

# THE HISTORY OF THE WORD FOR CACAO IN ANCIENT MESOAMERICA

Terrence Kaufman<sup>a</sup> and John Justeson<sup>b</sup>

<sup>a</sup>Departments of Anthropology and Linguistics, 3302 WWPH, University of Pittsburgh, 230 South Bouquet Street, Pittsburgh, PA 15260, USA

<sup>b</sup>Department of Anthropology, University at Albany, 1400 Washington Avenue, Albany, New York 12222, USA

## Abstract

The word *\*kakaw(a)* ('cacao', *Theobroma cacao*) was widely diffused among Mesoamerican languages, and from there to much of lower Central America. This study provides evidence establishing beyond reasonable doubt that this word originated in the Mije-Sokean family; that it spread from the Mije-Sokean languages in or around the Olmec heartland into southeastern Mesoamerican languages; that its diffusion into Mayan languages took place between about 200 B.C. and A.D. 400; and that it spread from a Mije-Sokean language in or near the Basin of Mexico into languages in the region. It shows that each of the arguments presented by Dakin and Wichmann (2000) against a Mije-Sokean origin is either unworkable, is based upon false premises, or is not relevant; and that their proposed alternative – that it originated in and spread from Nawa into other Mesoamerican languages – conflicts with the mass of evidence relevant to the issue.

This study also discusses the linguistic details of vocabulary for drinks made from cacao; shows that no proposed etymology for the word chocolate is correct, but agrees with Dakin and Wichmann that its proximate source is a Nawa form *chikola:tl*; and discusses the history of words for *Theobroma bicolor* ('Nicaragua chocolate tree; *pataxte*') and their use.

The linguistic data are pertinent to issues of intergroup interaction in pre-Columbian Mesoamerica, but do not shed light on the nature or the cultural context of the diffusion of cacao in Mesoamerica, nor on its uses.

This study addresses a problem in linguistic reconstruction that is relevant to work on lexical diffusion in Mesoamerica, and thereby to work on intercultural interaction that probably dates to Preclassic times. It focuses on the origin and spread of the widely diffused form *\*kakawa*<sup>1</sup> (and variants) as a word for *Theobroma cacao* in Mesoamerican languages (Figure 1). Its purpose is to show that Campbell and Kaufman (1976) are right in claiming a Mije-Sokean origin for this word, and why, and that Dakin and Wichmann

(2000) cannot be right in claiming that it originated in Nawa. In addition, it addresses aspects of the histories of a few other terms in the same semantic field, mainly Nawa *chokol=a:tl* for the drink chocolate and a variety of terms for *Theobroma bicolor* ("Nicaragua chocolate tree" [Kelsey and Dayton 1942:621]; in local Spanish, *pataxte*), and the range of uses of these terms. It discusses hieroglyphic attestations that provide our earliest documentary evidence of the range of applications of the word *kakaw*.

E-mail correspondence to: tzajinkajaw@aol.com

<sup>1</sup> Unless otherwise noted, all linguistic forms cited in this study have been collected by members of the Project for the Documentation of the Languages of Meso-America (PDLMA) or verified by the authors, even when these forms are also cited in Dakin and Wichmann (2000). Any data cited from a source where no independent verification was made or could be made are credited to that source.

This paper makes use of the following conventions:

Words in Mesoamerican Indian languages are cited according to the orthographic practices of the PDLMA, which in turn derive from those of the Proyecto Lingüístico "Francisco Marroquín" (PLFM); a version of these principles is officialized as the orthography of indigenous languages of Guatemala. This spelling system uses only ASCII symbols, spells phonemically (one phoneme per grapheme [which may consist of a group of ASCII symbols]), and follows Spanish and traditional Mesoamerican orthographic practice when it is not inconsistent. This means that /q/ is <q>, /k/ is <k>, /w/ is <w>, /š/ is <x>, /c/ is <tz>, /č/ is <ch>, /' is <'>, /ŋ/ "eng" is <nh>, /x/ is <j>, /i/ "barred i" is <u>, vowel length is <V:> or <VV>, etc. Phonemic spellings found in sources that do not employ the PDLMA's orthographic practices are respelled; pre-modern citations are presented in faithful transcriptions of the original spellings. The only forms that are not respelled are those whose pronunciation is not unambiguously or adequately indicated by their symbols. Such forms are

cited within angled brackets (e.g., <abc>). Against current custom, Kaqchikel vowels are transcribed as long and short, instead of tense (plain) and lax (with superimposed dieresis). The long-versus-short contrast is mostly (but not entirely) limited to final stressed syllables. The transcription used here makes structural statements simpler and facilitates comparison with the closely related K'ichee7 and Tz'utujil.

Transcriptions of Epigraphic Mayan data use the same orthography. These words are cited between angled brackets <...>, with logograms presented as English words written in capital letters, and syllabograms presented in lower case. Some CVC syllabograms are postulated. Transcriptions of signs in the same glyph group are connected by hyphens. Phonologically interpreted Epigraphic Mayan forms are spelled as if they were Ch'olan forms after the relevant cases of \*e: and \*o: were raised to /i:/ and /u:/, respectively, but before vowel length was lost—thus, no cases of /ä/ appear. These Ch'olan forms are cited between slashes /.../.

Language names are spelled according to a single set of orthographic principles, the same used to spell ordinary vocabulary items. They agree with the officialized Guatemalan spellings except in respecting local differences among languages in the spelling of long/tense vowels and in distinguishing glottal stop from glottalization. We spell the name of the extinct language Ch'olti7 in this way on the basis of the modern pronunciations of the names of Ch'ol and Ch'orti7. This language name is spelled <Cholti> in the original sources, which do not mark glottal stops or glottalization.

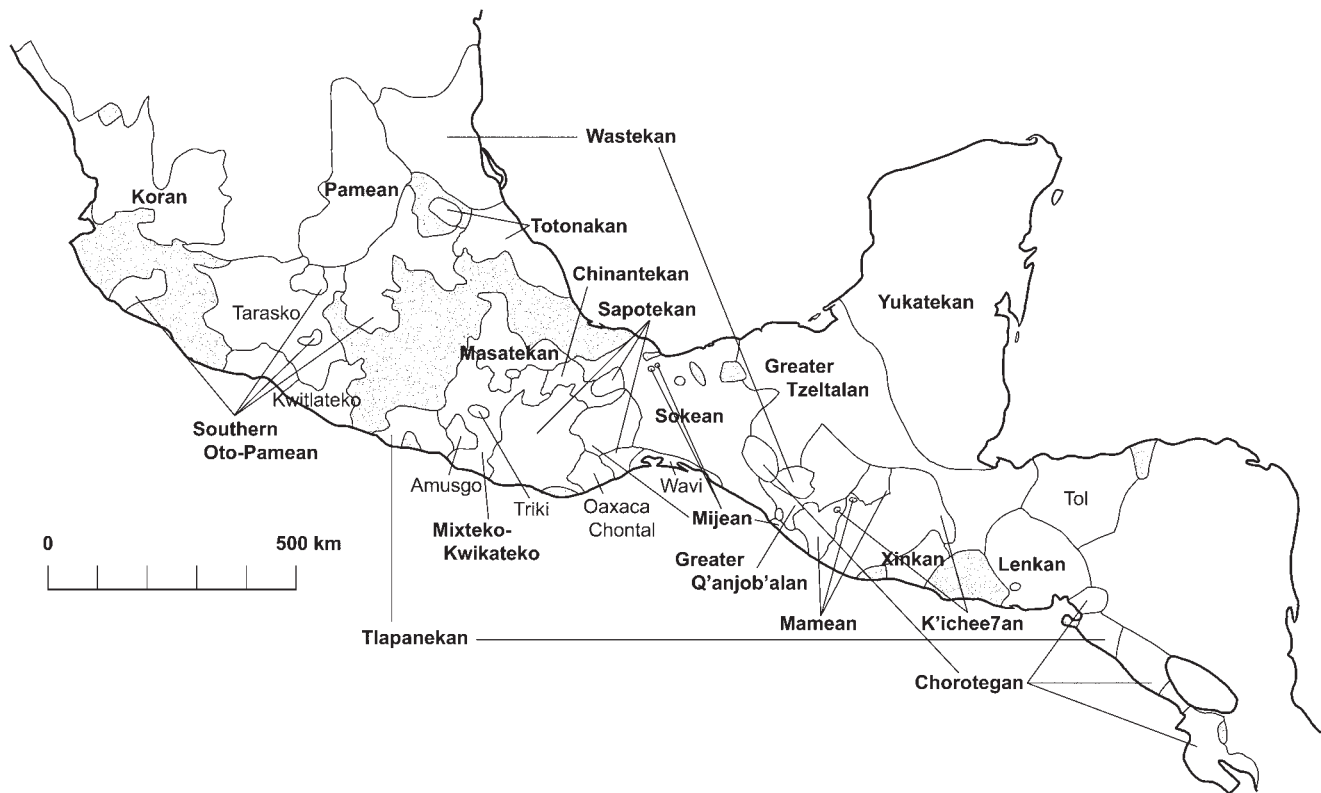


Figure 1. Languages of Mesoamerica, in their approximate locations as of A.D. 1500 (after Kaufman 1994). Except for isolates, individual Mesoamerican languages are not represented. Rather, the map groups these languages into the families or major subgroups of which they were members, which were individual languages between about 1200 and 600 B.C. (The locations of many were substantially different in that era from what is depicted here.) Areas of Nawa speech are shaded in gray. Individual languages (isolates) are specified in plain type; language families and subgroups are in bold.

Cacao (cocoa) was a major crop in pre-Columbian Mesoamerica (for a recent synthesis, see McNeil 2006). The kernel was ground and beaten with water, flavorings, and usually maize to make a drink, one version of which we know as chocolate. In historical times, among Mesoamerican Indians the pulp that surrounds the kernels inside the husk/pod has been and often is fermented to produce an alcoholic beverage. Aztecs, and arguably Teotihuacanos

and other pre-Columbian societies, made strong efforts to control the production and distribution of cacao. The kernel (cocoa bean) was at some point used as currency. In Xinka, for example, the word /tuwa/ means both 'cacao' and 'money'.

Cacao has long been grown in South America, lower Central America, and Mesoamerica. In Mesoamerica, archaeologically recovered remains of cacao have been dated to as early as 600 B.C.

It should be noted that these spellings of language names depart in many instances from those that are in standard use in *Ancient Mesoamerica*—and, indeed, from those that are most widely used in the field. We adopt these spellings because, as pointed out by B'alam Mateo-Toledo (2003:151), the orthographies chosen to represent language names "have political effects and an impact on issues of social and linguistic legitimacy in minority communities." The representation of the name is part of the representation of a linguistic identity. In Guatemala, representatives of indigenous communities have rejected colonial, Spanish-based spellings in favor of those designed by indigenous linguists after "a long period of work and struggle for language revitalization and recognition, self-definition, and definition of linguistic identity" (Mateo-Toledo 2003:152). A failure to use these spellings can be interpreted by indigenous people as reflecting political positions concerning their languages, their communities, and their human rights. There is no equivalent, national dialogue among indigenous communities in Mexico. However, these are definite issues in individual communities, and the Guatemalan model is having some impact, especially among Mayan languages of Mexico. Our writings convey our representations of indigenous people, to themselves, to their countrymen, to their governments, and to the world. We consider it more appropriate to refer to their languages using spellings that reflect viable orthographies for those languages than representations that were imposed by their conquerors.

In some cases, the name itself is different from one that is widely used in the literature. For example, the indigenous name Tol is used in preference to Jicaque, which in local Spanish means 'cannibal'. In the case of Sokean languages, the languages conventionally known as "Zoque" form a proper genetic subgroup of Sokean languages, so we reject Wichmann's extension of this term to members of the other proper subgroup, Gulf Sokean.

Most of the equivalences are obvious, but for completeness, we supply a full concordance of our usages alongside those that are either conventional or are used by Dakin and Wichmann (2000) (forms with suffixed -an are not listed separately): Amusgo (Amuzgo), Ayapa Gulf Sokean (Ayapa Zoque), Boruka (Boruca), Chiapaneko (Chiapanec), Chinantekan (Chinantecan), Eastern Mije (Lowland Mije [Guichicovi]), Honduras Lenka (Lenca-Guaxiguero), Kabékar (Cabecar), Kájita (Yaqui, Mayo), Kora (Cora), Mange (Mangue), Masawa (Mazahua), Mije (Mixe), Misteko (Mixtec), Nawa (Nahuatl), Salvador Lenka (Lenca-Chilanga), Sapoteko (Zapotec), Soke (Zoque), Sotepan Gulf Sokean (Sierra Popoluca, Sotepan Zoque), Tarasko (Tarascan), Taraumara (Tarahumara), Tepewa (Tepehua), Tlapaneko (Tlapanec), Tol (Jicaque), Totonako (Totonac), Warijiyo (Guarjijo), Wasteko (Huastec), Mobe (Guaymi), Western Mije (Highland Mije [Totontepec]), Wichol (Huichol), Xinka (Xinca), Yokot'an (Chontal Mayan), Yukateko (Yucatec, Maya), Yuta-Nawan (Uto-Aztec).

in Belize (Hurst et al. 2002); the earliest dates for the associated vessel types go back to 700–1000 B.C. in the Ulua Valley (Henderson and Joyce 2006:143). Cultivated and escaped *Theobroma cacao* is now widely distributed in lowland areas of Mesoamerica, and *Theobroma bicolor* grows uncultivated in some of these areas.

There is one widely attested term for ‘cacao’ whose distribution is largely the result of diffusion: Sokean \**kakawa*; Mijeán \**kakaw*; Nawa /*kakawa-tl*/; Masawa /*kakawa*/; general Mayan, Totonako, and perhaps Salvador Lenka /*kakaw*/; Paya [kaku]; and Tarasko /*khékua*/. Boruka, Tol, and Honduras Lenka have [kaw], and Mobe (also known as Waymí) has [ku]. These, we show, can be traced back to a form like /*kakaw*/, probably borrowed from Mayan. A proximate antecedent form pronounced something like [kVwa] is also reflected in Amusgo and possibly in Chinantekan. A variety of other terms is found in other languages, none of them widespread.

Throughout Mesoamerica, the peanut (which originates in South America) is named after ‘cacao’. Usually it is called ‘earth-cacao’, just as in British English it is called “ground-nut”. In northern Mesoamerica, where cacao is absent, ‘peanut’ can be called simply ‘cacao’ (e.g., Huasteca Nawa /*kakawa-tl*/, Totonako /*kakaw*/).

Nawa /*kakawa-tl*/ is the source of the Spanish *cacao*, which is pronounced /*kakáwa*/ (*cacahua*) in some regional types of Spanish. (Other pronunciations were current in Spanish in the early colonial period [Steinbrenner 2006:253].) The same Nawa form is also the source of Spanish *cacahuete* (Iberian/Peninsular Spanish *cacahuete*) ‘peanut’.

## DEMONSTRATING BORROWING

The basic topic of much of this paper is the determination of the reason for the similarity of form among words that have the same meaning. There are three possibilities: that the compared words, if part of the same language family, are native to that family, descending by normal transmission from an earlier, ancestral language into the descendant languages; that the word is diffused, with one or more languages that did not have a version of the word having adopted (“borrowed”) it from another language, whether or not in the same language family; or that the resemblances are simply due to chance and that the two words have no historical relationship.

There are standard methods for demonstrating either of the first two possibilities.

Demonstrating that a particular form is inherited within a language family requires strict adherence to the *comparative method*. The pronunciation of words changes in regular ways, and it must be demonstrated that the word in question conforms to all of the regularities otherwise known to characterize the language of which it is a part. In a small percentage of cases, to be sure, there are isolated departures from what is expected; in these cases, additional linguistic facts may help to show that the form is likely to have been inherited rather than borrowed, and these facts are likely to be peculiar to the languages or families of languages involved.

Demonstrating that a particular form has diffused *within* a language family depends on showing that the resemblances cannot have been inherited, because their similarities and differences in pronunciation do not conform to the regular sound changes that affected the individual members of the family and cannot be attributed to other known but less regular processes that operate in the individual languages or in ancestral forms of the language.

The presence of forms of closely similar meaning in different language families that cannot be shown to be closely related are generally borrowed, but demonstrating this depends on showing that there was a form of the word at some stage in the history of one of the languages that would descend normally to the attested forms of the word in the descendants of that language and that, on borrowing, it would descend normally to the attested form in the descendants of the language into which it was borrowed.

Words from one language when taken over into another typically undergo adjustments in pronunciation that result from a mismatch between the phonetics and phonologies of the two languages. In well-understood cases, these discrepancies usually follow what in retrospect can be regarded as predictable patterns, but in some cases, especially in early loans in a two-language contact situation, the borrowed forms are mangled. To establish a body of loans from one language into another, it is necessary to be able to demonstrate that there are a number of cases that conform in close detail in pronunciation and meaning. Once this has been accomplished, promising cases that depart in one way or another can be interpreted within the framework established by the clear cases.

The methods are best explained by example. Because of its relevance to the topic of this paper, we do so using examples of borrowings between Nawa and other languages and discuss alternative proposals for the direction of some of these borrowings.

The Nawa language group is the southernmost member of the Yuta-Nawan family. Yuta-Nawan has two branches: Northern and Southern Yuta-Nawan. Southern in turn consists of the Nawa group and the Sonoran languages. The Nawa group itself has two branches, Pochutec and General Nawa. Pochutec, an extinct language, is so scantily documented that the common ancestor of these two languages cannot be reconstructed in any detail. General Nawa consists of a large number of distinct forms, some of which are different enough to be considered different languages but most of which are different dialects of the same language. Because of the large amount of data available on many of these languages, by applying the comparative method, a substantial proto-Nawa vocabulary can be reconstructed, along with many of the structural properties of the language: its phonology (sound system), morphology (word structure), and syntax (phrase, clause, and sentence structure).

Detailed consideration of Yuta-Nawan data shows that Nawa entered Mesoamerica from the north, where other Yuta-Nawan languages are spoken. This is the consensus among historical linguists specializing in Nawa, notwithstanding the recent proposal by Hill (2001) that the family originated in central Mexico. An ancestor of proto-Nawa (pre-Nawa) was specifically influenced by Kora and Wichol, so its speakers must have spent some time in their vicinity.

Nawa differs from its closest relatives, the Sonoran subgroup of Yuta-Nawan languages, in large part by its Mesoamericanization through contact with languages of Mesoamerica. For example, Yuta-Nawan languages generally show object–verb order within the clause and many other features that are grammatically correlated with this, such as the placement of genitive phrases before the nouns they modify. In Nawa, these features have changed to verb–object order and correlated orders such as the placement of genitive phrases after the noun. A relic of the older Yuta-Nawan pattern remains in a genitive–noun order in some lexicalized (frozen) phrases, especially plant and animal names.

By applying the comparative method, a large number of lexical items can be reconstructed for proto-Nawa. Kaufman (1994–2004)

shows that a large number of them come from other Mesoamerican languages, chiefly from Mije-Sokean, Totonakan, and Wastekan (see Table 1). A number of these items are discussed later to illustrate generally applicable methods.

Several structural features of proto-Nawa can be shown to have come from these same languages.

Common to all forms of Nawa is a set of Mije-Sokean loans, at least one grammatical morpheme, and both morphological and syntactic patterns that result from Mije-Sokean grammatical influence.

In addition, structural patterns were adopted from Mije-Sokean that could not have been based on Wastekan or Totonakan.

Nawa developed morphologically complex verb words from an earlier Yuta-Nawan verb-phrase pattern involving clitics and auxiliary verbs. The only Mesoamerican languages that “could have supplied a clear and full model for the morphologization of the verb word in Nawa” are Mije-Sokean languages (see Kaufman 1994–2004 for details).

The Nawa system of locative relational nouns has close parallels with that of Mije-Sokean. They often occur with a generic locative suffix ( $\{+m\bar{u}7\}$  in Mije-Sokean,  $\{-k(o)\}$  in Nawa); when they govern pronoun “objects”, they are marked with possessive prefixes; they may be immediately postposed to a noun, forming a compound of which the relational noun is the head, and with the reading of a locative adpositional phrase. No other Mesoamerican language group has such structures.

The Nawa third-person possessive prefix  $\{i(:)-\}$  does not have a satisfactory Yuta-Nawan etymology, and proto-Mije-Sokean  $*7i+$  ‘third person ergative’ may be its source or may have influenced its development; proto-Mije-Sokean  $*7i+$  ‘third person ergative’ has also been borrowed by several Mayan languages.

Totonakan and Wastekan were also sources of structural features of proto-Nawa. For example, pre-Nawa developed a phoneme  $/t/$ . In Mesoamerica, this sound is found only in Totonakan and Nawa.

Several other structural features of proto-Nawa can be traced to these languages, although it cannot always be determined which of them was the source.

The adoption of so many Mesoamerican lexical items and structural features by the proto-Nawa stage shows that the breakup of Nawa into Eastern and Western branches took place in Mesoamerica—contrary to Dakin’s view (Dakin and Wichmann 2000:58) that proto-Nawa broke up far to the north, near Kora and Wichol.

## Discussion

Some of the forms presented in Table 1 are claimed by Dakin and Wichmann (2000) to have in fact been native to Nawa and to have been borrowed *into* the languages that we consider their sources. We discuss those cases in which their data are more or less accurate to help make the inference methods explicit. In no case is there a cogent case for the borrowing having been *from* Nawa; in most, there is a strong case that the borrowing was *into* Nawa.

The direction of borrowing of some words can be determined because the borrowed form shows sounds that could not have descended into the language from native resources. For example, there is a proto-Yuta-Nawan diminutive suffix  $*tsi$ , but this could not be the source of proto-Nawan  $*tzi(:)n$ , because  $*/tzi/$  regularly yields  $/chi/$  in Nawa, and  $\{-ch\}$  is in fact attested (frequently but not productively) as a diminutive suffix on Nawa nouns.

Contrastive phonology plays a major role in determining the direction of borrowing: the phonetics of the word in the source

language determines the way the word will be pronounced in the borrowing language. For example, glottalized consonants in Mayan languages are borrowed as unglottalized consonants in languages that lack phonetically glottalized consonants. Thus, a form such as Wasteko *net’etx*, with glottalized  $t'$ , is predictably borrowed into Nawa as *netech*. The reverse direction of borrowing is not feasible, because Mayans do not interpret Nawa plain consonants as glottalized; there is not a single plausible loan-word from Nawa into any Mayan language in which a Nawa plain consonant has been borrowed as a glottalized consonant. Similarly, the  $q$  of Totonako *sagat* would be borrowed into Nawa as  $k$ , yielding *saka-tl*; because Nawa did not have  $[q]$  as an allophone of  $/k/$ , Nawa *saka-tl* cannot have been the source of the Totonako word (contrary to Dakin and Wichmann 2000:68).

Wichmann (1998) and others make the mistake of supposing that (or operating as though) phonemic inventories rather than phonetic interpretation are a reliable basis for establishing directionality. This is illustrated by the example of proto-Mije-Sokean  $*tu(7)nuk$  and Nawa *to:tol-in*. The two can be compared after undoing the effects of Nawa reduplication, with Mije-Sokean  $n$  corresponding to Nawa  $l$ . Because Mije-Sokean languages have  $n$  but not  $l$  in native roots in ordinary vocabulary, Wichmann argues that the directionality of the borrowing is from the language that has  $/l/$  (Nawa) to the language that lacks  $/l/$  (Mije-Sokean). On the surface, this parallels the earlier discussion of glottalized versus plain consonants in Mayan versus Nawa. However, it is the *phonetic* realization of Mije-Sokean  $n$  that is at issue, and in many languages that lack phonemic  $/l/$ , phonemic  $/n/$  may have a range of pronunciations, including some that more closely approximate the  $[l]$  than the  $[n]$  of languages in which these are phonemically distinct. Since it is the phonetics of the source sound that is at issue, the  $n:l$  correspondence does not provide evidence that the borrowing was from Nawa rather than into Nawa. There is further evidence in other cases involving other languages that Mije-Sokean  $/n/$  was occasionally perceived as  $[l]$ .

Totonako *puuchuut* ‘silkcotton tree’ compares closely with proto-Nawa  $*po:cho:-tl$ . The vowel difference is not revealing, because both Totonako and Nawa have only one rounded vowel. Totonako  $u$  would be borrowed into Nawa as  $o$ , and Nawa  $o$  would be borrowed into Totonako as  $u$ . The linguistic distributions of the words are similar: one is found in Totonako but not in Tepewa and one in Nawa but not in other Yuta-Nawan languages. The internal diversity of Totonako and Nawa are comparable. There is little doubt that borrowing is involved, since the meaning is so narrow, and the agreement in meaning is precise.

To determine the direction of borrowing of this word, the most telling feature is the  $p$  that begins it. Any non-verb in Nawa with initial  $p$  is a strong candidate for a borrowing, because  $p$  regularly disappears in word-initial position, except in some words that habitually occur with prefixes so that the  $p$  is not usually initial. Dakin and Wichmann (2000:59b) attempt to support a Nawa origin by claiming that it is derived from a Nawa verb meaning ‘to card (cotton)’. The verb in question is *pochi:na* (for ‘to card [cotton]’; Dakin and Wichmann mistakenly give *po:chewa*, which in fact means ‘to get smoky’). There is no variation in vowel length of a Nawa morpheme when it occurs in different words, so *pochi-* is not plausibly an etymological source of *po:cho:-tl* within Nawa. Finally, this direction of borrowing makes sense. Nawas came from an area in which the silkcotton does not grow and only got a word for it on their arrival. Totonakos live in an area where silkcottons are found.

Table 1. Proto-Nawa lexical items borrowed from Mije-Sokean, Wasteko, and Totonakan<sup>a</sup>

	Source Form	Meaning	Nawa Borrowing
A. Mije-Sokean			
a	pMS *po7t	to grind (flour)	poto:-n-ik 'finely ground flour'
a	pMS *7a	canoe	a:=ka(:)l-li
a	pMS *kakawa	cacao	kakawa-tl
a	pMS *ko(-)pak	head	kopa:k-tli 'roof of mouth; also =kpa-k 'above'
a	pS *potz	to pile up	po:tzA
a	pS *naka	skin, leather	naka-tl 'flesh, meat'
a	pM *pus	to cut	pos=tekI (tekI by itself also means 'to cut'; te=pos-tli 'metal' (te-tl 'stone'))
b	pMS *kʉ7ak	footgear	kak-tli
b	pMS *tu(7)nuk	turkey	to:tol-in
b	pMS *(jaj)tzuku(7)	ant	tzi:ka-tl
b	pMS *sam	to heat	xami-tl Pochutec 'tortilla' general Nawa 'trivet, adobe brick'
b	pS *pata7	mat	petla-tl
b	pS *tʉ7pʉ	fish sp	to:poh-tli
b	pS *soki7	snail	xok-tli 'snail, shell'
c	pS *7une7	baby	kone:-tl 'child'
d?	pMS *na7aw	old man, husband	na:wal-li ~ na:wal-in 'shapeshifter' <sup>b</sup>
B. Wastekan			
a	Was kutxu7 < pM *quch 'hooked'	parrot	kocho(-tl)
a	Was 7ojox < pM *7ojx	breadnut	ohoxih-tli ~ ohox-tli
b	Was book < pre-Was *wo(:)k < pM *woq 'bubbling'	pulque	ok-tli
b	Was txotxob	hoof	chochol-li 'deer's hoof'
c	Was net'etx ~ nit'itx	striated, stacked	netech 'close together'
C. Totonakan <sup>c</sup>			
a	Tot túkay	spider	toka-tl
a	Tot ma:tzajtza	pineapple	ma:tzah-tli
a	Tot puuchuut	silk-cotton tree	po:cho:-tl
a	Tot -tziin	diminutive suffix	-tzi(:)n
a	Tot chichi7	dog	chichi
a	Tot saqat ~ seqet	grass	saka-tl
a/b	Tot xuunuk	cork tree	xo:no:-tl NPue Nawa xo:nok
b	Tot xuulh	jolote (fish)	xo:lo:-tl
b	Tot qaa7x	gourd bowl	kaxi-tl
b	Tot wajkat	crate	wahkal-li
b	Tot waapa	guapota, mojarra	#wapo-tl
b	Tot tiix	brother-in-law	te:x-tli
b	Tep lhpaw	avocado species	pa(:)wa(-tl)
b	Tep piipi7	man's elder sister	pih-tli <sup>d</sup>
b	Tot pi7si-y, pi7xi-y	to harvest (fruit)	
c	Tep p'us	to pick fruit	pixka 'to harvest, to pick (especially corn)'
c	Tot pa7ks-	to get well, to become cured, to form a scar	pah-tli 'medicine', pak-ti=nemI 'to be strong/healthy/happy'
c?	Tot xkuta7	sour	xoko-ke <sup>e</sup>

Notes: a = Forms that are effectively identical (given that Nawa has no contrast of [o] and [u]); b = forms with segmental differences that can be explained by the contrastive phonologies of the borrowing and source languages; c = likely borrowings with unsystematic phonological discrepancies or substantial differences in meaning; d = possible borrowings with unsystematic phonological discrepancies or substantial differences in meaning; pMS = proto-Mije-Sokean; pS = proto-Sokean; pM = proto-Mije-Sokean.

<sup>a</sup>A morphologically complex lexical item, proto-Mije-Sokean \**ʉk.7ʉy* 'to enter' (< \**ʉk* 'house'), is the model for Nawa *kal=aki* 'to enter' <*kal* 'house' + *aki* 'to be able to be inserted, to fit.'

<sup>b</sup>The semantics of the comparison between proto-Mije-Sokean \**na7aw* 'old man, husband' and Nawa *na:wal-* 'shapeshifter' is weak.

<sup>c</sup>Only four of the Totonako items in question are also found in Tepewa, suggesting that the diffusion from Totonako to Nawa occurred after the split between Totonako and Tepewa. This split took place probably no later than A.D. 300 and probably no earlier than 400 B.C.

<sup>d</sup>Nawa *pih-tli* 'man's older sister' has a proposed Yuta-Nawan etymology, but the putative source means 'younger sister.'

<sup>e</sup>The phonological overlap of Totonakan *xkuta7* and Nawa *xoko-* is not overwhelming, but the Nawa word cannot descend from proto-Yuta-Nawan or proto-Southern Yuta-Nawan.

One feature of this word that might suggest borrowing from Nawa is the seeming correspondence of final *t* in Totonako to final *tl* in Nawa. Totonako does not allow *tl* at the end of words, so had it been borrowed from Nawa, the final *t* of *puuchuut* is expected. Our account of the borrowing, from Totonako to Nawa, requires that the final *t* must have been reanalyzed in Nawa. This sort of reanalysis is not uncommon in language-contact situations.

The Nawa borrowing of Totonaka *puuchuut* illustrates a general regularity that words for local plants, when not expressed by neologisms created out of native resources, are adopted by newcomers to an area from people already living there unless they have previously wiped out the indigenous inhabitants. Dakin and Wichmann cite their misanalyzed example of *po:cho:-tl* as part of a rationale for their “perception that Nahuatl has received very few loans from other languages but has resorted to resources of the language to produce new descriptive terms.” This is highly unusual as a pervasive pattern in the case of names for new plants, and it is especially unusual that indigenous people would systematically adopt newcomers’ invented names for indigenous plants. Methodologically, it is not cogent to build systematically on an “impression” that is so much at variance with established trends. Such a claim requires compelling evidence of several unproblematic and unambiguous borrowings of native Nawa words, derived from Yuta-Nawan sources, for plants that they encountered in Mesoamerica, to provide support for such a borrowing in any ambiguous instance.

Several of the plant names listed in Table 1 show clear evidence that they are not native to Nawa—for example, the initial *p* of *pa(:) wa(-tl)* ‘avocado species’ is suggestive of borrowing. The Wasteko term for ‘breadnut’ shows a Wasteko innovation, the shift of proto-Mayan \**7ojx* to Wasteko *7ojox*; this shifted form is the basis for the Nawa loan. The word itself is no longer widely distributed within Nawa. However, it survives in place names from several parts of Mexico—for example, Ojitipa = *ohoxih-ti=pan* in San Luís Potosí and Ojitlán = *ohoxih-tla:n* in Oaxaca; it is also reflected by the word *ujushte* ‘breadnut’ in the local Spanish of southeastern Chiapas and Guatemala—a borrowing from Nawa *ohox-tli*, a variant of *ohoxih-tli*. The broad distribution of the word in Nawa shows that it was borrowed by Nawa, probably from Wastekan, at an early stage in its history. The borrowing between Wastekan and Nawa could not have gone the other way. The Mayan form in languages with a contrast between *h* and *j* have *j* in this word; had the word been borrowed from Nawa into Mayan, these languages would show *h* rather than *j*. (For conformity with Spanish orthography, in Mayan languages without a contrast between *h* and *j*, the sound is spelled as *j* even when, as in Wasteko *7ojox*, it is pronounced as [h].) In addition, some Mayan languages develop two syllables, with a repeated vowel, in words whose original shape was \**CVjC*, while original \**CVjVC* does not reduce to a single syllable. In summary, proto-Nawa *ohox-tli* would have been borrowed into Mayan languages as *7ohox*, with two syllables and with *h* rather than *j*, a contrast not found in Nawa. In terms of the antiquity of the forms, it is widely distributed within Mayan, a very diverse language family. In contrast, *ohox(ih)-tli* is found only in Nawa, with no other Yuta-Nawan attestation. Nawa is a weakly differentiated language group.

Proto-Nawa *wahkal-li* ‘crate; gourd bowl’, which Kaufman (1994–2004) cites as a borrowing from Totonako *wajkat*, is cited by Dakin and Wichmann (2000:69a) as, instead, a borrowing from Nawa into Totonako. Their evidence is the incorrect claim that it can be analyzed etymologically as consisting of a putative

Nawa root *wah* meaning ‘plank’ + *kal*, glossed as ‘box’ (really ‘house’ and, by extension, perhaps ‘container’); but ‘plank’ is not *wah* but *wapal-li*.

The Nawa word *xikal-li* ‘gourd dipper’ has been borrowed into several individual Mesoamerican languages (though not into any ancient ancestral languages) and into Spanish. Contrary to Dakin and Wichmann (2000:69a), it is not a native Nawa term but a borrowing from a Sapotekan language. For proto-Sapotekan, a root \**eka7* can be reconstructed. This root occurs with one of two different classifiers: with \**xi* prefixed, as \**xika7*, it refers to a gourd dipper or cup/bowl; with \**k* prefixed, as \**keka7*, it refers to a bottle-gourd. This shows that \**xika7* ‘gourd dipper’ is a native Sapotekan word for a gourd dipper. This form cannot have arisen by reanalysis of an existing Nawa word *xikal-li*, since when it appears with the classifier \**k*, the underlying vowel surfaces as *e* rather than *i*. The word is ancient in Sapotekan, which is far more diverse than Nawa, and is found in both branches of the family. Loan-word data correlated with archaeological data, presented later, show that Chatino and Sapoteko separated no later than about 200 B.C.

One anomaly in this borrowing is that, in *xikal-li*, Nawa has innovated a final *l* where perhaps final *h* (“saltillo”), corresponding to Sapotekan \**7*, might be expected, yielding *xikal-li* instead of *xikah-tli*. Possibly, *l* was substituted for *h* because most Nawa nouns with a CVCVC shape end in *l*, *x*, *tz*, or *ch*. Such a consideration may also account for the stem-final *l* of Nawa *wahkal-li* borrowed from Totonako *wajkat*. No more than two nouns (*nehmat* ‘right hand’, *i:lamat* ‘old woman’) end in *t* in Nawa, and in both cases some forms of Nawa drop the stem-final *t* or change it to *h*.

A more complex example is Totonako *saqat* ~ *seqet* ‘grass’, which compares closely with proto-Nawa \**saka-tl* ‘grass’. If this is a borrowing rather than a chance resemblance, then the relevant form to compare is *saqat*, since the *seqet* variant disagrees with \**saka-tl* in the vowel quality. There are two differences in the pronunciations involved. Totonako distinguishes between \**k* and \**q* in words with no mid vowels, so Nawa *saka-tl* would have been borrowed into Totonako as *sakat*, not *saqat*. This excludes Nawa as a possible source for the Totonako word. The only viable external source of the *q* in this word is Mijejan \**sokot*, since Mije-Sokean \**k* is borrowed into Totonako as *q* in words with *e* or *o*. However, apart from this item, the vowel variation found in *saqat* ~ *seqet*, is found only in native Totonako words. This variation is therefore evidence that the word is in fact native to Totonakan and suggests that the similarity with Mijejan is due to chance rather than to borrowing, which is also suggested by the unexplained discrepancy in vowels. In contrast, there is no evidence for any particular antiquity of the form in Nawa, as there is no convincing related form in any other Yuta-Nawan language. Finally, the final *t* of the Totonako form was reshaped to *-tl* on borrowing, as in the case of *po:cho:-tl*.

All of these borrowings occurred by the proto-Nawa stage, since they are widespread in the Nawa group and can be reconstructed to proto-Nawa, and most presumably arose earlier, since proto-Nawa is the earliest we can go on purely distributional grounds. However, some are demonstrably earlier than proto-Nawa, because they underwent changes that took place in pre-Nawa.

Sonoran and Yuta-Nawan \**u* shifted to *ɨ* in Kora, Wichol, and pre-Nawa. The vowel *ɨ* is a high back unrounded vowel, structurally but not phonetically equivalent to *u*. By proto-Nawa times, this \**ɨ* had shifted to *i*, except in a few words in Central Nawa where it shifted to *e*. An example is the word for ‘ant’. Two variants can be reconstructed for Mije-Sokean, proto-Sokean \**jaj=tzuku(7)* and proto-Mijejan \**tzuku(n)*. The first element, *jaj*, in the Sokean

compound recalls Mayan (proto-Mayan *\*ha7h*) words meaning ‘fly’ or ‘grub’, so this is probably its relevance here. The element *\*tzuku(7) ~ \*tzukun* goes back to proto-Mije-Sokean. Proto-Nawa *\*tzi:ka-* ‘ant’ must have arisen from a pre-Nawa *\*tzi:ka* or earlier *\*tzu:ka*, since proto-Yuta-Nawan *\*tsi* shifts to *chi* in Nawa. We have no evidence concerning the difference in the final vowels of pre-Nawa *\*tzi:ka* and Mije-Sokean *\*tzuku(C)*. The Mije-Sokean term was perhaps more plausibly *\*tzuku* or *\*tzuku7* than the Mije variant *\*tzukun*.

Similarly, proto-Southern Yuta-Nawan *\*t* became *tl* before *\*a*; afterward, some instances of Southern Yuta-Nawan short *\*a* became *e*. Both of these changes occurred before the proto-Nawa stage. These changes occurred after Sokean *\*pata7* ‘mat’ was borrowed by Nawa, since this word yields proto-Nawa *\*petla-tl*. These pre-Nawa loans from Mije-Sokean suggest that the influence of the Mije-Sokean elite language of the Basin of Mexico spread as far north as, say, Zacatecas by, say, the year A.D. 200. Another possible scenario might have Nawas arriving in the Basin of Mexico by circa A.D. 200 to have direct contact with Mije-Sokean elites in and around Teotihuacan (see “Culture-Historical Inferences”), but not yet to constitute an important presence in this new dwelling place.

### THE MIJE-SOKEAN HYPOTHESIS

Campbell and Kaufman (1976: 84) traced the word *\*kakaw(a)* back to proto-Mije-Sokean and the Olmec diffusion sphere. Justeson, Norman, Campbell and Kaufman (1985:23, 57–59) agree that this word originated in the Mije-Sokean family and argue that it spread from there to other Mesoamerican languages. In the case of Lowland Mayan languages (Ch’olan and Yukatekan), they suggest that the word spread in association with cultivated cacao or its products and entered Lowland Mayan languages during the Late Preclassic period. The basis for this scenario was as follows.

Using published materials available between 1959 and 1963, Kaufman (1963) reconstructed a vocabulary of about 600 proto-Mije, proto-Sokean, and proto-Mije-Sokean words and affixes. Among them was the reconstruction of *\*kakawa* as the proto-Mije-Sokean word for cacao, along with words for a large number of other lowland cultigens. Work by Kaufman and Campbell in the 1960s and 1970s showed that many of these Mije-Sokean words appeared in other Mesoamerican language families, in which they were not reconstructible to the earliest stages. In fact, Mije-Sokean vocabulary proved to be found in every language family in Mesoamerica, from Tarasko in the north to Xinkan in the south, and much of this influence was early enough that the terms are reconstructible to early stages in the histories of most of those families. *\*kakawa* was among these widely diffused Mije-Sokean words.

The diffused words are found in a variety of semantic domains. They include words for plants, animals, tools, food preparation, the calendar and numerical calculation, and kinship and other social roles. The most numerous of these borrowings are in names for plants and animals, especially of domesticated lowland plants and animals and of those wild plants and animals that are the names of days in the Mesoamerica. 260-day ritual calendar.

It should be noted that in assessing this body of evidence we exclude *Wanderwörter*, words that are found in similar form throughout Mesoamerican and whose ultimate source is unknown.

No other language families in pre-Columbian Mesoamerica had the impact that Mije-Sokean had, either in the range of languages they affected or in the number of items that were borrowed from them. The next biggest impact came from Nawa, but the number of

pre-Columbian Nawa loans does not approach the number of Mije-Sokean loans. Loans from Nawa began to be widely adopted late in pre-Columbian Mesoamerican history and date at least 1,000 years later than the Mije-Sokean loans. Nawa lexical material diffused into Mesoamerican languages even after the arrival of the Spanish as a consequence of their (varying) language policies. After massive borrowing by Spanish of Nawa vocabulary, some Nawa-origin lexical material has entered Mesoamerican languages from Spanish.

Some of the Mije-Sokean loans probably go back to the influence of Olmecs, while others are attributable to a post-Olmec era of Mije-Sokean influence. These results were presented by Campbell and Kaufman (1976) and, in the case of loans into Lowland Mayan languages, by Justeson et al. (1985). The Mije-Sokean loans into northern Mesoamerica are discussed by Kaufman (2000–2007), and the evidence is summarized by Kaufman and Justeson (2007). (Readers should note that the language labels [proto-Mije, proto-Sokean, proto-Mije-Sokean] associated with the various reconstructed forms were not correctly copied into the published version from the manuscript form of Campbell and Kaufman’s paper and should not be relied on for historical inference; reference should be made to Kaufman [1963] or, now, Wichmann [1995] for the correct historical level of reconstruction.)

Regarding *\*kakaw(a)* in particular, Justeson et al. (1985) show that this word was diffused in the post-Olmec era. In our own recent work (Kaufman and Justeson 2007) we have been able to narrow the period of this diffusion, in the case of Mayan, to between 200 B.C. and A.D. 400. It must have entered a Lowland Mayan language before A.D. 400, because at that time it is attested in Mayan hieroglyphic texts (Stuart 1988). The evidence that it diffused after 200 B.C. is more elaborate.

It must have diffused after the shift of Western Mayan *\*k(ʔ)* to proto-Greater Tzeltalan *\*ch(ʔ)* since it would otherwise have shown up as *chächäw\** in proto-Ch’olan and proto-Tzeltalan. By comparing several independent lines of evidence for the timing of the shift of *\*k(ʔ)* to *ch(ʔ)*, the date of this change can be shown to have occurred right around 200 B.C.

On epigraphic grounds, the shift of *\*k > ch* must have occurred before A.D. 200. One word that underwent the shift was Greater Tzeltalan *\*chij* < proto-Mayan *\*kehj* ‘deer’, and a word for the day Deer is spelled by the sign for the syllable /chi/ in the Late Preclassic Uaxactun murals. Other epigraphic evidence, perhaps less definitive, suggests that it took place before 100 B.C. Mora-Marín (2001:276, n. 180) identifies a verb form whose spelling is followed by the sign for /chi/ on the Late Preclassic text of the Dumbarton Oaks pectoral, which on stylistic grounds he dates between 300 and 100 B.C. The only etymologically viable analysis known to us is Kaufman’s 2001 suggestion that it functions as the reflex of the Greater Tzeltalan clitic *\*+ich* ‘already’ (< proto-Mayan *\*+ik*), which survives in Ch’ol as *äch* and which is attested epigraphically in the spellings <jj-chi> (mostly after words ending in *j*) and <yi-chi> (mostly after words ending in *y*).

The change also preceded the diversification of Greater Tzeltalan into Ch’olan and Tzeltalan branches. Greater Tzeltalan most likely broke up when Mayans moved into the highlands of Chiapas, which took place between about 200 and 100 B.C. (Clark 2000:54), and in any case, no later than this time. The glottochronological estimate for the diversification of Greater Tzeltalan is A.D. 100 or earlier. (See Mora-Marín [2001:46–50] for a more complete discussion of the relevance of recent archaeological dates for the entry of Mayans into this part of Chiapas to the timing of the diversification of Greater Tzeltalan.)

The change followed the borrowing of *#manik'* into Mayan as the name of the day Deer of the ritual calendar. This name shows up in colonial Ch'olan baptismal records as <manich> (Campbell 1988; Fox and Justeson 1982), so the word had to have been adopted in Greater Tzeltalan before the shift of *\*k'* to *ch'*. This brings more constraints to bear on the timing of the sound change, because this word is a borrowing from proto-Sapoteko *\*mma=ni7* 'animal, large quadruped'. This word itself has a foreign origin, because Sapoteko does not have *m* in native words. The morpheme *\*mma* was a borrowing of Sokean *\*mu7a* 'deer'. It did not displace the native Sapoteko word *\*kwe+tzina7* for deer' but must have maintained some kind of association with 'deer' to have been borrowed later for the corresponding day name by Lowland Mayans. Proto-Sapotekan had a word of the approximate shape *#nani* meaning 'animal', which in Oto-Mangean terms can be analyzed as a pre-Sapotekan nominalization (in *#na-*) of a root *#ni* 'alive'—thus, 'living thing'. The word *#nani* is reconstructible based on Zenzontepec Chatino *nya7nè*, Tataltepec Chatino *na7ni*, Yaitepec and Panixtlahuaca Chatino *7ni*, and Lachixío Sapoteko *náni*.

We suppose that some speakers of the ancestor of Sapoteko created *\*mma=ni7* by conflating the Sokean borrowing *\*mma* with the pre-existing Sapotekan form *#nani* 'animal'. The word *#nani* survived into Western Sapoteko and was replaced by *\*mma=ni7* 'animal' elsewhere. That *#ni* 'alive' may have still had some kind of independent existence is suggested by the fact that *ni+* is a preposed classifier for animals in some forms of Sapoteko.

The development of Greater Tzeltalan *#manich'* therefore involves five successive developments: the borrowing of *\*mu7a* by Sapotekos from Sokean; the addition of *=ni* within pre-Sapoteko to form *\*mma=ni7*; the borrowing of this word as *#manik'* into Greater Tzeltalan or Yukatekan; the diffusion of this word between the two of them; and the change of *\*k'* to *ch'* in Greater Tzeltalan. It is implausible that this series of changes, involving four different languages, would have taken place in less than a century. Accordingly, since the shift of *\*k'* to *ch'* took place before the diversification of Greater Tzeltalan, which in turn took place before 100 B.C., the first stage in the process took place before about 200 B.C.

The borrowing of *\*mu7a* occurred when speakers of Sokean were interacting with speakers of Sapoteko. In fact, this is but one reflection of interaction between Mije-Sokeans and Sapotekos. Sapoteko acquired several Sokean vocabulary items—notably, words for four out of ten animals that are names of days in the Mesoamerican ritual calendar: iguana/lizard, deer, dog, and macaw. In Sapoteko, they are simply the names of animals (the Sokean word for deer became the Sapoteko word for 'large animal' generally). Sapoteko also underwent some phonological changes under Sokean influence: a shift of accent from the last to the next-to-last (penult) syllable; a loss of nasality on vowels; and a change of *\*kw* to *p*. (This last change began in Sapoteko after it had started to diversify, because it did not apply to medial *\*kw* in Western Sapoteko.) These features illustrate a process that occurred repeatedly in Mesoamerica, which Kaufman calls "yokel anxiety": groups with aspirations to cultural prominence avoid or alter features of their pronunciation that are uncharacteristic of the speech of "civilized" groups. At least one grammatical morpheme was also borrowed, which could have come from either Mijejan or Sokean.

With one exception, these items are not found in Chatino and so postdate the breakup of proto-Sapotekan, and they are

reconstructible to proto-Sapoteko, so they date to the pre-proto-Sapoteko era of Sapoteko speech. Since the borrowing of *\*mu7a* into Sapoteko happened before the diversification of Greater Tzeltalan, the breakup of Sapotekan into Chatino and Sapoteko must have occurred before 200 B.C. Given this, the most plausible population process that might have been associated with this division is a century-long process of cultural consolidation, culminating with the establishment of state control at Monte Alban starting in Monte Alban I (500–200 B.C.), which began between 500 B.C. and 400 B.C. This process entails the creation of cultural boundaries at the limit of the unified territory. Glottochronology is consistent with this date, putting the diversification of Sapotekan at around 500 B.C. The exceptional case, proto-Sapoteko *\*kwe.7wa* and Zenzontepec Chatino *kō7mā* 'macaw', from proto-Mije-Sokean *\*7owa* 'macaw', may be an instance of diffusion from Sapoteko to Chatino, because all other forms of Chatino have a different term for 'macaw'. Less likely is the possibility that proto-Mije-Sokean *\*7owa* was diffused into proto-Sapotekan before the split between proto-Sapoteko and proto-Chatino.

There are several lines of evidence for the timing of Epi-Olmec contact with Sapoteko. The clearest evidence for contact begins during Monte Alban II, which dates from about 200 B.C. to A.D. 250. Hostile interaction between Epi-Olmecs and the Valley of Oaxaca is registered in the "conquest slabs" of Mound J at Monte Alban, which date to Monte Alban II. Caso (1947:23, 27–28) proposed that inverted heads below the "hill" logogram depict conquered persons or peoples, an interpretation that is now generally accepted. He also recognized that several of the inverted heads had Olmec-style facial features and headgear and suggested that they represented people from Chiapas. In further support of this association, what we identify as the personal name of a captive on Monte Alban Tablet 41 is spelled using an Epi-Olmec glyph. Our analysis of the most explicit dates in the Mound J tablets—those with a named year, ritual calendar date, and day of the lunation—show that the recorded captive-taking spans at least 111 years (Justeson and Kaufman 1996–2001; cf. Kaufman and Justeson 2004). The earliest of these events therefore dates no later than a century or so before the end of Monte Alban II and potentially to a century or so before it began—between about 300 B.C. and A.D. 150. The linguistic data, however, suggest that some of the interaction between Epi-Olmecs and Sapotekos was not hostile. This is consistent with data on the Sapoteko influence at Epi-Olmec sites. Archaeologically, Monte Alban II pottery was found by Drucker and Stirling at Cerro de las Mesas (Drucker 1942:84), and grayware from the Valley of Oaxaca begins to appear around 200 B.C. in western Chiapas (Evans 2004:223–224).

All of this evidence together indicates that the borrowing of Sokean *\*mu7a* 'deer' did not occur much before 300 B.C., the earliest plausible dating of the evidence for contact between Epi-Olmecs and Sapotekos. With a century or more separating the adoption of *\*mu7a* from the shift of *#manik'* to *#manich'* in pre-Greater Tzeltalan, the process would have ended between about 200 B.C. and A.D. 150. So the Greater Tzeltalan shift of *\*k'* to *ch'*, the last step in the process, probably took place no earlier than about 200 B.C. Greater Tzeltalan diversified after the *\*k'* to *ch'* shift. Since it is during the second century B.C. that Greater Tzeltalans began immigrating into the Chiapas highlands and the Grijalva Depression, where they became Tzeltalans, the best chronological estimate is that the *\*k'* to *ch'* shift took place right around 200 B.C.

It should be noted that other, less stringent constraints agree with these results. For example, Greater Tzeltalans and Sapotekans, who



were not in direct contact, were not in a position to have significant influence on one another until the development of significant state power in the Valley of Oaxaca after about 500 B.C., so the Mayan adoption of *#manik'* can be placed after that date.

Since the word *\*kakaw(a)* was adopted by speakers of one or more of the Greater Tzeltalan languages after the *\*k(ʔ)* to *ch(ʔ)* shift took place, this word was not borrowed until sometime after 200 B.C.

The evidence from lexical diffusion among Greater Tzeltalan, Yukatekan, Epi-Olmec, and Sapoteko, together with the epigraphic data from Ch'olan texts, therefore show that the word *\*kakaw* entered Greater Tzeltalan languages between 200 B.C. and A.D. 400.

The Lowland Mayan use of cacao as a beverage is attested archaeologically before the borrowing of the word *kakaw*, going back to at least 600 B.C. (Hurst et al. 2002). A native word for cacao, *\*pe:q*, existed in Mayan languages, including Greater Tzeltalan, before it was largely displaced by the diffused *kakaw*. The word *\*pe:q* survives in K'ichee7an with the meaning 'uncultivated cacao' (*Theobroma bicolor*), suggesting that the spread of the word *kakaw* may have been associated with the spread of cacao cultivation, of new practices associated with its cultivation or use (cf. Steinbrenner 2006:264–268), or of its rising economic and/or ritual importance. (Justeson et al. [1985:59] cite Miksicek [1983, personal communication 1983] as arguing from data on Pulltrouser Swamp that, in the Maya Lowlands, cultivated cacao first shows up in the Late Preclassic period. However, this is not reflected in the full publication [Turner and Harrison 1983], and means have not yet been found to distinguish pataxte, which is a domesticated cacao that grows untended in the wild, from cultivated cacao.)

Mayan *\*pe:q* evidently survived as *\*pe:k* in proto-Greater Tzeltalan, since it later shifted in Ch'olan to the pronunciation *\*pi:k* and to the meaning '8,000'. The change in pronunciation is due to a sound change that affected a common ancestor of the Ch'olan languages after they separated from Tzeltalan, and the semantic association derives from the use of gunny sacks to store large numbers of cacao beans, a practice documented

iconographically by the Late Classic period (see Stuart 2006: 190–191). Words for a gunny sack (of cacao beans) are used for '8,000' in several Mesoamerican languages. Besides Lowland Mayan *\*pi:k*, examples are Sokean *tzunu7* 'sack, bag, pocket, cap, 8,000' and Nawa *xikipi:l-li* 'gunny sack, 8,000'. In Sapoteko, the word *\*(kwe+)(s)su:7ti* 'bag, 8,000' comes to mean 'skirt', a semantic extension also attested in proto-Yukatekan *\*pi:k* 'skirt'.

In the Maya Lowlands, evidently, the meaning of this ancient word did not shift to include or become restricted to 'pataxte' with the adoption of the word *kakaw*. Instead, a contrasting Ch'olan term for pataxte, proto-Ch'olan *\*b'ahläm=te7*, appears to have developed—most likely after the adoption of *kakaw*, for reasons presented later.

At least four modern Sokean languages, belonging both to Soke proper and to Gulf Sokean, have words for at least two types of cacao, and there are at least two other plants that have *\*kakawa* in their names (see Table 2).

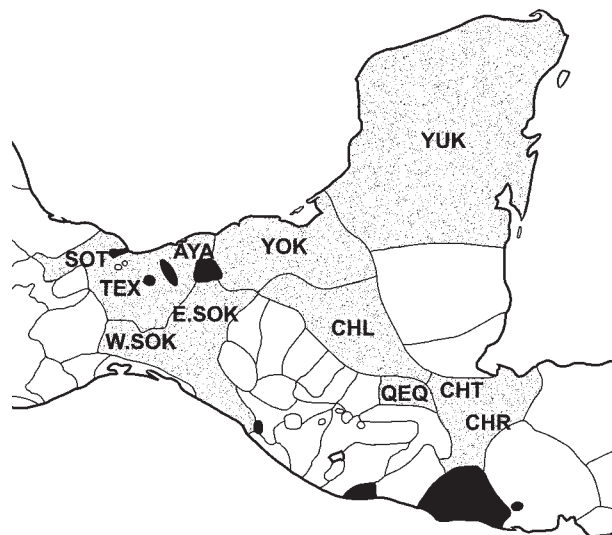
At least four Sokean languages have a word for pataxte. In three of them—colonial and modern Tecpatan Soke, Ayapa Gulf Sokean, and Soteapan Gulf Sokean—the word for pataxte contains the word 'jaguar' as a modifier of a word for a type of plant. This same practice is found in Lowland Mayan languages and in Q'eqchi7:

Ch'ol	b'ajrum=tye7	pataxte	Roberto Zavala, personal communication 2006
Ch'ol	b'ajlum=tye7	pataxte	Montejo et al. 2001
Yokot'an	baläm=te7	pataste	Keller and Luciano G. 1997; Montgomery-Anderson 2003
Ch'olti7	< balante > ~ < bahlamte >	"patate"	Morán ca. 1695
Yukateko	< balamte >	pataxte	Book of Chilam Balam of Chumayel (Roys 1933)
Q'eqchi7	b'a:lam kakaw~ kakaw b'a:lam	pataxte	Haeserijn 1979:53, 79; Sedat 1940:246

Table 2. Words for pataxte, for varieties of cacao, and for other plants incorporating the word for cacao in Mije-Sokean languages

A. "Generic cacao" ( <i>Theobroma cacao</i> )			
Santa María Chimalapa Soke	kakawa	cacao	
Tecpatan Soke	kakawa7	cacao	
colonial Tecpatan Soke	<Cacava>	[gloss missing]	
Soteapan Gulf Sokean	kaakwa	cacao	
Ayapa Gulf Sokean	ka:gwa~kak	cacao	
B. "JAGUAR CACAO" and analogous expressions = <i>Theobroma bicolor</i>			
colonial Tecpatan Soke	< Cangba >	patastle (cf. <Paa> 'Yerva')	
colonial Tecpatan Soke	< Tzutupue cangba > /tzujtzu + pʉ7 kanh=pa7/ < cangbu > /kanh puj/	patastle verde  cacao silvestre	
colonial Tecpatan Soke 1672		pataxte (como el cacao), coyol de tigre, cacao blanco (cf. <i>puj</i> 'seed')	
Tecpatan Soke	kanh=puj	tepecacacahu cf. <i>pujki</i> 'tepecacahu; papachote' [tree names]; cf. <i>nunta</i> 'real'	
Soteapan Gulf Sokean	kanh.kanh <i>pujki syn. nunta</i> <i>pujki</i>	pataste // large, pungent fruit (a relative of cacao)	
Ayapa Gulf Sokean	kanh=kak	pataste: tipo de cacao de fruta larga; cacao blanco; cacao natural del pueblo (cf. <i>na7-ti</i> 'real')	
Santa María Chimalapa Soke	tapunh kakawa <i>syn. na7-ti</i> kakawa	(lit., "white cacao") clase de árbol con flores como la de cacao; la corteza se usa para hacer el mecapan, y el mecapan	
Sayula Mije(an)	po7p kagaw		

Notes: *Theobroma bicolor*, *pataxte* or *cacao blanco* in Spanish, is usually labeled by a plant name with the modifier "jaguar." Soteapan Gulf Sokean *kanh.kanh* is apparently a reduplicated form of *kanh* 'jaguar.' Sayula Mije(an) *po7p kagaw* reflects the label "white"; the Santa María Chimalapa Soke term *tapunh* does not occur in any other contexts.



**Figure 2.** Mesoamerican languages with terms for *Theobroma bicolor* containing a word for jaguar. SOT = Soteapan Gulf Sokean; TEX = Texistepec Gulf Sokean; AYA = Ayapa Gulf Sokean; W. SOK = Western [Chimalapa] Soke; E. SOK = Eastern [Chiapas] Soke; YOK = Yokot'an [Chontal Mayan]; CHL = Ch'ol; CHT = Ch'olti7; CHR = Ch'orti7; QEQ = Q'eqchi7; YUK = Yukateko. Areas of Nawa speech are in black.

The spelling of the Ch'olti7 gloss is presumably a mistake for *pataste* or *pataxte*.

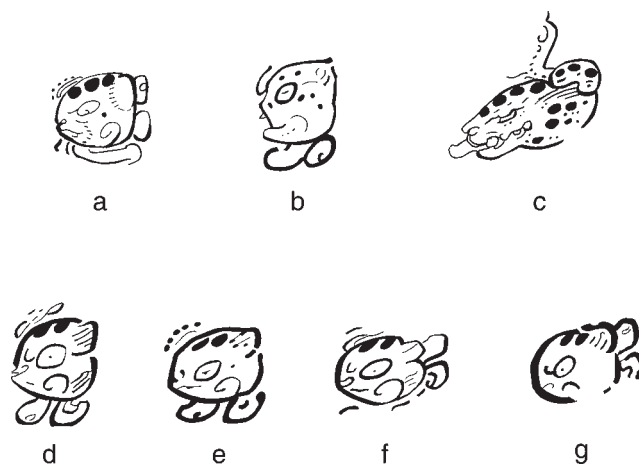
The languages that show this practice—Chiapas and Chimalapa Soke, Ayapa Gulf Sokean, the Ch'olan languages, Yukateko, and Q'eqchi7—form a single, continuous diffusion area (Figure 2), so these formations are clearly related. The Ch'olan form is reconstructible, to proto-Ch'olan *\*b'ahläm=te7*, since it is found in both branches of the subgroup.

Ch'orti7, the only other Ch'olan language, has a word *b'ajram=te7*, the expected form of the descendant of proto-Ch'olan *\*b'ahläm=te7* 'pataxte', but this word is glossed as a bush that is used to cure sprains. This could be a semantic shift or an independent formation.

Pre-*proto-Ch'olan* *\*b'ahläm=te7* is also reflected in Yukateko *b'ahläm=te7*. Whether this term is reconstructible to earlier stages of Yukatekan is unknown, since words for *pataxte* have not yet been found in Lakantun, Itzaj, or Mopan. The term appears to be a borrowing, since *\*b'ahläm=te7* 'pataxte' is reconstructible for proto-Ch'olan, while *=te7* 'tree' is not a native element in Yukatekan.

A similar term is Q'eqchi7 *b'a:lām kakaw* (also *kakaw b'a:lām*). Q'eqchi7 has borrowed many Ch'olan words, especially for plants and animals, but there are no known early borrowings from Q'eqchi7 into Ch'olan. It is therefore more likely that *b'a:lām kakaw* was borrowed into Q'eqchi7 from Ch'olan than into Ch'olan from Q'eqchi7. The formation is identical to that of Ayapa Gulf Sokean *kanh=kak*. Within Sokean, however, the diversity of the forms and the lack of documentation of words for *pataxte* in some languages leave the details of the historical development across this geographical area unclear.

One of these terms for *pataxte* may be recorded in Mayan hieroglyphic texts (Figure 3). It is most frequent on a set of "codex-style" vessels (K531, K1197, K1344, K1371, K1560, K4546). They seem



**Figure 3.** Proposed JAGUAR × CACAO glyphic conflations, from (a) K1344 and (b) K1182, compared with (c) an iconographic example of a jaguar head from another codex-style vase, K531, and with less definite examples of "jaguar cacao" glyphs from (d) K1371, (e) K4546, (f) K531, and (g) K1560.

to have been produced by a single scribe or scribal school, given the similarities in features of calligraphic style (as pointed out to us by David Stuart, personal communication 2005) and of sign choice (see Figure 4). In these cases, the skullcap of the CACAO sign appears to be marked with two or three jaguar spots in the same way as iconographic depictions of jaguars on the same vessels (Figure 3c). The two clearest instances (Figure 3a–b) are on K1344, in which the CACAO sign has whiskers and other mammalian features, the oval shape and shading of the jaguar spots is most pronounced, and there are three jaguar spots rather than two, and on K1560, in which jaguar spots appear on the cheek as well as the skullcap. Cases with two spots on the skullcap are found among these same vessels (Figure 3d–g). This usage is rare outside the texts by this scribal group, and their interpretation is not as clear. Totally unclear are instances in which a pair of blackened strokes appear together on the CACAO sign's skullcap. They may be an independent feature, but they could be a conventionalized reduction of a pair of jaguar's spots. Intermediate are instances that have two more clearly oval spots on the skullcap.

The most straightforward interpretation of these composite signs is as a spelling <JAGUAR × CACAO>, presumably for a Lowland Mayan term *b'ahlām kakaw* that is our postulated source for Q'eqchi7 *b'a:lām kakaw*. This Epigraphic Mayan term would appear to have predated Ch'olan + Yukateko *\*b'ahläm=te7*, since only the descendants of the latter term and not of *b'ahlām kakaw* survive in modern Lowland Mayan languages. The diffusion of *\*b'ahläm=te7* within Ch'olan can hardly have postdated the Classic Mayan collapse, so the spread of "jaguar cacao" as a term for *pataxte* must have taken place no later than the Late Classic period. In principle, the spelling could have persisted after *\*b'ahlām=te7* replaced *\*b'ahlām kakaw* in Ch'olan.

#### Summation

The widespread and early diffusion of the word *\*kakaw(a)* into a large number of Mesoamerican languages and language families is consistent with what is known about the diffusion of words for other cultigens from Mije-Sokean languages. In contrast, no

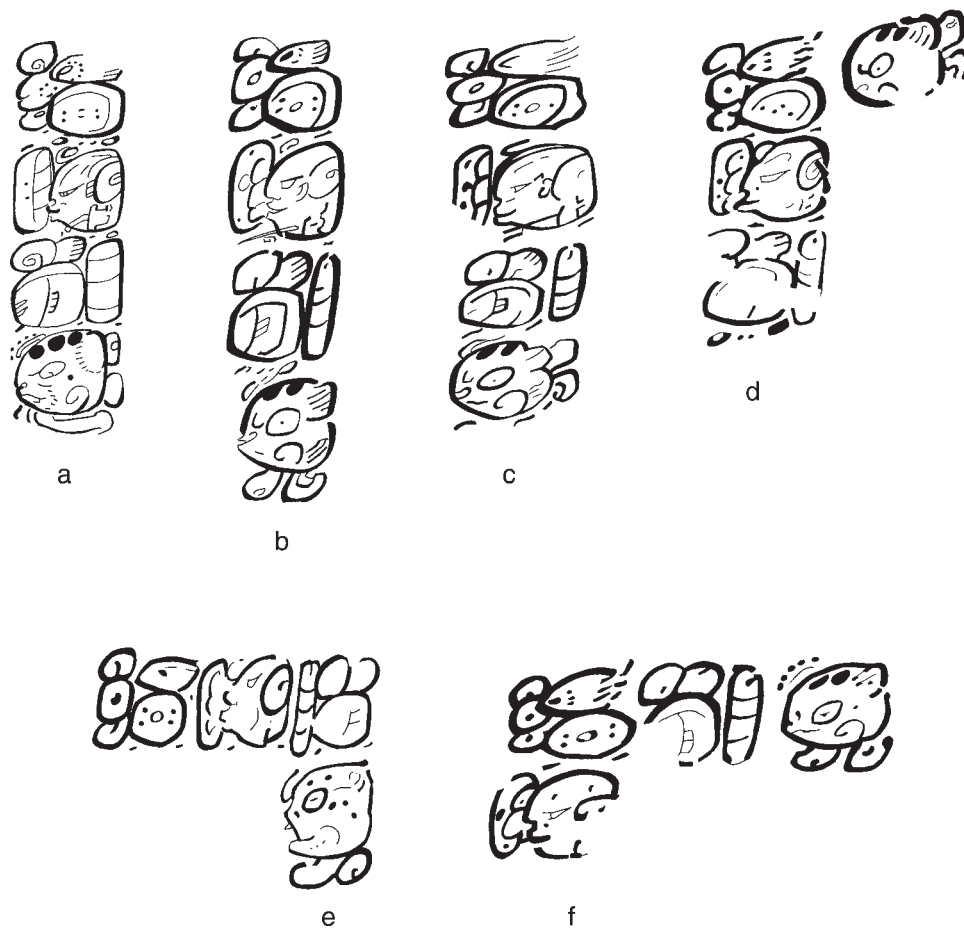


Figure 4. Stylistically similar texts from codex-style vases. (a) KI344; (b) KI371; (c) K531; (d) KI560; (e) KI182; (f) 4546. KI197, which is similar, is not illustrated because the sign for cacao seems to have been repainted.

words for cultigens are known to have diffused so widely at such an early date from any other language family. In addition, one of the prime areas of cacao cultivation, in the lowlands of Tabasco, was part of the (Mije-Sokean-speaking) Olmec heartland, where Gulf Sokean languages are still spoken today. These facts provide strong circumstantial support for the linguistically motivated hypothesis that this word is native to the Mije-Sokean family and diffused into other Mesoamerican languages, ultimately, from speakers of one or more Mije-Sokean languages. In fact, setting aside the recently contested case of *\*kakawa*, no currently reconstructed proto-Mijejan, proto-Sokean, or proto-Mije-Sokean term has a demonstrated foreign origin, so any alternative to a Mije-Sokean source requires substantial independent evidence.

#### EVIDENCE AGAINST A RECENT ALTERNATIVE HYPOTHESIS

In a recent article, Dakin and Wichmann (2000) claim that the word *\*kakawa* is not original in Mije-Sokean and that, instead, it developed in Yuta-Nawan and diffused into other Mesoamerican languages from the Nawa form. These conclusions are based on linguistic arguments: Wichmann's arguments that a Mije-Sokean origin is inconsistent with internal Mije-Sokean linguistic data, and Dakin's arguments that there is a satisfactory Yuta-Nawan etymology for the Nawa form. From these conclusions, they derive a

series of what would, if true, be important culture-historical inferences, culminating in the claim that the term diffused in association with the Teotihuacano diffusion sphere, and thus—based on this single word—that the dominant language of the city of Teotihuacan was Nawa. But they correctly emphasize (Dakin and Wichmann 2000:69) that their hypotheses hinge above all on the linguistic arguments.

We demonstrate in this paper that the linguistic arguments for Dakin and Wichmann's conclusions range from invalid to highly speculative by showing that there is positive evidence for and no difficulty with the hypothesis of a Mije-Sokean origin for the term *kakaw(a)*, and by showing that their proposed Nawa origin for the term is not possible. We further show that there is no alternative to a Mije-Sokean origin that is consistent with what is known about linguistic diffusion in Mesoamerica.

We close with a discussion of the culture-historical context of the diffusion of the term. In particular, we summarize arguments presented in more detail elsewhere (Kaufman 2000–2007; Kaufman and Justeson 2007) that Nawa cannot have had an influential role at Teotihuacan in its heyday, but that it is very likely that Teotihuacan was influential in spreading the word for cacao, and presumably the practices surrounding its use, in and around the Basin of Mexico and perhaps more widely.

The chief linguistic arguments presented by Dakin and Wichmann are (1) a “refutation” of Kaufman's reconstruction of *\*kakawa* as a bona fide Mije-Sokean word, arguing instead that it

diffused, separately into Mijejan and Sokean, from outside the family; and (2) an attempt to show that the Nawa form of the word was not borrowed but, rather, that it originated within Nawa, descending from native Yuta-Nawan vocabulary. Both of these arguments are fallacious. The authors have treated these words in isolation, making undemonstrated and, it turns out, false assumptions about both Mije-Sokean and Yuta-Nawan language history.

A secondary argument by Dakin and Wichmann is that the word *chokol*=*a:-tl* for ‘chocolate (drink)’ is also a native Nawa term. They discuss this word to raise and argue for the possibility that its diffusion throughout Mesoamerica was related to that of *kakawa*.

### General Features of Borrowings into and out of Nawa

Before evaluating Dakin and Wichmann’s specific arguments, they can be framed in the context of what is known about lexical diffusion in Mesoamerica more generally. Our view that Nawa *kakawa-tl* is a borrowing from Mije-Sokean, and not into it, is concordant with a number of facts about linguistic diffusion in Mesoamerica that are set forth in the previous section, while the contrary view is not consistent with what is otherwise known about diffusion from Nawa. Setting aside the contested case of the word for cacao, four empirical observations can be made about cleanly established borrowings between Nawa and other Mesoamerican languages:

(1) Nawa in general is heavily influenced by Mije-Sokean and Totonakan, and to a lesser degree by Wastekan in its pre-Nawa stage (see “Demonstrating Borrowing”). In contrast, neither Mije-Sokean in general nor Totonakan in general is much influenced by Nawa. Even Wasteko (a single language rather than a family of languages), which has had Mesoamericanized Nawa as a neighbor since at least A.D. 900, has relatively few lexical borrowings from Nawa. This makes Nawa look like a relative newcomer to Mesoamerica, not in a position to provide very early loans into languages throughout Mesoamerica.

(2) More specifically, proto-Nawan—the common ancestor of all forms of Nawa—shows substantial borrowing from Mije-Sokean languages in both vocabulary and grammar.

(3) Setting aside the contested case of the word for cacao, proto-Mije-Sokean does not show a single plausible instance of a lexical borrowing from Nawa or Yuta-Nawan, and neither does proto-Mijejan or proto-Sokean or any other genetic subgroup of Mijejan or Sokean. (Individual Mije-Sokean languages have borrowed some Nawa lexical material.)

(4) Setting aside the contested case of *\*kakawa*, there is no demonstrated instance of a Nawa loan into any Mesoamerican language that clearly predates the Late Classic period. In particular, no loans from Nawa have undergone sound changes characteristic of any genetic group of languages. Rather, Nawa loans in Mesoamerican languages reflect Nawa phonology as we know it from the sixteenth century and therefore cannot go back earlier than about A.D. 1000.

There are proposed counterexamples to these claims, mostly by Dakin, but none is cogent. (See “Demonstrating Borrowing”, earlier, for refutations of a sample of such proposals.)

Given points (2)–(4), compelling evidence is required to make a case for a word of Nawa origin having diffused into Mesoamerican languages at a substantially earlier time period, especially one that was borrowed as widely as *\*kakawa*.

There are two other general problems with the claim of widespread early borrowing of *\*kakawa* from Nawa.

(5) Nawa nouns are almost always borrowed in their unpossessed form and reflect the absolute suffix *-tl~tli~li* when this suffix is present in the Nawa model (Kaufman 2000–2007). In Mayan and Mije-Sokean languages in particular, then, one would expect to find something like *kakawat\**, rather than the attested *kakaw*, if this word were indeed a borrowing from Nawa. By way of illustration, we first discuss a set of borrowings into a single language, and then we discuss borrowings that are widespread in a particular language family.

Table 3 presents all of the Nawa loans in Kaufman’s data on Soteapan Gulf Sokean (Kaufman and Himes 1993–2005), which can serve as a typical instance of how the Nawa absolute suffixes—in particular, *-tl(i)*—appear in loans into other languages. In all 15 cases where the Nawa source has a word-final *-tl* or *-tli*, Soteapan Gulf Sokean has a straightforward reflex of it. In five cases, Soteapan Gulf Sokean has a reflex of *-tl*, *-tli*, or *-li* where present-day Gulf Nawa lacks the suffix (*taanajti*, *kukuji*, *lupuji*, *manteeka7t*, *xik7ipiili*). These borrowings into Soteapan Gulf Sokean presumably reflect an earlier stage of Gulf Nawa, though one that postdates the arrival of Spanish speakers.

There are 13 nouns of Nawa origin that have a fairly wide spread in Mayan languages. In each case, some languages are likely to have adopted the word directly from Nawa, while others probably received it from another Mayan language.

Three of these—*#mis* ~ *#mistuun* ‘cat’, *#masa:t* ‘deer’, and *#xunakat* ‘onion’—date from after the arrival of the Spanish.

Eight are found in the Guatemala highlands and bear the traits of Gulf Nawa or Pipil: *#koht* ‘eagle’, *#to:ch(in)* ‘armadillo’, *#karat* ‘frog’, *#chikiwit* ‘basket’, *#no:chti7* ‘prickly pear’, *#matzahti7* ‘pineapple’, *#nakatamal* ‘meat tamale’, *#nawal* ‘shape-shifter’. Whatever absolute suffix was present in Gulf Nawa is also present in the Mayan borrowing. Gulf Nawa /*kohti*/ ‘eagle’ has been shortened to make it look more Mayan.

Two Nawa loans into Mayan present interesting problems:

Nawa *tena:mitl* ‘fortified place’ (possessed theme *-tena:n*) appears in Mayan languages of Huehuetenango and in Wasteko as *#tena:m* ‘fenced-off area, town’. These forms seem to lack the Nawa absolute suffix. However, the possessed form of *tena:mitl* is Poss-*tena:n*, which ends in /*n*/, not /*m*/, so the possessed form was not the basis of the borrowing. It is plausible that, if the form *#tena:m* is borrowed directly from Nawa *tena:mitl*, it has been shortened by Mayans, and that is why the /*m*/ was kept. K’ichee7an languages borrow *tena:mitl* as *#ina:mit* ‘town’, and the absolute suffix is preserved. However, there is room for doubt as to whether the word is of ultimate Nawa origin, since it is found in only some varieties of Nawa and is not analyzable. While one may start by hypothesizing that the segment /*te-*/ corresponds to the root {*te*} ‘stone’, the sequence /-*na:m(i)*/ is not a known Nawa morpheme. Thus, the word may have originated outside Nawa. While there can be no doubt that *#tina:mit* comes from Nawa, the form *#tena:m* is only possibly of Nawa origin, not definitely so.

The last item constitutes an exception to the norm that nouns of Nawa origin are borrowed with a reflex of the Nawa absolute suffix, when one is in fact present in the Nawa original. The word *#xa:n* ‘adobe brick, wall’ occurs so widely in Mayan that a form *\*xa:n* would be reconstructed for proto-Mayan if the Nawa word *xa:mitl*, possessed-*xa:n* ‘adobe’ were not attested. The Mayan word definitely seems to reflect the possessed form of the Nawa noun. It is irrelevant that Nawa *xa:mitl* is in turn borrowed—from Mije-Sokean. The motive for borrowing ‘adobe’ in its possessed form is not obvious.

Table 3. Nawa loans into Soteapan Gulf Sokean and the fates of their absolutive suffixes

Nawa Source	Form in Soteapan Gulf Sokean	Gloss
A. Group I (6×): Nouns that have no suffix at all and are not borrowed with one. Words of this class are generally rare in Nawa but include all nouns referring to physical deformities.		
tzapa	chapa	'chaparro' // 'short in height'; 'enano' // 'dwarf'
	tzapu <del>p</del> xiny	'hombres encantados de la montaña' // 'bush dwarves' [also from /tzapa/]
cha:n.ej	chaanij	'chaneque' // 'forest spirit'
xote	xutyi	'caracolillo' // 'periwinkle'
we:weh	weewej	'abuelo' // 'grandfather'; 'señor grande, señor que tiene edad' // 'old man'
xipin (MEC)	xipin	'pito, pene' // 'penis'
a:ka=tzana (COX) <sup>a</sup>	chana	'zanate' // 'grackle'
B. Group II (3×): Absolute <i>-li</i> is now lacking after stem-final /l/ in polysyllables in Gulf Nawa; word-final /l/ went to /t/ in early loans into SOT because SOT lacked a phoneme /l/.		
ma:tla=yawal-li	maatayawat	'matayaguala/e' (MEC, PAJ have ma:ta=yawal) // 'kind of fishing net'
na:wal-li	naawat	'nagual' // 'shape-shifter'
yawal-li	yawat	'yagual' (MEC has yawal) // 'cloth ring'
C. Group III (15×): Nouns that end in the absolute suffix <i>-tl</i> ~ <i>-tli</i> ~ <i>-li</i> when unpossessed and are borrowed with a reflex of the suffix (except possibly for nouns that are typically possessed, of which there are no examples in these data).		
ka:ka:x-tli	kaakaxtyi	'cacaste' // 'crate'
xi:koh-tli	xiikujti	'abejón, jicote' // 'bumblebee'
no:ch-tli	nuuchtyi	'nopal, pita(ha)ya' // ' <i>Cereus</i> '
ta:nah-tli	taanajti	'tenate (hecho de palma)' // 'palm basket' (MEC, PAJ have ta:nah)
so:ya-tl	suyat	'palma' // 'palm tree'
a:-l = tepe:-tl	7aatpe7et	'Soteapan; pueblo; ciudad' // 'town'
epaso:-tl	7epasu7t	'epazote' // 'goosefoot'
teki-tl 'work'	tek7et	'obligado (sin querer)' // 'obligatorily'
te:ska-tl	teeskat	'espejo' // 'mirror'
tzitzimi-tl	tzitzimat	'abuela de Jomxuk' // 'Jomxuk's grandmother'
tzohmi-tl 'hair, fur'	tzujmity	'cobija' // 'blanket'; 'zarape' // 'poncho' [post-contact]
xikipi:l-li	xik7ipiili	'chiquipili' (type of lizard)
<coco> ⇒ *ko:koh-tli <sup>b</sup>	kukujti	'coco' // 'coconut' (PAJ has ko:koh)
<lobo> ⇒ *lo:poh-tli <sup>b</sup>	lupujti	person who transforms into a donkey
<manteca> ⇒ *mante:ka-tl <sup>b</sup>	manteeka7t	'manteca' // 'lard' (PAJ has mante:gah)
D. Group IV (2×): Nouns that take the absolute suffix <i>-in</i> in the singular.		
to:l-in	tuulin	'tule' // 'cattail'
kapol-in	kapulin	'capulín' // 'chokecherry'

Note: SOT = Soteapan Gulf Sokean (Kaufman and Himes 1993–2005); MEC = Mecayapan Gulf Nawa (Wohlgemuth 2000); COX = Coxcatlan Gulf Nawa (Kaufman 1969/1984–1993/2006); PAJ = Pajapan Gulf Nawa (Peralta 2002–2007).

<sup>a</sup>This Nawa form occurs both with and without *-tl*.

<sup>b</sup>From Spanish via Nawa. The source of these Nawa loans is Gulf Nawa as spoken in Mecayapan (Wohlgemuth, et al. 2000) and Pajapan (Peralta 2004), Veracruz. The Nawa forms are cited, though, as they occur in the more conservative *tl* dialects.

Two other proposed borrowings, from Nawa into Mayan, are not demonstrably early loans from Nawa. One is not demonstrably a loan, and one is demonstrably not early.

One word, \**ku:m* 'pot', is reconstructible for proto-Yukatekan, which broke up circa A.D. 1000. This word is likely to share a common history with Nawa *ko:mi-tl* 'pot, water jar'. This could conceivably be by borrowing from Nawa, but the word has no Sonoran etymology, and it is equally plausible that it entered both languages from another source. It may be entertained that the word entered Nawa from an ancestor of proto-Yukatekan, but it is not clear where the contact could have occurred that would make this a realistic option.

On the basis of Ch'ol *chikib'* and Ch'orti7 *chiki7*, Kaufman and Norman (1984:118) reconstructed \**chiki7* to proto-Ch'olan because it was formally possible to do so. As a historical statement, however, this was misleading in that the form *chiki7* reflects a contraction from Nawa *chikiwi-tl* that is otherwise known only in Gulf Nawa,

which probably did not become distinct from other forms of Nawa before A.D. 800, whereas epigraphic evidence and glottochronology both put the breakup of Ch'olan before that date. The Ch'ol and Ch'orti7 forms therefore represent independent borrowings from some kind of Gulf Nawa.

In summary, out of ten pre-Columbian borrowings from Nawa in Mayan languages, only one was borrowed without the absolutive suffix. These data bear on generalizations about tendencies in the process of borrowing Nawa nouns (see "Evidence that \**kakaw(a)* was not Borrowed from Nawa . . .", below).

(6) Cacao does not grow anywhere near the Basin of Mexico, nor farther north in areas from which Nawas entered Mesoamerica. It is completely plausible that Nawa would have borrowed its word for cacao from a language localized in the area in which it grows. This is in fact the norm when newcomers encounter unfamiliar plants and animals, whenever they do not devise neologisms using native resources. Nawa was strongly affected by one or

more Mije-Sokean languages in its vocabulary, its morphology, and its syntax. Apart from Mije-Sokean, early forms of Nawa show substantial influence only from Wasteko and Totonakan, and the word *kakaw* of Wasteko and Totonakan, if borrowed into Nawa, would have been expected to yield *kakaw-tli* rather than *kakawa-tl*. On the face of it, Mije-Sokean is the only plausible source for Nawa *kakawatl*.

Given these characteristics of lexical diffusion from Nawa, any proposal for a widespread lexical borrowing from Nawa in the Preclassic or Early Classic has to be approached with skepticism and requires compelling linguistic evidence to be accepted. We show below (“Refuting the arguments . . .”) that the evidence presented by Dakin and Wichmann for a Nawa origin of the word for cacao does not meet such a standard. Instead, their proposed Yuta-Nawan etymology is speculative, as is their attempt to deal with the absence of the Nawa absolute suffix—the only one of the linguistic issues that they address—and the supporting arguments for each of these proposals are invalid. In addition, nothing in the evidence provided by Dakin and Wichmann undermines the arguments that Mije-Sokean *\*kakaw(a)* was the source of the word for cacao.

Dakin and Wichmann’s proposed scenario for the diffusion of *kakawa* from Nawa involves the hypothesis that the word for cacao was diffused by Nawa-speaking Teotihuacanos who controlled the Mesoamerican trade in cacao. But Teotihuacan had a massive impact on other Mesoamerican societies. This level of impact would have been accompanied by a substantial impact on at least the vocabularies of many other Mesoamerican languages. Their historical scenario therefore requires a substantial body of early Nawa loans into Mayan, Xincan, Sapoteko, Wastekan, and probably Tarasko and Matlatzinkan languages. So substantial an early impact would be detectable in a large number of obvious and unproblematic candidates for early borrowings.

Dakin and Wichmann’s proposal for the diffusion of *kakawa* from Nawa in association with Teotihuacan’s interactions with other Mesoamerican societies raises a further chronological problem. They argue (Dakin and Wichmann 2000:57) that internal evidence within Mije-Sokean shows that “a word *kakawa* or one close to that in form was borrowed into the [Mije-Sokean] linguistic family, but at a time when it was still at an early stage of differentiation.”

The breakup of Mije-Sokean probably took place between about 1200 and 800 B.C. (1) Elsewhere (Kaufman 2000–2007; Kaufman and Justeson 2007), we provide evidence that the source of Mije-Sokean loans in central Mexican languages was a northern branch of Mije-Sokean that was in the Basin of Mexico. The loans reflect a vocabulary that was not differentiated between Mijejean and Sokean, so the speakers of this language probably arrived before or around the time of the breakup of Mije-Sokean proper. Archaeologically, their arrival can be dated to 1200 B.C. or earlier. (2) This constraint agrees with the glottochronological estimate of 1000 B.C. for the breakup of Mije-Sokean. (3) The previous section shows that a body of Sokean loans, including names of four animals associated with day names, entered pre-Sapoteko from Sokean, so Mijejean and Sokean must have been well differentiated at that time. One of these borrowings can be dated to before 200 B.C. (4) Epi-Olmec data from the La Mojarra stela (A.D. 157) and the Tuxtla Statuette (A.D. 162) show only two lexical retentions from proto-Mije-Sokean that are now associated with Mijejean but 37 words that are now found only in Sokean. This much differentiation

is likely to require at least 1,000 years to develop. Certainly, it does not occur within as few as 500 years of differentiation, after which dialects are always close enough to be inter-intelligible. The Epi-Olmec data therefore put the early differentiation of Mije-Sokean well before 300 B.C., and probably back to 800 B.C. or earlier.

Even the latest of these estimates is far too early for the diffusion of *anything* from Teotihuacan, which was not established until around 150 B.C., and whose non-local impact did not begin until the second century A.D. It is even further out of line with the known timing of lexical diffusion from Nawa.

In the remainder of this paper we show that, contrary to Dakin and Wichmann, the Mije-Sokean data are consistent with a Mije-Sokean origin for *kakawa* and that their alternative Yuta-Nawan origin in a reduplicated form, descending from a word for egg, is impossible.

### Consistency of the Mije-Sokean Evidence

Dakin and Wichmann (2000:57) claim that their analysis of Mije-Sokean data pertaining to the word for cacao demonstrates that “it is not possible to continue to attribute a Mixe-Zoquean origin to it. Instead we argue that *kakawa*—most likely pronounced *kàkawá* by its donors—entered [Mije-Sokean] from the outside . . .”. They make three claims in favor of the view that *kakawa*, or something similar, was borrowed into Mije-Sokean languages rather than having descended from a native Mije-Sokean lexical item. One is the claim that “morphemes consisting of three open syllables (CV.CV.CV)” are virtually unknown in Mije-Sokean, while they are common in Nawa.” (Verb roots are either monosyllabic ending in a consonant or disyllabic ending in a consonant, and most affixes are monosyllabic; these and other morpheme-structure constraints in Mije-Sokean languages mean that only noun roots can be at issue.) A second claim is that the stress pattern of proto-Sokean *\*kakawa* was demonstrably in conflict with Sokean stress rules. Both claims are untrue, and the arguments based on them would be invalid even if they had been true. Their third claim is the correct observation that while Sokean languages point to an ancestral *\*kakawa*, Mijejean languages point to an ancestral *\*kakaw*; these two forms cannot both descend normally from a single proto-Mije-Sokean form. Evidence cited later shows that there are numerous cases of discrepancies (though they are hardly prevalent) between proto-Sokean and proto-Mijejean reconstructions for what are incontrovertibly single etymologies. All three of Dakin and Wichmann’s claims are taken by them as evidence that *kakawa* was not native to the Mije-Sokean family.

This section addresses the reconstructibility of *\*CVCVCV(C)* trisyllabic roots in Mije-Sokean and Gulf Sokean evidence for initial stress in these roots, establishing the regular descent of Sokean forms from proto-Sokean *\*kakawa*. We address the evidence for the regular descent of all pre-Columbian Mijejean forms for ‘cacao’ from proto-Mijejean *\*kakaw*. This shows that the form was in Sokean before the proto-Sokean era and in Mijejean before the proto-Mijejean era. Finally, it addresses the nature of the discrepancy between proto-Sokean *\*kakawa* and proto-Mijejean *\*kakaw*.

Our discussion is based on Table 4, which presents our Mije-Sokean data on all of the trisyllabic roots that are reconstructible for proto-Mijejean, proto-Sokean, or proto-Mije-Sokean. Each reconstructible proto-Sokean trisyllabic form survives as a trisyllable in Soke but is reduced to a disyllable in Gulf Sokean.

Table 4. Trisyllabic \*CVCVCV roots reconstructible to proto-Mijejan, proto-Sokean, or proto-Mije-Sokean

	cacao	mamey ⇒ plantain	fire	guava	apompo (tree: <i>Pachira</i> ) Guiana chestnut	peccary
Proto-Mije-Sokean					*wakata	
Proto-Sokean	*kakawa	*sapane	*jukutɨ	*patajaC	*wakata	
San Miguel Chimalapa Soke	kakawa	xapane	jukutɨk	pataja7		
Santa María Chimalapa Soke	kakawa	sapane	jukutɨ	pataja	wakata	
Copainala Soke	kakawa7	sapane	jukutɨk			
Tecpatan Soke	kakawa7	sapane	jukutɨk			
Francisco León Soke						
Rayón Soke	kakwa					
Ayapa Gulf Sokean	[ka(:)gwa] (~ kak=)	xapne= (~ xap)	jukte	[pa:7da] (~ pa7danh=)	wakta (~ waktanh=)	
Soteapan Gulf Sokean	ka:kwa	sa:ɲni (xapan=chay 'hojas secas del platanar')	juktɨ	patanh	waakta	
Texistepec Gulf Sokean	kaga= (~ ka:k#)	Sapun	jugut	patanh		
Proto-Mijejan	*kakaw				*wakata	*7i:tzɨmɨ
Oluta Mijejan	kaka[7]w				wakata kuy	7i:tzɨmɨ
Sayula Mije(an)	kagaw				wa:gat 'guacalsuchi'	7i:chim
Western Mije (Totontepec)	kaku					7i:tzɨm
Eastern Mije (Guichicovi)	kugá:W [kaká]					[7ɨdzɨm]

Notes: AYA = Ayapa Gulf Sokean; SOT = Soteapan Gulf Sokean. The data cited here were collected by linguists working on the PDLMA, except data from Rayón Soke (from Harrison and Harrison 1984). The trisyllabic variant of Wichmann's (1995:464) incorrectly reconstructed proto-Mije-Sokean \**tuju7k(ay)* 'yesterday' would also agree with our rules in having initial stress. However, the correct reconstruction, \**tuju7k*, is a disyllabic word followed (optionally, in just one descendant) by a clitic {7ay}; as such, stress rules operate on \**tuju7k* alone. This root therefore does not provide data on trisyllabic roots in Mije-Sokean languages and is not included in the table.

*Mije-Sokean trisyllabic roots.* Dakin and Wichmann state that only two CVCVCV stems are reconstructible within Mije-Sokean, and that this in itself suggests that a word shape *kakawa* must be foreign. It is true that most Mije-Sokean noun stems are disyllabic, and most of the remainder are monosyllabic, while trisyllabic stems are rare. But this does not constitute evidence that trisyllabic roots in Mije-Sokean languages must have or are likely to have a foreign source. It is a commonplace that languages admit a variety of syllable types as roots or stems, and some are much rarer than others. As Table 4a shows, six trisyllabic roots are reconstructible in proto-Mijejan, proto-Sokean, or proto-Mije-Sokean. All of them, including \**kakawa*, begin with two open (CV or CV7) syllables. Empirically, setting aside the contested case of *kakawa*, none of these trisyllabic nouns has a recognized foreign source.

Note that CV7 syllables must be treated as open for present purposes. In Mije-Sokean languages, V7 has the same consequences as V for the assignment of stress. In Sokean languages that lengthen stressed short vowels in open syllables, where  $C\bar{V}CV > C\bar{V}:CV$  or  $C\bar{V}VCV$ , it is also the case that  $C\bar{V}7CV > C\bar{V}:7CV$  or  $C\bar{V}7VCV$ . Accordingly, V7 behaves as a syllable nucleus, not as a vowel-plus-consonant sequence.

*Gulf Sokean reflexes of Mije-Sokean trisyllabic roots.* Dakin and Wichmann further claim that the Gulf Sokean words for cacao must be loans into Gulf Sokean. Their reason is that Wichmann reconstructs penult stress for all native words in Sokean. Dakin and Wichmann therefore consider roots with stress on a different syllable to be non-native—hence, borrowed forms.

In 1963, Kaufman offered the following rules for stress in proto-Mije-Sokean (section 2262 of the manuscript, somewhat

paraphrased in the interest of updating terminology but not of updating the analysis, which was faulty in one respect):

Rules for stressing polysyllabic stretches excluding clitics in proto-Mije-Sokean:

a. nouns, adjectives, and numerals are stressed on the first syllable and also on the penult syllable if the latter is at least two syllables forward of the first syllable.

b. verbs are stressed on the penult syllable and also on the first syllable if the latter is at least two syllables earlier than the penult syllable.

[These rules are distinct because proto-Mije-Sokean verb words always contain at least two syllables, while non-verbs can be monosyllabic.]

If two or more syllables are stressed by the above rules, the last one is primary and the rest secondary.

The different rules for nonverbs and verbs may reflect their differing positions with respect to sentence stress, but this is pure conjecture.

From the vantage of 2007, these statements are too complicated. There is no need to refer to the lexical class of any word. As for the exclusion of clitics, a clitic is a word or affix distinctively lacking stress. It is by definition not within the scope of stress assignment—unless in the phonology of the language clitics are attached before stress assignment.

In proto-Mije-Sokean, proto-Sokean, and proto-Mijejan, finite verbs always occur with one suffix containing a vowel. They follow a rule of penult stress, which in an unenlightened way can be viewed as stem-final stress as long as no non-final inflexional

suffixes intervene between the verb stem and the word-final suffix. If they have three or more syllables, the first syllable will also be stressed. Non-verbs follow the same rule: penult stress, and initial stress in phonological words of three or more syllables. The penult stress is stronger than earlier word-initial stresses.

Such is Kaufman's restatement of what he formulated in 1963, and we take this to be the correct formulation.

As for Gulf Sokean, an innovation must be postulated. Trisyllabic words of the shape \**CVCVCV(7)* are stressed on the initial syllable and not (anymore) on the penult syllable.

The data under discussion are all nouns, but proto-Sokean verbs of the shape *CVCV<sub>1</sub>C<sub>2</sub>* were restructured as *CVCV<sub>1</sub>7* in Soteapan Gulf Sokean and Texistepec Gulf Sokean, but not in Ayapa Gulf Sokean. Proto-Sokean also had verb stems of the shape *CVCV<sub>1</sub>7*). Then the pattern *CVCV7* underwent in the several Gulf Sokean languages the same innovation in stress assignment before being restructured to the shape *CV:C* (in Soteapan Gulf Sokean and Texistepec Gulf Sokean) or *CV:7C* (in Ayapa Gulf Sokean):

GRAMMATICAL CATEGORY	UNDERLYING FORM	PRONUNCIATION
Incompletive	CVCV7-pa	[CV:CV7pa]
Completive	CVCV7-wɸ	[CV:CV7wɸ]
Imperative	CVCV7-ɸ7	[CV:CV7ɸ7]
Optative	CVCV7-7in	[CV:CV77in]
Dependent	CVCV7-e	[CV:CV7e]
	~CVCV7-i	[CV:CV7-i]
Nominalized	CVCV7-e	[CV:CV7e]
	~CVCV7-i	[CV:CV7-i]

All of these verb forms lost the second syllable (including the glottal stop), and from these contracted forms a new stem shape was generalized.

These forms show that in Gulf Sokean trisyllabic words of the shape \**CVCV7CV(C)* underwent the same innovation in stress placement as \**CVCVCV(C)* words. In contrast, trisyllabic words of the shape \**CVCVCCV(C)* show penult stress. In the case of verbs, this can be illustrated by any *CVCVC* verb stem in which the final consonant is not 7. For example:

GRAMMATICAL CATEGORY	UNDERLYING FORM	PRONUNCIATION
Incompletive	tokoy-pa	[tokóypa]
Completive	tokoy-wɸ	[tokóywɸ]
Imperative	tokoy-ɸ7	[tokó:yɸ7]
Optative	tokoy-7in	[tokóy7in]
Dependent	tokoy-e	[tokó:ye]
Nominalized	tokoy-e	[tokó:ye]

Wichmann (1995:88–89) outlines his views regarding stress in proto-Mijejan, proto-Sokean, and proto-Mije-Sokean. We find these rules to be overly complicated. Wichmann relies on syllable weight and the ability to identify roots, both of which are unnecessary. All that is required to predict stress in any Mije-Sokean language (with the possible exception of Sayula Mijejan [Rhodes

et al. 1994–2005] and Texistepec Gulf Sokean) is knowledge of which morphemes are clitics.

Wichmann (1995:68) gives the following rule: “Assign P[primary] S[tress] to the rightmost heavy syllable in the word string if any such syllables are present. (A heavy syllable contains V:, V:7, V7, or Vh as its nucleus.) Or else assign P[primary] S[tress] to the rightmost root. If that root is a polysyllabic nonverb, it receives stress on its penultimate syllable; if it is a polysyllabic verb it receives stress on its last syllable.” Leaving aside heavy syllables and roots, Wichmann's rule for the heaviest stress is the same as that of Kaufman (1963), although Wichmann does not acknowledge this. (He criticizes other features of Kaufman 1963). Nor does he notice that Kaufman's (1963) rule could have been formulated more simply, as we have done here. Wichmann does not envision more than one stress in polysyllabic words.

While it is true that penult stress is typical of Mije-Sokean languages, in some languages there are additional wrinkles. In Santa María Chimalapa Soke (see Kaufman and O'Connor 1994–2005), every word of three or more syllables is stressed on both the initial syllable and the penult syllable. The penult syllable *may* have a *slightly* higher prominence than the first syllable. In San Miguel Chimalapa Soke (see Kaufman and Johnson 1994–2005), however, in words of three syllables the penult is stressed, and if the first syllable is heavy, it receives secondary stress. If the word has four or more syllables and the first syllable is open, it is stressed and long and the second vowel is optionally (and usually) dropped. In both Chimalapa Soke dialects, vowels in stressed open syllables are allophonically long.

Dakin and Wichmann's evidence for antepenult (initial) stress on \**kakawa* is that the original penult syllable is lost in Gulf Sokean. This is cogent evidence for stress, because in roots whose structure and stress pattern can be established on language-internal grounds, it is only syllables known to be unstressed that are lost in Gulf Sokean. (Cross-linguistically, too, it is unstressed syllables that are most likely to be reduced.) They are also correct in limiting attention to the placement of stress in Mije-Sokean roots, not in arbitrary trisyllabic words, because (for a variety of reasons, not the same for each language) stress on words with more than one morpheme is predictable in Mije-Sokean languages.

Where Dakin and Wichmann go wrong is in claiming that trisyllabic roots in Sokean were regularly stressed on the penult syllable. An analogy from the penult stress of disyllabic roots is not valid, because this pattern can equally be treated as one of root-initial stress. What is required to establish the stress pattern on trisyllabic roots are data from trisyllabic roots themselves. Dakin and Wichmann's only evidence of this sort comes from the word for cockroach, reconstructed incorrectly by Wichmann as a root \**makoko*. This *word* indeed had stress on the penult syllable, as indicated by the preservation of that syllable in Gulf Sokean, but this is explained by the fact that its root is actually a disyllable \**koko7*, with a preposed optional proclitic \**ma*+ (see later), which, as a clitic, is not stressable.

Using a larger set of data, we show here that initial (or antepenult) stress was normal in Gulf Sokean roots of the shape \**CVCVCV(C)* and thus that the stress pattern of \**kakawa* is what is to be expected for native Mije-Sokean roots. There is no internal basis within the Mije-Sokean family for interpreting Sokean \**kakawa* as other than a native lexical item.

The Mije-Sokean data in Table 4 are relevant to words for cacao and to stress patterns in Mije-Sokean trisyllables. All five of the Sokean words in this table have a *CVCVCV* or *CVCVCVC* pattern



in Soke while having just two syllables in Gulf Sokean. To this extent, at least, these appear to be regular correspondences, pointing to the reconstructibility of these sets to proto-Sokean. The Soke forms are not predictable from the Gulf Sokean forms, while the Gulf Sokean forms are predictable from the Soke forms. It is therefore the latter that preserve the proto-Sokean forms of these words.

There are Mijejan forms that are cognate with one of these words, *\*wakata*.

### Discussion of Mije-Sokean Forms

This section discusses the details of the Gulf Sokean data.

*Cacao, mamey, fire, guava, and Pachira (apompo)*. These words show that proto-Sokean had trisyllabic noun stems of shape *CVCVCV(C)* and that, in proto-Gulf Sokean, these were stressed on the initial syllable. Generally speaking, Soteapan Gulf Sokean and Ayapa Gulf Sokean stress the first syllable and lose the second syllable, while Texistepec Gulf Sokean stresses the first syllable and loses the last syllable. Since Texistepec Gulf Sokean disagrees with Soteapan Gulf Sokean and Ayapa Gulf Sokean on which syllable is lost, this syllable reduction must have taken place after the breakup of proto-Gulf Sokean. In Totontepec Highland Mije (see Suslak 1996–2002), /ʔi:tzəm/ ‘peccary’, from proto-Mijejan *\*ʔi:tzəm*, also has initial stress. In Soke proper, all trisyllabic words are stressed on the penult syllable (and, as stated earlier, in Santa María Chimalapa Soke the first syllable is stressed, as well).

In some of these cases, Soteapan Gulf Sokean and Ayapa Gulf Sokean lengthen the stressed syllable, while in others they do not. There is too little data to provide a secure account for the presence or absence of vowel length in these forms. Possibly related is the fact that, in Soteapan Gulf Sokean, disyllabic roots serving as preponouns may lose their second vowel and may lengthen the first, stressed vowel.

Reduction of the second syllable is obvious in the Gulf Sokean words for cacao, mamey, fire, and *Pachira*, but the case of *\*patajaC* requires more extended explanation. The attested Soteapan Gulf Sokean form, *patanh*, has a short vowel in the first syllable, which can only result from an underlying synchronic *tt*. The only possible source for this *tt* that is consistent with the forms of cognates in Soke would be a cluster *tj*, resulting from a reduction of the middle syllable of *\*pataja(C)*, yielding pre-Soteapan Gulf Sokean *\*patja(C)*. In Ayapa Gulf Sokean, roots of the shape *CV:7CV* regularly arise from proto-Sokean and proto-Gulf Sokean *\*CVCV7*; Ayapa Gulf Sokean [pa:7da] therefore suggests a pre-Ayapa Gulf Sokean *\*pata7*, again presumably from *\*patja7*. Both Soke and Gulf Sokean therefore support reconstruction of word-final 7 in this form, and in particular require the reconstruction of *\*pataja7* for ‘guava’. Soteapan Gulf Sokean *patanh* in place of expected *pata7\** is consistent with a sporadic but fairly common phenomenon in individual Sokean languages whereby word-final 7 or *V* in one language corresponds to *nh* or *n* in another language. If the Texistepec Gulf Sokean form *patanh* is descended from an antecedent form *\*patajanh*, which may well be the case, we do not know the rules that would yield this form. Alternatively, the form may be a simple borrowing from Soteapan Gulf Sokean.

Ayapa Gulf Sokean has one word, *pa7tanh=kuy* ‘guava tree’, in which the word for guava is the first element in a compound. This compounding form has *nh* at the end of the root. In Soteapan Gulf Sokean, *patanh* occurs both in compounds and as an

independent word. Since the internal 7 of Ayapa Gulf Sokean *pa7tanh*= must originally have been root-final, the current *nh* at the end of the root must have developed in *pa7tanh=kuy* after the collapse of the unstressed penult syllable and, therefore, after the breakup of Gulf Sokean. In *waktanh=kuy*, Ayapa Gulf Sokean has *nh* at the end of the root in the compounding form of *wakta* < proto-Sokean *\*wakata* ‘Guiana chestnut’.

The form *kakwa* of Rayón Soke—a dialect of Chiapas Soke, which generally has *kakawa(7)*—seems to reflect the Gulf Sokean pronunciation and may be a borrowing from the direct antecedent of Ayapa Gulf Sokean [ka(:)gwa7], although perhaps not from the precise current location of Ayapa Gulf Sokean in Tabasco.

Table 5 provides forms that have the appearance of being trisyllabic roots but are in fact morphologically complex and, correspondingly, show different stress patterns.

*Cicada*. In Mije-Sokean languages (Table 5A), ‘cicada’ differs from the words in Table 4 in that it is a *Wanderwort*, a word that is widely diffused throughout Mesoamerica, and cannot be confidently attributed to any particular source. Like other such words, it also shows irregular sound correspondences and influence from sound-symbolic factors. In Tecpatan Chiapas Soke (Zavala 2000–2003), for example, one variant for cicada occurs as a verb root for the sound the cicada makes (*ʔiskitin=ʔiskitinapya te7 ʔiskitinh* ‘está cante y cante la chicharra’). The first two forms in Table 5A have analogs in Spanish *chiquirín*, *pichichi*, *pijiji*, and *pijja*, and in at least some other Mesoamerican languages. They are *Wanderwörter*. Even though they may be of Mije-Sokean origin, they can have been directly developed free of symbolic considerations only in some Mije-Sokean languages. All these words except for Texistepec Gulf Sokean /pe(:)7xe:xe7/, to the extent they are found in Gulf Sokean, show initial stress, as well.

*Tree duck and butterfly*. Like ‘cicada’, both ‘tree duck’ and ‘butterfly’ show irregular sound correspondences and influence from sound-symbolic factors. They sometimes show unusual phonology in the language where they are found.

The Soke forms for tree duck can go back to pre-Soke *\*pisisi7*. The Sayula Mijejan form /pi:xix/ is compatible with this if they all go back to proto-Mije-Sokean *\*pi:sisi7*. But the Texistepec Gulf Sokean and Oluta Mijejan (Zavala 1994–2004) forms are not compatible with this or with each other. Oluta Mijejan has /i7/ for *\*i:*; Texistepec Gulf Sokean has /e(:)7/ for *\*i:* and lengthens the second vowel, which would have to have been stressed in the antecedent form to be lengthened. If we postulate a proto-Mije-Sokean form *\*pi:+sisi7*, it would account for all of the forms cited, with the following wrinkles: (1) the proclitic *\*pi:+* was changed to [pi7] in Oluta Mijejan and Texistepec Gulf Sokean; and (2) the proclitic was fused with the root in Sayula Mijejan. This is thus not an originally trisyllabic form but a disyllabic “root” with a proclitic “prefix”.

The Sayula Mijejan form *xu+pe:p* for butterfly goes back to *\*su+pe:pV(7)*. The other forms are compatible with being derived from an antecedent form *\*su+pe:pe7*, with some extra “symbolic” modifications. This, again, is not a trisyllabic form but a disyllabic “root” with a proclitic “prefix”.

*Cockroach, thunder, whippoorwill, and ant*. These words also give the appearance of being trisyllables that show stress on the second syllable in Gulf Sokean but are not relevant to the analysis of Gulf Sokean trisyllabic roots because they are morphologically complex. Their roots are disyllabic. In ‘cockroach’ and ‘thunder’,

Table 5. Morphologically complex forms that resemble trisyllabic roots in some Mije-Sokean languages

	‘cicada’	‘tree duck’	‘butterfly’		
<i>A. Wanderwörter</i>					
Proto–Mije-Sokean	*si:kitiw	*pi: + sisi7	#su + pe:pe7		
Proto-Sokean					
San Miguel Chimalapa Soke		pixixi7	surupenhpenh7 [phonologically irregular]		
Santa María Chimalapa Soke		pisixi7	susepe7		
Copainala Soke	sikitin				
Tecpatan Soke	sikitinh~7iskitinh (~ sikitin)				cf. supupi7 ‘remolino’
Francisco León Soke	7ixktinh				
Rayón Soke					
Ayapa Gulf Sokean	tzyigilé:na	pisix			
Soteapan Gulf Sokean	xi:kyiny cf. xikityiny~xikitiinh ‘sound made by cicada’				
Texistepec Gulf Sokean	xekre:nh [phonologically irregular]	pe(:)7xe:xe7			
Proto-Mijean					
Oluta Mijean	xi:xkiti [phonologically irregular]	pi7xixi			
Sayula Mije(an)	xi:git	pi:xix	xu + pe:p		
Western Mije (Totontepec)	xiktiw				
Eastern Mije (Guichicovi)	xegediw (Wichmann 1995:344)				
For ‘cicada’, cf. Wavi /nchikiw/, which shows final w; cf. Tzeltal /xikirin/, which shows final n					
	‘cockroach’	‘thunder’	‘agouti’	‘ant’	‘nighthawk, nightjar, whippoorwill’
<i>B. Other morphologically complex forms</i>					
Proto-Mije-Sokean	*(ma+) koko7	*ma + jɯ:ɯC		*tzukuC	*pu7 + juyu7
Proto-Sokean			*junjɯɯ	*jaj = tzuku(7)	
San Miguel Chimalapa Soke	mako7		juyɯɯ	jajtzuku7	
Santa María Chimalapa Soke		mɯɯ7	juyɯɯ	jajtzuku	pú7júyu7
Copainala Soke	makoko7		junjɯɯ7	jajchuku7	
Tecpatan Soke	makoko7	(7anh = jɯy.kɯ7)	junjɯɯ7	jajchuku7	pu7yu7
Francisco León Soke				najchuku	
Rayón Soke				jajchuku	
proto-Gulf Sokean					*pukúyyu
Ayapa Gulf Sokean	ko7oko	mɯ:nye	junɯ:ye	jatztzu:ke	[pukúyyu] <sup>a</sup> ‘putput, cutcut’
Proto-Soteapan/Texistepec Gulf Sokean				*jajtzuk	
Soteapan Gulf Sokean		majɯ =	junɯ:yɯ	jajtzuk	pukúyyu
Texistepec Gulf Sokean			judɯ:y	jasuk	puuy
Proto-Mijean				*tzukuC	*pu7 + juyu
Oluta Mijean		mɯ:yɯ[7]w		tzu(7)kuti[7]k	pu7juyu
Sayula Mije(an)				jajtzuk	pu: + juyuy
Western Mije (Totontepec)				tzukɯn	
Eastern Mije (Guichicovi)	mugúk			tzukt	

Notes: The data cited here were collected by linguists working on the PDLMA, except data from Francisco León Soke (from Engel and Allhiser 1987) and Rayón Soke (from Harrison and Harrison 1984).

<sup>a</sup>The form [púgyu7] reported by Wichmann (1995:428) differs from the form collected by Suslak (Suslak et al. 1997–2007), and, among other problems, has two violations of Ayapa Gulf Sokean phonotactic constraints: the first vowel, which is stressed, should be long, and underlying word-final glottal stops should jump to the penult syllable. We therefore reject this form as inauthentic. The Gulf Sokean form \*pukuyyu is not directly cognate with proto-Mije-Sokean \*pu7 + juyu7, though it is obviously related to it or derived from it—by some irregular process. Because of its form, with a medial closed syllable, it cannot instantiate the Gulf Sokean initial stress on words of original shape \*CVCVCV(7).

this is because these words have the proclitic {ma+}, and as a clitic it cannot be stressed. The following root in each of these words is disyllabic, with stress on the first syllable of the root. The following data substantiate this analysis. Soteapan Gulf Sokean /7onhko = nak/~ /ma+ 7onhko = nak/ ‘type of frog/toad’, and Soteapan Gulf Sokean /saawa/ ‘wind’ alongside /ma+ saawa/ ‘windstorm’, show that {ma+} is some kind of optional modifier. Copainala Chiapas Soke (see Pye 1996–1999) /makoko7/ and Ayapa Gulf Sokean [ko:7go], both ‘cockroach’, show that {ma+} is optional in this cognate set. The word \*ma + jɯ:ɯC ‘thunder’ may include a nominalization of the verb root \*jɯ ‘to make a loud noise’, thus \*ma + jɯ:ɯ(7); that {ma+} in this set is a separate element

is anyway suggested by the differently derived Tecpatan Chiapas Soke 7anh = jɯy.kɯ7 ‘trueno’. (The n of the Ayapa Gulf Sokean form /mɯ:nye/ is unexplained, and all forms but that of Soteapan Gulf Sokean show a contraction of ma + jɯ ... to mɯ ... .) The San Miguel Chimalapa Soke form mako7 ‘cockroach’ is contracted from \*ma + koko7 in a so far unexplained way.

A Mije-Sokean word for whippoorwill (Spanish *tapacamino*) can be reconstructed as \*pu7 + juyu7 based on forms from Santa María Chimalapa Soke, Oluta Mijean, and Sayula Mijean. The initial syllable in Sayula Mijean is a proclitic; this property is masked in the Santa María Chimalapa Soke and Oluta Mijean reflexes. An innovated Gulf Sokean proximate form \*pukuyyu can be set up

with this meaning. It, however, has several unique features for a trisyllabic stem: (1) it is stressed on the penult syllable; (2) it loses the medial /k/ in Texistepec Gulf Sokean; (3) the final vowel drops in Ayapa Gulf Sokean; and (4) it is the only apparent trisyllabic noun stem under discussion with a closed medial syllable. Inasmuch as the second/penult syllable is closed in the Gulf Sokean form, that syllable bears the stress.

The case of 'ant' is similar. There are three reconstructible forms: Sokean *\*jajtzuku(7)*; Soteapan Gulf Sokean and Texistepec Gulf Sokean *\*jajtzuk*; and Mijean *\*tzukuC*. The Mijean form shows that *jaj=* in Sokean is a preposed element. Forms like [hah], meaning something like 'fly' or 'grub', are found in Mayan (from *\*ha7 h*). The Ayapa Gulf Sokean form, if not borrowed from Soke, shows the expected Gulf Sokean reflex of a *CVC.CVCV* form. The Soteapan Gulf Sokean and Texistepec Gulf Sokean forms reflect a common antecedent; that antecedent is phonemically like the Sayula Mijean form. Perhaps the simplest hypothesis to account for these unexpected similarities is to postulate that pre-Texistepec Gulf Sokean shortened *\*jajtzuku* [jájitzuku] to [jájitzuk], then both Sayula Mijean and Soteapan Gulf Sokean borrowed the Texistepec Gulf Sokean pronunciation. Later pre-Texistepec Gulf Sokean *\*jajtzuk* was mangled to yield *jasuk*. But *\*jajtzuku* would presumably not have been pronounced [jájitzuku] earlier in pre-Texistepec Gulf Sokean unless no morpheme boundary was perceived between *\*jaj* and *\*tzuku*. Otherwise, it should have patterned like *\*jun=jnyu*, which (unlike *\*kakawa*, *\*sapane*, *\*jukutu*, *\*patajaC*, *\*wakata*) was pronounced [juná:yu], not [jú:nøyu]\*, in proto-Gulf Sokean. Note that Ayapa Gulf Sokean *jatzzu:ke* and *jun:ye* reflect the same pattern, as if the words were compounds with a proclitic first element. The fact that Texistepec Gulf Sokean *jasuk* reflects [jájitzuk] while Texistepec Gulf Sokean *jun:y* reflects [juná:yu] remains anomalous. Since a form like [jájitzuk] is otherwise not known in Mijean, Sayula Mijean *jajtzuk* seems like a loan from Sokean. However, it may be that the Sayula Mijean form is borrowed from proto-Sokean *\*[jájitzuku(7)]*, and that the Soteapan Gulf Sokean and Texistepec Gulf Sokean forms are borrowed from Sayula Mijean.

It is not clear whether the common Sokean word for agouti was a trisyllabic root or not. The shape *CVCCVCV* of *\*junjnyu* may reflect that /jun/ is either a prebound or a proclitic. Even if this word consists of a single trisyllabic morpheme, the shape alone is enough to account for the different stress, since all of the trisyllabic roots with initial stress begin with an open (CV) syllable.

The data and analysis given in this section has shown that, among the six roots that are reconstructible to proto-Sokean and/or proto-Mijean with a *CVCVCV(C)* trisyllabic shape, all five that have Gulf Sokean reflexes had stress on their first syllable in these reflexes. Dakin and Wichmann's claim that Sokean trisyllabic roots had stress on the second syllable—a claim that they support with a single, misanalyzed form *\*ma+koko7* (spelled *\*makoko* by Wichmann) that does not have a trisyllabic root—is simply false. Rather than violating the regular stress patterns of Sokean trisyllabic roots, the data presented here establish that the stress pattern in the Gulf Sokean reflexes of *\*kakawa* agrees with that of every other reconstructible instance. Rather than casting doubt on the Mije-Sokean pedigree of *\*kakawa*, the evidence for stress on other Sokean trisyllabic roots supports the view that proto-Sokean *\*kakawa* was inherited normally in Gulf Sokean. Apart from the Rayón Soke form, which seems to reflect Ayapa Gulf Sokean developments and might be diffused from an Ayapa Gulf Sokean-like language, there is *no* evidence for diffusion of this term within Sokean. The data on Mije-Sokean trisyllabic

roots and their stress patterns are therefore consistent with Campbell and Kaufman's (1976) arguments for a Mije-Sokean origin of this term and add to the body of evidence brought forward in the section "The Mije-Sokean Hypothesis" in support of that conclusion.

#### Pre-Columbian Mijean Forms were Inherited from Proto-Mijean

Wichmann (1995:343–344) argues that the proto-Oaxaca Mije form corresponding to proto-Sokean *\*kakawa* is *\*kakaw*. In fact, all Mijean (not just Oaxaca Mije) forms are consistent with a reconstruction of proto-Mijean *\*kakaw*, except for one highland Mije form, Mixistlán [kaká:wa], which derives from regional Spanish *cacahua*. The highland and lowland Mije forms of approximate shape [kaká:w] cited by Wichmann (1995:343–344) are not Spanish borrowings, as he states, but follow regular Mije developments from *\*kakaw*. Wichmann's evidence for borrowing from Spanish is that the final vowel in these forms is long and stressed. In Totontepec Highland Mije, it is true that descendants of proto-Mijean *\*CVCVC* forms received initial stress, as in *káku*, which Dakin and Wichmann recognize as descending from proto-Mijean *\*kakaw*. However, in other forms of Mije, *\*CVCVC* forms receive final stress. In all Mije, stressed final syllables and monosyllabic stressable words insert [h] before the final *C*, and this inserted [h] goes to vowel length before resonants. Thus *\*kakaw* ⇒ *kakáw* ⇒ *kakáhw* ⇒ *kaká:w*. Dakin and Wichmann (2000:57a) go further than Wichmann (1995) in erroneously suggesting that the [kaká:w] forms do not develop from the same form as Totontepec Highland Mije /kaku/. Oluta Mijean *kakaw* [kaka7w] shows that this form also descends from one with a final consonant.

Wichmann (1995:343–344) cites the following data (language labels follow his terminology):

#### North Highland Mije

Totontepec Highland Mije [káku]

#### South Highland Mije

Tlahuitolpetec [kakó:w]

#### Midland Mije

Mixistlan [kaká:wa] {< Sp}

Juquila, Jaltepec, Puxmecatan [kəgá:]

Matamoros [kəgá:W]

Atitlan [kagá:w]

#### Lowland Mije

Coatlan [kə7əga:]

Camotlan and Guichicovi Lowland Mije [kəgá:]

#### Sayula Mijean kágaw

#### Oluta Mijean kaka7w

The Mijean form *\*kakaw* has nothing to do with the fact that in Mije proper, in Sayula, and in Tapachula word-final proto-Mije-Sokean *\*V* and *\*V7* are dropped in polysyllables. (The development in Sayula and Mije may be a shared change; that in Tapachula is not.) When these Mijean languages drop a final vowel, the phonological reflexes are different from those in

words that never had a final vowel; it is shown here that the surviving Mije forms are consistent with the reconstruction of proto-Mijeian *\*kakaw*. Oluta does not drop word-final vowels, and the phonetic [ʔ] before the final consonant of Oluta Mijeian *kakaw* [kakaʔw] is regular for Oluta Mijeian roots of two or more syllables descending from proto-Mijeian and proto-Mije-Sokean consonant-final forms but not for those descending from roots ending in *V* or *V7*.

Accordingly, except for one late loan from Spanish, the Mijeian forms are consistent with direct inheritance from proto-Mijeian *\*kakaw*. They supply no evidence for any pre-Columbian borrowing that postdates the breakup of proto-Mijeian.

#### The Discrepancy between Proto-Sokean *\*kakawa* and Proto-Mijeian *\*kakaw*, and Its Source

The only remaining puzzle, then, is that proto-Mijeian *\*kakaw* and proto-Sokean *\*kakawa* are not identical. Note that under the model we propose—that *\*kakaw* ~ *\*kakawa* is native to Mije-Sokean—borrowing between Mijeian and Sokean does not account for this difference. Had Mijeian *\*kakaw* been borrowed into Sokean, it would be expected to have retained its pronunciation in Sokean, which has a large number of *\*CVCVC* noun roots—far more of them than the number of its *\*CVCVCV* roots. So such a borrowing is implausible. Similarly, although there were few *\*CVCVCV* roots in proto-Mijeian, it is unlikely that an ancestor of proto-Mijeian would have reduced a trisyllabic noun root *\*kakawa* to two syllables on being borrowed into a form of Mijeian predating proto-Mijeian. Not only are there at least two solid proto-Mijeian *\*CVCVCV* roots—*\*wakata* ‘Guiana chestnut’ and *\*ʔi:ʔzʔmʔ* ‘peccary’—but *CVCVCV* is within the range of phonotactic shapes of complex words in proto-Mijeian.

The remainder of this section explores a viable alternative to borrowing, for which there are numerous parallels: that the differences reflect an ancient heritage of this word within the Mije-Sokean family that dates back to proto-Mije-Sokean.

Very often for a given meaning, Mijeian and Sokean have completely different morphemes or combinations of morphemes. Sometimes, though, the forms are clearly related phonologically, but show discrepancies that forestall the reconstruction of a single phonological form to proto-Mije-Sokean. Examples currently known to us are provided in Table 6. Note again that borrowing between an ancestor of proto-Mijeian and an ancestor of proto-Sokean, in either direction, cannot account straightforwardly for any of these differences.

*Long and short forms of the same root.* Both between branches and within the same language, a single root may occur with both a shorter and a longer form, with the longer form having an extra vowel (plus or minus glottal stop) at the end (cf. Wichmann 1995:80–88). There are more cases than the ones cited here.

Proto-Sokean *\*kʔ(7)*=*tzʊs* ‘digit nail’ [under corner]

Copainala Chiapas Soke *maks*=*chʊs tʔa7* ‘flint’ [four-cornered stone]

Santa María Chimalapa Oaxaca Soke *tzʊsu* ‘corner’

*\*jʔp* ‘nose’ (Mi); ‘mouth’ (So)

Copainala Chiapas Soke, Tecpatan Chiapas Soke *jʔpʊ* ‘jaw, chin’

Table 6. Non-identical but phonologically related reconstructions for Sokean and Mijeian

Gloss	Sokean	Mijeian
A. Mijeian is longer (12 ×)		
bottlegourd	pok	pokok
grass	so7k	sokot
pillar	kom	komom
frog	najk	nakak
meat	sis	sisi
snake	tzajin	tzan7ay
man	jaya	jay7aw ‘person’
person	pʊn	pʊnA ‘who?’
shit <sup>a</sup>	tin	tʔ:n7.i
sweet potato	mʊn	mʊni
to kindle	no7	no7k
road	tuw	tu7aw
B. Sokean is longer (6 ×)		
cacao	kakawa	kakaw
eye	witʔm	win
sand	po7oy	pu7ʔ
ant	jajtzuku[7]	tzuku(n)
coral bean	tzentzen	tzejtzE
C. No difference in length but parallel segments differ (18 ×)		
inside	joj	jo:t
head	kopak	ko7+ pak
baby	7une7	7unak
elder (brother)	7atzi7	7ajtzi
to measure	ki7ps	kips
to make cord	tʔ7ps	tʔps
anona (fruit)	yati7	7a:ti
scorpion	kakwe7(n)	kaHpe(n)
toucan	katzi(7)	ka:tʔ
to recognize	7is=pʔk	7is=kap
to breathe	jej	sej
turtle	tuki	tuka
good	7oye	7oya
to change	kak	kek
cojolate (bird)	wʔku7	wʔ:ki
to sleep	mow	ma:j7
to suckle	tzu7tz	tzi7tz
elder sister, aunt	tʔʔzʔ	tʔʔ7ʔ
iguana	tʔʔʔ7	tʔʔʔj

<sup>a</sup>*\*tin* is the proto-Mije-Sokean word for ‘shit’, as can be seen in the proto-Mijeian *\*tintʔay* ‘gut’ (i.e. “shit-vine”); proto-Mijeian *tʔ:n7.i* ‘shit’ is a nominalization of proto-Mije-Sokean *\*tʔ:n7* ‘to shit’.

Proto-Mijeian *\*kakaw* ‘cacao’

Proto-Sokean *\*kakawa* ‘cacao’

Whenever there is evidence, the long forms are seen to be derived from the short forms. The short forms are not truncated.

Consider the following set of forms, apparently based on *\*pok*:

Proto-Mije-Sokean *\*pokʔi7* ‘ankle’

Santa María Chimalapa Soke *poki7* ‘ankle’

Santa María Chimalapa Soke *poki7* ‘ankle’

Tecpatan Chiapas Soke *poki7* ‘ankle’

- Oluta Mijeán *po7ki* ‘ankle’
- Proto-Sokean *\*poj̥k* ‘bottle gourd’  
 Tecpatan Chiapas Soke *pok* ‘bottle gourd’  
 San Miguel Chimalapa Soke *pojok* ‘gourd container for seed (for planting)’  
 Ayapa Gulf Sokean *pok* ‘gourd bowl’  
 Soteapan Gulf Sokean *pok* ‘bottle gourd’
- Proto-Mijeán *\*pokok* ‘bottle gourd’
- Proto-Sokean *\*po7k* ‘knot’  
 Tecpatan Chiapas Soke *po7k* ‘knot (in a rope, on a tree, on your head)’  
 Santa María Chimalapa Soke *po7k* ‘knot (in tree or rope)’  
 San Miguel Chimalapa Soke *po7k* ‘knot (in a rope or tree), lump’  
 Soteapan Gulf Sokean *po7k* ‘trunk (of tree)’
- Soke *\*pok.pok* ‘round’  
 Santa María Chimalapa Soke *pok.pok* ‘circular’ (like a plate or the rim of a bucket)  
 San Miguel Chimalapa Soke *pok.pok* ‘puffed up’  
 Tecpatan Chiapas Soke *pok.pok* ‘calf of leg’  
 cf. Soteapan Gulf Sokean *pok.pok* ‘oriole’ [probably not connected]
- Western Soke *\*po7ojk*  
 Santa María Chimalapa Soke *po7ok* ‘egg’  
 San Miguel Chimalapa Soke *pojo7k* ‘egg; ballock’
- Tecpatan Chiapas Soke *pok.a7* ‘egg’

In the case of the word for egg, there is no problem with the fact that Chimalapa Soke has a final consonant and Copainala and Tecpatan Chiapas Soke has a final V7. The problem is why there is extra (laryngeal) material in the middle of the Western Soke forms. Compare also

- Tecpatan Chiapas Soke *pok.o7* ‘elephant-ear tuber’  
 Ayapa Gulf Sokean *po[ɣ]ok* ‘to roll up; to roll’  
 Soteapan Gulf Sokean *pook* ‘cornstalk’  
 Soteapan Gulf Sokean *pookon* ‘reed’

All of these suggest that they come from a root *\*pok* that meant something round or spherical or cylindrical.

These examples are a few among many illustrating that individual Mije-Sokean languages use a variety of strategies for creating new lexical items by extending an existing stem in some way. This is illustrated in the cases discussed here—Copainala and Tecpatan Chiapas Soke *ju7pɨ* ‘jaw, chin’ <proto-Sokean *\*ju7p* ‘mouth’ <proto-Mije-Sokean ‘nose’; Santa María Chimalapa Soke *tzusu* ‘corner’ <proto-Sokean *\*tzus*. In each case, the longer form is an expansion of the shorter form; in none of the cases known to us is there evidence for truncation of an originally longer form. Like any natural language, proto-Mije-Sokean would have had some lexical items with more than one pronunciation, distributed by geography, social class, or style; some would have been produced as extensions of existing forms. In addition, the Mijeán

and the Sokean branch must each have taken some inherited words and extended them to form new words or new variants of old words. Many of the forms presented in Table 6 could have arisen, and likely did arise, in this way.

It is therefore consistent with what we know of variation in Mije-Sokean generally to hypothesize that Sokean *\*kakawa* is an expanded version of an original proto-Mije-Sokean *\*kakaw*, or that *\*kakaw* and *kakawa* co-existed in proto-Mije-Sokean and that in each branch only one variant survived. (It is not plausible that Mijeán *\*kakaw* is truncated.) This being the case, nothing in the Mije-Sokean data requires us to conclude that a word for cacao was borrowed into any Mije-Sokean language in pre-Columbian times.

This kind of variation raises an issue regarding comparative reconstruction within Mije-Sokean that is incorrectly handled by Wichmann (1995): only in the rarest cases can there be any doubt about where to reconstruct medial and final vowels in Mijeán, Sokean, or Mije-Sokean etymologies. Even if no cognate for a given Mijeán etymology is found in Oluta, so that all Mijeán cognates of a proto-Mijeán form that ended in V or V7 may end in consonants, there is almost always evidence that a vowel had been there. It is also possible to determine whether the vowel was a front vowel (*i* or *e*), or some other vowel (*ɥ*, *a*, *u*, or *o*): if it was a front vowel Mije Proper (but not Sayula) palatalizes the consonant that had preceded that vowel, or raises (“umlauts”) the vowel of the original penult syllable; if it was a non-front vowel, the preceding consonant is not palatalized, and there is no “umlaut”. Some languages, such as Soteapan Gulf Sokean and San Miguel Chimalapa Soke, and probably some forms of Mije, drop some unstressed medial vowels, but Sayula Mijeán and Oluta Mijeán do not. When San Miguel Chimalapa Soke drops a medial vowel, this leaves a trace in the lengthening of the preceding vowel; when Soteapan Gulf Sokean drops a medial vowel, the preceding vowel is often but not always lengthened.

Word-final vowels dropped also in Texistepec Gulf Sokean. This change affects only word-final vowels; word-final /V7/ is not affected, and word-final proto-Mije-Sokean *\*7* is preserved in Texistepec. The vowel-drop in Texistepec may have diffused from the neighboring town of Sayula. Before about 1,000 years ago, there was no regular phonological process of dropping final vowels in any Mije-Sokean language.

The relevance of this in the current context is that any item that can be reconstructed phonologically is going to point clearly to there having been a final vowel, or to having not ended in a vowel. Often enough to be interesting, it turns out that Sokean has final V and Mijeán does not, or vice versa, or even between one language and another that clearly preserve final V, one has it and one does not.

*Morphological analyzability.* If *\*kakaw* or *\*kakawa* were analyzable as being composed of more than one meaningful unit in one language family but not in another, that would contribute to evidence for the origin of the word in the family in which it is analyzable. However, there is no known possible morphological analysis of *\*kakaw(a)* within Mije-Sokean. The only seemingly obvious possibility would be partial reduplication, with *ka* preposed to a base form *kaw* or *kawa*, but such a hypothesis must be rejected because in Mije-Sokean languages it is the root that comes first and its partial (or complete) replication that follows. The first syllable is not a reduplication of the second.

In Mije-Sokean, generally speaking, -VC and -CVC reduplication is done with CVC roots in the formation of derived verbs.

Complete reduplication of *CVC*, *CVCV*, and *CVCVC* shapes is found in derived nouns and adjectives.

Verbal reduplication of a hypothetical root *kaw* would yield *kaw-aw* or *kaw-kaw*, rather than, for example, the Mijejan form *\*kawkaw*. Reduplication of a hypothetical nonverb root *kawa* would yield *kawa-kawa*, rather than, for example, the Sokean form *\*kakawa*. No known reduplication process in any Mije-Sokean language could produce *kakaw* or *kakawa* from any hypothetical base form.

We therefore find no obvious way to analyze *\*kawkaw(a)* morphologically in Mije-Sokean. Morphemes of the shape *CVCVC* are fairly common in proto-Mije-Sokean, proto-Sokean, and proto-Mijejan, and proto-Sokean has at least five simple roots, excluding *\*kakawa*, that are shaped *CVCVCV(C)* (see Table 4) and for which no morphological analysis seems possible. The lack of analyzability of *\*kawkaw(a)* therefore does not constitute evidence against the Mije-Sokean hypothesis. The occurrence of two *ka* sequences does not in itself require an interpretation of reduplication as a process. In a language with 66 different CV syllable shapes, a non-negligible proportion of words having two or more syllables will begin with two identical CV sequences, as in the case of proto-Sokean *\*koko7* ‘cockroach’, cited earlier, in addition to proto-Mije-Sokean *\*mumu* ‘all’, *\*nu:nu7* (> proto-Sokean *\*nunnu7* ‘atole’, proto-Mijejan *\*nu:nu* ‘tortilla’), and *\*tujtu* ‘beyond/plus five’; proto-Sokean *\*toto7* ‘*Ficus* sp., amate fig’; proto-Soke *\*tatzu* ‘younger sister’; proto-Gulf-Sokean *\*meme* ‘butterfly’, *\*na7na7* ‘gum (tree)’, *\*nono* ‘mushroom; tree ear’, *\*nunu* ‘(woman’s) breast’; proto-Mijejan *\*se:se* ‘(small) fish’, *\*sisi* ‘meat’, *\*totok* ‘butterfly’. The issue would have been relevant only if *\*kawkaw* or *\*kakawa* were analyzable in a language that, on other grounds, had proven to be a *viable* candidate for the language from which *\*kawkaw(a)* spread to other Mesoamerican languages.

#### Refuting the Arguments for a Nawa Source of *\*kakawa* in Mesoamerican Languages

*Dakin and Wichmann’s Nawa etymology is invalid.* The Nawa and Yuta-Nawan data cited in this section are from Kaufman’s field notes (Kaufman 1981, 2001).

Forms containing the noun stem /*kakawa*/ are the following in Huasteca Nawa (other forms of Nawa contain similar items):

kakawa-tl ‘cacao’; ‘peanut’

i-kakawa-h ‘his cacao’; ‘his peanut’

i-kakawa-yo ‘its thick bark’

kakawa.ti.k ‘hollow’

tla:l=kakawa-tl ‘peanut’ (literally, ‘earth-cacao’)

Dakin and Wichmann (2000) hypothesize that an early Yuta-Nawan word meaning ‘egg’ is the source of Nawa /*kakawa*-tl/; that [k*akawa*] arose as a CV- reduplication of a pre-Nawa *\*kawa* ‘egg’; and that [k*akawa*] would have originally meant ‘egg-like thing’. The semantics of this hypothesis are not implausible, given the shape of the cacao pod, although in Mesoamerica the pod is analogized rather to an ear of maize in its husk (e.g., Molina’s <cacahuacentli> ‘maçorca de cacao’). Dakin and Wichmann acknowledge that there is no attested Nawa word *kawa(-tl)\** and thus no internal Nawa evidence for the

analysis, which depends entirely on the plausibility of relating Nawa [k*akawa*] to Sonoran forms. (Southern Yuta-Nawan is Sonoran plus Nawa.)

The Sonoran data they cite (Dakin and Wichmann 2000:59) are:

Warijiyo *ka7wá* ‘egg’

Taraumara *ka7wá* ‘to lay eggs’

Kájita *kava* ‘egg’

Eudeve *aa]kabo[ra’a* ‘egg’

These data are sufficient to reconstruct proto-Sonoran *\*kava*. Contrary to Dakin and Wichmann’s assumption, however, a Southern Yuta-Nawan form *\*kava* (which would yield Sonoran *\*kava*) cannot be the source of Nawa *kakawa-tl*. The reason is that their claim that postvocalic Yuta-Nawan single *\*p* becomes *w* in Nawa is wrong. Instead, postvocalic Yuta-Nawan single *\*p* (like initial *\*p*) changes to [v] in Southern Yuta-Nawan, which shifts to [h] in Koran and Nawa, and this is subsequently lost in Nawa.

YN *\*səpɸ.* ‘cold’ [n] ⇒ PSYN *\*seve-ta* ⇒ Nawa *se:-tl*  
‘cold’ [a] PSYN *\*se-seve-ka* [a] ⇒ Nawa *se-se:-k*

YN *\*napo-tɬ* ‘prickly pear’ ⇒ PSYN *\*navo-tɬi-ta* ⇒ Nawa *no:ch-tli*

YN *\*tapun-tɬi* ‘rabbit’ ⇒ PSYN *\*tavu-tɬi-ta* ⇒ Nawa *to:ch-tli*

YN *\*tɬi:pɸH* ‘bitter’ ⇒ PSYN *\*tɬi-tɬi:vu-ka* ⇒ Nawa *chi-chi:-k*

YN *\*pi:pah* ‘tobacco’ ⇒ PSYN *\*vi:va-ta* ⇒ Nawa *i(:)ya-tl*

These data show that postvocalic Yuta-Nawan *\*p* = Son *\*[v]* does *not* survive as /*w*/ in Nawa but disappears, and the resulting vowel sequence merges into a single long vowel (except that *\*iva* > *\*iha* > *iya*, perhaps not passing through the stage *\*ia*). Thus, although there is a Sonoran etymon *\*kava* ‘egg’, a putative Southern Yuta-Nawan form *\*kava* can not yield [k*awa*] in Nawa, and a hypothetical reduplicated noun deriving from a proto-Southern Yuta-Nawan *\*kava* would have shown up in Nawa as *kaka:-tl\**, not as *kakawa-tl*. Dakin and Wichmann’s proposed Yuta-Nawan origin for the Nawa word *kakawa* is simply not possible.

There are two etymologies that might give the false impression that Sonoran medial [v] survives as [w] in Nawa. These are instances in which Sonoran [v] follows a rounded vowel. As usual, Sonoran medial [v] shifted to [h] in Koran and Nawa; subsequently in Nawa, after a rounded vowel this [h] became [w], after which the rounded vowel desyllabified and disappeared (this may have happened especially after /*k*/):

YN *\*kopa* ‘forehead’ ⇒ pSYN *\*kova* ⇒ Nawa *kwa:(yi)* ‘head’  
pSYN *\*ma:kova* ‘five’ ⇒ Nawa *ma:kw:i-li*

Unlike postvocalic Yuta-Nawan *\*p*, which disappeared in Nawa (except as noted earlier), Yuta-Nawan (and Southern Yuta-Nawan) medial *\*w* did survive, as *w*:

YN *\*konwa* ‘snake’ ⇒ Nawa *kowa:-tl*

YN *\*sunwa* ‘woman’ ⇒ Nawa *siwa:-tl~sowa:-tl*

YN *\*təwa* ‘to see’ ⇒ Nawa *itwa~itta~ita*

pSYN *\*ku7awi* ‘tree’ ⇒ Nawa *kwawi-tl*

In sum, Nawa *kakawa-tl* cannot possibly be derived from an ancestor of proto-Sonoran *\*kava* ‘egg’, as Dakin and Wichmann claim. This leaves no evidence either in Nawa or more broadly in Yuta-Nawan for a Nawa origin of Mesoamerican words for cacao.

*Evidence that \*kakaw(a) was not borrowed from Nawa into early Mesoamerican languages.* The previous section shows that the word *\*kakaw(a)* did not originate in Nawa. This section shows that, once it was adopted by Nawas, it did not pass from them to speakers of any other Mesoamerican language at an early date. The evidence comes from the lack of any trace of the Nawa absolute suffix *-tl* in the borrowed forms. The section “General Features of Borrowings into and out of Nawa” shows in some detail that Nawa words are almost always borrowed in their absolute form. Dakin and Wichmann (2000:67b) acknowledge that this is an issue: “*kakawa-tl* is always borrowed without the so-called absolutive [*sic*] suffix”. They further acknowledge that it would be a “serious” difficulty, were they unable to provide a viable rationale for the absence of this suffix in the borrowed forms. They seek to overcome this problem by arguing that it could have been borrowed in a possessed form, which would lack the absolute suffix. Their rationale for borrowing the word in a possessed form is a supposition that, in the diffusion of cacao as a commodity, it “would have been an object more likely to have been discussed in possessed form”.

This claim is pure speculation, for which Dakin and Wichmann provide no evidence. Direct linguistic evidence would be a demonstration that other borrowed Nawa words for commodities regularly show up without absolute suffixes. In fact, other Nawa words for commodities are borrowed in the absolute, and not in the possessed, form. For example, mirrors were made from materials that were not found everywhere and were traded to areas that lacked them. Nawa *te:ska-tl* ‘mirror’ is borrowed as Soteapan Gulf Sokean *teeskat*.

Dakin and Wichmann’s imaginative scenario to account for the borrowing of the word for cacao in an unexpected form is a “just-so” story; and they do not simply propose that one instance of borrowing of the word into some language showed this peculiarity, but that it was “always” borrowed in this form. According to their proposal, then, every language that borrowed the word for cacao from Nawa would have to have borrowed it from the possessed form of the Nawa word. Without serious evidence to support this idea, the uncharacteristic lack of a final *t* in the postulated repeated borrowings of the word for cacao remains a serious inconsistency of the Nawa hypothesis with the data on the borrowing of this word.

Another problem with this particular speculation is that, although Dakin and Wichmann are not entirely explicit on the point, they clearly mean to contrast the presence of the suffix *-tl* in the absolute form of *kakawa* with the absence of any suffix in its possessed form (“Once a noun is possessed in Nahuatl—always by means of a prefix—the absolutive [i.e., absolute] suffix is dropped” [Dakin and Wichmann 2000:68]). But it is not correct that the possessed form of *kakawa* is unsuffixed. As Canger (personal communication, 2006) points out, the possessed forms—*i:-kakawa-w* ‘his cacao’, *no-kakawa-w* ‘my cacao’, *mo-kakawa-w* ‘your cacao’, *to-kakawa-w* ‘our cacao’, *a:n-kakawa-w* ‘y’alls cacao’, *i:n-kakawa-w* ‘their cacao’—actually end in a consonant, *w*, which is a suffix marking some nouns as being in the possessed state. Although we might imagine that a final *w* is phonetically easier to eliminate than a final *t* or *tl*, the fact is that there is no form of the Nawa word for cacao that ends in a vowel.

Also unsubstantiated are the nonlinguistic presuppositions of Dakin and Wichmann’s scenario that “cacao was a trade item and the word must have been diffused along trade routes in situations of trade negotiation” (Dakin and Wichmann 2000:68). It is not known whether cacao was in fact traded as a commodity at the time its (earlier) borrowing is first documented in Mayan hieroglyphic texts. There are many alternatives to a trading explanation for the earliest diffusion of cultivated cacao. Most generally, the processes that fostered enough of an increase in the demand for cacao—for example, an increase in the use of the beverage made from cacao in elite interactions—would need to have preceded any substantial trade in cacao beans, if any such trade developed as early as Dakin and Wichmann’s hypothesis requires. The spread of the word could as easily have accompanied these processes of diffusion rather than the subsequent trade, if any. Or it could have accompanied the spread of the practice of cacao cultivation rather than the distribution—trade-based or otherwise—of the products of cacao cultivation.

#### SHORTER FORMS DIFFUSED IN MESOAMERICA

[*kakaw*] outside Mijejan

In Mayan, /*kakaw*/ is the typical form. It is borrowed, not native, because a proto-Mayan *\*kakaw* would not have preserved [k] in all languages. Mayan tolerates disyllabic noun roots (though they are relatively few) but does not have native trisyllabic roots nor roots ending in vowels. The model for the diffused forms could have been either [*kakaw*] or [*kakawa*]. The borrowing reached Greater Tzeltalan after the change *\*k > ch* had run its course, so sometime after 200 B.C. Mayan has a reconstructible root *\*pe:q*, now referring to uncultivated cacao (Spanish *pataxte*); that Greater Lowland /*pi:k*/ expresses the numerical value of 8,000 (and proto-Yukatekan *\*pi:k* ‘skirt’; cf. “The Mije-Sokean Hypothesis”, above) suggests that the Mayan term *\*pe:q* also referred to cultivated cacao at one time.

Totonako, Tepewa, Salvador Lenka, and Paya all have forms like /*kakaw*/. The Paya and Salvador Lenka forms plausibly have spread from Mayan, but possibly from Mijejan, or even from pre-Sokean if proto-Sokean *\*kakawa* developed after the proto-Mije-Sokean stage. The Totonako and Tepewa forms more likely spread from Mije-Sokean.

Totonako /*kakaw*/ ‘peanut’ (Kaufman field notes, 2003)

Tepewa /*kakaw*/ ‘cacao’ (Susan Smythe, personal communication 2005)

Salvador Lenka [*k’á:gaw*] = /*k’akaw*/ ‘cacao’ (Lehmann 1920; the anomalous initial *k’* may well be an error on Lehmann’s part)

Paya [*kaku*] ‘cacao’ (Holt 1999a)

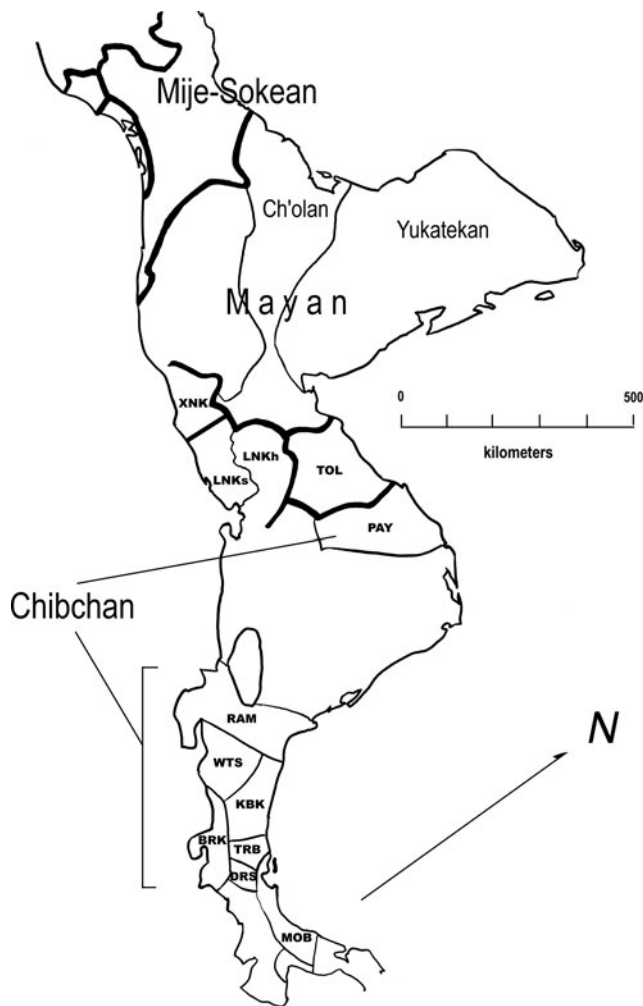
#### Forms like *kaw(a)*

Most of the following forms, all from Lower Central America, resemble the Mije-Sokean *\*kakaw(a)*; some seem to reflect a model [*kaw*], which is not known in Mesoamerica proper. These forms are listed in the geographical order in which they occur (Figure 5):

West to East in El Salvador and Honduras

Mayan *kakaw*

Tol [*khaw*] (Holt 1999b)



**Figure 5.** Languages of lower Central America that have words for cacao derived, ultimately, from Mije-Sokean, showing their locations as of about A.D. 1500. To provide a sense of the linguistic geography when the word for cacao was diffusing, intrusive groups that reached their contact-period locations after A.D. 500 have been removed, and their territory has been divided among remaining adjacent groups. We have no data on words for cacao from Misumalpan languages, which divide Paya from the rest of the Chibchan family. *Chibchan*: BRK = Boruka; DRS = Doraske; MOB = Mobe; PAY = Paya; RAM = Rama; TRB = Terraba; WTS = Watuso. *Other*: LNKH = Honduras Lenka; LNKs = Salvador Lenka; TOL = Tol; XNK = Xinka. After Kaufman and Justeson (2006:Figure 6.2).

Honduras Lenka \*[kayaw] > /kaw/ (Lehmann 1920)

Pipil *kakawa-t* (Campbell 1985; inherited from proto-Nawa, which had borrowed *\*kakawa* from Mije-Sokean by about A.D. 500)

Salvador Lenka /k'akaw/ (or /kakaw/) (Lehmann 1920)

North to South on the Pacific Coast

(Chiapaneko *nuusi* [Fernández de Miranda and Weitlaner 1961:53])

(Chorotega *nyuusi* [Fernández de Miranda and Weitlaner 1961:53])

(Sutiaba *uusi*) (Lehmann 1920)

North to South on the Atlantic Coast (Paya and Rama are Chibchan)

Paya *kaku* (Holt 1999a)

[Sumu-Ulwa] (forms not found)

[Mískitu] (forms not found)

Rama *kuk* (Colette Grinevald, personal communication circa 2000)

West to East in Costa Rica (all of these are Chibchan languages)

Proto-Watuso-Boruka *\*kshú7* (Constenla 1981:373)

Watuso *kaju*: (Constenla 1981:373)

[Kabékar tsirú] (Dakin and Wichmann 2000:74)

Térraba *kó* (Dakin and Wichmann 2000:75)

[Bribri sirú] (Dakin and Wichmann 2000:74)

Boruka *kaw7* (Constenla 1981:373)

Doraske *koa* (Dakin and Wichmann 2000:74)

Waymí *ku* (Dakin and Wichmann 2000:74)

Besides forms like [kaw], Chorotega and Sutiaba, whose speakers probably invaded the region circa A.D. 800-900 (Salgado 1996: 303; Steinbrenner 2006:257) and circa A.D. 1200, respectively, share a form like [(ny)uusi]. Tlapanekan (= Tlapaneko + Sutiaba) and Chorotegan form a node on the Oto-Mangean family tree, but since no form for cacao has been identified in Tlapaneko, the Sutiaba form could be a borrowing from Chorotega. Given the Chiapaneko form, *\*nuusi7* can be reconstructed for proto-Chorotegan and rolled back to central Mexico, where Chorotegan originated.

Kabékar and Bribri share a form like [(t)sirú]. Its further connections are unknown to us.

Salvador Lenka, Paya, and Rama reflect the typical Mayan pronunciation [kakaw], and this is plausibly their immediate source.

Going back to forms like [kaw], the following scenario may be envisioned:

First of all, the Chibchan data are limited geographically to only some of the Chibchan languages in Central America, and they are not found in any of the Chibchan languages of South America. The distribution does not correspond to a genetic subgrouping within Chibchan. Constenla's reconstruction *\*kshú7* may more properly be treated as a formula subsuming the phonological regularities between the Watuso and the Boruka forms *as if* they were cognate, even though they are due to diffusion. The rest of the Chibchan forms cited represent most of the branches of the stock, but there is no "cacao"-like form in any South American language that is not the result of colonial-period diffusion. Except for the Rama form, which cannot be fully accounted for, we can postulate that the phonological antecedent for all these forms is something like [kahaw]. Central American Chibchan languages show the development [káhaw] (with first syllable "stress") > [*\*kshú7*] > *kaju*:, *káw7*, *kó*, *ku*, *kuk*. Tol may have developed [kaháw] (with second syllable "stress") to [khaw] or borrowed its word from one of the Chibchan languages with a (possibly intermediate) Boruka-like form [káw7]. Honduras Lenka may have done likewise.

The postulated Central American antecedent [kahaw] would have been borrowed from the general Mayan form /kakaw/. If the first intermediary into Central America from the Mayan area had been Honduras Lenka, there would be an explanation of the shift of medial /k/ to [h], because in Honduras Lenka, single intervocalic /k/ is pronounced [ɣ], which does not sound like [k] in a language having only [k] and [h] but not [ɣ] or [g]. If Honduras Lenka is the intermediary, why it should have simplified [kayaw]



to [kaw] is not crystal clear, but it may be observed that Spanish *vacas* yielded Honduras Lenka /waš/ ‘cattle’ through an intermediate form [wayaš], so this is a plausible internal development in Honduras Lenka.

Salvador Lenka /k’akaw/ (maybe /kakaw/) and Paya /kaku/ could be direct borrowings from Mayan with no Honduras Lenka intermediary, but Mayan languages are quite far away, and there are viable alternative explanations. The [g] of Salvador Lenka [k’á:gaw] could reflect the [ɣ] of an earlier Honduras Lenka [kayaw] rather than Mayan [k]. Similarly, since Paya borrows proto-Mije-Sokean \*pa:=ju7 ‘coyote’ as /paku/, it is more likely that it borrowed an antecedent Central American form [kahaw] as /kaku/ than that it made a far more distant borrowing from Mayan.

Altogether, the evidence from Central American languages does not clearly support an antecedent form like [kaw]; in fact it more strongly suggests an antecedent form like [kahaw], borrowed from Mayan /kakaw/, maybe specifically via Honduras Lenka.

There are two Oto-Mangean words that resemble [kakawa] or [kawa], both found in Oaxaca: (1) Proto-Chinanteko has \*kwá:7 ‘case; peeling; pod; shell’ (Rensch 1989:50, no. 163); cf. also \*kwé:7 ‘bark, peeling’ (Rensch 1989:50, no. 160). In no Chinanteko language does a form like [kwa] or [kwe] actually mean ‘cacao’. (2) Amusgo has /teh šuah/ (literally, “bean cacao”) (tone pattern is low, mid) ‘cocoa bean’, plural /teh nguah/ (Tapia 1999:216). There is a class of nouns in Amusgo that take the prefix {tz-} in the singular and {n-} in the plural; given this, and the fact that the underlying sequence //tz-k// is realized as š in Amusgo, the underlying form of the word for cacao can be seen to be //-kuah//, with the singular //tz-kuah// realized as /šuah/, and the plural //n-kuah// realized as /nguah/. Both the Chinanteko and the Amusgo forms can derive from an antecedent [kVwa]. The identity of the first vowel cannot be determined.

Dakin and Wichmann (2000:74) cite Chocho /ka:kau7/ (from Mock (1977), which could be borrowed from Spanish. No other Oto-Mangean words for cacao reflect [kakawa] or [kawa].

The Chinanteko and Amusgo forms seem as though they could reflect specifically [kawá]. Both of these forms lack the initial syllable [ka] of \*kakawa, and an antecedent \*kawa could yield both of them. The [kawá] form has stress or prominence on the second vowel. This may reflect habitual accentual patterns: in Oto-Mangean languages (apart from the Mije-Sokean-influenced Sapoteko, Misteko, and Kwikateko), polysyllabic words have highest prominence on the last syllable. In the earliest stages of Oto-Mangean languages, lexical stems had one or two syllables; any lexical material in antepenult position is a clitic or a classifier. Hence, if a form like proto-Sokean \*kakawa were taken into an early stage of an Oto-Mangean language (at least 2,000 years ago), something would have to be done with the first /ka/. If it did not correspond in a meaningful way with an existing proclitic or classifier in the target language, it might be eliminated. We suggest that this is indeed what happened to \*kakawa in Amusgo and what might have happened in Chinanteko if proto-Chinanteko \*kwá:7 ~ \*kwé:7 (dating before about 1,500 years ago) is a borrowed word (pre-proto-Chinanteko underwent sound changes whereby antecedent CVCV forms were reduced to CCV, and the initial cluster was subject to simplification in certain cases).

The kaw(a) forms discussed in this section might suggest that there was originally an “unreduplicated” form meaning ‘cacao’ drifting around that was reduplicated to produce the form \*kakawa. The discussion here shows that this would be an

unnecessary assumption and that these forms provide no viable evidence that \*kakaw(a) was a reduplicated form.

## THE NAMES FOR THE DRINK CHOCOLATE

In Mesoamerica, there is a related set of words for chocolate, the drink made from ground cacao kernels mixed with water and seasonings, that come from four sources: Nawa *chokola:tl*, its borrowing into Spanish as *chocolate*, Nawa *chikola:tl*, and its borrowing into regional/substandard Spanish as *chicolate*. The Nawa form is made up of a first element of uncertain origin, “chokol” or “chikol”, plus {a:} ‘water’.

### Evidence Concerning the History of Nawa *chikola:tl*~*chokola:tl*

All suggestions so far offered for the origin of the first element are unsatisfactory. Nawa *ch* in native words should occur only before *i* (Campbell and Langacker 1978). If the earliest Nawa form was *chikola:tl*, the form *chokola:tl* could have developed from it by assimilation of the first vowel to the second. If the earliest form was *chokola:tl*, {chokol=} is perhaps borrowed. Evidence for either of these two possibilities is lacking, though the first is more likely, and Dakin and Wichmann argue for it, adding the consideration of additional data that may not be relevant. This is not to say that we (Kaufman and Justeson) have a perfectly obvious derivation for *chikola:tl*~*chokola:tl*, because we do not. In any event, from the forms of the borrowings listed here, it appears highly likely that the languages that have it received the loan from Nawa.

Dakin and Wichmann (2000:62b) cite *chikola:tl* for the Nawa towns of Ocotepéc (Morelos), Ameyaltepec (Guerrero), Cuetzalan (Veracruz), and Rafael Delgado aka San Juan del Río (Veracruz). This form is also found in North Puebla Nawa (Una Canger, personal communication 2005), a type of Central Nawa. In the case of Cuetzalan Nawa, Dakin and Wichmann (2000:62b) point out that *chokola:t* is said now, but the older form is said to have been *chikola:t*. While Dakin and Wichmann cite data from Rafael Delgado, attributed to uncited work by Tuggy, as having *chikola:tl*, the more recent PDLMA data from Rafael Delgado (Romero 1999–2002) has both *chokola:tl* and *xikola:tl* (not *chikola:tl\**). Dakin and Wichmann argue plausibly that *chikola:tl* was the original pronunciation of this word. The pronunciation *xikola:tl* and its implications need investigation.

The form *chokola:tl* was documented by Kaufman in 1978 for Santa María Izhuatlán. Huasteca Nawa, both West and East (Kaufman 1969/1984–1993), has *chokola:tl* ‘chocolate’; but in Chontla Huasteca Nawa the word means ‘caldo de tripa de puerco’ (soup/broth made from swine gut). This suggests that a Nawa morpheme {chokol=} (or {chikol=}) combined with {a:} ‘water’ has a generic meaning with at least two applications. Unfortunately, the more generic meaning of {chokol=} is not easy to divine. Pipil, which has been separate from other forms of Nawa since circa A.D. 900, has *chukula:t* (Campbell 1985:200). (Fowler [2006:310] cites Sampeck 2005 for an archaeologically based date of circa A.D. 1200 for the arrival of Nawa speakers in the Izalcos region of El Salvador. This does not square with the linguistic data, if we assume that the Pipils arrived in Izalcos about the time when they became linguistically separate from other forms of Nawa—specifically, that of the Southern Gulf coast—which was probably around A.D. 900. If, however, they were in Chiapas or Escuintla for a while before A.D. 1200, the archaeological and lin-

guistic chronologies would not be in conflict.) This raises the possibility that the form *chokola:tl* is at least 1,000 years old, but it could also be a borrowing from the Central Nawa speakers (“mejicanos”) brought into Guatemala by the Spanish after 1525. In light of the other data discussed in this study (both above and below), the latter possibility seems more likely.

Even if {*chokol*=} is assimilated from {*chikol*=}, the closest comparanda in Nawa are *chiko* ‘bent in a half-circle’ and *chihkol-li* ‘thing bent in a half-circle’, neither of which is in fact {*chikol*}. Dakin and Wichmann’s (2000:63–66) hypothesis is that *chihkol-li* meant ‘cacao-beater’ in some kinds of Nawa, but *chihkol-li* has a preconsonantal /h/, and *chikol=a:-tl* does not. On the one hand, there is no straightforward evidence for a Nawa word *chikol-li*\*; on the other hand, *chiko* and *chihkol-li* must be related, both reflecting the meaning ‘bent, hooked’, and this is not consistent with Dakin and Wichmann’s hypothesis, which derives *chihkol-li* and their hypothetical *chikol-li*\* from a supposed Yuta-Nawan \**ci*’ ‘small/pointed stick’ (Dakin and Wichmann 2000:63–64) plus \**ku*- ‘tree, pole’ (Dakin and Wichmann 2000:64–65). These difficulties render their proposed etymology unconvincing.

Their proposed Yuta-Nawan etymology in particular is untenable. Dakin and Wichmann (2000:63–66) argue for a Yuta-Nawan etymology for this word by attempting to analyze between 15 and 20 polysyllabic Yuta-Nawan words into monosyllabic roots, with the aim of providing evidence for putative proto-Yuta-Nawan elements \**ci*- ‘pointed stick’, \**ku*- ‘tree, pole’; and \**ri* ‘noun derivational suffix’. This section is of interest mainly as an illustration of Dakin’s long-term research program of etymologizing Nawa lexical items of two or more syllables as compounds made up of two or more monosyllabic roots. This type of analysis is not employed by most other Yuta-Nawanists, who recognize a limited number of monosyllabic roots in each language and in proto-Yuta-Nawa itself—the majority of roots being disyllabic. This analysis into monosyllabic roots is not required for the morphological analysis of the lexicon into its constituent morphemes in the individual languages. As is to be expected if these items are not in fact composed of monosyllabic elements, the meanings associated with the parts of Dakin’s proposed compound words do not often bear a compositional relationship to the meaning of the hypothesized compound, and the resulting monosyllabic elements typically lack finely focused semantic specificity.

The case at hand illustrates these problems. There is practically no identity of structure or gloss between any two Yuta-Nawan languages in the cited vocabulary, just partial overlap. There is one example only of an apparently plausible cognate—one with a close semantic matching—between two Yuta-Nawan languages:

Kawaiisu *či-ku-li* ‘stirrer’ (cited from Dakin and Wichmann)

Nawa *chikol=a:-tl*, which *might* say literally ‘stirring stick water’

However, this comparison is not valid, because Kawaiisu /u/ corresponds not to Nawa /o/ but to Nawa /i/. The proposed Yuta-Nawan etymology must be rejected.

If there were a Nawa word *chikol-li*\* ‘stirring stick’, then Dakin and Wichmann’s hypothesis that *chikol=a:-tl* meant ‘stirring stick water’ (their own term is ‘beater-drink’) would be plausible and would have no opposition from us. It would not, however, support their supposition that the diffusion of the word \**kakawa* and that of the name for chocolate were related in any way.

Further, Dakin and Wichmann’s claim that Mayan forms like [chukul] that mean stirring stick are borrowings from Nawa is false. These forms descend from a noun \**tuuk.ul* ‘stirring stick’ that can be reconstructed for Greater Q’anjob’alan and K’ichee7, at least, and that is derived from a Mayan verb \**tuk* ‘to mix, stir’ that can be reconstructed from Eastern Mayan and Greater Q’anjob’alan languages (Kaufman with Justeson 2003:395).

In any case, why the above combination in Nawa should mean both ‘powdered roasted cacao whipped/shaken/stirred/frothed with water and seasonings’ and ‘chitterling soup’ is by no means clear. We may remind ourselves of how unexplainable some lexical formations are by considering the name of the storm god *Tla:l=o.k*, which literally translated is ‘one who lies on the ground’, or the dwelling of the blessed after death *Tla:l=o.k.a:n*, which literally is ‘place of lying on the ground’.

#### The Diffusion of \**chikol=a:-tl* and its Relationship to that of \**kakaw(a)*

Dakin and Wichmann (2000:62) argue that the distribution of *chikola:-tl* and *chokola:-tl* across Nawa suggests that *chikola:-tl* originated within Eastern Nawa, and that *chokola:-tl* developed from it within Western Nawa based on a propensity of the latter for “vowel harmony” (by which they mean assimilation between vowels of adjacent syllables). As stated in the previous section, we agree with them that the form *chikola:tl* is likely to have been the original pronunciation of this word. Their further conclusions about its dialect history, however, are speculative, not secured by either linguistic or culture-historical data.

1. The assimilation of vowels in adjacent syllables is not a regular process (or rule) in *any* Nawa dialect. Rather, it is basically sporadic and occurs not only in Western Nawa but also (contrary to Dakin and Wichmann) in Eastern Nawa. (Some forms of Western Nawa have rules for the assimilation of the short vowels in inflexional prefixes, but rules or outcomes of this type are not general in the lexicon.) In addition, *chokola:tl* is found in Eastern Nawa dialects, including Huasteca Nawa dialects, Santa María Izhuatlán Nawa, and Pipil. While we consider it likely (because of specific, known culture-historical data) that some of the forms with *o* probably spread into some forms of Nawa and, perhaps, even of Spanish, it is not at all certain that all of this diffusion was due to this influence. The origin *chokola:-tl* from *chikola:-tl* therefore cannot be reliably attributed to either the Eastern or the Western branch of Nawa.
2. Even if the assimilated form did happen to originate in one of the Western dialects, this provides no evidence whatsoever that the earliest *chikola:-tl* originated in an Eastern dialect rather than in a Western dialect, nor that its diffusion into other Mesoamerican languages was from an Eastern dialect. Not only do some modern Western dialects have the *chikola:-tl* form, as Dakin and Wichmann observe, but since *chokola:-tl* is a later development, *chikola:-tl* could have developed in a Western dialect and spread from there to other Nawa dialects and/or to other Mesoamerican languages before changing their own pronunciation of the word to *chokola:-tl*.

While Dakin and Wichmann raise this Eastern-origin scenario quite tentatively—“Could it be that the *čikola:tl* form is an Eastern Nahuatl form . . . ?” (Dakin and Wichmann 2000:67)—the status of this speculation is elevated to the status of a fact on its next mention:

The reasons for focusing on the Pipil as the group most likely to have been responsible for the dispersal of the word *kakawa-tl* are

historical as well as linguistic. Pipil descends from the Eastern Nahuatl dialect, whose speakers, *as we have seen*, also created the word *čikola:-tl*. It is reasonable to suppose that these two words share their center of dispersal. [Dakin and Wichmann 2000:67b; emphasis added]

They go on to address possible objections to this possibility that involve differences in the grammatical forms and regional distributions of the two words.

Note that the dispersal that they had argued for on page 62—albeit based on the incorrect claim of a restriction of assimilatory processes to Western dialects—had been for dispersal among Nawa dialects. Here, however, Dakin and Wichmann incorrectly present their conjecture as having been a conclusion that it was from Eastern Nawa that *chikola:-tl* was spread to other Mesoamerican languages. (The first sentence of the quoted passage is also misleading in that they in fact present no linguistic argumentation for the conclusion that Pipils were responsible for the claimed spread of *kakawa-tl* to other Mesoamerican languages.)

Dakin and Wichmann nowhere explain why they consider it “reasonable to suppose that these two words share their center of dispersal”—here clearly presented as an assumption, not a conclusion. The only answer that we have been able to divine, making full use of the context of this claim, is that they imagine these words to have diffused together—from the same cultural group as part of a single cultural process, if not at the same time.

Nawa *chikola:tl* yields:

Sayula Mijean *chikúla:t*

Kora *tzikura:*

Chayuco Misteko *sikula*

San Mateo del Mar Wavi *chikolət* (perhaps via Spanish)

From Spanish *chocolate*:

Huichol *sikurá:ti*

Mitla Sapoteko *chikulahd~chigulahd*

Besides the forms listed here, Dakin and Wichmann (2000:62) cite *i*-forms for San Juan Colorado Misteko, Tlaxiaco Misteko, Warijiyó, Chamorro, Asturian Spanish, Catalan, and Dutch.

Nawa *chokola:-tl* ‘chocolate’ yields:

Zinacantán Tzotzil *chukul 7at* ‘chocolate drink’

Oaxaca Chontal *-tzugulalh*

The incidence of other Mesoamerican languages borrowing the Nawa form *chokola:tl* is slim, indeed.

From Spanish *chocolate*:

Juchitán Sapoteko *dxuladi/chu+lati!7/* ‘chocolate’ < [chukulá:ti]

Pajapan Gulf Nawa (Peralta 2002–2007) *chokola:teh*.

Coxcatlan Huasteca Nawa (Kaufman 1969/1984–1993) *chokola:teh*

The last two forms suggest that the version of chocolate made and drunk by Spanish speakers tends to get its name adopted even by the people who invented the word and the drink in the first place. It is in fact the case that, at the present time, Nawa speakers often think that a Spanish word of Nawa origin is actually the source of

a Nawa word they use natively; following up on this false impression, they often adopt the Spanish pronunciation of a word and forsake the Nawa pronunciation.

It is quite clear that the pre-Columbian diffusion of this Nawa word, in either form, was extremely limited compared with that of *kakaw(a)* in pre-Columbian times. Certainly, they could not have diffused together. It is also extremely unlikely that the term for chocolate and the term for cacao diffused through Mesoamerica anywhere near the same time period: *\*kakaw(a)* was borrowed into Lowland Mayan languages before A.D. 400 (and not from Nawa) and shows the results of sound changes affecting whole subgroups of Mayan languages. The Nawa word for chocolate was borrowed in diverse forms into a variety of Mesoamerican languages, with no internal evidence in any instance for substantial antiquity. No distributional or linguistic evidence directly suggests that the diffusion of these words was related in any way. It is in fact not reasonable to simply “suppose” that they were.

#### The Timing of the Origin of *\*chikol=a:-tl*

A further wrinkle is that neither form of the Nawa word is even attested from the first decades of the Spanish colonization. It is not found in Alonso de Molina’s *Vocabulario en Lengua Castellana y Mexicana* and *Vocabulario en Lengua Mexicana y Castellana* (neither in the 1551 edition nor in the expanded 1571 edition) or Bernardino de Sahagún’s *Historia general de las Cosas de Nueva España* (1577). Given the lateness of its first citations, reasonable doubt may be entertained as to whether the Nawa forms *chikola:tl* and/or *chokola:tl* even existed in pre-Columbian times. Indeed, Corominas [1980–1983:2:385–386] and many other students of the history of Spanish do not believe that Nawa *chokola:tl* existed in the pre-contact period.

The drink made from cacao certainly existed in pre-Columbian times, but, among other things, it may simply have been called by the name of its principal ingredient. In Zinacantán Tzotzil, /kokow/ refers both to the seed and the drink (Laughlin 1975: 176). This is probably true in other forms of Tzotzil and in Tzeltal, in which /kokow/ (Tzotzil) or /kakaw/ (Tzeltal) is translated both as ‘cacao’ and ‘chocolate’. In Q’anjob’al of Santa Eulalia, /kakaw/ means both ‘cacao’ and ‘chocolate’ (Mateo-Toledo, personal communication 2005).

In Nawa itself, the word for cacao could also be used to name drinks that were made from it. In sixteenth-century Nawa, Sahagún (1558–1561, 1577) refers to the drink made from cacao as /kakawa-tl/ (“cacao”) (see Anderson and Dibble 1951–1982: 8.39–40, 11.119–120; Sullivan and Stiles 1988:202). Sahagún (1558–1561) has the following passage in a section dealing with feasting by nobles:

#### SAHAGUN’S SPELLING

#### ANALYSIS

<xuchiayo cacauatl in tiqui	xo:chi=a:.yoh kakawa-tl in ti-k-i-h
quauhnechuyo cacaoatl in tiqui	kwaw=nekw.yoh kakawa-tl in ti-k-i-h
vei ynacazyo cacaoatl in tiqui	we:yi i:-nakas.yoh kakawa-tl in ti-k-i-h
amo tle nelihqui cacaoatl in tiqui	amo tle:n nel.iw.ki kakawa-tl in ti-k-i-h
vlló cacaoatl in tiqui	ol.loh kakawa-tl in ti-k-i-h
temecaxuchio cacaoatl in tiqui	te=meka=xo:chi.yoh kakawa-tl in ti-k-i-h
tlachichioalcacaoatl in tiqui	tla.chi.chi:wa.l=kakawa-tl in ti-k-i-h
vcyo cacaoatl	ok.yoh kakawa-tl

vctli	ok-tli
vino	wi:noh
chilo cacaoatl	chi:l.loh kakawa-tl
tlixuchiyo cacaoatl	tli:l=xo:chi.yoh kakawa-tl
yuluxuchiyo cacaoatl.>	yo:lo:=xo:chi.yoh kakawa-tl.

cacao having flower water is what we drink;  
 cacao having honey is what we drink;  
 cacao having big ears is what we drink;  
 cacao having nothing mixed in is what we drink;  
 cacao having rubber (?latex) is what we drink;  
 cacao having stone-vine-flower is what we drink;  
 elaborated cacao is what we drink—  
 cacao having pulque,  
 pulque,  
 wine,  
 cacao having chilli,  
 cacao having soot-flower (vanilla),  
 [and] cacao having heart-flower (magnolia).

Sahagún 1577 also refers to <quahnejujo cacaoatl> /kwaw=nek.w.yoh kakawa-tl/ ‘honeyed cacao’, <xochiocacaoatl> /xo:chi.yoh kakawa-tl/ ‘flowered cacao’, <chichiltic cacaoatl> /chi:chi:l.ti.k kakawa-tl/ ‘red cacao’, <vitztecolcacaoatl> /witz= tekol=kakawa-tl/ ‘thorn-charcoal cacao’, <xochipalcacaoatl> /xo:chi=pa.l=kakawa-tl/ ‘flower-painted (i.e. pink or orange-colored) cacao’, <tiltic cacaoatl> [sic] /tli:l.ti.k kakawa-tl/ ‘black cacao’, and <itztac cacaoatl> [sic] /ista:k kakawa-tl/ ‘white cacao’—all referring to kinds of chocolate (the drink), not kinds of cacao. Molina (1571:Nawa-to-Spanish section 10v, column b) has <cacauatl> /kakawa=a-tl/ (“cacao water”) ‘beuida de cacao’ (cited also in Siméon 1885:56b). Molina (1571:161a) lists <xochiayo cacauatl> /xo:chi=a:yoh kakawa-tl/ (“flowerwater-having cacao”) ‘beuida de cacao con ciertas flores secas y molidas’ (drink made from cacao with certain dried ground up flowers). Molina (1571:Spanish-to-Nawa section 22r, column a) lists ‘cacao, beuida’, meaning by this “cacao—a drink”, “the drink called cacao”, showing that for Molina, chocolate was called simply *cacao* in his use of Spanish. On page 19r of the Spanish-to-Nawa section, Molina lists ‘beuida de cacao con mayz’ (drink made from cacao with maize): <cacauatl> /kakawa=a-tl/ (“cacao water”); ‘beuida de cacao con axi’ (drink made from cacao with chilli pepper): <chillo cacauatl, chilcacauatl> /chi:l.loh kakawa-tl/ (“peppery cacao”), /chi:l=kakawa-tl/ (“pepper cacao”); ‘beuida de cacao folo’ (drink made from cacao alone): <atlanelollo cacauatl> /ah tla-nelo:l.loh kakawa-tl/ (“unmixed cacao”). All these are cited as well in Molina’s 1551 edition (right-hand column on page 34), which also cites ‘beuida de cacao compuefta con flores’ (drink made from cacao put together with flowers): <xochiayo cacauatl>, <xochayo cacauatl> /xo:ch(i)=a:yoh kakawa-tl/ (“flower-watery cacao”).

Bierhorst (1985:54) cites <cacahuaoctli> /kakawa=ok-tli/ (“cacao pulque”) from the *Cantares Mexicanos*, which date from

circa 1582 and lack the word <chocolatl> or <chicolatl>. The word /kakawa=oktli/ is probably the same as the /ok.yoh kakawa-tl/ mentioned in Sahagún (1558–1561).

In the Nawa poems *Romances de los Señores de la Nueva España* (from 1582) there are two instances of *kakawatl* naming the drink chocolate:

[a] <ma xocon cua in cacahuatl, in cacahuaxochitl: ma ya on ihua in>  
 ma: xo-k-on-kwa in kakawa-tl, in kakawa=xo:chi-tl; ma: ya on-i-wa in  
 ‘may you eat the cacao, the cacao flower; may it already be drunk’  
 [poem 5, lines 13-15 (Garibay 1993:9)]

[b] <o ya noconi izquicacahuatl xochitl>  
 o ya no-k-on-i iski=kakawa-tl xo:chi-tl  
 ‘oh, already I drank toasted cacao, the flower’  
 [poem 55, line 30];

In the Spanish document *Relación de Juan Bautista de Pomar (Tezcoco, 1582)* (Garibay 1993:193), the following passage occurs:

Su bebida de los poderosos era cacao  
 (The drink of the powerful was cacao).

Fray Diego Durán’s *Historia de las Indias de Nueva España* and his *Islas de la Tierra Firme*, written before his death in 1586, do not use the word *chocolate*, but they do use the word *cacao* 16 times in reference to the drink and 27 times for the seeds of or for ground-up cacao (Durán 1965).

Hernando Ruiz de Alarcón’s 1629 *Treatise on the Heathen Superstitions that Today Live among the Indians Native to this New Spain* (cf. Andrews and Hassig 1984:132; Coe and Whittaker 1982:188), which reports on magical practices by Nawa speakers in Guerrero, refers to the drink made from cacao by the Spanish word *cacao*. The kind of Nawa found in Ruiz de Alarcón is from the same general dialect group as the Nawa of the Basin of Mexico, which can be called Central Nawa.

The foregoing is evidence suggesting that, in the Nawa of the Basin of Mexico—and perhaps in Central Nawa generally—no word *chikola:tl* or *chokola:tl* existed in the sixteenth century or before and that drinks made from cacao were referred to as *kakawatl* or by expressions that included it. The words *chikola:tl* and *chokola:tl* may have arisen in a peripheral type of Nawa, at an undeterminable date, and only spread later to Central Nawa.

The first known use of the Nawa word *chokola:tl* is cited in Corominas (1980–1983:2.385–386) from Francisco Hernández (1959 [1577]) as <chocollatl>. In Chapter lxxxvii (“cacahoquá-huitl árbol del cacao”), where Hernández discusses cacao and its varieties (Hernández 1959 [1577]:2:303–305), a good deal of information is presented. On page 304, he names four varieties of what he considers to be the same basic plant, which must be *Theobroma cacao*: <quauhcacahoatl> /kwaw= kakawa-tl/ “tree cacao”, <mecacacahoatl> /meka=kakawa-tl/ “vine cacao”, <xochicacahoatl> /xo:chi=kakawa-tl/ “flower cacao”, and <talcacahoatl> /tla:l=kakawa-tl/ “earth cacao” (not the homophonous ‘peanut’, discussed later). On pages 304–305, he considers adding <quahpatlachtl> /kwaw=patlach-tli/ “broad/flat tree” to this group. Sahagún 1577 distinguishes three types of cacao by the color of the fruits (Anderson and Dibble 1951–1982:8:39). Hernández (1959 [1577]:2:304) says that “hacen tambien de ella una bebida” (they also make a drink from it [cacao]), but he goes on to describe four different drinks made with maize (maize dough, as Clavijero’s description makes clear)

and <cacahoatl> /kakawa-tl/, by which term he specifically refers to the kernel:

1. <atextli> /a:=tex-tli/ (“water flour”) ‘pasta aguada’: ground <cacáhoatl> mixed with ‘grano indio’ (maize)—for refreshment and nourishment; also as an aphrodisiac;
2. [unnamed]: made from kernels of <cacahoapatlachtl> /kakawa=patlach-tli/ (“broad/flat cacao”), <cacáhoatl>, and ‘grano indio’ (maize)—for nourishment and refreshment;
3. <chocóllat> /chokol=a:-tl/: made from an equal number of kernels of <pochotl> /po:cho:-tl/ (*Ceiba*) and <cacahoatl>, with ‘grano indio’ (maize)—drunk lukewarm as a fattener and as a medicine for tuberculosis;
4. <tzone> /tzon.ɛh/ (“hairy”, “furry”): equal parts of roasted ‘grano indio’ (maize) and <cacáhoatl>—for refreshment and nourishment, not as medicine.

Before describing these four drinks, he mentions another use: a drink made from the cacao kernel alone that medicinally serves to reduce heat in the body (Hernández 1959 [1577]:2:305).

Other plants reported by Hernández that contain the element {kakawa} include <quauhcacáhoatl> /kwaw=kakawa-tl/ “tree cacao” (Hernández 1959 [1577]:2:305) and <tlalcacáhoatl> /tla:l=kakawa-tl/ (“earth cacao”) ‘peanut’ (Hernández 1959 [1577]:2:306–307; both are apparently different from the terms of identical form listed above); <iztactlalcacáhoatl> /ista:k tla:l=kakawa-tl/ ‘white peanut’ (Hernández 1959 [1577]:2:307); <cacahoaxóchitl> /kakawa=xo:chi-tl/ “cacao flower” (Hernández 1959 [1577]:2:307–308; in Huasteca Nawa this is *Hamelia patens*, ‘scarletbush’), <cacahoapatli> /kakawa=pah-tli/ “cacao medicine” (Hernández 1959 [1577]:2:308). None of these is identified as a kind of <cacáhoatl>. (The acute accents on Nawa words cited by Hernández were probably added by the Latin-to-Spanish translator.)

The next known citation of the word *chokola:tl* in Nawa is from Clavijero (1780), who cites Nawa <chocolatl> with the gloss ‘alimento hecho con almendras de cacao y semilla del árbol llamado *pochotl*, en partes iguales [a food made from cacao kernels (“almonds”) and the seed of a tree called *po:cho:tl* (*Ceiba*, silk-cotton tree), in equal parts]’ (Siméon 1977:107).

Clavijero (cited in Santamaría 1992:412), who was born in Veracruz, writes as follows of drinks made from cacao:

Con el cacao formaban varias bebidas comunes, y entre ellas las que llamaban *chocolatl*. Molían igual cantidad de cacao y de semilla de *pochotl*; ponían todo junto en una vasija, con una cantidad proporcionada de agua; allí lo meneaban y agitaban con el instrumento de madera llamado *molinillo* en español; hecho esto, ponían aparte la porción más oleosa que quedaba encima. En la parte restante mezclaban un puñado de pasta de maíz cocido y lo ponían al fuego hasta darle cierto punto, y después de apartado, le añadían la parte oleosa y esperaban a que se entibiase para tomarlo. ... Los mejicanos solían perfumar su chocolate y las otras bebidas de cacao, o para realzar su sabor, o para hacerlas más saludables, con *tlixochitl* o vainilla, con flor de *xochinacatzli*, o con el fruto del *mecaxochitl*, y las dulcificaban con miel como nosotros hacemos con azúcar.

[From cacao they made several common drinks, among them the ones[!] they called *chokol=a:-tl*. They ground the same amount of cacao and the seed of *po:cho:-tl*; they put it all together in a vessel with an appropriate amount of water; there they stirred and shook it with the wooden tool called *molinillo* in Spanish; having done this, they set aside the oiliest part that came to the top. In the remaining part they mixed a handful of corn dough,

and they put it [the preparation] on the fire until it reached a certain point; after taking it off the fire, they added the oily part and waited till it was lukewarm to drink it. ... The Mexicans were accustomed to perfume their chocolate and the other drinks made from cacao, either to bring out their flavor, or to make them healthier, with *tli:l=xo:chi-tl* (“soot flower”) or vanilla, the flower of *xo:chi=nakas-tli* (“flower ear”), or with the fruit of *meka=xo:chi-tl* (“vine flower”), and they sweetened it with honey like we do with sugar.] [All of the English past-tense verbs are in the imperfect in Spanish and would more accurately be rendered in English as “used to VERB” or “would VERB”.]

Clavijero’s description (from 1780) matches that given by Hernández (from circa 1580) in a great many (not all) of the details. Though Hernández’s descriptions of drinks made from cacao do not specify any sweeteners (and this has wrongheadedly been made much of by historians of chocolate), Sahagún (1577) does refer to “honeyed cacao”.

Apparently because it is not attested in the earliest sixteenth-century sources, etymologists of Spanish (see, Corominas 1980–1983:2:385–386) seem universally to agree that *chokola:tl* is not a genuine Nawa word. Rather, they think *chocolate* is a Spanish word created by Spanish speakers through the mangling of a Nawa word or expression. As a result, etymologists of Spanish have proposed a great number of hypothetical origins for this word. Not a single one of the etymologies suggested by members of this brotherhood seems to have a chance of being correct. By way of illustration, we discuss proposals by Corominas (1980–1983:2:385–386). He discusses and discards several proposed etymologies of *chokola:tl* and *chocolate* that depend on Spanish speakers’ mangling the pronunciation of some no longer extant Nawa expression, and then this mangled word being borrowed into Nawa in the guise of a Nawa word. Then he offers his own suggestion. According to Corominas, since chocolate was made from equal amounts of cacao (Nawa *kakawa-tl*) and ceiba (Nawa *po:cho:-tl*) kernels, the Nawas probably called it *pocho-kakawá-atl* “ceiba cacao drink”. This was contracted to *cho (ca)cahuatl*, and this, in turn, would have been mangled into *chocolatl* in the mouths of Spaniards. He also contemplates that a form something like *xochayocacahuatl\** (which is a misspelling of Molina’s <xochiayo cacauatl>) could be the source of *chocolatl*, again through mangling in the mouths of Spanish speakers.

This proposal is so fraught with speculation that it hardly merits serious discussion. In the interest of explicitness, we observe the following specific difficulties. First, the premise is incorrect. There were many types of cacao drinks, as discussed above, and not all were made with ceiba seeds. Furthermore, there is no evidence of any such word as *po:cho:=kakawa-a:-tl\** in Nawa, and deriving *chokola:tl* from it requires four separate manglings: the loss of two non-adjacent syllables, the initial syllable *po* and one of the *ka* syllables; the change of *a* in the remaining *ka* to *o*; and the change of *w* to *l*, a substitution not found in established Spanish borrowings from Nawa. Deriving *chokola:tl* from Molina’s <xochiayo cacauatl> requires loss of three syllables (*xo*, *chi*, and *ya*) and the following segment *y*, along with the changes of *a* to *o* and *w* to *l*. These are wholly implausible hypotheses.

The fact is that the Nawa forms *chokola:tl* and *chikola:tl* are not plausibly the re-Nawatizations of Spanish *chocolate* and *chocolate*. There is no evidence that any Spanish word of the shape /..Vte#/ was ever borrowed into Nawa ending with /..V:tl/, and it

is clear from the discussions of both Hernández and Clavijero that they thought that *chocolatl* was just as much a Nawa word as all the other Nawa terms they brought under discussion.

The first occurrence of the Spanish word *chocolate* is found in Book 4, Chapter 22, of Joseph (José) de Acosta's *Historia Natural y Moral de las Indias* (1590): "El principal beneficio defte cacao es, vn breuaje que hazen q <ue> llaman Chocolate, que es cofa loca lo que en aquella tierra le precian, y algunos q <ue> no eftan hechos a el, les haze aſco [The main benefit of this cacao is a drink that they make that they call chocolate, which in that land they prize like mad: some who are not used to it are disgusted by it]." An English translation of Acosta's work from 1604 provides the first citation of *chocolate* in the Oxford English Dictionary. The Spanish word *chocolate* also occurs in the Motul dictionary of Yukateko as the gloss for <chacau haa>; this work is thought to have been produced around 1590.

This evidence also shows that in the Spanish of central Mexico, chocolate was called *cacao* until well into the seventeenth century, and we know of no evidence of the word *chocolate* or *chicolate* being used there at that time, even though José de Acosta, who lived in both Mexico and Peru, was using the word *chocolate* by 1590. We must suppose that his usage in Spanish was simply different, for reasons that we are not at the moment able to determine, though his place of writing, his place of origin, and his social-group affiliation are all possibly relevant. Its use in the Motul dictionary suggests that *chocolate* may have been used by 1590 in Spanish in the Yucatan.

The use by Hernández of Nawa *chokola:tl* (spelled <chocollat>) shows that the word existed in Nawa by 1577, but it does not show where in the Nawa-speaking world it was used. Hernández collected information in various parts of Mexico, as well as in Peru and the Philippines. The fact that the next known citation of this Nawa word is in Clavijero (1780) shows that there are serious gaps in our documentation of this Nawa word.

#### Epigraphic Mayan Evidence about Cacao, Pataxte, and Chocolate

As in the case of colonial Nawa, several Mayan languages use their word for cacao as a word for the drinks made of cacao. This is true in Epigraphic Mayan as well. This section explores the linguistic expressions used in Classic Mayan texts to refer to such drinks. Our conclusions are in general agreement with Stuart (2006), in part because both were completed after his 2005 workbook and in part because Justeson and Stuart discussed many of the issues before our respective papers were written. We differ on the interpretation of a few individual terms in the passages to be discussed, but mainly in our grammatical analyses and correlated issues involving relationships among ingredient terms. Our interpretation of elements of the phrase makes use of mostly the same data as that cited by Stuart (2005) in his workbook, although we originally extracted it from Mora-Marín's (2003) database.<sup>2</sup>

The key sign that represents 'cacao' is a depiction of a fish, sometimes rendered as the head of a fish with fins at the back. It has long been known that this sign represents the syllable /ka/ in some contexts, in which it may be transcribed as <ka<sub>2</sub>> (<ka>

is the transcription of a separate syllabogram). Stuart (1988) demonstrated that this sign is regularly used in spellings of the word for cacao in the Primary Standard Sequence (PSS) on Classic Mayan vases. The attested spellings show the following pattern of variation, ordered by their approximate relative frequencies in Mora-Marín's (2003) database:

- 80% <ka<sub>2</sub>-wa>
- 10% <ka-ka<sub>2</sub>-wa>
- 5% <ka<sub>2</sub>>
- 5% <<sup>2</sup>ka<sub>2</sub>-wa>

The spellings <ka<sub>2</sub>-wa> and <ka<sub>2</sub>> show that this sign is generally treated as a logogram for cacao in these contexts; we transcribe it here as <CACAO>. The spelling <ka-ka<sub>2</sub>-wa> can be treated as a fully phonetic spelling or as a logographic spelling with phonetic complements, <ka-CACAO-wa>. The spelling <<sup>2</sup>ka<sub>2</sub>-wa> shows that Mayan scribes made use of the sign in two logically distinct values in this context. Similar variation is found in other logosyllabic writing systems.

These spellings of the word for cacao almost always occur in a noun phrase within the part of the PSS that refers to the drinking vessel of a prominent person, used for some kind of cacao. This phrase usually has the following formulaic structure:

*y-uk' .ib' ta/ti* NOUN<sub>1</sub> (and NOUN<sub>2</sub> . . .) NAME

which should be translated as 'NAME's drinking vessel for (a drink made of) NOUN<sub>1</sub> (and NOUN<sub>2</sub> . . .)'.<sup>3</sup>

The word *kakaw* is usually the head of the noun phrase specifying what the drink was made of. Because Epigraphic Mayan had no regular way to express "and", this meaning and parsing mostly has to be supplied by the reader. Grammatically, a modifier of an expression parsed as a conjunction of two phrases can apply only to the first of those phrases.

About 10% of these phrases consist of a mere noun phrase rather than a prepositional phrase. About 5% consist of a string of two or more prepositional phrases; each of these prepositional phrases applies to *y-uk'-ib'* but cannot qualify one another. The word *kakaw* is to be understood as the head only of the prepositional phrase that it ends and not of any preceding or following prepositional phrase.

We have investigated the structure of references to drinks made from cacao in Mayan hieroglyphic texts, using as our corpus mostly the texts transcribed by Mora-Marín. To anticipate our conclusions, these drinks are referred to by their ingredients. The most commonly mentioned ingredients are (1) cacao (<CACAO-wa>, etc. /kakaw/); (2) "tree cacao" (e.g., <TREE-(7e-)le CACAO-wa> /te7-e:l kakaw/); and (3) maize (<MAIZE>, <7i-MAIZE> /7ixi:m/). In some cases, the drink itself may be referred to as *kakaw*.

If we restrict our attention to those texts that make explicit reference to cacao, the following observations summarize these occurrences:

- (1) In one case (on K2777), the spelling of the word for cacao is used with a final syllabogram <la>, which spells a suffix {-a:l},

<sup>2</sup> For some time after the breakup of proto-Greater Tzeltalan, Ch'olan must have retained a vowel-length distinction that was later lost. We do not know in

what era this distinction was lost in Ch'olan, but Ch'olan words are represented in this section with the vowel length that Ch'olan had at some point.

seemingly reflecting an adjectival use as a modifier of a subsequent word (Stuart 2006:191–192).

(2) In all other PSS texts, the word for cacao, when present, is the final word in the phrase referring to the contents of the vessel. In about 10% of cases, it is not preceded by a modifier but immediately follows the word *y-uk'.ib'* ‘his drinking vessel’. Typically, it is preceded by one or two modifying nouns or adjectives. In about 8% of cases, three or more modifiers precede it.

(3) With one exception, there is no statistically significant tendency for any pair of modifiers of ‘cacao’ to occur together. The exception is that there is a strong tendency for a word spelled by a sign <MAIZE>, depicting the head of the young maize god, to occur in the same noun phrase as *te:7* or *te7-e:l* (spelled <TREE>, <TREE-le>, or <TREE-7e-le>). In about 80% of the instances in which one of these occurs, the other occurs as well, and the order is always with ‘maize’ before *te7-e:l*. Often they occur together in the same glyph block. It seems that they are (or might be) part of a single expression.

Martin (2006) provides an interpretation of this association in terms of a mythological complex that he identifies and analyzes through its iconography. One key piece of this iconographic complex is a set of depictions of a cacao tree bearing ears of corn. Martin suggests that this tree is named by the glyphic sequence <MAIZE TREE> in the noun phrases referring to the contents of drinking vessels in the PSS. Stuart (2006:197) proposes that <MAIZE> is to be read as *ixim* (transcribed by us *7ixi:m*) based on two cases in which the sign is preceded by <7i> (K791, K8764). Both Martin and Stuart read <MAIZE TREE> literally as <7iximte'> ‘maize tree’. (In our representation, this would be *7ixi:m=te:7*.)

In our view, Martin has provided substantial evidence for the existence of an iconographic and mythic complex associating maize with cacao trees. However, he does not provide detailed argumentation for the reading as *iximte'* of the glyphic phrase discussed above. We show here that this reading is inconsistent with the range of variation in spellings of this phrase, and we provide an alternative interpretation that is consistent with all of the data known to us. Before doing so, however, it is relevant to point out that the expression *7ixi:m=te:7* is widely found in Mayan languages of the highlands, sometimes referring to the breadnut (Spanish *ramón*)—a fruit that is eaten in times when the maize harvest is poor—and otherwise to a large, very tall tree that grows in the rain forest. (In most Mayan languages the breadnut is called by a form descending from proto-Mayan \**7ojx*.) Among Lowland languages, it is found only in late colonial Yukateko where it refers to a kind of bush; and in modern Ch'orti7, where it refers to a wild tree, *Karwinskia calendroni*, whose leaves are used medicinally (Stuart 2006:198–199, citing a personal communication from Johanna Kufer). In existing Mayan languages, these terms do not name a mythological “maize tree” or define “maize tree cacao” as a variety of cacao. The most we can squeeze out of all this is that maize is a highly evocative plant and has spawned several plant *names* that are morphologically the same but do not otherwise refer to biologically related plants.

In spite of their strong statistical association and consistent word order, the two terms ‘maize’ and *te7-e:l* appear to have separate relevance to the chocolate drink. This is certainly the case in a text in which the two seem to occur separately as modifiers of ‘cacao’ <yu-k'i-b'i MAIZE CACAO-wa TREE-le CACAO-wa> (K5857)—that is, in the expression “his drinking vessel [for]

maize (and) [plain/default] cacao (and) tree-type cacao”. Their separate relevance is further suggested by the fact that ‘maize’ and either *te:7* or *te7-e:l* each occurs about 20% of the time without the other when ‘cacao’ follows. These examples do not support an interpretation of <MAIZE TREE> or <MAIZE TREE-le> as spelling a single (compound) lexical item.

(4) In 90% of the texts mentioning cacao, there are noun and/or adjective modifiers before that word. The norm in these cases is for the string of modifiers to begin with a preposition, indicating that the vessel is “for” a drink made with the stated type of cacao. In case there are no modifiers, the norm is for no preposition to occur (the word ‘cacao’ is directly preceded by the preposition in a few cases, although only in a minority of instances in which this word occurs without modifiers).

(5) In a few instances, more than one phrase or term is preceded by a preposition (the word <yu-ta-la> or <yu-ta> is preceded by one of these prepositions in most of these cases, and *kakaw* is involved in about half of them):

K7190	ta tzi-hi	ti CACAO-wa
K2704	TA yu-ta-la	?TA CACAO-wa
K1558	TA yu-ta-?la	TA CACAO-wa
K8245	TA yu-ta-la TREE-?	ti CACAO
K1092	TA CACAO-wa	ta tzi ...
K5016	ta yu-ta	ta MAIZE TREE-?
K2323	ta MAIZE	?ta yu-ta-la-ja CACAO-wa
K8719	ta yu-ta	ta ?MAIZE TREE-le CACAO-wa

In these situations, the prepositional phrases are separate statements, and cannot be conjoined to modify the word ‘cacao’. The same thing occurs when the word ‘cacao’ does not appear in the text, including cases that are followed by nominal phrases and so do not involve truncation of the portion of the text that names the contents, as in:

K4572	TA yu-ta	ti tzi-hi
K4909	TA yu-ta-la	ti tzi+hi
K8728	TA yu-ta-la	ti tzi-hi

Details of sign execution and text content suggest that these texts, along with K7459, are not independent cases.

(6) With this background, it is possible to arrive at a straightforward interpretation of the meanings of many of these expressions that is also consistent with the detailed accounts of cacao and cacao-based drinks from colonial Nawa sources.

The term *kwaw=kakawa-tl* “tree cacao” refers to a particular variety of the cacao plant mentioned by Hernández for colonial Nawas (see “The Timing of the Diffusion of *chikol=a:tl'*”, above). In fact, all of the Nawa terms in which a word for a general class of plants (not specific plants) is used as a modifier of *kakawa-tl* were used to distinguish varieties of cacao. “Tree cacao” is also known in Sokean: *kuy kakawa7* in Tecpatan Soke refers to the *cacaté* “tipo de fruta como nuez silvestre que es amarga [a kind of fruit like a wild (wal)nut which is bitter]” (the fruit resembles an almond, and it is toasted, salted, and eaten), while in Sotepan Gulf Sokean, it refers to cacao. We take *te:7 kakaw* and *te7-e:l kakaw* to have been a variety of cacao plant,

“tree cacao”, and we interpret almost all cases of *te7-e:l* in these texts as meaning “tree-type”, even if the head of the construction has been left out.

A viable alternative interpretation entertained by Mora-Marín (2003:Figure 1) is that *te7-e:l kakaw* refers to ‘forest (i.e., uncultivated) cacao’. The comparative support makes “tree cacao” seem to us a more attractive hypothesis, but in either case, the term would designate a kind of cacao plant, not a kind of drink.

Since half of the vessels in which *kakaw* takes a modifier use the term *te7-e:l kakaw* to refer to the contents of these vessels, and no frequent modifier of *kakaw* is in complementary distribution with it, this must have been a common type of cacao for use in these vessels. It could have been the type of cacao that was implicit when the word *kakaw* was used without *te:7* or *te7-e:l* as a modifier, but it is also possible that *kakaw* and *te7-e:l kakaw* labelled contrasting varieties of *Theobroma cacao*. Further evidence for this interpretation is provided below, under point (7).

It should be noted that the word *kakaw* may have referred, and probably did refer, to a range of related varieties of *T. cacao*. It was no doubt possible to make these varieties explicit, but readers would have had the knowledge to infer what varieties were intended as default interpretations of these terms in particular contexts. We today lack this knowledge, so certain types of information will not be readily accessible to us.

Nawa sources have no word that literally translates as “maize cacao” or the like. However, in these sources, all of the terms in which words for foods or food ingredients were used as modifiers of *kakawa-til* were used to refer to drinks made of cacao that used those ingredients. We suggest that the use of the word for maize was such a usage. A similar case is *nal kakaw*, attested in the phrase <yu-k’i-b’i TA-?? na-la CACAO-wa> on an Early Classic vessel (Stuart 2006:192; Mora-Marín, personal communication 2005, interprets some of the MAIZE logograms as spelling *nal*).

As in the case of *te7-e:l*, half of the vessels in which *kakaw* takes a modifier have MAIZE in these phrases. This suggests that cacao mixed with maize was a very common type used in these vessels. Most of the drinks made from cacao by the proto-historical Nawas were prepared from ground-up roasted cacao kernels combined either with maize dough or with ground-up roasted maize kernels. This includes most of the cacao-based drinks distinguished terminologically by Nawas that were discussed in the previous section. This practice continues to the present day in indigenous communities in Mesoamerica, including among Mayans, Mije-Sokeans, and Nawas, and was doubtless general in proto-historical Mesoamerica.

It may be noted that *7ixi:m* is the only one of the frequent terms in these phrases that precedes *kakaw* and that does not occur with a *-V:l* suffix. All of the other terms show variation (although in the case of *tzi:h* ~ *tzih-i:l*, the form with a *-V:l* suffix is rare).

Under the interpretations presented here, the phrase <yu-k’i-b’i ta MAIZE TREE(-le) CACAO(-wa)> can be read *y-uk’.ib’ ta 7ixi:m* [and] *te7-e:l kakaw* and can be translated, ‘his drinking vessel for [a drink made from] maize [and] tree cacao’.

The order of these terms is consistent in always placing all ingredient terms before the only recognized plant-type modifiers.

Color terms were used among the Nawas in names of types of cacao and in names of other kinds of trees that include the root *kakawa*. Sahagún mentions that some types of cacao plants were distinguished by the colors of their fruit. On these analogies, <k’an-na CACAO-wa> seems more likely to be spelling *k’an kakaw* (‘yellow cacao’ or ‘ripe cacao’) and less likely *k’a:n kakaw*

(‘prized cacao’) in <ta yu-ta k’an-na CACAO-wa> (K625), and the expression might refer either to a variety of cacao from which a drink is made, or it could be a general modifier of an ingredient, describing the state (ripe) of the cacao kernel.

(7) Two other words that occur in phrases referring to the contents of the drinking vessels are <tzi> ~ <tzi-hi> ~ <tzi-hi-li>, and <yu-ta<sub>1</sub>(-la)>. MacLeod and Reents-Budet (1994:118) identify this word with pre-Cholan \**tzi:h* (which, however, means ‘unripe, uncooked/raw’ rather than ‘fresh’, as she translates it). This appears to be a feasible interpretation, and we know of no other root whose distribution in the Mayan family suggests that it could be cognate with a Lowland Mayan form consistent with the phonetic spellings. Unlike the other terms considered so far (with the exception of *k’an* if for ‘ripe’ rather than ‘yellow’), the word *tzi:h* refers neither to a variety of the cacao plant nor to an ingredient in the drink. In the frequent phrase *tzi:h(-i:l) kakaw* ‘raw cacao’, the word *kakaw* must refer to the cacao kernel or pulp—thus, to cacao as an ingredient and not as the name for a drink.

Stuart (2006:188) notes that <yu-ta<sub>1</sub>(-la)> occurs in a possessed form <7u-yu-ta-la> on a carved vase illustrated by Dütting (1992:Figure 17); another possible instance is on K8088, which seems to be <7u-yu-ta<sub>2</sub>-la>. This shows that we are dealing here with a noun *yut* (presence of vowel length unverifiable). We know of no prior interpretation of this term that is consistent with Stuart’s observation. We raise the possibility that <yu-ta<sub>1</sub>(-la)> may imply the ability to confer good fortune. In colonial Yukateko, <yut> or <yutal> (whose possessed form is <u yutal>) names the bezoar, a stone found in the stomachs of some deer when they are butchered. This type of stone is believed to confer good luck on its owner even by present-day Mesoamerican Indians. If this is the correct connection for *yut(-a:l)*, the word may imply either the power of the bezoar or the use of it in the preparation of the cacao drink. From descriptions provided to Kaufman by some present-day Indians, the bezoar is not really a stone but more like a hair ball and is made of organic matter. Others say it is indeed a stone.

Further progress in the interpretation of these words depends on distributional analysis in terms of other words that appear in the same phrases. Such analysis is complicated by the fact that there is a general absence of multiple terms within the phrases under discussion, apart from *7ixi:m* and *te7-e:l*.

In 75–80% of the texts in which two or more terms precede the word for cacao, *7ixi:m* and *te7-e:l* are among them. Only 10% of texts that mention cacao have two or more added terms when *7ixi:m* and *te7-e:l* are not both among them. Since it is rare for more than two added terms to occur together, except in the case of *7ixi:m* and *te7-e:l* with one another, the occurrence of any one term is usually negatively correlated with the occurrence of any other. As a result, most of these terms can tell us little or nothing about any uninterpreted terms in these phrases.

This statistical pattern, however, may explain one seeming anomaly in the data: 80% of references to *te:7 kakaw* and *te7-e:l kakaw* are preceded by *7ixi:m* as a modifier, but only about 15% of references to *kakaw* are preceded by it. Apart from *7ixi:m* and *te:7* ~ *te7-e:l*, the presence of one modifier reduces the likelihood of the use of any other modifier, so just the opposite effect would be expected. Accordingly, *te7-e:l kakaw* is more strongly associated with a specification of maize as an ingredient than is *kakaw*. It is known that maize was typically mixed with *kakaw*, so this cannot be the reason for the difference; rather, it is presumably due to a far greater susceptibility of *kakaw* than of *te7-e:l kakaw* to the use



of *tzi:h* and *yut(-a:l)*. This supports the view that *kakaw* and *te7-e:l kakaw* are names for two different varieties of *T. cacao*.

A more involved set of inferences shows that neither *yut(-a:l) kakaw* nor *tzi:h kakaw* refers to a variety of the cacao plant.

In several texts, the word *tzi:h* alone refers to the contents of the drinking vessel. In many of these texts, a reference to the owner follows, so that these expressions do not involve truncation, which sometimes occurs in the PSS when space is lacking for completing the text.

When *tzi:h* and *yut(-a:l)* occur in the same phrase, they can occur in either order. Otherwise, each is usually the first of the modifiers in any phrases in which it occurs that specify the contents of the vessel. The only exception is in K2323, in which *ta 7ixi:m* is followed by *ta yut(-a:l) kakaw*—that is, in which *yut(-a:l)* is the first element after the preposition in a prepositional phrase in the content phrase. The order of terms in the phrases specifying contents are therefore largely consistent overall and may be fully consistent within a prepositional phrase.

The order among adjectives is likely to relate to the semantics or relevance of these modifiers; in many languages, semantically defined classes of adjectives occur in a fixed sequence except when departures have a specific contextual motivation. The distributional properties of *tzi:h(i:l)* and *yut(a:l)* indicate that these words precede both ingredient and plant-type modifiers of *kakaw*. If ingredient modifiers precede plant-type modifiers, as suggested above, then *tzi:h(-i:l)* and *yut(-a:l)* are not plant-type modifiers. They could both be ingredient modifiers, or they could both be modifiers of a third category that itself precedes both ingredient and plant-type modifiers. Another likely example of such a modifier in the PSS is <7a-ch'a> for proto-Greater Tzeltalan \*7ach' 'new' or pre-Ch'olan \*7a:ch' 'wet' in K8713 <ti 7a-ch'a ka<sub>2</sub>-CACAO-wa> (Zender 2002).

As with the other terms in these phrases, when the word *tzi:h (-i:l)* appears with *kakaw*, it is most often the only modifier of *kakaw*. When another modifier appears with *tzi:h*, it is almost always *te7-e:l*; *yut(-a:l)* does not occur in these cases. When *tzi:h(-i:l)* appears without *kakaw*, usually none of the other terms occurs with it, but *te7(-e:l)* or *yut(-a:l)* appears in about a third of these cases. In about a third of the phrases with *te7-e:l* but not *7ixi:m*, *tzi:h* occurs before it; *tzi:h* rarely or never occurs when *7ixi:m* appears without *te7-e:l* (and rarely accompanies the phrase *7ixi:m te7-e:l*). One possible interpretation is that cacao was more often uncooked than was maize in these drinks. Another is that the descriptive terms associated with the word *kakaw* might affect a reader's default interpretation of which particular variety of cacao was being alluded to.

In contrast, *yut(-a:l)* appears at about the same rate before *7ixi:m kakaw*, before *te7-e:l kakaw*, and before *7ixi:m te7-e:l kakaw*—in each case, about once in four instances. This word *yut(-a:l)* has one somewhat puzzling but potentially revealing association. Sixteen (70%) of the 23 cacao vessels in Mora-Marín's database whose owner is labelled <7i-tz'a-ti> are among the 77 with *yut(-a:l)* as the first modifier, while only seven are among 124 that lack it. This association is statistically significant ( $p=0.001$ ). In colonial Yukateko, the adjective <idzat> is attested, with the meanings 'clever, crafty, wise'. There may be a connection between possession of a lucky charm and cleverness.

(8) Another possible modifier in the phrase specifying contents, each time preceding <CACAO-wa>, is <LORD TREE-le> ~ <LORD-wa TREE-le> ~ <LORD TREE>. The first word is most likely *7aja:w* 'lord'. It could form a compound

with *te7-e:l*—thus, *7ajaw=te:7* or *7ajaw=te7-e:l*. A Mayan etymon \*7a:ja:w=tye:7 'white sapote (*Casimiroa edulis*)' is known from all Central Mayan branches, but it names a highland fruit tree. The corresponding word in Yukateko has a different referent. We do not know whether white sapote is a plausible ingredient of any cacao drink.

Along with the Central Mayan \*7a:ja:w=tye:7 'zapote blanco, matasano // white sapote (*Casimiroa edulis*)', as attested in Tzeltal *7aja[w]=te7* and Tzotzil *7aj=te7*, we have the Ch'ol form *käk=te7=pa7* ('cacao tree [of] gully') 'zapote de agua (fruta paracida a la del zapote)'. The terminological connection between certain types of sapote with both *7aja:w* and *kakaw* suggests that <LORD TREE> as an ingredient of a cacao drink might reflect the Central Mayan term \*7a:ja:w=tye:7, even though it is a highland tree.

In any case, it is feasible to analyze *te7-e:l* as the same one that specifies a variety of cacao, and to treat <7AJAW(-wa)> as a general modifier meaning 'lordly'. In all cases, <LORD> is the first element of the sequence, consistent with being a member of the same class as *tzi:h(-i:l)* and *yut(-a:l)*.

(9) Another possible rare modifier is <k'in>, which appears at most once in the corpus, on K3472. The sign is not entirely clear, and Stuart (personal communication 2005) points out that parts of this vase have been repainted. The sign at issue occurs in a spelling <ta k'i-?k'in>. Repainting seems to have affected parts of the preceding and following glyph blocks, since the ubiquitous fine lines due to root damage do not cross thickly applied black paint outlining parts of some signs, but we see no evidence of repainting of the <ta k'i-?k'in> sign group. The identification of the uncertain sign as <k'in> is due to Mora-Marín (2003), who marks the transcription as uncertain. However, the surviving details appear to be consistent with <k'in> and not with any other sign that is known to take a <k'i> superfix. If the sign's identification is valid, it seems likely to spell the Lowland Mayan word \*k'ihn 'hot'. This would make it a general modifier appropriate to the drinks made with cacao. It could, however, simply represent the word \*k'i:n 'sun', which would have to be a once conventional but no longer current usage as the name of a type of cacao plant or drink.

(10) Stuart (2006:194) identifies place names as another class of modifiers for 'cacao', suggesting that they indicate cacao from particular locations (so far, all in the eastern Peten). There are two clear examples and one likely example. One of these is the only modifier of *kakaw* in its phrase, and two of them immediately follow *yuta:l* and precede *kakaw*.

In summary, we distinguish three classes of terms that precede *kakaw* in the phrases describing the contents of the drinking vessels: a class whose members appear immediately before the word *kakaw* and specify a biological variety of cacao (III); a class of nouns whose members appear before biological variety labels and specify ingredients of the drink made with cacao (II); and a general class consisting of all other terms, all of them seemingly general types of modifiers (I). They can be charted as follows:

I	II	III
<i>yut(-a:l)</i>	<i>7ixi:m</i>	<i>te7-e:l</i>
<i>tzi:h(-i:l)</i>	( <i>7aja:w=te:7-el</i> )	/0/
<i>7ach'</i>		<i>k'an</i> 'yellow'
place name(-V:l)		( <i>k'i:n</i> 'sun-type')
<i>k'an</i> 'ripe'		

(or *k'a:n* ‘prized’)

(*7aja:w*)

(*k'ihn* ‘hot’)

Only one of these modifiers seems to require that the ‘cacao’ referred to in the PSS is the drink itself rather than the ingredients—<k'in> *k'ihn* (if for ‘hot’ and not ‘sun’)—and this may not be a correct reading of the original text. Certain of these modifiers seem to require that the cacao be an ingredient rather than the drink: <tzi(-hi)> *tzi:h* ‘raw, uncooked’; probably place names (presuming that they are places where a particular kind of cacao comes from); and <k'an> (if it is not for *k'a:n* ‘prized’). Other mentions of cacao provide no evidence that shows whether *kakaw* in those instances refers to an ingredient or to a prepared drink.

Table 7 presents the attested combinations of *kakaw* with its associated terms.

On a few cacao vessels (/y-uk'.ib'... *kakaw*/) we find the expression <JAGUAR × CACAO>, a superimposition of jaguar

Table 7. Epigraphic phrases referring to cacao and cacao drinks

A. Varieties of cacao	
1. <i>kakaw</i> ‘ <i>Theobroma cacao</i> ’	
(a) <i>kakaw</i> (without modifiers indicating variety, with a default reading unknown to us)	
<i>kakaw</i>	‘cacao’
<i>tzi:h kakaw</i>	‘raw/unripe cacao’
<i>7a(:)ch' kakaw</i>	‘fresh/new/wet cacao’
(b) <i>k'an kakaw</i> (1 ×; perhaps ‘yellow <i>Theobroma cacao</i> ’, if not ‘ripe cacao’ [no parallel in Nawa])	
<i>yut k'an kakaw</i>	‘lucky yellow/ripe cacao’
(c) <i>k'in</i> (or <i>k'ihn</i> ) <i>kakaw</i> (1 ×; perhaps ‘sun <i>Theobroma cacao</i> ’, if not ‘hot cacao drink’)	
(d) <i>te7-e:l kakaw</i> <sup>a</sup> (21 ×; ‘tree cacao’ [parallel to Nawa /kwaw=kakawa-tl/])	
<i>tzi:h te7-e:l kakaw</i> [6 ×]	‘raw/unripe tree cacao’
<i>7aja:w(=)te7-e:l kakaw</i> [1 ×]	‘lordly tree cacao’ or ‘white sapote and cacao’
(e) Varieties of cacao modified by place names and untranslated phrases	
<wi-ti-ki> <i>kakaw</i> (1 ×)	‘Copan(?) cacao’
<ko-xo-ma mu-lu> <i>kakaw</i> (1 ×)	‘? ? cacao’
<5-kab'> <i>kakaw</i> (1 ×)	‘Ixtutz cacao’
<SA7(-la) <i>kakaw</i> > (2 ×)	‘Naranjo cacao’
B. Combined ingredients of cacao drink	
<i>7ixi:m</i> [and] <i>kakaw</i> (15–17 ×)	‘maize with cacao’
<i>7ixi:m</i> [and] <i>te7-e:l kakaw</i>	‘maize with tree cacao’
C. Possible reference to drinks, secondarily listing ingredients	
<i>k'ihn kakaw</i> (1 ×)	‘hot cacao drink’
This term might be /k'in <i>kakaw</i> ‘sun cacao’ rather than ‘hot cacao’ and therefore be a variety of <i>Theobroma cacao</i> .	

<sup>a</sup>To EpM *te7-e:l kakaw* may be compared: (1) Yokot'an (Chontal Mayan) *te7el kākāw* ‘guoquo, gogo’ = *Salacia belizensis* ‘spindle tree’ [family]; *Fusanus Hippocratacea* [family] – an exact match to EpM /*te7-e:l kakaw*/; (2) Tecpatan Chiapas Soke *kuy=kakawa7* ‘cacaté (tipo de fruta como nuez silvestre que es amarga) = *Oecopetalum mexicanum*; and (3) Sotapan Gulf Sokean *kuy kaakwa* ‘cacao’, ‘cacahua’. *Salacia belizensis* and *Oecopetalum mexicanum* are different genera, but both bear the name ‘tree cacao’ in some languages.

markings on the <CACAO> logogram, which we interpret as *b'ahlam kakaw* ‘pataxte’ on the basis of forms with this meaning in Q'eqchi7 and in Ayapa Gulf Sokean that literally mean ‘jaguar cacao’; see ‘The Mije-Sokean Hypothesis’, above. All examples are of *7ixi:m te7-e:l b'ahlam kakaw*. We do not know whether *te7-e:l b'ahlam kakaw* is the Epigraphic Mayan word for pataxte or for a ‘tree-type’ variety of pataxte. In either case, it may relate to the origin of the later Ch'olan term, \**b'ahlam=te:7*, for pataxte. All of the examples refer to the drinking vessel being for a mixture of maize with pataxte, but since the few examples known to us seem to come from a single scribe or scribal school, the lack of variation in these rare references to pataxte is probably misleading. Other examples with dark markings on the face or body of the CACAO logogram, which may also be intended to represent a word for pataxte, occur with a broader range of ingredients.

#### A Proposed Mayan Source for Nawa *chokola:tl*

Dávila Garibi (1939) proposed that the Nawa word *chokola:tl* for the drink chocolate originated, in part, in a Mayan language. This proposal was accepted and elaborated by Coe and Coe (1996:118–119) with data from Yukateko and colonial Kaqchikel. It has been uncritically accepted by some scholars—for example, by Tedlock (2002:170)—but it is demonstrably false. It depends on a misunderstanding of the Kaqchikel sources and on a lack of understanding of the history of a Yukateko word meaning ‘hot’.

*Yukateko* ‘chocolate’ = ‘hot water’. Coe and Coe (1996:118) cite and endorse Dávila Garibi’s (1939) suggestion that Spaniards combined a putative ‘Maya’ (Yukateko) <chocol> (meaning ‘hot’) + ‘Aztec’ <atl> (meaning ‘water’) to form a new word for chocolate. The linguistic elements of their argument are (1) the fact that the colonial Yukateko term for the drink chocolate is attested in the earliest sources as <chacau haa>, an expression with the literal content ‘hot water’; (2) the erroneous belief that there existed a Yukateko form <chocol>\* meaning ‘hot’; (3) the speculation that there might have been a Yukateko expression <chocol haa>\* with the same literal content and thus the same application as the attested <chacau haa>; and (4) the speculation that Spaniards who knew enough Yukateko and Nawatl substituted <atl> for <haa> in this hypothetical form to produce a Nawa neologism *chokol=a:tl*. Finally (5), Coe and Coe (1996:118) cite a supposed K'ichee7 *chokola'j* ‘to drink chocolate together’ as somehow relating to this proposal, while admitting that how the terms could relate is unclear (in fact, this is really a Kaqchikel word, Coe and Coe’s phonological representation of it is faulty, and it means ‘to do something in common’).

Coe and Coe (1996:pp) cite Miguel León-Portilla (personal communication to Coe and Coe) as considering this a ‘reasonable explanation’. Tedlock (2002:170) endorses Coe and Coe’s opinion that the word *chocolate* incorporates a Mayan word *chokol* meaning ‘hot’. Nonetheless, the only factual element of this proposed etymology is the first point—namely, that there was a Yukateko form <chacau haa> ‘chocolate’ that is literally ‘hot water’.

There are several difficulties with this proposal, but the most important involves the supposed Mayan word *chokol*\*. (Except as noted, data cited in this section are from Bricker et al. 1998, for Yukateko; Ulrich and Ulrich 1976, for Mopan, Hofling with Tesucun 1997 for Itzaj; and Canger 1969 for Lakantun.) Modern Yukateko has a word *chokoh* ~ *chokow* ‘hot’; underlyingly it is

/chokow/, as demonstrated by the derived form /chokwil/ ‘heat’. This is in fact the same word as the Motul’s <chacau>. Proto-Mayan \**tiqaw* ‘hot’ is reflected in every branch of the family but Wastekan. Sometime in the ancestry of proto-Yukatekan, the \**t* shifted to *ch*; this is a regular change in Yukatekan that occurred whenever \**t* (or \**ty*) was followed by \**i* or \**e*. It also underwent the Yukatekan shift of \**q* to *k* (which is incidentally found in all Mayan languages that migrated into the lowlands out of the Mayan homeland in highland Guatemala). It cannot be determined which change happened first. Sometime after the change of \**t* to *ch*, pre-Yukatekan \**chiqaw* or \**chikaw* underwent vowel assimilation, leading to proto-Yukatekan \**chakaw*. Short *a* before final *w* sporadically changes to *o* in some Mayan languages. Yukateko *chokow* and Itzaj *chokoh* arose in this way from earlier \**chakaw*, while the proto-Yukatekan vowels descend normally to Lakantun *chäkaw* and Mopan *chäkaj*. (Justeson 1985 notes that the shift of \**w* to *j* is also attested in Mopan *käkaj* from \**kakaw*. This may be a regular change, as no other disyllabic words surviving in Mopan have word-final *w*. Final *w* in this word also shifts to *h* in Yukateko and Itzaj, although it resurfaces as *w* in some prevocalic contexts.)

The timing of the development of *chokow* in Yukateko can be roughly determined through the occurrences of this word in the sources cited in Table 8. These citations suggest that forms like *chokow* were well established but still in variation with *chakaw* in the mid- to late eighteenth century, while there is no evidence for any form but *chakaw* in the seventeenth century. To this extent, then, a suggestion that Yukateko *chokow* might have had a role in the origin of the Nawa word *chokola:tl* would be anachronistic. As for *chokol\**, no such form exists in the meaning ‘hot’ in any Mayan language.

The Yukatekan historical development behind *chokow* and *chokoh* is sufficient to eliminate this proposal, but there are other difficulties with it. One is its consequence that Nawa *chokola:tl* is a Nawatization of the Spanish word *chocolate*; we show in “Evidence concerning the History . . .”, above, that this is implausible. Another difficulty for Dávila Garibi’s hypothesis is created by Dakin and Wichmann’s (2000:62) plausible argument that the Nawa term was originally *chikola:tl*, thereby rendering inconsistent the comparison between the Mayan vowels and the Nawa vowels and rendering irrelevant the compared words themselves.

One feature of this proposal—the notion that the word for this Mesoamerican drink came from Spaniards and was then widely adopted by Nawas—is of at most minor concern. Such a process would provide a lexical distinction that did not fully or customarily exist in indigenous languages between cacao as a plant/kernel and the drinks that are made from it. There are many parallels, such as the adoption by English of Old French words for animals as food (*beef, pork*) in contrast to native words for animals on the hoof (*cow, swine*). Nonetheless, Coe and Coe (1996) attempt a rationale for Spaniards’ creating a word for this drink, a word that by chance was then widely adopted by Mesoamericans (specifically by Nawas and only occasionally by other Mesoamericans). They note the existence of a Nawa word that could have such a meaning, citing Molina (1571) for Nawa <cacahuatl> /kakawa=a:tl/ ‘cacao water’ (which they misspell as *cacahuatl*). They go on to suggest that this word was eliminated because the first two syllables remind one of *caca*, the Spanish word for ‘poop’ (Coe and Coe 1996:119). This argument is invalid because Spanish borrowed from Nawa the words *cacao, cacahuatl* ‘cacao plantation’, and *cacahuate* ‘peanut’ (cited in the same connection in Dakin and

Table 8. Colonial citations of Yukateko forms descended from *chakaw* ‘hot’

Motul dictionary (Ciudad Real ca. 1590)	
<chacau>	cosa caliente o calurosa y que tiene calor, y la calentura o calor
<chacau haa>	chocolate
San Francisco dictionary (seventeenth century, before 1690; Michelon 1976)	
<chacau>	caliente
<chacau haa>	agua caliente y chocolate
<Chacau haa; chacau>	agua caliente
<Chacaui>	calentura; calor
Beltrán de Santa Rosa (1746; completed 1743)	
<chocou>	lo caliente
<chochocou>	lo poco caliente ó lo tibio
<Chacaui>, <chocuil>	calentura
<Chocouhaa>	agua caliente
<Chucua>	chocolate
Ritual of the Bacabs (after 1779; Roys 1965)	
<cha[ca]u haa>	chocolate
<u chacau haail>	his chocolate
Book of Chilam Balam of Chumayel (between 1782 and 1828 [Gunsenheimer 2001:6])	
<chucua>	chocolate
Book of Chilam Balam of Tizimin (between 1824 and 1837 [Edmonson 1982:185])	
<chucuaa>	chocolate
Pío Pérez (data collected in the 1830s [Owen 1970:iii])	
<choco>	caliente
Bricker et al. (1998)	
/chokow/~chokoj/	hot
/chokwil/	heat

Wichmann 2000:62a), all involving the same morpheme /kakawa/, as in /kakawa=a:tl/; and Spanish also borrowed unrelated words containing this sequence, such as *cacalote* ‘crow’.

*Colonial Kaqchikel evidence.* In support of a Mayan role in the development of this term, Coe and Coe (1996:118) also cite a putative K’ichee7 word *chokola’j* ‘to drink chocolate together’. This is really a Kaqchikel word; it is a misspelling of <chocolaah>, and it means ‘to do something in common’. Tedlock (2002:170) also refers to K’ichee7 and Kaqchikel *chokola’j* ‘drink chocolate together’, citing Ximénez (1993), Varea (ca. 1600), and Coto (1983 [1656]); however, he appears to have gotten his information from Coe and Coe (1996) and not from the above-cited sources, since the spelling and gloss he uses come from Coe and Coe, who got them wrong. Tedlock also accepts Coe and Coe’s incorrect assertion that the *chocol* of *chocolatl* comes from a Mayan source (and further claims, without providing supporting evidence or a reference to anyone else who has made such a claim, that Spanish *cacao* came from a Mayan source rather than from Nawa). Dakin and Wichmann (2000:74) also cite this form, in the spelling <čokola>, as a borrowing of Nawa *chocola:tl*.

The reality is this. Colonial Kaqchikel had a noun <chocola> (pronounced something like /chokola/) and a denominal transitive verb <chocolaah> (something like /chokola7-a-; /-:j/ is an inflexional suffix). The structure of <chokolaah> is ‘to do <chocola>’. Because of the pragmatic applications of these words, as cited in Varea (ca. 1600) and discussed in this section, together no doubt with the shape of the word, some Mesoamericanists have drawn the false conclusion that <chocola> names in some way a drink made from cacao. To make this clear, we need to refer to the meanings of these and all related Kaqchikel words in the colonial sources. To anticipate our basic conclusion, these forms are based on a root {chok} that means ‘gathered together’.

ANONYMOUS (ca. 1578; SMAILUS 1989:2:166)

<tin choc 3ab ah apon> /ti+ n-chok q’ab’ aj 7apo:n/ (I “gather” hand X arrive) [vt] ‘llamar con la mano’ (to wave to someone to come hither)

<tika chocolaah> /ti+ qa-chok-ol-a7-a-j/ ‘we all do it to her together’ [Kaufman] [vt] ‘a una hazerse todos’

<tika chocolaah ru banic> /ti+ qa-chok-ol-a7-a-j ru-b’a:n-i:k/ ‘we all do its doing together’ [Kaufman] [vt] ‘hazer alguna cosa todos juntos’ (for everyone all together to do something)

<chocolaam r ahil vay> /chok-ol-a7-a-m r-aj-il way/ (the reckoning of the tortillas has been done together) [pcp < vt] ‘escote en el comer pagando cada uno su parte’ (contribution to the meal everyone paying his share)

VAREA (ca. 1600)

<chocola> ‘es cacao junto q<ue> dan veinte a cada uno, y lo beben entre todos’ (this is jointly shared cacao which is given in the quantity of 20 [kernels] to each, and they all drink it [the chocolate] together).

<tikaban kachocola, tikapopolih rukumic> /ti+ qa-b’an qa-chokola7, ti+ qa-popol-i-j ru-qu:m-i:k/ ‘we make/do our shared thing [chocolate], we communalize its drinking’ [Kaufman].

<chocolaah> ‘hacer algo de comun, como ir muchos a cavar mi[!] heredad y despues las de los demas compañeros’ (to do something in common, like if many come to till my land and then we go and till everybody else’s land). The forms cited are <tu-chocolaah> ‘he works it in common’, and <tika-chocolaah> ‘we work it in common’.

<tika-chocolaah rih hun ixok> /ti+ qa-chokola7-a-j r-i:j ju:n ixoq/ is glossed ‘hacersele muchos a una’; the Kaqchikel phrase means ‘we all have sex together with one woman’.

This meaning for <chocolaah> ‘to do something in common’ is supported by a further example cited by Varea, who introduces it as follows: “Si uno combida oy a muchos y mañana a otros, o a los mesmos” [If a person invites many people today and tomorrow others, or the same people] <tika-chocolaah he ruvaixic kavay, xaki kalo3obal ki> /ti+ qa-chokola7-a-j je7 ru-wa7i-x-ik qa-way, xa qi qa-loq’-ob’al qi/ ‘we share thus the eating of our tortillas, as well as our prized/bought things’.

COTO (1983 [ca. 1656]:105, 285–286)

<chocola> /chok.ola7/ (nombre) ‘una bebida q[ue] haçen en común, juntando el cacao para ella, en que da cada uno [veinte] granos. Y, despues, lo juntan y muelen, y lo beben en común’ (a drink that they make together, gathering the cacao for it, in which each one gives twenty kernels. After that, they gather it and grind it, and drink it in a group).

<ti ka ban ka chocola> /ti+ qa-b’an qa-chok.ola7/ ‘let us do our group/shared activity’ [Kaufman]; ‘hagamos n[uest]ra junta de cacao’ (let us do our cacao group event).

<ti ka chocolaah> /ti+ qa-chok.ola7-a-j/ ‘let us do it together’ [Kaufman]; ‘juntar así el cacao y beberlo’ (to gather thus the cacao/chocolate and drink it).

<ti ka chocolaah ru kumic k’uqiya> /ti+ qa-chok.ola7-a-j ru-qu:m-i:k q-uk’.i=ya7/ ‘let us do together the drinking of our beverage’.

<hun qu’ix moque vi, yx alabon, yx çamahoma, xa ti moçih ri i vay ti chocolaah> /ju:n k+ ix+ mok.e:7 wi, ix+ alab’-o:n, ix+ samaj-om-a:7, xa t+ i-motzi-j ri+ i-way t+ i-chok.ola7-a-j/ ‘gather together at once you boys, you workers, just gather y’all’s tortillas [and] y’all do it together!’ [Kaufman]. Para deçirles a los trauajadores, o a los muchachos, q[ue] se juntan i coman juntos (In order to tell the workers, or the boys, that they should gather together and eat together).

<ti chocolaah> /t+ i-chok.ola7-a-j/ ‘y’all do it together!’ [Kaufman]; ‘para lleuar entre dos o más vna cosa, carga o vanco’ (to carry/take along a thing, load, or bench with the participation of two or more people).

<ka chocolaam lo3oh> /qa-chok.ola7-a-m loq’-o:j/ ‘we have done buying in common’ [Kaufman].

<ti ka chocolaah ru lo3ic vleu, vacax ..> (o otra cualquier cosa q[ue] compran de común) /ti+ qa-chok.ola7-a-j ru-lo:q’-i:k ule:w, wa:kax, ... / ‘let us do in common the buying of land, cattle, etc.’ [Kaufman].

<chocolaah> [este verbo] lo vsan, también, para yr de común a haçer algo, como a cabar la millpa de algún amigo, y q[ue] allí los del chinamital an juntado para regresarlos (they use this verb also for going as a group to do something, such as digging up the cornfield of some friend, where those of the community have joined together to repay the favor).

<ti ka chocolaah, ti ka mo[tz]ih (r’ih) ka chenoh> /ti+ qa-chok.ola7-a-j, ti+ qa-motzi-j (r-i:j) qa-chen-o:j/ ‘let us do it in common, let us gather (on) our first weeding of the cornfields’ [Kaufman]; ‘juntémonos para labrar n[uest]ras milpas en común, oy la de vno, y otro día de otro’ (let us join together to till our cornfields, today that of one, another day that of another).

<x-qui chocolaah v’ih> /x+ ki-chok.ola7-a-j w-i:j/ ‘they acted in common against me’ [Kaufman]; ‘se an juntado contra mí para haçerme pleito’ (they have gotten together against me in order to pick a fight with me).

<ti ka chocolaah r’ih hun ixok> /ti+ qa-chok.ola7-a-j r-i:j ju:n ixoq/ ‘let us do it [have sex] together against/with a woman’ [Kaufman]; ‘[dicen esto] para deçir q[ue] vna muger es común a todos, o q[ue] todos en común la conosçieron’ (they say this in order to say that a woman is shared by all, or that all had sex with her in a group).

<xaki at qui chocolaam achiha, qahola> /xa qi at+ chok.ola7-a-m achij-a:7, k’ajol-a:7/ ‘the men/boys have simply shared you[r] favors’ [Kaufman]; ‘[dicen esto] para afrentarla’ (they say this in order to insult her).

In Ximénez’s (1993) combined vocabulary of Kaqchikel, K’ichee7, and Tz’utujil (ca. 1700), of the forms cited below, only the first one is conceivably not Kaqchikel:

<chocoh> /chok-o:j/ [sv<vt] ‘bodas o convites’ (wedding or party)

Kaq <chocol> /chok-ol/ [stat<P] ‘estar por orden’ (in order/turn)

Kaq <tin chocola> /ti+ n-chok-ol-a7/ [vt] ‘juntar comida o bebida para comerlo entre muchos’ (to gather food or drink in order to consume it among many)

Kaq <tin chocolaah> /ti+ n-chok-ol-a7-a-j/ [vt] lo mismo, e.g. ‘juntar comida o bebida para comerlo entre muchos’

Brasseur de Bourbourg (1961 [1862]:200) cites for K’ichee7 the root shape <choc> /chok/ as a verb ‘alquilar’ [to hire] and also as a (possibly different) verb ‘llamar, convidar’ [to call, invite]’. Brasseur’s sources are not identified.

Sáenz de Santamaría (1940:97), whose authority is primarily Varea (ca. 1600) (and whose orthography is often garbled), cites:

<choqola, ru> /chok-ol-a7/ sust. ‘banquete popular a que cada uno contribuye con 20 granos de cacao’ (feast of the people to which each [participant] contributes 20 cacao kernels).

<choqolaaj, tu> /t+ u-chok-ol-a7-a-zj/ v. act. ‘convidar a la gente a un banquete de los llamados *choqola*; llamar a la gente para algún trabajo de comunidad’ (to invite the people to a banquet called *choqola*; to call together the people for a communal task).

<choqolaax, ti> /ti+ chok-ol-a7-a-x/ v. pas. ‘ser convidado por el pueblo; ser reprendido por el pueblo’ (to be invited by the [towns]people; to be reprimanded by the [towns]people).

Brinton (1885), whose authorities are unknown, cites <chocola> /chok-ol-a7/ adj ‘in common, communal’.

Tz’utujil of San Juan La Laguna has the transitive verb *chok* ‘encargarlo’ (to commission it, to invite someone to do one a favor) and the corresponding nominalization *chook-ooj* ‘encargo’ (commission). Tz’utujil seems to be the only present-day K’ichee7an language to preserve a reflex of the root {chok}.

Pérez and Hernández (1996:77–78) give these examples:

n+ in-b’e na pa chok-oj n-pantaloon  
‘tengo que ir a encargar mi pantalón’  
‘I have to go to arrange to have some trousers made for me’

ja n-ata7 x+ b’e-r-chok-o7 r-xajajb’  
‘mi papá se fue a encargar sus zapatos’  
‘my father went to arrange to have some shoes made for him’

chi+ b’an-oj jun chook-ooj x+ in-pi wí  
‘vine a hacer un encargo’  
‘I came to make an arrangement to have something done’

Clearly, the noun that lies behind <chocolaah> refers to some sort of shared activity among quite a few people. The noun <chocola> is glossed by Varea as ‘es cacao junto q<ue> dan veinte a cada uno, y lo beben entre todos (this is jointly shared cacao which is given in the quantity of 20 [kernels] to each, and they all drink it [the chocolate] together)’. All that we get from this about chocolate is that Varea called it *cacao* in Spanish around 1600. The reference to chocolate is simply to provide an instance of what a shared item might be; because chocolate was so important a festive drink for so many kinds of occasions, it was apparently a typical or default application of the term /chokola7/, but this term in no way directly expresses the meaning ‘chocolate’.

Varea (ca. 1600) and Coto (1983 [ca. 1656]) both cite <chocola> as a noun, and inform us that its default application is to the consumption of chocolate in a festive context. They also inform us that the verb <chocolaah> refers to several people acting as a group for a common purpose. Brinton’s source for his adjective “communal” has not yet been identified.

We may provide a unitary analysis of all of these forms by starting with a positional root {chok} ‘gathered together’; from this, a transitive verb *chok* ‘to call over’ or ‘to arrange to have somebody do something for one’ is formed, with zero derivation; from the transitive verb *chok*, a nominalization *chok-o:j* ‘invitation’ or ‘arrangement’ is formed; from the root {chok}, a stative adjective *chok-ol* ‘according to turns’ is formed; from the transitive verb *chok*, a derived transitive verb *chok-ola7* ‘to invite to share’ is formed;

from the transitive verb *chok-ola7*, a noun *chok-ola7* ‘a sharing’ is formed; from the noun *chok-ola7*, a transitive verb *chok-ola7-a* ‘to invite to share’ is formed.

This set of words has only an incidental connection with chocolate, and this is because a drink made from cacao was prized and shared at festive occasions.

As for the *real* words for “chocolate” in colonial Kaqchikel, we can cite the following from Coto (1983 [ca. 1656]):

<vqiya> /uk’.i=ya7/ (“drink-ing water”): BEBIDA, generalmente. . . . Tóbase por el chocolate batido, PUTZULE, y otras bebidas, aunq[ue] muchas dellas tiene sus nombres particulares. (uk’.i=ya7: Drink in general. It is taken to be whipped chocolate, pozole, and other drinks, although many of them have their special names.)

<hoqham>. Otra q[ue] haçen de cacao molido y hecho masa para lleuar [a] camino o a sus millpas, y después lo deslíen en agua. (Another [drink] that they make from cacao that is ground and made into dough/paste to carry on the road or to their cornfields, and later they dissolve it in water). The analysis of <hoqham> is unclear.

<3utuh> /q’utu:j/. Otra q[ue] haçen del cacao batido, sacando la manteca, q[ue] es lo q[ue] beben. . . . Ésta sirue a los días de sus fiestas y conbites. (Another [drink] that they make from whipped chocolate, removing the fat, which is what they drink. . . . This [drink] is used on the days of their festivals and invited gatherings.)

<3utuh> /q’utu:j/ o <3utum ya> /q’utu:m ya7/. La bebida que así haçen de cacao batido. (The drink that they make thus from whipped cacao.)

These words are derived from the verb <tin 3ut> /ti+ n-q’ut/ ‘batir con cuchara o con la mano, como baten ellos su bebida de cacao’ (to beat/whip with a spoon or by hand, like they beat/whip their cacao drink).

<pulim ya> /puli:m ya7/. Otra que haçen de cacao. (Another [drink] that they make from cacao.)

This word is derived from the verb <tin pulih> /ti+ n-puli-j/ ‘haçer la tal bebida, o lleuar la masa hecha para desleírla en el camino’ (to make such a drink, or to carry the dough/paste along with one in order to dissolve it [in water] on the road).

<pulim ya>, <pu3um ya> /puq’u:m ya7/. Synónomos deste género de bebida. (Synonyms for this type of drink.)

The second word is derived from the verb <tin pu3> /ti+ n-puq’/, which “sig[nifi]ca lo mesmo q[ue] <tin pulih>.”

<aca puzu>. ‘Bebida, otra q[ue] haçen del cacao seco en las mesmas maçorcas, q[ue] para este fin las guardan sin quebrar’. (A drink, another one that they make from dry cacao in its own pods, that they set aside without breaking for this very purpose.) Coto cites the example <aca puzu chic ru na3 cacao, ti ka kum> /<aca puzu> +chik ru-na:q’ kakaw, ti+ qa-qum/ ‘ya este cacao está bueno y seco, bebámoslo’. (The cacao kernels are “aca puzu”; let’s drink them). The analysis of <aca puzu> is not clear.

From these examples, it is clear that colonial Kaqchikel had several words for drinks made from cacao, all of them transparent as to morphological formation, and none of them similar to words in non-K’ichee7an languages.

Table 9. Cacao-pataxte couplets in colonial K'ichee7an ritual texts.

From the Po:pol Wu:j

Passage A, where the hero twins Ju:n Aj=B'uh and Ix=B'a:l=ke:j are repaying the rat for some useful news.

Christenson (2004:lines 3237–3244)

<ta xquiya cut recha cho are cui recha ri ixim zaquil ic, quinaç, pec, cacou>  
ta x+ ki-ya7 +k'ut r-e:cha:7 ch'o:h,  
are7 +k'u ri r-e:cha:7;  
ri+ ixi:m, saki:l;  
i:k, kinaq';  
pe:q, kakow.

Then therefore they gave the rat his food,  
they therefore [gave him] his food;  
the maize, [and the] squash seeds;  
[the] chilli [and the] beans;

[the] *Theobroma bicolor* [and the] *Theobroma cacao*

Passage B, describing the creation of mankind.

Christenson (2004:lines 4892–4903)

<quehe cut xequicot vi rumal ri vriquitahic vtzilah huyub nohinac chi quz  
tzatz chi 3ana hal zaquihal tzatz naipuch chipec chi caco maui ahilan tulul  
caux quinom tapal, ahache cab>

keje7 +k'ut x+ e7-ki7.kot wih  
r-uma:l ri+ u-riq-itaj-i:k utz + ilaj juyub',  
no:j-inaq chi+ kus,  
tzatz chi+ q'an + a jal, saq + i jal,  
tzatz nay puch chi+ pe:q, chi+ kakow  
ma +wi ajila:n tulul, k'awex,  
q'inom, tapa7l,  
aja=che7, ka:b'.

Thus therefore they rejoiced  
because of the finding of good mountains  
filled with “kus/deliciousness”,  
thick with yellow ears of maize, white ears of maize,  
thick also with pataxte (*Theobroma bicolor*) and with cacao (*Theobroma cacao*);  
not counted are zapote/marmalade fruit (*Pouteria mammosa*), anona/  
sweetsop (*Annona squamosa*),  
jocote/hogplum (*Spondias* spp.),  
nance/pickle tree (*Byrsinoma crassifolia*),  
matasano/white sapote (*Casimiroa edulis*), and honey.

From the Rab'ina:l Achih

Passage A, part of a commentary by the Aj=Ux-a:b' and the

Aj=Poqom-a:b' to Oye:w Achih, the Wrathful Man:

Breton (1994:lines 600–623 [Folios 16–17])

< ] manare varal coh... //  
etar. vi. quc val.. que. nuqahol co vi... //  
xe V3 3azutz xeza3 3izutz chima[t]ch... //  
mal teu chimachmal horon... //  
chuxe umuhibal raxon chuxe 3anal //  
pe3. u3anal qaco.o chuxe 3aná puva3... //  
chuxe za3ipuva3 xepich xeccat quc... //  
val.. quc nucahol arena rival arenu... //  
qahol mana qoqax caquiri3 vi huper... //  
cha3ap tzu3bal que xaetzulic qoc ulo3 //  
hoqal pe3 hoqal qacoo rumal eah... //

pich eahcot chuxe 3ih chuxeza3.>  
ma +na re wara:l k+ oj-et-ar wih  
k-u:k' w-a:l, k-u:k' nu-k'ajo:l  
k'oh wih  
xe:7 [q' ]eq+a su:tz', xe:7 saq+i su:tz'

Table 9. Continued

chi+ mach-m-al te:w, chi+ mach-m-al joron  
ch+ u-xe:7 u-muj-ib'al rax-o:n  
ch+ u-xe:7 [u-]q'an-a:l pe:q, u-q'an-a:l kakow  
ch+ u-xe:7 q'an-a puwaq, ch+ u-xe:7 saq-i puwaq  
x+ e7-pi:ch x+ e7-k'o:t  
k-u:k' w-a:l, k-u:k' nu-k'ajo:l  
are7 na ri w-a:l, are nu-k'ajo:l  
ma +na k'oh k'ax ka+ ki-riq wih  
ju=pe:r ch'aca7p tzuq.b'al k-eh  
xa e7 tz'u7l-ik k+ o:k ul-oq  
jo7=k'a:l pe:q, jo7=k'a:l kakow  
r-uma:l e7 aj-pich, e7 aj-k'ot  
ch+ u-xe:7 q'i:j, ch+ u-xe:7 saq

Is it not right here that we are established  
with my male('s) offsprings and with my female('s) offsprings  
who are under the black clouds [and] the white clouds  
with the tingling frost [and] with the tingling cold  
under the shade of the green feathers  
under the ripeness of the pataxte [and] the ripeness of the cacao  
under the gold [and] under the silver  
they are bedecked [and] they are engraved  
with my female('s) offsprings [and] with my male('s) offsprings?  
They are my female('s) offsprings [and] my male('s) offsprings.  
There is no difficulty for them to find one piece,  
a remainder of food for them,  
they are just becoming tough,  
a hundred [kernels of] pataxte [and] a hundred [kernels of] cacao,  
because they are bedeckers [and] they are engravers  
under the sun [and] under the dawn.

Passage B, in which K'ichee7 Achih speaks for the thirteenth time:

Breton (1994:lines 2173–2201 [Folio 53])

< . . . qachachibarinutzih chuvach. qah.. //  
chuvach ulev! la are vaibal la vá 3atzare. //  
uholom nu3ahau qanuvilo qanumu3uh. la... //  
mana. roquicam quehetabanel tza3anitza- //  
nel. chirech vauba3il nuvi uba3il nuholom qo- //  
tmtachrih qotimtachuvach tachi3ah apano3. //

chinuhuyubal. chinuta3ahal tza3atitza- //  
bal tare oqal. pe3. oqal qaqó. qumal. //  
val. qumal. nuqahol chinuhuyubal. //  
chinuta3ahal . . . >  
ka+ cha7 chi+ b'a ri+ nu-tzi:j  
ch+ u-wach ka:j, ch+ u-wach ule:w  
la+ are7 wa7.b'al la+ wa7  
qatz are7 u-jolo:m nu-qaja:w  
ka+ nuw-il-o, ka+ nu-muqu:-j  
la+ ma +na r-okika:m  
keje7 ta  
b'an.e:l, tz'aqan.isa.n.e:l  
chi+ r-e:ch wa7  
u-b'aq-i:l nu-wi7, u-b'aq-i:l nu-jolo:m  
k'oti:-m ta chi+ r-i:j,  
k'oti:-m ta ch+ u-wach  
ta chi+ qa:j apan-oq  
chi+ nu-juyub'-a:l, chi+ nu-ta'qaj-a:l  
tz'aqat.isa.b'al ta r-eh  
jo7=k'a:l pe:q, jo7=k'a:l kakow  
k-uma:l w-a:l, k-uma:l nu-k'ajo:l  
chi+ nu-juyub'-a:l, chi+ nu-taq'aj-a:l

Continued

Continued

Table 9. Continued

Here is what my word says  
on the face of the sky [and] on the face of the earth:  
Now this dish of yours—  
it is the skull of my father.  
I see it, I look at it.  
Now it would not be fitting  
in thus being made [and] being fulfilled  
for this, the bone of my head [and] the bone of my skull  
to be removed from in back,  
and to be removed from in front  
that it should go down  
in my mountain [and] in my plain  
as the fulfilment of  
a hundred [kernels of] pataxte [and] a hundred [kernels of] cacao  
because of my female('s) offsprings [and] because of my male('s) offsprings  
in my mountain [and] in my plain.

From the Annals of the Kaqchikels, Chapter 6 (Maxwell and Hill 2006)  
<nabey na xu4am rikan vukama3>  
Nab'e:y na x+ u-k'am r-i:qa.7n wu:q=ama:q'.  
First, he bestowed the burden of the seven tribes.

<4ate 4a xu4am chic rikan ahlalab>  
K'ate k'a x+ u-k'am chik r-i:qa.7n aj=lab'.a:l.  
At once, then, he bestowed the burden of the warriors.

Continued

## CACAO AND PATAXTE

Two different kinds of cacao were used to make drinks in pre-Columbian times: the cultivated *Theobroma cacao*, called *kakawa-tli* in Nawa, *kakow* (earlier) and *kakoh* (later) in K'ichee7, *kakaw* in Q'eqchi7 and in Yukateko, and *cacao* or *cacahua* in Spanish; and the wild *Theobroma bicolor*, called *kwaw=patlach-tli* and/or *kakawa=patlach-tli* in Nawa, *pe:q* in K'ichee7, *kakaw b'a:lam~b'a:lam kakaw* in Q'eqchi7, *b'ahlam kakaw* in Epigraphic Mayan, *b'aHlam=te7* in Yukateko, *\*b'ahläm=te7* in proto-Ch'olan (with descendants in Yokot'an, Ch'ol and Ch'olti7), and *pataxte* or *pataste* in Spanish.

Both Spanish terms come from Nawa. *Pataste~pataxte* seems as if it would come from a Nawa form *patlach-tli*, but no such simple form seems to be attested in Nawa. However, Pipil provides *pa(:)tach* 'pataxte' (Campbell 1985:380, 771), and *ku:patach* (Campbell 1985:297), which is cognate with Hernández's <quauhpatlachtli>.

Given that the Nawa word for *Theobroma bicolor*, *kwaw=patlach-tli*, contains the stem {patlach} 'flat, flattened, wide', it is worthy of note that the Mayan word *\*pe:q* '*Theobroma bicolor*' (and possibly also *Theobroma cacao*) may have as its root a form *\*peq* 'flat', as suggested by the Yukateko, Kotoke, and Tzeltal forms cited below. A Mayan etymon *\*peq* 'flat' is not in general well supported, but *sapo* 'toad' (in proto-Mayan, *\*peq*) is widely used in Mesoamerican Spanish as a metaphor for squat (low and wide) people.

YUKATEKO (CIUDAD REAL 2001 [ca. 1590]:485–486 [MOTUL DICTIONARY]).

<pec> cuenta para cosas redondas, circulares, como hostias, panes, tortillas

<peca<a>n> cosa puesta de plano o e llano, y no de lado ni en pie, y lo llano de la espada o cuchillo, etc.

<pec cab>~<pec cabal> cosa puesta de llano

Table 9. Continued

<xa4a ruyon xit puak 3u3uraxon 4ubul chactit>  
Xa k'a ru-yo:n xi:t, puwaq, q'u:q'-u ra-x.o:n, k'ub'u:l, #chak=tit  
Just jade, [precious] metal, quetzal feathers, trogon feathers [and] ?red  
?feathers

<ru4in 4a 4,ibanic 4otonic quiyanic>  
r-uk'i:n k'a tz'ib'a-n-ik, k'ot-on-ik, kiy-an-ik;  
along with writings, carvings, weavings;

<xul, bix, 4hol gih, may 3ih pek cacouh>  
xu:l, b'i:x; ch'ol=q'i:j, may=q'i:j; pe:q, kakow;  
flutes, songs; 260-day calendars, solar calendars; pataxte, cacao;

<xa ruyon 3inomal xrikah pe pa Tullan>  
xa ru-yo:n q'inom.:al x+ r-i:qa-j pe pa Tula:n.  
just riches they carried forth from Tulan.

(In this passage, *tula:n* and *xi:t* are loans from Nawa *to:lla:n* and *xiwitl* (> Gulf Nawa *xi:t*), respectively; *kakow* is a loan from Mije-Sokean; #*chaktit* is a loan from somewhere not yet determined, possibly within Mayan.)

A second passage is cited in Maxwell and Hill (2006) that is not found in the material published in Brinton 1885 or Otzoy 1999:

Ma:=ni ju:n pe:q, ma:=ni ju:n kakow, x+ u-ya7 Xajil-a:7.  
Not one kernel of pataxte, not one kernel of cacao did the Xajils give

YUKATEKO (BRICKER ET AL. 1998:212)

*pek* [T] vt to fold, to hem

*pek* [P] pv to stretch out at full length

MOCHÓ [KAUFMAN DATABASE (1967-1968)]

*peq-An* 'no bien plomeado'

TUSANTEKO (KAUFMAN 1967–1968)

*pe:q* 'hoja para envolver tamalitos' (such leaves are always flat and broad, like those of banana/plantain or Heliconia)

TZELTAL COPANABASTLA (ARA 1620:85V)

<pecan> /pe{h}k.an/ act. 'poner llano algo'

<pequel> /pek.el/ 'puel'to assí'

From the descriptions in Nawa and Nawa-oriented sources cited in the previous section, we know that sometimes both types of *Theobroma* were combined in a single drink. References to pataxte are regularly paired with references to cacao in the colonial poetic and ritual texts in which we have sought this term (Kufner and McNeil 2006:99). K'ichee7 and Kaqchikel poetic and ritual texts often pair *kakow* and *pe:q* (see Table 9). In these passages, *pe:q* and *kakow* are always named together, with *pe:q* coming first. Their uses are not specified (but Tedlock [2002] argues from context and from ethnographic considerations that some such references relate to drinking chocolate at a wedding or betrothal). The only example we have found in such texts in which this is *not* the case is in a proper name, /i)x=kaka:w/ 'Lady (or Small) Cacao'. (The independent noun meaning 'cacao' is /kakow/; the form /kaka:w/ corresponds to what would be the possessed form of /kakow/, though here it does not have that function. The Yukateko term for pataxte, *b'aHlam=te7*, is found [once] in the Book of Chilam Balam of Chumayel (Roys 1933:36, 111) paired

with *kakaw*: <cabal chac bolay balam y cacao balamte>, the point of which, however, is obscure.

## CULTURE HISTORICAL INFERENCES

A word pronounced something like *kakaw* or *kakawa* was borrowed widely, into Mayan languages in southern Mesoamerica, into some languages of lower Central America, and into several languages in and near the Basin of Mexico; and it was borrowed early, probably between 200 B.C. and A.D. 400, into a Lowland Mayan language. Many other words that are reconstructible within Mije-Sokean for culturally important cultigens spread widely in Mesoamerica. A Mije-Sokean origin for the diffusion of this term fits into what is known of the diffusion of such terms in Mesoamerica and is characteristic of no other language family in Preclassic or Classic Mesoamerica.

This study has shown that *\*kakaw(a)* has an unimpeachable pedigree as a native Mije-Sokean word. In particular, Gulf Sokean cognates reflecting initial stress are consistent with the stress patterns on all other trisyllabic roots in all Gulf Sokean languages. Wichmann's arguments against the reconstructibility of this word to proto-Sokean and proto-Mije-Sokean are vitiated by being based on a very incomplete set of data on Sokean trisyllabic roots and by a flawed analysis of the data on stress in Gulf Sokean cognates of proto-Sokean trisyllabic roots. The linguistic evidence is unambiguous: there is no viable alternative to a Mije-Sokean origin for this term.

On the face of it, Dakin's counterproposal that Nawa *kakawa-tl* is the source of the word *kakawa* and *kakaw* in other Mesoamerican languages is implausible. Nawa nouns such as *kakawa* that take the absolute suffix *-tl* when unpossessed normally show up, when borrowed into other Mesoamerican languages, with final *-t* (when the borrowing language tolerates word-final /t/), as in Sotepan Gulf Sokean [7a:ttébet] /7aattep7et/ for 'town' from Nawatl *a:-l=tepe:-tl*. Besides the putative case of *kakawa*, no individual Nawa loan-word that was widely diffused in Mesoamerica and that takes the *-tl(i)* suffix in Nawa is characteristically found without a reflex of this suffix in the borrowed form (just one such loan is widely found among Mayan languages), yet this word for cacao was *never* borrowed with it. In addition, no Nawa loan word clearly predates the end of the Classic period, and clearly there is not an early body of loans of cultigens from Nawa.

Especially strong evidence is needed in such a situation to establish a case for diffusion from Nawatl, but such evidence is not forthcoming. Dakin and Wichmann present an argument that the Nawa term descends by reduplication from a Southern Yuta-Nawan root *\*kava* 'egg', reconstructible for the Sonoran, though not for the Nawa, branch of Southern Yuta-Nawan. But there is no Nawa form *kawa-tl\** to undergo reduplication, and we show that a Southern Yuta-Nawan form *\*kava* cannot possibly have yielded *kawa* in Nawa (rather, it would have yielded *ka:*). This Nawa etymology for the term is not simply implausible; it is invalid. Dakin and Wichmann's proposed Nawa origin for this term must be rejected.

What, then, was the cultural context of the spread of the word *kakaw(a)* in Mesoamerica? Dakin and Wichmann (2000:67–68) associate it with the Teotihuacan diffusion sphere, but their reasoning is untenable. They begin with the erroneous assumption that it diffused from Nawa, and this assumption vitiates their entire argument. The next two premises are correct. Temporally, the word had certainly entered a Lowland Mayan language before A.D. 400, the time of its first known attestation in hieroglyphic texts. As for

geography, Dakin and Wichmann assume that Nawas were in or near the Basin of Mexico during the Early Classic period. We agree with this, but Kaufman (1994–2004/2007 and above in the section “Demonstrating Borrowing”) has demonstrated from the impact of Mije-Sokean, Totonakan, and Wastekan on the vocabulary and grammatical structure of proto-Nawa that it was an ancestor of proto-Nawa that first had a homeland in northern Mesoamerica, while Dakin and Wichmann suppose that it was Eastern Nawa that first arrived there. Drawing these three premises together, Dakin and Wichmann seek a cultural origin that was prominent enough before A.D. 400 to be responsible for the diffusion of cacao and its associated vocabulary and that was in the core area of Nawa occupation of Mesoamerica. They see only Teotihuacan as meeting these criteria; based on this, they identify Nawa as the dominant language of Teotihuacan and Teotihuacan influence as the vector for the spread of the word for cacao throughout Mesoamerica.

Without the prop of the demonstrably false premise that the ultimate source of the word *kakawa* in other Mesoamerican languages was Nawa, Dakin and Wichmann's entire argument fails, because it is the origin of the word among Nawa speakers that provides the geographic localization. But even had their proposed etymology for the origin of *kakawa* been viable, it would not have been enough to make a case for Nawas being major players at Teotihuacan. It takes a body of evidence, not a single loan-word—even if the arguments for the borrowing were methodologically unproblematic—to provide believable argumentation regarding the linguistic identity of a prehistoric culture. In the case of Teotihuacan, the language or languages of its elite classes must have had a serious impact on the vocabulary, and potentially on the grammar and pronunciation, of many other languages in Mesoamerica during the period from A.D. 100–500. This means that there should be a substantial number of early loan-words into languages in and around the Basin of Mexico and a substantial but smaller number farther afield—in particular, in Mayan languages, around Kaminaljuyu, and along the Pacific coast of Guatemala. As has long been known, Nawa languages had no such impact until several centuries later; the evidence is summarized in the section “Evidence Against a Recent Alternative Hypothesis”. Accordingly, Nawas could not at that time have been culturally influential in Mesoamerica, and whether or not there were Nawas living in and around Teotihuacan, they could not have played a major political, economic, or religious role in the city's public affairs.

Nonetheless, like Dakin and Wichmann but for different reasons, we also associate at least part of this diffusion with the regional influence of Teotihuacan. We now know that the Mayan borrowing of the word cannot be put back as early as the Olmec era, as proposed by Campbell and Kaufman (1976), but, rather, to the Epi-Olmec era—to the Late Preclassic or the Early Classic period. Given their timing, the loans into Mayan could indeed be associated with Teotihuacanos, and thus with the northern branch of Mije-Sokean, but the Epi-Olmecs are a viable alternative, since they, but not the Teotihuacanos, lived in or near areas where cacao was cultivated.

For the loans in and around the Basin of Mexico, however, a Teotihuacano source is very likely. Kaufman (2000–2007, 2001; Kaufman and Justeson 2007) shows that there was a massive diffusion of Mije-Sokean vocabulary into languages in the Basin of Mexico and its immediate surroundings (Figure 6). The borrowing into Totonakan was truly massive, about 50 items. By current count, between eight and 17 Mije-Sokean words were borrowed into each of Nawa, Tarasko, Otomian, Matlatzinkan, and



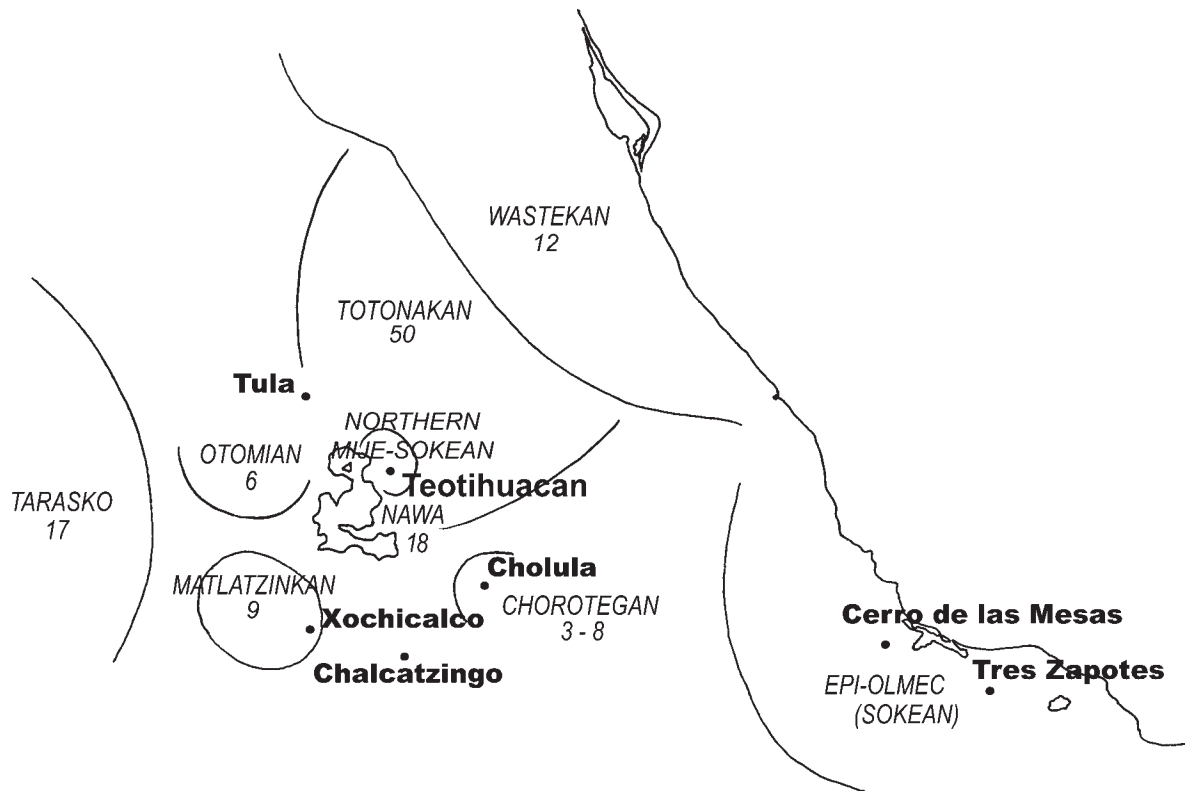


Figure 6. Numbers of Mije-Sokean loan-words into languages of northern Mesoamerica, showing estimated locations of these languages around A.D. 500. The inferred region of Northern Mije-Sokean also included speakers of Totonakan, which surrounds it; the localization of Nawa, which probably arrived in Mesoamerica during the Early Classic period, is less secure than the locations of the other groups. Matlatzinkan becomes Matlatzinka and Tlawika (Ocuilteko); Otomian becomes Otomí and Masawa; Totonakan becomes Totonako and Tepewa. After Kaufman and Justeson [2007:Figure 3].

possibly Chorotegan. Farther afield, 11 words were borrowed into Wasteko.

From the locations of the languages with the greatest numbers of loans, the center of this diffusion can be localized among or adjacent to Totonakans and more involved with speakers of Tarasko and Nawa than with speakers of Wasteko. This places them in their greatest concentration in or near the eastern half of the Basin of Mexico—thus, in the vicinity of Teotihuacan—and also, probably, throughout the southern half of the Basin.

This geographic analysis of the northern Mije-Sokean loans leads us to propose that one of the languages of Teotihuacan was a “northern” branch of the Mije-Sokean family. It probably left Olmec country no later than the time of the separation of Mijejan and Sokean from one another (ca. 1000 B.C.; see “Evidence Against a Recent Alternative Hypothesis”, point 6), since the loans now unique either to Mijejan or to Sokean are proportionally about equal (10–20%). These immigrant populations may be recognized at Early Preclassic sites in the Basin of Mexico, beginning around 1200 B.C. At Coapexco, in the southeast, they occur in all contexts and all functional components of the artifact assemblage, including utilitarian artifacts (Tolstoy 1989:98). Tolstoy (1989:98) makes their immigrant status clear, stating that the Olmec features “(1) appear suddenly; (2) appear early; (3) appear together; (4) pervade general refuse, all households, and many sectors of activity; and (5) seem most abundant at the time of their first appearance. Their subsequent history, in fact, is one of fairly rapid fading or transformation and replacement by new elements. . . .”

In spite of their assimilation to local material-culture practices, these Mije-Sokean speakers evidently remained linguistically and probably socially distinct. Centuries later, at Teotihuacan, the loan-word evidence suggests that they were the elite at the site and probably coexisted there with speakers of a Totonakan language. (Mije-Sokean had a more massive impact on Totonakan than on any other language or language group in Mesoamerica.)

One of the words that was diffused into several of these languages was *kakawa*. It shows up in Totonakan, Nawa, Tarasko, and Masawa.

The word *kakawa*, then, is a quintessential representative of the distribution of Mije-Sokean loans into Mesoamerican languages: with substantial borrowing into Mayan languages in the south; few borrowings by Oto-Mangean languages in Oaxaca; and borrowing into several languages in and around the Basin of Mexico. This pattern provides further support for the Mije-Sokean origin of this term. Given the localization of the center of diffusion of northern Mije-Sokean loans, it is quite probable that the word *kakawa* diffused in this area in association with the regional influence of Teotihuacan. A detailed account is provided elsewhere (see Kaufman 2007; Kaufman and Justeson 2007).

This study has also shown that the same word, in the form \**kakaw*, diffused from southern Mesoamerica into lower Central America, where it underwent a series of modifications characteristic of the phonologies of the borrowing languages. On geographical grounds, the most likely proximate source of this borrowing was Mayan.

The linguistic data tell us that it was speakers of Mije-Sokean languages who were influential in the diffusion of the word for cacao throughout Mesoamerica, both in the north and in the south. They do not, however, inform us on the nature of the intercultural interaction that was the basis for foreigners' adoption of this word. It is plausible that the term diffused in association with the cultivation of cacao (cf. Justeson et al. 1985:59); in the Mayan case, this would account for the preservation in K'ichee7an of an ancient word \**pe:q* in reference to uncultivated cacao. However, it is also possible, especially in northern Mesoamerica, that the word diffused in connection with the processing of cacao or, more likely, with a rising importance of its use—perhaps in a ritual context or perhaps through an economic importance, for example, as money.

Linguistic analysis also demonstrates that proposals for a partly Mayan origin of the Nawa word *chokola:tl*~*chikola:tl* are untenable, and, in agreement with Dakin and Wichmann (2000), that

the term almost certainly originated within Nawa. It is, however, implausible that it could have diffused along with the word \**kakaw(a)*. Forms based on *chokola:tl* and *chikola:tl* are found in few indigenous languages and only in limited dialects of them. In fact, the term may not have existed in pre-Columbian times, as it is unattested in compendious sources, such as Molina's *Vocabulario* and Sahagún's *Primeros Memoriales* and *Historia general*, before 1577. Until this time, and still today in many languages, the word for cacao was also used for drinks made from it.

The uncultivated pataxte (*Theobroma bicolor*) was on certain occasions used together with cacao (*Theobroma cacao*)—at least, in proto-historic central Mexico (e.g., Hernández's second drink made from cacao, described above)—and is textually associated with cacao in Highland Mayan literary contexts. Several varieties of cacao proper (*Theobroma cacao*), distinct from pataxte, were known and distinguished lexically in colonial Nawa sources and probably also in Epigraphic Mayan.

## RESUMEN

La palabra \**kakaw(a)* ('cacao', *Theobroma cacao*) se había difundido ampliamente entre las lenguas mesoamericanas precolombinas, y de Mesoamérica a la Centroamérica inferior.

Este estudio ofrece evidencias que establecen sin duda razonable que esta palabra tiene su origen en la familia lingüística mixe-zoqueana—que de las lenguas mixe-zoqueanas en el hogar de los olmecas se extendió a otras lenguas del sureste de Mesoamérica, y a algunas lenguas mayances entre 200 a.C. y 400 d.C., y que se extendió desde una lengua mixe-zoqueana hablada en la Cuenca de México hasta en otras lengua de esa región.

Este estudio demuestra que cada uno de los argumentos ofrecidos por Dakin y Wichmann (2000) en contra de un origen mixe-zoqueana o no funciona, o se basa en conceptos falsos, o le falta relevancia, y que la alternativa que pro-

ponen ellos—que originó en el nahua y que del nahua se extendió en otras lenguas mesoamericanas—está en desacuerdo con la preponderancia de las evidencias relevantes al asunto.

Este estudio también discute los detalles lingüísticos de terminología relacionada a bebidas hechas de cacao; demuestra que ninguna etimología propuesta para la palabra "chocolate" es correcta, pero está de acuerdo con Dakin y discute la historia de palabras para 'pataxte' (*Theobroma bicolor*) y sus usos.

Los datos lingüísticos son relevantes a cuestiones de interacción entre grupos etnolingüísticos en tiempos precolombinos, pero no revelan la naturaleza del contexto cultural de la difusión del cacao en Mesoamérica ni de sus usos.

## ACKNOWLEDGMENTS

We thank the many archaeologists and linguists who asked us to write this paper. Una Canger and Thomas Smith-Stark read an earlier draft and offered helpful comments; Gabriela Pérez Báez read a nearly final version and offered helpful comments; Judie Maxwell provided us with access to a pre-publication draft of her edition of the *Annals of the Kaqchikels* (Maxwell and Hill 2006). Louise Burkhart helped us in tracking down sixteenth century

Nawa references to chocolate drinks. We thank three anonymous reviewers, who took us to task for the lack of detail in some of our argumentation, which we have dutifully attempted to rectify.

We thank William Fowler, Cameron McNeil, and John Byram for permission to publish a summary and extract of this study as chapter 6 of *Cacao in Mesoamerica: a Cultural History of Cacao* (McNeil 2006).

## ABBREVIATIONS

YN	Yuta-Nawan
pSYN	proto-Southern Yuta-Nawan
*ABC	a reconstructed form
ABC*	a non-occurring incorrect form
H	a laryngeal ( <i>h</i> or 7)

+abc, abc+	a clitic (enclitic, proclitic)
–abc, –abc	an inflexional affix (suffix, prefix)
.abc, abc.	a derivational affix (suffix, prefix)
=abc, abc=	a bound root (postpound or prepound)
{abc}	a morpheme

## REFERENCES

- Acosta, Joseph (José) de  
1590 *Historia natural y moral de las Indias*. Casa de Iuan de León, Seville.
- Anderson, Arthur J.O., and Charles E. Dibble (translators)  
1951–1982 *Florentine Codex: General History of the Things of New Spain*, 13 vols. University of Utah Press, Salt Lake City.
- Andrews, J. Richard, and Ross Hassig (translators and editors)  
1984 *Treatise on the Heathen Superstitions that Today Live among the Indians Native to this New Spain, 1629*, by Hernando Ruíz de Alarcón. University of Oklahoma Press, Norman.
- Ara, Fray Domingo de  
1620 *Vocabulario de lengua tzeldal segun el orden de Copanabastla*,

- facsimile ed. (1986), edited by Mario Humberto Ruz. Universidad Nacional Autónoma de México, Mexico City.
- Beltrán de Santa Rosa, Fray Pedro  
1746 *Arte del idioma maya reducido a sucintas reglas y semilexicon yucateco*. Viuda de D. Joseph Bernardo de Hogal, Mexico City.
- Bierhorst, John  
1985 *A Nahuatl-English Dictionary and Concordance to the Cantares Mexicanos*. Stanford University Press, Stanford.
- Brasseur de Bourbourg, Charles Étienne  
1961[1862] *Gramática de la Lengua Quiché*. Editorial del Ministerio de Educación Pública, Guatemala City. [originally published in 1862 as *Gramática de la lengua Quiché/Grammaire de la langue Quiché... avec un vocabulaire... Collection de Documents dans les Langues Indigènes*, Paris]
- Breton, Alain  
1994 *Rabinal Achi: Un drame dynastique maya du quinzième siècle (Édition établie d'après le Manuscrit Pérez)*. Société des Américanistes and Société d'Ethnologie, Nanterre.
- Bricker, Victoria, Eleuterio Po'ot Yah, and Ofelia Dzul de Po'ot  
1998 *A Dictionary of the Maya Language as Spoken in Hocabá, Yucatán*. University of Utah Press, Salt Lake City.
- Brinton, Daniel G.  
1885 *The Annals of the Cakchiquels: The Original Text with Translation, Notes, and Introduction*. Library of Aboriginal American Literature, no. 6. n.p., Philadelphia.
- Campbell, Lyle  
1985 *The Pipil Language of El Salvador*. Mouton, Berlin.  
1988 *The Linguistics of Southeastern Chiapas*. New World Archaeological Foundation Papers, 50. New World Archaeological Foundation, Provo, UT.
- Campbell, Lyle, and Terence Kaufman  
1976 A Linguistic Look at the Olmecs. *American Antiquity* 41:80–89.
- Campbell, Lyle, and Ronald W. Langacker  
1978 Proto-Aztec Vowels. *International Journal of American Linguistics* 44:85–102, 197–210, 262–279.
- Canger, Una  
1969–1970 San Quintín Lakantun slip file. 1435 entries. (Electronic version 1995).
- Caso, Alfonso  
1947 *Calendario e escritura en las antiguas culturas de Monte Albán*. Talleres Gráficos de la Nación, Mexico City.
- Christenson, Allen  
2004 *Popol Vuh, Volume II: Literal Poetic Version*. O Books, Winchester, England.
- Ciudad Real, Antonio de  
2001 [ca. 1590] *Calepino maya de Motul: edición crítica y anotado por René Acuña*. Plaza de Valdés, Mexico City.
- Clark, John E.  
2000 Los pueblos de Chiapas en el formativo. In *Las culturas de Chiapas en el periodo prehispánico*, edited by Dúrdica Šeqota, pp. 37–59. Consejo Estatal para la Cultura y las Artes de Chiapas, Tuxtla Gutierrez, Chiapas. Mexico.
- Clavijero, Francisco Xavier (“Francesco Saverio Clavigero”)  
1780 *Storia Antica del Messico*. A Spanish version in 4 volumes was published in 1826 as *Historia Antigua de México* (or Méjico).
- Coe, Sophie D., and Michael D. Coe  
1996 *The True History of Chocolate*. Thames and Hudson, New York.
- Coe, Michael D., and Gordon Whittaker  
1982 *Aztec Sorcerers in Seventeenth Century Mexico: The Treatise on Superstitions by Hernando Ruiz de Alarcón*. Institute for Mesoamerican Studies Publication No. 7. Institute for Mesoamerican Studies, Albany, NY.
- Constenla, Adolfo  
1981 *Comparative Chibchan Phonology*. Unpublished Ph.D. dissertation, Department of Linguistics, University of Pennsylvania, Philadelphia.
- Corominas, Joan (Joan Coromines)  
1980–1983 *Diccionario crítico etimológico castellano e hispánico*. Gredos, Madrid.
- Coto, Tomás de  
1983 (1656) *Vocabulario de la Lengua Cakchiquel v[el] Guatimalteca, nuevamente hecho y recopilado con summo estudio, traujo y erudición*, edited by René Acuña. Universidad Nacional Autónoma de México, Mexico City.
- Dakin, Karen, and Søren Wichmann  
2000 Cacao and Chocolate: A Uto-Aztec Perspective. *Ancient Mesoamerica* 11:55–75.
- Dávila Garibi, Ignacio  
1939 *Nuevo y más amplio estudio etimológico del vocablo chocolate y de otros que con él se relacionan*. Emilio Pardo y Hijos, Mexico City.
- Drucker, Philip  
1942 *Ceramic Stratigraphy at Cerro de las Mesas, Veracruz, Mexico*. Bureau of American Ethnology Bulletin 141. Smithsonian Institution, Washington, DC.
- Durán, Fray Diego  
1965 *Historia de las Indias de Nueva España, e islas de la tierra firme*, 2 vols., edited by Ángel María Garibay Kintana. Editorial Porrú, Mexico City.
- Dütting, Dieter  
1992 The Hieroglyphic Texts of Chichen-Itza. *Baessler-Archiv* 41: 167–205.
- Edmonson, Munro  
1982 *The Ancient Future of the Itza: The Book of Chilam Balam of Tizimin*. University of Texas Press, Austin.
- Engel, Ralph, and Mary Allhiser de Engel  
1987 *Diccionario zoque de Francisco León*. Instituto Lingüístico de Verano, Mexico City.
- Evans, Susan Toby  
2004 *Ancient Mexico and Central America: Archaeology and Culture History*. Thames and Hudson, New York.
- Fernández de Miranda, María Teresa, and Roberto J. Weitlaner  
1961 Sobre algunas relaciones de la familia mangué. *Anthropological Linguistics* 3(7):1–99.
- Fowler, William R.  
2006 Cacao Production, Tribute, and Wealth in Sixteenth-Century Izalcos, El Salvador. In *Chocolate in Mesoamerica: A Cultural History of Cacao*, edited by Cameron L. McNeil, pp. 307–321. University Press of Florida, Gainesville.
- Fox, James A., and John Justeson  
1982 A Cholan Calendar in the Gates Collection. Manuscript on file, Department of Anthropology, University at Albany, Albany, NY.
- Garibay Kintana, Ángel María  
1993 *Poesía náhuatl I: Romances de los señores de Nueva España; Manuscrito de Juan Bautista de Pomar, Tezcoco 1582*. Universidad Nacional Autónoma de México, Mexico City.
- Gunsenheimer, Antje  
2001 *Geschichtsradiierung in den yukatekischen Chilam Balam-Büchern: Eine Analyse der Herkunft und Entwicklung ausgewählter historischer Berichte*. Inaugural-Dissertation, Rheinischen Friedrich-Wilhelms Universität zu Bonn.
- Haeserijn V., Padre Esteban  
1979 *Diccionario k'ekchi'-español*. Editorial “Piedra Santa”, Guatemala City.
- Harrison, Roy, and Margaret B. de Harrison  
1984 *Vocabulario zoque de Rayón*. Instituto Lingüístico de Verano, Mexico City.
- Henderson, John, and Rosemary A. Joyce  
2006 Brewing Distinction: The Development of Cacao Beverages in Formative Mesoamerica. In *Chocolate in Mesoamerica: A Cultural History of Cacao*, edited by Cameron L. McNeil, pp. 140–153. University Press of Florida, Gainesville.
- Hernández, Francisco  
1959 (1577) *Historia natural de Nueva España*, 2 vols. Universidad Nacional Autónoma de México, Mexico City (original in Latin).
- Hill, Jane  
2001 Proto-Uto-Aztec: A Community of Cultivators in Central Mexico? *American Anthropologist* 103:913–934
- Hoffling, Charles Andrew, with Felix Fernando Tesucún  
1997 *Itzaj Maya-Spanish-English Dictionary/Diccionario Maya Itzaj-Español-Ingles*. University of Utah Press, Salt Lake City.
- Holt, Dennis  
1999a *Pech (Paya)*. LINCOM Europa, Munich.  
1999b *Tol (Jicaque)*. LINCOM Europa, Munich.
- Hurst, W. Jeffrey, Stanley M. Tarka Jr., Terry G. Powis, Fred Valdez Jr., and Thomas R. Hester  
2002 Cacao Usage by the Earliest Maya Civilization. *Nature* 418: 289–290.

- Justeson, John  
1985 Yucatecan Phonological History. Manuscript on file, Department of Anthropology, University at Albany, Albany, NY.
- Justeson, John., and Terrence Kaufman  
1996–2001 Warfare in the Preclassic Inscriptions of Monte Alban. Manuscript on file, Department of Anthropology, University at Albany, Albany, NY.
- Justeson, John, and Terrence Kaufman  
2007 The Epi-Olmec tradition at Cerro de las Mesas in the Classic Period. In *Classic Veracruz: Cultural Currents in the Ancient Gulf Lowlands*, edited by Philip J. Arnold and Christopher A. Pool. Dumbarton Oaks, Washington, DC, in press.
- Justeson, John, William Norman, Lyle Campbell and Terrence Kaufman  
1985 *The Foreign Impact on Lowland Mayan Language and Script*. Middle American Research Institute Publication 53. Tulane University, New Orleans.
- Kaufman, Terrence  
1963 Mixe-Zoque Diachronic Studies. Manuscript on file, Department of Anthropology, University of Pittsburgh.  
1981 Comparative Uto-Aztec Phonology. Manuscript on file, Department of Anthropology, University of Pittsburgh.  
1967–2003a Motozintla Kotoke Lexical Database, 5,845 entries.  
1967–2003b Tuzantán Kotoke Lexical Database, 4,752 entries.  
1969/1984–1993. Huasteca Nawa Lexical Database, 13,166 entries.  
1994 The Native Languages of Meso-America. In Christopher Moseley and R. E. Asher, editors, *Atlas of the World's Languages*, pp. 34–45. Routledge, London.  
1994–2004/2007 The History of the Nawa Language Group from Earliest Times to the Sixteenth Century: Some Initial Results. Manuscript on file, Department of Anthropology, University of Pittsburgh. (The 1994 version is available online at <http://www.albany.edu/pdlma/Nawa.pdf>).  
2000–2007 Olmecs, Teotihuacaners, and Toltecs: Language History and Language Contact in Meso-America. Manuscript on file, Department of Anthropology, University of Pittsburgh.  
2001 Language Contact in Preclassic Meso-America and the Languages of Teotihuacán. Paper presented at the 66th Annual Meeting of the Society for American Archaeology, New Orleans.
- Kaufman, Terrence, and Valerie Himes  
1993–2005. Sotapan Gulf Sokean Lexical Database, 14,283 entries.
- Kaufman, Terrence, and Heidi Anna Johnson  
1994–2005. San Miguel Chimalapa Oaxaca Soke Lexical Database, 14,530 entries
- Kaufman, Terrence, and John Justeson  
2004 Zapotec. Appendix 3 in *Encyclopedia of the World's Ancient Languages*, edited by Roger D. Woodard, pp. 1109–1111. Cambridge University Press, Cambridge.  
2007 The Epi-Olmec Language and Its Neighbors. In *Classic Veracruz: Cultural Currents in the Ancient Gulf Lowlands*, edited by Philip J. Arnold and Christopher A. Pool. Dumbarton Oaks, Washington, DC, in press.
- Kaufman, Terrence, with John Justeson  
2003 *A Preliminary Mayan Etymological Dictionary*. Available online at <http://www.famsi.org/reports/01051/index.html>.
- Kaufman, Terrence, and William M. Norman  
1984 An Outline of Proto-Cholan Phonology, Morphology, and Vocabulary. In *Phoneticism in Mayan Hieroglyphic Writing*, edited by John S. Justeson and Lyle Campbell, pp. 77–166. Institute for Mesoamerican Studies Publication No. 9. State University of New York at Albany, Albany NY.  
1994–2005 Santa María Chimalapa Oaxaca Soke Lexical Database, 16,738 entries.
- Keller, Kathryn C., and Plácido Luciano, G.  
1997 *Diccionario Chontal de Tabasco*. Serie de vocabularios y diccionarios indígenas “Mariano Silva y Aceves”, 36. Instituto Lingüístico de Verano, Tucson.
- Kelsey, Harlan P. and William A. Dayton  
1942 *Standardized Plant Names*, 2d ed. J. Horace McFarland Company/American Joint Committee on Horticultural Nomenclature, Harrisburg.
- Kufer, Johanna, and Cameron McNeil  
2006 The Jaguar Tree (*Theobroma bicolor* Bonpl.). In *Chocolate in Mesoamerica: A Cultural History of Cacao*, edited by Cameron L. McNeil, pp.90–104. University Press of Florida, Gainesville.
- Laughlin, Robert M.  
1975 *The Great Tzotzil Dictionary of San Lorenzo Zinacantán*. Smithsonian Contributions to Anthropology, Number 19. Smithsonian Institution, Washington DC.
- Lehmann, Walther  
1920 *Zentral-Amerika*, 2 vols. Dietrich Reimer, Berlin.
- Martin, Simon  
2006 Cacao in Ancient Maya Religion: First Fruit from the Maize Tree and Other Tales from the Underworld. In *Chocolate in Mesoamerica: A Cultural History of Cacao*, edited by Cameron L. McNeil, pp. 154–183. University Press of Florida, Gainesville.
- Mateo-Toledo, B'alam  
2003 The Use of Languages' Names: The Mayan Case. *International Journal of American Linguistics* 69:151–153.
- Maxwell, Judith M., and Robert M. Hill II (translators)  
2006 *Kaqchikel Chronicles: The Definitive Edition*. University of Texas Press, Austin.
- Maya Society  
1935 (ca. 1695) *Arte y diccionario en lengua Chol'tí: A Manuscript Copied from the Libro Grande of Fr. Pedro [Francisco] Moran of about 1625*, facsimile ed. Publication No. 9. Maya Society, Baltimore.
- MacLeod, Barbara, and Dorie Reents-Budet  
1994 The Art of Calligraphy: Image and Meaning. In *Painting the Maya Universe: Royal Ceramics of the Classic Period*, by Dorie Reents-Budet, pp. 106–163. Duke University Museum of Art, Durham, NC.
- McNeil, Cameron L. (editor)  
2006 *Chocolate in Mesoamerica: A Cultural History of Cacao*. University Press of Florida, Gainesville.
- Michelon, Oscar (editor)  
1976 *Diccionario de San Francisco*. Akademische Druck- und Verlagsanstalt, Graz.
- Miksicek, Charles H.  
1983 Paleocology and Subsistence at Pulltrouser Swamp: The View from The Float Tank. Report to the National Science Foundation, Washington, DC.
- Miller, Wick R.  
1988 *Uto-Aztec cognate sets*. Printout of computerized database, May 1988. Manuscript on file, Department of Anthropology, University of Pittsburgh.
- Mock, Carol (compiler)  
1977 *Chocho de Santa Catarina Ocotlán*. Archivo de Lenguas Indígenas del Estado de Oaxaca, No. 4. Centro de Investigación para la Integración Social and the Colegio de México. Mexico City.
- Molina, Alonso de  
1571 *Vocabulario en lengua castellana y mexicana / Vocabulario en lengua mexicana y castellana*. Casa de Antonio de Spinoza, Mexico City.
- Montejo López, Bernábe, Jorge Guzmán Gutiérrez, Rubén López López, Enrique Jiménez Jiménez, and Ernesto Martínez López  
2000–2001 *Säkläjib' Ty'an Ch'ol*. Gobierno del Estado, Tuxtla Gutiérrez, Chiapas, Mexico.
- Montgomery-Anderson, Brad  
2003 Yokot'an (Chontal Mayan) Lexical Database, 10,102 entries.
- Mora-Marín, David  
2001 The Grammar, Orthography, Content, and Social Context of Late Preclassic Mayan Portable Texts. Ph.D. dissertation, Department of Anthropology, University at Albany.  
2003 Primary Standard Sequence Database. Foundation for the Advancement of Mesoamerican Studies, Inc., Project Report 02047. Available online at <http://www.famsi.org/reports/02047>.
- Morán, Fray Francisco ca. 1695 See Maya Society 1935.
- Motul Dictionary. See Ciudad Real.
- Otzo Calí, Simón  
1999 *Memorial de Sololá*, facsimile ed. Comisión Interuniversitaria Guatemalteca de Conmemoración del Quinto Centenario del Descubrimiento de América, Guatemala City.
- Owen, Michael G.  
1970 *A Short Maya-Spanish and Spanish-Maya Dictionary*. Copy Mart, Seattle.
- Peralta, Valentín  
2002–2007 Pajapan Gulf Nawa Lexical Database, 10,750 entries.
- Pérez Mendoza, Francisco, and Miguel Hernández Mendoza  
1996 *Diccionario T'utujil*, edited by Juan Felipe Dayley. Proyecto Lingüístico “Francisco Marroquín”, Antigua, Guatemala.

- Pye, Clifton  
1996–1999 Copainalá Chiapas Soke Lexical Database, 9,084 entries.
- Reilly, Ehren and Cathy Berezna  
1994–2004. Texistepec Gulf Sokean Lexical Database, 4,556 entries.
- Rensch, Calvin R.  
1989 *An Etymological Dictionary of the Chinantec Languages*. Summer Institute of Linguistics, University of Texas, Arlington.
- Rhodes, Richard, Terrence Kaufman, and Dennis Holt  
1994–2005 Sayula Mijejan Lexical Database, 14,683 entries.
- Romero, Sergio  
1999–2002 Zongolica Nawa Lexical Database, 6,837 entries.
- Roys, Ralph L.  
1933 *The Book of Chilam Balam of Chumayel*. Carnegie Institution of Washington Publication 435. Carnegie Institution of Washington, Washington DC.  
1965 *The Ritual of the Bacabs: A Maya Book of Incantations*. University of Oklahoma Press, Norman.
- Saenz de Santa María, Carmelo  
1940 *Diccionario Cakchiquel-Español*. Tipografía Nacional, Guatemala City.
- Sahagún, Bernardino de  
1558–1561 *Primeros Memoriales*, paleography of Nahuatl text and English translation by Thelma D. Sullivan (1997). University of Oklahoma Press, Norman.  
1577 *Historia general de las Cosas de Nueva España*, translated by Arthur J. O. Anderson and Charles E. Dibble (1951–1982). University of Utah Press, Salt Lake City.
- Salgado González, Sylvia  
1996 *Social Change in a Region of Granada, Pacific Nicaragua (1000 B.C.–1522 A.D.)*. Unpublished Ph.D. dissertation, Department of Anthropology, University at Albany, Albany, NY.
- Sampeck, Katherine  
2005 *An Archaeology of Conquest and Colonization: A Comprehensive Regional Survey of the Río Ceniza Valley, Department of Soconusco, El Salvador*. Unpublished Ph.D. dissertation, Department of Anthropology, Tulane University, New Orleans.
- Santamaría, Francisco J.  
1992 *Diccionario de Mejicanismos*. Editorial Porrúa, Mexico City.
- Sedat S., Guillermo  
1955 *Nuevo diccionario de las lenguas k'ekchi' y española, en dos partes. n.p.*, Chamelco, Guatemala.
- Siméon, Rémy  
1977 *Diccionario de la lengua nahuatl o mexicana*. Siglo Veintiuno, Mexico City. (Translation of *Dictionnaire de la langue nahuatl ou mexicaine* [1885]. Imprimerie Nationale, Paris.
- Smailus, Ortwin  
1989 *Vocabulario en lengua castellana y guatemalteca que se llama cakchiquel chi. Análisis gramatical y lexicológico del Cakchiquel colonial según un antiguo diccionario anónimo*, 3 vols. Bibliothèque Nationale de Paris, Fonds Américaine No. 7. Wayasbah, Hamburg.
- Steinbrenner, Larry  
2006 Cacao in Greater Nicoya: Ethnohistory and a Unique Tradition. In *Chocolate in Mesoamerica: A Cultural History of Cacao*, edited by Cameron L. McNeil, pp. 253–270. University Press of Florida, Gainesville.
- Stuart, David  
1988 The Rio Azul Cacao Pot: Epigraphic Observations on the Function of a Maya Ceramic Vessel. *Antiquity* 62:153–157.  
2005 Sourcebook for the 29th Maya Hieroglyphic Forum, March 11–16, 2005. Department of Art and Art History, University of Texas at Austin.
- 2006 The Language of Chocolate: References to Cacao on Classic Maya Drinking Vessels. In *Chocolate in Mesoamerica: A Cultural History of Cacao*, edited by Cameron L. McNeil, pp.184–201. University Press of Florida, Gainesville.
- Sullivan, Thelma  
1988 *Compendio de la gramática nahuatl*. Instituto de Investigaciones Históricas, Universidad Nacional Autónoma de México, Mexico City.
- Sullivan, Thelma and Neville Stiles (translators)  
1988 *Thelma D. Sullivan's Compendium of Nahuatl Grammar*, edited by Wick R. Miller and Karen Dakin. University of Utah Press, Salt Lake City.
- Suslak, Daniel  
1996–2002 Tototepic Highland Mije Lexical Database, 7,661 entries.
- Suslak, Daniel, Giulia Oliverio and James Fox  
1995–2005 Ayapa Gulf Sokean Lexical Database, 5,562 entries.
- Tedlock, Dennis  
2002 How to Drink Chocolate from a Skull at a Wedding Banquet. *RES: Anthropology and Aesthetics* 42:165–179.
- Tapia García, L. Fermín  
1999 *Diccionario amuzgo-español: el amuzgo de San Pedro Amuzgos, Oaxaca*. Centro de Investigaciones y Estudios Superiores en Antropología Social, Mexico City.
- Tolstoy, Paul  
1989 Coapexco and Tlatilco: Sites with Olmec Material in the Basin of Mexico. In *Regional Perspectives on the Olmec*, edited by Robert J. Sharer and David C. Grove, pp. 85–121. Cambridge University Press, Cambridge.
- Turner, Billie L., and Peter D. Harrison  
1983 *Pulltrouser Swamp: Ancient Maya Habitat, Agriculture, and Settlement in Northern Belize*. University of Texas Press, Austin.
- Ulrich, Mateo, and Rosemary de Ulrich  
1976 *Diccionario bilingüe: Maya-Mopán y Español, Español y Maya-Mopán*. Instituto Lingüístico de Verano, Guatemala City.
- Varea, Francisco de ca.  
1600 *Calepino en lengua Cakchiquel*. Manuscript on file at the American Philosophical Society, Philadelphia.
- Wichmann, Søren  
1995 *The Relationship among the Mixe-Zoquean Languages of Mexico*. University of Utah Press, Salt Lake City.  
1998 A Conservative Look at Diffusion Involving Mixe-Zoquean Languages. In *Archaeology and Language II: Archaeological Data and Linguistic Hypotheses*, edited by Roger Blench and Matthew Spence, pp.297–323. Routledge, London.
- Wolgemuth Walters, Joseph Carl, Marilyn Minter de Wolgemuth, Plácido Hernández Pérez, Esteban Pérez Ramírez, and Christopher Hurst Upton  
2000 *Diccionario Náhuatl de los municipios de Mecayapan y Tatahuicapan de Juárez, Veracruz*. Instituto Lingüístico de Verano, Mexico City.
- Ximénez, Francisco  
1993 *Arte de las tres lenguas kaqchikel kiché y tzutujil*. Academia de Geografía e Historia, Guatemala City.
- Zavala Maldonado, Roberto  
1994–2004 Oluta Mijejan Lexical Database, 5,720 entries.  
2000–2003 Tecpatán Chiapas Soke Lexical Database, 11,712 entries.
- Zender, Marc  
2002 Comments on the Inscription of the Williams College Vessel (K8713). Available online at <http://www.mayavase.com/8713/com8713.html>.

