Ilia Peiros
Santa Fe Institute

Uto-Aztecan comparative linguistics and etymological databases

The paper presents a general survey of Uto-Aztecan historical studies. It discusses the genetic classification of the family, phonological reconstructions of individual daughter languages and of Proto-Uto-Aztecan itself. Lexical material used in the paper is taken from electronic databases created for the "Evolution of Human Languages" Project (Santa Fe Institute).

Keywords: Uto-Aztecan languages, lexicostatistics, comparative linguistics, etymological databases.

The project "Evolution of Human Languages" was launched in 2002 by Murray Gell-Mann and Sergei Starostin. One of its main aims is to build a network of etymological databases for all language families of the world, in which, for each known family, we plan to have:
— a system of phonological correspondences connecting many, if not all, languages within this family;
— an extensive etymological corpus supporting these correspondences;
— a lexicostatistical classification based on reliable etymologies.

After this type of work was performed for some better known families (results are available at starling.rinet.ru), it became clear that in many cases, instead of a full-scale study of every language family, we can only conduct a preliminary survey based on existing material.

The current paper briefly presents the results of such a survey on Uto-Aztecan (UA), one the better known language families of Northern America. Due to the nature of the paper, no detailed discussion of literature and alternative views is included.

The family consists of about 60 languages. Its generally assumed classification is given in [Campbell 1997: 134]:

Uto-Aztecan
    Northern Uto-Aztecan
      Numic
          Western
              Northern Paiute (Oregon, Idaho, Nevada)
              Monache (California) [obsolescent]
          Central
              Shoshone, Panamint [obsolescent], Comanche [obsolescent] (Nevada, Utah, Wyoming, Oklahoma)
          Southern
              Southern Paiute (Utah, Nevada, California, Arizona)
              Ute, Chemehuevi [obsolescent] (Utah, Colorado, California, Arizona)
              Kawaiisu [obsolescent] (California)
      Tübatulabal [moribund] (California)

Takic

    Serran: Serrano [moribund], Kitanemuk [extinct] (California)

    Cahuilla [moribund?], Cupeño [moribund] (California)

    Luiseño-Juaneño [obsolescent] (California)

    Gabrielino [extinct], Fernandeño [extinct] (California)

Hopi (Arizona)

Southern Uto-Aztecan

Pimic (Tepiman)

    Pima-Papago (Upper Piman) (Arizona, Sonora)

    Pima Bajo (Lower Piman, Nevome) (Sonora)

    Northern Tepehuan, Southern Tepehuan (Sonora, Durango, Jalisco)

    Tepecano [extinct] (Jalisco)

Taracahitic

    Tarahumaran

        Tarahumara (Chihuahua)

        Guarijio (Chihuahua, Sonora)

    Tubar [extinct?] (Chihuahua)

    Cahitan (Yaqui-Mayo-Cahita) (Arizona, Sonora)

    Opatan

        Opata [extinct] (Sonora)

        Eudeve [extinct] (Sonora)

Corachol-Aztecan

    Cora-Huichol

        Cora (Nayarit)

        Huichol (Nayarit, Jalisco)

    Nahuan (Aztecan)

        Pochutec [extinct] (Oaxaca)

        Core Nahua

            Pipil (El Salvador, extinct in Guatemala and Nicaragua)

            Nahuatl: many dialects (Mexico).

"[Languages] known to have fewer than 10 speakers are specified as 'moribund'; languages with more than 10 but fewer than 100 speakers are labeled 'obsolescent'" (p. 133).

An additional list of extinct languages which may have belonged to UA is given on pp. 133–134 of Campbell's book as well as in other sources, like.

In a somewhat impressionistic way, typical of the rest of the book as well, Campbell does not justify his classification. The only modern UA classification that is based on complete, explicit presentation of the evidence is W. R. Miller's paper from 1984.[1] Using lexicostatistics, Miller (p. 13) comes to the following conclusion: "thus there seem to be five groups that are more or less coordinate with each other: Numic, Tübatulabal, Takic, Hopi and SUA" (South Uto-Aztecan); this is supported by the matrix of 'cognate density'. Further investigation of this matrix (see the table on p. 114 below) suggests, however, that lexicostatistics allows to generate a different classification.

Let us reduce Miller's matrix to a compressed version, where each transparently obvious group is represented by the lowest percentage that one of its members shares with other groups. This version looks as follows:

---

[1] Hale's earlier lexicostatistical classification [Hale 1958–59] also presents the data.

	*Numic	*Takic	Hopi	*Sonoran	*Aztecan
*Numic	×	22	22	23	14
*Takic	22	×	26	24	18
Hopi	22	26	×	26	24
*Sonoran	23	24	26	×	29
*Aztecan	14	18	24	29	×

It can be seen that *Numic and *Aztecan form two ends of the spectrum, while the other groups are somewhere in between. Normally, *Sonoran should be grouped with *Aztecan, but it shows much higher figures with *Numic. Such discrepancies are a clear indication that we are dealing with the effects of undetected borrowings within related language groups.

Restricting ourselves to reconstructed intermediate 50-item proto-lists[2], we can overcome this obstacle and obtain the following Uto-Aztecan classification (utza.dbf[3]):

Uto-Aztecan
1. *Numic
2. Main group
   a. Central:
      (i) Tubatulabal, Hopi
          Kitanemuk
      (ii) *Cupan
   b. Southern:
      (i) *Sonoran
      (ii) *Aztec

The weakest point of this classification is that it places Hopi close to Tübatulabal; this may also reflect later contacts between the two, not detected so far.

According to glottochronological calculations[4] based on the complete 100-item lists for attested languages, the Proto-UA language began to disintegrate approximately 4,500 — 4,200 years ago. It should be mentioned, however, that calculations based on Sergei Yakhontov/Sergei Starostin's 35-item list reveal an entirely different age — only 3,500 years. This is surprising, since usually calculations based on the shorter list generate slightly older dates than those based on the longer list. The reasons for such a discrepancy remain unknown.

Numic languages

The Numic group is usually divided into three branches, with two languages each — Southern: Kawaiisu and Ute (Chemehuevi, Southern Paiute and Ute); Central: Comanche and Shoshoni (Panamint and Shoshoni dialects); Western: Mono and Northern Paiute.

This classification is accepted by Miller (1984), but the matrix of similarities given is his paper suggests a different tree structure:

---

[2] See [G. Starostin: 2009].

[3] All of the databases mentioned in this paper will be available on EHL's website (http://starling.rinet.ru) in the near future and can also be obtained upon request from the author.

[4] If not indicated otherwise, all the lexicostatistical and glottochronological results have been generated with the aid of STARLING, a computerized linguistic environment developed by S. Starostin.

	Mon	NPa	Pan	Sho	Com	Kaw	Che	Spa	Ute
Mono	×	77	59	58	57	52	50	53	52
N.Paiute	77	×	58	58	58	56	55	58	57
Panamint	59	58	×	87	79	54	61	62	59
Shoshoni	58	58	87	×	88	55	58	62	61
Comanche	57	58	79	88	×	49	54	59	59
Kawaiisu	52	56	54	55	49	×	75	79	76
Chemehuevi	50	55	61	58	54	75	×	86	78
S.Paiute	53	58	62	62	59	79	86	×	87
Ute	52	57	59	61	59	76	78	87	×

Numic (more than 49%)

    A. Central-Northern Numic (more than 57%)

        1. Western: Mono and Northern Paiute (77%)

        2. Central: Panamint, Shoshoni, and Comanche (more than 79%)

    B. Southern (more than 75%)

        3. Ute: Chemehuevi, Sothern Paiute, and Ute (more than 78%)

        4. Kawaiisu

The same classification is obtained using S. Starostin's 35-item list (numi.dbf).

Southern Numic languages

Phonologically, the Ute languages are very similar to each other. Southern Paiute, studied in [Sapir 1930], has the following system:

Initial consonants:

p-	m-	w-	
t-	n-		
c-		y-	s-
k-			
kʷ-			
ʔ-			h-

Word-medial position (single consonants and clusters):

-v-	-pp-	-mp-	-mm-		-w-	
-r-	-tt-	-nt-, -nč-	-nn-	-n-		
-c-	-cc-	-nc-			-y-	-ss-
-ɣ-	-kk-	-ŋk-		(-ŋ-)		
-ɣʷ-	-kkʷ-	-ŋkʷ-		-ŋʷ-		
-ʔ-						-h-

Final consonants: none

Vowels (short and long):

i, i:	ɨ, ɨ:	u, u:
e, e:		o, o:
	a, a:	

The language is well known for its "consonant gradation". Sapir (1930, 62) talks about spirantization (S), gemination (G), and nasalization (N):

S: *uɣʷi-vi* 'grass'+Abs

G: *kučča-ppɨ* 'ashes'+Abs

N: *mɨyɨ-mpɨci* 'gopher'+Abs

Morphophonemic [Sapir 1931], accentual [Miller 1980], and other explanations were suggested for these phenomena, but the real reason for "gradations" seems to be purely historical. A simple reconstruction may be based on the following considerations:

— the stem in Southern Paiute is usually formed by two syllables and thus may have the form of (a) CVPPV, where PP is any geminated stop, affricate or ʔ; (b) CVPV, where P is any other consonant or a cluster of nasal + stop;

— in most cases Gemination is associated with the first type of stems, whereas Spirantization is mainly found within the second: (a) CVPPV + G, (b) CVPV + S;

— therefore, Spirantization and Gemination, predetermined by the consonants in the second syllable, must have developed later in the history of the languages;

There are, however, also other cases, each with a significantly smaller number of examples: (c) CVPPV + S: *kɨcci-vi* 'saliva' Abs, (d) CVPV + G: *movi-ppɨ* 'nose' Abs. Historically these can be regarded as showing traces of final consonants: *h (or *s) and *ʔ. Nasalization can be explained as a trace of a lost final nasal: (e) *CVPVN, *CVPPVN.

As a result, we can identify five types of stems in Southern Paiute:

(a) *CVPPV#: *kučča# > kučča-ppɨ* 'ashes'

(b) *CVPV#: *uɣʷi# > uɣʷi-vi* 'grass'

(c) *CVPPVh: *kɨccih > kɨcci-vi* 'saliva'

(d) *CVPVʔ: *moviʔ > movi-ppɨ* 'nose'

(e) *CVPVN, *CVPPVN: *mɨyɨN > mɨyɨN-mpɨci* 'gopher', *wikkuN > wikku-mpɨcc* 'buzzard'

The same three types of "consonant gradation" are found in the other Ute languages as well as Kawaiisu.

The Kawaiisu phonological system, defined on the original stems of the language, is found in [Zigmond et al. 1990, pp. 5–6]:

p	v	b	m	w	
t	r	d	n		
c	ʒ				s
č	ǯ				š
k	ɣ				
kʷ	ɣʷ				
ʔ					h

i, i:	ɨ, ɨ:	u, u:
e, e:		o, o:
	a, a:	

The Proto-Southern Numic system of consonants is reconstructed as follows:

Finals:

*#	*-ʔ	*-h	*-N

Initial position:

*p	*m	*w	
*t	*n		*s
*c		*y	
*k	*ŋ		
*kʷ			
*ʔ			*h

Medial position:

*-b-	*-pp-	*-m-	*-mm	*-w-	
*-d-	*-tt-	*-n-			*-s-
*-ʒ-	*-cc-			*-y-	
*-g-	*-kk-	*-ŋ-			
*-gʷ-	*-kkʷ-				
	*-ʔ				*-h-

The supporting correspondences are given in the following table:[5]

Kawaiisu	Proto-S. Numic	S. Paiute	Chemihuevi	Ute	Note
p	*p-	p	p	p	
v	*-b-	v	v	v	
p	*-pp-	pp	p	p	
b	*-mb-	mp	mp	p	< *N-pɨ
m	*m-	m	m	m	
m	*-m-	mm	m	m	
w	*w-	w	w	w	
w	*-w-	#	#	#	
w		ŋʷ	w	w	-VN, ɨwV
t	*t-	t	t	t	
r	*-d-	r	r	r	

---

[5] The sources are as follows: Kawaiisu — [Zigmond, et al. 1990]; Southern Paiute — [Sapir 1931]; Chemihuevi — [Press 1979]; Ute — [Givon 1979].

Kawaiisu	Proto-S. Numic	S. Paiute	Chemihuevi	Ute	Note
t	*-tt-	tt	t	t	
n	*n-	n	n	n	
n	*-n-	n	n	n	
c	*c-	c	c	č	
ʒ, ǯ	*-ʒ-	c	c	č	
c	*-cc-	cc	c	č	
s, š	*s-	s	s	s	
s, š	*-s-	ss	s	s	
y	*y- ,	y	y	y	
y	*-y-	y	y	y	
k	*k-	k	k	k	
g	*-g-	g	g	g	
k	*-kk-	kk	k	k	
ɣ	*-nk-	ŋk	ŋk	k	
n	*-ŋ-	ŋ	ŋ, (u+m+u)	#	
kʷ	*kʷ-	kʷ	kʷ	kʷ	
ɣʷ	*-gʷ-	w	gʷ	ɣʷ-, ɣ(+ɨ)-	
kʷ	*-kkʷ-	kkʷ	Kʷ	kʷ	
ʔ	*ʔ-	#, ʔ	#, ʔ	#	
ʔ	*-ʔ-	#, ʔ	ʔ	ʔ	
h	*h-	ʔ, #	h	ʔ, #	
h	*-h-	#	h	#	

The etymological data supporting these correspondences can be found in the database numiet.dbf. The vocalic system of Proto-Southern Numic requires further investigation.

Central Numic languages

All Central Numic languages have six vowels (i, e, ɨ, a, u, and o) that can be short or long. The systems of initial consonants of modern Central Numic languages are identical to those of Southern Numic, while in the medial position there occur many more phonemes (single consonants or clusters).

Medial position consonants of Shoshoni [SOD]:

b	pp	mb	ʔb	m	(mm)	hm		w	hw			
d	tt	nd	ʔd	n	nn	hn	ʔn				s	ss
ʒ	cc	nʒ						y	hy	ʔy		
g	kk	ŋg		ŋ		(hŋ)						
gʷ	kkʷ	ŋgʷ		ŋʷ								
	ʔ										h	

It is possible that all clusters containing ʔ, *h* and nasal consonants are not actually found within roots, but are rather the results of various morphophonemic processes.

Another noticeable Central Numic feature is the existence of an additional type of "consonantal gradation". Thus, in Panamit one finds the following types of articulation:

— spirantized: [naɣavaʔa] 'on the bighorn' (naka 'bighorn', paʔa 'on')
— geminated: [tuappaʔa] 'on the son' (*tuaʔ 'son', paʔa 'on')
— nasalized: [pɨyɨmbaʔa] 'on the duck' (pɨyɨn 'duck', paʔa 'on')
— aspirated: [haińčiɸaʔa] on the friend' (haińčih 'friend', paʔa 'on')

(Miller et al. 2005: 415)

Apparently the "aspirated gradation" corresponds to the final *-h* that is reconstructed for Southern Numic.

The following reconstruction of Proto-Central Numic is based only on my analysis of four languages / dialects: Panamint [Dayley 1989], Shoshoni and Western Shoshoni [SOD], Comanche [Robinson and Armagost 1990].

The sets of initial and final consonants are identical to those of Proto-Southern Numic, with the exception of *ŋ-, not found in PCN.

The set of PCN medial position consonants is as follows:

*-b-	*-ʔb-	*-pp	*-mb-		*-mm-	*-hm-	*-ʔm-	*-w-	*-hw-	
*-d-	*-ʔd-	*-tt-	*-nd-	*-n-	*-nn-	*-hn-	*-ʔn-		*-ss	
*-ʒ-	*-ʔʒ-	*-cc	*-nʒ-					*-y-	*-hy-	*-ʔy-
*-g-	*-ʔg-	*-kk	*-ŋg-	*-ŋ-		(*-hŋ-)	*ʔŋ			
*-gʷ-		*-kkʷ-	*-ŋgʷ-	(*-ŋʷ-)						
	*ʔ								*-h-	

It should be noted that some of the reconstructed medial consonants may in fact be clusters found at morpheme boundaries.

The following correspondences support this reconstruction:

Proto-C. Numic	Panamint	Shoshoni	W. Shoshoni	Comanche	Notes
*p	p	p	p	p	
*-b-	p	b	p	b / hp, p	
*-ʔb-	p	:ʔb, (:b)	p	bV̱, :bV̱	
*-pp-	pp	p	pp	hp	
*-mb-	mp	mb	mp	:p, p	
*m-	m	m	m	m	
*-ʔm-	m	:m, m	m	m	
*-m-	mm	mm	m	m	
*-hm-	hm	hw	hm	Hm	
(*-)	m	ʔw, :ʔw	:w	ʔm, :mV̱	2 examples
*w-	w	w	W	w	
*-w-	w	#	#	w	
*-hw-	hw		hw, h	h	

Proto-C. Numic	Panamint	Shoshoni	W. Shoshoni	Comanche	Notes
*t-	t	t	t	t	
*-d-	t	t	t	r	
-ʔd-	t	:ʔd, ʔd	t	:r, (e:t)	
*-tt-	tt	t	t	ht	
*-nd-	nt	nd	nt	t	
*n-	n	n	n	n	
*-nn-	nn	nn	n	n	
*-n-	n	n	n	n	
*-hn-	hn	hn	hn	hn	
*-ʔn-	nn	:ʔn		ʔn, :n	
*c	c	c	t	c	
*-ʒ-	c	ʒ, ǯ	c	c	
*ʔʒ	c	:ʔʒ, :ʒ	c	c	
*-cc-	cc	c, č	cc	hc, :c	
*-nʒ-	nc	nʒ, nǯ		c	
*s-	s	s	s	s	
*-ss-	s	s, š	s	s	
*y-	y	y	y	y	
*-y-	y	y	y	y	
*-hy-	hy	hy	hy	hy	
*-ʔy-	y	ʔy	y	:y, y	
*k	k	k	k	k	
*-g-	k	g	k	k	
*-kk-	kk	k	kk	hk, (:k)	
*-ŋg-	ŋk	ng	nk	k , :k	
-ʔg-	kVʔ	:ʔg	k	:k	
*-ŋ-	ŋ	n	n	n	
*-ʔŋ-	ŋ	nn		n	
(*-hŋ-)	hm	hn	hm	Hm	1 example
*kʷ	kʷ / k(+u)	kʷ / k(+u)	kʷ / k(+u)	kʷ / k(+u)	
*-gʷ-	kʷ	gʷ / (u+)g	ku / (u+)k	kʷ / (u+)k	
*-kkʷ-	kkʷ	/ (u+)k	kʷ /	kʷ / (u+)hk	
*-ŋgʷ-	ŋkʷ	ngʷ	nkʷ	kʷ	
(*-ŋʷ-)	ŋʷ			w	1 example
*ʔ-	#	#	#	#	
*-ʔ-	ʔ	ʔ	ʔ	ʔ	
*h-	h	h	h	h	
*-h-	#, h	h	h	h	

The following reconstruction of Proto-Numic is based on a comparison of Proto-Central and Proto-Southern systems. A significant gap, however, is the absence of data on Western Numic languages.

A preliminary reconstruction of the vowel system seems to be quite simple (it is not quite clear if we need to reconstruct long vowels):

*i	*ɨ	*u
		*o
	*a	

The consonantal system of Proto-Numic is as follows:

Initial consonants:

*p	*m	*w	
*t	*n		*s
*c		*y	
*k			
*kʷ			
*ʔ			*h

Final consonants:

*#	*-ʔ	*-h	*-N

Medial consonants:

*-b-	*-pp	(*-mb-)	*-m-	(*-hm-)	*-w-		
*-d-	*-tt-		*-n-	*-hn-			*-s-
	*-cc				*-y-	*-hy-	
*-g-	*-kk	*-ŋg-	*-ŋ-	*-hŋ-			
*-gʷ-	*-kkʷ-		*-ŋʷ-				
	*ʔ						*-h-

Some of the medial combinations seem to have been clusters, found on morpheme boundaries. It is possible that further investigation will fill in a few gaps in the reconstruction, such as an unexpected absence of *-ʒ-.

Correspondences supporting the reconstruction are:

*Proto-Numic	PCN	PSN	Notes
*p-	*p-	*p-	
*-b-	*-b-	*-b-	
*-pp-	*-pp-	*-pp-	few examples
(*-mb-)	*-mb-	*-mb-	alternatively: *-N-pa
*t-	*t-	*t-	
*-d-	*-d-	*-d-/-c-	c — next to *i

*Proto-Numic	PCN	PSN	Notes
*-tt-	*-tt-	*-tt-	few examples
*c-	*c-	*c	
*-cc-	*-cc-	*-cc-	
*k-	*k-	*k-	
*-g-	*-g-	*-g-	
*-kk-	*-kk-	*-kk-	
*-ŋg-	*-ŋg-	*-ŋg-	
*kʷ-	*kʷ-	*kʷ	
*-gʷ-	*-gʷ-	*-gʷ-	
*-kkʷ-	*-kkʷ-	*-kkʷ-	
*m-	*m-	*m-	
*-m-	*-m-, *-mm-	*-m- / *-w-	
(*-hm-)	*-hm-	*-m-	1 example
*n-	*n-	*n	
*-n-	*-n-, *-nn-	*-n-	
*-hn-	*-hn-	*-n-	
*-ŋ-	*-ŋ-	*-ŋ-	
*-hŋ-	*-hŋ-	*-ŋ-	
*-ŋʷ-	*-ŋʷ-, *-m-	*-w- / #	# after *o
*w-	*w-	*w-	
*-w-	*-w-	*-w-	few examples
*y-	*y-	*y-	
*-y-	*-y- / #	*-y- / #	# after *i
*-hy-	*-hy-	*-y-	
*s-	*s-	*s-	
*-s-	*-s-, *-ss-	*-s-	
*h-	*h-	*h-	
*-h-	*-h-, #	*-h-, #	
*ʔ-	*-ʔ-	*ʔ-	
*-ʔ-	*-ʔ-	*-ʔ-	

Examples supporting these correspondences are collected in numiet.dbf. Further investigation is needed to finalize the set of PN medial consonants and firmly distinguish them from other possible morpheme-boundary clusters.

The reconstructed system allows for further 'fine tuning'. E. g., one can suggest that the two types of stops and affricates actually represent an opposition between plain and glottalized consonants: *p ~ *p', *t ~ *t', etc., which is neutralized in the initial position.

The EHL Numic etymological database (numiet.dbf) consists of more than 600 entries. Most of these represent modified versions of etymologies found on such lists as [Iannucci 1972] and [Miller 2003].

Glottochronological calculations suggest that the disintegration of Proto-Numic may have begun about 2,500 years ago.

Takic languages

Miller (1984) distinguishes two branches of Takic languages:

a. Serrano-Gabrielino
   (1) Serranan: Serrano, Kitanemuk
   (2) Gabrielino (Gabrielino, Fernandino)
b. Cupan
   (1) Cupeño, Cahuilla
   (2) Luiseño

His lexicostatistical matrix, however, does not support any reasonable classification of Serrano and Gabrielino, probably due to the quality of the wordlists:

	Gabrielino	Serrano	Cahuilla	Cupeño	Luiseño
Gabrielino	×	45	42	34	38
Serrano	45	×	50	38	35
Cahuilla	42	50	×	65	50
Cupeño	34	38	65	×	48
Luiseño	38	35	50	48	×

It is possible to improve the results for Serrano by including forms only from Kroeber (1907) and by adding Kitanemuk words collected by Harrington [Anderton 1988]. A new matrix (based on taki.dbf) is:

	Kitanemuk	Serrano	Cahuilla	Cupeño	Luiseño
Kitanemuk	×	75	50	50	50
Serrano	75	×	61	58	52
Cahuilla	50	61	×	80	61
Cupeño	50	58	80	×	61
Luiseño	50	52	61	61	×

Some differences in the percentages of lexical similarities between the languages are mainly due to differences in the wordlists. Nevertheless, the overall pattern remains the same, suggesting that Miller's classification (without Gabrielino) may be accepted:

Takic (more than 52%)
   a. Kitanemuk + Serrano (75%)
   b. Cupan (more than 61%)
         Cahuilla + Cupeño (80%)
         Luiseño

The three Cupan languages are quite well described. The system of Luiseño can be presented as follows:

Vowels (spelled as i, u, e, a, o):

i		u
ɛ		ɔ
	a	

According to Bright (1965), all vowels may be short / long and stressed / unstressed, but certain structures predetermine the choice of length and / or strength. For example, 'in final accented position, only long vowels occur' (p. 344); 'when an initial syllable has a long vowel, it is always accented' (p. 345).

Consonants:[6]

p	v	m	w	
t		n	l	s
			r	ṣ
č			y	
k		ŋ		x
kʷ				kʷ
q				
qʷ				
ʔ				h

Elliott (1999: 14–15) writes that (all?) consonants between two short vowels are either long or accompanied by ʔ if the first vowel is stressed.

My analysis of Takic comparisons suggests that the vocalic length distinction in Luiseño is of later origin[7]. The rules to eliminate it are as follows:

— the root may consist of one or two syllables: CV, CVC(C), CVC(C)V, CVC(C)VC(C);

— in **nouns**, monosyllabic roots always have long vowels unless the final consonant is ʔ: *ku:ŋ* 'husband', *wiʔ* 'fat, grease';

— in disyllabic words stress can be placed on the first or the second syllable. If a stressed syllable ends in a consonant[8], the vowel is long, otherwise the vowel is short: *ʔáana* 'ant', *ʔamúu* 'mescal plant, Agave deserti', *ʔéčvu-š* 'left hand', *qawíʔ* 'forehead';

— **verbs** are predominantly monosyllabic with short vowels. Disyllabic verbs follow the same rules as nouns: *ʔuláʔq=i-* 'to sew (vt.)'.

When modern Luiseño forms do not follow the suggested pattern, we assume that they may have recently lost a consonant: *čáay-i-* 'to sift, winnow' < ***čáʔay-i-* (cf. Cahuilla *čáʔa-l* 'sieve, trap'), *qíin-i* 'to plow' < ***qíyin-i* (cf. Cahuilla *qíyne-* id.).

---

[6] I was not able to complete a list of consonantal clusters acceptable in Luiseño roots.

[7] For other views on stress in Cupan languages see Hill and Hill 1968, Munro 1990, etc.

[8] One should mention, however, that in some cases it is not clear how to divide a word into syllables, cf. *kúumal* 'poker' < *** kúu=mal* (*ku-* 'fire').

The Proto-Cupan system is reconstructed as followed:

Consonants:

*p	*v	*m	*w	
*t		*n	*l	*s
*č			*y	
*k		*ŋ		*x
*kʷ				
*q				
*qʷ				
*ʔ				*h

Vowels:

*i	*ɨ	*u
*e (=*ɛ)		*o (=*ɔ)
	*a	

The correspondences behind these proto-phonemes are mostly trivial.

The EHL etymological database (takiet.dbf) currently contains 562 entries that reproduce, with some modifications, Miller's list of Proto-UA etymologies [Miller 2003], with additional comparisons from [Munro 1990] and some other sources. No direct search for new etymologies has been conducted.

Glottochronology suggests that the Proto-Takic language may have begun to disintegrate about 3,200 years ago.

Sonoran languages

A list of Sonoran languages can be found in [Miller 1984]:

1. Tapiam: Upper Piman (Papago, Pima, Nevome), Lower Piman, Northern Tepehuan, Southern Tepehuan (Southern Tepehuan, Tepecano)
2. Taracahitian
   (a) Tarahumara (several dialects)
       Guarijio (Highland, Lowland)
   (b) Opatian: Opata, Eudeve, Jova?
   (c) Cahita (Mayo, Yaqui)
3. Tubar
4. Corachol: Cora, Huichol.

The Opatian languages are extinct and I was not able to obtain wordlists for any of them.

Tapiam. A brief description of Papago phonology can be found in [Mathiot 1973: 113–114]:

	Bilabial	Dental	Alveolar	Palatal	Velar	Glottal
Tense stop	p	t		č	k	
Lax stop	b	d	ḍ	ǯ	g	ʔ
Fricative			s	š		h
Nasal	m	n				
Lateral			l			
Glide	w					

Vowels (short and long):

i	ɨ	u
		o
	a	

Stress is predictable.
The systems of other Tepiman languages are listed in [Bascom 1965].

A reliable reconstruction of Proto-Tepiman was conducted by Bascom in 1965 and is the main source of the Tepiam database (tepiet.dbf, with more than 300 etymologies). Additional lexical data are available only for Papago (Saxton & Saxton 1983; Mathiot 1973) and Nevome (Pennington 1979). Papago is now sufficiently well represented in the database. Nevome, known only from a dictionary composed in the XVIIth or XVIIIth centuries, has not been studied.
The following system is proposed for Proto-Tepiman:

Root structure: CVCV with no consonantal clusters.
Consonants (initial and medial):

*p	*b	*m	*v	
*t	*d	*n	*r	*s
*k	*g			
*ʔ	*h			

Vowels (may be short or long):

*i	*ɨ	*u
		*o
	*a	

This reconstruction is based on the following set of phonological correspondences:

Tepiman	N. Tepehuan	S. Tepehuan	U. Pima	L. Pima
*p	p	p	p	p
*b	b	b	b	b

Tepiman	N. Tepehuan	S. Tepehuan	U. Pima	L. Pima
*m	m	m	m	m
*v	v, #	v	v	v
*t	t / tʸ (i+, +i)	t / tʸ (i+, +i)	č / t (+i,ɨ,u)	t / č (i+, +i)
*d	d / dʸ (i+, +i)	d / dʸ (i+, +i)	ǯ / d (+i,ɨ,u)	d / ǯ (i+, +i)
*n	n / ń(i+, +i)	n / ń(i+, +i)	ń/ n (+i,ɨ,u)	ń, n (i+, +i)
*r	r / l (i+, +i)	r / lʸ (i+, +i)	ḍ / l (+i)	r / l (+i)
*s	s / š (i+, +i)	s / š (i+, +i)	ṣ̌ / s (+i)	ṣ̌ / s (i+, +i)
*k	k, h	k	k, h	k
*g	g, #	g, ʔ	g	g
*ʔ	#	ʔ	ʔ,#	ʔ,#
*h	h, #	h, #	h, #	h, #
*i	i	i	i	i
*ɨ	ɨ	ɨ	ɨ	ɨ
*u	u	u	u	u
*o	o	o	o	o
*a	a	a	a	a

These correspondences are supported by reconstructions, collected in tepiet.dbf.

Tarahumaran. The two main sources for Tarahumaran languages are a dictionary of Central Tarahumaran [Hilton et al. 1993], and a grammar and dictionary of Highland Guarijio [Miller 1996].

The Guarijio system of phonemes is as follows:

Consonants (initial and medial):

p	m	w	
t	n	l	s
č		y	
k	g		
ʔ			h

Two types of medial clusters are found in the roots: -ʔ + consonant and -h + consonant.

Vowels:

i		u
e		o
	a	

Stress is usually placed on the second (last) syllable of the root, but there are a significant number of exceptions.

The system of Tarahumara is quite similar, as is the reconstructed Proto-Tarahumaran. Minor issues of the reconstruction are not discussed here.

The Tarahumaran database (taraet.dbf, more than 350 etymologies) is mainly based on Miller's materials. It should be mentioned, however, that the database includes only a small portion of possible etymologies for this group.

Mayo (Cahita). Lexical corpora for several closely related Mayo languages / dialects are found in three dictionaries: Mayo [Collard and Collard 1962], Yaqui [Estrada et al 2004] and Yoeme [Molina et al. 1999]. Only the Yoeme dictionary has been used for the time being. The phonological system of this variety is as follows:

Consonants (initial and medial):

b^w	p	m	v	w	
	t	n	R	l	s
				y	
	k	g			
	ʔ				h

No consonant clusters are found within root morphemes.

Vowels:

i		u
e		o
	a	

All vowels can be short and long; vowel length can sometimes fluctuate under certain conditions, probably caused by morphological factors that are not yet fully understood.

A small Mayo database was created with the sole aim of checking the validity of Yoeme forms used in comparisons with other Sonoran languages.

Corachol. For my study, two sources of Huichol data were available: a small dictionary of the language published as [McIntosh & Grimes 1954] and a manuscript of a much bigger dictionary prepared in 1982 [Grimes et al. 1982], which had never been published. It is this manuscript that is used as my primary source of information. The phonology of the language may be presented as follows:

Consonants (initial and medial):

p	m	W		
t	n		r	
c		Y		s
k				
k^w				
ʔ				h

No consonantal clusters are found.

Vowels (all vowels can be short or long):

i	ɨ	u
e		
	a	

Tones: H(igh) (V́), L(ow) (unmarked). The origin of tones is not yet fully understood.

The Cora data is taken from [McMahon & McMahon 1959], but has not yet been properly analyzed.

A few additional extinct Sonoran languages are known with various degree of detalization. None of them have been used in the reconstruction, but some of their words mentioned in various etymological sources are given in the database.

The Proto-Sonoran system and a list of etymologies (about 350 entries) were produced by [Lionnet 1985]. The phonological system of the proto-language is as follows:

Consonants (initial and medial):

*p	*m	*w	
*t	*n	*l	*s
*č		*y	
*k			
*kʷ			
*ʔ			*h

Proto-Sonoran *l is not found in the initial position.
The suggested reconstruction is based on the following set of correspondences:

Proto-Sonoran	*Tepiman	Yoeme	Huichol	*Tarahumara
*p-	*p, *v	v, p	h	#
*kʷ	*b	bʷ	kʷ	*w, *kʷ(+u)
*w	*g, v	v, w	w	*w
*t	*t	t	t, č	*t
*č	*s	č	č	*č
*s	*h	s	s	*s
*h	*ʔ	h	#	*ʔ
*ʔ	*ʔ	#	#/ Vʔ-	#
*k	*k	k	k	*k
*m	*m	m	m	*m

Proto-Sonoran	*Tepiman	Yoeme	Huichol	*Tarahumara
*n	*n	n	n	*n
*l	*r, *d	l	l	*l
*y	*d	y	y	*y

Vowels:

*i	*ɨ	*u
		*o
	*a	

The vocalic correspondences are straightforward.

Although, as has already been mentioned, some Sonoran languages distinguish between long and short vowels, this distinction is not reconstructed for the proto-language.

Proto-Tepiman roots with long vowels normally correspond to Proto-Tarahumara roots with consonantal clusters. This observation permits to reconstruct Proto-Sonoran clusters as *-ʔC- and *-hC-. The root structure in the proto-language, best preserved in Guarijio, is *CV(H)CV. Further study is needed to clarify the origin of long vowels in Yoeme and Huichol.

The Proto-Sonoran etymological database (sonoet.dfb, about 400 entries) consists of all the acceptable comparisons from Lionnet's and Miller's sources. Additional research has also been conducted to find missing reflexes in various languages of the family.

Miller (1984) has identified four equal branches (groups) of the Sonoran languages. An alternative interpretation of his matrix, however, gives us the following tree:

Sonoran (34% and more):
    A. Papago, Nevome, N. Tepehuan (79% and more)
    B. 39% and more:
        1. 51% and more:
            (a) Guarijio and Tarahumara (83%);
            (b) Opta and Eudeve (73%);
            (c) Mayo and Yaqui (93%);
        2. Tubar
        3. Corachol (58%)

This classification is confirmed by my own lexicostatistical study based on 35 and 50 item-lists (sono.dbf). The 100-list gives, however, a different picture:

	Nevo	Papa	Guar	Tara	Yoem	Cora	Huic
Nevome	×	93	56	52	59	46	50
Papago	93	×	53	49	57	46	48
Guarijio	56	53	×	81	66	51	54
Tarahumara	52	49	81	×	61	47	51
Yeome	59	57	66	61	×	53	59
Cora	46	46	51	47	53	×	74
Huichol	50	46	54	51	59	74	×

The differences in classifications are presumably caused by etymological misinterpretations of the (so far) poorly understood data from Cora. Without it, the classifications are identical:

A. *Tepiam
B   1a. Tarahumaran
    1b. Yeome
    2. Huichol

Glottochronology suggests that Proto-Sonoran began to disintegrate about 3,200 years ago.

Aztecan languages

Miller classifies the Aztecan group into two branches: a. Pochutec, b. General Aztecan: Pipil, Aztec (Classical Aztec. Tetelcingo, Zacapoaxtla, and others). The extinct language Pochutec, known from a short article by Boas (1917), was not included in Miller's matrix. Nevertheless, lexicostatistical calculations support his view:

35-item matrix:

	N.Pu	Tet	Pip	Zac	Mec	Poc
N.Puebla	×	94	94	85	81	75
Tetelcingo	94	×	97	88	84	78
Pipil	94	97	×	91	88	79
Zacapoaxtla	85	88	91	×	81	76
Mecayapan	81	84	88	81	×	78
Pochutec	75	78	79	76	78	×

A detailed discussion of modern Aztec dialects can be found in [Canger 1988]. The comparative phonology of the family is well-known [Dakin 1982]. In 1978 Campbell and Langacker published the following set of Aztecan vocalic correspondences (p. 94):

	C&L	Pochutec	Class.Azt.	Tetelcino	Zacapoaxtla	Pipil
1	*a	e	a	a	a	a
2	*a:	a	a:	ɔ	a:	a:
3	*o	o	o	o	o	o
4	*o:	u	o:	u	o:	u:
5	*i	i	i	I	i	i
6	*i:	i	i:	i	i:	i:
7	*ɨ	o	i	I	i	i
8	*ə	o,e,#	ə	e	e	ee
9	*e:	e	e:	ie	e:	e:

Dakin (1983) has shown that correspondence No. 7 is only found before consonants and thus does not reflect an additional Proto-Aztec vowel. A restricted distribution, how-

ever, is not a sufficient argument against a reconstruction, so the proto-system can be interpreted as:

*i (5)	*ɨ (7)	
*e (8)		*o (3)
	*a (1)	

*i: (6)		
*e: (9)		*o: (4)
	*a: (2)	

The Proto-Aztecan consonantal system, suggested by Dakin [1979: 50], is formed by 15 consonants:

*p	*m	*w	
*t	*n		
*λ		*l	
*c			*s
*č		*y	*š
*k			
*kʷ			
*ʔ			

This reconstruction is based on the following set of correspondences:

Proto	Pochutec	Class.Azt.	Tetelcino	Zacapoaxtla	Pipil
*p	p, (b)	p	p	p	p
*m	m	m	m	m	m
*w	w	w	w	w	w
*t	t, (d)	t	t	t	t
*n	n, ń	n	n	n	n
*λ	t, (d)	λ	λ	t	t
*l	l	l	l	l	l
*c	c	c	c	c	c
*s	s	s	s	s	s
*č	č	č	č	č	č
*š	š	š	š	š	š
*y	y	y	y	y	y
*k	g, (k)	k	k	k	k
*kʷ	g, (k)	kʷ	kʷ	kʷ	kʷ
*ʔ	h	ʔ	h	h	h

The phonemes *č and *š are often found in proximity with *i and *i:; this indicates their secondary origin.

The Aztec database (azttet.dbf) contains about 350 entries, which obviously represents only a small portion of the proto-language lexicon. Glottochronological calculations suggest that the disintegration of Proto-Aztecan began about 2,000 years ago; Kaufman [2001] talks about 500 CE.

Hopi

The dialects of Hopi form a separate branch of UA. Whorf [1946] identifies four of them (a claim that I am not able to verify through lexicostatistics): First Mesa (or Whorf's Polacca), Mishongnovi (or Whorf's Toreva), Shipaulovi (or Whorf's Sipaulovi), Third Mesa (or Whorf's Oraibi). The best known is the dialect of the Third Mesa [Hopi 1998]. Its phonological system is as follows:

Consonants:

p	v	m	w		
t		n	l		s
c				z̦ / r	
kʸ		ń	y		
k		ŋ			
kʷ		ŋʷ			
q	qʷ				
ʔ					h

The opposition of velars and uvulars is lost before front vowels.
Vowels (long and short) (the vowels are spelled as i, e, u, ö, a, o):

i	ɨ	
ø		o
ε	a	

Stress is predictable.

Tones: the Third Mesa dialect has developed tonal distinctions on long vowels, diphthongs, and vowel + sonorant sequences (*l, m, n, ŋ, ŋʷ*). The tones are either falling or level. The falling tone corresponds to one of several sequences in the Mishongnovi dialect recorded by Whorf [Manaster-Ramer 1986]: (a) vowel + preaspirated consonant; (b) vowel + voiceless sonorant; (c) vowel + *h*.

Proto-Uto-Aztecan

The resulting Proto-UA consonantal system can be reconstructed as follows:

Initial consonants:

*p	*m	*w	
*t	*n		*s
*č		*y	
*k	*ŋ		
*kʷ		*ɣʷ	
*q			
*qʷ			
*ʔ			*h

Medial consonants:

*p	*b	(*mb)	*m	(*w)	
*t	*d		*n	*l	*s
*č	*ǯ			*y	
*k	*g	*ŋg	(*ŋ)		
*kʷ	*gʷ		*ŋʷ	(*ɣʷ)	
*q	*G	(*ŋG)			
*q			(*Nʷ)		
*ʔ					

Proto-consonants in brackets are not supported by sufficient number of examples.

The system of final consonants is probably the same as that of Proto-Numic.

The consonantal correspondences are supported by etymologies in utatet.dbf. This database consists primarily of comparisons from Miller's list of 2003. The list, published on the Internet, is Hill's revised and expanded version of Miller's computerized database for Uto-Aztecan cognate sets of 1988. The list was rekeyed and made available for the purposes of the EHL project by Nikolayev; I have converted the list to the current database format, double-checked the actual language forms and often expanded the etymologies, using recently published dictionaries and other data sources.

The set of Proto-UA consonantal correspondences is as follows:[9]

*Proto UA	*Sonor	*Aztec	Tubatulaba	*Takic	Hopi	*Numic	Notes
*p-	*p	*p, #V:	p	*p	P	*p	
*-b-	*p	*p	p, b	*v	v, p	*b	
(*-p-)	*p			*p		*pp	no good exx.
(*-mb-)	*hp		mb				1 example
*m-, *-m-	*m	*m	*m	*m	m	*m	
(*w-)	*w	#	w	*w	#	*w	1 example
(*-w-)	*w	*w	w	*w	w	*w, #	
*t-	*t	*λ, *t	t	*t	t	*t	
*-d-	*t, ht	*t, *ɬ	d, l	*l	t	*d	
*-t-	*ht	t	l	l	t	*tt	
*n-, *-n-	*n	*n	n	*n	n	*n	
*-l-	*l	*l	n	*l	n	*hn	
*s-, *-s-	*s	*š, *s	s	*s	s	*s	
*č-	*č	*č, *c		*c	c	*c	
*-ǯ-	*č, *ǯ	*c	y, dᶻ	*y	*y	*y, *ǯ	
*-č-	*č, *hč	*č	c	*c		*y	

*Proto UA	*Sonor	*Aztec	Tubatulaba	*Takic	Hopi	*Numic	Notes
*y-, *-y-	*y	*y	y	*y	y	*y	
*k-	*k	*k	k	*k	k, kʸ	*k	
*-g-	*k	*k	g	*k	*k	*g	
*-k-	*k	*k	*k	*k	k, kʸ	*kk	
*-ŋg-	*hk		*ŋg	*k	k	*ŋg	
(*ŋ-)	*n	*n	n	*ŋ	ŋ	*n, *ŋ(ʷ)	no good exx.
(*-ŋ-)	*n		n	*n, ŋ	n, ŋ	*ŋ	
*-ŋʷ-	*m	n	m	m	m	*ŋʷ	
(*-Nʷ-)	*hw		ŋʷ, kʷ	*w, *ŋʷ	ŋʷ		no good exx.
*kʷ-	*kʷ	*kʷ	w	*kʷ	kʷ	*kʷ	
(*-gʷ-)	*kʷ	*w			kʷ	*gʷ-	1 example
(*-kʷ-)		*kʷ			kʷ	kkʷ	1 example
(*ɣʷ-)	*w	#	w	*w	l	*w	1 example
(*-ɣʷ-)	*w	*kw	w	w	l	#	
*q	*k	*k	h	*q	*q	*k	
(*-G-)	*hk	*k	h	*q	q	*g	1 example
*-q-	*k	*k	h	*x	q	*g	
(*-ŋG-)	hk	k	ŋ	*x	q	*ŋg	1 example
*qʷ-	*kʷ	*kʷ	w	*qʷ	*kʷ	*kʷ	
(* -qʷ-)	(*p)	*kʷ		*qʷ			1 example
*ʔ	*ʔ	#Vʔ, #	ʔ-, #	*ʔ	#	#	
*-ʔ-	#, *ʔ	#, *ʔ	ʔ-, #	#, *ʔ	#	#	
*h	*h	#Vʔ	h	*h	h	*h	

This reconstruction is, naturally, far from being complete: some proto-phonemes are not supported by a sufficient number of examples, while others demonstrate unexplained double reflexes in daughter-languages. Much more work is needed to obtain a detailed proto-UA reconstruction and a better etymological dictionary. Some etymologies in utatet.dbf may not belong to the Proto-UA level, but rather represent other chronological levels, areal roots, or even borrowings. These problems require further investigation.

It is suggested that the homeland of the Proto-UA family was located somewhere in Arizona and neighboring areas of New Mexico and Mexico (data in [Flower 1983]), based on an analysis of plant and animal names). Jane Hill [2001] made an attempt to demonstrate that the speakers of Proto-UA were agriculturalists, but her etymological conclusions are not well supported, since none of the proposed "agricultural" etymologies can be attributed to the level of the proto-language.

	Mn	NP	Pn	Sh	Cm	Ka	Ch	SP	Ut	Tb	Gb	Sr	Ca	Cu	Lu	Ho	Pp	Nv	NT	Gu	Ta	Op	Eu	Ma	Yq	Tu	Hu	Co	CA	Te	Za	Pi
Mono	×	77	59	58	57	52	50	53	52	39	26	26	29	28	26	33	23	24	25	26	23	26	28	27	29	28	25	25	18	19	17	16
N.Paiute	77	×	58	58	58	56	55	58	52	42	26	24	27	27	27	32	26	26	28	29	27	29	27	28	30	27	24	22	18	18	17	15
Panamint	59	58	×	87	79	54	61	62	59	37	26	24	27	24	25	27	25	24	26	23	21	21	23	25	26	27	23	22	16	16	15	14
Shoshoni	58	58	87	×	88	55	58	62	61	38	26	24	27	24	24	23	26	26	26	23	21	20	22	24	26	28	23	23	16	16	15	14
Comanchi	57	58	79	88	×	49	54	59	59	35	23	21	24	23	22	22	23	23	23	24	21	20	26	27	24	27	21	22	14	14	13	12
Kawaiisu	52	56	54	55	49	×	75	79	76	39	24	26	27	26	24	31	26	24	27	24	22	20	24	25	29	28	23	21	16	17	16	16
Chemehuevi	50	55	61	58	54	75	×	86	78	42	27	28	31	30	27	33	28	26	28	24	22	26	26	27	26	27	23	22	15	15	16	15
S.Paiute	53	58	62	62	59	79	86	×	87	39	26	27	31	31	27	31	28	26	30	25	23	24	24	25	29	30	24	23	16	16	17	16
Ute	52	52	59	61	59	76	78	87	×	40	27	27	29	28	26	32	30	27	29	28	26	23	26	27	30	31	25	23	16	17	18	17
Tübatulabal	39	42	37	38	35	39	42	39	40	×	40	35	38	37	34	38	35	35	37	36	32	33	27	28	35	33	32	30	24	25	26	24
Gabrielino	26	26	26	26	23	24	27	26	27	40	×	45	42	34	38	29	35	24	26	29	28	26	35	35	28	24	24	19	20	20	21	21
Serrano	26	24	26	24	21	26	28	27	27	35	45	×	50	38	35	29	27	27	30	34	34	31	26	29	32	28	28	21	22	22	24	23
Cahuilla	29	27	27	27	24	27	31	31	29	38	42	50	×	65	50	31	31	30	32	34	33	33	30	33	35	29	34	24	23	24	20	20
Cupeñõ	28	27	24	24	23	26	30	31	28	37	34	38	65	×	48	31	28	27	29	29	26	29	29	26	26	26	26	23	19	20	19	18
Luisenõ	26	27	25	24	22	24	27	27	26	34	38	35	50	48	×	26	25	24	26	28	28	24	25	28	28	23	27	22	19	19	24	24
Hopi	33	32	27	23	22	31	33	31	32	38	29	29	31	31	26	×	32	33	33	32	28	33	35	34	36	30	28	26	24	24	31	30
Papago	23	26	25	25	23	26	28	28	30	35	35	27	31	28	25	32	×	85	79	44	41	40	45	43	45	40	41	34	29	30	31	30
Né'vome	24	26	24	26	24	24	26	26	27	35	24	27	30	27	24	33	85	×	79	47	42	44	47	45	47	41	43	34	29	30	32	29
N.Tepehuan	25	28	26	26	23	27	28	30	29	37	26	30	32	29	26	33	79	79	×	47	42	40	45	49	49	46	42	35	30	30	29	33
Guarijío	26	29	23	23	24	24	24	25	28	36	29	34	34	29	28	32	44	47	47	×	83	55	59	58	58	48	51	42	32	32	33	34
Tarahumara	23	27	21	21	21	22	22	23	26	32	28	34	33	26	28	28	41	42	42	83	×	54	52	51	51	44	48	36	33	34	35	38
Opata	26	29	21	20	20	20	26	24	23	35	26	31	33	29	24	33	40	44	40	55	54	×	73	53	55	42	48	35	39	38	39	40
Eudeve	28	27	23	22	26	24	26	24	26	27	35	26	30	29	25	35	45	47	45	59	52	73	×	61	62	51	49	42	40	40	39	40
Mayo	27	28	25	24	27	25	27	25	27	28	35	29	33	26	28	34	43	45	49	58	51	53	61	×	93	51	48	45	38	38	37	39
Yaqui	29	30	26	26	24	29	26	29	30	35	28	32	35	26	28	36	45	47	49	58	51	55	62	93	×	53	51	46	39	39	38	39
Tubar	28	27	27	28	27	28	27	30	31	33	24	28	29	26	23	30	40	41	46	48	44	42	51	51	53	×	41	39	36	35	35	35
Huichol	25	24	23	23	21	23	23	24	25	32	24	28	34	26	27	28	41	43	42	51	48	48	49	48	51	41	×	58	39	37	33	35
Cora	25	22	22	23	21	22	21	22	23	30	19	21	24	23	22	26	34	34	35	42	36	35	42	45	46	39	58	×	37	35	33	35
Cl.Aztec	18	18	16	16	16	16	15	16	16	24	20	22	23	19	19	24	29	29	30	32	33	39	40	38	39	36	39	37	×	85	81	77
Tetelcingo	19	18	16	16	14	17	15	16	17	25	20	22	24	20	19	24	30	30	30	32	34	38	40	38	39	35	37	35	85	×	85	81
Zacapoaxtla	17	17	15	15	13	16	16	17	18	26	21	24	20	19	24	24	31	31	32	29	33	35	39	37	38	35	33	33	80	85	×	77
Pipil	16	15	14	14	12	16	15	16	17	24	21	23	20	18	24	30	30	29	33	34	38	40	39	39	39	35	35	35	77	81	77	×

Matrix from Miller 1984

Literature

ANDERTON, Alice J. (1988). *The Language of the Kitanemuks of California*. PhD. Dissertation. University of California, Los Angeles.

BASCOM, Burton William. 1965. *Proto-Tepiam (Tepehuan-Piman).* Ph.D. dissertation, University of Washington.

BOAS, F. 1917. El dialecto mexicano de Pochutle, Oaxaca. *IJAL* 1:9–44.

BRIGHT, William, and Jane H. HILL. 1967. The linguistic history of the Cupeno. *Studies in Southwestern Ethnolinguistics*, ed. By Dell Hymes, pp.351–371

BRIGHT, William. 1965. Luiseño Phonemics. *IJAL* 31:342–345.

BRIGHT, William. 1965. Luiseño Phonemics. *IJAL* 31:342–345.

CAMPBELL, Lyle, and Ronald W. LANGACKER. 1978. Proto-Aztecan vowels. *IJAL* 44:85–102, 197–210, 262–279.

CAMPBELL, Lyle. 1997. *American Indian Languages*. Oxford — New York: Oxford University Press.

CANGER, Una. 1988. Nahuatl dialectology: A survey and some suggestions. *IJAL* 54:28–72.

COLLARD, H. y E. COLLARD. 1962. *Castellano-mayo, mayo-castellano*. Istituto Linguistico de Verano.

DAKIN, Karen. 1979. Phonological Changes in Nahuatl: the Tense / Aspect / Number Systems. *IJAL* 45:48–71.

DAKIN, Karen. 1982. *La Evolutión Fonológica del Protonáhuatl*. Colección Lingüística indígena. Mexico: Universidad Nacional Autónoma de Mexico.

DAKIN, Karen. 1983. Proto-Aztecan vowels and Pochutec: An alternative analysis. *IJAL* 49:196–203.

DAYLEY, Jon P. 1989. *Tumpisa (Pamamint) Shoshone Dictionary*. UCPL 115. Berkeley and Los Angeles: University of California Press. Electronic edition 1998.

ELLIOTT, Eric B. 1999. *Dictionary of Rincon Luiseño*. Ph. D. dissertation.

ESTRADA FERNÁNDEZ, Zarina et al. 2004. *Diccionario yaqui-español y textos: obra de preservación lingüística*. Hermosillo and México, D.F.: Universidad de Sonora and Plaza y Valdés Editores.

FELIPE S. Molina, Herminia VALENZUELA, David Leedom SHAUL. 1999. *Yoeme-English English-Yoeme : with a comprehensive grammar of Yoeme language*. New York: Hippocrene Books.

FOWLER, Catherine S. 1983. Some lexical clues to Uto-Aztecan prehistory. *IJAL* 49:224–57.

GIVON, T., ed. 1979. *Ute Dictionary*. Preliminary ed. Ignacio, Col.: Ute Press.

GRIMES, Jose. 1981. El Huichol. MS

HALE, Kenneth. 1958–59. Internal diversity in Uto-Aztecan: I. *IJAL* 24:101–7, II. *IJAL* 25:114–21.

HILL, Jane H. 2001. Proto-Uto-Aztecan: a community of cultivators in Central Mexico? *American Anthropologist* 103:913–934.

HILL, Jane H, and Kenneth C. HILL. 1968. Stress in Cupan (Uto-Aztecan languages). *IJAL* 34:233–241).

HILL, Jane H., and Rocinda NOLASQUEZ. 1973. *Mulu'wetam: The Fist People*. Banning, Calif.: Malki Museum Press.

HILTON, Simón K. *Diccionario Tarahumara*. Istituto Linguistico de Verano.

IANNUCCI, David. 1972. *Numic historical phonology.* Ph.D. dissertation, Cornell University.

Hopi-English Dictionary of the Third Mesa Dialect. 1998. Tucson: The University of Arisona Press.

KAUFMAN, Terrence. 2001. *The History of the Nawa language group from the earliest times to the sixteenth century: some initial results*. MS

KROEBER, A.L. 1907. Shoshonean dialects of California. *UCPAAE* 4:65–165.

LAMB, Sydney M. 1964. The classification of the Uto-Aztecan languages: a historical survey. *UCPL* 34:106–25.

LIONNET, André. 1985 Relaciones internas de la rama sonorense. *Amerindia*, 10.

MANASTER-RAMER, Alexis. 1993. On Lenition in some Northern Uto-Aztecan Languages. *IJAL* 59:334–341.

MANASTER-RAMER, Alexis. 1986. Genesis of Hopi Tones. *IJAL* 52:154–160.

MATHIOT, Madeleine. 1973. *A Dictionary of Papago Usage*. CV.1. Bloomington: Indiana University.

MCINTOSH, Juan B. y Jose GRIMES. 1954. *Vocabulario Huichol-castellano. Castellano-huichol*. Istituto Linguistico de Verano.

MCLAUGHLIN, John E. 1992. Counterintuitive Solution in Central Numic Phonology. *IJAM* 58:158–181.

MCMAHON, Ambrosio y Miaria Aiton de MCMAHON. 1959. *Cora y Espanol*. Istituto Linguistico de Verano.

MILLER, Wick R. 1967. *Uto-Aztecan Cognate Sets*. UCPL 48. Berkeley and Los Angeles: University of California.

MILLER, Wick R. 1982. Southern Paiute and Numic Final Feature. *IJAL* 48:444–449.

MILLER, Wick R. 1984. The classification of the Uto-Aztecan languages based on lexical evidence. *IJAL* 50:1–24.

MILLER, Wick R. 1996. *Guarijío: gramática, textos y vocabulario*. Mexico: Universitad nacional autónoma de México.

MILLER, Wick R. 2003. *Uto-Aztecan Cognate Set*. Revised and expanded by Kenneth C. HILL (ms.).

Miller, Wick R., Dirk Elzinga, and John E. McLaughlin. 2005. Preaspiration and Gemination in Central Numic. *IJAL* 71:413–44.

Miller, Wick R., James L. Tanner, and Lawrence P. Foley. 1971. A Lexicostatistic Study of Shoshoni Dialects. *Anthropological Linguistics* 13/4:142–164.

Munro, Pamela. 1983. Selected studies in Uto-Aztecan phonology. *IJAL* 49:277–99.

Munro, Pamela. 1990. Stress and Vowel Length in Cupan Absolute Nouns. *IJAL* 56:217–250.

Pennington, Campbell W., ed. 1979 *Vocabulario en La Lengua* Nevome, The Pima Bajo of Central Sonora, Mexico. Vol. 2. Salt Lake Sity: University of Utah Press.

Press, Margaret. 1979. *Chemehuevi: Grammar and Lexicon*. UCPL. 92. Berkeley and Los Angeles: University of California Press.

Robinson, Lila W., and James Armagost. 1990. *Comanche Dictionary and Grammar*.

Sapir, Edward. 1931. *Southern Paiute Dictionary.* Publications of the American Academy of Arts and Sciences 65, no. 3.

Sapir, Edward. 1931. *The Southern Paiute language,* Parts I, II, and III. AAASP 65.1–730.

Saxton, Dean, and Lucille Saxton. 1983. Dictionary: *Papago and Pima to English, English to Papago and Pima*. Tucson: University of Arizona Press.

Seiler, Hansjakob, and Kojiro Hioki. 1979. *Cahuilla Dictionary*. Banning, Calif.: Malki Museum Press.

SOD: *Shoshoni Online Dctionary*, at: www.shoshonidictionary.com.

Starostin, George. 2010. Preliminary lexicostatistics as a basis for language classification: a new approach. *Journal of Language Relationship* 3:79–116.

Voegelin, C. F. 1958. Working dictionary of Tubatulabal. *IJAL* 24:221–28.

Voegelin, C. F., F.M. Voeglin, and Kenneth Hale. 1962. Typological and Comparative Grammar of Uto-Aztecan: I (Phonology). *IJAL Memoir* 17.

Whorf, Benjamin. 1946. The Hopi language, Toreva dialect. *Linguistic Structures of Native America*, ed. by H. Hoijer, 158–183. New York : The Viking Fund, Inc.

Zigmond, Maurice L., Curtis G. Booth, and Pamela Munro. 1990. *Kawaiisu. A Grammar and Dictionary with Texts.* UCPL 119. . Berkeley and Los Angeles: University of California Press.

Статья представляет собой краткий обзор наработок по исторической фонологии юто-ацтекской языковой семьи. Основными предметами обсуждения являются генетическая классификация языков этой семьи, реконструкции фонологических систем праязыков ее отдельных ветвей и собственно праюто-ацтекская реконструкция. Лексический материал, на основании которого разработаны соответствующие реконструкции, взят из электронных баз данных, созданных в рамках проекта «Эволюция языка» (Институт Санта-Фе).