

**Tlapaneko-Sutiaba, OtoMangean, and Hokan:**  
where Greenberg went wrong

Terrence Kaufman  
University of Pittsburgh  
Jan-Mar 1990

Prepared for **Language and Prehistory in the Americas:**  
**a Conference on the Greenberg Classification**  
Boulder, Colorado  
March 22-25, 1990.

revised December 1990  
revised June 1991  
some editing 21 September 2006  
edited 2-10 October 2106

Copyright 1990, 2006, 2016 by Terrence Kaufman

Note to reader. Although Allan Taylor (U Colorado Boulder), who convened the Greenberg conference, undertook to edit the volume that emerged from that conference, and although in Fall 2006 he thought he had a commitment from Stanford University Press to publish the volume, the deal was not concluded, and this article remained unpublished until now.

It should be noted that for some language and group names I now use different spellings:

Popolokan => Masatekan  
YutaNawa => Yuta-Nawan  
OtoMangean => Oto-Mangean  
Mesoamerica => Meso-America

## **Abstract**

This article/study has several aims: (a) to show that Tlapaneko-Sutiaba belongs to OtoMangean [OM] (not to Hokan as stated in Greenberg 1987) and to outline some results of comparative OM studies by the author [TK] since 1982; (b) to give an outline of/preliminary report on comparative work done on Hokan by TK since 1987; (c) to give evidence that suggests that OM and Hokan may be genetically related; but (d) to claim that Sapir, Oltrogge, and Greenberg have NOT shown any special relationship between Tlapaneko-Sutiaba and Hokan: the only valid comparison is between proto-Hokan and proto-OM.

## **0. Introduction**

In 1925, Edward Sapir tried to show that Sutiaba [FN 1] (there spelled Subtiaba) and Tlapaneko [FN 2] [Tlapaneko and Sutiaba form the Tlapanekan group of languages] were Hokan languages. For the (at the time) better-documented Sutiaba, Sapir used data collected by Walther Lehmann and published in 1920. In this two-part article Sapir laid out a good deal of data supporting the reconstruction of a sizable number of grammatical markers for Hokan. Sapir's effort to show that Tlapanekan was Hokan, however, was unsuccessful, and the phonological changes he had to assume in the prehistory of Tlapanekan to make it comparable to what he thought of as the rest of Hokan were elaborate and often not very plausible. The weakness of Sapir's case for Tlapanekan as Hokan was not apparent to many readers, who were dazzled by Sapir's evidence for Hokan grammatical markers.

Scholars both before and after Sapir 1925 saw that Tlapanekan belonged in OtoMangean (or a precursor of the OM hypothesis). Those writing after 1925 had to make some comment on Sapir's 1925 paper.

Already in 1864 Orozco y Berra proposed a "Mixteco-Zapoteco" linguistic stock consisting of Mistekan, Amusgo, Sapotekan, Popolokan, and Tlapaneko.

Weitlaner 1941 and Radin 1944 hypothesized that since Tlapanekan was OtoMangean, OM and Hokan were probably genetically related.

Swadesh 1960 presented grammatical evidence linking Tlapanekan to Mangean.

Rensch 1977 presented lexical evidence relating Tlapanekan to OM, and remarked that if Sapir 1925 was right, OM and Hokan must be related, without taking any position on the correctness of Sapir's hypothesis of a relationship between Tlapanekan and Hokan.

"The claim that Tlapanec -- and therefore, Subtiaba -- is clearly related to the already recognized branches of Otomanguean, however, is not necessarily to deny its Hokan affiliation. If the Tlapanec-Otomanguean hypothesis is accepted, there are at least two possible views regarding the Tlapanec-Hokan hypothesis: (a) that Tlapanec is not genetically related to the Hokan languages; (b) that Otomanguean (including Tlapanec) is a previously unrecognized branch of Hokan-Coahuiltecan." [77]

Oltrogge 1977 compares Jicaque (Tol), Sutiaba, and Tequistlatec (Chontal) and claims that they are all genetically related and form a genetic unit, a claim that I find to be impossible, as subsequent argumentation will show. Oltrogge seems to think that these three languages are more likely to be part of Hokan than otherwise.

"The possibility that Supanec [= Tlapanekan TK] may be Otomanguean (while not necessarily ceasing to be Hokan) has just recently been proposed by Rensch (1977)." [10]

"There are three logical possible solutions, therefore, to the problem of affiliation of Jicaque-Subtiaba-Tequistlatec:

- (1) that Jicaque-Subtiaba-Tequistlatec is related exclusively to the Hokan languages ...
- (2) that Jicaque-Subtiaba-Tequistlatec is related exclusively to the Otomanguean lgs ...
- (3) that Jicaque-Subtiaba-Tequistlatec is somehow related to both the Hokan and Otomanguean phyla, as Rensch 1977 has already suggested." [40]

Sua\*rez 1986 presented grammatical evidence relating Tlapanekan to OM, and made remarks similar to those made by Rensch.

"[E]s oportuno presentar datos que deben disipar las dudas sobre la filiación otomangué del tlapaneco y en consecuencia también de subtiaba; esto implicaría la invalidez de la hipótesis de Sapir sobre la filiación hoka del subtiaba-tlapaneco en tanto no se conecte todo el grupo otomangué con el hipotético hoka." [267-268]

The above point is exactly the one I also make here: Tlapanekan is OtoMangean; if Tlapanekan is to be compared to Hokan, then all of OM must be compared to Hokan. Sua\*rez's inferences are based on Rensch's phonological model, and to that extent are shaky; in any case the present study presents much more data in support of the genetic membership of Tlapanekan in OM.

Neither Rensch nor Sua\*rez expressed any particular enthusiasm for the possible genetic connexion between OM and Hokan. Rensch is not nearly as forceful as he should be in asserting the OM affiliation of Tlapanekan. Oltrogge fails to recognize the absurdity of his postulated Jicague-Subtiaba-Tequistlatec genetic group. Both Rensch and Oltrogge point out that the task of reconstructing proto-Hokan and then comparing it to proto-OtoMangean would involve tremendous amounts of work.

It is therefore notable, though perhaps not surprising, since he wishes to emulate Sapir and abandon none of his good ideas, that Greenberg 1987 assigns Tlapanekan to Hokan. It is surprising, however, that Greenberg does not discuss the fact that the OtoMangean connexion of Tlapanekan has been repeatedly asserted with supporting evidence; nor does Greenberg discuss and reject a connexion between OM and Hokan. What is really surprising is that he goes on to group OM with YutaNawa (UtoAztecan, YN) and Tanoan, and accord this set of languages a major status within his Amerind construct: Central Amerind.

Now Greenberg's OM data comes from Rensch's 1966 dissertation (published in 1976), and consists of proto-OtoMangean reconstructions with supporting forms but no reconstructed glosses. In my 1983ms article I show that Rensch's reconstructions are seriously flawed, and outline what I believe to be an adequate model of proto-OM phonology and a diachronic phonology for the whole stock, branch by branch. Inasmuch as Rensch's OM reconstructions are valid only occasionally, the comparisons made by Greenberg between OM, YN, and Tanoan rely for the OM

component on essentially chance resemblances.

The purposes of this study are:

[a] to provide some conclusive grammatical evidence showing that Tlapanekan is genetically a member of the OM stock with a special connection to Mangan within that stock. This evidence is drawn from the verb markers: tense, aspect, mood, valence-changers, and nominalizers. Additional grammatical evidence is available, as well as lexical and phonological evidence, but to present and discuss it here would be gilding the lily, and use up too much space.

[b] to examine the question of whether OM and Hohan may be genetically related, whether or not such a relationship is a special node under an even higher grouping [see Kaufman 1988.53].

[c] to urge that long-range comparison make use of proto-languages whenever language families/stocks are being brought into the comparison. This means that if the reconstruction is not available, one has to be made.

Two other tasks that logically relate to the above and are potentially doable, have not been undertaken here:

[d] to ask why Sapir 1925, Oltrogge 1977, and Greenberg 1987 thought Tlapanekan was related to Hohan.

[e] to ask why Greenberg thinks OM is related to UA and Tanoan.

### **1. Comparative work on OtoMangan and Hohan by Kaufman**

Like others who pursued a graduate career at Berkeley, I have always been interested in the Penutian and Hohan hypotheses, without thinking they were very convincing in the absence of successful reconstruction. I have kept abreast of comparative work within the constituent families of both Hohan and Penutian since 1958. Beginning in the first part of 1987, I started looking at the literature on comparative Hohan with a view to seeing whether reconstruction was feasible. My motivation was first, that I have been doing my own comparative work on OtoMangan since 1981; and second, that I knew that Greenberg had assigned Tlapanekan to Hohan in his forthcoming book, and I intended to write something demolishing that assignment, since the work of Swadesh, Rensch, and Sua\*rez had quite clearly shown Sapir to be in error, although

none of these scholars made this point very forcefully. I also wanted to know if there was any evidence at all for supposing that Tlapanekan and Hokan might have a special genetic connexion. The result of my comparative work in Hokan is that I believe that there is a valid genetic grouping containing MOST (not all) of the languages usually included within the Hokan hypothesis, and that there are at least 600 proposed Hokan etymologies (mostly discovered by Sapir and Karl-Heinz Gursky) that are probably valid. I can also offer a set of sound correspondences, a partial theory as to the phonological structure of proto-Hokan, and a moderately detailed (though incomplete) model of the morphosyntax of proto-Hokan. It must be stressed that all Hokan reconstructions I offer are still tentative.

So far some criticism has been voiced (which I will not attribute to anyone in particular) that I merely 'assume' that the Hokan languages are genetically related, and that my results are 'speculative' and 'unconvincing'. I will comment on these points briefly. I do not merely 'assume' the the Hokan languages are related to each other. Evidence assembled by other scholars, when carefully examined, convinces me of the relationship. In order to independently verify the evidence for their relationship, I had to assume they were relatable. The data was subjected to the test of regular sound correspondences and detailed points of agreement in grammar. When the evidence was sifted, it became clear that certain languages were NOT relatable -- namely Tonkawe, Karankawe, and Chumash. And some languages are so poorly documented that although they fit my developing model of Hokan, their relationship may not ever be demonstrable in an air-tight way -- Eselen, Yeme\*an (Komekrudo), and Yue\* (Kotoname) (Kaufman 1994.38). To get started, I had to 'assume' a set of comparable entities. Once the data have been sorted into GOOD STUFF and JUNK, however, 'assuming' is no longer at issue. I have not tried to 'convince' anyone that proto-Hokan is real, or that it had a particular phonological or morphological structure or lexical content. I have merely reported, without supporting evidence, what have been the current states [1989, 1989msb] of my conclusions about comparative Hokan. It is misguided (or at least redundant) to call my comparative Hokan work 'speculative' -- all historical inference is speculative, or at least must start out that way. It is fair to call my hypotheses 'unsupported' (in print) . But I have pledged to provide the evidence for my proposals, along with argumentation to support them. Until this is done, I certainly do not expect anyone to be 'convinced'. On the other hand, I do predict that if anyone delves into Hokan comparative matters as deeply as I have before I get my

evidence before the public, he will find my publicly-presented hypotheses to be fruitful and explanatory. I certainly do not demand credit for what I haven't done.

The OtoMangean and Hokan reconstructions that I have made (OM: 1982-1988; Hok: 1987-1989) have been made without reference to each other or to any other reconstructions for other language families. Although it may seem difficult to avoid contamination, I have little trouble focusing on the data of one language family at a time. If I become convinced of a special relationship between OM and Hok I will probably use insights from one side to see if they will be profitable on the other.



TABLE 1

COMPARATIVE WORK ALREADY ACCOMPLISHED BY TK ON OTOMANGEAN AND HOKAN

stock	OtoMangean	Hokan
time depth	ca. 6500 years	ca. 8000 years
etymologies	ca. 500	ca. 600
regular correspondences	yes	yes
comparative phonology	all but tones	educated guess
person markers	etymologies	etymologies
numerals	etymologies	etymologies
verb inflection	yes	hypotheses
noun inflection	---	hypotheses
deictics	---	hypotheses
NP	hypotheses	hypotheses
VP	hypotheses	hypotheses
historical scenario		
phonology	yes	not yet
verbs	yes	not yet

**OtoMangean breakdown** (this is my considered opinion about the classification of OtoMangean)

Western OM

OtoPame-Chinanteko

OtoPame: Otomi\*, Masawa, Matlatzinka-Okwilteko,  
Pame, Chichimeko

Chinanteko

Tlapaneko-Mangean

Tlapaneko-Sutiaba

Mangean: Chiapaneko, Mange

Eastern OM

Popolokan-Sapotekan

Popolokan: Masateko, Iskateko, Chocho-Popoloka

Sapotekan: Sapoteko, Chatino

Amusgo-Mistekan

Amusgo

Mistekan: Misteko, Kwikateko, Triki

**Hokan breakdown** (every item in the last column is a genetic group or otherwise unaffiliated single language; groups in [...] are regional rather than genetic).

Northern Hokan

[Sonoma]

Pomoan family

[Northern California]

Chimariko

Yanan language area

Karuk

Shasta-Konomihu family

Achumawi-Atsugewi family

[Basin]

Washu

[Coast]

Eselen

Salina [FN 3]

Southern Hokan

[Southwest]

Yuman-Kochimi\* family

Seri

[Texas-Mexico]

Pahalat (AKA Coahuilteco) [FN 4]

Komekrudoan (Yeme\*an) family or language area: Mulato, Yeme\*, Mia\*kan, Mamulike

Kotoname (= Yue\*)

[Meso-America]

Chontal [FN 5]

Tol (AKA Jicaque) [FN 6] language area

No classification of Hokan that identifies mid-level groupings is yet clear to me, except a possible North:South split.

## 2. Notes on the History of OtoMangean Classification [FN 7]

### 2.1. Early Classifications

The classification of the OM languages that I propose here, insofar as it concerns groupings above the family level, was done in ignorance of any cross-family classification schemes produced before 1959. Until 1947, with the appearance of Swadesh's 'The Phonemic Structure of Proto-Zapotec', classifications dealing with the OM languages or some subset of them were based on the inspection of lexical materials. This was followed by reconstructions for Otomian (Newman and Weitlaner 1950), Mistekan (Longacre 1957), Popolokan (Gudschinsky 1959), Mangean (Ferna\*ndez and Weitlaner 1961), OtoPamean (Bartholomew 1965ms), Chinantekan (Rensch 1963ms) and others. The systematic testing of parts of the OM hypothesis began with Gudschinsky's **Proto-Popotecan** (1959), followed by Ferna\*ndez and Weitlaner's **Sobre Algunas Relaciones de la Familia Mangué**, capped by Rensch's **Comparative Otomanguéan Phonology** (1966, pub. 1976).

Classifications from before 1959, however, are interesting to look at because they give us an idea of what can be accomplished in the generating of hypotheses without reconstructing. Jime\*nez Moreno 1962 contains a useful survey of the classification of the OM languages since 1864.

Orozco y Berra 1864 defines a "Mixteco-Zapoteco" group including EOM plus Tlapanec and two non-OM languages 'Teco' of Michoaca\*n, and 'Pupuluca' of Guatemala. In 1875 Pimentel defines a "Mixteco-Zapoteco" family consisting of several EOM languages plus Chinanteko. In 1905 Belmar defines a "Mixteco-Zapoteco" family containing all the OM languages except Tlp and CM.

This is the first inkling of OM as we know it, the two earlier treatments hardly overflowing the bounds of EOM. Nothing like Belmar's extended grouping was talked about again until 1920. Between 1901 and 1911 appeared three works that recognized EOM under three different names: Brinton 1901 ("Tzapotec-Mixtec"), Leo\*n 1902 ("Mixteco-Tzapotecan"), and Thomas & Swanton 1911 ("Zapotecan").

In 1920 Lehmann initiated a new tack on grouping the languages we know to be OM. He grouped CM, OP, Pn and Tri into an unnamed genetic group. In 1924 Rivet proposed an "Otomi" group made up of CM, Chn, OP, Pn+Tri, and Sn. Schmidt 1926 recognized an Otomi-Mangue group made up of CM, Chn, OP, Pn, and Tri and a Mixteco-Zapoteco group made up of Chn, Mis-Kwi, Ams, and Sn. Note that he classifies Chn in each of the two groups. This approach adumbrates OM from an expanded WOM angle, but WOM as I define it was never recognized as a special set of languages before the present study.

Sapir 1929 suggested that Rivet's "Otomi" group (= Schmidt's "Otomi-Mangue" group = Lehmann's unnamed group (which Sapir renames "Otomian")) plus Chn combined with Schmidt's "Mixteco-Zapoteco" might form a larger valid genetic group. [He says, as though to avoid direct responsibility, 'Some connect Chinantec, Mixtec-Zapotec, and Otomian in one great linguistic stock, **Mixtec-Zapotec-Otomi**']. This hypothesis, which became the orthodox position on the make-up of OM, that is OM without Tlp, lasted till Rensch 1977.

Although the connection of Tlapaneko with other languages we now know to be OM was proposed in Orozco y Berra 1864, there was a general exclusion of Tlapaneko from OM following Sapir 1925, and the less bold may perhaps be forgiven for not trying to swim upstream: but note that at this point we pass from science into politics.

In 1941 Weitlaner suggested connections between Tlapaneko, Hokan, and OtoMangean, as well as between OtoPamean and YutaNawan. In 1944 Radin included Tlapaneko once again in OM. He reestablished EOM as a genetic grouping. He believed that Sapir had shown than Sutiaba was related to Hokan, that Tlapaneko belonged in [unnamed] OM, and that OM and Hokan were related; this suggestion was supported by Weitlaner 1945, but the proposal was generally ignored.

## 2.2. Recent Classifications

In 1959 and 1960 Swadesh suggested reorganizing several aspects of OM. He recognized EOM, calling it first "Macro-Mixtecan", then "Popoloca-Zapotecan", then "Oaxacan" (which I concede might be as good a term as my EOM). He thought Chn to be only distantly related to OM. He thought Tlp and CM to form a genetic group perhaps not related to OM at all in any meaningful or interesting way, except as "Amerindian". He thought Wavi [usually spelled Huave] was related to Oaxacan, especially closely to Sapn. Swadesh was (as I think I show in Kaufman 1988ms) right that Tlp and CM form a special genetic group. It cannot be said that Swadesh made very clear what his intentions were as regards the subclassification and membership of OM. In two articles published in 1964 he made explicit proposals about the reconstruction of OM and its high-level divisions and branches. He included Wavi in OM. He never tried to account for anything in Chn; while in 1960 he linked CM and Tlp, and excluded both from "OM", in 1964 he included CM in OM, yet never mentioned Tlp. Nevertheless, Swadesh had good insights about the possible evolution of the OM languages, and identified a number of etymologies that the procrustean ideas of the members of Longacre's school (virtually all OtoMangeanists associated with SIL) made opaque to them. Swadesh did not carry out comparisons that were systematic or deep enough to enable him to arrive at mature and reasonably watertight syntheses. And his work has unfortunately and undeservedly been largely ignored. For a while both Longacre and Rensch accepted Swadesh's proposal that Wavi is OM, but little has been heard of that hypothesis recently, and I am not impressed by it.

In 1977 Rensch showed Tlp to be OM, and additional evidence has been accumulating at a satisfyingly rapid rate.

Several aspects of the grouping of OM languages into families proved to be hard to deal with until the 1950's, due probably to scanty and inaccurate data that were difficult to analyze. Between Mechling 1912 and Stresser-Pe'an 1952, Tri was grouped with Pn, although sometimes tentatively. During this same period Mis-Kwi was always grouped with Ams. In Stresser-Pe'an 1952, Tri is grouped with Mis-Kwi and Ams. Only since Longacre 1957 is Amusgo claimed (correctly) to be less closely related to Mis-Kwi than Tri is.

### 3. A Demonstration of the OtoMangean Affinity of Tlapaneko-Sutiaba

All of the verb markers of Tlapaneko-Sutiaba are of OM origin or have OM cognates; Tlapanekan does not lose or change the functions in any way untypically of OM languages generally, and Tlapanekan is especially closely related to Mangean within the OM stock. The evolution of the other OM families is given in order to illustrate the typically OM characteristics of Tlapanekan in context.

I sketch here a model of the linguistic diversification of the OM stock in terms of phonological developments, and developments of the markers for tense, aspect, mood, valence-changers, and nominalizers (TAMVN). Obviously, the study of pronouns, nouns, syntax, and vocabulary will provide a much more fleshed-out model and undoubtedly some fairly serious changes will have to be made. Nothing is cherished about this reconstruction except the firm conviction that a great amount of detail can be reconstructed for pOM and its intermediate descendants. The changes that will have to be made will probably involve adding new categories and filling out the distributions of known categories, rather than forcing the abandonment of categories proposed here (except perhaps some of those based in only one language each of just two families).

In the sketch to follow, phonology will be dealt with only for the divisions and branches, not the families and isolates. These latter facts are largely covered in Kaufman 1983ms. They could profitably be covered here as well, but space does not permit. [See sections 4 for an outline of pOM phonology; phonology is discussed there to allow a more direct comparison between OM and Hokan phonologies].

Preparatory to outlining the developments in the various divisions, branches and families of OM, I will recapitulate here (in Table 2) the reconstruction of the markers and their distributions that I attribute to pOM. # indicates that the marker was a separately accented word in pOM, + that the item was a proclitic. When the accentual properties of an item have not been determined neither symbol is found. Items marked with ! are limited to "Central OM". Items limited to WOM or EOM are not listed here.

TABLE 2

NOMINALIZERS (derivational)

\*nau#           NOMINALIZER  
\*ma#            ADJECTIVIZER

VALENCY CHANGERS (derivational)

\*ce#            CAUSATIVIZER  
\*ta+/#          IMPERSONAL 1  
\*hi+/#          IMPERSONAL 2  
!\*wai           IMPERSONAL 3 (if not Tlp-CM spread to Pn)

AUXILIARIES (position 1)

\*tau            COME to

PLURAL SUBJECT (position 2)

!\*+o            PLURAL SUBJECT (if not Tlp-CM spread to Pn)

POSITION 3 TAM MARKERS

\*i+/# ~ ZERO    INDEFINITE INDICATIVE  
\*wau+            DURATIVE 1  
\*xi+             DURATIVE 2  
\*kWe#            COMPLETIVE  
\*ki# ~ \*ka#      POTENTIAL  
!\*ha             IMPERATIVE (if not Tlp-CM spread to Pn)  
!\*tau            OPTATIVE (if not Tlp-CM spread to Pn)



POSITION 4 TAM MARKERS

4a	
* (Y)ti+/#	PERFECT
*ni+/#	REMOTE PAST/LONG.AGO
4b	
*ka#	RECENT PAST/RECENTLY
*nma#	'ALREADY'
*na#	'NOW'
*ma#	FUTURE
4c	
*kai#	PROGRESSIVE; BE
!*ni#	BEING (if not Tlp-CM spread to Pn)

NEGATORS (position 5)

*sa(u)# ~ *si	NEGATIVE 1
*tau# ~ *tta#	NEGATIVE 2
!*7au#	NEGATIVE 3 (if not WOM spread to Pn)
*ya#	NEGATIVE 4

The chart below (Table 3) lays out the TAMVN markers that it has been possible to reconstruct for pOM, EOM, WOM, OP-Chn, Tlp-CM, Pn-Sn, and Ams-Mn.

TABLE 3

THE TAMVN STRING

	POS. 4	POS. 3	POS. 2	POS. 1	
NEGATOR	TENSE 4a ADVERB 4b HIGHER VERB 4c	ASPECT MODE	PLURAL   AUXILIARY   SUBJECT	DERIVATION	
*sa(u) # NEGATIVE 1	*(Y)ti+/# PERFECT	*i+/# ~ 0 INDEFINITE	*+o PLURAL SUBJECT	*tau COME.TO	*nau# NOMINALIZER
*tau# ~ *tta# NEGATIVE 2	*ni+/# LONG.AGO	*wau+ DURATIVE	*kka PLURAL SUBJECT (EOM)	*kwi# GET.UP.TO (WOM)	*ma# ADJECTIVIZER
*7au# NEGATIVE 3	*mi# IMPERFECT (OP-Chn)	*xi+ DURATIVE			*ce# CAUSATIVIZER
*ya# NEGATIVE 4	*ka# RECENTLY	*kWe# COMPLETIVE			*kWkWa (u) CAUSATIVIZER (WOM)
*ci# NEGATIVE 5 (WOM)	*nma# ALREADY	*ki# ~ *ka# POTENTIAL			*se# CAUSATIVIZER (EOM)
	*na# NOW	*tau OPTATIVE (COM)			

NEGATOR	POS. 4 TENSE 4a ADVERB 4b HIGHER VERB 4c	POS. 3 ASPECT MODE VERB 4c	POS. 2 PLURAL   AUXILIARY   DERIVATION SUBJECT	POS. 1
	*ma# FUTURE	*ha IMPERATIVE (COM)		*ta+/# IMPERSONAL 1
	*mu# HYPOTHETICAL (WOM)			*hi+/# IMPERSONAL 2
		*rV# STATIVE (WOM)		*wai IMPERSONAL 3 (COM)
	*kai# ONGOING			
	*ni# BEING (COM)			
	*kWa# GO/COME TO (WOM)			

NOTE to editor: such terms as PROGRESSIVE and REMOTE PAST will not fit on the chart. Thus, synonyms have been used.

### 3.1. A Sketch of OM Diversification Based on the Reconstruction of Verb Morphology

#### 3.1.1. Western OtoMangean

##### From Proto-OtoMangean to Western OtoMangean

In phonology WOM merged \*ai with \*a, thus leaving an eight vowel nucleus system \*/i e a au o u ia ea/. There is still uncertainty over whether pOM \*/r/ was [r] or [θ], but if the latter, WOM has [r] (Tlp has [t], which is easily derivable from [r], though equally easy to derive from [θ]). WOM has not lost any of the TAMVN markers postulated for pOM, and it has added three. These are: in position 1 \*k<sup>w</sup>i# 'GET UP TO'; in position 4 \*mu# (or \*mo#) 'HYPOTHetical' and \*k<sup>w</sup>a# 'GO TO'; also \*ci#, an additional 'NEGative 5'. This new NEG may have had the same uses as \*ya#, since there is no language in which both survived. It appears that the TAM categories of WOM are heavily weighted to TENSE, though this is doubtless partly an artifact of traditional grammatical labelling.

##### From Western OtoMangean to OtoPamean-Chinanteko

In phonology there are two changes. First, \*ia > u. Second, radical monosyllabism (all polysyllabic strings lose all but one of the vowels: resultant consonant clusters are simplified in various ways). The following TAMVN markers are lost:

- \*ce CAUSATIVE (deriv.)
- \*tau COME to (AUX)
- \*ki POTENTIAL allomorph (position 3)

Two markers new in OP-Chn have been identified:

- \*rV# STATIVE/PERFECT (position 3)
- \*mi# IMPERFECT (position 4a)

### **From OtoPamean-Chinanteko to OtoPamean**

Of surviving TAMVN markers only \*ci 'NEG 5' is lost. No other changes have been identified. This suggests that if OP-Chn is a valid node in the OM family tree, Chn can be derived from a properly reformulated OP. In fact, there may be no reflex of \*ci# in Chn. Thus OP preserves a reflex of all but three of the pOM and pWOM markers reconstructed in this study. This is remarkable, because OP data was not initially used in making the reconstruction.

### **From OtoPamean-Chinanteko to Chinanteko**

The following phonologically independent TAMVN markers are lost:

- \*ma ADJECTIVIZER (deriv.)
- \*+o PLURAL SUBJECT (position 2)
- \*k<sup>w</sup>e COMPLETIVE (position 3)
- \*ka POTENTIAL allomorph (position 3)
- \*(Y)ti PERFECT (position 4a)
- \*mu HYPOTHETICAL (position 4b)
- \*kai PROGRESSIVE (position 4c)
- \*ya NEGATIVE 4

The following cliticized markers have been lost, probably through phonological attrition:

- \*ta IMPERSONAL 1 (deriv.)
- \*hi IMPERSONAL 2 (deriv.)
- \*i INDEFINITE (position 3)
- \*wau DURATIVE (position 3)
- \*xi DURATIVE (position 3)

\*nau NOMINALIZER has gone through the function changes  
NOMINALIZER > ADJECTIVE > STATIVE > PERFECT.

## From Western OtoMangean to Tlapaneko-Mangean

Tlp-CM lost the following:

- \*xi DURATIVE (position 3), merging it with its synonym \*wau
- \*ka RECENT PAST (position 4b)
- \*ya NEGATIVE 4

It lost inherited \*ta IMPERSONAL 1 and \*hi IMPERSONAL 2,  
and acquired a new IMPERSONAL 3 \*wai

It added a new CAUSATIVE \*k<sup>w</sup>k<sup>w</sup>a.. ~ \*k<sup>w</sup>k<sup>w</sup>au..

Tlp-CM also added the following new TAMVN markers:

- \*ha.. IMPERATIVE (position 3)
- \*tau.. OPTATIVE (position 3)
- \*ni# 'BEING' (position 4c)

These four along with \*wai IMPERSONAL3 seem to have originated in Tlp-CM and spread to Pn, since these are the only places they occur. Since \*ni# occurs only in CM and Pn, it may be supposed that CM itself was the point of spread of these features into Pn. And \*ni# may never have occurred in Tlp.

The OM languages that lack explicit IMV and OPT markers typically use POT to express the imperative and OPT functions. In the negative volitives the COMpletive marker may be used. This may explain the fact that markers that once meant COMP have come to mean POT in some languages, and vice versa.

### **From Tlapaneko-Mangean to Tlapaneko**

Tlp loses the following:

- \*tau 'COME to' (AUX)
- \*kWi 'GET UP to' (position 1)
- \*tau 'COME to' (position 1)
- \*wau DURATIVE (position 3)
- \*kWe COMPLETIVE (position 3)
- \*kWa 'GO to' (position 4c)

It changes the function of \*tau OPTATIVE to IMPERATIVE.

### **From Tlapaneko-Mangean to Mangean (Chiapaneko)**

Chp loses the following TAMVN markers:

- \*ma ADJECTIVIZER (deriv.)
- \*ni REMOTE PAST (position 4a)
- \*ma FUTURE (position 4b)
- \*nma 'ALREADY' (position 4b)
- \*na 'NOW' (position 4b)
- \*sa NEGATIVE 1
- \*ci NEGATIVE 5

Several changes occur in the surviving TAMVN markers:

- \*tau OPTATIVE (position 3) > FUTURE 1
- \*wau DURATIVE (position 3) > FUTURE 2 or CONDITIONAL
- \*kWe COMPLETIVE (position 3) > am-]pi NEGATIVE IMPERATIVE
- \*kai 'BEING' (position 4c) > dummy AUX (or possibly PROGRESSIVE AUX)
- \*ni 'to BE' (position 4c) > RELATIVIZER

### 3.1.2. Eastern OtoMangean

#### From Proto-OtoMangean to Eastern OtoMangean

In phonology, EOM merges the central vowels with the front vowels of corresponding height: \*ia > i, \*ea > e. This results in a seven vowel nucleus system for EOM: \*/i e ai a au o u/. There are no losses or definitive gains in the set of TAMVN markers. Though \*se# CAUSATIVE (deriv.) may have been innovated at this level, it may also be a Sn innovation. A new marker \*k<sup>w</sup>e#, a classifier for animals, is brought into use at this stage. Since all EOM languages have such a pattern, an alternative CAUSative structure - CAUSAUX plus POT plus VERB - may have been innovated at this time.

In EOM languages most of the surviving TAM markers have been incorporated into the TAM string, but some have been excluded and survive as adverbs or higher predicates. Those that are incorporated seem to have almost entirely aspect and mood functions, and tense functions only incidentally or not at all. There may have been a general drift in EOM to either exclude tense or leach it of its specifically temporal content. After the above-mentioned changes in EOM, it is fair to say that tense categories as such do not exist in EOM languages, though durative, completive, and potential are often labeled present, past, and future in non-technical writing and by linguists who follow conventional Spanish-oriented terminology. Conceivably, of course, one or another EOM language might today be adequately analyzed as showing tense phenomena rather than aspectual phenomena for the items in question. I doubt that such meanings, if present, would invalidate my generalization.

#### From Eastern OtoMangean to Popolokan-Sapotekan

In phonology Pn-Sn merges \*au with \*a, and pOM \*θ > t. The following TAMVN markers are lost:

- \*ma ADJECTIVIZER (deriv.)
- \*ka POTENTIAL allomorph (position 3)
- \*ma FUTURE (position 4b)
- \*tau NEGATIVE 2



The function of \*(Y)ti (position 4a) goes through the following change: PERFECT > PROGRESSIVE. There was already a PROGRESSIVE, \*kai, in the language, and the two were not necessarily fused into a suppletive set. Since Sn has \*Yti+ HABITUAL, perhaps this meaning should be postulated for Pn-Sn in order to make the Sn outcome more straightforward. Then we would have HABITUAL > PROGRESSIVE in Pn, something that is less plausible semantically than the inverse change.

### **From Popolokan-Sapotekan to Popolokan**

In Pn the following TAMVN markers are lost:

- \*ta IMPERSONAL 1 (deriv.)
- \*tau 'COME to' (AUX)
- \*i INDEFINITE (position 3) [this loss could perhaps be motivated phonologically]
- \*ni 'REMOTE PAST/LONG AGO' (position 4a)
- \*na 'NOW' (position 4b)

The CAUSATIVE function is restructured as 'to DO' (AUX) plus VERB preceded by POTENTIAL. The derivational causative, whether \*ce or \*se, disappears.

Pn has innovated the marker \*ce- FUTURE & COMPLETIVE (position 3).

Pn seems to have borrowed the following from Tlp-CM, specifically CM:

- \*wi IMPERSONAL 3 (deriv.)
- \*to PLURAL SUBJECT (position 2)
- \*tau OPTATIVE (position 3)
- \*ha IMPERATIVE (position 3)
- \*ni 'IS'/BEING (position 4c).
- \*7au NEGATIVE 3

If these markers are not borrowed, then either all these morphemes are pOM but lost except in the languages that show them, or else the OM languages should be grouped differently than I have done. The data presented here will (I hope) show why I do not like either of those alternatives. I will concede, however, that I do not consider the issue to be closed.

Except for \*ni and \*ma all surviving markers become prefixes/proclitics.

### **From Popolokan-Sapotekan to Sapotekan**

Sn lost the following TAMVN markers:

\*wau DURATIVE (position 3)  
\*xi DURATIVE (position 3)  
\*ka 'RECENTLY' (position 4b)  
\*ya NEGATIVE 4.

It changed the function of \*Yti PROGRESSIVE to DURATIVE to HABITUAL.

All surviving TAMVN markers become proclitic, except for \*nma 'ALREADY', \*sa 'NO MATTER'/NEGATIVE1.

### **From Eastern OtoMangean to Amusgo-Mistekan**

In phonology \*c and \*s merged (probably as [s]). The following markers are lost:

\*nau NOMINALIZER (deriv.)  
\*hi IMPERSONAL 2 (deriv.)  
\*tau 'COME to' (AUX)  
\*kai PROGRESSIVE (position 4c)

The CAUSATIVE function is restructured as CAUSATIVE AUXILIARY plus POTENTIAL plus VERB. The CAUSative marker \*se+ may come from POM \*ce# or from EOM \*se#.

The OPTATIVE function is represented by \*na+ plus POTENTIAL plus VERB. This may be an innovation in structure, and is probably an innovation in function.

\*kWe+ COMPLETIVE shifts its function to POTENTIAL, thus becoming a suppletive fellow allomorph of \*ka+ ~ \*ki+.

\*i- INDEFINITE comes to form a suppletive allomorphic set with \*wau- and \*xi-, all meaning DURATIVE.

### **From Amusgo-Mistekan to Amusgo**

The following markers are lost in Ams:

- \*ma+ ADJECTIVIZER (deriv.)
- \*ta- IMPERSONAL 1 (deriv.)
- \*ka+ POTENTIAL allomorph (position 3)
- \*ni+ REMOTE PAST (position 4a)
- \*ma+ FUTURE (position 4b)
- \*ka# RECENT PAST (position 4b)
- \*sau# NEGATIVE 1
- \*ta# NEGATIVE 4

Ams makes these changes in surviving markers:

- \*ma# 'ALREADY' > PROGRESSIVE
- \*i- ~ \*wa- ~ \*xi- DURATIVE > HABITUAL
- \*na+ plus POTENTIAL plus VERB: OPTATIVE > FUTURE
- \*kWe# ANIMAL CLASSIFIER is changed to ka# or \*ki# depending on the dialect.

### **From Amusgo-Mistekan to Mistekan**

Mn loses \*ma# 'ALREADY'.

\*Yti+ COMPLETIVE > pMistekan \*nti- STATIVE.

ki- changes function from POTENTIAL to COMPLETIVE.

\*kWe# ANIMAL CLASSIFIER is replaced by \*tu7#.

It should be readily apparent that a slightly reformulated pMn could be made to include Amz.

### **3.2. Tlapaneko and Sutiaba**

Tlapaneko-Sutiaba is a language area with about 8 centuries of internal diversification. It is made up of three emergent languages, Sutiaba or Maribio (an outlier in Nicaragua), Azoyu\*, and general Tlapaneko or Yopi (These latter two in Guerrero, Mexico). The boundary between the latter two may be fuzzy.

Sources of data on Tlapaneko-Sutiaba are Lehmann 1920, Radin 1933, Schultze-Jena 1938, Sua\*rez 1983, and Weathers 1976.

Phonological notes: phonetically voiced plosives are to be analyzed as phonemically single, while phonetically voiceless plosives are phonemically double.

#### **3.2.1. Tlapanekan TAMVN Markers**

Following (Table 4) is an outline of the TAMVN markers of Tlapaneko and Sutiaba (Mal = Malinaltepec, Azo = Azoyu\*, Sut = Sutiaba; JS = Jorge Sua\*rez, SJ = Leonhard Schulze-Jena, PR = Paul Radin, WL = Walther Lehmann). The following positions in the TAMVN string can be identified, working leftward from the first root: DERIVATION, PLURAL SUBJECT, TAM, NEGATIVE, IMPERATIVE, ADVERB/AUXILIARY, and ENCLITIC.

TABLE 4

MALINALTEPEC

AZOYU

SUTIABA

DERIVATIONAL MARKERS

NOMinalizer

na- [SJ;JS 491-493]  
found on a few nouns

ADJective former

mi- ~ ma- ADJ  
[JS 214-215, 469-473,  
483-489]

mi- ~ ma- ADJ  
[PR 51]

mi- (~ ma-) ADJ  
[WL 959-960]

CAUSativizer

ci- ~ ce- ~ ca- ~ c- ~ s-  
CAUS  
[JS 394-403, 527-530]  
[SJ 349-350, etc]

ci- ~ ce- ~ ca-  
CAUS [PR 48]

pa- CAUS  
[PR 48, 55]

PASSivizer

wa- (M) (See TAM ZERO)  
PRES/STAT/PCP PASS  
[SJ;JS 89-90, 211-214]

ni-wa- (See TAM ni-)  
PRET PASS [JS,SJ]

ma-(wa-) (See TAM ma-)  
FUT PASS [SJ]

ka-  
PERF PASS [SJ]

PLURAL SUBJECT

-o- ~ -u-[JS 192-193, SJ passim]

## MALINALTEPEC

## AZOYU

## SUTIABA

## TENSE-ASPECT-MODE POSITION

na- PRES [JS 192]  
[SJ HAB PRES]

na- DUR PRES  
[PR 47,54]

na- PRES  
<V marked short>

n-o- ~ n-u- PL SUBJ

i7- EMPH PRES  
[JS 192;  
SJ IMMED PRES]

∅- STAT PRES  
[SJ: some verbs  
take no PRES mkr]

ni- PAST [JS 192]  
[SJ PRET]

ni- PERF DUR  
[PR 47,54]

ni- PRET  
<V unmarked>

ga- ~ gi- (H) POT [JS 193]  
[SJ INDEF FUT]

ga- FUT  
[PR 54]

ga- FUT  
<V unmarked>

ma- ~ mi- FUT [JS 192]  
[SJ DEF FUT]

ma- INDEF PRES  
[PR 47,54]

ma- 'gerundial'  
occurs post-AUX  
<V marked short>

mo- ~ m-u- PL SUBJ

ra- 'CONCOMITANT ACTION'  
[JS 193]

## MALINALTEPEC

## IMPERATIVE POSITION

a- IMV  
 precedes ci- NEG PRES,  
 ga- POT, and ZERO  
 [SJ;JS 193]

## NEGATIVE POSITION

ra- (M) NEG STAT  
 w/ ZERO PRES [SJ;JS 193]

r-o- (M) PL SUBJ

ci- (H) NEG PRES  
 w/ ZERO PRES [SJ;JS 192]

ni- (H) 'no longer'  
 w/ ZERO PRES [JS 193]

ta- ~ ti- (H) NEG PAST  
 w/ ZERO PAST [SJ;JS 192-193]

t-o- ~ t-u- (H) PL SUBJ

nda: (L) 'there isn't any' [JS]

## AZOYU

a- IMV SG  
 [PR]

ci- (H)  
 [PR 48]

## SUTIABA

da- IMV occurs  
 before verb stem  
 <V marked short>



MALINALTEPEC

AZOYU

SUTIABA

-s^a- ~ -s^i- (H) NEG FUT  
[SJ;JS 192]  
metathesizes with ma- FUT

-s^o- ~ -s^u- (H)  
PL SUBJ  
[ma-s^o/u- ~ mo-s^o/u-] (H)

a- NEG occurs  
before TAM  
markers  
and ADV/AUX  
<V marked long>

#### ADVERB and AUXILIARY POSITION

##### Proclitic Adverbs

da- ~ dV-  
INDEF PRES  
[PR 47]

do- 'on-goingly'  
PRES PCP  
bef TAM mkr  
including ZERO  
<V marked long>

ka- PERF PASS [SJ]  
does this go here?

gi- PERF PASS,  
PASS PCP  
bef TAM mkr  
including ZERO  
<V marked long>

MALINALTEPEC

AZOYU

SUTIABA

s<sup>^</sup>i- 'already'  
PERF & PLUPERF  
bef TAM mkrs  
incl ZERO  
<V marked long>

ba- MOMENTANEOUS  
PRES [PR 47]

Independent Adverb or frozen Auxiliary

gisto' "PROG" ADV  
<1st V marked  
long, 2d short>  
bef TAM mkrs

Auxiliary

-ka FUT AUX [SJ]  
= 'to go'  
bef ga- POT  
and ZERO TAM

-ka ~ -ga [PR 47]  
FUT AUX  
before ZERO TAM

-hka FUT AUX  
bef ga- POT

ENCLITIC POSITION

+ma 'already'; PERF  
added to verb with ni-  
[SJ;JS]

#### Notes to Table 4

-o- ~ -u- attached to whatever TAM marker precedes it marks plural subject of a transitive verb (thus what is being marked is plural agent).

While in Mal IMV a- precedes NEG, in Sut NEG a- precedes IMV da-, and IMV precedes VERB.

In Mal after higher verbs the main verb can take various of the TAM markers, or none at all (in the latter case, this may be ZERO PRES). See JS 80-82, 353-359.

In Mal NEG and TAM categories are intertwined. ra-, ci-, and ni- seem to occur with a ZERO PRES marker; ta- may represent a fusion of PAST and NEG; and in the FUT the TAM precedes the NEG. In the last case it appears that metathesis has occurred. In all OM languages but Mal Tlapanec, NEG precedes any TAM markers. While NEG is typically an independent particle, TAM markers are typically proclitic or prefixes, as we have seen at length already.

The allomorphs mi- of FUT and s<sup>^</sup>i- of NEG FUT result from contraction of \*ma# FUT plus \*i+ INDEF and \*sa# NEG plus \*i+ INDEF.

### 3.2.3. OM Etymologies of Tlapanekan TAMVN Markers

The following Tlp-Sut TAMVN markers have OM etymologies (Forms cited are found only in Mal unless otherwise indicated):

OM DERIV

na- NOM is from pOM \*nau# NOM. It does not seem to be productive.

mi- ~ ma- ADJ (Mal,Azo,Sut) is from pOM \*ma# ADJ. The variant ma- is much less common.

cV- CAUS (Mal,Azo) is from pOM \*ce# CAUS or pOM \*cai 'to do'. The various allomorphs never contain a rounded vowel. The fact that a variant with [s] (/c/ before C) occurs suggests that the basic allomorph is /ci-/ (since /i/ is the least sonorous vowel), but the ce- and ca- variants (as well as different ideas about sonority) might better support /ce-/ as the basic variant. It is not clear whether this prefix is productive.

pa- (/ppa-/) CAUS (Azo) is from WOM k<sup>w</sup>k<sup>w</sup>a(u).. CAUS  
(See Chp pa- ~ po-).

wa- (M) PASS is from pOM \*wai.. IMPERS 3.

ka- (/kka-/) PERF PASS may be related to pOM \*kkai# PROG ~ 'to be', a higher verb in OM position 4.

OM 2

-o- ~ -u- PL SUBJ of tv is from pOM \*+o PL SUBJ.

OM 3 & 4

- na- HAB PRES (Mal,Azo,Sut)  
is from pOM \*na# 'now', position 4.
- i7- EMPH/IMM PRES  
is from pOM \*i+ INDEF PRES, position 3.
- ∅- STAT PRES  
may be from the ZERO allomorph of pOM \*i+ INDEF PRES, position 3.
- ni- PRET (Mal,Azo,Sut)  
is from pOM \*ni# 'long ago'/remote past, position 4.
- ga- (~ gi-) (H) /ka- ~ ki-/ POT/INDEF FUT (Mal,Azo,Sut)  
is from pOM \*ki# ~ \*ka# POT, position 3.
- ma- DEF FUT (Mal) = INDEF PRES (Azo) = 'gerundial' (Sut)  
is from pOM \*ma# FUT, position 4.

IMV (OM 3)

- a- IMV (Sut)  
is from pOM \*ha.. IMPV, position 3.
- da- IMV (Mal,Azo) /ta-/  
is from pOM \*tau.. OPT, position 3.

ENCLITIC

- +ma 'already' ~ PERF  
is from pOM \*nma# 'already', position 4.

OM NEG

ta- (~ ti-) (H) /tta- ~ tti-/ NEG PAST (Mal)  
seems to be a conflation of pOM \*tau# ~ \*tta# NEG 1 and pOM \*(Y)ti+ PERF,

position 4.

nda: (L) /nta:/ `there isn't any'  
is from pOM \*tau# ~ \*tta# NEG 1, as above.

-s^a- (H) NEG (w/ FUT)  
is from pOM \*sa(u)# NEG 2. It has switched places with ma- DEF FUT.

a- NEG (Sut)  
is from pOM \*\*7au# NEG 3.

ci- NEG PRES is cognate with Chn 7a-]cyi `it is not' < WOM \*ci#.

## DISCUSSION

Tlp-Sut has restructured the OM TAMVN string in the following ways:

Where pOM has a string NEG-ADV/T-AM-PLSUBJ-AUX-DERIV-VERB  
POSITION           4    3    2    1    ∅

Tlp-Sut has a string ADV/AUX-NEG-TAM-PLSUBJ-DERIV-VERB-ENCL

This has been achieved by removing some items from OM position 4 (kka- PERF PASS, +ma PERF) and then combining what was left in position 4 with position 3 to form the Tlp-Sut TAM position. The category INC AUX has been eliminated. Some ADV/AUX elements have been created in Tlp-Sut out of elements not otherwise known in OM. The above describes what Sut has. Mal has changed the system further. Mal has taken the morpheme string NEG - IMV and changed it to IMV - NEG - TAM; it has taken the morpheme string NEG - FUT and changed it to FUT - NEG; and it has taken the morpheme string NEG - PAST and fused the two morphemes into a set of alternating allomorphs.

#### **3.2.4. Summary on Tlapanekan TAMVN Markers**

The foregoing has been a demonstration that Tlapaneko-Sutiaba belongs to the Tlapaneko-Mangean branch of the Western division of the OtoMangean stock. Further support is found in the following section, where comparative OtoMangean phonology is sketched. This is the first step in demolishing Greenberg's assignment of Tlapaneko-Sutiaba to Hokan. I suggest that no demonstration linking Tlapaneko-Sutiaba to Hokan in the way I have linked it to OtoMangean is feasible or even conceivable.

#### **4. Phonological correspondences linking Tlapanekan to OtoMangean [FN 8]**

##### **4.1. Outline of Proto-OtoMangean Phonology**

I outline here my current view of comparative OM phonology. This was developed mainly between Spring 1982 and Spring 1983. The following Table (Table 5) shows the pOM phonemes and their distributions.



TABLE 5

PROTO-OTOMANGEAN PHONOLOGY

Primary data source:

Calvin Rensch, Comparative Otomangean Phonology.

University of Pennsylvania dissertation, 1966 (published 1976).

Proto-OtoMangean phonemes.

Consonants (C):

K:	t	c	k	kW	ʔ	} :H
S:	[θ]	s	x	xW	h	
	[ʔ]					
R:	l	r				
N:	n			m		
W:		y		w		

<u>Vowels (V) :</u>			
high	i	ia	u
mid	e	ea	o
low	ai	a	au
	front	central	back rounded
	unrounded	or back	
		unrounded	

Tones (T) :

Not yet worked out.

---

Table 6 shows the attested pOM 2-consonant clusters. As many as 3 preposed elements may occur with a given root-initial consonant; such more complex combinations have not been displayed.

TABLE 6

ATTESTED TWO-CONSONANT pOM CLUSTERS:  
ONSET PRECEDED BY ONE PREPOSED ELEMENT

[ ]: clearly attested in one branch only

<u>Alone</u>	<u>YC</u>	<u>NC</u>	<u>hC</u>	<u>7C</u>	<u>TT</u>
t	Yt	nt	ht	7t	tt
c	Yc	nc	hc	7c	cc
k	Yk	nk	hk	7k	kk
kW	[YkW]	nkW	hkW	7kW	kWkW
s	Ys	ns => nc	hs	?	
x	Yx	nx	#	?	
xW	[YxW]	nxW	#	?	
r	Yr	[nr]	?	?	
l	Yl	[nl]	hl	?	
n	Yn	#	hn	7n	
m	?	#	hm	7m	

y	#	ny	hy	7y
w	Yw	nw	hw	7w

---

### OtoMangean Consonants

Besides the 9 consonants (\*\*t, \*\*k, \*\*kW, \*\*s, \*\*n, \*\*y, \*\*w, \*\*h, \*\*7) reconstructed by Rensch, I find 6 more: \*\*c, \*\*x, \*\*xW, \*\*m, \*\*l, \*\*r [FN 8']. While I accept that syllable-initial consonants could be preceded by \*\*n, \*\*h, \*\*7, and \*\*y or \*\*Y (though \*\*y is phonetically a palatalizing feature) and that a vowel could be followed by \*\*n, \*\*h, or \*\*7, I also find that solitary \*\*h and \*\*7 can begin syllables. I postulate a different order for preconsonantal and postvocalic elements from that proposed by Rensch, namely \*HnyC and Vnh7 (CR has YnHC and h7n). I also postulate initial \*\*C<sub>1</sub>C<sub>1</sub> on the basis of Tln, OP and perhaps Sn evidence, though the first C may in fact reflect diverse consonants (other than \*h, 7, n, y).

Tln and Chn, which both have p ≠ kW, suggest that \*\*p may have to be reconstructed for pOM. One would like to avoid this, although Chn \*p, which is rare, occurs in a few OM etymologies, and Tln \*kW, which is also rare, also occurs in a few OM etymologies.

Table 7 presents selected OM consonant and morph(eme) structure correspondences. For certain reconstructions involving H + C, I have cited CR's hypotheses (in square brackets) where I have not yet worked out my own.

TABLE 7  
 SELECTED OTOMANGEAN CONSONANT AND MORPH (EME)  
 STRUCTURE CORRESPONDENCES

OM	Tln	OP	CM	Mn	Ams	Chn	Pn	Sn
t	[d]	t/n	t	t	t	t	t	t
c	c	c	s	s	c	s	c	c
s	s <sup>^</sup>	s	s	s	c	s	s	s
k	[g]	k	k	k	k	k	k	k
x	[g]	h	h	x	k;h[1]	h	h	k
h <sup>[2]</sup>	h?	h	h	∅	h	H[3];h	h	h
kW	[b]; [gW]	p/m;ko	p	kW;ku	kW	kW;p?	kW;ku	kW
n	n	n	n	n	n	n	n	n
l[4]	r	r	r	r	h[3];l[8]l	l	l	l
r[8]	[d]	r	r	r	h	r	t?	t
w	w	w	w	w	w;m[6]	w	w;m[6]	w
Y	Y	Y	Y	Y	Y	Y	Y	Y
7	(?)	7	7[7];0[8]7	7	7	7	7	0[3];7[8]
x <sup>w</sup>	[b]; [gW]	hw	hw	hW	kW	hw;hm[6]	hw	kW
hw	hw?	hw	hw	w	wh	hw;hm[6]	hw;hm[6]	w
m	m	m	m	m	m	m	m	kW
nw	m	nw	m	m	m	m	m	KkW

<sup>1</sup> in clusters

<sup>2</sup> non-final

<sup>3</sup> \_\_V

<sup>4</sup> 6 examples

<sup>5</sup> 8 examples

<sup>6</sup> nasal

<sup>7</sup> V\_\_

<sup>8</sup> elsewhere

(continues on next page)

OM	Tln	OP	CM	Mn	Ams	Chn	Pn	Sn
Ky	c^	K	K	K	K <sup>y</sup>	Ky	K	tY;k
sy	s^	s	s	s	c	sy	s	s^
ly[9]	(?)	ry	r	y	y?	ly	y	l
ry[10]	(?)	ry	r?	y?	y?	r?	tY	tY
//Yn//	(?)	n	n?	n	nY	ny	nY	n
//Ny//	(?)	ny	ny	n~ = ny	yV?	ny[10]	nY	n
NK	NG	nK	nG	nK	nG	G	nG	nK
nx	ng	(n)h	(?)	nx	(?)	g; 9[6]	nh?	nk
nxW	mb	nhw	np	?	nKW?	gW	(?)	nkW
KK	KK	KK	(?)	(?)	(?)	(?)	(?)	KK
7K	7G	K7;b	(?)	[7CV] [11]	K7	K	K7	7K
7R	7R	7R	(?)	[7CV] [11]	7R	7R	7R	7R
hK	hK	Kh	hk=>h	[hCV] [11]	Kh	K	Kh	hK
hR	hR	Rh	(?)	[hCV] [11]	hR	hR	hR	hR
Vn#	V&	V&	V	V&	V&	V&	V&	V&
Vh#	(?)	Vh	(?)	[Vh] [11]	Vh[12]	Vh[12]	[hV] [11]	Vh
C <sub>1</sub> VC <sub>2</sub> V	CVCV	(C <sub>1</sub> )C <sub>2</sub> V	CVCV	CVCV	(C <sub>1</sub> )C <sub>2</sub> V	C <sub>2</sub> V	C(V)CV	CVCV
HNCyV	HNCyV	NKHV;HRV	nGV	7nCV	HNGyV	HGyV	HNCyV	7NCyV

---

<sup>6</sup> nasal

<sup>9</sup> 12 examples

<sup>10</sup> 1 example

<sup>11</sup> CR

<sup>12</sup> ballistic

---

## OtoMangean Vowels

Besides the 4 vowels (\*\*i, \*\*e, \*\*a, \*\*u) postulated by Rensch, I posit 4 diphthongs and an additional vowel: \*\*ia, \*\*ea, \*\*au, \*\*ai, and \*\*o (= \*\*ua?).

Nasal vowels are fairly rare in all OM branches, and V nasality often does not agree across branches, or even within the same branch. The most natural assumption, to me, is that in many cases nasality represents a suffix that was differentially present in different ancestors, at various time levels.

Table 8 shows the recurrent vowel correspondences I have identified. Two points should be borne in mind:

(a) CM and Tln data is still very skimpy (i.e., the number of cognates found is very small); consequently reconstruction of the pOM system or of individual OM etyma can hardly be crucially based on them.

(b) Those OM vowels whose normal Chn reflex is \*a also have \*ea, \*u and \*ia reflexes. Rensch had to resort to some fancy footwork to derive these reflexes, and I have not yet succeeded in accounting for them, either. I think they may be attributable to some disappeared enclitics.

There may be other, rarer, V correspondences conditioned by preceding C or by postposed elements other than \*n.

TABLE 8

## RECURRENT OTOMANGEAN VOWEL CORRESPONDENCES

**pOM	*Tln	*OP	*CM	*Mn	Ams	*Chn	*Pn	*Sn
i 53 exx	i	i	i;u <sup>1</sup>	i	i/e	i	i	i
ia 15 exx	u;CY]i	o	e;i;u? <sup>2</sup>	i	i/e	u	i	i
e 58 exx	e;CY]i	e	e;i	e;CY]i	e/3	ɨ;CY]i	e;CY]i	e;CY]i
ea 20 exx	u;CY]i	a	e;i(a) <sup>2</sup>	e;CY]I	e/3;o[n	u	e	e;CY]i
ai 15 exx	a	a	a	e	a;e	a(6)	e	a
a 18 exx	a	a	a;e	a	a	a(6,u,ɨ)	a	a
au 62 exx	a	o	u	u	o/2	u	u	u
u 37 exx	u?	o	u	u	o/2	u	u	u
o 30 exx	o?	u	o	u	o	a	u	o?



---

<sup>1</sup> 2 exx; nasal?

<sup>2</sup> 2 exx

---

[In Table 8, <3> = aesc, <2> = open o, <6> = schwa]

TABLE 9  
SYMBOLS USED FOR INDETERMINANCIES

---

I	i or ia	A	a or au
E	e or ea	U	o or u
E <sub>2</sub>	i or e	O <sub>2</sub>	au or u
E <sub>3</sub>	ia or ea	A <sub>2</sub>	ai or a
E <sub>4</sub>	e or ai	HK	x or hk
V <sub>f</sub>	i, ia, e, or ai	X	h or x
S	s or c	M	m or nw
L	l or r		

Hypotheses about OM word and phrase structure:

- a) Roots are typically monosyllabic.
- b) There are prefixes, both syllabic [not yet identified] and nonsyllabic (n, ʔ, h, y, K).
- c) Stems can be formed by compounding two monosyllabic roots.
- d) There are proclitics of tense and aspect found in front of verbs.
- e) There are (proclitic) noun classifiers found in front of nouns.
- f) There are syllabic enclitic particles of shapes  $**hV(n)$ ,  $*hV(n)ʔ$ , and  $*ʔV(n)$ .  
(Or  $**xV$ , etc.)
- g) There are also postposed elements  $**n$ ,  $**ʔ$ , and  $**h$ .
- h) Stress falls on the (last) vowel of the root.

The consequences of these hypotheses are that stems and tight phrases can be of one, two, three, or four syllables, with the following stress patterns (I use # after proclitics and + before enclitics):

- 1) `CV (root)
- 2) `CV+XV (root + enclitic)
- 3) {
  - CV' CV (prefix + root, root + root)
  - CV#' CV (proclitic + root)
  - CV' CV+XV (above + proclitic)
- 4) {
  - CV#' CV+XV (proclitic + root + enclitic)
- 5) CV#CV' CV (proclitic + prefix [or root] + root)
- 6) CV#CV' CV+XV (proclitic + prefix [or root] + root + enclitic).

To these hypotheses I add the following observations. Some branches (Mn) lost the vowel of the enclitic. Some languages (Mas, probably CM, maybe Tln) shifted the stress to the enclitic. Some languages (Sap, Mis, Kwi) shifted stress to the first syllable of stems. Some languages (Ams, Chn, OP) are so desperately monosyllabic that all 6 patterns listed above show a monosyllabic reflex. Some of these may (God forbid!) have shifted stress before undergoing syllabic crunching, so that the surviving vowel would not be the root vowel. Sap, Chp and Tln seem most promising for identifying the fuller forms of all these syllabic elements.

## 5. Notes on Comparative Hokan

### 5.1. A Series of Hypotheses about Proto-Hokan Phonological Structure.

This section outlines my hypotheses about proto-Hokan phonological structure. [FN 9]

**Phonemes.** I postulate that proto-Hokan had the following set of phonemes and/or contrasting phonological entities (some of which may turn out to be analyzable as clusters):

c o n s o n a n t s									v o w e l s	
p	t.	t	c	c <sup>^</sup>	k <sup>y</sup>	q	kW	ʔ	i	u
p'	t.'	t'	c'	c <sup>^</sup> '	kY'	q'	kW'		e	o
p''	t.''	t''	c''	c <sup>^</sup> ''	kY''	q''	kW''		a	
f		θ	s	s <sup>^</sup>	xY	x.	xW	h	length /:/	
	r								stress /' /	
	l	lY								
m	n	nY								
			y			w				

NOTE: \*t is postulated to have been apicodental, and in my handwritten notes I always write it with a half-circle underneath. Lack of a convenient way to keyboard it has led me to omit that diacritic in this document with no loss of information, but it should be kept in mind that I consider it to be a marked segment in contrast to apicoalveolar \*t., which I categorize as unmarked.

**Diphthongs.** clusters of vowel + semivowel are found in virtually all Hokan languages, and the most commonly occurring are /ay/, /aw/, /uy/, and /iw/. It appears that without positing /ey/, /ow/, /oy/, /ew/, /iy/, and /uw/ (that is all the other possible combinations), we cannot plausibly account for all the vowel correspondences among the Hokan languages. Working out the details, however, will take some time. If we need no more than four simple vowels, then the potential diphthongal contrasts /uy/:/oy/, /uw/:/ow/, and /ow/:/aw/ would disappear, providing much more work for /aw/, /uw/, /iw/, and /ew/, and the long diphthongs.

### **Phoneme Alternations.**

**Consonantal Symbolism.** In the etymologies I have found acceptable, there are numerous instances of multiple or discrepant reflexes among the apical resonants \*l, \*n, and \*r, and among the labial resonants \*w and \*m. Since alternations among these phonemes (or their reflexes) are found actively in a number of Hokan languages, and lexically frozen in others, I attribute alternation among these phonemes to proto-Hokan. The alternation among apicals seems to be associated mainly with size and respect symbolism, i.e. neutral:small:large, neutral:nice:nasty. Possibly \*l is neutral, \*n is little/nice, and \*r is big/nasty. It does not necessarily follow that all instances of \*n and \*r exhibit symbolism. \*n is an extremely common phoneme. \*r is not very common, but I would not expect all of its occurrences to be accountable for by the effects of sound symbolism. I do not so far have a theory about the conditions for alternation among the labials, but \*w may be neutral and \*m 'little/nice'.

**Vocalic Ablaut.** In otherwise acceptable etymologies, there exist discrepancies among the indicated proto-Hokan vowels of the following three types, the more frequent being cited first in each case: both \*a and \*o, both \*i and \*a, both \*i and \*u. The first case is extremely common, the second less so, and the third very infrequent. These alternations seem to occur mainly in verbs and nouns that plausibly derive from them.

There is in Yana considerable similarity to the postulated proto-Hokan \*a ~ \*o and \*i ~ \*a alternations. Yuman provides analogues for both the \*a ~ \*o and \*i ~ \*u alternations of pHok. Shasta provides a morphophonemic analog for the pHok \*i ~ \*a alternation, and frozen analogs for all three of the postulated pHok alternations.

It has not escaped my notice that in the daughter languages discussed here the vowels that alternate are just the three "apex" vowels /i a u/. I am not willing, however, to concede at this point that this suggests that pHok had just three vowels. The languages in question could all have undergone mergers of distinct pHok vowels. The issue is open, though, and I could change my mind.

**Fronting/Palatalization.** In a sizeable number of etymologies there is multiple or conflicting evidence for the reconstruction of the following pairs of contrasts: \*t. and \*tʰ, \*c abd \*cʰ, \*q and \*kʸ, \*s and \*sʰ, \*x. and \*xʸ. In the reconstructions generally the first member of each alternating set is lexically more frequent. My hypothesis is that the second, lexically less frequent, member of each set is derived from the first member by a phonological process of fronting (or palatalization). I do not believe this fronting to be a case of ablaut, but to involve affixation, usually of a derivational/lexical nature (rather than being morphosyntactic). It often affects what seems to be the first consonant of a root, and thus it can function like a prefix. At the historical phonological state we find proto-Hokan in, it does not seem feasible to identify the affix as a preposed or postposed segment /y/, but as our understanding of proto-Hokan develops, it may become feasible. As an interim measure, I will call this feature /J/, order it as a prefix to the consonant it affects, and stipulate that it may represent more than one grammatical category, without being able to say more on the topic at present. I will write the alternations referred to as \*(J)t, \*(J)c, \*(J)k, \*(J)s, and \*(J)x respectively. I would expect \*l and \*lʸ, \*n and \*nʸ to enter into a parallel set of alternations, but so far instances of these alternations have escaped my attention. From the above formulation, it does not necessarily follow that all instances of \*t, \*cʰ, \*kʸ, \*sʰ, and \*xʸ contain morpheme boundaries between the two phonological elements, nor is it necessarily entailed that the set of proto-Hokan phonological contrasts should or can be reduced by reanalyzing \*t as \*Jt, \*cʰ as \*Jc, \*kʸ as \*Jk, \*sʰ as \*Js and \*xʸ as \*Jx. I confess, however, that the idea is attractive. Since \*θ has among its reflexes mostly [y], it seems likely that \*θ was a palatalized segment (although other possibilities come to mind). There is no palatalized/fronted segment to match \*r, unless it is \*θ. Both \*f and \*θ have voiced as well as voiceless reflexes in one or another language, but \*r has no known voiceless reflexes. The idea should be kept in mind in making further refinements of proto-Hokan phonology.

**Multiple Reconstructions.** In a small number of instances, in a given etymology there is evidence for reconstructing more than one phonologically similar proto-Hokan phoneme. Some of the more common types of discrepancy are \*plain vs. \*glottalized, \*plain vs. \*aspirated, \*aspirated vs. \*glottalized, \*s vs. \*ç, \*s\* vs. \*c\*.

These discrepancies will have to be explained. My current view is that they are probably irregularities that will be explained on a case-by-case basis rather than being clues to a more subtle understanding of proto-Hokan phonology, but time will tell.

### **Phoneme Distributions** (Phonotactics).

C = consonant  
V = vowel  
H = laryngeal /h ʔ :/  
Y = semivowel /w y/  
\$ = syllable boundary

I postulate the following basic phonological structure for a proto-Hokan lexical item of one to three syllables that is not a compound:

$$\#([C(x)]V[H])\$C(x/w)V(H)(Y)(C)\$(+CV)$$

The minimum structure is CV. The C may = /ʔ/

A vowel may be long, or followed by /h/ or /ʔ/, even though another consonant follows in the same syllable.

A word may contain an enclitic (or unstressable) final syllable.

If a word contains two syllables the first may be a prefix.

A syllable-initial plain (not glottalized, not aspirated, and probably not fronted) plosive

may be followed by /x./. This suggests that glottalization and aspiration might be in origin /ʔ/ and /h/ after plosive. However, in the daughter languages glottalized plosives do not seem amenable to a /C + ʔ/ analysis, and for aspirated plosives only Chimariko to my knowledge seems hospitable to a /C + h/ analysis. Nevertheless, between them Silver (1976), Jacobsen (1959, 1976), Talmy, and Moshinsky (1976) have suggested that the aspirated stops of Yana[n], Washu, Pomoan, Atsugewi, and Chimariko have come primarily from consonant clusters (Jacobsen 1976.234-236). It should be noted that I explicitly postulate both aspirated plosives and plosive + /x./ clusters for proto-Hokan, and find them to be in contrast. I would not be disappointed if it were eventually feasible to analyze the pHok aspirated plosives as clusters of plosive + /h/.

A syllable-initial obstruent may apparently be followed by /w/, but there are only two cases where such a reconstruction has seemed unavoidable: #(a)s<sup>h</sup>wa 'fish', and #iHpwa 'tail'. I would rather look forward to more such clusters than to find ways around reconstructing consonant + w. If no more are forthcoming I will consider ways to avoid positing these clusters.

In a very preliminary way, all possible /VY/ and /V:Y/ clusters seem to be required for proto-Hokan to account for the observed vowel +/- semivowel correspondences, but I would look favorably on an effective explanation of pHok \*[e:] and \*[o:] as underlying /ay/ and /aw/. If, for example, it could be shown that both \*[ay] and \*[a:y], \*[aw] and \*[a:w] were not required, and that \*[e] and \*[o] were correlated with adjacent consonant qualities and/or the qualities of vowels in near-by syllables, an argument for a three-vowel system with length and all possible V(:)(Y) combinations might begin to seem plausible. Shirley Silver and Margaret Langdon have both expressed a preference for such a formulation.

**Stress in Comparative Perspective.** A good hypothesis about **stress** eludes me. Most Hokan languages have free stress, but some have predictable stress. Many languages show vowel dropping of syllables that must have been unstressed when they were dropped. I tentatively hypothesize that in proto-Hokan stress could occur on any syllable of a lexical item (except the by definition unstressable enclitics), but I suspect that a deeper understanding of the morphological patterning of proto-Hokan and its evolution in the various daughters might render much of the stress phenomena predictable on a combination of morphological and segmental



grounds.

**An Alternative View of Proto-Hokan Phonological Structure.** If all the simplifications contemplated in the preceding paragraphs should pan out, we could postulate the following phonological system for proto-Hokan:

P H O N E M E S

c o n s o n a n t s						v o w e l s	
p	t	c	k	kW	ʔ	i	u
p'	t'	c'	k'	kW'		a	
f	θ	s	x	xW	h	length	/:/
	r						
	l						
m	n						
	y		w				

D I S T R I B U T I O N S i n t h e s y l l a b l e

C(h/x/w) (J)V(H) (Y) (C[h/x] [J])

(J is the fronting feature which was probably a postposed [or preposed] /y/)

## A P O P H O N I C   A L T E R N A T I O N S

**a ~ u; i ~ a; u ~ i; l ~ n ~ r; w ~ m**

**Cover Symbols.** In many etymologies the set of languages that the etymon survives in does not allow for complete specificity in the reconstruction. Rather than state all possibilities in such reconstructions, I have devised a set of cover symbols standing for the most frequent specific ambiguous cases, and the reconstructions have been symbolized in those terms. In the instances of ambiguity not covered by these symbols, I have stated the alternative possible reconstructions.

\*P = \*p or \*f  
\*Y = \*y or \*θ  
\*R = \*l or \*r  
\*L = \*l or \*lY  
\*N = \*n or \*nY  
\*C = \*c or \*c^  
\*T = \*t. or \*t  
\*K = \*q or \*kY  
\*S = \*S or \*s^  
\*X = \*x. or \*xY  
\*I = \*i or \*e  
\*U = \*u or \*o  
\*X! = \*x., \*xY, or \*xW  
\*K! = \*q, \*kY, or \*kW

There are four additional cover symbols that stand for conflicting attestations rather than ambiguity; that is, some languages support one reconstruction, while others support a different reconstruction, but I am not convinced that there was more than one pronunciation of the etymon

in proto-Hokan.

\*E = there is evidence for both \*e and \*i

\*O = there is evidence for both \*o and \*u

\*A = there is evidence for both \*a and \*e

\*H = there is evidence for \*h and/or \*ʔ and/or \*: [: = vowel length]

**On Polysyllabic Reconstructions.** If straightforward phonological comparison leads to the reconstruction of lexemes of more than two syllables, I think that it is appropriate to appeal to the following heuristic: Lexemes of three or more syllables are probably (a) morphemically complex, (b) borrowed, (c) "morphemized", or (d) invented. That is, monosyllabic and disyllabic morphemes are to be expected, tri- and multisyllabic morphemes are not. Many of the reconstructions of Sapir and Gursky have three syllables, and many of Haas's reconstructions have more than three syllables. A theory of Hokan word structure must be devised to account for and enable us to decompose multisyllabic reconstructions that seem otherwise valid. In the absence of such an analytic and explanatory framework, multisyllabic reconstructions ought to remain suspect.

## 5.2. Hokan Comparative Grammar [FN 10]

I have worked out quite a number of specific hypotheses about the grammatical structure of proto-Hokan, of which I will outline several here.

### NOUNS

The NOUN STEM is made up of a ROOT optionally followed by a NOMINALIZER (4 exx), or a FIRST-ORDER NOMINAL SUFFIX (1 ex) and a SECOND-ORDER NOMINAL SUFFIX (3 exx). In Pomo and Chimariko a noun root may be preceded by another noun root to form a noun-noun compound with the first noun modifying the second; Salina and Yuma have no such compounding, and the status of such compounds in proto-Hokan is still in doubt.

(ROOT) **ROOT** {                    } | (NOMLZ) |  
  | (NOM<sub>1</sub>) | (NOM<sub>2</sub>)

The NOUN WORD consists of a NOUN STEM plus up to two preposed optional grammatical markers, a POSSESSED STATE PREFIX [Pom, Sal, Yum] (3 exx), and a PROCLITIC CLASSIFIER (6 exx), one optional POSSESSED STATE SUFFIX (1 ex) and one obligatory CASE SUFFIX (8 exx). The classifier has become a prefix in many languages. It is possible that future study may lead to a slight modification in the order statement about prenominal inflexions. The cases are both locative and relational, but the only relational cases that have etymologies are limited to Southern Hokan and encode switch-reference (**same subject** versus **different subject**).

(CLASS) (POSS) **NOUNSTEM** (POSS) CASE

## VERBS

The VERB STEM is made up of a ROOT optionally preceded by a STATIVIZER (2 exx), a CAUSATIVIZER (4 exx), or an incorporated INSTRUMENTAL PREPOUND [Pom, Chi, Yan ("primary verbs"), ?Sha, pre-Kar, Ach, Ats, Wsh, Yum, Ser, pre-Cho] (9 exx) and optionally followed by a FIRST-ORDER VERBALIZER (3 exx) and an incorporated DIRECTIONAL POSTPOUND [Pom, Chi, Yan, Kar, Sha, Ats, Wsh, Yum, Cho] (18 exx + 9 from NC).

Both Northern and Southern languages have instrumental preponds. Only Salina seems definitely to lack them.

Both Northern and Southern languages have directional/locative postponds. Again, only Salina seems definitely to lack them.

It is not clear whether there is evidence for simple noun incorporation or verb root compounding. No incorporation is found in ?Pom, Sal, or Yum. V-N incorporation is found in Yan, but is probably an innovation reflecting its VO syntax. Langdon 1988 shows that compound verb stems whose first member is not necessarily instrumental occur in Yana, Shasta, Atsugewi, and Washu.

(STAT) | (CAUS) } **ROOT** (VRBLZ<sub>1</sub>) (DIR)  
(INSTR) |

The VERB WORD (or VERB COMPLEX) consists of a VERB STEM plus up to three preposed optional grammatical markers and up to 5 postposed grammatical markers, of which one is obligatory. The preposed markers are [-1] PLURALIZER PREFIX (2 exx), [-2] FUTURE PROCLITIC (1 ex), [-3] TEMPORAL SUBORDINATOR PROCLITIC (2 exx). The postposed markers are [+1] PASSIVIZER SUFFIX (1 ex) or SHIFTER (2 exx), [+2] ANDATIVE SUFFIX (2 exx: Northern Hoka only), [+3] obligatory TAM<sub>1</sub> SUFFIX (8 exx), [+4] TAM<sub>2</sub> SUFFIX (2 exx), [+5] TAM<sub>3</sub> ENCLITIC (3 exx). VERB-TAM order is found generally in Hoka: Pom, Chi, Yan, Kar, Wsh, Sal, Yum, Cho.

(TEMP (FUT) (PL) **VERBSTEM** (PASS) (ANDAT) TAM<sub>1</sub> (TAM<sub>2</sub>) (TAM<sub>3</sub>) SUBORD)

## ADJECTIVES

Some adjective-like words act like nouns and some act like verbs. It is likely that proto-Hokan had both kinds, and while it is possible that adjectives as a class had no independent existence, is equally possible that there were two kinds of adjectives which were neither nouns nor verbs, as there are in Nahua, Mayan and Bantu languages, to name just three cases.

## NEGATIVES

There are four widespread etyma with 'negative' meaning. This is quite a few, but not so many as to arouse suspicion; OtoMangean has three or four, Mayan has at least two, and IndoEuropean has two. If proto-Hokan indeed had four negative morphemes, they must have had at least partially different syntactic distributions.

## 7. Comparing OM and Hokan

I now offer a brief comparison of OtoMangean and Hokan, which I suggest supports the hypothesis of a genetic connection between the two stocks. Comparison are made both of typological patterns (phonemes, phonotactics, and morphological patterning) and of potentially cognate morphemes.

### 7.1. Phonological Comparisons

Phonemes.

Proto-OtoMangean phonemes.

#### Consonants (C):

K:	t	c	k	kW	ʔ		
S:	[θ]	s	x	xW	h	}	:H
	?						
R:	l	r					
N:	n			m			
W:		y		w			

#### Vowels (V):

high	i	ia	u
mid	e	ea	o
low	ai	a	au
	front	central	back

#### Tones (T):

Not yet worked out.

Proto-Hokan Phonemes (revised version).

c o n s o n a n t s						v o w e l s	
p	t	c	k	k <sup>w</sup>	ʔ	i	u
p'	t'	c'	k'	k <sup>w</sup> '		a	
f	θ	s	x	x <sup>w</sup>	h	length	/:/
	r						
	l						
m	n						
	y		w				

proto-Hokan apophonic alternations.

**a ~ u; i ~ a; u ~ i; l ~ n ~ r; w ~ m**



We now compare the OM and Hokan systems point-for-point.

OM	--	t	c	k	kW	7
Hok	p	t	c	k	kW	7

OM	--	r =	θ	s	x	xW	h
Hok	f	r	θ	s	x	xW	h

OM	m	n	l	y	w
Hok	m	n	l	y	w

OM	hC	7	YC	Vh
Hok	c''	C'	CY	V:

OM	i	ia	e	ea	ai	a	au	o =	ua	u
Hok	i	--	iy	--	ay	a	aw	uw		u

Note that Hokan \*p and \*f lack equivalents in pOM, and that pOM \*ia and \*ea lack direct equivalents in pHok; otherwise the systems are quite similar.

Phonological canons.

Proto-OtoMangean distributions in the syllable.

Minimum syllable:

T  
CV

Any C, including H, may begin a syllable.

Maximum syllable:

T  
(H) (n) (Y) CV(n) (h) (ʔ)

When C is H, H and perhaps Y may not precede.

//YC// was [cY] phonetically; I usually write it CY.

C<sub>1</sub>C<sub>1</sub> possibly also may begin a syllable, probably without other preposed elements. The first C may represent various consonants other than those already identified (i.e., n, h, ʔ, y).

Proto-Hokan distributions in the syllable.

C(h/x/w) (J)V(H) (Y) (C[h/x] [J])

(J is the fronting feature which was probably a postposed [or preposed] /y/)

Comparison of syllable canons. The formulas for pOM and pHok syllables are reordered in order to facilitate comparison. Reordered elements are marked with !.

OM	(H)	(n)	(Y)	C	V	(n)	(h) (ʔ)	T!
Hok	(h/x/w)		(J)!	C	V	(Y)	(H)	C[h/x] [J]

Morpheme/Lexeme structure.

OtoMangean: A goodly number (ca. 50) of two-syllable OM-level reconstructions are feasible on available data.

Hokan: Straightforward phonological comparison often leads to the reconstruction of lexemes of two syllables.

**Discussion.** The generous reader will note a large amount of typological similarity between the OM and Hokan phonological systems.

## 7.2. Morpheme comparisons.

The following morpheme comparisons are the result of a single pass through my OM and Hokan cognate files. Hokan etymologies were matched to the approximately 500 OM etymologies. Glosses were matched, but neither file was ordered by the glosses of the reconstructions. The OM file is ordered alphabetically by the pOM reconstructed form, and the Hokan file is ordered by semantic fields. Consequently, a few potential semantic matches may have escaped me; I estimate the maximum number of overlooked semantic matches at 40.

Proto-Hokan \*θ has been transcribed here as \*θ<sup>y</sup>, given the evidence (Kaufman 1989.99-100) that it may have been a palatalized segment. Proto-Hokan front \*t has been rendered here as \*t<sup>y</sup>, since I hypothesize that there was a back:front = plain:palatal contrast with all consonants except labials and labiovelars.

OM reconstructions refer to the proto-OtoMangean level unless otherwise noted. The Hokan 'reconstructions' are preceded by # because the phonological reconstruction of proto-Hokan is still in progress and subject/liable to a serious reinterpretation. These Hokan forms are found in both Northern and Southern languages (as defined here) unless otherwise noted.

In these comparisons no effort has (so far) been made to weed out 'pan-Americanisms' (items that are so widely spread throughout the New World that they cannot serve to define a special linguistic group among Amerindian languages -- whether these similarities are due to relationship or diffusion is not relevant to their lack of decisiveness in classification).

The phonological correspondences that up to this point have been recognized are for the most part essentially equivalent. Since OM basically lacks [p], corresponding to Hokan [p] OM has [kW] (and [k] before rounded vowels). Affricates and the corresponding spirants occasionally match, and [y] may correspond to a palatalized consonant.

### 7.2.1. Grammatical Morphemes and Patterns.

OM is verb-initial, and its TAM markers are preposed to the verb. Hoka is verb-final, and its TAM markers are mostly suffixed to the verb. Whatever their current status, these markers, when and if cognate, were probably VP-level particles placed according to the VO or OV patterns generally prevailing in the syntax. Noun markers that appear to be comparable are preposed in both OM and Hoka.

In the material compared in the following table (Table 10), the suggested Hoka reconstructions are preceded by # to indicate that the comparative phonology has still not been fully worked out.

Special symbols in reconstructions: + means clitic boundary; = means compound. In glosses + means 'in addition to the gloss given in the left-hand column'. See 4.1 and 5.1 for the meanings of the capital letters, which stand for ambiguities in the available data for reconstructing specific phonemes.

TABLE 10

MEANING	OTOMANGEAN	HOKAN
PRONOMINALS & DEICTICS		
first person	* (n) an	#nYa; #Ha
second person	* (l) i	#nYi
anaphoric; distal	* (h) o	# (h) U
this/that	*sa [EOM] 'thus'	#sa 'yon'; #si 'this'
proximal	*te 'here'	#Ti 'this, here'
proximal	*ya 'now'	#ya 'this, here'
NOMINALS		
demonstrative	*la 'def. art.; rel.'	#La 'he/she/they'
noun prefix	*l- 'plural/mass' [-1 position]	#l- 'absol. mass/pl.' #lY- 'absol. count/sg.' [-2 position]
noun prefix	*c- 'sg.?' [-1 position]	#c <sup>h</sup> -, C <sup>h</sup> - [-2 position]

noun prefix	*sYi- 'possessed' [-1 position]	#s^i- [-2 position]
noun prefix	*i- 'generic' [-1 position]	#Hi:- 'intimately possessed' [-1 position]
TWO	*wi	#-wi [i ~ a] 'dual & plural of NP/PN/Dem'
TOGETHER	*ka +u 'and, with'	?#x.aK'a [N]

#### NUMERALS

TWO	*ha +u	#hag''u; ?#(q-)x.ow(a)
THREE	*hau(n)	#x.a [a ~ o]

#### VERBALS

adjectivizer	*ma+ [-1 position]	#m(a)-[derivational]
TAM	*i+ ~ ∅ 'indef. indic.' [-3 position]	#-i [N] 'pres./impv.' [+3 position]
TAM	*xi+ 'durative <sub>2</sub> ' [-3 position]	#x(J)a 'fut./opt.' [+3 position]
TAM	*tau 'optative' [-3 position]	#-t <sup>y</sup> a 'desid./pol. impv.' [+3 position]

TAM	*t(Y)i+ `perfect' [-4 position]	#-aT' `completive' [+3 position]
TAM	*ni+ `remote past' [-4 position]	#-nI [N] `remote past' [+3 position]
TAM	*ki# ~ #ka# `potential' [-3 position] &	#Ka- `impv./fut.' [-2 position] ~
	*ma# `fut.' [-4 position]	#Kam+ [S] `fut.'
NEG	*sa# ~ *si# ~ *sau# `neg <sub>1</sub> '	#sey [-5 position]
NEG	*tau# ~ *tta# `neg <sub>2</sub> '	#T'V [-5 position]

PARTICLES

NEG	*ku(a) [EOM] `negative'	#kY''u(wa) `negative'
-----	-------------------------	-----------------------



### 7.2.2. Lexical Comparisons

TABLE 11

MEANING	OTOMANGEAN	HOKAN
ARM	*k <sup>w</sup> a '+hand, wing, muller'	#px.aL '+shoulder, knee'; #Pam [ʔasp] [NC] '+hand'
ASHES	*Hma '+smoke, soot, black'	#ma 'to burn' + nom.
BABY	*we7=ne	#(u)wa(y) [S] 'child, offspring' #qawi [N] '+child, boy, girl, small' (cf. #qu 'small, little, young, baby')
BACK	*cU '+tail, rear'	#(i-)Suw [w ~ m] [NC] 'baackbone'
BELLY	*ko 'stomach'	#Cuqu(N) [n ~ l] [N]; #(7)qUw(L) [ʔu ~ a] '+guts'
BIRD	*sai [EOM]	#ŝa 'eagle'
BLACK: see NIGHT		
BONE	*nt(Y)a '+horn'	##i(H)θ <sup>y</sup> a
to CUT, BREAK	*ka	#kYatY [a ~ o] '+to bite'; #q''a(w) '+to cut'; #qaP '+ to split'
BREAST [f]	*ci '+to nurse; drop'	#[i/u]c''i '+milk'; #[i/a]ci(c) 'to suck'
BREAST [f]	*co '+milk' *co(n) 'woman'	#coc' [MA] FEMALE ?#(a)Cu(wV) 'grandmother'

BROTHER	*ku=Xu [WOM] 'man's...'	#7ap"o [NC]
BUG	*ka(h) 'ant, insect'; *kau 'ant'; *hko 'louse, fly'	?#iK'ey 'flea, louse'
to BURN	*(H)me7 'to set fire to'	#ma
to CARRY	*ka '+to bring, take'	#a*Hka [N]; #kY'[a ~ o] [S]
to CHEW	*cYa [OP-Chn]	#c (J)a 'to bite'; 'chewing gum' ?#uc'i [NC] 'to chew'
CLOUD	*xWI	#[a/i-]kW"ey '+rain'
to COVER/CLOSE	*cA [EOM] 'to cover'	#c^'aK [?a ~ o] [N] 'to close'
CROW	*ka	#q(x.)a:q [symbolic]
to DAWN	*(n)xea(n)	#(a)xYa7 'morning, daylight'
DAY	*x <sup>w</sup> e '+name'	#ipe*
EARTH	*lYau [WOM] 'dirty' *yo 'earth, sand, powder'	#yax. [a ~ o] [NC] 'earth, sand, dirt'
to DO	*7ta [OP-Chn] 'work'	#(a)tYa '+to make'
DUST	*ca *se	#tY''a [N] '+dirt' SAND #T'as [N] 'sand' (T'a = tY'a 'dirt'); ?#s^iqx.o7 [N] 'earth' (cpd.?)

EYE	*lo `face`; *(n)tea `+face`	#[u/i-]θ <sup>y</sup> iw `eye, face`; ?#C <sup>h</sup> iw [NC]
FATHER	*ta	#tYatY(a) [symbolic]
FEMALE:	see BREAST	
FIRE	*yu	#iyu [N] `(to catch) fire`
FISH	*ka	#KaL [K ~ k'] [N] `+trout, eel`
FOOD	*ma	#ama `+to eat`
to GIVE	*k(Y)e `+to pay, lend`	#[a/i]qa [a ~ i]
to GO	*ha7 [EOM]	#ha
to GO/COME	*kWa [OP-Chn] `to go`	#i7pa [i ~ a] `to arrive`
GRANDMOTHER	*k <sup>w</sup> au(n) [EOM] `aunt`	#k <sup>w</sup> a:(wV) `MoFa`
GRASS	*ca [OP-Chn] `hay, fodder`	#x <sup>y</sup> aCa
GRASS/HAY	*isYi	?#qx.atsi(r) [NC] (cpd.?)
HAND	*tYaH	#i-tYa(lY) `+arm, shoulder`
HAND	*hlYo [COM]	#iL [NC] `to do with...`; #Ila [NC] `to carry`
HEART	*(n)se7ea(n) `+guts`	#(X)a*sa(L) [a* ~ o], [l ~ n]

HEART	* (7)wa `+stomach'	#i-waHy ~ yaHy; #-C-wa(:y) [NC] `chest'; #uwa [NC] `liver'
to HIT	*(n)pah [symbolic]	#pa; also #(u)PakY", #PaK, #PaT, #Pa*L(a)
HOLE	*sau `+open'	#CulY [C ~ S]
HOUSE	*(n)ku=wa7	#a/i:-wa*
INSIDE	*ki [EOM] `+under, stomach'	#Ki `at'
to LAUGH	*t(Y)e	?#asiy +o [i ~ a]
LEAF	*k <sup>w</sup> a `+bark, skin'	#p <sup>h</sup> A
LEG	*Caku	#Ku `+foot'
LEG	*t(Y)e	?#sey [N] `+foot'
LIVER	*ya	#e*ya [S]
LONG	*kau `+big, tall, high'	#aqolY; #qam `+big/large, tall'
LONG/FAR	*(n)tU	*To [S] `+tall'
LOUSE	*awe	#aHkWi
LOUSE	*7tYi	#(a)c'(J)i(n)

LUMINARY	*sa `moon'	#s^a `star, clear, shining'
MALE	*caH	#(Ka:)c^a `brother'
MAN/HUSBAND	*[X]wi `+lord'	?#awa [NC] `+person'
MOTHER	*mau [WOM]	#ma [S] `+breast, milk'; [symbolic] #[a/i-]ma:(wV) `FaMo'; #ima*tYV `+woman, breast, milk' (#tYa `woman, female'); #mam [a~ o] `woman, female'; #mari [NC] `woman'
MOUTH	*Co7wau [EOM]	#(h)a:wa [N]
NIGHT/BLACK	*(n)t(Y)u(n) `black'; *to [Amz-Mn] `soot'	#tYu [N] ~ Tiy= [S] `night'
NIGHT/STAR	*la `+sleep'; *lYai [EOM] `black'	?#(x.)alVm [N] `night', pitch dark'; ?#al7a `sun; asterism' ?#n <sup>y</sup> il <sup>y</sup> [S] `black, to write'
NOSE	*Nyo [WOM]; *tYo(h) [EOM]	#yam [NC]; #(ya)x.u [a ~ o]
ONE	*kWia [EOM]	#pey; ?#(xi)Pu `first'
RABBIT	*ku[h/w]a [WOM]	#KulY [lY ~ nY] `jackrabbit'
RED	*kWah [EOM]	#(a-)xWa*(-T') `+blood'
ROAD	*LYau	#(a)nYa [S] `+path'

ROOT	*ma `+vine'	#i-ma*(y) `+sinew, hair'; #s(J)ama `+sinew'
to ROT	*tu [EOM]	#TuPi [NC] `rotten'
to RUN	*ka(u) `+to go, walk'	#KanY [N] or #KaL(Vw)
SALT: see SWEET		
to SAY	*coh; *(e)cauh	#CHan [N] [C <sup>h</sup> ~ C'] `to speechify'
SAND: see DUST		
SEED	*ce `+cacao'	#yic [SW] `+fruit'
SHAMAN	*(C)uHwa(H) `+witch(craft)'	#q'uH or #q'u7wa [N] `shamanry; poisoning song'
to SHOUT	*7ya [EOM]	#ya: `to speak, tell'; #yaN `throat'
to SHOUT	*x <sup>w</sup> a `to yell, say'	#pa [a ~ o] `to holler, say'; #wa(c) `+ to cry'; #kWa [S] [a ~ o] `to talk, say'
to SING	*sai `music'	#[u/i]sow [S] [s ~ s <sup>^</sup> ]
to SIT	*hkU	#K'uHLA [l ~ n] [NC] `+ to stay, dwell' (sg. subj.)

to SIT	*nA [WOM] `seated'	#(i)na [N] `+to live, be' (anim. du. subj.)
SKIN	*luwa `+leather' (*l- noun prefix?)	#(kW"-)aHwalY `+bark' (*kW"i- impers./indef. possessor)
SKY	*ma	#ema7(y) `+up, above'; #ma [N] `to climb, (go) up'
SMALL	*ci; *7(n)tYia	#c"i [S] `small'; #-c^'i [N] `dim. suff.'; #qiC'i `+little'
to STAY	*cau `+to leave'	#c^u [NC] `to sit, dwell, lie on ground'
to STINK	*me `skunk, onion'	#mes(J)e(w) `+to smell'
STONE	*ka `grindstone, metal'	?#[i/a-]xYa*7
STRONG	*kWa `+hard'	#p"a [NC]
SWEET	*sYe; *(se)hre(n)7 [EOM] `salt'; *ihru [WOM] `salt'	#(i)si +yV `salt' (cf. Dakota [mni* sku*ya] `salt', lit. `sweet water')
TOAD	*tu [symbolic/areal]	#got <sup>y</sup> o [N] `frog, tadpole' (also found outside Hokan)
TOOTH	*yau [EOM]	#[i*/a-]θ <sup>y</sup> a(w) [?a ~ o]

to TRADE	*ʔwai 'market, to pay'	#wiy [i ~ a] '+to buy, sell'
TURKEY	*tu	#Tu [S] [areal]
to TWIST	*kWe 'to weave, mat, loom'; *kWi [Amz-Mn] 'to wind, coil, wrap'; *Xwi [WOM] 'braid, hair'; *wa 'to weave'	#kW'inY 'to twist, twirl, twine'; #(i)pic'i 'to weave'; #wi(K) [NC] 'to weave'; #iwiL [w ~ m] 'to roll, twist'
to WASH	*k <sup>w</sup> e	#pa[q <sup>h</sup> /ç] [a ~ o]

### 7.3. Loose Ends in the OtoMangean-Hokan Comparison

Semantic matchings that did not yield plausible comparisons/cognates.

to AWAKEN, BAD, BASKET, BEANS, to BITE, BITTER, to BOIL, BUTTERFLY, CACTUS, to CUT, to DANCE, DARK, DEEP, DEER, to DIE, to DIG, DOG, DRY, DYE, to FEAR, FENCE, to FIGHT, to FILL, FIVE, FLAT, FLEA, FLY, FOREHEAD, FOX, HAIR, HEAD, HEAVY/FAT, GRINDSTONE, ICE/FROST, KNEE/ELBOW, to KNOW, to LEARN, LEFT, LIGHTNING/THUNDER, LIZARD, to LOOK, MAN, MEAT, MOON, MORE/MUCH, MOUNTAIN, MUSHROOM, NAME, NECK, NET, OLD, PERSON, PINE, to PLAY, to PULL, PUS/MUCUS, RAIN, to RAISE, RAT, RAW, to ROAST, ROUND, to RUB, to SAY, to SEE, to SEW, to SHAKE, SHIT, SHORT, SMOOTH, SNAKE, SPIT, SQUIRREL, STRING/FIBER, to TELL, THORN, to TOAST, TOBACCO, TONGUE, TREE, UNRIPE/GREEN, WARM/HOT, WATER, WHAT?, WHITE, WIND, WING: 88 items

Among the available reconstructed proto-OM and proto-Hokan forms there are about 250 near-identical glosses. There are at least 115 (given above) that have nearly identical glosses and show detailed phonological similarities. This is a rather high degree of similarity within the comparable material. There are detailed similarities in the grammatical markers used with both nouns and verbs. There is thus a prima facie basis for pursuing the question of a genetic relationship between OM and Hokan.



Since in OM many of the verb markers were clitic particles that were positioned by syntactic rules rather than morphological rules, their positions in pOM and pHok are not necessarily identical; but they show ordering properties that reflect VO (right-branching) vs. OV (left-branching) syntax, and one system can be envisioned as a systematic transformation of the other. Although the pHokan verb markers that appear to be suffixes correspond to proclitics in pOM, in the common ancestor they would have been clitic particles positioned by syntactic rather than morphological rules. The syntax of the common ancestor was more likely to have been OV-based than VO-based.

Non-genetic similarities? A few of the comparisons cited here may involve diffusion (TURKEY) or sound symbolism (CROW, FATHER, to HIT, MOTHER) or both (TOAD).

Brevity of compared forms. Roots in both OM and Hok are mostly monosyllabic. This is true of many Amerindian protolanguages, e.g. Athabaskan, Mayan, Tanoan, etc. Disyllabic stems in both OM and Hok are usually either compounds or contain a prefix. Unfortunately, pOM shows no syllable-closing consonants except \*n \*h \*ʔ. In Hokan, on the other hand, syllables may end in most of the consonants that they can begin with. pOM had tones (perhaps three of them, although this has not been worked out); it is plausible that an earlier elimination of syllable-final plosives and spirants may have resulted in the phonologization of tonal contrasts in pOM, and this is a hypothesis I intend to explore in the future.

Speculation: If OM and Hok are genetically related, what is the tree that describes the relationship? Inasmuch as the reconstruction of pHok is still very preliminary, and an internal subgrouping of Hokan cannot be made yet, is it essentially idle to contemplate this issue, but a few pertinent remarks can be made.

[1] Hokan has greater internal time depth (ca. 8000+ years) than OM (ca. 6500 yrs).

[2] pHok shows syllable-final consonants that may be the analogs of pOM tone contrasts. pHok has a few more phonological contrasts and phoneme combinations than pOM. Thus the common ancestor may have looked/sounded more like pHok than pOM.

[3] OM has VO syntax and Hokan has OV syntax, except in a few cases where VO seems to be innovative and due to diffusion. The word morphology of pHok seems to have been OV-based, and even the Hok languages with VO syntax have not reorganized their word morphology (to the extent that it goes back to pHok). VO syntax is an areal trait in Meso-America, and while universal and perhaps original in OM, is probably not original in the common ancestor of OM and Hokan.

[4] Given that the common ancestor may well have been typologically more like pHok than pOM, it does not follow that the known pOM structure can always be derived by rule from the known cognate pHok structure.

In Kaufman 1990.62 I stuck my neck out and suggested that OM belonged with Hokan in some version of Greenberg's Northern Amerind. This was in the context of the evidence presented above (which I had already for the most part collected) as well as a lack of support for Gr's Central Amerind, discussed below. Whether such an entity as Northern Amerind (with or without Almosan-Keresiouan: some Soviet long rangers would like to remove it [p.c. Sergej Nikolaev, Dmitrij Les^c^iner]) is justifiable remains to be seen.

The present version of my OM-Hokan comparison is still preliminary. Some of the steps in long range comparison that I outline in Kaufman 1990.24-25 have not yet been carried out. Among the kinds of further work that need to be pursued are the following:

1. Refine the Hokan reconstruction, replacing # with \*. Pore over and clean up the etymologies assembled so far.
2. Search for additional recurrent phonological correspondences that show more divergent reflexes. Reorganize both the OM and the Hok etymological collections so that they are more directly comparable, e.g. order by English gloss or according to a standardized semantic field listing.
3. Experiment with a reconstruction that would yield both pOM and pHok forms.

## 8. Greenberg's Central Amerind Compared to Kaufman's OtoMangean-Hokan

Greenberg's Central Amerind (CA) is hypothesized by him to contain OtoMangean, YutaNawa (UtoAztecan), and Tanoan. (See Greenberg 1987 for more detail).

As support for a genetic connexion between OM and Hokan I offer 25 grammatical comparisons and 90 lexical comparisons, for a total of 115 sets (as above).

Gr offers 90 'etymologies' (call them comparisons) unique to Central Amerind [CA], and 45 more CA comparisons that are listed in his Amerind 'etymologies', for a total of 135 putative cognate sets. This is the smallest number of proposed cognates for any of Gr's groups. The highest number is found in Penutian, which has 394 sets, followed by Chibchan-Paezan with 324 sets. Macro-Carib, with 143 sets, and macro-Panoan, with 141 sets, are nearly as low in supporting data as Central Amerind. But the low numbers for Gr's CA is surprising, since both YutaNawa (Uto-Aztecan) and OtoMangean are apparently well-studied.

In Gr's 90 exclusively CA comparisons OM is actually cited only 71 times. Of the 45 CA sub-entries in the 'Amerind dictionary', OM is cited only 38 times. The Amerind etymologies do not, of course, support the special subdivisions of Gr's Amerind, such as CA. These lower-level groupings can only be supported by the exclusive etymologies. In any case, Gr cites OM in only 109 comparisons. His phonology and semantics is often very loose. I cite 115 comparisons, of which 25 are grammatical. I do not intend to show that OM and Hokan form a subgroup within Amerind, since it is not my purpose here to support or deny a more inclusive genetic unit like Amerind. My purposes are simply to suggest that

- [a] there is prima facie evidence for a genetic relationship between OM and Hokan, and
- [b] the evidence linking OM and Hokan is better than Gr's evidence linking OM and UA + Tanoan.

It is better because it is more extensive, the phonology and semantics are more tightly controlled, more grammatical data is cited, and very importantly, the OM reconstructions are more adequate/accurate than those of Rensch.

I have shown above how Tlapanekan belongs to OM rather than to Hokan. I have also given some evidence suggesting a genetic connexion between OM and Hokan.

## 9. Further Work

Certain chores of a book-keeping nature have not been performed here, though a complete understanding of the literature on Tlapanekan and its genetic connexions require them:

1. Of the evidence Sapir (1925) offered for a Tlapanekan-Hokan affiliation, which data point to a Hokan-OM connexion?
2. Ditto for Oltrogge 1977.
3. Ditto for Greenberg 1987.
4. How good would Greenberg's case for "Central Amerind" be if adequate pOM reconstructions had been used in the comparison?

Questions 1-3 would require a comparison of this evidence with the equivalencies I have suggested between pOM and pHok. Question 4 would presumably simply identify pan-Americanisms (or "Amerind"-level etymologies in Greenberg's terms).

These tasks are in the present context not necessary, because I consider the genetic position of Tlapanekan to be established: it is most closely related to Mangan, thence to OtoPamean-Chinanteko, thence to Eastern OtoMangan; OM may be genetically related in a special way to Hokan. The four works mentioned above, however, do not compare OM and Hokan directly.

Among further tasks which might well be carried out are to answer the following questions:

[a] In which of my 115 comparisons do Tlapanekan forms show up in the OM etymologies?

[b] Which of my 115 comparisons overlap with comparisons made in Sapir 1925?

[c] Did Sapir's comparison of Sutiaba-Tlapaneko with Hokan involve a meaningful subset of the comparisons just presented?

[d] What would be the result of comparing my (adequate) OM reconstructions with adequate/accurate YutaNawa reconstructions?

## 10. Concluding Remarks

I repeat that when reconstructions can be made, that is, when language families rather than individual languages are being compared, reconstructions must be done before the comparisons can be evaluated. Greenberg [public remarks] points out that he tried to do this when he used Rensch's reconstructions for pOM, and that he was unlucky enough to choose the wrong reconstruction to use. This does not obviate the fact that Tlapaneko-Sutiaba was wrongly placed in Hokan, and I suggest that even with Rensch's faulty OM reconstructions, Greenberg would have found more comparables between OM and Hokan than he did between OM and YN, if he had made a serious effort. I could be wrong, though.

I would like to close with an observation about Greenberg's focus on classification rather than reconstruction. Greenberg claims that his classifications are 'valid'. This stretches the meaning of 'valid' as I understand it. It suggests that alternative classifications using his methods would not be as well supported as the ones he offers. It also suggests that the ones he offers should be accepted, and can reasonably be used in drawing further inferences. In my view, Greenberg's classifications are hypotheses which must be tested. Some may be supported by further study; many certainly will not (The present study provides an example of the second possibility). In Kaufman 1990 I have listed a large number of hypotheses concerning the classification of South American languages offered by other scholars that conflict with those of Greenberg. Neither Greenberg's hypotheses or these latter hypotheses deserve to be used in drawing further inferences, and this rules out, among other things, correlations with untested hypotheses in other fields of the study of man. Only the recovery of linguistic history through reconstruction and the devising of diversification models can be used to correlate with the assured results of archeology, physical anthropology, and climatology. Valid classification follows upon reconstruction; probabilistic classification, which is untested, precedes reconstruction.

## ABBREVIATIONS

### 1. Of language, group, and family names

Ach-Ats	Achumawi-Atsugewi
Ams	Amusgo
Ams-Mn	Amusgo-Mistekan
Ats	Atsugewi
Chm	Chimariko
Chn	Chinantekan
Cho	Chontal
Chp	Chiapaneko
Ch-Po	Chocho-Popoloka
Cht	Chatino
CM	Mangean (Chiapaneko-Mange)
EOM	Eastern OtoMangean
Esl	Eselen
Hok	Hokan
Isk	Iskateko
Kar	Karuk
Kom	Komekrudoan
Kot	Kotoname
Kwi	Kwikateko
Mas	Masateko
Mis	Misteko
Mn	Mistekan
Mng	Mange
Msw	Masawa
Mts-Okw	Matlatsinka-Okwilteko
OM	OtoMangean
OP	OtoPamean
Otm	Otomi*

Pah	Pahalat
Pam	Pame
Pn	Popolokan
Pn-Sp	Popolokan-Sapotekan
Pom	Pomo[an]
Sal	Salina
Sap	Sapoteko
Ser	Seri
Sha-Kon	Shasta-Konomihu
Sn	Sapotekan
Sut	Sutiaba
Tln	Tlapanekan
Tlp	Tlapaneko
TM	Tlapanekan-Mangean
Tri	Triki
WOM	Western OtoMangean
Wsh	Washu
Yan	Yana[n]
YN	YutaNawa[n]
Yum	Yuman

## 2. Of grammatical categories

absol[utive  
ADJective  
ADVerbial  
ANDATive  
art[icle  
AUXiliary  
CAUSative  
CLASSifier  
DEFinite  
DERIVation  
desid[erative  
EMPHatic  
ex(x)                example(s)  
FUTure, fut[ure  
HABitual  
IMMediate  
IMPERsonal  
IMperative  
INDEFinite  
INSTRument  
NEGative  
NOMinalizer, NOMinalizer  
NP                    noun phrase  
opt[ative  
PASSive  
PCP                    participle  
PERFect  
PLUPERFect  
PLural, pl[ural



pol[ite  
 POSSESSive state  
 POTential  
 PRESEnt  
 PRETErite  
 PROGressive  
 rel[ativizer  
 SinGular, s[in]g[ular  
 STATive  
 SUBJect  
 SUBORDinator  
 TAM                    tense-aspect-mood  
 TEMPoral  
 Tense  
 VP                    verb phrase

CONVENTIONS

X+, +X	proclitic, enclitic
X-, -X	morphosyntactic (inflectional) affix
X>, >X	morphosyntactic class-shifting affix
X., .X	lexical (derivational) affix
X=, =X	prebound, postbound

## NOTES

1. **Subtiaba** is a place-name in Nicaragua that is pronounced locally /sutia\*ba/. This comes from Nawa /xo:te=a:-pan/ 'periwinkle river on'. Variants of this name are **Jutiapa** in both Honduras and Guatemala, and **Soteapan** in Veracruz, Mexico. The first <b> in **Subtiaba** was never pronounced.
2. The spellings I use for American Indian language names differ somewhat in individual cases from those commonly in use. /k g w s h/ are always so written, no matter how they are spelled in Spanish or English. /c^/ is <ch>, /c/ is <ts>, and /s^/ is <sh>. Only five vowel qualities are distinguished. See Kaufman 1994.
3. **Coahuilteco** is a barbarism in Spanish. The gentilic adjective from Coahuila is **Coahuilense** in Spanish. **Coahuila** does not have a Nawa origin. Pahalat was spoken primarily in southern Texas north of the Rio Grande river, and NOT in Coahuila. Pahalat was spoken by many different ethnic groups. Pahalat is practically the only name that is repeated in more than one source.
4. **Salina** is a single language. It is not appropriate for linguists to use the suffix -an on names of single languages.
5. **Chontal** is ambiguous, since it is derived from a Nawa word meaning 'foreigner', and is used for more than one language: however, the term **Tequistlateco** is a barbarism. It cannot be formed from Tequisistla\*n, the place-name of which it is supposed to be the gentilic adjective.
6. The term **Jicaque** is undesirable, since in regional Spanish it means 'savage Indian'. The native name is **Tol**, and this term is now widely used in Honduras, where Tol is spoken.
7. Sections 2 and 3 are from Kaufman 1988ms.
8. Section 4 is based on Kaufman 1983ms.

8'. Since reconstructions for OM branch reconstructions are preceded by a single asterisk, reconstructions to the proto-OtoMangean level are preceded by a double asterisk.

9. Section 5 is based on Kaufman 1989.

10. Section 6 is based on Kaufman 1989msb.

## References

- Bartholomew, Doris 1965ms. The Reconstruction of Otopamean (Mexico). PhD Dissertation in Linguistics, University of Chicago. viii + 445pp.
- Belmar, Francisco 1905. Lenguas de Me\*xico: Familia Mixteco-Zapoteca y sus Relaciones con el Otomi\*. Familia Zoque-Mixe, Chontal, Huave y Mexicano. Mexico.
- Brinton, Daniel G. 1901. The American Race. New York.
- Ferna\*ndez de Miranda, Mari\*a Teresa & Roberto J. Weitlaner 1961. 'Sobre algunas relaciones de la familia mangué'. Anthropological Linguistics 3.1. [whole number, 99pp.]
- Greenberg, Joseph H. 1987. Language in the Americas. Stanford: Stanford University Press. xvi + 438pp.
- Gudschinsky, Sarah C. 1959. Proto-Pototecan: A Comparative Study of Popolocan and Mixtecan. Indiana University Publications in Anthropology and Linguistics no. 5. Baltimore: Waverly Press. viii + 118pp.
- Jacobsen, William H., Jr. 1959. 'Washo and Karok: an approach to comparative Hokan'. IJAL 24.195-212.
- 1976. 'Observations on the Yana stop series in relationship to problems of comparative Hokan phonology'. Hokan Studies 203-236.
- Jime\*nez Moreno, Wigberto 1937. Mapa Lingüi\*stico de Norte y Centroamerica
- 1962. 'Estudios Mixtecos' pp54-88, in Vocabulario en Lengua Mixteca por Fray Francisco de Alvarado. INI and INAH.

Kaufman, Terrence 1983ms. 'New perspectives on comparative Otomanguean phonology'.

----- 1988ms. OtoManguean Tense/Aspect/Mood, Valence-changing, and Nominalization Markers

----- 1989. 'A research program for reconstructing proto-Hokan: first gropings'. pp 50-168 in Papers from the 1988 Hokan-Penutian Languages Workshop, ed. Scott DeLancey.

----- 1989msa. 'Early OtoMangue homelands and cultures: some premature hypotheses'. Read at Pacific Linguistics Conference, Eugene. Published (1990) in University of Pittsburgh Working Papers in Linguistics 1.91-136.

----- 1989msb. 'Some hypotheses regarding proto-Hokan grammar'. Read at Hokan-Penutian Workshop, Tucson, July 1989.

----- 1989msc. 'Long-range and mid-range comparison; the cases of Hokan and OtoMangue'. Read at SSILA summer meeting, Tucson, July 1989.

----- 1990. 'Language history in South America: what we know and how to know more', pp13-73 in Amazonian Linguistics: Studies in Lowland South American Languages, ed. Doris Payne. Austin: University of Texas Press.

----- 1994. 'The native languages of Latin America', pp31-89 in Atlas of the World's Languages, ed. Christopher Moseley & R. E. Asher.

Lehmann, Walther 1920. Zentral-Amerika, 2 vols. Berlin.

----- 1920. 'Die Subtiaba-Sprache', pp932-969 in Zentral-Amerika.

Leo\*n, Nicola\*s 1902. 'Familias Lingüi\*sticas de Me\*xico", Anales del Museo Nacional, orig. series, 7.279-335.

- Longacre, Robert E. 1957. Proto-Mixtecan. Bloomington: Indiana Univeristy Research Center in Anthropology, Folklore, and Linguistics, no. 5. vii + 195pp.
- Mason, J. Alden 1940. 'The Native languages of Middle America', pp 52-87 in: The Maya and their Neighbors. New York.
- Mechling, W. H. 1912. 'The Indian Linguistic Stocks of Oaxaca, Mexico', AA n.s. 14.643-682.
- Mendizabal, Miguel O. de and Jimenez Moreno, Wigberto 1937. Distribucio\*n Prehispa\*nica de las Lenguas Indi\*genas de Me\*xico.
- 1941. revision of the above.
- Moshinsky, Julius 1976. 'Historical Pomo phonology'. Hokan Studies 55-76.
- Newman, Stanley & Robert Weitlaner 1950a. 'Central Otomian I: Proto-Otomi reconstuctions'. IJAL 16.1-9.
- 1950b. 'Central Otomian II: Primitive Central Otomian reconstuctions'. IJAL 16.73-81.
- Oltrogge, David F. 1977. 'Proto Jicague-Subtiaba-Tequistlateco: A comparative reconstruction', pp 1-52 in Two Studies in Middle American Comparative Linguistics. Arlington: SIL/University of Texas at Arlington.
- Orozco y Berra, Manuel 1864. Geografi\*a de las Lenguas y Carta Etnogra\*fica de Me\*xico.
- Pimentel, Francisco 1875. Cuadro Descriptivo y Comparativo de las Lenguas Indi\*genas de Me\*xico.

- Radin, Paul 1933. 'Notes on the Tlappanecan Language of Guerrero', IJAL 8.45-72.
- 1944. 'The Classification of the Languages of Mexico', Tlalocan 1.3.259-265.
- Rensch, Calvin R. 1963ms. Proto-Chinantec Phonology. MA Thesis in Linguistics, University of Pennsylvania. 234pp.
- 1966 (pub. 1976). Comparative Otomanguean Phonology. PhD Dissertation in Linguistics, University of Pennsylvania. [published 1976. Bloomington: Indiana University Language Science Monographs 14. 350pp.]
- 1977. 'Classification of the Otomanguean languages and the position of Tlapanec', pp 53-100 in Two Studies in Middle American Comparative Linguistics. Arlington: SIL/University of Texas at Arlington.
- Rivet, Paul and Marcel Cohen 1924. Les Langues du Monde. Paris.
- Sapir, Edward 1925. 'The Hokan affinity of Subtiaba in Nicaragua', AA 27.402-435, 491-527.
- 1929. 'Central and North American Languages', Encyclopaedia Britannica, 14th ed. 5.138-141. London & New York.
- Schmidt, Wilhelm 1926. Die Sprachfamilien und Sprachenkreise der Erde. Heidelberg.
- Schultze-Jena, Leonhard 1938. Indiana III. Bei den Azteken, Mixteken und Tlapaneken der Sierra Madre del Sur von Mexiko, pp111-373. Jena.
- Silver, Shirley 1976. 'Comparative Hokan and the Northern Hokan languages'. Hokan Studies 193-202.
- Stresser-Pe\*an, Guy 1952. 'Famille Otomang', pp1078-1084 of Les Langues du Monde.

- Sua\*rez, Jorge 1983. La Lengua Tlapaneca de Malinaltepec. UNAM. 641pp.
- 1986. 'Elementos gramaticales otomangues en tlapaneco', pp 267-284 in Language in Global Perspective, ed. Benjamin F. Elson. Dallas: Summer Institute of Linguistics.
- Swadesh, Morris 1959. 'Mapas de Clasificacio\*n Lingüi\*stica de Me\*xico y las Ame\*ricas', UNAM, Mexico DF.
- 1960. 'The Oto-Manguen hypothesis and Macro-Mixtecan', IJAL 26.79-111.
- 1964a. 'Interim notes on Oaxacan phonology'. SWJA 20.168-189.
- 1964b. 'Algunos problemas de la lingüi\*stica otomangue'. Anales de Antropologi\*a 1.91-123.
- Thomas, Cyrus and J. R. Swanton 1911. Indian languages of Mexico and Central America, BAE-B 44. Washington DC.
- Weathers, Mark 1976. 'Tlapanec 1975', IJAL 42.367-371.
- Weitlaner, Roberto J. 1941. 'Los Pueblos no Nahuas de la Historia Tolteca y el Grupo Linguistico Macro-Otomangue', RMEA 5.2-3.249-269.
- Weitlaner, Roberto J. 1945. 'Paul Radin's Classification of the Languages of Mexico', Tlalocan 2.1.65-70.



xxx

AA = American Anthropologist, new series

Hokan Studies. edited by Margaret Langdon and Shirley Silver. 1976. The Hague: Mouton. 373pp.

IJAL = International Journal of American Linguistics

RMEA = Revista Mexicana de Estudios Antropológicos

SWJA = Southwestern Journal of Anthropology