, Ch'r *ihtz'in*, Tze *ihtz'in*, Ixh *itz'in*, Awa *iitz'in*, Q'eq *iitz'in*-, Toj *ihtz'in*-.

**ikahq'* 'nephew, cousin': Hua *itzak'*, Chl *ichak'*, Chn *ichak'*, Tzo *ichok'*, Pqch *ikaaq'*, Tz'ut *ikaaq'*.

**ikaan* 'uncle': Hua *itzaan*, Chl *ichan*, Chn *ichan*, Tze -*ichan*, Tz'ut -*ikaan*, K'ich *ikaan*, Q'eq *ikan*.

inuhp* 'ceiba': Hua *unup*, Tzo *inup*, Ixh *inup*, Awa *unu**p*, Pqch *inuup*, K'ich *inuup*, Q'eq *inup*, Moch (Tuz) *inuup*, Q'an *inup*, Aka *inop*, Jak *inup*, Toj *inip*, Chu *inup*.

ityaaj* 'greens': Tzo *itah*, Tek *itzaj*, Mam *itzaaj*, Ixh *itza*, Awa *itza**j*, Pqch *ichaaj*, Kaq *ichaaj*, Tz'ut *ichaaj*, K'ich *ichaaj*, Q'eq *ichaj*, Q'an *itaj*, Aka *ita*, Jak *ita*, Toj *itah*, Chu *itaj*.

ixi'm* 'corn': Yuc *ixí'im*, Mop *ixi'im*, Chl *ixim*, Chn *ixim*, Ch'r *ixim*, Tzo *ixim*, Tze *ixim*, Tek *ixi'm*, Ixh *ixi'm*, Awa *ixi'n*, Pqch *ixii**m*, Kaq *ix**ii**n*, Tz'ut *ix**ii**m*, K'ich *ix**ii**m*, Q'eq *ixim*, Moch *ixi'm*, Q'an *ixim*, Aka *ixim*, Jak *ixim*, Toj *ixim*, Chu *ixim*.

**jalab'* 'loom': Lac *häräb'*, Chl *halb'*-, Tzo *holob'*, Tze *jalab'*, Moch *halb'*-, Q'an *jalb'*-, Jak *halb'*-, Toj *halab'*, *jalb'*-.

**jayum* 'to yawn': Mop *hayam*, Tz'ut *jayam*, Moch *hayum*.

jaawan* 'woman's son-in-law': Itz *hab'ä**n*, Mop *hab'**ä**n*, Hua *hawan*, Chl *hawän*, Chn *hawän*, Tzo *hawan*, Tze -*jawan*.

[19] In the Tuz dialect of Mocho, pM **V*h*–* yields V.

**jukuub'* 'canoa': Chl *hukub'*, Chn *hukub'*, Ixh *jukub'*, Kaq *jukuu'*, Tz'ut *jukuu'*, K'ich *jukuub'*, Q'eq *jukub'*, Moch *hukuub'*, Q'an *hukub'*, Aka *hukub'*, Jak *hukub'*, Chu *jukub'*.

**kahlam* 'jaw': Ch'r *kahram*, Tzo *kolom*-, Q'eq *kaalam*, Jak *kalam*-.

kahwoq* 'thunder': Yuc *cháak*, Itz *chaak*, Mop *chaak*, Hua *tzook*, Chl *chahk*, Chn *chawäk*, Tzo *chauk*, Tze *chahwuk*, Awa *kyooq*, Pqch *kahoq*, Tz'ut *ka**woq*, K'ich *kawoq*, Q'eq *kaaq*, Q'an *ka'eq~kaq'eq*, Toj *ch**a**wuk*.

kakaw* 'cocoa': Itz *ka**kaw*, Mop *käk**ä**h*, Hua *kakaw*, Chl *käkäw*, Chn *käkäw*, Tzo *kokow*, Tze *kakaw*, Tek *kakaw*, Q'eq *kakaw*, Q'an *kakaw*, Jak *kakaw*-, Toj *kakaw*, Chu *kakaw*.

**kaqahj(l)* 'palo mulato': Yuc *chakah*, Itz *chikah*, Mop *chikah*, Hua *tzakah*, Chl *chäkahl*, Ch'r *chakahr*, Tze *chakaj*, Q'eq *kaqaj*, Moch *kaqaah*, Jak *kajah*, Toj *chakah*.

**keleem* 'young': Yuc *-kelem*, Itz *kelen*-, Mop *kelem*, Chl *kolem*, Tzo *kerem*, Tze *kerem*, Ixh *chelem*, Moch *kereem*, Toj *kerem*, Chu *kelem*.

keenaq'* 'beans': Hua *tzanakw'*, Tzo *chenek'*, Tze *chenek'*, Tek *keena**q'*/*kiinq'*, Mam *cheenaq'*, Pqch *kenaq'*/*kinaq'*, Kaq *kinaq'*, Tz'ut *kinaq'*, K'ich *kinaq'*, Q'eq *kenq'*, Toj *chenek'*.

**kohtom* or **qohtom* 'coati': Itz *koton*-, Mop *koton*-, Chl *kohtom*, Tzo *kotom*, Tze *kohtom*, Toj *kohtom*. [If the Toj form is a loan, then this set does not attest to a pM stem.]

**koob'aan* 'a kind of chili': Itz *kob'an*, Mop *kob'an*, Ixh *kob'an*, Tz'ut *koob'aan*, K'ich *koob'aan*, Aka *kob'an*, Chu *kób'an*.

**kooyaa'* 'tomato': Chl *koya'*, Chn *koya'*, Mam *-koo'ya*, Awa *-kooya'*, Kaq *-koyaa'*, Tz'ut *-kooyaa'*, Toj *koyah*.

kuhkay* 'firefly': Yuc *kóokay*, Mop *kukay*, Hua *kukay*, Chl *ku**k*-, Chn *-kukay*, Ch'r *kuhkay*, Tzo *kukay*, Tze *k**u**kay*, Tek *kuukuy*, Mam ***kuukxh***[20], Ixh *kukuy*, Q'an *kukuy*, Chu *kukay*.

**k'aak'as* 'bad, evil': Yuc *k'àak'as*, Itz *k'ak'as*, Mop *k'ak'as*, Hua *k'ak'ath*.

[20] Bold type for the entire form here as well for Mam *k'ol* (see *k'olol below) indicates unexpected loss of the second syllable

**k'e'wex* 'soursop': Yuc *k'é'ew*, Chl *k'ewex*, Tzo *k'ewex*, Tze *k'ewex*, Tek *k'iwixh*, Mam *ch'wix*, Ixh *ch'evew*, Tz'ut *k'ewex*, K'ich *k'awex*, Moch *k'iwix*, Q'an *k'ewex*, Aka *k'ewex*, Jak *ch'iw*, Toj *k'ewex*, Chu *k'ewex*.

k'exohl* 'namesake': Mop *k'exuul*, Chl *k'exo**l*, Ixh *ky'**e'**x*/*ch'**e'**x*, Awa *ky'**e'**x*, Kaq *k'exeel*, K'ich *k'exol*, Q'an *k'exel*, Aka *k'exel*, Chu *k'exul*.

k'olol* 'oak': Chl *k'olol*, Tze *k'olol*, Mam ***k'ol,* Ixh *k'ol*, Awa *k'ol*, Q'an *k'olol*, Toj *k'olol*, Chu *k'olol*.

k'utub'* 'finger span': Hua *ch'uchub*, Tzo *ch'utub'*, Tze *ch'utub'*, Mam *-k'achub'*, Ixh *k'uchub'*, Awa *k'uchub'*, Kaq k'utuu**b'-, Q'eq *k'utub'*, Q'an *q'uchub'*, Aka *k'uchub'*/*q'uchub'*, Jak *k'uchub'*. [The Hua form is possibly a loan from Tzeltalan.]

**k'u'k'um* 'feather': Yuc *k'ú'uk'um*, Itz *k'uk'um*, Chl *k'uk'um*, Tzo *k'uk'um*, Tze *k'uk'um*, Q'eq *k'uk'um*, Moch *k'uk'um*, Aka *k'uk'un*, Jak *k'uk'um*, Toj *k'uk'um*, Chu *k'uk'um*.

**kyitahm* 'peccary': Yuc *kitam*, Itz *kitam*, Mop *kitam*, *chitam*, Chn *chitam*, Ch'r *chitam*, Tzo *chitom*, Tze *chitam*, Ixh *chicham* [K 1974], Moch *chitaam*, Q'an *txitam*, Aka *txitam*, Jak *txitam*, Toj *chitam*, Chu *chitam*.

ky'ajahng* 'rope': Yuc *k'áan*, Itz *k'aan*, Mop *k'aan*, Hua *tz'aah*, Chl *ch'ahan*, Chn *ch'ahan*, Ch'r *ch'a'n*, Tzo *ch'ohon*, Tze *ch'ajan*, Mam *ky'ijaaj*, Ixh *k'aa*, Awa *ky'ajaaj*, Q'eq *k'aam*, Moch *ch'ajaang*, Q'an *tx'an-*, Aka *tx'a**n*, Jak *tx'**a**ng*, Chu *ch'**a**ng*.

lajuhng* 'ten': Yuc *lahun*, Itz *la**hun-*, Hua *l**aa**juj*, Tzo *lahun-*, Tze *lajun-*, Tek *l**aa**juj*, Awa *lajuj*, Kaq *lajuuj*, Tz'ut *lajuuj*, Moch *lah**u**ng*, Q'an *lajon-*, Aka *laan-*, Jak *lahung-*, Toj *lahun-*, Chu *lajn-*.

**looqoq* 'mud': Hua *lukuk*, Moch *looqoq*, Toj *lokok*.

mahtahn* 'gift': Yuc *máatan*, Itz *matan*, Mop *matan*, Chl *mahtan*, Chn *matä**n*, Tzo *moton*, Tze *mahtan*, Q'eq *maatan*, Moch (Tuz) *mataan*[21], Q'an *matan*, Aka *matan*, Jak *matan*, Toj *mahtan*.

[21] In the Tuz dialect of Mocho, pM **V*h*–* yields V.

mahtzahb'* 'eyebrow, eyelash': Yuc *máatzab'*, Itz *matza'*, Mop *matza'*, Hua *matab*, Chl *mä**tzab'*, Ch'r *mahtzo'*, Tzo *motzob'*, Tze *m**a**tzab'*, Q'eq *maatzab'*/*m**a**tzab'*, Toj *m**a**tzab'*, Chu *matzab'*. [A possible collective approach to the phonologically irregular reflexes in this set is to hypothesize that an irregular form, **matzahb'*, developed at some early point and subsequently diffused. An etymon of this shape would account for reflexes in Chl, Tze, Q'eq, and Toj.]

**majan* 'lad': Ch'r *ma'n*, Q'an *mahan*, Aka *maan*. [This set does not meet the distributional criteria for a pM word. It is included since it shows a special correspondence.)

majaan* '(a) loan': Yuc *mahàa**n*, Itz *m**äha**n*, Mop *maan*-, Chl *mahan*, Chn *mahan*, Tze *majan*-, Q'an *mahan*, Aka *maan*, Toj *mahan*, Chuj *majan*.

meeb'aa'* 'orphan, widow': Chl *meb'a'*, Tek *meeb'a'*, Mam *meeb'a*, Ixh *mee**b'a'*, Awa *meeb'a'*, Kaq *meb'aa'*, Tz'ut *meeb'aa'*, K'ich *meeb'aa'*, Moch *meeb'aa'*, Q'an *meb'a'*, Aka *m**ee**b'a'*, Jak *meb'a'*, Toj *meb'a'*, Chu *meb'a'*.

**muuxan* 'a kind of plant': Itz *muxan*, Ixh *moxan*, Tz'ut *muuxan*, Q'eq *mox*, Chu *moxan*.

pahay* 'skunk': Yuc *páay*, Itz *paay*, Mop *paay*, Chl *pahäy*, Ch'r *pa'y*, Tzo *poy*, Tze *pahay*, Kaq *paar*, Tz'ut *paar*, K'ich *paar*, Q'eq *paar*, Q'an *pay*, Aka *pay*, Jak *pay*, Toj *pa'a**y*, Chu *pay*.

**patah* 'guayaba': Mop *päta*, Chl *päta*, Chn *patá*, Ch'r *patah*, Tzo *potoh*/*potow*, Tze *pata*, Q'eq *pata*, Moch *patah*, Aka *pata'*, Jak *pata'*, Toj *patah*, Chu *pata*.

pataan* 'tribute': Yuc *patan*, Chn *patan*, Tzo *patan*, Tze *patan*, Kaq *pata**n*, K'ich *pataan*, Moch *pataan*, Q'an *patan*, Aka *patan*, Jak *patan*, Chu *patan*.

pehtaq* 'prickly pear': Tzo *peto**k*, Tze *pehtak*, Tek *peechaq*, Moch *peechaq*, Aka *petxaj*, Jak *petxaj*, Toj *pehtak*, Chu *pechak*.

peteht* 'spindle': Yuc *peché**e**ch*, Itz *pech**ee**ch*, Mop *pech**ee**ch*, Chl *pete**h**t*, Ch'r *pete**h**t*, Tzo *petet*, Tze *petet*, Q'eq *pete**e**t*, Q'an *petet*, Toj *petet*, Chu *petet*.

pixaan* 'soul': Yuc *pixàa**n*, Itz *pixan*, Mop *pixan*, Chn *pixan*, Tzo <pixan->, Q'an *pixan*-, Aka *pixan*-, Jak *pixan*-, Chu *pixan*.

**pojow* 'pus': Ch'r *po'w*, Tzo *pohow*, Tze *pojow*, Tek *poj*, Mam *poj*, Ixh *poow*, Pqch *poj*, Kaq *puj*, Tz'ut, *puj*, K'ich *puj*, Q'eq *poj*, Moch *pohow*, Q'an *pohow*, Aka *poow*, Jak *pohow*, Toj *pohow*, Chu *pojow*.

**poqoq* ‘dust’: Tzo *pukuk*, Ixh *poqoq*, Pqch *pooq*, Tz’ut *poqoq*, Q’eq *puquq*-, Moch *puquq*, Q’an *poqoq*, Aka *pojoj*, Jak *pojoj*, Chu *pokok*.

puumuuy* ‘a kind of dove’: Ch’r *pumuy*, Kaq -*pumuuy*, Tz’ut -*puumuuy*, Moch *pu**m*, Jak *pum*.

**pu’huy* ‘bird sp.’: Yuc *pú’uhuy*, Itz *puhuy*, Mop *puhuy*, Q’eq *puhuy*. [If the Q’eq form is a loan, then this set does not attest to a pM form.]

**q’anaal* ‘fatness’: Hua *k’anaal*, Tz’ut -*q’anaal*, K’ich -*q’anaal*.

sahb’in* ‘weasel’: Yuc *sáab’in*, Itz *sab’in*, Mop *sab’in*, Hua *theben*, Chl *sahb’in*, Ch’r *sahb’in*, Tze *sahb’in*, Jak *sah**b’in*, Toj *sahb’en*.

**sakiil* ‘squash seed’: Yuc *sikil*, Itz *sikil*, Mop *sikil*, Ch’r *sakir*, Tzo *sakil*, Tze *sakil*, Tek *skiil*, Ixh *sachil*, Kaq *sakiil*, Tz’ut *sakiil*, K’ich *sakiil*, Q’eq *sakil*, Moch *saqiil*, Aka *sachil*, Jak *sachil*, Toj *sakil*, Chu *sakil*.

same’t* ‘comal’: Chl *semeh**t*, Chn *semet*, Ch’r *semet*, Tzo *semet*, Tze *samet*, Ixh *sem**e**tX*, Kaq *seme’t*, Tz’ut *same’t*, Moch *s**a’**m*.

sanik* ‘ant’: Yuc *síi**nik*, Itz *sinik*, Mop *sinik*, Hua *thanitz*, Chl *x**i**nich’*, Chn *x**i**nich’*, Ch’r *x**i**nich*, Tzo *x**i**nich*/*h**i**nich*, Tze *xanich’*, Tek *sanik*, Mam *sniky*, Ixh *sanich*, Awa *snik*, Pqch *sinik*, Kaq *sanik*, Tz’ut *sanik*, K’ich *sanik*, Moch *sanik*, Q’an *sanik*, Aka *sanik*, Jak *sanik*, Toj *san’ich*/*sanich*. [The vowel *i* of the first syllable in the Chn, Chl, Ch’r, and Tzo forms is suggestive of borrowing from Yucatecan, where the change **a* > *i* is a regular process.]

sib’aq* ‘soot’: Yuc *sab’ak*, Itz *säb’ä**k*, Mop *säb’**ä**k*, Chl *sib’ik*, Chn *sib’ik*, Tzo *sib’ak*, Tze *sib’ak*, Moch *sib’aq*, Toj *sib’ak*, Chu *sib’ak*.

sihna’ng* ‘scorpion’: Yuc *síina’ng*, Mop *sina’an*, Hua *thiniy*, Chl *si**Nan*, Chn *sina*, Ch’r *s**i**nam*, Mam *sii’n**a**n*, Awa *siina’j*, Kaq *sina’j*, Tz’ut *sihna’y*, K’ich *sina’j*, Moch *siina’m*, Aka *sinam*, Chu *sina’an*.

siijoom* ‘soap’: Chl *sihom*, Tzo *sihom*, Tze *sijom*, Pqch *sijoh**m*, Tz’ut *siijoom*.

tahab’* ‘tumpline, twenty’: Yuc *táab’*, Itz *taab’*, Chl *tahb’*-, *täb’*-/*tab’*, Ch’r *ta’b’*, Tzo -*tob’*, Tze -*tah**b’*, Q’eq *taab’*, Toj *tahab’*.

**taq’aang* ‘savanna’: Yuc *chak’an*, Itz *chäk’an*, Mop *chäk’an*, Tek *chq’aaj*, Kaq *taq’aaj*, Tz’ut *taq’aaj*, K’ich *taq’aaj*, Q’eq *taq’a*, Moch *chaq’aang*, Jak *chaq’ang*.

**toq'ohr* 'willow': Hua *tok'oy*, Tzo *tok'oy*, Tze *tok'oy*, Tek *toq'ooy*, Moch *toq'ooch*, Jak *taq'oy*, Toj *tok'oy*. [Despite showing expected syllable nuclei, the Tek form is clearly a loan by phonological criteria. The expected form for Tek is *choq'oot*.]

tyaqiing* 'dry': Yuc *tikin*, Mop *tikin*, Chl *täkin*, Chn *ti**kin*, Ch'r *takin*, Tzo, *takin*, Tze *takin*, Mam *tzqiij*, Pqch *chiqiij*, Kaq *chaqiij*, Tz'ut *chaqiij*, K'ich *chaq**i**j*-/*chaq**i'**j*, Q'eq *chaqi*, Moch *taqiing*/*taq**i'**ng*, Q'an *taqin*, Aka *tajin*, Jak *tajing*, Toj *takin*, Chu *taking*. [The vowel *i* of Chn *tikin* suggests borrowing from a Yucatecan language, where the change **a* > *i* is a regular process unlike what is found in Cholan in general.]

tyaq'ahng* 'ripe, cooked': Yuc *tak'an*, Itz *täk'an*, Mop *täk'an*, Chl *täk'an*, Chn *ta**k'an*, Ch'r *tak'an*, Tzo *tok'on*, Tze *tak'an*, Tek *tzq'**a'**j*, Mam *tzq'aaj*, Ixh *tzaq'*, Kaq *chaq'aaj*, Q'eq *chaq'*, Moch *taq'aang*, Toj *tak'an*.

**tzimaah* 'guacal': Hua *tima'*, Chl *tzima*, Ch'r *tzimah*, Tzo *tzima*, Tze *tzima*, Tek *tzimaa'*, Mam *tzii'ma*, Ixh *tzima*, Awa *tzimaa'*, Kaq *tzimaay*, Tz'ut *tzimaay*, K'ich *tzimah*, Moch *tzi'm*, Q'an *tzima*, Aka *tzima*, Jak *tzima*, Toj *tzima*, Chu *tzima*.

tz'uunu'n* 'hummingbird': Yuc *tz'unú'un*, Itz *tz'unu'un*, Mop *tz'unu'un*, Hua *chunun*, Chl *tz'uñun*, Chn -*tz'unu'*, Ch'r *tz'unun*, Tzo *tz'unun*, Tze *tz'unun*, Tek *tz'uunu'm*, Ixh *tz'unu**n*, Awa *tz'uun**u**n*, Pqch *tz'un**u**n*, Kaq *tz'un**uu**n*, Tz'ut *tz'un**u**n*, K'ich *tz'un**u**n*, Q'eq *tz'unun*, Moch *tz'uun**u**n*, Q'an *tz'unun*, Aka *tz'unun*, Jak *tz'unun*, Toj *tz'unun*, Chu *tz'un**u**n*.

**umaa'* 'mute': Chl *uma'*, Chn -*uma'*, Tzo *uma'*, Tze *uma'*, Moch *umaa'*, Q'an *uma'*, Aka *uma'*.

xangab'* 'sandals': Yuc *xanab'*, Itz *xa**nab'*, Mop *x**a**na'*, Chl *xänäb'*, Ch'r *xanab'*, Tzo *xonob'*, Tze -*xanab'*, Tek *xajab'*, Mam *x**a**jab'*-, Ixh *xaab'*, Awa *x**a**b'*, Pqch *xijab'*, Kaq *xaj**aa**b'*, Tz'ut *xaj**ah**b'*, K'ich *xajab'*, Q'eq *xaab'*, Moch *xangab'*, Q'an *xanab'*-, Aka *xanab'*-, Jak, *xangab'*, Toj *xanab'*, Chu *xangab'*.

xuulub'* 'horn': Yuc *xu**lub'*, Mop *xulub'*, Chl *xulub'*, Tzo *xulub'*, Tze *xulub'*, Moch *xuulub'*, Toj *xulub'*.

**yahlang* 'below, under': Yuc *yáalan*, Itz *yalam*, Mop *yalam*, Hua *alam*, Tzo *olon*, Tze *ahlan*, Moch *aalang*, Q'an *yalan*, Aka *yalan*, Jak *yalang*, Chu *yalang*. [The Hua form is probably a loan from Yucatecan.]

References

Brown, Cecil H. and Søren Wichmann. 2004. Proto-Mayan syllable nuclei. *International Journal of American Linguistics* 70:128-186.

Fox, James Allan. 1978. Proto-Mayan accent and palatalization of velars. Unpublished Ph.D. dissertation, University of Chicago.