<table>
<thead>
<tr>
<th>Title</th>
<th>Author</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Lakhota Definite Articles and Topic Marking</td>
<td>Traci S. Curl</td>
<td>1</td>
</tr>
<tr>
<td>Switch-Reference and the Identification of Small Pro in Hidatsa</td>
<td>John P. Boyé</td>
<td>17</td>
</tr>
<tr>
<td>Cherokee Cities: The Word Boundary Problem</td>
<td>Marcia Haag</td>
<td>33</td>
</tr>
<tr>
<td>Paragraph-Level Switch-Reference Markers in Chickasaw Conversation</td>
<td>Cynthia Walker</td>
<td>45</td>
</tr>
<tr>
<td>Ojitlán Chinantec Phonology and Morphology</td>
<td>Monica Macaulay</td>
<td>71</td>
</tr>
<tr>
<td>Glottalization in Nuu-chah-nulth in Optimality Theory</td>
<td>Eun-Sook Kim</td>
<td>85</td>
</tr>
<tr>
<td>On the Relationship Between Mixe-Zoquean and Uto-Aztecan</td>
<td>Søren Wichmann</td>
<td>101</td>
</tr>
</tbody>
</table>
ON THE RELATIONSHIP BETWEEN
MIXE-ZOQUEAN AND UTO-AZTECAN

Søren Wichmann
University of Copenhagen

Abstract: The paper seeks to establish evidence for the relationship between Uto-
Aztecan and Mixe-Zoquean, a relationship which was proposed by Whorf (1935), but
never before substantiated by actual comparisons. In order to reduce the factor of chance
semblances proto-forms are compared. Given the several thousand years of non-
proximity of the two language families, borrowing cannot be used used as an
explanation for the similarities in vocabulary. In several instances the comparisons
reveal regular correspondences involving non-identical segments, correspondences of a
sort that is expected in true genetic relationships.

0. Introduction

This article will differ from earlier attempts to find external relatives to the MixeZoquean (MZ)
languages both methodologically and materially. Previous attempts (see Wichmann 1994: 238-
243 for a list and brief discussion) often do not cite any evidence at all. If they do, regularity of
correspondences is not taken very seriously, the number of cognates given is small, and single
languages rather than protolanguages are compared. In the present work, however, it is taken for
granted that all segments should correspond, therefore any deviation from the expected pattern
will be noticed. To alert the critical awareness of the reader, cognates will be cited as I go along
and not collected in an appendix. Moreover, for the first time proto-languages are compared. The
results are at variance with McQuown (1942), Swadesh (e.g. 1956), Greenberg (1987), Witkowski
and Brown (1978), Brown and Witkowski (1979), and others who have thought
MixeZoquean to be closest related to Totonacan and Mayan. I side more closely with Whorf
(1935), who included UtoAztecan (UA) in his “Macro-Penutian” stock along with (almost
certainly in the extended sense of Sapir 1929, see summary of personal communication with
Whorf in Mason 1962 [1940]), Totonac, Kiowa, and Mayan. As is well known, Whorf never
published any supporting evidence. Others who did try to supply evidence for the Penutian relation-
ship of MixeZoquean unfortunately did not include UtoAztecan in their comparisons. These
are the works of Freeland (1931) and Hymes, (1964). The relationship of MixeZoquean to UtoAz-
tecan is a hypothesis which is hard to ignore, as the reader shall soon see. It seems to be a mere
coincidence that there has been no scholar familiar with both language families who has been
sufficiently attracted to the hypothesis so as to check it in detail. Thus the relationship has been
largely neglected up to this point.

UtoAztecan proto-forms are from Kaufman (1981) (abbreviated TK) when not otherwise
noted. This author builds on previous studies, notably by Voegelin, Voegelin and Hale (1962),
Miller (1967), lames (1972), Campbell and Langacker (1978), Bright and Hill (1967), and

Kansas Working Papers in Linguistics 24:2, pp. 101-113
Bacon (1965). These works will hereafter be referenced by the abbreviations VVH, M1, DL, C/1, B/H, BAS. Later important contributions are Miller (1987) and Lionnet (1985) (abbrevi-
ated M2 and AL). My own expertise is in MZ and I am therefore not in the position to judge the
validity of Kaufman's reconstructions. But they are expressedly said to be conservative and are
well argued for and embrace both length and final features unlike all previous studies but Whorf
(1935). TK has some 473 UA reconstructions plus some additional reconstructions at lower
nodes in the family tree. About 10% of these reconstructions enter into comparisons with MZ
forms in the present article.

Occasionally I cite from VVH, M1, or the sources of M2. This will be because the
reconstructed form is not included in TK (which does not purport to be an exhaustive lexical
collection), not because the shape of their proto-form happen to better suit the MZ form. It
should be kept in mind that Miller never offers true reconstructions, but only cognate sets
indexed by some important, recurring sounds. All MZ lexical reconstructions are from
Wichmann (1995), which includes a comparative dictionary. "MZ" abbreviates the MZ proto-
language, "M" the immediate ancestor of Mixean (one of the two main branches), "Z" the
ancestor of Zoquean (the other main branch), and "OM" is the parent of Oaxacan Mixean, a group
of languages within the Mixean branch. The reconstructions of grammatical morphemes are from

The cognates are cited in an order that focuses first on reflections of the vowel system, then
on changes involving consonants. The symbol "*" and "#" represent a glottal stop and a velar
nasal, respectively. Correspondences signalled by V, C, and () are not accounted for. Other things
that look like they are not accounted for will be explained by sound laws to be revealed as I go
along.

1. Correspondences Involving Changes

Correspondences signalled by V, C, and () are not accounted for.

UA *a(?): MZ *a(?)

1) **/ahyV 'good' > UA *ahy MZ *?oya
2) **/at:CV(h) 'house' > UA *attif MZ *?oci
3) **/ikas(a) 'not warm' > UA *ikia 'shade, be cool' MZ *ikos 'epid'

UA *e(?): MZ *e(?)

1) **/epaV 'sister' > UA *epe MZ *epa
2) **/a:CV(h) 'face' > UA *a:ti MZ *a:ci
3) **/khiV 'not big' > UA *ekhi MZ *ekhi

UA *e(?): MZ *e(?)

1) **/eka 'he' > UA *ek MZ *ek
2) **/epeV 'strength' > UA *epe MZ *epe
3) **/a:CV(h) 'face' > UA *a:ti MZ *a:ci
4) **/sa 'room' > UA *sa MZ *sa

(4) **ko?/V 'die' > UA *ko?/i/o 'to die, kill'  
Z *ka? 'to die'  
Cf. also UA *ko?:ya 'to kill', MZ *ya:k-a 'to kill'  
(5) **ma(?)/C 'grind' > UA *mataa 'quern' (final a must be an old suffix)  
MZ *mo?c 'to grind'  
(6) **ma(?)sah 'deer' > UA *maasu:oh  
Z *m?ah  
(7) **na(?)/V 'burn' > UA *na'ui 'to burn, kindle'  
Z *no? 'to light, set fire to'  
(8) **pa:k(a) 'reed' > UA *pa:kkaa  
MZ *?o:ki(wi)?/r  
(9) **to(?)ka 'old woman' > SUA *?ooka  
MZ *?oko 'grandmother'  
(10) **naCV 'put' > Tak *tuvV 'put'  
OM *to?o 'basket woven of palm'  
(11) **yi(?)/k 'heat; desire' > UA *ya'a 'to desire'  
MZ *yo:k 'to get warm'  
UA initial kWi-; MZ ko-  
(12) **kWit 'fall' > UA *kWi:n 'shut'  
(final /a/ a suffix?)  
MZ *kot 'fall'  
UA initial 9-; MZ w-  
(13) **Wan? 'cry' > UA *9ao 'to cry'  
MZ *wan? 'to sing; want'  
(14) **Woy 'roll' > UA *9o: 'to bend back'  
MZ *wai 'to roll'  
(15) **Wa(?)-ci 'little root' > UA *9a:/i 'root'  
Z *wai 'root'  
(contains **-ci dim)
MZ initial p- : UA zero

(16) **pa(b)-ci *older brother* > UA *pa-ci (*ci is dim.)
    MZ *fa-ci

(17) **po(k-i) *rolled* > UA *po(k-i) *pavel*
    MZ *fe: -i *folded or rolled (as paper or cigar)

(18) **pwi(i) *sleep* > Numic *piw* *to sleep; sleep*

(19) **pu(3)ku *dog* > UA *punku *dog*
    MZ *vku *dog* or *dog*

(20) **paC(a) *leaf* > Cyp *pala *leaf*
    MZ *ay *leaf"

MZ initial c- : UA s-

(21) **cik *cut* > UA *sik *cut* (M)
    MZ *cik *husk, harvest, peel"

(22) **cok(o) *wet ground* > UA *sok, *cok *ground, earth* (M)
    Z *koko *soaked"

(23) **cum(a) *tie* > UA *suma (M)
    MZ *cum"

L/A modal s- : MZ, ? between like vowels, y between low-high configurations, Ø elsewhere

(24) **kuhsi *wood* > UA *kuhsi *wood*
    MZ *kyu *tree"

(25) **kWisi *grab* > LA *kWis *to take, catch*
    MZ *ki? *hand, arm"

(26) **nasi(t) *product of burning* > UA *nasi *ashes*
    Z *nay *wax"

Set cited already: (6) 'deer.'
UA medial -hÇ- : MZ -Ç-

(27) **nVhy`V 'to name'  >  Num *ni̯iyá 'to name, call'
MZ *ni̯iy 'name'

Two additional sets have been cited earlier which bear evidence to this correspondence: 'good'
(1), 'wood' (24). A third follows: 'grandfather' (43).

MZ final k : UA zero

(28) **nukWak 'smell'  >  UA *nuf vá 'smelly'
Z *tukuk 'smell bad'

(29) **ko-pak 'head'
(pertaining to head + bone)  >  UA *kopa 'forehead'
MZ *ko-pak 'head'

(30) **ko-ta(ç)k 'neck'  >  UA *kutaa 'neck'
MZ *ko-: pertaining to head
MZ *tak(as) 'walking stick'

(31) **maç:inVk 'child'
 >  UA *ma:nà 'female child'
MZ *mank 'son-daughter'
(Cf. also for UA Cuaño *man-da 'diminutive' and for
MZ Textespec Popoluca *man-da? 'plural diminutive')

(32) **to:koç(ç)k 'protuberance in leg area'  >  UA *to:ko: 'knee'
Z *to:koi 'heel'
(metathesis unaccounted for)

Other example, already cited: (11) 'beast; desire'.

MZ final ĕ : UA zero

(33) **weč 'to comb'  >  UA *we:n, *wen (M wi-09)
MZ *weč 'comb, drib'

MZ final s : UA zero

(34) **ki:si 'bite'  >  UA *ki:si 'to bite'
MZ *ki:s 'pull; bite'

(35) **hVkos 'not warm'  >  UA *hiká 'shade, be cool'
MZ *hokos 'tepid'
MZ final y : UA zero

(36)  "makat("y) 'give' > UA "makaa 'to give, hit'  
      MZ "ma?ay 'to sell'

(37)  "t(-)kay 'put' > UA "tika 'to put, lay down'  
      MZ "tiky 'to enter'  
      (Vowel-leveling)

(38)  "win-cay 'umbilical cord' (?),  
       or 'umbilical cord' (?)) > UA "win(CV) 'string'  
       MZ "win 'eye, person', *cay 'rope'

Other examples are 'to roll' (14) and (MZ morpheme-final y : UA zero) 'reciprocally' (82).

MZ contractions

(39)  "niya-kWîsV 'breathe' > UA "niya-kWîsV 'suffocate'  
       MZ "ni:s 'suffocate'

(40)  "to(a) wi 'burn' > UA "to(a)wi 'to be hot'  
       MZ "toy 'to burn, hurt'

The set 'give', cited earlier (36), is an unparalleled example of **k > ** in MZ between vowels.

2. Correspondences of Identity

UA *i : MZ *i

(41)  "wil(a) 'whistle' > UA "wilu (M-457)a  
       M *wil-wil

UA *i : MZ *i

(42)  "ti(n)wîV 'namesake' > UA "tiwa 'name'  
       Z *tiwi 'relative'

UA *a(:) : MZ *a

(43)  "aapV 'GrFa' > UA "aapi 'grandfather'  
       MZ "apu 'grandfather'

(44)  "a(a)wV 'mouth' > UA "aawV 'to tell' (V probably a suffix)  
       MZ "aw 'mouth'
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>**mV:ca(?) 'moon'</td>
<td>&gt;</td>
<td>UA *mi:ca 'moon'</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>**ya(h) 'die'</td>
<td>&gt;</td>
<td>UA *ya 'die' (M-132)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>**su(?)n- 'central inner organ, seat of emotions'</td>
<td>&gt;</td>
<td>UA *su:na 'heart, middle'</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>**jumu</td>
<td>&gt;</td>
<td>Cup *jumu- 'put on hat'</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>**soho 'tree sp.'</td>
<td>&gt;</td>
<td>UA *soho 'cottonwood'</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>**to?k 'spread, weave'</td>
<td>&gt;</td>
<td>UA *to?ka 'spider'</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>**tsohV(?)m- 'cough'</td>
<td>&gt;</td>
<td>Numic *tsoh(?)m- 'to cough'</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>**to(?)CV 'stick'</td>
<td>&gt;</td>
<td>Sonoran *uto- 'to get stuck, stick'</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>**tsoC 'swell'</td>
<td>&gt;</td>
<td>UA *to 'stomach' (M)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>**woh? 'to bark'</td>
<td>&gt;</td>
<td>Numic (?) *wohi, etc. 'bark, yell, howl' (IAN-274)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Additions from Miller (1987). Reconstructed forms need to be checked and projected phonologically back to UA.
(56) *?avi 'tasty' (BAS-315)
    Z 'som' tasty

(57) *'oimirai 'to walk around' (BAS 318)
    MZ 'aoy' go (and have returned)

(58) UA *bi 'yes' (MIL-481)
    MZ 'bi: 'yes'

(59) **sv(·)w-V 'to day'
    > *sivi 'now, today' (BAS-194)
    MZ 'sivi: day, sun'

(60) **cu(·) 'night'
    > *cu, *co 'disappear' (IAN-258)
    MZ 'cu: 'night'

(61) **ak(V)
    > UA *aki 'river' (MIL-348)
    MZ 'aka 'riverbank'

(62) **hot(a)
    > *hota 'to dig' (IAN-34)
    M 'hot' to dig a hole'

(63) **ho(·)n...
    > *ho(·)limi 'bat'
    MZ 'ho:n 'bied'

(64) **hū...
    > *hū-l 'pull out' (B-H)
    MZ 'hū: 'take out'

(65) **kaka
    > UA *kaka 'sweet' (MIL-427)
    MZ 'kakawa'

(66) **kap...
    > oo-kawadk, ya-kapó 'plano'
    M *kap-kap empty'

(67) **mak '10'
    > *makoi '10' (LJC-135)
    MZ 'mak:Y '10'

(68) **ko(·)m...
    > *komV 'pitcher, jug, pot' (C/L-127)
    MZ 'kom: pot in'

(69) **kum...
    > *kuna, *kuna 'husband' (MIL 504a, 504b)
    Z 'komi: master'

(70) **kWana
    > *kWana 'smelly' (IAN-78)
Sierra Popoluca kin 'smell'

(71) **pu(n)C() > UA punci 'eye' (TK)
    (eye, seed associated in IAN)
    Z *puh 'seed'

(72) **sapa 'fruit meat' > *sapa 'meat' (LJO-232)
    Z *sapa 'plantain'

(73) **sam > *sami 'adohe' (LJO-230) (contains suffix -i 'resultative')
    MZ *sam 'heat'

(74) **so:jn.. > sh. so:n 'many', hp. so: 'all, many' (M-so-14)
    Z *sonc 'much, many'

(75) **suh.. > *sa(w)ah 'breathe' (IAN-187)
    MZ *su'h 'blow'

(76) **suy.. > B/H *suji- 'sting'
    MZ *suy 'sweat, fish with hook'

(77) **tak.. > *taka 'fruit' (LJO-269)
    Z *tak 'cacao'

(78) **pak > UA *pak 'shirt' (MIL-377)
    Sayula Popoluca pak 'rake (cloth)' 

(79) **pa 'wild' > UA *pa 'mountain sheep' (MIL-369)
    MZ *pa- prefix: 'wild'

3. Grammatical morphemes

(80) **-i 'deverbalizer'
    (resultative) > UA *-i derives nouns
    meaning 'result of action'
    from verbs
    MZ *-i/e derives nouns and adjectives 'meaning product
    of action' from verbs

The existence of this suffix explains the following set in which the UA cognate is the derived
form, and the MZ cognate the basic one:

(81) **hul()j 'tie' > UA *huuk 'bunchgrass'
    M *huk 'to tie together (cargo, etc)'
Other grammatical morphemes are:

(82) **na(-y)- 'reciprocal' > UA *na- 'twice, double; reciprocal'
MZ *nay- 'reciprocal'

(83) **hu 'demonstr.' > UA *hu 'that'
MZ *hu 'question marker'

(84) **pi(?) 'distrib.' > UA *pis 'mediopassive, distributive particle'
(M)Z *pi? 'distributive'

(85) **pi(?) 'participle' > UA *pis 'mediopassive, distributive particle'
MZ *pi? 'human participial or deverbal'

(86) **-m(-) 'plural' > UA *-m
Z *-m (requires that *-tam 'plural' be segmentable)

(87) **-ci 'diminutive'

See the sets 'little roof' (15), 'elder brother' (16)

(88) **ka(-b) 'negation' > UA *ka 'negative'
MZ *kach 'no'

4. Conclusion

A possible reaction to evidence such as that presented above is to focus the attention on forms that are widespread (and thus possible borrowings), onomatopoeic/symbolic forms, and non-evident semantic equations. Along with monosyllables, lexical correspondence types that have possible non-genetic explanations such as those just mentioned are often considered weak as evidence for establishing genetic relationships. Although I agree with this, I would like to stress that the method of weakening an argument about a genetic relationship by weeding out such possible non-genetic forms is much more warranted in the gather-as-many-look-alikes-as-possible approach than in the present approach. When proto-languages are compared and attention is paid to strict correspondences, actual falsification in fact becomes possible: the proper refutation method here is to point out cases where sound laws are not adhered to. After all, also proto-languages may have symbolic forms and monosyllables. Finally, in this particular instance apparent correspondences cannot be explained as loan words, since the two language families appear to have non-adjacent homeland and have generally continued to be spoken in different areas. Thus, it is has been proposed that Proto-Aztecan was spoken in Arizona and northern Mexico, possibly extending into Southern California (Fowler 1983), whereas the home-land of the Mixe-Zoqueans is generally assumed (Campbell and Kaufman 1976, Wichmann 1999) to be
close to the Mexican Gulf Coast.

Many things need to be refined. I have sacrificed one half of my cognate collection for the sake of making things work phonologically. I hope, however, that approximately the same amount may be regained on the basis of strict criteria. These new sets may include perhaps twice as many grammatical morphemes as we already have. Though things look promising, the sort of evidence that we really need is, for instance, to be able to explain in detail how MZ increased its vowel system with an e and how vowel length works. We should look more at verb roots and grammar. Success at explanations in these areas would have me finally convinced of a genetic relationship between UA and MZ. I do think, however, that the preceding pages provide a promising start.

NOTES

1. The present work is a slightly revised version of a paper originally prepared for circulation among participants at the 13. International Congress of Anthropological and Social Sciences, Mexico City, July 29-August 5, 1993. I gratefully acknowledge critical comments from Lyle Campbell.

REFERENCES


