Type: Book Chapter

Introduction

Author(s): Brian D. Stubbs
Source: *Exploring the Explanatory Power of Semitic and Egyptian in Uto-Aztecan*
Published: Provo, UT; Grover Publications, 2015
Pages: 1-64

Introduction

Explanatory power is the ability of a hypothesis to explain what is otherwise not explainable. Explanatory power is what linguists look for to identify the best among competing theories to explain what happens in language. For a century, the answers to many unresolved questions in Uto-Aztecan linguistics eluded Uto-Aztecan (UA) specialists. While the language ties in this title may seem unseemly to some, they provide more explanatory power to the unknowns of UA than many might be comfortable with initially. So take your time. This study is exploratory, a work in progress toward answers, not yet having them all. Nevertheless, if the ties are valid, then ignoring them is like finding written records of Proto-Indo-European (PIE) and then ignoring those PIE records in comparative Indo-European studies. A valid key can provide instant progress to what would otherwise take decades or be impossible.

Uto-Aztecan is a Native American language family of some 30 related languages, mostly in the western United States and Mexico, from the Utes in the north to the Aztecs in the south, with Hopi, Pima, and others between (map on page 41). Some 1500 correlations between UA and three Near-Eastern languages, consistent with the linguistic comparative method (pages 9, 16), create a case at least as viable as the first accepted treatise establishing each Native American language family.

Knowing how unwelcome such a proposal would be in the linguistic community and being a peace-loving recluse by nature, I have been in no hurry to invite the avalanche of controversy upon me. However, equally risky is pressing my luck in postponing a presentation that should reside on this side of the mortal divide. So as youth becomes a more distant memory, I end the four-decade delay to share these findings, which, as both a Semiticist and a Uto-Aztecanist, I could not help but notice during three decades of writing the reference book *Uto-Aztecan: A Comparative Vocabulary* (UACV, Stubbs 2011), favorably received by Uto-Aztecanists, though no two UA specialists will agree on all aspects and reconstructions, as Kenneth Hill notes in a favorable review in the *International Journal of American Linguistics* (Hill 2012). After any comparative work, adjustments follow, and this work has a few adjustments to that 2011 work. A case not valid unravels with scrutiny, while truth is further substantiated with time, accumulating more and more support. So this work is not the final word, but an introductory offering. Let each consider all the data, then decide for oneself. Anything less is not a fair assessment. The strength of a case for language ties lies in the quantity and quality of the similarities, so to short either disqualifies a partial review as a fair judgement.

While this study is intended for linguists, Semiticists, and Egyptologists—and therefore includes the linguistic rigor demanded by the comparative method—it is also designed to be accessible to the astute and interested lay person by including explanations and (1.1) an introduction to linguistics (language science), which linguists can skip; (1.2) a brief outline of Semitic languages, which Semiticists can skip; (1.3) an even briefer word about Egyptian, which Egyptologists can skip; (1.4) and an introduction to UA, which even UA specialists should not skip. As the number who are specialists in all those areas approaches zero, most would benefit by perusing some of them. Of course, those lightly interested can skip them all, simply look at the pronunciation table, the Near-Eastern forms in bold, the associated UA forms, and get out of it what they may. However, for a better understanding, one is encouraged to read and refer to the introductions not within one's specialties. Sections 2, 3, 4, and 5 focus mainly on consonant correspondences of the 1500+ parallels; section 7.1 addresses vowel correspondences; section 7.3 treats grammatical and morphological parallels.

After Sapir (1913, 1915) established Uto-Aztecan as a viable family of related languages, Voegelin, Voegelin, and Hale (1962) produced the first numbered list of 171 cognate sets (groups of related words, page 13). Klar (1977) brought the Chumash languages to clarity with 168 sets. Taylor (1963) established Caddoan (a language family of the central plains), assembling 107 cognate sets. Hale (1962, 1967) did the definitive study for Kiowa-Tanoan with 99 sets. This work's proposal may better compare to tying two distant language families, as did Haas (1958) by ending four decades of controversy in uniting Algonkian-Ritwan, an eastern U.S. family with a west coast family, by means of 93 sets. Chamberlain (1888) began the union of Catawba with Siouan via 17 comparisons, and Siebert (1945) secured it with mostly morphological correlations, as not enough clear cognate sets were known at the time to establish correspondences (Campbell 1997, 140). Thus, the going rate is between 50 and 200 sets to establish most Native American language families. So this case of 1500 sets merits proportionate consideration.

Some characteristics of UA are different or not at all like Egyptian or Semitic, but reflect influences rather typical of Amerindian language families, which we would expect of a transplant from the outside into the Americas. One example is suppletion in singular vs. plural verb forms. That is, one verb is used for

singular subjects and an entirely different word is used when the subject is plural, while suppletion is nearly non-existent in Semitic or Egyptian. A score of such pairs in UA show such influences on UA. Semitic conjugation morphology (patterns of how verbs are conjugated) is not productive in UA, but hundreds of fossilized forms of both the suffixed/perfective conjugation and the prefixed/imperfective conjugation are found in UA. (See Introduction to Semitic 1.3, for Semitic conjugation morphology; see 1.12, for productive vs. fossilized, still producing forms vs. fixed and no longer producing forms.)

In contrast to differences, other grammatical features align and substantial amounts of Uto-Aztecan vocabulary produce consistent sets of sound correspondences (1.11, p. 13) between UA and the Near-Eastern languages, with each treated as a separate entity. For example, among the consistent patterns of sound correspondences, some 40 examples show Hebrew b corresponding to p of Proto-Uto-Aztecan (PUA); i.e., Hebrew / Phoenician b > PUA *p (> means 'became' or 'changed to'; < means 'changed from'; * marks a proto-form or original sound or word as reconstructed by linguists. So Hebrew b > PUA *p means Hebrew b changed to what linguists see as originally *p in UA). The following matches are a few from among many more examples of each sound change, and, of course, are naturally abbreviated from the fuller data and explanations found in the numbered paragraph sets. Non-linguists may want to read at this point the introduction to linguistics (1.1, p. 13) and the introduction to Semitic (1.2, p. 27). Verbs in Semitic consist of three consonants (bṣq, for example) subject to a variety of vowelings for different aspects, conjugations, adjectives, and nouns (C = any consonant, an unknown consonant):

Semitic b	> Uto-Aztecan *p
(527) baraq 'lightning'	> UA *pïrok / berok 'lightning'
(528) byt / bayit / beet 'spend the night, house'	> UA *pïtï; Tr bete 'house'
(528) byt / bayit / beet 'spend the night, house'	> UA *pïtï 'lie down, spend night'; Num *payïC 'go home'
(528) bytu 'spend the night, plural'	> UA *pïtu 'lie down, spend the night, plural'
(531) Hebrew boo' 'coming (used as 'way to')	> UA *pooC 'road, way, path'
(534) Hebrew batt 'daughter'	> UA *pattï 'daughter'
(550) Aramaic bəsár 'flesh, penis'	> UA *pisa 'penis'
(559) Semitic *bakay; Syriac baka' 'cry'	> UA *paka' 'cry'
(532) Arabic bṣr 'see'; baaṣirat 'eye'; Hebrew *booṣer(et)	> UA *pusi 'eye'
(535) Aramaic bəquuraa 'livestock'	> UA *pukuN 'domestic animals'
(540) Hebrew bṭḫ / *baṭiiḫ 'trust(ed)'	> UA *piciwa 'believe' (ṭ > c (=ts))
(552) bṭn 'be pregnant'	> UA *puca 'pregnant' (ṭ > c (=ts))
(553) bṣq 'swell'	> UA *posa 'swell'
(556) bayṣa(t) / beeṣa(t), pl: beeṣoot 'egg, testicle'	> UA *pïyso 'testicle'
(558) bwṣ / byḍ 'be white'; buuṣ 'white linen'	> UA *pos 'white': Tb poosït~'opoos 'be white'
(562) -bbiiṭ 'look'	> UA *pici / *pica 'look, see' (ṭ > c (=ts))

The other voiced stops also devoice, that is, Semitic b, d, g > UA p, t, k; also Semitic q > k:

(606) dubur 'buttocks, rear'	> UA *tupur 'hip, buttocks'	
(607) dobɛr 'pasture, vegetation'	> UA *tupi 'grass, vegetation'	
(1484) dwr / duur 'go round, turn, revolve'	> UA *tur 'whirl, roll, twist'	
(1103) dakka 'make flat, stamp, crush'	> UA *takka 'flat'	
(1279) *yagar 'hill, heap of stones'	> UA *yakaC / *yakaR (AMR) 'nose, point, ridge'	
(608) gdʕ 'cut off'	> UA *katu' 'cut, wound'	
(1014) qədaal 'neck, nape of neck'	> UA *kutaC 'neck'	(*q > k)
(1023) tqn 'make straight, set, lay down'	> UA *tïkaC 'put lying down, stretched/spread flat'	(*q > k)
(1089) Hebrew qippod 'hedgehog'; Arabic *qunpuð 'hedgehog'	> UA *kïNpa 'prairie dog'	(*q > k)
(864) *quuppoot 'baskets, pl'	> UA *koppo 'basket'	(*q > k)
(74) Hebrew təbuu'at 'produce from the land'	> UA *tïpï'at / *tïpat (AMR) 'pinion nut'	

Proto-Semitic đ (> Arabic đ, Aramaic d), corresponds to UA *t:

(616) Aramaic dakar 'male'	> UA *taka 'man, person'
(617) Aramaic diqn-aa 'beard / chin-the'	> UA *tï'na 'mouth'
(618) Aramaic di'b-aa 'wolf-the'	> UA *tï'pa 'wolf'
(620) unattested f. pl: *đabboot(ee^{y}) 'flies'	> UA *tïpputi 'flea'

Semitic ’aleph or glottal stop ’ > w in UA (which change also occurs in Arabic), or other times both a glottal stop and adjacent round vowels occur, perhaps ’ causing vowels to round (o, u):

(566) ’ariy / ’arii ‘lion’ > UA *wari ‘mountain lion’
(567) Hebrew ya’amiin-o ‘he believes him/it’ > UA *yawamin-(o) ‘believe (him/it)’
(569) Hebrew ’egooz ‘nut tree’ > UA *wokoC ‘pine tree’ (C = unknown consonant)
(571) ya’ya’ / yaa’ayaa’ ‘(be) beautiful’ > Ls yawáywa, Sr yï’aayï’a’n ‘be pretty, beautiful’
(572) Hebrew ’iiš ‘man, person’ > UA *wïsi ‘person’
(574) Hebrew ’išaa / ’ešɛt / **’išt-** ‘woman, wife of’ > UA *wïCti ‘woman, wife’ (C = unknown consonant)
(577) Semitic ’aas- ‘myrtle willow’ > UA *wasV ‘willow’
(579) Arabic pa’r- ‘mouse’ > UA *pu’wi(N) ‘mouse’
(581) Hebrew ’arṣ-aa ‘earth-ward, down’ > UA *wïcï ‘fall’
(575) kama’- ‘truffle(s)’ > UA *kamo’- ‘sweet potato’
(truffles are also edible fleshy appendages to a root system, as are potatoes)
(596) ’arnab ‘hare’ > UA *wa’na ‘rabbit net’
(576) ’ata^{y}, *’atii-; Syriac ’ita / ’ɛta ‘come’ > UA *wic ‘come’ (t > c(ts) by high vowels like i, u)
(871) ’pl / *tu’pal ‘be dark, go down (sun), f’ > UA *tu’pa > *cuppa ‘be dark, (fire) go out’ (t > c, by u)
(872) ’pl / *yu’pal ‘be dark, go down, m’ > UA *yu’pa > *yuppa ‘be dark, black, (fire) go out’
(873) ’pl / *yu’pal ‘be dark, go down, m’ > UA *yu’pa(l) > Aztecan *yowal, CN yowal-li ‘night, n’
Aztecan branch regularly loses a single -p-
(1110) Aramaic ’ard-aa’ ‘mushroom-the’ > UA *witto’oC ‘mushroom’
(1331) ’ikkaar ‘plowman, tiller of ground’ > UA *wika 'digging stick'
(1333) Hebrew m’n / *me’’an ‘refuse’ > Hp meewan- ‘forbid, warn’

Semitic initial r- > t- in UA:

(600) r’y / raa’aa ‘see, v’ > UA *tïwa ‘find, see’
(603) Aramaic rima / rimə-taa ‘large stone-the’ > UA *tïmï-ta ‘rock’
(604) Aramaic rə’emaan-aa / reemaan-aa ‘antelope-the’ > UA *tïmïna 'antelope’
(99) rakb-u ‘they mounted, climbed’ > UA *tï’pu / *tïppu ‘climb up’
(889) Aramaic rakbaa / rikbaa ‘upper millstone’ > UA *tïppa ‘mortar (and/or) pestle’

Loss of Semitic final -r, without effect on the preceding vowel:

(565) makar ‘sell’ > UA *maka ‘give, sell’
(616) dakar ‘male’ > UA *taka ‘man, person’
(550) Aramaic bəsár ‘flesh, penis’ > UA *pisa ‘penis’
(1331) ’ikkaar ‘plowman, tiller of ground’ > UA *wika 'digging stick'

Semitic initial voiceless pharyngeal ħ > UA *hu, or w/o/u, and non-initially ħ > w/o/u:

(672) ħbq ‘pass air, break wind’ > UA *hupak- ‘stink’ (*q > k)
(673) ħnk ‘train, dedicate’; ħanukkaa ‘dedication, consecration’ > Ca huneke ‘to take an Indian bath’;
Yq húnak-te ‘show, direct, raise (young)’
(671) ħmm ‘heat, bathe, wash’ > UA *huma ‘wash, bathe’
(1040) ħml ‘carry, lift, pick up’ > UA *homa ‘take, carry, pick up’

The Semitic voiced pharyngeal ʕ > UA w/o/u, that is, some form of rounding:

(677) ʕagol ‘round’ > UA *wakol 'round(ed)’
(676) paqʕ- ‘whiteness, species of fungus’ > UA *pakuwa ‘mushroom, fungus’ (*q > k)
(683) ʕmṭ ‘cloud over, become dark’ > UA *(w)umaC / *(w)ïmaC ‘rain, be cloudy / overcast’
(686) ʕɛrwaa ‘nakedness, genitals’ > UA *wowa ‘vulva, vagina’
(1197) Hebrew ʕaaqeeb ‘heel, footprint’ > UA *woki ‘track, footprint’ (*q > k)
(747) Aramaic / Syriac ṣibʕ- ‘finger’ > UA *sipwa ‘finger’
(876) dʕk, impfv: -dʕok (< *-dʕuku) ‘(fire) go out’ > UA *tuka / *tuku / *tuki 'fire go out, dark, black, night’
(900) nʕm ‘be lovely, good, beautiful’ > UA *numa / *noma ‘good, well, pretty’
(1289) šgʕ, Hebrew məšuggaʕ ‘raging, mad’ > Nahuatl šiikoaa ‘be jealous, angry’
(94)ršʕ ‘act wickedly, be guilty’ > UA *tasawa ‘be/do bad’

Many phonemes (sounds) remain much the same, such as t, k, p, m, n, etcetera:
(52) Hebrew mukkɛ ‘smitten’ > UA *mukki ‘die, be sick, smitten’
(769) *taqipa (sg), *taqipuu (pl) ‘overpower’ > UA *takipu ‘push’
(750) tmh ‘in awe, fear, speechless’, Syriac təmah > UA tuma’ / tu’mï / tehmat / tïhmï ‘be silent, afraid’
(755) Hebrew kutónet ‘shirt-like tunic’ > UA *kutun ‘shirt’
(754) Hebrew participle pone ‘turn to, look’ > UA *puni ‘turn, look, see’
(851) Hebrew panaa-w ‘face-his’ > UA *pana ‘cheek, face’
(852) pl construct paneeʸ- (< *panii) ‘face, surface of’ > UA *pani 'on, on surface of'
(1339) šippaa ‘make smooth’ > UA *sipa / *sippa ‘scrape, shave’
(56) šɛkɛm / šikm-, Samaritan šekam ‘shoulder’ > UA *sïka ‘shoulder, arm’, Numic *sikum ‘shoulder’
(57) *siggoob ‘squirrel’ > UA *sikkuC ‘squirrel’
(563) sapat ‘lip’ > UA *sapal ‘lip’
(879) šwy / šawaa ‘broil, roast’ > UA *sawa ‘boil, apply heat, melt'
(1138) Hebrew šor ‘navel’; Arabic surr ‘navel cord’ > Sr ṣuur ‘navel’
(13) snw ‘shine, be beautiful’ > Hopi soniwa ‘be beautiful, bright, brilliant, handsome’
(890) kann ‘shelter, house, nest’ > UA *kanni (NUA) ‘house’ > *kali (SUA) ‘house’
(903) khh, kehah ‘be inexpressive, disheartened’ > UA -kïhahï- ‘sad’
(1045) Hebrew *moškat ‘bracelet, fetter, belt’ > Tb mohkat-t ‘belt’
(1105) kali / kulyaa ‘kidney’ > UA *kali ‘kidney’
(1409) Aramaic kuuky-aa’ ‘spider-the’ > UA *kuukyaŋw ‘spider’; Hopi kòokyaŋw ‘spider’

Semitic emphatic or pharyngealized ṣ > s in UA:
(892) ṣanawbar ‘type of pine tree’ > UA Sh sanawap-pin ‘pine tree’
(901) ṣb’ / ṣby / ṣəbee ‘wish, want, seek, delight in’ > UA *supiC ‘like, want’
(1173) mwṣ ‘suck’ > UA *mos ‘suck’
(1350) ṣd’ / ṣdi ‘grow rusty’ > UA *sïta / *sïti ‘red’

Semitic emphatic or pharyngealized ṭ > c (ts):
(770) ṭwy / ṭawaa ‘spin (thread)’ > Nahuatl cawa ‘spin’
(771) ṭʕm ‘taste, eat’ (plural participle ṭoʕmiim) > UA *cu’mi 'suck, sip, kiss'
(772) ṭame’ ‘(be) unclean’, ṭum’a(t) ‘uncleanness, filthy mass’ > UA *co’ma ‘mucus, have a cold’
(832) *sarṭoon ‘scratcher, crab’ > UA *saCtun >*sicu/*suttu ‘claw, fingernail, crab, scratch’

Sometimes the c lenites (weakens) one more step to s:
(778) ṭibbuur ‘navel’ > NP sibudu; Cr sipu; Hp sipna / sivon- ‘navel’

Semitic-p distinguishes x from ḫ, as in pre-exilic Hebrew, thus Semitic *x > UA k:
(1088) *xld ‘burrow’, xuld / *xild-aa’ ‘mole-the’ > UA *kita ‘groundhog’
(630) *xole ‘be sick, hurting’ > > UA *koli ‘to hurt, be sick’
(631) xmr ‘to ferment’; *xamar ‘wine’; Arabic ximiir ‘drunkard’ > UA *kamaC 'drunk'
(632) *xnk ‘put around the neck’ > UA konaka ‘necklace, string of beads’
(634) *xaṣr- > xaṣṣ ‘hip, haunch, loins’ > UA kaca ‘hip’

Clusters like -m’-, -’m-, -qm-, that is, m with either ’ or q > ŋ in NUA:
(1246) Old Canaanite sim’al ‘left’, *ha-sim’al ‘the-left’ > Tb aašiŋan ‘left side’ (l > n in NUA)
(1012) šeqma(t) / šiqma(t) ‘sycamore tree’ > UA *sïŋŋa(C) 'cottonwood or aspen tree'
(1144) ’lm ‘be grieved’ > Hebrew ’almaanaa ‘widow’ > UA *o’mana / *oŋani ‘sad, suffering’

Clusters with -r- as 2nd consonant show -Cr- > -Cy-, especially -gr-, -qr- > -ky-, or -gra / -qra > Hopi -kya:
(1130) Aramaic pagr-aa ‘corpse-the’ > Hopi pïïkya ‘skin, fur’
(1403) Syriac šigr-aa ‘drain, ditch, gutter-the’ > Hopi sikya ‘small valley, ravine, canyon with sloped sides’
(1405) šqr ‘fair, yellow to red’, Arabic šuqra ‘fair complexion, blondness, redness’ > Hopi sikyà ‘yellow’
(743) *tamar; Aramaic tuumr-aa ‘palm tree-the’ > UA *tu’ya ‘palm tree, sp’

Proto-Semitic *z > c(ts) in UA:
(1116) Hebrew zɛ́pɛt (< *zipt-) / zaapet ‘pitch’ > UA *copï ‘pitch, resin’
(87) Arabic ʕgz / ʕagaza ‘to age, grow old (of women)’ > Tr wegaca- 'grow old (of women)’

Egyptian terms in UA exceed 400 and have the same sound correspondences as the above Semitic. Egyptian did not include written vowels, only the consonants. Sometimes the vowels are hinted at in transcriptions from other languages, or from Egyptian’s later forms in Demotic and Coptic, but generally only the consonants are certain. Sometimes the Coptic term is listed along with the Egyptian term, but do not regard Coptic as involved in the Egyptian-to-UA tie, because the Egyptian-to-UA sound correspondences differ from the Egyptian-to-Coptic correspondences. In fact, UA preserves the Egyptian phonology better than Coptic usually does, though UA is two more millennia removed. Coptic is simply listed for hints at vowels or to show Uto-Aztecan’s better preservation (7.6, p. 347):

Egyptian	Uto-Aztecan
(115) sbk / *subak ‘crocodile’	> UA *supak / *sipak ‘crocodile’ (b > p)
(116) -i ‘old perfective/stative verb suffix’	> UA -i ‘intransitive / past / passive/ stative verb suffix’
(117) -w / -iw ‘passive verb suffix’	> UA -wa / -iwa ‘passive verb suffix’
(124) tks ‘pierce’	> UA *tïkso ‘pierce, poke’
(125) km ‘black’	> UA *koma ‘dark, gray, brown, black’
(126) nmi ‘travel, traverse’	> UA *nïmi ‘walk around’
(129) wnš, pl wnšiw ‘jackal’	> UA *wancio / woncia ‘fox’
(131) šm ‘go, walk, set out, leave’	> UA *sima ‘go, leave’
(219) iqr ‘skillful, excellent, capable, intelligent’	> UA *yikar ‘knowing, intelligent, able, good’
(221) wr ‘great (in size/importance), wrw ‘greatest’	> UA *wïru ‘big’
(222) wnx ‘be clothed, roll of cloth’	> UA *wanaC ‘cloth, clothing’
(136) win ‘thrust aside, push away, set aside’	> UA *wina 'throw down/out, spill, empty'
(253) spd ‘sharp, be sharp pointed’	> UA *sipaC 'point'
(255) sqd ‘slope (of pyramid)’	> UA *sikiC ‘slanted (terrain), side’ (q > k)
(210) twt ‘sandal(s)’	> UA *tuti ‘sandal(s)’
(339) t′-ɧimat ‘the-wife’; Coptic hime	> UA *tïhima 'spouse'

Note again Egyptian b > UA p, as in Semitic-p above:

(132) sbq ‘calf of leg’	> UA *sipika ‘lower leg’	(b > p)
(133) sbty ‘enclosure’	> UA *sapti ‘fence of branches’	(b > p)
(134) qbb ‘cool; calm, quiet, cool breeze’	> UA *koppa ‘quiet, calm’	(b > p)
(137) bbyt ‘region of throat’	> UA *papi ‘larynx, throat, voice’	(b > p)
(138) bši ‘spit, vomit’, bšw ‘vomit, vomiting’	> UA *piso-(ta) ‘vomit’	(b > p)
(139) bnty ‘breast’	> UA *pitti / *piCti ‘breast’	(b > p)
(141) bit ‘bee’	> UA *pitV > *picV 'bee, wasp'	(b > p)
(142) bik ‘falcon’	> UA *pik ‘hawk species’	(b > p)
(154) sb’ ‘star’	> UA *sipo’ > *si’po ‘star’	(b > p)

Also Egyptian x > UA *k, as in Semitic-p above:

(170) txi ‘be drunk, drink deep’, txw ‘drunkard’	> UA *tïku 'drunk’
(294) xpš ‘foreleg, thigh’	> UA *kapsi ‘thigh’
(295) xpd ‘buttock’	> UA *kupta ‘buttocks’
(295) xpdw ‘buttocks’	> UA *kupitu ‘buttocks’
(171) sxn / zxn ‘kidney fat, pancreas’	> UA *sikun ‘kidney’
(174) sxt ‘field, country, pasture, willow’	> UA *sakat / *sakaC ‘grass, willow’
(178) x’yt / h̲’yt ‘disease, slaughter, corpse-heap’	> UA *ko’ya ‘die, pl subj; kill, pl obj’
(247) xr ‘fall’	> UA *kuri ‘fall’, UA *kara ‘fall’
(320) xpx ‘rob’	> UA *kïpïk ‘take, grasp’
(224) wxd ‘be painful, sick, suffer, endure’	> UA *okotï ‘be in pain, suffer, sorrow’
(452) xt ‘fire, heat’	> UA *kut ‘fire’

Egyptian initial pharyngeal ḫ > UA *hu, and non-initially ḫ > w/o/u:

(180) ḫbi ‘be / make festival’ > UA *hupiya ‘sing, song’
(181) ḫnqt ‘beer, drinkers’ > UA *hunaka ‘drunk, alcohol’
(182) ḫtp / hotpe ‘be gracious, peaceable, set (sun), bury’ > UA *huppi ‘peaceable, go down, sink, dive’
(187) ḫw’ ‘foul, putrid, stink, vi’ > UA *hu’a / *hu’i ‘break wind, stink’
(188) nḫbt ‘nape of the neck, yoke’ > UA *nohopi > nopi ‘hand, arm’
(189) nḫb ‘to harness, to yoke’ > UA *noopi ‘carry on back’
(397) ḫti ‘smoke, vapor’ > UA *uti 'dew, vapor, frost’
(415) ḫnn ‘penis’ > UA *huna ‘penis’

Egyptian glottal stop ’ > w, or glottal stop next to round vowels, ’ probably causing vowels to round (o, u):

(147) m’i ‘lion’; Coptic mui > UA *mawiya ‘mountain lion’
(148) t’yt ‘shroud’ > UA *tawayi ‘cape-like garment’
(198) ḏ’rt ‘bitter gourd’ > UA *sawara 'gourd'
(205) ṯ’y ‘male, man’ > UA *tawi > *tïwi ‘man, male’
(322) q’i ‘tall, high’; q’yt ‘high land, hill’ > UA *kawi ‘mountain, rock’
(515) ’xi ‘sweep together’ > UA *wak / *wok ‘sweep, comb, brush’
(150) t’ ‘earth, land’; Coptic to > UA *tïwa / *to’o ‘sand, dust’
(151) i’w ‘old man’; i’wi ‘be aged’ > UA *yo’o ‘old’
(153) s’ ‘son’ > UA *so’o ‘child, son’
(259) sṯ’ ‘jar, jug’ > UA *soto’i ‘jar’
(258) sṯ’ ‘drag, pull, pull out, draw’ > UA *(piC)-sutu’a ‘(behind)-pull, drag’
(154) sb’ ‘star’ > UA *sipo’ > *si’po ‘star’
(157) iṯ’ ‘take, carry, steal’ > UA *itu’i > i’tu ‘steal, take’
(370) ḫ’ ‘behind, around’ > UA *huwï ‘around’
(431) b’k / b’kt ‘document’ > UA *po’ok ‘mark, write, tattoo’ (b > p)

Egyptian ḏ corresponds to Semitic ṣ, and thus Egyptian ḏ > UA *s, like Semitic ṣ > UA *s also:

(200) ḏbt / *ḏubat ‘brick, adobe brick’ > UA *supa ‘adobe’
(199) ḏb’ ‘to clothe, garment, clothing’ > UA *sipu’ > *si’pu ‘slip, skirt, shirt, clothing’
(198) ḏ’rt ‘bitter gourd’ > UA *sawara 'gourd'
(197) ḏʕb ‘coal-black’, ḏʕbt ‘charcoal’ > UA *so’opa ‘black, dark’
(194) ḏ’i ‘pierce, transfix’ > UA *so’a/*so’i 'pierce, sew, shoot arrow'
(390) ḏwt ‘mosquito, gnat’ > UA *suti ‘mosquito, gnat’

Egyptian initial r- > UA t-, though Tarahumara retains r-:

(164) rn ‘young one, of animals’ > UA *tana ‘offspring’
(165) rwi ‘dance, v’ > UA *tawiya / *tuwiya > *tuya ‘dance’
(169) rmṯ ‘man, person’ > UA *tïmati ‘young man’: Tr ŕemarí, Eu temáci-
(167) rwḏ ‘cord, bow-string’ > UA *tïsa ‘rope’
(337) r’-ib ‘stomach’ > NUA *to’i ‘stomach’ / SUA *toCpa ‘stomach’

Egyptian pharyngeal ʕ > UA *w/o/u:

(163) rʕ / rʕw ‘sun’ > UA *tawa / *tawi ‘sun, day’
(162) šʕy ‘sand’; Coptic šoo > UA *siwa(l) ‘sand’
(262) ʕnt ‘nail, claw’ > UA *wati 'claw, fingernail'
(400) sʕr ‘thorn bush(es)’ > UA *sawaro ‘saguaro cactus'
(426) ʕnr(t) ‘flint’ > UA *wi’naC ‘flint’
(464) ʕq ‘enter’ > UA *waka/u ‘enter’
(475) sw ‘it, pronoun’ (is) p’ʕt ‘quail’ > UA *supa’awi ‘quail’

Like the devoicing of Egyptian b > UA *p, so also is the devoicing of Egyptian d > UA *t, and g > *k:
(268) dwn ‘stretch, straighten; Coptic town > UA *tuna ‘straight’
(269) dqr ‘fruit’ (> Coptic tiče / jiji) > UA *taka(C) ‘fruit’
(270) dbḫ ‘ask for’ (Coptic toobh) > UA *tïpiwa / *tïpiN 'ask'
(271) dm ‘be sharp, sharpen’; Coptic toom > UA *tama / *tomo ‘be sharp, sharpen’
(272) dmi (dmr) ‘touch’ > UA *tam ‘touch’
(273) dw’ ‘rise early’; dw’w / dw’yt ‘morning’; Coptic to’we > UA *to’i ‘rise, come up/out’
(395) ngg ‘gander/male goose’ > *nakï 'goose' (devoicing of g > k)

Egyptian cluster *-m’- > UA *-mw- > -ŋ- in three items widespread throughout Uto-Aztecan:
(280) ḫm’ / ḫm’t ‘salt’ (> Coptic hmu) > UA *omwa > *oŋwa / *oŋa ‘salt’
(281) sm’ ‘lung’; pl: sm’w ‘lungs’ > UA *somwo > *soŋo ‘lungs’
(284) qm’ ‘create, beget (of father)’ > UA *kumwa > *kuŋa ‘husband’ (q > k)

Other clusters and parallels:
(332) qrḫt ‘serpent, partner’ (*qarḫat >) > UA *koŋwa ‘snake, twin’ (q > k)
(384) inqt ‘net’ > UA *ikkaC / *iCkaC ‘carrying net’ (q > k)
(391) ishb ‘jackal, fox’ > UA *isap / *isa’apa ‘coyote’
(398) k’p ‘cover, close (eyebrows/eyelids) > UA *kuppa / *kuCpa 'close (eyes)’
(434) g’p ‘cut’ > UA *kappi 'break, cut' (devoicing g > k)
(381) wrt ḫq’w ‘buzzard’ > UA *wirhukuN 'buzzard, turkey vulture'
(404) ḫ’dt ‘basket’ > UA *huCta ‘basket’
(426) ʕnr(t) ‘flint’ > UA *wi’naC ‘flint’
(263) šwt ‘shade, shadow’ > Nahuatl seewal-li ‘shade’
(264) šmrt ‘large bow’, pl šmrwt > -samaaloo-t of Nahuatl koo-samaaloo-tl ‘rainbow’
(267) twr ‘reed’ > Nahuatl tool-in ‘cattails, reeds’;
(266) šnw / šni ‘hair, grass’; šni ‘encircle, cover’ > UA *soni / *sono ‘grass, blanket’
(331) qny ‘be yellow’; qnit ‘yellow(ness)’ > Cp kenekene’e- ‘yellow’ (q > k)
(333) qd ‘go round, turn, spin’ (> Coptic koote) > UA *koti / *kuri ‘turn, go around’ (q > k)
(446) qm’ ‘fight’; qm’tyw ‘enemies’ > UA *kïma’a / *kïmma(n)ci ‘different, enemy’ (q > k)
(409) nk ‘copulate’ > UA *naka ‘copulate, cover’
(468) ’wt ‘length’ > UA *otï / *utu / *uta ‘long, tall’
(470) t’-imnti ‘the west’ > UA *tïmïnïmïn ‘north, west’ (reduplicated)
(519) wpi ‘open, separate, divide’ > UA *wopa ‘divide’

The above 105 Egyptian-UA matches are but 25% of the 400+ listed in the Egyptian section.

The above Semitic and Egyptian parallels in UA both have the same sound correspondences, apparently spoken or used by the same group of people. However, in contrast to those two, a separate sizable set of data suggest another contributing Semitic element, with a different set of sound correspondences in which Semitic b > UA *kw, though the Tepiman branch of UA, and Eudeve, Opata and some Nahuatl dialects actually have b from Semitic b, all corresponding to presumed UA *kw. This Semitic-kw language is more Phoenician-like, while the Semitic-p language is more Aramaic-like, which differences are discussed periodically throughout the book. The data of the Semitic-kw language are what I noticed first, and because the Hebrew b > UA *p group were exceptions to the correspondences noticed first (Hebrew b > UA *kw), I ignored them for years, but kept them in the back of my mind (not a safe place), until I noticed Egyptian similarities (in UA) whose sound correspondences with UA aligned with those exceptions: that is, Egyptian b > UA *p also, as well as another 40 examples of Semitic b > UA *p. Not until then did it occur to me that we seem to have two separate Semitic entities that merged in UA—a Phoenician-like Semitic-kw (Sem-kw) wherein Semitic b > UA *kw, and an Aramaic-like Semitic-p (Sem-p) in which Semitic b > UA *p. Furthermore, the Sem-p speakers seemed to know some Egyptian as well; that is, the Sem-p and the Egyptian in UA have the same sound correspondences. The data show the two languages (Sem-kw and Sem-p) to have separate sets of correspondences for other phonemes (basic sounds) as well, the Sem-p being consistently parallel to the Egyptian correspondences.

Below are examples of data and sound correspondences from the Phoenician-like Semitic-kw wherein Semitic b > UA *kw:

(4) Hebrew baašel ‘boiled, cook, ripen’ > UA *kwasïC ‘cook, ripen’
(5) Hebrew bááśaar ‘flesh, penis’ > UA *kwasi ‘tail, penis, flesh’ (r > y/i)
(6) Hebrew baalaʕ ‘swallow’ > UA *kwïluC ‘swallow’
(7) Semitic *bahamat ‘back’ > UA *kwahami ‘back’
(24) bky / bakaa^{y} ‘cry’ > UA *kwïkï ‘cry’ (from Semitic-kw)
(19) barr- ‘land (as opposed to sea)’ > UA *kwiya / *kwira 'earth' (r > y/i)
(27) brm ‘worn out, weary, bored with’ > UA *kwiyam ‘be lazy, do lackadaisically’ (r > y/i)
(1457) Arabic ṣabba ‘pour, drip, overflow’ > UA *cikwa ‘rain
(11) Hebrew -dabber ‘speak’ > UA *tïkwi ‘say, talk, speak’ (r > y/i)
(26) Hebrew bɛn ‘son’; pl: bənee^{y} ‘children (of)’ > Nahuatl *konee 'child, offspring’:

As in the Egyptian and the Semitic-p contributions, so also in the Semitic-kw, ħ > hu or w/o/u:

(78) Hebrew ħeṣ ‘arrow’ > UA *huc ‘arrow’
(79) Hebrew ħmr ‘cover with, smear on’ > UA *humay ‘smear, spread, rub, paint’ (r > y/i)
(80) Hebrew ħbb ‘rub off, wash’ > UA *uppa ‘bathe, wash, rub’
(81) Hebrew ħabéret ‘wife’ > UA *hupi 'woman, wife' (r > y/i)
(82) Hebrew ħzy / ħazaa ‘see, behold, look’ > UA *husi / *h^{w}asi ‘look, peek at’
(658) ħbl ‘bind’, *-ħabbil ‘bind’ > NUA *wïkkwiN- ‘wrap around, coil’
(853) Aramaic ħippušit-aa ‘beetle-the’; Arabic *xunpusaa’ / xunpus > UA *wippusi ‘beetle’
In the next section are three more examples (83, 84, 85).

Semitic-kw ṣ > UA c (ts):

(83) Hebrew ṣrħ ‘cry, roar’ > UA *cayaw ‘yell’
(84) Hebrew ṣmħ, imperfective: yi-ṣmaħ ‘sprout’ > UA *icmo ‘sprout’
(85) Hebrew ṣlħ ‘rush, v’ > UA *coloa ‘flee, run’
(899) ṣinw-, pl aṣnaa’ ‘twin, one twin’ > UA *cono’o 'twin(s)'
(29) ṣəbii > ṣəvii ‘gazelle’ > Hopi cöövi- ‘antelope’
(86) ṣʕq ‘shout, call out, cry (out)’, ṣəʕaaqaa ‘yell, call, n’ > UA *coaka ‘cry’
(28) ṣurṣur ‘cricket’ > UA *corcor ‘cricket’
(78) ħeṣ ‘arrow’ > UA *huc ‘arrow’

As in all three languages, the voiced pharyngeal ʕ > w/o/u:

(88) ʕlq ‘stick, adhere’, ʕalaqat ‘leech’ > UA *walaka ‘snail’ (of similar slimy adhering texture)
(89) śeeʕaar ‘hair’; Arabic šaʕr / šaʕar ‘hair’ > UA *suwi ‘body hair’ (r > y/i)
(92) yáʕar ‘wood, forest, thicket’ > UA *yuwi / yuyi ‘evergreen species’ (r > y/i)

Unlike its associated rounding in Semitic-p, the Semitic-kw glottal stop ’ is not rounded and often lost:

(991) Hebrew ni-qra’ ‘he/it is called/named’ > UA *nihya ‘call, name’
(587) ’argaamaan ‘purple, red-purple’ > UA *aNkaC ‘red’
(1214) Hebrew mee-’ayn ‘from where?’ > Tb maa’ayn ‘where from’
(1055) ’aamaqqət-aa ‘lizard-the, n.f.’ > UA *makkaCta(Nka) ‘horned toad’
(591) ’adaamaa / ’^{a}daamaa ‘earth’ > UA *tïma ‘earth’
(592) Hebrew ’abneṭ, pl: ’abneṭ-iim ‘sash, girdle’ > UA *natti ‘belt’
(1054) raqbubit ‘moth, decayed, moth-eaten’ > UA *…kupïpika / *(C)Vkupïpika 'butterfly'

Non-initial -r- > Semitic-kw -y-, and tends to raise and front the preceding vowel (V > i):

(62) srq / saraq ‘to comb’ > UA *siyuk / *ciyuk ‘to comb’ (r > y/i)
(65) mrr ‘pass, go, walk’ > UA *miya ‘go’ (r > y/i)
(64) Semitic krr / krkr ‘go in circles, dance’ > SP kiya ‘have a round dance’ (r > y/i)
(19) barr- ‘land (as opposed to sea)’ > UA *kwiya / *kwira 'earth' (r > y/i)
(27) brm / baram ‘worn out, weary, bored with’ > UA *kwiyam ‘be lazy, do lackadaisically’ (r > y/i)
(79) Hebrew ħmr ‘cover with, smear on’ > UA *humay ‘smear, spread, rub, paint’ (r > y/i)
(81) Hebrew ħabéret ‘wife’ > UA *hupi 'woman, wife' (r > y/i)

Final or non-initial -l in Semitic-kw tends to raise and front vowels (V > e, i):
(1225) Hebrew ’abaal ‘truly, indeed’ > Tr abe ‘yes, an emphatic’
(54) Hebrew taapel ‘whitewash’; Aramaic ṭəpel ‘plaster’ > UA *tïpi ‘white clay’
(1321) Hebrew ħargol, Arabic *ħargal / *ħurgul ‘locust’ > Tr urugi-pari ‘type of grasshopper’
(798) Hebrew ’akal ‘(he/it) ate’ (perfective) > UA *’aki ‘open mouth, eat, take/put into one’s mouth’
(797) Hebrew *yo’kal ‘(he/it) eats’ (imperfective) > UA *yï’ïki 'swallow, taste, finish'
Number 797 (-l raising -a- > -i-) is in contrast to Semitic-p *tukkaC wherein final -l has no raising effect.
(796) Hebrew *to’kal ‘(she/it) eats’ > UA *tukkaC > Num *tïkkaC 'eat'

Such a tripartite combination I first considered suspect until the quantity for each grew to more than sufficient to allow each to stand on its own strength, as each dimension has 400-700 sets. Should we ignore the strength of a case of 1500 similarities? Or should we be fair and consider the data when a few hundred items support each dimension of the tripartite scenario? If one simply cannot bear the thought of the three, then pick only one of the groups, any one of which yields 400 to 700 items. Ought a correlation of 400 sets be ignored? Even 400 sets is three times what most Native American language families were founded on.

Admittedly, this may sound incredible at this point, as truth often does at first, but working through the data will diminish doubt. So read with an open mind and consider the quantity and quality of the evidence. Perhaps this first edition contains enough loose ends to serve as some consolation for those who do not like the idea of such possibilities. In fact, **several words of caution** are in order:

(1) First of all, linguists would look dimly on a tripartite collection of languages to propose an Old World tie with a Native American language family. Linguistically, each of those three has to stand on its own merit, independent of the other two. Yet the numbers of similarities for each are enough data for each one of the three to do exactly that—serve as a valid case each in and of itself (400 to 700 similarities for each).

(2) Anthropologists and linguists are wary and weary of hearing about proposed ties between Semitic or Egyptian and New World languages—about 300 years’ worth of weary. Most such claims have been bogus to borderline or amateurish at best, somewhat justifying linguists’ wariness in light of claims void of sound methodology, that is, lacking what linguists have found to be established principles and patterns for verifying language relatedness: rules of sound change that create consistent sound correspondences, hundreds of vocabulary matches consistent with those sound correspondences, and some grammatical and morphological alignments, which sum constitutes **the comparative method**. Thus, the language similarities in this work are presented within such a framework of sound correspondences, etc. In fact, the Semitic or Egyptian forms proposed to underlie the UA forms often answer questions and explain puzzles in UA that Uto-Aztecanists have not yet been able to explain; and explanatory power is a cherished quest among linguists. While the finds do seem significant, some details remain to be worked out.

(3) Given the amount of Egyptian vocabulary in UA, we might expect to find and may yet identify more Egyptian grammatical patterns in UA. However, if the Egyptian phrasing in UA is reduced as much as many Egyptian phrases are reduced in Coptic (a late form of Egyptian dating to 2,000 years ago), then such identifications would be a challenge (if even possible), requiring time, not to mention requiring a greater depth of familiarity with UA languages and Egyptian than yet exists in any single mind. Many living languages reduce as drastically. In American English, one often hears ‘hwəjədu?’ for ‘what did you do?’ Therein -j- is the phonological reduction of the final -t of ‘what’, the whole of ‘did’, and the y- of ‘you’—some of three words (-t did y-) reduced to one consonant (-j-).

Often as drastic was the change from Egyptian to Coptic: Egyptian *iwr-ti* became Coptic εετ (eet) ‘pregnant’ (Loprieno 1995, 78); the i/y is not obvious, nor anything w- or r-like; so practically nothing of the stem ‘pregnant’ (iwr) is left, only a long vowel and the t of the stative suffix. Egyptian *r-di.t iri.f sḏm* became Coptic *e-t-ref sotem* ‘to cause that he may do hearing’—a reduction of eight consonants (*r-di.t iri.f*) to (*etref*) three consonants and two vowels (Cerny and Groll 1993, 155), though three of the original eight consonants are vowel-like or semi-vowels. Egyptian *tw.i m nʕy r sḏm* ‘I am in going to hear’ (= I shall hear) became Coptic *tinasotm*, or *tw.i m nʕy r > tina* (Cerny 1976, 104), eight segments (sounds) to four. Adding to the challenge is that the time depth between Late Egyptian and Coptic is half the probable time depth in this problem: if UA is partially from Egyptian, the Egyptian in the UA languages is now being recorded at a time depth a millennium or two greater than the time depth between Late Egyptian and Coptic. Yet UA preserves many vowels and details better than Coptic does (see 7.6).

On the other hand, **these data explain many things previously unexplained in UA:**

(1) The phonology of medial (middle) consonant clusters is a huge problem in UA itself, and Semitic and Egyptian shed light on many of those clusters and help explain the mutual effect of adjacent consonants on each other. See 7.2 on consonant clusters.

(2) Regarding PUA *p, Uto-Aztecanists agree on each UA language's reflex that corresponds to PUA *p. (A language's reflex is its corresponding sound which the proto-sound changed to.) However, five UA languages—Tarahumara, Mayo, Yaqui, Arizona Yaqui, and Eudeve—show both initial b and p corresponding to PUA *p. This split is usually ignored as an inconvenient inconsistency in these languages. However, the initial b forms in these languages correspond to Egyptian b or Semitic b of Semitic-p, and the initial p forms in these languages to Semitic/Egyptian p. How can such an alignment be coincidental? For the various UA forms of b vs. p to match Semitic/Egyptian b vs. p is significant. (See 6.2)

(3) PUA initial *t (at the beginning of words) corresponds to the initial t of most UA languages, except for Tarahumara initial r. So if PUA *t became Tarahumara r, then where does Tarahumara initial t come from? The data in this work suggest that Semitic/Egyptian initial r became t, so in most UA languages initial r and initial t merged to look like PUA *t, but Tarahumara kept them separate. Thus, 6.1 clarifies the Tarahumara r vs. t puzzle, which see.

(4) Other matters in 6.3, 6.4, 6.5, 6.6, and 6.7 are also explained by these language ties.

Significant is the language parallel of Yiddish, the language of the Jewish peoples of Central Europe. Uto-Aztecan and Yiddish are both Semitic infusions into non-Semitic areas, where each (as a minority people) borrowed heavily from the languages of the larger surrounding peoples. Originally coming out of Palestine, many Jews sojourned in Greece, Rome, and elsewhere along the northern Mediterranean, then some among them expanded into central Europe, where their original Hebrew-and-Aramaic idiom borrowed mostly from German, but also from Slavic and other languages of their successive environments through which they traveled and periodically settled (Kriwaczek 2006, 40-48; Harshaw 1990, 5-7). Thus, Yiddish is a transplant and very much a language mix (like English and many languages are). Estimates generally have 15-20% of Yiddish being from the original Hebrew-Aramaic vocabulary, and 80-85% borrowed from German, etc. Similarly, only 15% of Old English continued into modern English; the other 85% was lost, being replaced by words from French, Latin, and other languages from which we English speakers borrowed (Baugh and Cable 55). While the details of Uto-Aztecan's prehistory may yet require lifetimes to unlock, Uto-Aztecan seems to have a higher percentage of its basic vocabulary from Near-Eastern languages than Yiddish has. For example, Yiddish pronouns are all from German, whereas most UA pronouns match Semitic (see section 3 on pronouns). Most Yiddish body-part terms are from German—kop (head), oig (eye), oi'er (ear), hant (hand), hartz (heart), k'nee (knee), fus (foot), etcetera—while a higher percentage of UA body-part terms, animal terms, and basic nouns of nature match Semitic or Egyptian (see section 7.4).

The two forms of Semitic are both Northwest Semitic, though often quite distinguishable, but not always. Two separate sets of sound correspondences distinguish most of the vocabulary as noted previously, but not all. The exact nature of each remains to be clarified. While Semitic-kw exhibits Phoenician-Hebrew like features and Semitic-p has Aramaic-like features and vocabulary, it also has Hebrew-like features. These kinds of unique sets of features are typical of related languages. For example, the language of the Book of Job is unique: though labeled Hebrew, it contains features more Arabic-like and Aramaic-like than the Hebrew of the other authors. The language of the Nabateans, though primarily an Aramaic dialect, was also more Arabic-like than other Aramaic dialects. So any diffused offshoot can be expected to be a unique combination of features.

Regarding the Aramaic leaning of the Semitic-p, some scholars (Young 1993, 54-62, 85-86) note that Aramaic did influence the dialects of ancient Israel, especially northern Israel. What is not known is the degree or extent, though it may have been more significant or pervasive than presently known. These data may be relevant to that void in present knowledge. Marsha White (1997), in a review of Young 1993, summarizes Young's substance more clearly and concisely than either I or Young could: "Young … suggests that Biblical Hebrew goes back to the adaptation of the pre-Israelite Canaanite prestige language…. Thus, from the beginning of Israelite history there were two linguistic strata: literary/formal and dialectical/colloquial. This situation of diglossia persisted throughout pre-exilic Israelite history…. The best explanation for … so many Aramaisms in the early literary language is that they were in the lower (i.e., spoken) form of the language, and that Archaic Biblical Hebrew was open to elements from the underlying

dialects. The strong presence of Aramaisms in the oldest Biblical Hebrew undermines the theory that Aramaisms equals late" (White 1997).

This all aligns well with the likelihood of Aramaic substrata serving as underlying dialects to the literary language of Canaanite/Hebrew, perhaps throughout the Northern Kingdom's centuries. What language did the mothers of the Israelites (Leah and Rachel) speak? Aramaic! In addition, Aramaic was somewhat a lingua franca throughout most of the area through most centuries. So did the Israelites really set aside Aramaic upon entering Canaan? Or did they adopt degrees of bilingualism while adding the Phoenician/Canaanite literary language? The latter is likely nearer the case in some areas, if not most. Yet many UA features match reconstructable Hebrew/Phoenician better than they match other Semitic languages:

	Uto-Aztecan		Hebrew	Arabic	Aramaic	Akkadian
(1)	***-ima** (pl suffix)	Semitic masc pl:	***-iima**	-uuna/-iina	-iin	-uu
(904)	***-te** (pl suffix)	Semitic fem pl:	***-ooteey**	-aat	-aat	-aat
(2)	***na-**	reciprocal/passive:	***na-**	in-	--	
(3)	***yasipa**	'sit / dwell'	***yašiba**	waθaba	yəθeb	

The UA basic vocabulary in this work are numerous: body parts, plant and animal terms, nouns of nature (sun, moon, star, sky, rock, water, etc (see 7.4). A considerable amount of Semitic morphology or fossilized parts of Semitic verb conjugations are found in UA. Below are three groups.

(1420) Semitic nwr 'to make/become light' with infinitive and imperfective: -nuur(u), and perfective naar: UA has both in Eu nurú 'to dawn, become light' and Tbr nare 'to dawn, become light'.

Uto-Aztecan has four separate forms from the verb bky /bakaa 'to cry, weep':
(559) Semitic-p bky/ bakaa 'he cried, wept'; Syriac bakaa / baka' > UA *paka' 'cry'
(24) Semitic-kw bky/ bakaa 'he cried, wept'; Hebrew baakaa > UA *kwïkï / *o'kï 'cry'
Because bilabials as first segment in a cluster disappear (-bk- > -k-), the imperfective 3rd person masculine singular *ya-bkV 'he/it weeps' with imperfective prefix originally *ya- (later yi-) also matches UA *yakka
(560) Semitic *ya-bka^{y} 'he/it weeps, cries, m.sg.' > UA *yaCkaC > *yakka / *yaka 'cry'
(561) Semitic *ta-bka^{y} 'she/it weeps, cries, f.sg.' > UA *takka > NP taka 'cry'.

So Northern Paiute has both the masculine 3rd sg of *ya-bka > yakka and the feminine 3rd singular *ta-bka > UA *takka 'cry' (and geminates/doubles the middle consonant in both as well), and also has the perfective stem in UA *paka' of Semitic-p and also *kwïkï/*o'kï of Semitic-kw.

Uto-Aztecan also has three separate forms from the Semitic root ktš 'grind': the imperfective verb stem in most languages, a perfective qittel in Yaqui, and a noun 'grindstone' in most languages:

	Hebrew root ktš 'grind'	UA
(1094)	impfv -ktoš (< *-ktusu) 'pound, grind'	*tusu 'grind' with loss of 1st C in a cluster
(615)	*kitteš (< *kittaš) 'grind'	Yq kitte / kittasu 'grind'
(614)	makteš 'mortar, grinding stone'	*ma'ta 'mortar, grinding stone' and Ca mataš

Of interest is the denominalized verb Ca mataš 'crush, squash, vt' showing final -š and a medial cluster or geminated *-tt-.

In addition, many unusual semantic combinations in Semitic and Egyptian are preserved in the corresponding UA sets. Besides the examples below, many more are at 7.5.

(283) Eg qm' 'create' and 'mourn' > UA 'make, create' and 'mourn'
(332) Egyptian qrḫt 'serpent', Egyptian qrḫ 'friend, partner' > UA/Nahuatl koŋwa 'snake, twin'
(406) Egyptian b' 'ram, soul' > UA *pa'a 'mountain sheep, all living beings'
(98) Hebrew rqʕ 'stamp, beat out (metal), spread out'; Hebrew raaqii^{a}ʕ 'extended surface, expanse, sky'
> UA *tukuN- in * tukuN-pa 'sky' and 'metal' in the Takic languages.
(994) Ls qáya/i- 'blow down (a tree)' (which is the same result as 'uproot')
and Ls qáya/i- 'heal' are listed as separate verbs in the Luiseño dictionary, though phonologically identical, yet the corresponding Syriac verb ʕqr also means both 'uproot' and 'heal' (ʕəqar or -ʕqar > qayV).

Stress in UA prehistory is a complex issue, which the data in this work may have some potential to help clarify. In *Uto-Aztecan: A Comparative Vocabulary*, I wrote "In the reconstructions I do not deal with vowel length, only vowel quality and consonants. Figuring out PUA vowel length may fill another lifetime, but not mine. Reduced consonant clusters with compensatory vowel lengthening underlie some long vowels in UA, raising doubts about vowel length until the medial clusters are clarified. That and changing stress patterns—causing vowel lengthening with stress, or shortening or syncope without stress, in the various branches and languages through the layers of time—make the puzzle of PUA vowel-length quite unappealing to me, if not presently impractical" (page 1). Likewise in this work, only vowel quality, but not vowel length, is represented in the UA reconstructions, though I will say the following about stress.

Proto-Semitic *bas$_2$ar 'flesh' > Hebrew bááśaar 'flesh, penis'; Aramaic bəśár 'flesh'; Arabic bašar. Note that in UA the originally stressed vowels retain their quality, while the unstressed vowels do their typical unstressed schwa-like behavior, which in UA is V > ï or i. Hebrew's stress on the first syllable shows Semitic-kw (Hebrew/Phoenician) bááśaar 'flesh, penis' > UA *kwasi 'tail, penis' (5); and Aramaic's stress on the 2nd syllable has Semitic-p (Aramaic-like) bəśár > UA pisa 'penis' (550). In both cases the originally stressed á remains a, but unstressed a > i in both cases, regardless the present or intervening stress patterns of the various languages' reflexes. See also Hopi in 174, and stress-related details in 611, 933, 1015, 1056.

Works establishing language relationships often include only matches of reconstructible forms with *identical* meanings and later are added matches of probable, but less than identical meanings. However, (1) I cannot assume the luxury of such a lifespan; and (2) am tired of writing huge, detailed reference works after 30 years of doing so; and (3) I care not to exclude probabilities to be added later in yet another huge detailed reference work. So, if the reader prefers, (s)he can toss the 100 or so of less than identical meanings, and consider only the other 1400 matches. However, I include from the start what I consider reasonable, and will leave it to coming generations to do whatever debating and sorting they think best. Nevertheless, I do identify those sets with [iddddua] meaning 'if desired, delay differing definitions until acceptance'.

Nevertheless, the less-than-identical semantic inclusions have changed meaning in understandable ways: (734) Hebrew mə-ṣuudat 'net, prey' i.e., game > UA *masat / *masot 'deer';
(720) Hebrew nebɛl 'skin-bottle, skin' in the common phrase of Hebrew nebɛl yayin 'skin of wine';
Syriac nbl / nʼbl > Classical Nahuatl noʼpal-li 'prickly pear' often used to make alcoholic beverage;
(675) Hebrew ɧnp 'limp'; Arabic ɧnp 'have distorted foot, be curved, pigeon-toed, walk bow-legged with toes inward' (like turtles, badgers, and bears) > UA *hunap- 'badger, bear'; Arabic uses this stem for 'tortoise' and 'chameleon' while the UA match is 'badger' and 'bear' all having similar turned-in feet;
(724) Semitic parʕoš 'flea (jumper)' (< Semitic verb prʕš 'jump') > UA *parʼosi / *paroʼosi 'jackrabbit'; the jackrabbit, like the flea, is also a jumper, and in UA *paroʼosi 'jackrabbit' we see all 4 consonants and 2 identical vowels in two of the most extraordinary jumpers of the animal kingdom.

I express thanks and admiration for many fellow Uto-Aztecanists. Beyond founders of comparative UA, like Edward Sapir, Kroeber, Whorf, Hale, the Voegelins, and Wick Miller, several contemporaries continue. Alexis Manaster-Ramer (AMR) through the 1980s and 1990s published several illuminating insights that I am not sure anyone else would have figured out. Manaster-Ramer (and Bright 1993) noticed consonant clusters, like the -p- in *kapsi 'thigh' (294 Egyptian xpš 'thigh') that everyone else had missed for a half century of reconstructing *kasi. He noticed many final consonants, like -R- in *yakaR 'nose, ridge' (1279 Aramaic *yagar 'hill'). His figuring out *tw > kw (1991d, 1992d, 1993a) is also impressive, and *-c- > NUA -y- (1992a), etc. As Serrano (Sr) may best preserve PUA phonology, we are indebted to Kenneth C. Hill (KCH) for his founding works in Sr (grammar and dictionary); his noticing Sr's "pharyngealized and retroflex" vowels is impressive. White Mesa Ute (WMU) also has strong pharyngealization. His noting the pharyngealized vowels or rounding with retroflex in Sr (not as apparent in other UA languages) is regularly significant to Semitic pharyngealization. Ken Hill also revised and added to Miller's huge 1988 work. Other major contributors to comparative UA include Jane Hill, Pamela Munro, Jeffrey Heath, David Shaul, Jason Haugen, William Merrill, Karen Dakin, Zarina Estrada Fernández, Lyle Campbell, Ronald Langacker, Andrés Lionnet, Terrence Kaufman, Jose Luis Moctezuma Zamarron, and Catherine Fowler. Ronald Langacker (1976b, 1977a) and Jason Haugen (2008) have also authored excellent books on UA grammar. The above and other linguists, too many to mention, have contributed dictionaries, grammars, and articles on individual UA languages. Many linguists in Mexico continue to add valuable documentation to UA languages in Mexico. Knowing the arduous load of life-long linguistic labors, I laud all the above and many other researchers (see bibliography) with deepest respect for their many valuable contributions.

1.1 Some Basics of Linguistics (Language Science)

1.11 Language Families and Similarities by Coincidence, Contact, or Descent

A language family is a group of related languages, descended from the same parent language. The parent language may be a well known language like Latin whose descendants are Spanish, Portuguese, French, Italian, and others, or it may be an ancient proto-language, unknown except as reconstructed by linguists. Knowing how languages and sounds typically change, linguists can examine a group of related languages descended from a common parent language and reconstruct many words and features of that ancient parent language, though unknown and unwritten. Such a hypothesized parent language is called a proto-language. Thus, Proto-Uto-Aztecan (PUA) is the hypothesized ancient parent language of the approximately 30 Uto-Aztecan languages. Likewise, the parent language of most European languages and of several Asian languages that have been demonstrated to be related is called Proto-Indo-European. The first step is to demonstrate relatedness, thoroughly treated in Campbell and Poser, 2008.

When two languages have similar words with similar meanings, those similarities can be due to (1) chance / coincidence, (2) contact—that is, neighboring languages usually borrow words from each other, which borrowings are called loanwords—or (3) common descent from a common source or parent language.

Coincidence: When randomly comparing any two languages, chances are that 1% or even 3% of their vocabularies can yield chance similarities. The shorter the words and the fewer the number of sounds, the higher is the probability of chance similarities. For example, 15 consonants (C) and 5 vowels (V) may yield 75 CV patterns (15 x 5; C = any consonant; V = any vowel) or 1,125 CVC patterns (15 x 5 x 15) or 5,625 CVCV patterns. When comparing the basic vocabularies of say 2000 words in two languages with short morphemes (parts with meaning) of CVC length and limited phonological inventories (number of sounds), two matches by coincidence are likely. When adding those with "kind of similar" sounds, like b and p, or d, t, and r to count as matches, then 20 or so (1%) are likely. Languages with longer words and more sounds provide lower percentages of probability for chance similarities; nevertheless, any two languages can and usually do have some similarities by coincidence.

Contact: the number of loanwords between neighboring languages depends on how long they are neighbors, the people's attitudes toward their neighbors and their languages, political dominance, and such things. For example, even though English belongs to the Germanic branch (sub-language family) of Indo-European (the larger language family), the words on a page of written English are typically about half loans—many from Latin, when Latin was the Medieval language of academia and English was not allowed in the schools, and even more from French, when the Norman French ruled England for three centuries, and some from Greek and other languages.

Cognates are the related words in related languages, as those words descended from the same proto-form or original ancient word. Related languages yield several of these descended sets of related words, and each set of related words is called a **cognate set**, a set of related words descended from the same proto-word.

All living (spoken) languages are always changing. Though slow, the change is inevitable. After a population separates, the languages of the separated groups gradually change. Some meanings change, some features of grammar change, and some words lose sounds and/or change other sounds, and some words are replaced. In spite of the inevitable change, linguists have found that in related words the sounds change in consistent ways. For example, Proto-Indo-European (IE) *p remained p in Latin and Greek, but consistently changed to f in Germanic. When a number of words or cognate sets exemplify each sound change with a consistent pattern of sound change, with few exceptions, that pattern sets up what is called a **sound correspondence**: that is, Germanic f corresponds to Greek p, or IE *p > Greek p (> means 'became' or 'changed to'), also IE *p > Latin p, and IE *p > Germanic f. Likewise, IE *k > Greek k, > Latin k, > Germanic h. That is, because sounds do not change randomly, but in consistent patterns, the same sound will change the same way in the same language in the same phonological environment (environment of surrounding sounds). When two languages exhibit a decent percentage (say 10% or more) or a sizable number (say 100 or more) of their respective vocabularies to be similar in meaning and to establish a consistent system of **sound correspondences**, usually amounting to dozens or hundreds of relatable words, then the chance of such a sizable correlation of similarities happening by chance is zero, and the two languages or that group of those languages' similarities are deemed due to descent from a common origin.

Another way of saying "correspond to" is that Germanic f **reflects** (corresponds to) IE *p, or that f is the Germanic **reflex** of IE *p. A **reflex** can be a corresponding sound or a corresponding word: so father is the English **reflex** (cognate) of IE *pater, and f is the English **reflex** (sound correspondence) of IE *p.

Some Indo-European Cognate Sets and Sound Correspondences

English	hound	water	thou	daughter	tooth	heart	foot	father	knee	two	three
German	hund	wasser	du	tochter	zahn	herz	fuss	vater	knee	zwei	drei
Greek	kuon	hudor	su	thugater	dont-	kardia	pod	pater	gonu	duo	treis
Latin	kanis		tu		dent-	kord-	ped-	pater	genu	duo	tres
Sanskrit	śvan	udakam	tuvam	duhitar	dant-		pad	pitar	janu	duva	trayas
Hittite	--	watar	tuk	--		kart	pata		kenu	twi	tri

(Cambell 1999, 137-41; Beekes 1995, 208)

An asterisk (*) marks a hypothetical original or earlier form as reconstructed by linguists, an unattested form that the attested descendant forms derived from. One can see above in the cognate sets for 'foot' and 'father' that an original Indo-European *p consistently changed to f in English; and an original *t changed to th, as in 'thou,' 'tooth,' and 'three'; and Indo-European *k > h in the Germanic languages as is apparent in words for 'hound' and 'heart'. However, Indo-European *p, *t, *k remained p, t, and k in Latin; so the results of those sound changes provides a set of sound corredpondences between Latin and English:

Proto-Indo-European	*p	*t	*k
Latin	p	t	k
English	f	th	h

Similarly, for every pair or group of related languages, a system or set of sound correspondences will emerge. One might also notice a larger pattern—that the stops (p, t, k) generally became their corresponding fricatives (f, th, h)—such that all three patterns or systems constitute a larger pattern or system: stops > fricatives (for stops and fricatives, see pp. 18-19). Such multi-tiered patterns and systems of systems are typical of language change. And because linguists have found sound correspondences or consistent sound change to be a principle between related languages, they require that in order to prove a genetic or common-descent relationship between languages, one must establish the sound correspondences, as well as some grammatical or morphological similarities.

The lexical (word) comparisons between Semitic and Uto-Aztecan, as well as between Egyptian and UA, yield a consistent set of sound correspondences, as consistent as has been established for many other language families and a little more consistent than occurs within UA itself, as these ties explain many of the medial consonant clusters that have remained mostly mysterious to Uto-Aztecanists to date. Nevertheless, all language families yield a few apparent exceptions, though for some, an explanation is found later.

Glottochronology is the study of the **rates of language change**, or more specifically, rates of word retention (words kept) vs. replacement (words lost by substitution) over time. Two languages recently separated would still have a great majority of their words in common. For example, the recent separation (ca. 700 years ago) of the Apachean branch of Athapaskan has Navajo and the Apache languages generally retaining 93% or more of their vocabulary in common. In contrast, the Indo-European languages separated several millennia ago and share much smaller percentages of vocabulary, though enough to assure their relatedness. However, linguists find that **rates of language change** are subject to many variables, most of all the **type and intensity of contact** with other languages. For example, Icelandic, isolated in the Atlantic, did not change from its Old Norse ancestor as fast as Norwegian did in being more subject to other close and neighboring European languages.

Comparative size of neighboring languages matters. The Native American languages in the U.S. are tremendously outnumbered; thus, many became moribund (nearly dead) in two or three generations. Consider languages spoken by immigrant families: German, Dutch, and Italian immigrants to the United States may or may not learn English; their children are often bilingual, knowing their parents' language and the more prevalent language English; however, their grandchildren are often monolingual speakers of English, who may or may not understand what their immigrant grandparents say. **Political or cultural dominance** of a language may allow the language of a minority to have more influence than expected. The Norman French conquered England in 1066; though fewer in number, their political dominance in Middle English brought more French into English than the 15% of Old English that survived into modern English.

1.12 Morphology (Word Formation) and Syntax (Word Order)

A **morpheme** is a unit of meaning, and **morphology** is the study of how morphemes combine to form words or larger units of meaning. Just as a phoneme is a segment of sound or the smallest unit of sound (consonant or vowel), a **morpheme** is the smallest unit of meaning. For example, typical morphemes in English are cat, mouse, -ness, -ful, -less, un-, dis-, and -er, in words such as use-ful, use-less, use-ful-ness, dis-heart-en-ed, un-settle-ed, un-fruit-ful, and wash-er. Morphemes can be undividable words, prefixes, or suffixes. Prefixes and suffixes are both affixes that can be combined to the front or back of a stem respectively. Irresistable contains four morphemes. Re-sist literally means 'stand back' or in order of occurrence 'back-stand'. With the suffix -able added, re-sist-able means one is 'able to stand back or stay away from something'. The Latin prefix in- (meaning not) **assimilates** or changes to **ir-** before words beginning with r. So ir-re-sist-able breaks down to not-back-stand-able. Likewise, irrevocable means not-back-call-able or not able to call back.

Some morphemes or rules for morpheme combining are **productive** and some are not. A process or phenomenon in language that still happens readily is said to be **productive**, that is, it still produces new forms. If a previous language rule is no longer in effect, but the results of the once existent rule are apparent, then those resultant forms are **fossilized** forms. For example, prefixing *with-* 'against' to verb forms was once a productive rule in older English, but no longer is; nevertheless, we have a number of fossilized forms resulting from that once existent rule: withstand; withhold; withdraw.

By '**rule**' linguists mean a mechanism of language usage that native speakers use to structure their language, whether consciously aware of it or not. In fact, most of what native speakers know about how they create language is subconscious knowledge. They are not even aware of most of the rules that they use to create language. For example, consider the following misuses:

*Her saw he.
*After them beat we in tennis, us treated they to dinner.
*The tracks were hard for I to see, but me followed they until him appeared and scared I to death.

These are simple reversals of subject vs. object pronoun forms, yet most five-year-old preschoolers do not make such mistakes. At the very beginnings of learning a language, a two or three-year-old toddler may say something like "me want a cookie," but usually by four or five, their subconscious minds have figured out what the subject forms are, what the object forms are, where the subject slots are, and where the object slots are, and get it all 95% right without any formal education. About 4th grade the formal instruction begins and is repeated for eight consecutive years until they arrive in college, where I ask them what the grammatical subject is of a sentence on the board, and a handful know consciously. So by age 20, their conscious minds cannot remember how to identify the subject after several years of teaching their conscious minds, yet their subconscious minds knew by age five before they even started school and never forgot. For several other examples of subconscious language knowledge see "The Subconscious Mind's Role in Language Acquisition" in *Morsels for the Mind* (Stubbs 2009) and "The Language Instinct" (Steven Pinker 1995).

Besides common vocabulary revealing consistent systems of sound correspondences, related languages normally have some similar patterns of morphology or share morphological correspondences as well. A Germanic characteristic that disappeared from English shortly after the Middle English period was **conjugated verb forms**. These were still productive ('alive and well') in the early seventeenth century when the King James scholars translated the Bible. Note how similar the conjugated verb forms of earlier English are to those of German:

I	bind	ich	binde
thou	bindest	du	bindest
he	bindeth	er	bindet

Verb conjugation patterns are part of a language's morphology, but sometimes tend to be simplified over time and often eliminated, as they were in English. Something similar might be expected to happen to Navajo over the coming decades. The conjugation patterns of Navajo verbs are more complex than any Indo-European language. That complexity and Navajo's extensive contact with English combine to make

such a simplification likely. In fact, I have heard that in some areas or among some younger speakers, such simplifications are already underway. The Semitic languages also have specific verb conjugation morphology, which is no longer productive in UA, but have left hundreds of fossilized forms in UA.

For another example of shared morphology in the larger Indo-European language family, note the similarity of the primary verb endings in Sanskrit, Hittite, Greek, Latin, and Gothic, an East Germanic dialect of about A.D. 900 (Beekes 1995, 232):

	Sanskrit	Hittite	Greek	Latin	Gothic (Germanic)
I (verb)	-mi	-mi	-mi	-m	-m
You (verb)	-si	-si	-si	-s	-s
He (verbs)	-ti	-ci-	-ti	-t	-t

The conjugation of the IE verb *be* also shows morphological correspondences (Campbell 1995, 318):

	Sanskrit	Hittite	Greek	Latin	Gothic	Enlgish
I am	asmi	—	eimés	sum	im	am
He is	ásti	estsi	estí	est	ist	is
They are	sánti	asantsi	eisí	sunt	sind	Spanish: son

The second row (he is) is the source of English *is* (from Germanic *ist*) and Spanish *es* (from Latin *est*). We can also see in that same line of forms that final sounds are progressively left off over time. The older languages have the longer forms.

Syntax refers to the order of words and morphemes. An example is the **basic word order** of main parts of a sentence. The basic word order of English is subject-verb-object (SVO). Other languages have very different word orders. Consider these parallel sentences in five languages:

English: The tall man ate a red apple with a knife.
Spanish: El hombre alto comió una mansana roja con (un) cuchillo.
Navajo: hastiin néz bilasáana łichí'igii beesh yee yiyíiyáá'
man tall apple red knife with it-he-ate
White Mesa Ute: pa'átïm ta'wáč aká-ġar apïs tïkkái wiíč-Im
tall man red apple ate knife-with
Hebrew: 'akal ha-'iiš hag-gaboah 'et hat-tappuax ha-'adom bə-sakkiin
ate the-man the-tall the-apple the-red with-knife

In contrast to the word order of English (SVO), the word order of Navajo is subject-object-verb (SOV), and Hebrew is usually (VSO), but can be any order, and Aramaic is often verb-final (SOV). Besides basic order of verbs, subjects and objects (SVO, SOV, VSO), some languages put adjectives before nouns, like English and Ute, while others put adjectives after nouns, like Spanish, Navajo, and Hebrew.

Interestingly, VO languages generally have **prepositions**, as do English, Spanish, German, Hebrew, Arabic, and Samoan, while verb-final languages (OV) generally have **postpositions** as do Navajo, Ute, and many Native American languages. The preposition vs. postposition phenomenon relates to OV vs. VO word order, in that these relating words often connect verbs and their objects, thus coming between them. So we frequently see verb-preposition-object in SVO languages, and object-postposition-verb in SOV languages.

Like Old English, German, Navajo, Semitic, Spanish, and many Indo-European languages, conjugated verbs are part of the morphology of many languages. In UA we see many fossilized remnants of the Semitic verb conjugations, though not any full or productive systems of Semitic conjugations. For example, from the Hebrew root ktš 'pound (in a mortar), grind' are three very differently shaped items:

Hebrew	UA
impfv -ktoš (< *ktusu) 'pound, grind'	*tusu 'grind' with loss of 1st C in a cluster (1094)
unattested *kitteš (< *kittaš) 'grind'	Yq kitte / kittasu 'grind' (615)
makteš 'mortar, grinding stone'	*ma'ta 'mortar, grinding stone' (614)
	especially *mattas > Ca mataš 'to crush, squash, vt'

1.13 Historical Linguistics and the Comparative Method

The science of linguistics has various branches. Applied linguistics applies linguistic insights to facilitate second language learning; theoretical linguistics deals with competing theories of grammar and explores how the mind creates language; socio linguistics focuses on how language usage varies in various social contexts. Historical linguistics deals with the histories of languages or how languages change over time. Thus, language relatedness and studies in language families and how the related languages have changed from the original or proto-language all belong to the realm of historical linguistics, also called diachronic linguistics. **Synchronic** has to do with one-time (syn 'one' + chron 'time'); so a synchronic view of a language is a snapshot of it as a cohesive entity at one point in time. **Diachronic** refers to two different times on a spectrum, or comparing the changes in a language from this time to that time. Some features of language can be explained synchronically as the language exists at any given point; other features are better understood diachronically wherein some history of the language clarifies matters. As historical linguists compare related languages and map the changes of the various languages over time, their work is necessarily diachronic in nature. Their systematic comparisons that establish languages as related in a language family are called **the comparative method**.

The comparative method consists of (1) establishing a system of sound correspondences for (2) a sizable quantity of vocabulary; (3) identifying morphological parallels, and to lesser degrees, (4) similarities in syntax and (5) unusual semantic combinations. Syntax is limited in possible options—OV vs. VO, noun-adjective vs. adjective-noun, etc—and syntax can change quickly. Thus, categories (4) and (5) are less applicable than the first three. Yet the Egyptian and Semitic in UA provide numerous examples in all categories except (4) as syntactic options are simply not numerous, whatever the language.

The strength of the comparative method was impressively demonstrated in the discovery of Hittite. Based on evidences in the IE languages known at the time, a Swiss linguist named Ferdinand de Saussure in 1879 **reconstructed** certain laryngeals (guttural-like consonants) in the proto-forms of some IE words. (A **reconstruction** of a **proto form** is what linguists theorize the original form of a word to have been in the **proto-language** or the ancient parent language from which the later known languages are descended.) In other words, he theorized that those laryngeal consonants had existed in some original IE words even though those sounds did not clearly exist in any of the daughter languages known at the time. In 1906, the capital of the ancient Hittite Empire was discovered. In 1915, Hrozny, a Polish linguist, deciphered the Hittite language inscribed on thousands of clay tablets, and Hittite was found to be an IE language. (The Hittite word for water is watar and knee is kenu.) Not only was Hittite found to be an IE language, but Hittite contained the laryngeals that Saussure, by the comparative method, had predicted decades earlier as being in the original Proto-Indo-European language (Beekes 1995, 101-2; The New Encyclopedia Britannica 1997, 608).

Besides establishing language families, the comparative method helps to discern branches within a language family and to trace details of language change. One can imagine that an ancient unified people did not separate into 30 different groups at once, but at first there may have been a two- or three-way split, then some time later additional split-offs occurred, and so forth—thus, the creation of **branches** within a language family. For example, the Germanic branch of IE consists of English, German, Dutch, Icelandic, and most Scandinavian languages, except Finnish. The Germanic languages are more closely related to each other than they are to the other IE languages. The Italic or Latin branch of IE consists of Spanish, French, Italian, Portuguese, and others. Many languages of India are descendants of Sanskrit as the Indic branch.

Branches are often identified by **shared innovations** or shared retentions. A shared innovation is a new change that a branch shares among the branch languages, but not with the other languages of the family. For example, an innovation of the Germanic branch is that the voiceless stop series (p, t, k) became fricatives (f, θ, h). Shared innovations in UA are that O'odham, Pima, and the Tepehuan languages of the Tepiman branch all have **g** corresponding to ***w** of the rest of UA, and **d** corresponding to ***y** of the rest of the family. When a branch of languages all share a feature or qualtiy that the rest of the language family does not have, then it follows that that group of languages developed that feature after leaving the main body of the language family, but before splitting into the various languages of that branch.

Along with all the niceties and usual consistencies revealed by the comparative method, a few inconsistencies, exceptions, and unresolved difficulties plague most languge families. As Salmons (2012, 111) notes in *A History of German*, "we expect, as we saw earlier, for sound change to be regular, but we find messiness in real historic data." Sometimes a subset of irregularities are later explained by a special

phonological environment or some other explanation that moves them from the "exception" pile to the "explained" pile, but such discoveries take time and only if a mind sufficiently insightful to see what no one has seen before happens along to reduce what remains mysterious. For example, after Jacob Grimm (1822) published the first Germanic sound shift, a group of unsettling exceptions continued ruining the aspired order, until Karl Verner (1877) figured out the explanation for some of the exceptions … but more than a *half century* later! May the progress of this work be granted equally spacious leniency! Yet an army of linguists works on Indo-European versus the sole soul in the proposed language tie of this work.

1.14 Phonology: Sounds, Sound Change, and Sound Correspondences

Phonology is **the study of sounds in language, their changes and effects on each other**. An understanding of phonology clarifies many mysteries about language. Our mouths produce consonant sounds by affecting the airflow in primarily three ways: the voicing vs. voiceless option, the manner of restricting the airflow, and the place in the mouth where that restriction happens. Thus, consonants are categorized by three features: voicing, place of articulation (contact in mouth parts), and manner of articulation:

Voicing can be perceived by putting fingers on both sides of the "Adam's apple" and saying a slow elongated *aaasssaaa*. Because all English vowels are voiced, one can feel the vocal cords vibrate while saying the voiced vowels *aaa….aaa*, but the vibration or voicing stops in the middle while saying the long voiceless *…sss…*; in contrast, when saying *aaazzzaaa*, the vibration never stops, because z is voiced. One can feel the vibration while saying **voiced** consonants (z, j, b, v, d, g, m, n), but there is no vibration, that is, no voicing while saying **voiceless** consonants (s, š/sh, č/ch, f, p, t, k).

Sounds are also classified by the **place of articulation** or the place where the airflow is most restricted. **Bilabial**s (p, b, m, f, v) are pronounced with the two lips. English f and v are actually pronounced with the top teeth and lower lip, but are close to bilabials. **Dentals** touch the tip of the tongue at or between the teeth (θ as in think, đ as in there). For **alveolar**s the tongue touches the alveolar ridge—the hard ridge behind the upper teeth (t, d, s, z, n). To do **palatals**, the tongue curves close to the soft palate curving behind and up from the harder alveolar ridge (š, ž, č, j). **Velar**s put the back of the tongue against the back of the roof of the mouth (k, g). **Uvular**s (q) are further down the back of the throat from velars. We do not have uvulars in English, but Arabic uvular q vs. velar k are apparent in Arabic qalb 'heart' vs. kalb 'dog'. **Pharyngeal**s, such as the voiceless and voiced pharyngeal fricatives of Arabic are articulated at the pharynx, even further down the back of the throat than uvulars.

The **manner of articulation** is a third feature of consonant sounds. For **stops**, the airflow is stopped (p, b, t, d, k, g). For **fricative**s, the airflow is not stopped, but produces friction at the greatest restriction in the vocal tract (s, z, f, v). An **affricate** is a combination of stop plus fricative (c or ts = t + s; č/ch = t + š/sh as in kitchen), that is, it starts as a stop but quickly releases into a fricative: so t and ts(c) and s are the voiceless alveolar stop, affricate, and fricative. In contrast, d, dz, and z are the voiced alveolar stop, affricate, and fricative. For **nasal**s, the airflow passes through the nose while the oral tract is closed at the lips (m), the alveolar ridge (n), or at the velum for the velar nasal (ŋ as in sing) with the back of the tongue in a position for saying k. The **liquid**s are l and r in English. The **glides** are y and w, slight closures of the vocal tract in the same positions in which the vowels i and u are pronounced; thus, they are also called semi-vowels. A simplified consonant chart follows:

Consonants

		bilabial	dental	alveolar	palatal	velar	uvular	pharyngeal	glottal
stops	voiceless	p		t		k	q		'
	voiced	b		d		g			
fricatives	voiceless	f	θ	s	š(sh)	x		ħ	h
	voiced	v	đ	z	ž(zh)	ġ		ʕ	
affricates	voiceless			c(ts)	č(ch)				
	voiced				ǧ/j				
nasals		m		n	ñ	ŋ			
liquids				l, r					
glides		w			y				

The phonetic description of a consonant consists of voicing, place of articulation, and manner of articulation—in that order. Therefore, p is a voiceless bilabial stop; g is a voiced velar stop; s is a voiceless alveolar fricative; j is a voiced palatal affricate; etc. All nasals, vowels, liquids, and glides are voiced in English, but not necessarily in other languages. For example, Ute has some voiceless vowels and Navajo has both a voiced l and voiceless ɬ.

We mentioned earlier the larger pattern that the IE voiceless stops (p, t, k) became voiceless fricativces (f, θ, h) in Germanic. We also mentioned the sound changes in Tepiman of Proto-Uto-Aztecan (PUA) *y > d, and PUA *w > g. As a larger pattern, the UA glides (y, w) became voiced stops (d, g) in the Tepiman branch, doing contact at the roof of the mouth where the glides come closest (w has lip rounding in front, but like u, the back of the tongue comes close to the velum where g is pronounced).

In Semitic exists a series of pharyngealized consonants. Besides the actual pharyngeals ʕ and ħ, described below, Semitic also has the emphatics or pharyngealized ṭ and ṣ. In contrast to a regular t, the pharyngealized ṭ of Semitic is pronounced with the tongue sounding as if retroflex, mainly because the back of the tongue is approximating the pharyngeal position, which affects the vowels, darkly coloring them and drawing them to the back, as in Arabic.

Sounds not discussed below are pronounced (more or less) like English:

ṣ Biblical Hebrew pharyngealized or emphatic ṣ (ṣade) is here symbolized with ṣ. The Hebrew ṣ became c (ts) in the Hebrew Semitic-kw of UA and in modern Hebrew, but became s in Semitic-p. UA ṣ is said to be retroflex.

d Egyptian d corresponds to Hebrew ṣ, and both Egyptian d and Hebrew ṣ of Semitic-p became or correspond to UA *s, though often Coptic t.

š is the sh sound of English ‘shave’ and ‘dish’; the š of Hebrew also corresponds to UA s.

c represents ‘ts’ as in ‘hits’.

č is the ch sound of ‘chop’, an allophonic variant of PUA *c (ts) above.

’ represents the Semitic aleph or glottal stop, as in English ə’o (uh oh) ‘woops’ and ə’ə ‘no’; the glottal stop also became w/o/u in UA (and became w in Arabic sometimes as well), and sometimes both a glottal stop and w (-’w- or -w’-), or round vowels adjacent to ’: o’o/u’i.

ʕ represents the Semitic ʕ (called ʕayn), a voiced pharyngeal fricative, not in European languages; it occurs twice in SaʕudiʕArabia; it has become a form of rounding (w,o,u) in UA, which is a natural change. In fact, evidence suggests that the pharyngeal ʕ was associated with rounding in Phoenician also (see page 56)

ħ is a voiceless pharyngeal fricative, a very guttural h (often transcribed as ḥ) not found in European languages; at the beginning of a word it became hu/ho in UA. Like the other pharyngeal (ʕ), ħ also became w/o/u, a form of rounding, in non-initial positions. Interestingly, x and ħ merged and became the same sound in Hebrew between 300 BC and Christ’s time when they both became ħ, but were different before 300 B.C. (Kutscher 1982, 13-18; Sáenz-Badillos 1993, 81). They are still separate in Arabic. The Semitic-p in UA shows the pre-300 BC distinction: the pharyngeal ħ appears as rounded forms, while the velar x remains k-like.

x is a voiceless velar (or uvular in Semitic) fricative or soft k, as in German nacht; x became *k in UA generally.

r of both Hebrew and Egyptian changed to UA *t at the beginning of a word. When not beginning a word, r remained r in some UA languages, but changed to y/i more often in Sem-kw; r > y/i is also common in languages world wide. Interchanges between r and l are also common in the Near East and in UA. In fact, Egyptian had only r that represented both the l and r of Coptic.

b of Hebrew became UA *kw (in dageshed positions: word initial or geminated/doubled) in the Semitic-kw contribution, but became UA *p in Semitic-p’s contribution to UA.

b, d, g devoiced and became p, t, k generally, another common change in languages world wide, since p is the voiceless counter-part of b, t of d, and k of g.

ṭ of Semitic is a pharyngealized or emphatic t, in which the tongue is rather retroflex or the back of the tongue approximates a pharyngeal.

ŋ is a velar nasal, the ng sound in sing.

t of Egyptian, i.e., the underlined t was originally different from t, but not for very long, since even in Egyptian, and consistently in UA, Egyptian t merged with and became t in UA (and in Egyptian).

-C is an unknown consonant that causes gemination or doubling of the next consonant. In UA, -C means a final feature (an underyling consonant) that doubles the next consonant, another common feature in many languages: like coC/com ‘with/together’ + labor > collaborate; com ‘with’ + sonare ‘sound’ > consonant.

Vowels are defined by the tongue's relative position to the roof of the mouth in a high-to-low, front-to-back grid: one can feel the tongue's blade near the top and front of the mouth when saying high-front i.

	front	central	back
high	i	ï	u
	I		U
mid	e	ə	o
	ɛ		
low	æ	a	

Thus, *i* is a high front vowel; *o* is a mid back rounded vowel; *a* is the low central vowel; *u* is a high back rounded vowel; *ï* is a high central vowel not found in English, but is common in Ute, Hopi, and many Native American languages. The vowel symbols have the following values: the i in machine, I in sit, e in they, ɛ in set/pet, æ pat/sat (for each one the jaw drops lower though they are all pronounced in the front of the mouth. In the middle are *ə* in rut, *a* in saw. At the back are u in blue, U in book/hood, o in goal/bowl/sole/soul. For those knowing Spanish, pronounce the 5 main vowels like Spanish, which is the original Latin pronunciation.

Vowel shifts happen in many language families. English changed the original Latin vowel values, some of them in a **vowel shift**, shifting the vowels clockwise: o > a (as in top), a > æ/e (tap/tape), e > i (keep). Uto-Aztecan also does some vowel shifts. For example, Cora (Cr) and Huichol (Wc) shifted some Proto-Uto-Aztecan vowels counter-clockwise: PUA *u > ï, PUA *o > u. Classical Nahuatl (CN) shifted *u one more slot: PUA *u > ï > i. So in CN, PUA *u and PUA *i merged (became the same sound) to CN i, so that CN i can be from either PUA *i or *u.

It is also worth noting that i and y are largely equivalent, perhaps a difference in length and/or intensity, but produced with the tongue in the same position. Say aaaiiiaaa slowly, then aia faster each time, and soon it sounds like aya. Likewise, aauuaa speeded up to aua a few times begins to sound like awa. So w and u are essentially the same sound, just as i and y are.

The English **plural suffix** -s exhibits three forms: -s, -z, -əz. A subconscious rule predicts when each of the three occurs. The rule is that (1) final voiceless sounds take voiceless -s: tops, pots, cakes; (2) final voiced sounds take voiced -z: tabs, pods, rags, rams, cans, laws, seas; and (3) final sounds similar to the -s (alveolar and palatal fricatives and affricates) require the intervening schwa vowel ə to separate the two similar sounds; otherwise, how would we make kiss plural—by adding a third *s* and pronouncing the three s's (kisss) as a real long *sss* sound? Examples of -əz include kisses, wishes, witches, judges, quizzes. The reason that the last has the form -əz instead of -əs is because vowels are voiced in English, so the sound before the s/z is the vowel ə, a voiced sound which results in voiced z.

The same rule applies to **possessives** of the form **apostrophe plus s** (-'s): Kate's hair, the rope's strength, the cake's frosting (-s); but Bob's book, Brad's cat, the dog's house, Tom's house, the car's door, Celinda's sorrel (-z); and for the sibilants (s and č-like sounds): Mitch's cat, the mouse's hole (-əz). **Third person singular present** tense verb forms also require suffixed -s, which also abide by the same rules: he stops, licks, writes, and laughs (-s); but she sobs, swigs, hides, loves, runs, hurls, sees, and believes (-z); and he wishes, she kisses, he squeezes (-əz), and they live happily ever after.

This shows that systematic patterns govern most of what happens in language. All three suffixed -s morphemes in English obey the same phonological rules and are entirely predictable according to specific patterns known only subconsciously by most speakers. Indeed, every language is a system of systems.

A similar rule governs whether the -ed suffixed to past tense regular verbs takes on a sound like -d, -t, or -əd. When the end of the word is voiceless, the -ed becomes voiceless -t: hopped, baked, missed (mist). When the end of the word is voiced, the -ed remains voiced -d: grabbed, hugged, freed, judged, called, crammed, bulged. When the word ends with a sound articulated (pronounced) at the same place as d (-d or -t), it requires an intervening vowel to sound like -əd: roasted, plodded, plotted, and greeted.

1.15 Sound Changes and How Sounds Change

Assimilation is often the force encouraging sound change. Sounds change, but in natural ways, which are usually explainable and are seen repeatedly in language families around the world. Assimilation is when one sound becomes 'similar to' another in some way. In fact, the word *assimilation* itself is from Latin ad 'to' + similis 'like', but when combined, ad-simil…> assimilate, because the -d- when next to -s- becomes

-s- also, becoming similar-to the s by becoming another s. Very often doubled letters in English are from two different sounds next to each other wherein usually the first becomes like the second, precisely because it is next to it. For example, the Latin prefix *in-* 'not' remained *in-* for indecent, insufficient, and incomplete, but the alveolar nasal (n) of *in-* changed to a bilabial nasal (m) when next to bilabial p in imperfect and impossible (n > m/_p; that means n changes to m before p), becoming similar to the bilabial. The *in-* prefix was entirely assimilated before l and r, merely doubling the following consonants as in illegal, illegible (n > l/_l), irregular, and irreverent (n > r/_r). Similarly, Aramaic *'illaa* 'if not, except, unless' derives from Aramaic *'in* 'if' + *laa* 'not': 'in-laa > 'illaa 'if-not'.

Similarly, Latin com- 'with' assimilates the m to the point of articulation (place of pronunciation) of the next consonant when compounded (put together with another morpheme): a couple of examples are com 'with' + sonare 'sound' > consonant 'with sound' (m > n/_s, because n, like s, is an alveolar); and com 'with/together' + laborare 'work' > collaborate 'work with/together'.

Similarly in UA, a nasal as first consonant of a cluster often assimilates to the second consonant of that cluster (linguists use N to represent any nasal or a general nasal), so
*-Nk- > -ŋk- (the nasal N becomes velar nasal ŋ, assimilating to the velar stop k);
*-Np- > -mp- (the nasal becomes bilabial nasal m, assimilating to the bilabial stop p);
*-Nt- > -nt- (the nasal becomes alveolar nasal n, assimilating to the alveolar stop t);

The above examples show that adjacent sounds tend to affect each other, that is, assimilate to each other or become similar to each other in some way or in all ways. Another example occurs in Semitic. In Arabic qatala 'he killed' and Hebrew qaaṭal 'he killed', this cognate pair has a discrepancy in two different kinds of non-corresponding t's: a regular t and the emphatic or pharyngealized ṭ. Both languages have both, but what happened is that in certain conjugations, such as the prefix/imperfective conjugation the q and t are adjacent or next to each other: Arabic ya-qtulu, Hebrew yi-qṭol. The q and ṭ are similar in being pharyngealized deep-throated, more guttural sounds, so as they came into contact with each other, the original -qt- cluster (as we see in Arabic) assimilated to become -qṭ- in Hebrew, and thus Hebrew changed an original -t- > -ṭ- due to assimilation in the frequent clustering of -qt-.

In the above examples, we see that the environment surrounding a sound is what often triggers (causes) a sound to change. In linguistic lingo **C** means any consonant or an unknown consonant, and **V** is any or unknown vowel. Word and morpheme structures can thus be represented as CVC, CVCV, CVCCV, etc. When a consonant is between two vowels (VCV) it is said to be intervocalic, inter- 'between' vocal- 'vowel'. Two consonants together (VCCV) are called a consonant cluster (see more on clusters below).

Vowels may also assimilate or become similar to adjacent consonants—wa > wo—and similar to vowels on the other side of consonants: suka > saka. Vowels assimilate to consonants quite often in UA. For example, Semitic baraq 'lightning' > Mayo berok 'lightning' changes the 1st vowel from a > e, raising and fronting to the place of contact of r in anticipating r. Likewise, the 2nd vowel changes from a > o, moving to the mid-back vowel o, closer to where the uvular q is pronounced in anticipating it. Another instance of the uvular q changing a vowel to a back round vowel is Semitic daqal 'kind of palm tree' > UA *taku 'palm tree'. In Semitic-kw especially, liquids l and r tend to raise the vowels before them or the vowels which are anticipating them (Semitic basar > UA *kwasi 'tail'), whereas Semitic-p does not (Aramaic bəsar > UA *pisa 'penis'; Aramaic dakar > UA *taka 'man').

A vowel may also partially assimilate to preceding or following vowels: suka > soka. One may notice on the vowel chart that *o* (mid back round vowel) is halfway from *u* (high back round vowel) to *a* (low central vowel), so a change in a vowel sequence of u-a > o-a is partial assimilation. Or two vowels may level each other in a compromise—u-a > o-o; a-i > e-e—where both vowels assimilate toward each other, becoming the vowel between the two. (See the vowel chart on page 12 and notice that *o* is between *u* and *a*; and *e* is between *a* and *i*.)

Consonant harmony is when one consonant becomes like another, though separated by vowels. Consonant harmony happens often enough in Uto-Aztecan: for example, Hebrew 'ari 'lion' > UA *wari > Tubar wawi 'mountain lion'. Other examples of consonant harmony are the three Tr variants—Tr ɾata-góbutu/ ɾata-gógutu / ɾata-bobutu 'have a fever'—and (853) Arabic *xunpusaa' / xunpus 'beetle'; Aramaic ƀippuušiit 'beetle, n.f.' > UA ***wippusa > *pippusi** 'stink beetle': Ch wiposat '13-line beetle'; Mn pipóísi/piboisi 'stink beetle'; NP pipuzi 'stink beetle'; Sh pippusi 'stink beetle'. Ch reflects the original initial consonant (w), from which the others harmonized the 1st consonant to the 2nd consonant (w-p > p-p). In addition, the UA vowels too are identical to Aramaic *-i-u-i.

Palatalization is also very common in Uto-Aztecan and in languages worldwide. For example, the alveolar t often becomes palatalized to č (ch) or c (ts) before high vowels and especially high front vowels i or e, during which the tongue is close to the palate (t > č or t > c/_i). Latin *-nate* of innate keeps its *-t-* sound, but in nation, with a following *i*, it palatalizes to *-š-*. Similarly in irritate and irritation, rotate and rotation, dictate and dictation. In Uto-Aztecan, any high vowel—i, ï, u (see top line of vowel chart, p. 20)—causes palatalization of t > č or t > c in some UA languages.

Many sound changes, if not most, are due to what might be called laziness or changes toward easier pronunciation. Assimilations make differing sounds more similar and therefore easier to pronounce, so making pronunciation easier could be viewed as laziness. An example is a change from contact to approximation or near contact, but not quite. The flap r, which involves the tongue's contact with the alveolar ridge, sometimes changes to almost contact or to y/i. The liquids becoming y/i (r > y/i; l > y/i) happens often enough. In English creoles, Dickerton (1981, 61) lists three English creoles in which 'for' became fo, fi, and foe. In Italian, many l > i, as in blanco > bianco. Lyle Campbell (1977, 97-100) shows Proto-Mayan *r > y in several Mayan languages. Also Hebrew r > UA y/i in Semitic-kw. German -r and British English and some Northeast U.S. dialects say -r as a vowel approaching the place of -r contact in a high vowel, though not quite as front as y/i, almost the high-central vowel ï of UA: German hier [hiï]; English better [bettï]. Likewise, Semitic l became y in some Ethiopic languages due to Cushitic influence (Kapeliuk 2002, 311). Other examples of change from contact to approximation are the intervocalic stops becoming fricatives: -b- > -v-, -k- > -x-, -t- > -θ-.

Another frequent change toward the easier is the change of the low vowel *a* > *ə*, because the mid-central vowel (ə) does not require the mouth to open as wide as is necessary for the lower vowel (a). In fact, any vowel *V* > *ə*, as mid-central ə is probably the easiest vowel to pronounce, because it is in the middle both directions, between high and low, and between front and back (also called the schwa vowel, the schwa in dud, sun). A prolonged utterance of əəəəəəəə does not make one sound very smart because it approximates what might come out when one is asleep with the mouth slightly open during a voiced exhale: əəəəəəəə.

Vowel centralization is, in fact, common in many languages, and involves (usually) unstressed vowels becoming centralized. One can see in the vowel chart that the vowel ə, is the mid-central vowel, the most central of all vowels, and that is exactly the vowel that most unaccented vowels become in English words of 3 or more syllables. Consider photograph and photography.

phótográph > fotəgræf
photógraphy > fətagrəfi

In phótográph the 1st and 3rd vowels are stressed and thus keep their more-or-less original values o and æ, but the unstressed 2nd vowel changes from o > ə. However, adding another syllable (-y) changes the stress pattern so that the 2nd and 4th vowels are stressed and keep their values, while the 1st and 3rd vowels both become unstressed and both become ə. Similarly, some UA languages tend to centralize unaccented vowels to UA's most central vowel ï, or sometimes to i, as i also does the stressless schwa role in UA too.

A **hyphen** signifies that something else exists in the direction of the hyphen. The prefix in- 'not, opposite' has a hyphen where the other morpheme follows. The English plural suffix -s has a hyphen on the front side to show that it comes at the end of the noun, with the word in front of it. Intervocalic consonants (between-vowel consonants) may be depicted as -r- because vowels are on both sides of it.

Lenition is a weakening of a consonant or partial loss of its definite qualities. Lenition often affects consonants between vowels. The sequence apa > aba has voiceless p becoming voiced b, because the vowels on both sides are voiced, which helped the intervening voiceless p become voiced b; likewise, aka > aga and ata > ada. These kinds of changes happened in UA and happened in the participles' change from Latin -atus > Spanish -ado. These changes are also an assimilation: the voiceless stops became voiced stops similar to the voiced vowels around them. Another common intervocalic change is frication of a stop, changing a stop to a fricative. It happened to the intervocalic Hebrew stops: -b- > -v-, -d- > -đ- (as in the), -g- > -ġ-, -p- > -f-, -t- > -θ- (as in thin), -k- > -x-. In UA, the intervocalic environment caused changes that included both frication and voicing of the originally voiceless stops, that is, voiceless stop -p- > -v-, a voiced fricative, and *aka > aġa, and *ata > ara, changing t to a Spanish flap r. Between vowels, a natural pattern of sound change is for voiceless stops to become voiced, then the voiced stops become fricatives, then the voiced fricatives disappear. The last step happened in the change from Latin to Spanish: Latin credere > creer 'believe' of Spanish, Latin legere > leer 'read'. Also Latin ego > eo > yo 'I' because e is close to i/y.

Occasionally changes go the other way, from less intense to more intense. For example, while v > w is frequent enough, the change of w > v also occurs. In Hebrew, w came to be pronounced v in some Hebrew dialects and thus in Modern Hebrew also. The name of Adam's wife Eve was originally Hewa; thus, w > v. The English name Eva at least keeps the vowels, Eve even lost the pronunciation of the last vowel as well. I have also heard some Arabic speakers pronounce Arabic w as v. Also in UA is evidence for some *w > v, to be discussed later.

Loss of sounds over time is also frequent, especially at the beginnings and ends of words or morphemes, like the initial k and final silent e of knife, both of which used to be pronounced. All the silent e's when found at the ends of English words used to be pronounced, but they became silent or lost, though still written. Similarly, at the beginnings of words, the h in honor, hour, herb, and all initial-h words in Spanish, like hablar, hermano, etcetera, all became silent. Loss of final sounds happens in Semitic languages too. Arabic 'akala 'he ate' and Hebrew 'aakal 'he ate' show the loss of a final vowel in Hebrew. In fact, Hebrew lost most short final vowels of an earlier *-iima > -iim 'Hebrew plural suffix'; *ta-ktušu > ti-ktoš 'she pounds/grinds in a mortar'; etc. Hebrew also lost final consonants sometimes. Arabic 'akalat 'she ate' and Hebrew 'aklaa 'she ate' show loss of final t in Hebrew and loss of the middle vowel. Arabic reflects Proto-Semitic better than other Semitic languages in most ways.

Consonant clusters (groups of consonants clustered without vowels between them) may also tend to be reduced to one consonant, such as the loss of the gh sound in the cluster of -ght- in English daughter vs. German tochter (both pronounced) and Greek thugater (consonants separated, not clustered), and the loss of gh/k in night and Spanish noche vs. German nacht and Latin nokt-. We no longer pronounce the -gh- in night, but we still say the -k- in nocturnal, as an English loan from Latin. Examples of consonant loss in cluster reductions in UA include Hebrew makteš 'grinding stone' > UA *ma'ta 'grinding stone'. Many UA languages have intervocalic *-p- > -v-. That happens in Hopi, the Numic languages, and others. So when we see a -p- between vowels, it is due to an underlying consonant cluster being reduced to -p-, but showing -p- (instead of -v-) because of -Cp- or the cluster strengthening the -p-: Egyptian **ḫotpe** 'peace' > Hopi **hopi** 'peace, peaceable' at (183); otherwise, *hopi > hovi. Also Aramaic **ḫippušit** 'beetle, n.f.' > UA ***wippusi** 'stink beetle' (853). The Arabic cognate **xunpus** shows a consonant cluster *-np- which always doubles the 2nd consonant in Hebrew and Aramaic (-pp-): Proto-Semitic/Arabic *-nC- > -CC-; thus, Semitic *xunpus / ḫippušit > UA *wippusi is a lengthy (6-segment) match. The -p- in Ch means original *-pp- in UA, and the vowels are identical to Aramaic *-i-u-i (853).

Relative to consonant clusters, the phonology (patterns of pronunciation) of some languages do not allow clusters. For example, 'Merry Christmas' in traditional Hawaiian is 'meli kalikimaka' because Polynesian languages do not normally allow consonants to cluster, and so the kr- and -tm- clusters of **Chr**is**tm**as are separated by vowels in the Hawaiian expression. Spanish does allow clusters, but has limits on initial clustering possibilities. For example, Spanish 'creer' starts with a cluster kr-, but English 'study' and Spanish 'estudiar' show that English allows initial st- clusters, while Spanish traditionally has not. One may also hear native Spanish speakers say a helping vowel before an initial -st- cluster, like 'estreet'. In the English word 'strengths' [strɛŋθs], one vowel amidst six consonants separates two clusters of three consonants each, which shows that English has an unusual tolerance for almost intolerable clustering compared to many languages. However, the loss of initial k- in English 'knee', 'know', and 'knife' means that even cluster-tolerant English has difficulty with initial kn-. We have no trouble with the same cluster between vowels (sickness, blackness), but initial kn- is more problematic.

Some languages' phonology systems prevent speakers from ending a word with a consonant or with certain consonants. In the merger of the Semitic-p and the Semitic-kw in UA, one or both may have developed a phonology that had all or most words ending with a vowel, because UA adds a vowel to many Semitic forms that would otherwise be consonant final. Yet that is one among many matters for future study.

Consonant clusters often lose the first consonant, sometimes doubling the second. We have already seen examples in English in-legal > illegal, in-responsible > irresponsible. Originally and in written English, debt has a consonant cluster, but the first consonant became silent and only the 2nd is pronounced. Liquids (l and r) are very prone to be lost or absorbed thusly: e.g., Latin ursus 'bear' > Spanish oso. English 'walk' and 'talk' and 'salmon' all have silent l as first consonant in consonant clusters. Similarly, the -l- was often lost as first consonant in a cluster in the change from Semitic to Uto-Aztecan also: Hebrew śəlaaw 'quail', pl: salwiim; Syriac salway 'quail'; Arabic salwaa 'quail'; Samaritan šalwi > UA *solwi 'quail': CN sool-in 'quail'; Mn sowi' 'pigeon'. So Mn lost -l- as first segment in the cluster. Latin ex- 'out' in English loans

sometimes remains intact: ex-tract, ex-cept; but other times the -x- is absorbed in the cluster and only e- remains: e-mit, e-merge, e-lect, and e-radicate. Another example is English a/an. The original form is *an*, which remains *an* before a vowel (an apple, an iron), but before a consonant the pronunciation of the *n* over time became absorbed or assimilated to the following consonant, that is, -n- was lost as first consonant in the cluster; thus, (a dog (< *an dog), a cat (< *an cat). Another example is Hebrew qadqod 'head, skull' and Assyrian qaqqadu, the latter having assimilated the cluster *-dq- > -qq-. Also similar is Semitic qarqara > UA *qaqqara 'quail'. Such happens repeatedly in many languages throughout the world.

Compare the following Arabic and Hebrew forms:

	Arabic	Hebrew	Uto-Aztecan
daughter	bint	batt	(*pattï 'daughter' 534)
spike of grain	sunbul	šibbolet	(*suNkwu > suŋu 'corn' 828)
wheat	ɦinṭat	ɦiṭṭaa	--
beetle	xunpusaa'	ɦippušit	(*wippusi 'beetle' 853; note Hebrew ɦ > w)

One can see a pattern of *-nC- remaining - nC- in Arabic, but *-nC- > -CC- in Hebrew; thus, the 1st consonant of the cluster was absorbed to double the 2nd, or the 1st entirely assimilated to the 2nd. Similarly, in UA a cluster tended to obscure the 1st C and double the 2nd: *-Ct- > -tt-, *-Ck- > -kk-. Thus, Ca mataš 'crush, squash, vt' is from UA *mattas, because a single intervocalic -t- > -l- in Ca; and Hebrew makteš 'grindstone' matches very well what may have become a denominalized verb (1.17) in Ca mataš 'crush' with *-kt- > -tt-.

Another frequent result of consonant clusters is that the 1st C of the two may become a glottal stop, in a change between remaining and disappearing, but not completely disappearing by leaving a trace of its existence in the form of a glottal stop (ʾ). In English, for example, dictate has a cluster pronounced *-kt- when pronounced carefully, but in normal rapid speech, it is often pronounced as -ʾt-. Mountain is often said mau̲ʾn, the t > ʾ and the underlined vowels are nasalized. Similarly, 'written' is often pronounced rIʾn. In mountain > mau̲ʾn, the nasalized vowels are from the nasal n before the t, while rIʾn has no nasal before the t and does not have its 1st V nasalized. The first consonant becoming a glottal stop happens often in UA as well: we already mentioned Hebrew makteš > UA *maʾta 'grinding stone'.

Some consonants (like ʾ, nasals and liquids) in some langauges tend to be anticipated or fronted (put further in front from their original place). An English example is the biblical Aramaic name of ʕabed-nəgo, for which many English speakers say abindigo, with the n anticipated before the d from its original place after the d. Glottal stops are frequently anticipated in UA: e.g., Egyptian **sbʾ** 'star' > UA *siʾpo 'star': Wr soʾpóri; Tr seʾporí. UA anticipates the glottal stop, yet reflects all three consonants, whereas Coptic siu 'star' reflects only one, though it is also from Egyptian sbʾ 'star' (see 154).

Another route to vowel loss is **accent or stress** patterns. For example, Latin fábuláre stressed the 1st and 3rd vowels, and the lack of stress on the 2nd and 4th vowels helped them both become silent in the changes from Latin to Spanish and Portuguese:

Latin fábuláre > fablar > hablar > ablar (Spanish)

Latin fábuláre > fablar > falar (Portuguese)

Losing the 2nd V caused two originally separated consonants to become a consonant cluster (Latin fábuláre > fablar). Then in that cluster, the 1st consonant was lost or assimilated to the 2nd in Portuguese, similar to what we have talked about and seen in several other examples above. In Spanish, the cluster remained intact, but the initial f > h > ø (ø means zero or nothing, that is, f became h, then h became silent or disappeared). The current spelling of Spanish is hablar; however, h is silent in Spanish, so the first and last sounds of Latin fabulare were lost, as well as the middle unaccented vowel. Because *h* is a rather weak consonant, it often becomes silent or disappears in language change.

These kinds of changes happen in many to most languages. In Uto-Aztecan, stems of CVCVCV often lose the middle V, reducing to CVCCV, then the medial (middle) consonant cluster also reduces to one consonant. This phenomenon is common in Syriac and other Aramaic dialects as well. For example, Syriac kawkab 'star', when taking on the definite article suffix -aa 'the', loses the middle vowel in Syriac kawkb-aa 'star-the' because of stress patterns similar to what we have talked about.

1.16 Pronouns

Pronouns are often portrayed in paradigms like the following:

	Singular			Plural		
	subject	object	possessive	subject	object	possessive
1st person	I	me	my/mine	we	us	our(s)
2nd person	you/thou	you/thee	your(s)	you	you	your(s)
3rd person	he/she	him/her	his/her(s)	they	them	their(s)

Besides persons (1st person speaker, 2nd person spoken to, 3rd person spoken about), number can vary as well. Many languages have singular, dual, and plural, in which case plural is three or more, like Navajo and the Semitic languages (not related). Likewise, Old English had ik (I), wit (we two), and we (3 or more). Pronoun systems with three numbers often simplify to two numbers. Old English gave up its dual to make 'we' mean two or more. Navajo is in process of often having its dual cover for plural in some cases.

Many Amerindian languages, including a few Uto-Aztecan languages, have two 'we' pronouns: we-inclusive is I-and-you, to include the person(s) spoken to, and we-exclusive is I-and-he/they, to exclude the person(s) spoken to. Semitic langauges do not have the inclusive-exclusive distinction, nor does Egyptian, while many Amerindian language families do.

1.17 Nouns Become Denominalized Verbs

Most languages make nouns from verbs and make verbs from nouns, though some do so to a greater degree than others. In English we have 'hoof it' for 'walk'; and 'she mirrors her mother's behavior' for 'she behaves like her mother' from the noun 'mirror'; and 'he bicycled to Bluff' for 'he rode/pedaled a bicycle to Bluff'. These are called denominalized verbs because a nominal (noun) is made to serve as a verb. Even 'pedal' is a denominalized verb from the noun 'pedal'. The term de-nominal verb means 'from-noun verb'.

In the change from Semitic to Uto-Aztecan, many nouns were denominalized to become verbs. In fact, Uto-Aztecan *kuppa 'shine (as stars)' is a denominalized verb from the noun mentioned above: Syriac kawkb-aa 'star-the' > UA *kuppa 'shine (as stars)' wherein the consonant cluster *-kb- > *-pp- as we talked about above, and the vowel *a* assimilated to *w* in *-aw- > -u-.

1.18 Language Contact, Influence, Loanwords, and Mixing

Languages in contact influence each other. The type and intensity of the contact determines how they influence each other and how much. A few languages enjoy relative isolation, like Icelandic isolated in the Atlantic, though none escapes all outside influences. In fact, most languages are subject to various influences over time, and sometimes so intensely or suddenly that changes happen fast. For example, many Native American languages in the United States are dead or dying due to the overwhelming dominance of English. Sometimes the tribe survives, but as English is learned, a bilingual generation or two eventually raises a generation of monolingual English speakers, then as the older native speakers pass on, so does the language. The numbers of speakers of Native American languages in Latin America are generally more numerous, partially because in Latin America the mandatory requirements to attend school and learn Spanish are more lax or non-existent. Bilingual education in the U.S. can help provide some basics and an appreciation for the language and culture, but it does not produce native speakers.

One factor in language influence is numbers. When a small population dwells amidst a much larger population, the influence is usually proportionately imbalanced. As in our previous example, the nation of 300,000,000 English speakers contributed to the loss of some native languages, yet some of the native languages contributed loanwords to the much larger language despite the huge discrepancy in numbers/influence. Moccasin, tomato, and coyote are loanwords into English from Native American languages, the latter two from Nahuatl (Aztec), a Uto-Aztecan language.

A second factor in language influence is the relative perceived status of each language, that is, the relative cultural, political, or international superiority. The language of a people perceived to be culturally superior usually does more influencing than being influenced and is often called a superstratum to languages receiving their influence. For example, at one time, Latin was the language of learning and English was not

allowed in the schools; and during that time, many Latin loanwords were borrowed into English, most of our bigger, more academic words. The once pervasive status of Greek and Latin in academia are apparent in our medical terminology. We say cardiac arrest instead of heart-stop, five syllables instead of two, all due to previous perceptions of status. Greeks were once the dominant culture; thus, much Greek vocabulary was borrowed into Latin. Then the Romans became politically dominant, whether cultural or not, and so the rest of Europe borrowed much Latin, along with the Latin versions of their Greek loans already in Latin. While most borrowing between languages happens gradually, sometimes it is sudden and massive, more like a sudden mixing of languages.

Language mixes also exist. Spanglish or border Spanish are terms often applied to the frequent mixing of English and Spanish, but usually by those who know both languages and can speak either when needed. Sometimes the language mixing becomes fixed and becomes an actual language—English, for example. Modern English is a language mix of Old English and Norman French. Only 15% of Old English survived into modern English (Baugh and Cable 55), yet we still call it a Germanic language because most of the most basic words are Germanic, that is, from Old English, which was a Germanic language; e.g., body parts like head, hand, eye, and common nouns of nature like earth, water, etcetera, are Germanic. However, take almost any page of written English, look up the words to find their origin, and about half of any page or paragraph comes from French or Latin, if not more than half. In 1066 the Norman French conquered England and imposed their French as the language of the new rulers on their new land. For the next three centuries, the rate of French loans into English happened to such an extent that every generation of about 10 generations must have shaken their heads at the next generation's demolition of "proper" English, though the head-shakers did their share of damage, perceived by the generation preceding them. During this language mixing, English lost the case endings of nouns and the conjugation of verbs. Many irregularities of strong verbs in Germanic became "regular" verbs (with -ed past tense): shaved replaced shove as the past tense of shave; clomb became climbed; and hundreds more. In the Midwest, many are familiar with "clumb" as a past tense of climb—yesterday I clumb a tree. Most would count it as outback bad English, when in fact it is straight from Old English clomb (past tense) and is more original than the 'climbed' that we say today. In fact, those who first said 'climbed' were wrong until most were saying it, then 'clomb' became wrong. Nevertheless, the intensity of the contact during French rule in England caused English to change rapidly, and to end up as quite a language mix of Old English and French. Yet that kind of mixing of languages and peoples happens regularly. In fact, the Norman French themselves were a mixture of at least four peoples: the Viking (Germanic) Norseman (source of Norman) who settled their area of France, and they mixed with the French, who descend from the Celtic Gauls, the Germanic Franks, and the Romans who brought the Latin language which in that area became French. UA is also a language mix, as shall be seen later.

Such mixing happens often among Native Americans as well. In my classes, I ask my Navajo students how many of them have all four grandparents' being Navajo. Few raise their hands. Then I ask how many have one or two grandparents who are of another tribe or ethnic group. Most raise their hands. Most have one or two grandparents who are Ute or Hopi or Walapai or Sioux or Hispanic or Irish, etc.

Besides words being borrowed, language influences alter the grammar of a language as well. These grammatical changes are sometimes harder for native speakers to identify or even perceive, because, as we said previously, we mostly do grammar subconsciously, and so when bilingualism is prevalent in a border area between languages, the subconscious grammatical patterns of the two tongues can and do influence each other slowly enough that native speakers are hardly aware. For example, English *whom*, as accusative (object) form of *who*, is nearly dead as a last survivor of the Old English case system, yet most English speakers do not know how to use it and so do not, or if they do, they often use it incorrectly, because the case system in which it fits or which used to be part of the language, has all been lost for centuries.

This is all very applicable to a hypothesized arrival of Mediterranean speakers in ancient America, because the languages would differ enough that it is to be expected that such an arrival in a very different language environment would change very much. The derivational detail being lost would not be surprising, just as the Germanic case endings were lost in Middle English. The simplification or loss or fossilization of some verb conjugations would be expectable, just as English lost most of its verb conjugations.

1.2 A Brief Introduction to the Semitic Languages

Hoping to introduce Semitic in a few pages is rather presumptuous, since a 400-page book better suits such an effort. In fact, each Semitic language needs 400 pages. Good compact books on Semitic include Bennett (1998) and Rubin (2010), and more involved are Goldenberg (2013) and Lipinski (2001). Regardless, some basic features of Semitic warrant a few words in a work dealing extensively with Semitic.

The Semitic language family first divided into West and East Semitic. East Semitic is essentially Akkadian, which later developed into Assyrian (north) and Babylonian (south) in Mesopotamia. The Semitic family tree's branching thereafter may ever lack consensus, but mostly following Rubin (2010, 3-6), let us consider that West Semitic divided into Ethiopic (languages spoken in or near Ethiopia), Modern South Arabian (a different branch than Arabic) consisting of six languages spoken in Yemen and Oman, and Central Semitic. Central Semitic then divides into Arabic, Northwest Semitic, and Ṣayhadic, also called Old South Arabian or Epigraphic South Arabian, a group of dialects found in inscriptions in western Arabia from 1000 or 700 BC to AD 600 (Rubin 2010, 13-14; Goldenberg 2012, 15-16). Regarding Arabic, Classical Arabic is the language of the Qur'aan, and, though not an ancestor, is like a sister to the parent language(s) of the various Arabic dialects spoken today. The Northwest Semitic languages referred to in this study are Hebrew / Phoenician / Canaanite (different names or dialects of the same language), and Aramaic / Syriac, and Ugaritic. Aramaic periodically gained and waned as a frequently dominant language, lingua franca, or international language in the Fertile Crescent areas of the Near East. Aramaic developed into many dialects, Biblical Aramaic (books of Daniel and some of Ezra), Jewish Aramaic, Syriac, Samaritan, Mandaic, and several others, including many modern Aramaic dialects surviving to this day.

The Semitic languages have remained in relatively close contact with each other for millennia and thus retain many morphological similarities. The Semitic languages are very verbally based with only a few basic original nouns not easily associated with a verb root, as most nouns are derived from verbs. The tri-consonantal roots change shapes for various conjugations, participles, and nouns.

1.21 Semitic Verbs and Conjugations

Semitic verbs or verbal roots mainly consist of three-consonants. Four-consonant roots occur as well, such as Semitic prʕš 'jump'. Very often two-consonants seem to underlie related roots. Using 1 and 2 for those two consonants, related roots take forms like 12y (gly), 1w2 (gwl/gyl), 122 (gll), 1212 (glgl). Semiticists have also noticed that two consonants with whatever 3rd consonant often have related meanings; for example, many roots with pr… as the first two consonants generally have meanings like separate, part, divide: prd 'detach, separate, divide'; prṭ 'open wide, split'; prk 'crush, grind, break apart'; prm 'tear apart'; prs 'divide, separate, break bread'; prṣ 'split, make a breach, spread'; prq 'take away, split, part (ways), fork'; prś 'spread, stretch out'; pry 'produce/bear fruit/child (something separates from its producer, e.g., mother or tree)'. In Semitic roots, changing vowel patterns alter the shape of the root for a variety of structures and purposes, some also taking prefixes and suffixes for person and aspect.

Semitic verb conjugation patterns consist of two primary categories: one is a suffix conjugation or perfective (pfv) conjugation, because it usually expresses past tense or perfective (completed action or relative past) in Central Semitic and the persons doing the verb are revealed in the suffix (Arabic katab-ta 'wrote-you'); the other is a prefix conjugation or an imperfective (impfv) conjugation, because it usually expresses imperfect (not completed) aspect, i.e., usually present or future, and the subjects doing the verb are expressed in the prefix (Arabic ta-ktubu 'you-write/are writing).

The basic verb, in Hebrew, is called the qal (easy/light) conjugation. Arabic best reflects the Proto-Semitic form *CaCaCa (C = any consonant), while the other Semitic languages have lengthened, shortened, or lost a vowel or two:

Arabic kataba 'he wrote'

Aramaic/Syriac kətab 'he wrote' (shortened the 1st vowel and lost the 3rd)

Hebrew kaatab 'he wrote' (lengthened the 1st vowel and lost the 3rd)

Akkadian kataabu 'he wrote' (lengthened the 2nd vowel).

Uto-Aztecan also has many of these 3rd sg forms *CaCaC(a), the last consonant/syllable sometimes lost:
At (79) Hebrew **ḫmr** ‘to cover or smear’ (with s.th.) > UA ***humay** ‘smear, spread, rub, paint’ >
Ca húmay 'smear, paint, vt’; Cp hume- ‘spread a liquid or s.th. fine’. (ḫ > hu in UA, and r > y)
At (645) Semitic ḫabala ‘corrupt’; Hebrew -ḫabbel ‘ruin’ > Hopi hovala ‘waste s.th. of value, squander’.

For abbreviations of the UA languages, see the introduction to UA. The sound changes are covered in detail in the body of the sets, though we may here list some of the less obvious in parentheses. For example, both of the first two (79, 645) begin with the pharyngeal ḫ, which became UA *hu, or ho in Hopi. Also, when the 3rd consonant is y or ’ in Semitic (CCy/CC’), it is often not apparent in Semitic’s perfective *CaCay > CaCaa, but sometimes is in UA, as in the next example:
At (559) Hebrew **bky/ bakaaʸ** ‘cry, weep’ (perf stem); Syriac bakaa / **baka’** > Hopi pak- ‘cry’;
Tb pahaa’at / ’apahaa’ 'cry, bawl, howl' (Tb h < *k); Ktn paka’ ‘ceremonial yeller, clown who shouts all day to announce a fiesta’.
Of interest is that the Syriac form actually shows the glottal stop, often only used as a long vowel place holder; yet the glottal stop in Tb and Ktn show the glottal stop pronounced, aligning with Aramaic/Syriac more than with the Hebrew and Arabic terms lacking that glottal stop. Another pfv form is
At (565) Hebrew **mkr / maakar** ‘sell (he sold)’ > UA *makaC ‘give’ in all of UA; UA *na-maka ‘sell’

Hebrew’s first long vowel (kaatab) can be shortened when a suffix draws the stress/accent toward the end, as in Hebrew kətab-tem ‘wrote-you pl’. Many such vowel variations occur in Semitic, especially in Masoretic Hebrew (Old Testament Hebrew) which is a dialect of Hebrew not necessarily representative of all dialects in all centuries, to be discussed below. So Masoretic vowelings should not always be taken as absolute or as original. A more complete table of the pronoun suffixes to the verbs of Akkadian, Hebrew, Syriac, and Arabic is further below, but let us now continue our examples of Semitic with comparable fossilized forms in Uto-Aztecan.

In addition to the more common *CaCaCa, some Semitic verbs are voweled as *CaCiCa, as also in Arabic CaCiCa, sometimes Hebrew CaCeC and Aramaic CəCeC. Examples follow:
(3) Northwest Semitic *yašiba ‘sit, dwell’ > UA *yasipa ‘sit, dwell’ (yaašab in Masoretic Hebrew)
(769) Hebrew **tqp** ‘to overpower, v’; Aramaic(J) təqef ‘be strong’; the 2nd vowel of Aramaic shows Proto-Semitic *taqipa (sg), *taqipu (pl), exactly as UA *takipa and *takipu ‘push’.
Of interest is that while *yašiba reflects the 3rd person singular, the 3rd person plural *yašibuu is seen in the Tepiman branch of UA in ST daivu and TO dahivup, both pl forms (Tep d < *y, Tep h < *s, Tep w/v < *p).

All the above exemplify the perfective/suffix conjugation. The imperfective/prefix conjugation is
Arabic: ’a-ktubu ‘I-write’; ta-ktubu ‘you-write’; ya-ktubu ‘he-..’; na-ktubu ‘we-..’; ya-ktubuuna ‘they..’
Hebrew: ’e-ktob ‘I-write’; ti-ktob ‘you-write’; yi-ktob ‘he-…’; ni-ktob ‘we-…’; yi-ktəbuu ‘they…’

Again, the Arabic forms are more original, and note the last Hebrew (they) form loses the round vowel (o > ə) to shortening in Masoretic phonology, but is preserved in Arabic. One can also see that the 1st and 2nd consonants are clustered in the impfv stem (Arabic -CCuCu, Hebrew -CCoC, or -CCaC for some verbs). And since clusters often lose the 1st consonant in UA, the UA fossilizations of the imperfect often lack the 1st consonant. In Uto-Aztecan are many fossilized impfv qal forms, some with the prefix + impfv stem, others with only the impfv stem:
(1094) Hebrew **ktš** ‘pound, pound fine, grind’; impfv: -ktoš < ***-ktušu** with loss of 1st C in the cluster > UA ***tusu** ‘grind’ in most UA languages.

Besides impfv stems like Arabic ya-CCuCu / Hebrew yi-CCoC with the stem vowel *u/o* in the impfv stem, some verbs have a stem vowel of *a*, as in Hebrew yi-CCaC / Arabic ya-CCaCu. A prominent example of each is Hebrew ya-ʕ^{a}qob ‘he grabs the heel, deceives’ (Jacob) and Hebrew yi-ṣḥaq ‘he laughs’ (Isaac).

Another example of that impfv stem vowel is Arabic labisa, impfv: (ya)-**lbasu** ‘put on, wear’ and Hebrew lbš, impfv pl: (yi)-lbašuu. In this Semitic-kw item, the cluster absorbs the 1st consonant to dagesh (double) the 2nd as if -bb- > kw:
(50) from Hebrew lbš, impfv: -lbaš- ‘put on (garment), clothe (oneself)’: impfv stem vowel is -a-, as in UA: -lbaš > kwasu; pl would be yi-lbašu > UA *kwasu ‘dress, shirt, put on clothes’ in most of Numic.
(749) also Hebrew tmh, impfv: **-tmah** ‘be astounded, dumbfounded, v’ > UA ***maha** ‘fear’:
Wr maha- ‘be afraid’; Yq máhhae; AYq mahai ‘scared’; Tr mahá; CN mawi ‘be frightened’.

Some fossilized imperfective forms in UA include the prefix. For example, the previously noted perfective of Semitic/Syriac baka' 'cry' > UA paka' 'cry' has as its impfv Arabic ya-bkiy, Hebrew yi-bke. Considering that bilabials disappear as first consonant in a cluster (see 294-300), then the imperfective stem with the 3rd sg prefix yi-bke / *ya-bka would look like UA *yaka 'cry' which is exactly what we find:

(560) Semitic *ya-bka^{y} 'he/it cries' > Hebrew yi-bke$^{(y)}$ > UA *yaka / *yaCka / *yakka 'to cry, sg'

(561) Semitic *ta-bka^{y} 'she/it cries' > Hebrew ti-bke$^{(y)}$ > NP taka (< *takka) 'cry, vi'.

The first (560 UA *yakka 'cry') appears in many UA languages; the second (561 *takka 'cry') appears in Northern Paiute; so NP has both the 3rd masculine sg impfv *yakka and the 3rd feminine sg impfv *takka.

Certain consonants cause variant vowelings in Semitic. For example, the initial aleph or glottal stop of Semitic 'kl 'eat' has the usual perfect *'akal (798), but the impfv with prefixes results in 3rd fem sg imperfective Hebrew to'kal 'she/it eats' (796).

(798) Hebrew **'akal** '(he) ate (perfect), *to'kal 'she/it eats'; *yo'kal 'he/it eats'
> UA *'**aki** 'open mouth, eat, take/put into one's mouth' (In Semitic-kw, final -l raises vowels)

(796) Hebrew 'akal '(he) ate', ***to'kal** 'she/it eats' > UA/Numic ***tïkkaC** 'eat'
(Numic ï < UA *u, which corresponds to Hebrew o; so all matches, the doubled medial consonant from the *-'k- cluster and a final underlying consonant from final -l:
Hebrew *to'kal > Numic *tïkkaC. (In Sem-p, final -l does not raise the preceding vowel)

The participle of the Hebrew qal conjugation is *CooCeC, which corresponds to UA *CuCiC. A number of such *CuCiC forms appear in UA:

(754) Hebrew pny / panaa^{y} 'turn, turn and look, look'; participle **pone** > UA ***puni** 'turn, look'

Besides the qal or basic verb, all Semitic languages also have an intensive conjugation, usually doubling the middle consonant: Arabic CaCCaCa; Hebrew CiCCeC, called the qittel form in Hebrew, whose original form and UA form are usually *CiCCaC. We saw Hebrew ktš in the impfv -ktuš above; below is an apparent intensive of the same ktš in the intensive *CiCCaC form:

(615) Hebrew **ktš** 'pound, pound fine, bray, v'; kaataš (perfect qal); unattested *kitteš < *kittaš would be the qittel form: Yq **kitte / kittasu** 'grind, mash'. Some suggest that the final -su of the Yq form is another morpheme; even if so, kitta is striking enough, since we seldom see 3rd consonants in UA.

The general meaning of the intensive in Semitic is intensification, continuative, causative, distributive, or repetitive action; interestingly a consonant doubling or syllabic reduplication in UA languages is also employed for intensification, continuative, distributive, or repetitive action. Moving on, the imperfect of this intensive is Arabic yu-CaCCiCu and Hebrew/Aramaic yə-CaCCeC. The imperfective intensives are also well represented in UA:

(11) Hebrew impfv -dabber (< *-dabbir) 'to speak' (qittel) > **UA *tïkwi** 'say' (*-bb- > -kw-)

(809) Hebrew qittel impfv stem **-hattel (< *-hattil**) 'to mock' > **UA *'ati / *ata** / ***aCti** 'laugh'

(907) Arabic ğassa 'touch, feel'; Hebrew gšš 'touch'; perfect qittel: giššeš 'grope';
Hebrew qittel impfv: *-gaššiš > Ls ŋési 'touch lightly, graze, vt'; Cp ŋíse 'scratch, vt'. It may be due to s.th. else, but interestingly the Ls and Cp forms align with the impfv and pfv qittel forms.

Most Semitic languages also have a causative: cause someone to do s.th. Hebrew forms are often represented with the consonants q-ṭ-l, which we simplify to q-t-l, which are more original anyway. These basic causative forms are as follows:

	perfective	imperfective	participle
Hebrew	hiqtiil / hiqtal-(ti), etc	ya-qtiil, ta-qtiil, etc	maqtiil
Arabic	'aqtala / 'aqtal-(tu)	yu-qtilu	muqtilu
Aramaic	'aqtel	y-aqtel	maqtel

From the root slm 'peace', the Arabic causative is 'aslama 'cause peace'; the verbal noun is Islaam, and the participle is muslim 'one who causes peace, peace-maker'. UA forms resemble the Hebrew causatives: hiCCiiC, hiCCaC. Examples of that causative in UA are
At (1354) Hebrew hi-kbad- > UA *hipaca 'sweep' (d > c(ts)),
At (810) Hebrew hikkiir 'recognize, know, know how to' > Tr iki- 'know, be aware of'

At (1293) Hebrew hiśkiil, hiskal- ‘to understand, comprehend, make wise’ > CN iskal ‘to train’;
CN iskal-ia ‘be discreet, prudent’
At (567) Hebrew ya’amiin ‘he believes/trusts/stands firm’ > UA *yawamin ‘believe’ (’ > w)

The passive of the causative—be caused to do s.th.—in Hebrew is called the huqtal or hoqtal (huCCaC / hoCCaC) with a participle of muqtal. If the 3rd consonant is -y, then the forms are huCCe and muCCe. An example from a common Hebrew stem of a muCCe form is UA *mukki ‘be sick, die’ aligning with the participle of Hebrew mukke ‘smitten’ (52) and furthermore, Tb hookii ‘deceased grand-relative after death’ aligns with the Hebrew pfv hukke, a slight vowel discrepancy o/u; yet even in Hebrew the form is called both huqtal and hoqtal because both vowels happen among huqtal / hoqtal forms.

Also frequent enough in UA are the passive/stative adjectives / nouns, such as CaCiiC (qariib 977); and a form denoting noun of occupation or habit, i.e., noun who does the verb CaCCaaC (śannaa’ 756).

The Semitic Cohortative/Volitive -a Verb Suffix in Uto-Aztecan

A certain suffix of the Semitic imperfective (impfv) verb is **-a**, and merits mention as it seems to appear in Uto-Aztecan frequently enough. Cohortative and volitive are terms having to do with ‘will’ and ‘wanting to do’ the verb it is suffixed to. The cohortative -a in Hebrew signifies encouraging a cohort (group) to do something or a wish/wanting/suggestion that they do something, as in let’s … In Biblical Hebrew, the cohortative is limited to 1st person: let us do (s.th.) , or let me (do s.th.) or I shall (with more emphatic intention). However, in other Northwest Semitic languages closely related to Hebrew, the cohortative is not limited to 1st person. This -a vowel is related to the Arabic subjunctive -a, which signifies any potential action. This Semitic volitive -a at times can apply to a high percentage of subordinate clauses. (Blau 2010, 207; Lipinski 2001, 360-363) And the syntax of Semitic languages often allows much higher percentages of subordinate clauses than are typical in European languages.

This -a suffix is often used with verbs of motion, as in Hebrew neeləkaa ‘let us go!’ (1st pl, from Lipinski 2001, 363), and UA *yĭNka ‘enter’ (go in) from Hebrew yeelka (3rd sg) is exactly the same root as Lipinski uses in his example, but with 3rd person yee- prefix instead of 1st person pl nee-. Many other examples of this -a suffix permeate the Semitic-UA data.

Semitic Pronoun Morphology on Verbs

Semitic pronominal morphology on verb conjugations (pronominal is the adjective from of pronoun) consists of pronoun morphemes prefixed to the imperfective (not-completed/present/future) verb forms and other pronoun morphemes suffixed to the perfective (completed/past) verb forms:

Verbal Pronominal Suffixes of Some Semitic Languages:
Suffix verb conjugation (usually perfect/past) pronoun forms suffixed to *CaCaC-:

	Akkadian	Hebrew	Syriac	Arabic	
I verbed	-aaku	-tii	-eet	-tu	
you masc sg	-aata	-taa	-t	-ta	
you fem sg	-aati	-t	-t	-ti	
he	-	-	-	-a	
she	-at	-aa	-at	-at	
we	-aanu	-nuu	-nan	-naa	
you masc pl	-aatunu	-tɛm	-toon	-tum	
you fem pl	-aatina	-tɛn	-teen	-tunna	
they masc pl	-uu	-uu	-uun	-uu	
they fem pl	-aa	-uu	-een	-na	(Goldenberg 2012, 85)

The bound pronominal prefixes to verbs in the prefix conjugation (usually imperfect/present/future) are shown below. Some person forms also include a suffixed element (like -uu plural), though the prefixes are the primary indicators of person:

	Akkadian	Hebrew	Syriac	Arabic (classical)
I verb	a-	ʾɛ-	ʾ-	ʾa- / ʾu- -(u)
you masc sg	ta-	ti-/tɛ-/tə-	t-/te-	ta- / tu- -(u)
you fem sg	ta- -ii	ti-/tɛ-/tə- -ii	t- -iin	ta- / tu- -ii(na)
he verbs	i-	yi-/yɛ-/yə-	y-	ya- -(u)
she verbs	ta-	ti-/tɛ-/tə-	t-	ta- / tu- -(u)
we verb	n-	ni-/nɛ-/nə-	n-	na- / nu- -(u)
you pl masc	ta- -aa	ti-/tɛ-/tə- -uu	t- -uun	ta- / tu- -(u)
you pl fem	ta- -aa	ti-/tɛ-/tə- -naa	t- -aan	ta- / tu- -na
they masc	i- -uu	yi-/yɛ-/yə- -uu	n- -uu(na)	ya- / yu- -uu(na)
they fem	i- -aa	ti-/tɛ-/tə- -naa	n- -aan	ya- / yu- -na

(Goldenberg 2012, 86-87)

One can readily see the similar morphology among the Semitic conjugated verbs. While most Semitic verbs contain three consonants, Semitic (and Egyptian) have occasional quadrilateral verbs (of 4 consonants), such as Semitic prʕš ‘jump’ from which the Semitic noun parʕoš ‘flea (jumper)’ derives as a ‘jumper’. (Note UA *par’osi / *paro’osi ‘jackrabbit’ which is also a jumper and shows all four consonansts and both vowels.)

1.22 Semitic Pronouns

While presenting the Semitic pronominal affixes on verb conjugations, let us also look at the Semitic independent pronouns and the suffix pronouns. The independent pronouns for Akkadian, Hebrew, Syriac, and Arabic follow. Those found in or relevant to UA forms are in bold. See UA pronouns (101-114).

	Akkadian	Hebrew	Syriac	Arabic (classical)
I	anaaku	ʾanooki / ʾani	ʾenaa /(i)**naa(’)**	ʾanaa’
you masc sg	atta	**ʾattaa**	ʾatt	ʾanta
you fem sg	atti	ʾatt	ʾatt	ʾanti
he	šuu	**huu**	**huu**	**huwa**
she	šii	hii	hii	hiya
we	niinu	(ʾa)naħnuu / ʾaanuu	ħnan	naħnu
you pl masc	attunu	**ʾattɛm**	ʾattoon	**ʾantum**
you pl fem	attina	ʾatteen(aa)	ʾatteen	ʾantunna
they masc	šunu	**heem**(maa)	hennoon	**hum**
they fem	šina	heen(naa)	nenneen	hunna

(Goldenberg 2013, 82; Lipinski 2001, 306-7)

The Semitic oblique or suffix pronouns are used as possessors, objects, and subjects (as in his/your giving me/it). Oblique generally refers non-subject pronouns, i.e., object (of verb), dative (to/for whom given/done), and/or possessive pronouns. Again, forms appearing in UA or relevant to UA are in bold:

	Hebrew	Syriac	Arabic (classical)
I	**-ni / -i**	**-ii** / **-ay**	**-ni / -i**
you masc sg	**-kaa** / -aak	-aak / -**ay**k	-ka
you fem sg	-eek / -aak	-eek / -**ay**k	-ki
he	-(aa)**huu** /-**aaw** /-**oo**	**aaw(hi**)	**-hu/-hi**
she	**-haa / -aa**(h)	**-eeyh / -hi**	-ha
we	-nuu	-an / -**ay**n	-naa
you pl masc	**kɛm**	-koon /-**ay**koon	-kum
you pl fem	kɛn	-keen /-**ay**keen	-kunna
they masc	**hɛm / -aam**	hoon /-**ay**hoon	hum
they fem	hɛn / -aan	heen /-**ay**heen	hunna

(Goldenberg 2013, 88; Lipinski 2001, 314-15)

1.23 Semitic Sound Correspondences

Some Proto-Semitic consonants remain unchanged across the Semitic languages (l, r, m, n, y, which will not be listed), while others undergo changes worth noting. Though an additional proto-consonant or two have been proposed and debated, the generally accepted Semitic sound correspondences are as follows:

Proto-Semitic	Arabic	ESA	Ugaritic	Hebrew	Aramaic	Akkadian	(ESA = Epigraphic South Arabian)
*b	b	b	b	b	b	b	
*p	f	f	p	p	p	p	
*g	ǧ	g	g	g	g	g	(Arabic ǧ = j, from Proto-Semitic *g)
*k	k	k	k	k	k	k	
*q	q	q	q	q	q	q	
*t	t	t	t	t	t	t	
*d	d	d	d	d	d	d	
*đ	đ	đ/d	z	z	d	z	(đ = th as in the)
*z	z	z	z	z	z	z	
laryngeals / pharyngeals							
*’	’	’	’	’	’	’/ø	(ø = zero, no sound, disappeared)
*h	h	h	h	h	h	’/ø	
*ʕ	ʕ	ʕ	ʕ	ʕ	ʕ	’/ø	
*ġ	ġ	ġ	ġ	ʕ	ʕ	’/ø	
*ḫ	ḫ	ḫ	ḫ	ḫ	ḫ	’/ø	
*x	x	x	x	ḫ	ḫ	x	
sibilants (s-like consonants)							
*θ	θ	š	θ	š	t	š	
*š / s_1	s	š	š	š(s_1)	š	š	
*ś / s_2	š	ś	š	ś(s_2)	s	š	(ś = originally lateral fricative, ≈ voiceless ɬ)
*s / s_3	s	s	s	s(s_3)	s	s	
emphatic / pharyngealized consonants							
*ṭ	ṭ	ṭ	ṭ	ṭ	ṭ	ṭ	
*ṣ	ṣ	ṣ	ṣ	ṣ	ṣ	ṣ	
*ṯ	ẓ	ṯ/ṣ	ḏ	ṣ	ṭ	ṣ	(ṯ = emphatic interdental fricative)
*Ļ	ḏ	Ļ	ṣ	ṣ	ʕ	ṣ	(Ļ= emphatic lateral fricative)

(Bennett 1998, 68-71; Goldenberg 2013, 68; Lipinski 2001, 112-157)

1.24 Masoretic Hebrew

Masoretic Hebrew is the dialect(s) of the Hebrew Old Testament (OT) text as voweled by the Masoretes about AD 600-700. The original texts or various books of the OT were written with only consonants, like most Semitic languages, and were composed at different times, roughly ranging in date from 1200 to 300 BC. So some 1000 to 1800 years after the consonantal texts were written, the Masoretes developed a system for writing vowels and some consonant variations. The consonant variations from Proto-Semitic and probably early Hebrew to Masoretic Hebrew are that the stops became fricatives or spirants following vowels: b > v, p > f, k > x, t > θ, etcetera, but at the beginning of the word, or when doubled, or following a consonant, b remains b, p > p, etc. The same spirantization occured in Aramaic dialects as well. However, the Uto-Aztecan forms from Semitic do not show such spirantizations in Proto-Uto-Aztecan forms, though some spirantization happened later in some UA languages, like *p > v in some Northern Uto-Aztecan languages. Because UA does not come from a later spirantized Hebrew, but from earlier non-spirantized Semitic forms, we will not include those later spirantizations when citing Hebrew and Aramaic forms, because the spirantization was not original and is not apparent in early UA reconstructions. Arabic spirantized a couple of consonants—*p > f and *g > ǧ/j—changes from Proto-Semitic *p and *g, but again, parallels with UA do not reflect those changes.

1.25 Semitic Cognates

Semitic Cognates are the similar words or groups of related words in the Semitic languages; each group of related words descends from its ancient predecessor or ancestor proto-word. For example, from Proto-Semitic ***đi'b** 'wolf' (Bennett 1998, 60) are descended Hebrew **zə'eb** 'wolf', Arabic **đi'b** 'wolf', Syriac **di'b-aa** 'wolf-the', and Aramaic di'b-aa 'wolf-the'. Initial Proto-Semitic *đ corresponds to Hebrew z, Arabic đ, Syriac d, and Aramaic d; thus, those consonants begin the respective forms in those languages; the glottal stop (', 2nd consonant) and b (3rd consonant) remain the same in those languages. This set (Semitic *đi'b wolf) has a cognate in most Semitic languages (note UA *tï'pa 'wolf'); however, sometimes cognates appear in less than half the languages, such that the once-existing cognate did not survive or continue in all languages. This happens in all language families: some cognates continue prevalent or well represented in most lanugages, while others become sparsely represented, that is, may surface in only two or three languages, or may disappear altogether.

In this connection, sometimes the corpus or extent of an ancient language's vocabulary or cognates can hardly be known. The ancient Akkadian or Assyrian vocabulary is known to be rather voluminous as extracted from extensive records. The vocabularies of thriving modern languages with numerous native speakers, like the various Arabic dialects, can be quite thoroughly known as well. However, some ancient languages, whose records are limited, leave a proportionately limited amount of information behind and so our knowledge of them is similarly limited. For example, the ancient Epigraphic South Arabian languages (a different branch of Semitic than the Arabic dialects) are known only by a limited number of inscriptions on rock, and are limited in content and style to legal transactions, declarations of events, tombstones, and the sort, but are lacking a rich literature or lengthy narratives with extensive amounts of language. Though a little better known than Epigraphic South Arabian, **Biblical Hebrew** is also a limited corpus. The Israelites' dialects changed through time, from Moses to Jeremiah, as all living languages always do, and each book is but a snapshot (not a photoalbum) of that author's dialect in that century. So we know very little when considering all the dialects of all the centuries. The Book of Job, for example, represents its own unique dialect, and has many words which occur only once in the Old Testament (OT), though most books have theirs too. So if the whole OT has many words that made it into the text only once, how many other thousands of words in the spoken language missed out on gaining a single appearance in the OT?

A few inscriptions of ancient Hebrew also exist, but the Hebrew Old Testament text is by far most of what we know about classical or pre-exilic Hebrew (spoken before the exile or before the destruction of Jerusalem in 587 BC). After the Jewish captivity in Babylon, where Aramaic was spoken and where survivors became Aramaic speakers, Hebrew changed and much of its richness and former vocabulary had to have been lost. In fact, the post-exilic Biblical books of Daniel and parts of Ezra are written in Aramaic, not Hebrew. So what percent of the Israelite's pre-exilic spoken Semitic is found in the Masoretic Hebrew text? Would it exceed 10% or 20%? What percent of a pocket English dictionary is found in our Old Testament translation of that Masoretic text? That cannot be a high percentage either, let alone compared to the multi-volume Oxford English Dictionary. Consider, for example, that a Hebrew word for 'squirrel' does not occur in the Hebrew Old Testament text, yet the spoken language certainly had words for squirrel, and UA has three words for squirrel aligning with what would be the Hebrew cognate of Arabic and Aramaic words for squirrel. Arabic **singaab** 'squirrel' would correspond to Hebrew ***š/siggoob** 'squirrel' to which UA ***sikkuC** 'squirrel' corresponds perfectly (C means an underlying consonant that doubles the next consonant, and devoicing g > k, and rising of o > u, all typical of the Semitic to UA sound changes; see number 57). Arabic **qarqađ**aan 'squirrel' > UA ***qoŋi-** 'squirrel' does very well for 5 segments (segments are consonants or vowels) and **qarqađ** is the essence of the word, -aan being a noun augment of sorts: the cluster *-rq- > -ŋ- in Northern UA, which tends to nasalize liquids (change r and l to n or ŋ) and the velar nasal (ŋ) from a liquid and guttural (back consonant) cluster is all quite natural. Like words for squirrel, many other words and verbal conjugations would have been in the spoken language, but not be in the OT text.

Two factors limit our knowledge of the pre-exilic language: besides (1) a relatively small amount of the whole language finding its way into the Israelites' texts while the language was known, (2) even their knowledge of their language deteriorated after the exile, parts becoming unrecoverable within two or three generations. Future discoveries of additional ancient texts is always possible, but as matters now stand, we know only a small percentage of the ancients' conversational vocabularies. The Bible's retention of ancient

Hebrew may approximate the 15% retention (or 85% loss) of Old English in later English after French became the dominant language in English speakers' lives from 1066-1300+.

Whenever another language of a language family is discovered, it is invariably a unique combination of features, some of which are typical and expectable and others not so typical or expected. For example, the Nabatean language, though officially considered an Aramaic dialect, is more Arabic-like than other Aramaic dialects. The language in Job has leanings that are more Aramaic- and Arabic-like than the other books of the Hebrew OT text. So to find a peculiar combination of features in UA, some more Aramaic-like and some more Arabic-like, but all fused into a basic Hebrew conjugation system, is actually quite typical of any newly discovered relative to a group of relatives. To find cognates that match an Akkadian word or an Arabic word or an Aramaic word, but without an attested (verified) Biblical Hebrew cognate should not be thought strange at all. That is how cognates work, in any language family. Each relative has its surprise cognate contributions as well as its random voids.

1.26 'The' in Semitic

'The' in Hebrew and Arabic is a prefix, reconstructing to something like *hal-, though *han- has also been proposed. The -l- is absorbed / assimilated to double the next consonant in Hebrew: hay-yeled 'the-boy'; ham-melek 'the-king'; haš-šaloom 'the-peace'. Various ha-/hi-/a- noun prefixes sporadically appear in UA as noun prefixes, though it is unclear what their original meaning and purpose were, yet they resemble fossilized ha- prefixes, sometimes changing the vowel ha-/hi-, though Hebrew itself also sometimes changes the vowel ha-/he-. These may more often be nouns from Sem-kw. The Arabic article al- lost the h, but keeps the l- before some consonants—al-malk 'the-king', al-walad 'the-boy'—but assimilates before other consonants—as-salaam 'the peace', ađ-đakar 'the-male/man'.

Most interesting, however, are the Aramaic forms, which are abundantly apparent in UA. All Aramaic dialects suffix 'the' to their definite nouns: -aa 'the' is suffixed to masculine nouns and -taa 'the' suffixed to feminine nouns (feminine -taa is actually from feminine -t- + -aa): for example, malk-aa 'king-the', malkə-taa 'queen-the' and this definite the- form is often the citation form or the more common form of the noun. In fact, Goldenberg (2012, 133) says that in Syriac "the historically definite forms became the normal forms of nouns, unmarked for definiteness." The feminine definite suffix (UA *-ta) became part of the citation form in UA as well, though droppable when possessed as in Semitic also. We see -aa fossilized on many UA nouns that were masculine nouns in Semitic, and -taa is still productive as the general absolutive suffix on UA nouns in many branches of UA. Examples of masculine -aa are
Aramaic pagr-aa 'corpse-the' > Hp pïïkya 'skin, fur' (from dead animal) vs. Hebrew (hap-)pɛgɛr
Syriac šigr-aa 'drain, ditch-the' > Hp sikya 'small valley, ravine, canyon with sloped sides'
Aramaic(J) rəʼemaan-aa / reemaan-aa 'antelope-the' > UA *tïmïna 'antelope' (604)
Aramaic di'b-aa 'wolf-the' > UA *tïʼpa 'wolf' vs. Hebrew (haz-)zə'eb 'the-wolf' (618)
Aramaic(J) diqn-aa 'beard-the, chin-the'> UA *tïʼna 'mouth' vs. Hebrew (haz-)zaaqaan 'beard/chin'(617)

Even more interesting is that these suffixes -aa' and -taa' in written Aramaic actually end with a glottal stop, which either was never pronounced, only signifying the vowel -aa, or ceased being pronounced in the various Aramaic dialects, but in UA these suffixes often actually end with a glottal stop in Numic and Takic: Aramaic kookb-aa' 'star-the' > UA *kuppaa' > Serrano kupaa' 'to shine (as of the stars)' (1274)
Syriac ʼaamaqqət-aa 'lizard-the, n.f.' > UA *makkaCta 'horned toad': NP makaca'a 'horned toad' (1055)

Verbal Nouns are used in Hebrew and Arabic much more frequently than is customary in English. For example, for a narrative in Genesis 44:30-31, the King James English has five finite verbs: "when I **come** … and the lad be not with us; seeing that his life **is** bound up in the lad's life … when he **seeth** that the lad **is** not with us, he shall **die**." Yet the Hebrew has only one verb at the end "he'll die" but three verbal nouns and two verbless equative/copula constructions: "As/at my coming … and the lad not with us, his soul bound (adj) to his soul … as/at his seeing the lad not, he will die." Thus, Semitic often employs many verbal nouns more conveniently translated as verbs in English (Stubbs 1996c). So not surprisingly, we find many verbal nouns in UA: e.g., gəlom > UA kolom 'wrap' (934), Hebrew *ra'oot(-aa) 'seeing (it), to see (it), infinitive/ verbal noun' > UA *ta'uta 'find' (100), etc.

1.3 A Brief Introduction to Egyptian

As all living languages are always changing, Egyptian, over its 4,000-year history, also underwent stages of development from Old Egyptian (3100-2100 BC) to classical Middle Egyptian (2100-1600 BC), Late Egyptian (1600-600 BC), and then Demotic, beginning about 650BC and overlapping with and closely related to Coptic, which began being written with the Greek alphabet, and thus with vowels. This last stage of Egyptian, Coptic, continued in use more than 1,000 years, and is still the liturgical language of the Coptic Christian Church today (Allen 2010, 1). Each stage exhibited major and minor changes from its predecessor In fact, as details emerge, we should be able to identify the time or stage of the Egyptian from which the Uto-Aztecan infusion originated. Relevant to that eventuality, it is important to note that "Old Egyptian and Late Egyptian are historical phases of a single dialect, or closely related ones, likely from the north, while Middle Egyptian, chronologically between those two, represents a separate dialect, most likely southern in origin. In the history of the language, therefore, Middle Egyptian somewhat interrupts and obscures the presumably direct evolution of Old Egyptian into Late Egyptian" (Allen 2013, 4). The Egyptian element in Uto-Aztecan is closely associated with the Semitic-p; that and other factors suggest an Israelite group was likely the bearer of both. If Israelite, keep in mind where the Israelites were in Egypt? In the north, the Delta area. So when the UA Egyptian element exhibits both Old Egyptian and Late Egyptian features, such may be significant. My premature sense of the matter is that UA is mostly of that Old-plus-Late Egyptian duality. The prefixed articles of Late Egyptian (pV-, tV-, nV-) are in UA and at least two Old Egyptian features. Tarahumara's plural prefix *i- / *ip- matches Old Egyptian i(p…) as the beginning of plural demonstrative pronouns (these/those); see explanation at 121. A second matter of Old Egyptian in UA is that the UA stative suffix -i is in all eight branches of UA and is the oldest form (-i) of the stative suffix in Egyptian as well (see 116), though it later changed to -w in Middle Egyptian (Allen 2010, 206-7; Gardiner 1969, 234-8). UA has both -i and -wa, and some UA languages, like Hp and Tb, have both *-i-wa, as Egyptian sometimes shows both together also.

Two Egyptian stative/passive features are pervasive throughout Uto-Aztecan. In fact, one is called the old perfective from Old Egyptian and was also used as a stative, though the stative dimension continued through all stages of Egyptian even to Coptic. Stative structures present resulting states of verbs. For example, in English we have 'I do' (present) and 'I did' (past), but 'is done' expresses a present state resulting from a past action. Similarly, in Egyptian a final vowel -i at the end of verbs is the form of both the old perfective (past-tense like) and the stative (Allen 2000, 201; Gardiner 1969, 234-8). Likewise, every branch of Uto-Aztecan has exactly the same feature in which the final vowel of a transitive verb is changed to -i in order to signify the corresponding stative, intransitive, or passive verb. A few examples from 116:
Guarijio has transitive verbs ending in -a with corresponding intransitive verbs in -i (Miller 1996, 130):

Wr co'a 'put out fire'; Wr co'i 'be no fire';
Wr wela 'put upright/standing'; Wr weri 'be upright/standing';
Wr mo'a 'put pl obj's inside'; Wr mo'i 'enter, pl subj's';
Wr sa'wa 'cure s.o., alleviate s.th.'; Wr sa'wi 'be alleviated, go away';

Tarahumara also has such pairs of verbs' (Hilton 1993, 139):

Tr mana 'put, place, set'; Tr mani 'be (in/at a place), exist';
Tr bi'wá 'clean it'; Tr bi'wí 'be(come) clean';
Tr čiwá 'stick s.th., vt'; Tr čiwí 'be stuck, vi';

Classical Nahuatl also has such pairs of verbs (Sullivan 1988, 171):

CN tla-tema 'fill, place s.th.'; CN temi 'be full, be lying down';
CN tla-kotona 'break s.th.'; CN kotoni 'be broken';
CN tla-mana 'put s.th. on the floor'; CN mani 'be stretched out, extended';
CN tla-toma 'undo s.th.'; CN tomi 'be undone'; and so does Tbr:
Tbr towa 'leave s.th. behind, vt'; Tbr towi/tovi 'stay, remain, vi'.

In some UA languages, the final -i vowel is the perfective dimension of Egyptian's old perfective:

Cm -i 'completive suffix on verbs' (Charney 1993, 142-3).
TO -i 'perfective is marked by a final vowel change to -i' (Langacker 1977, 131);
Op -i 'perfective changes final -a to -i' (Shaul 2003, 25);
Eu -i 'the final stem vowel changes to final -i for the Eu preterite [past] in many, if not most Eu verbs, vs. Eu -a-n 'present indicative verb ending':

Eu hipra-n ‘watch over, care for’ vs. preterite: hipri ‘watched over, cared for’;
Eu maka-n ‘give’ vs. preterite: maki ‘gave’;
Eu taha-n ‘burn’ vs. preterite: tahi ‘burned’

The other Egyptian passive frequent in UA is the Egyptian suffix -w which aligns with UA *-wa ‘passive suffix’ and sometimes Egyptian -iw which matches UA *iwa. Remember that Egyptian shows only consonants, not vowels; thus, Egyptian -w and UA *-wa match well. See details at set number 117.

We must state clearly that Ancient Egyptian writing did not show vowels, only the consonants, though the consonants y and w sometimes represented the vowels i and u, respectively.

Reduplication was used in Older Egyptian for pluractional (more intense or frequentative) and imperfective verbs: wn ‘was’ vs. wnn ‘is, being, imperfective’; pr ‘came forth’ vs. prr ‘be coming forth’; and wn ‘walk’ and wnwn ‘walk to and fro’; from Egyptian fx ‘loosen’ are fxfx ‘totally release’ and fxx ‘loosen totally’; dbn ‘go around’ and dbndbn ‘go around and around’ (Bendjaballah and Reintges). Egyptian verbs with 5 consonants are always a reduplication of the 2nd and 3rd consonants: k’p ‘cover’ and k’p’p ‘cover up’; nhmhm from nhm ‘yell’; ḏdydy from ḏdy; sometimes a full reduplication: nḏdnḏd from nḏd (Allen 2010, 157). The most common kind of reduplication is doubling the 2nd of two consonants: wn > wnn; ẖzi > ẖzz; (Satzinger 2014).

Reduplication is also used in Uto-Aztecan for a similar array of uses. Langacker (1977, 128) notes that “virtually every UA language displays verbal reduplication of some kind, and in some cases a variety of patterns.” Reduplication can be found in UA to signify types of plurality, plural verbs, repetitive, continuative, distributive, durative, and intensive aspects of verbs, and for imperfective verb stems (Langacker 1977, 128-131).

A few other Egyptian grammatical structures are apparent in UA as well. The masculine pa-, feminine ta-, and plural na- article (‘the’) prefixes are found here and there as fossilized forms in a number of UA languages. See set number 369. The Egyptian structure *noun-pw* ‘he is a/the noun’ is found to a somewhat limited degree, but in several UA languages. See set 122.

Raymond Faulkner’s (1962) Middle Egyptian-English dictionary is the usual standard or the best available in English. However, Rainer Hannig’s (1995) Egyptian-German dictionary is three decades more recent, has more entries/words from more documents, and includes Late Egyptian and more semantic nuances, etcetera. They are the two Egyptian dictionaries regularly cited in this work, are among the best that are available, and are abbreviated in this work as Egyptian(F) and Egyptian(H), respectively. Coptic is a descendent of Egyptian and has the advantage of exhibiting vowels, some of them hinting at the more ancient vowels. Our primary source for Coptic terms is Jaroslav Cerny’s (1976) *Coptic Etymological Dictionary*. Other works, such as Antonio Loprieno’s (1995) *Ancient Egyptian: A Linguistic Introduction*, and James Allen’s (2013) *The Ancient Egyptian Language: An Historical Study,* and others listed in the Egyptian-and-Coptic bibliography are cited periodically as well.

1.4 Introduction to the Uto-Aztecan Languages, Branches, and Abbreviations

Uto-Aztecan (UA) is a language family of some 30 languages in the western United States and Mexico (map page 41). This book is based on the author's reference work—*Uto-Aztecan: A Comparative Vocabulary* (UACV 2011)—with some adjustments and many additions.

Any comparative work in Uto-Aztecan (UA) is a work in progress, not a finished product. The size of UA and the regular emergence of new materials guarantee that any comprehensive comparative effort is but a new horizon for viewing the next, but hardly finishable. Yet many a linguist's life work finds its final resting place in forgotten files due to (1) lack of time to finish it, despite the potential value to future researchers; (2) uncertainty about certain details, perhaps 3%, though the other 97% would have benefited all else studying the matter; and/or (3) not relishing the prospect that condemnations of the 3% may seem louder than commendations of the 97%. So let the latest from three decades of doing UA be made available lest it be lost to flame or file filler should I exit without warning. Publishing, despite its pretense of completion, is as often only the latest draft of endless endeavor. The original hope of finishing such an undertaking before one's own undertaking gradually gives way to time's reminder that no one gets everything right the first time, or even the last time in mortal exertions the magnitude of a language family, and our assumptions about when the last time might be are regularly erroneous, as we hardly get glimpses of our hourglasses. The tragic unpredictable passing of our mentor Wick Miller in May 1994 is an example.

Wick Miller was an example in several ways: he was open, cordial, and encouraging. He was not demeaningly critical, perhaps a tad animated at times, but generally friendly as a team-player in our cooperative progress in UA. As founder of the Friends of Uto-Aztecan organization, he was a friend to Uto-Aztecanists and devoted most of his life to UA. Miller's 1988 computerized database of potential cognate sets exemplifies his openness. He knew it was a compilation of rough-draft brainstorming in need of sorting, revision, etcetera, but he shared it openly—opening himself to an egoless vulnerability for the sake of progress, being more interested in our progress in knowledge than in his being right all the time. In that spirit is this work offered. Errors, loose ends, and uncertainties are certain, but some UA matters may remain unresolved even if one could spend three lifetimes on them, for many more than that have already been devoted to UA and to the reconstruction of Proto-Uto-Aztecan (PUA).

In the UA reconstructions I do not deal with vowel length, only vowel quality and consonants. Figuring out PUA vowel length may fill another lifetime, but not mine. Reduced consonant clusters with compensatory vowel lengthening underlie some long vowels in UA (CVCCV > CVVCV; see page 63), raising doubts about vowel length until the medial clusters are clarified. That and changing stress patterns—causing vowel lengthening with stress, or shortening or syncope without stress, in the various branches and languages through the layers of time—make the puzzle of PUA vowel-length quite unappealing to me, if not presently impractical. UACV also continues Miller's (1967, 1988) tradition of including sets found in only one branch. Rejecters (page 32) of Northern-Uto-Aztecan (NUA) and others of Southern Uto-Aztecan (SUA) make two-branch sets possibly from PUA, and one-branch sets are worth listing, since a reflex from another branch often appears later, though they can hardly be considered from PUA until such support surfaces. A few loans are listed if entering UA early enough to be found in multiple branches. As Miller (1988, 1) notes, "loans are of as much historical interest as inherited forms."

Edward Sapir (1913, 1915) was the first to apply the comparative method sufficient to establish Uto-Aztecan as a viable language family, after Buschmann, Brinton, Kroeber, and others helped lay the foundations for Uto-Aztecan studies, by identifying the three previously accepted branches—Shoshonean (NUA), Sonoran, and Aztecan. A five-letter surname that looms as large as Sapir's in UA contribution needs no further abbreviation, so sets from Sapir's founding works (1913, 1915) are cited as Sapir. A half century later, Voegelin, Voegelin, and Hale (1962) produced 171 cognate sets to further establish the sound correspondences and phonology of UA. Not long afterwards, Wick Miller (1967) published *Uto-Aztecan Cognate Sets*, containing 514 cognate sets. Miller continued working in UA and his last update (1988) of some 1185 potential cognate sets is herein abbreviated M88. Kenneth Hill (2006) has done much good work in sorting and revising M88, combining some sets, redistributing others, adding new reflexes to existing sets, and adding cognate sets of his own discovery, totaling more than 1200 sets. Hill's revision of M88 is herein abbreviated KH/M06. Besides the usual cognate collections, Kenneth Hill's *Serrano Dictionary* (in progress) has many comparative notes on other Takic languages, Tübatülabal, Hopi, and often Numic languages, i.e., most of NUA, so for sets with a Serrano reflex, it is another comparative resource for NUA, here cited as

KH.NUA. Stubbs (2011) then produced *Uto-Aztecan: A Comparative Vocabulary* with 2700 sets. Ronald Langacker (1976b, 1977a) and Jason Haugen (2008) have authored excellent books dealing with UA grammar. Through the 1980s and 1990s, Alexis Manaster Ramer (AMR) proved most prolific in his outpouring of insightful contributions to UA studies by means of more articles than are easily retrievable, until his illness. His and the works of Dakin, Campbell, Canger, Casad, Estrada Fernandez, Fowler, Heath, Jane Hill, Langacker, Lionnet, Munro, Shaul, Seiler, Steele, the Voegelins, Zamarron, and others—works both published and unpublished, like Kaufman's 1981 draft manuscript *Comparative Uto-Aztecan Phonology*—all constitute a corpus somewhat daunting for mere mortals to master.

As is the nature of research, this author's works also build on the good work of many others; thus, I am greatly indebted to the excellent output of scores of scholars before me. The 2011 work was finally made available after previous mentions (Stubbs 2000a, 2003) in spite of one lifetime being a few short of what is needed to do it. Though it doubles the number of previously known sets, the new are mostly smaller sets, as most of the larger ones, easier to find, have long been identified in previous works. Nevertheless, UACV (2011) adds some 1400 new UA cognate sets, adds new reflexes to previous sets, expands the number of branches for many sets, includes a phonology section treating features of UA comparative phonology (most of it here also), and provides discussion on salient questions in some sets, but mainly marshals an enlarged database and some new perspectives for furthering UA research.

In addition to strict cognate sets, Miller's works and UACV include (1) a score or more that may be early loans into UA and so are not cognate sets from PUA; and (2) another couple of hundred sets do not yet have the multi-branch representation needed to be properly counted as being from PUA. However, many times I and others, starting with single-branch sets, have found cognates in other branches that turn many single-branch sets into multi-branch sets. So single-branch sets are well worth listing in a comparative database designed to facilitate comparative research. (3) Many more UA words in a single UA language correspond well with Near-Eastern forms, which may count as Near-Eastern-UA cognates, but are obviously not UA cognate sets as at least two UA forms are needed to be a UA set.

Before diving into the minutia of comparative Uto-Aztecan (from which one may never return, if set on solving all), consider a bigger picture. As a relatively recent science, comparative linguistics first provided a flurry of impressive results in Indo-European. The more accurate recording of more Native American languages enabled similar bursts of impressive progress in Native Americana by the likes of Boas, Sapir, Kroeber, and Bloomfield. Their graduate students produced another generation or two of prominent comparativists; however, the number of those doing diachronic/comparative work seems on the wane, though a growing number of Mexican linguists are now passing U.S. output in comparative UA. The decrease in U.S. output may be partly due to: first, after the more obvious basics were established and caution resumed reign to rein in the macro-phyla momentum, progress slowed in the fine-tuning of the less obvious, which required deeper digging and other investments filling larger percentages of a lifetime. Second, the decrease in comparativists seems to have coincided with the Chomsky-initiated tidal wave of grammatical theories that swept the linguistic landscape and perhaps washed away a host of potential comparativists into the seeing of grammatical theory as the new wave to ride. I did theory too, before getting hooked on historical, for after a language family's more apparent tenets are established, further solutions can seem so deeply buried in data (data possibly unavailable) that comparative progress can turn into comparative composting; that is, progress often becomes mired in stewing over seeming unsolvables. Nevertheless, let an invitation be extended, that a few more linguists involve themselves in comparative research.

Returning to UA, the comparative effort (UACV) is assembled in hopes of helping Uto-Aztecanists postpone composting. The Near East tie answers many questions previously puzzling. If not accepted, then we can return to miring in our meager gains, loosely called "progress" for the sake of encouragement in a field where all but a handful have turned from comparative research to realms offering more hope of closure than reconstructing a large language family can possibly provide.

In short, the 2700 sets of these studies are intended to facilitate comparative research in UA and serve as a new plateau or expanded database. Adding to and refining this body of data will be an ongoing process by the author and any willing to join the cooperative effort. Other viable cognate sets, new reflexes to existing sets, enlightening discussion, and feedback are welcome, and will be credited to the contributor in future editions, and should be emailed to uanist@yahoo.com (Brian's email).

Table 1: The Preceding Cognate Collections in Chronological Order and Their Abbreviations
(Branch cognate collections are abbreviated as the initial(s) of author surname(s) dot branch; only the six in bold address the whole language family)

Sapir	**Sapir's "Southern Paiute and Nahuatl: a Study in Uto-Aztecan" (1913, 1915)**
VVH	**Voegelin, Voegelin, and Hale's *Typological and Comparative Grammar of UA* (1962)**
B.Tep	Burton Bascom's *Proto-Tepiman* (1965)
M67	**Wick Miller's *Uto-Aztecan Cognate Sets* (1967)**
BH.Cup	William Bright and Jane Hill's "The Linguistic History of the Cupeño" *IJAL* 33 (1967)
HH.Cup	Jane Hill and Kenneth Hill's "Stress in the Cupan Languages" *IJAL* 34 (1968)
I.Num	David Iannucci's *Numic Historical Phonology* (1972)
CL.Azt	Campbell and Langacker's "Proto-Aztecan Vowels," *IJAL* 44 (1978)
Fowler83	Catherine Fowler's "Lexical Clues to UA Prehistory" *IJAL* 49 (1983) and her fieldnotes
L.Son	Andrés Lionnet's *Relaciones Internas de la Rama Sonorense* (1985)
M88	**Wick Miller's 1988 Computerized Database of Uto-Aztecan Cognate Sets (1988)**
Munro.Cup	Pamelo Munro's "Stress and Vowel Length in Cupan Absolute Nouns" *IJAL* 56 (1990)
KH.NUA	Kenneth Hill's *Serrano Dictionary*, with comparative notes relevant to NUA (2001)
KH/M06	**Kenneth Hill's *Miller's Uto-Aztecan Cognate Sets*: *revised and expanded by KCH* (2006)**
UACV	**Brian Stubbs' *Uto-Aztecan: A Comparative Vocabulary* (2011)**

Table 2: The Uto-Aztecan Languages and Their Abbreviations

Mn	Mono	Hp	Hopi	Eu	Eudeve
NP	Northern Paiute	Tb	Tübatülabal	Op	Opata
		Ls	Luiseño	Tbr	Tubar
TSh	Tümpisha Shoshoni	Ca	Cahuilla	Yq	Yaqui
Sh	Shoshoni	Cp	Cupeño	AYq	Arizona Yaqui
WSh	Western Shoshoni	Sr	Serrano	My	Mayo
Cm	Comanche	Gb	Gabrielino	Wr	Guarijio
		Ktn	Kitanemuk	Tr	Tarahumara
Kw	Kawaiisu	TO	Tohono O'odham	WTr	Western Tr
Ch	Chemehuevi	UP	Upper Pima/Pima Alto	Cr	Cora
SP	Southern Paiute	Nv	Nevome	Wc	Huichol
WMU	White Mesa Ute	LP	Lower Pima/Pima Bajo	CN	Classical Nahuatl
NU	Northern/Uintah Ute	PYp	Pima de Yepáchic	Pl	Pipil
CU	Colorado Ute	PYc	Pima de Yécora	HN	Huastec Nahuatl
		NT	Northern Tepehuan		
		ST	Southern Tepehuan		

Table 3: The Branches of the Uto-Aztecan Language Family and Their Abbreviations

Mn	Western Numic (Num/WNum)	Hp	single-language branch	Eu	Opatan within TrC
NP	Western Numic	Tb	single-language branch	Op	Opatan within TrC
		Cp	Takic, Cupan (Cup within Tak)	Tbr	TaraCahitan (TrC)
TSh	Central Numic (Num/CNum)	Ca	Takic, Cupan (Cup within Tak)	Yq	Cahitan within TrC
Sh	Central Numic	Ls	Takic, Cupan (Cup within Tak)	AYq	Cahitan (Cah) in TrC
Cm	Central Numic	Sr	Takic (Tak)	My	Cahitan (Cah) in TrC
		Gb	Takic (Tak)	Wr	TaraCahitan (TrC)
Kw	Southern Numic (Num/SNum)	Ktn	Takic (Tak)	Tr	TaraCahitan (TrC)
Ch	Southern Numic	TO	Tepiman (Tep)	WTr	TaraCahitan (TrC)
SP	Southern Numic	Nv, UP	Tepiman (Tep)	Cr	Corachol (CrC)
WMU	Southern Numic	PYc	Tepiman (Tep)	Wc	Corachol (CrC)
NU	Southern Numic	PYp	Tepiman (Tep)	CN	Aztecan (Azt)
CU	Southern Numic	LP	Tepiman (Tep)	Pl	Aztecan (Azt)
		NT, ST	Tepiman (Tep)	HN	Aztecan (Azt)

The Branches of Uto-Aztecan

Miller (1984) and Cortina-Borja and Valiñas (1989) tallied the number of lexical agreements between UA languages using Swadesh's 100-word list, with 12 substitutions. Cortina-Borja and Valiñas added six languages to Miller's and analyzed the data differently. Table 4 presents most of those data:

Table 4: Lexical Correlations between Uto-Aztecan Languages

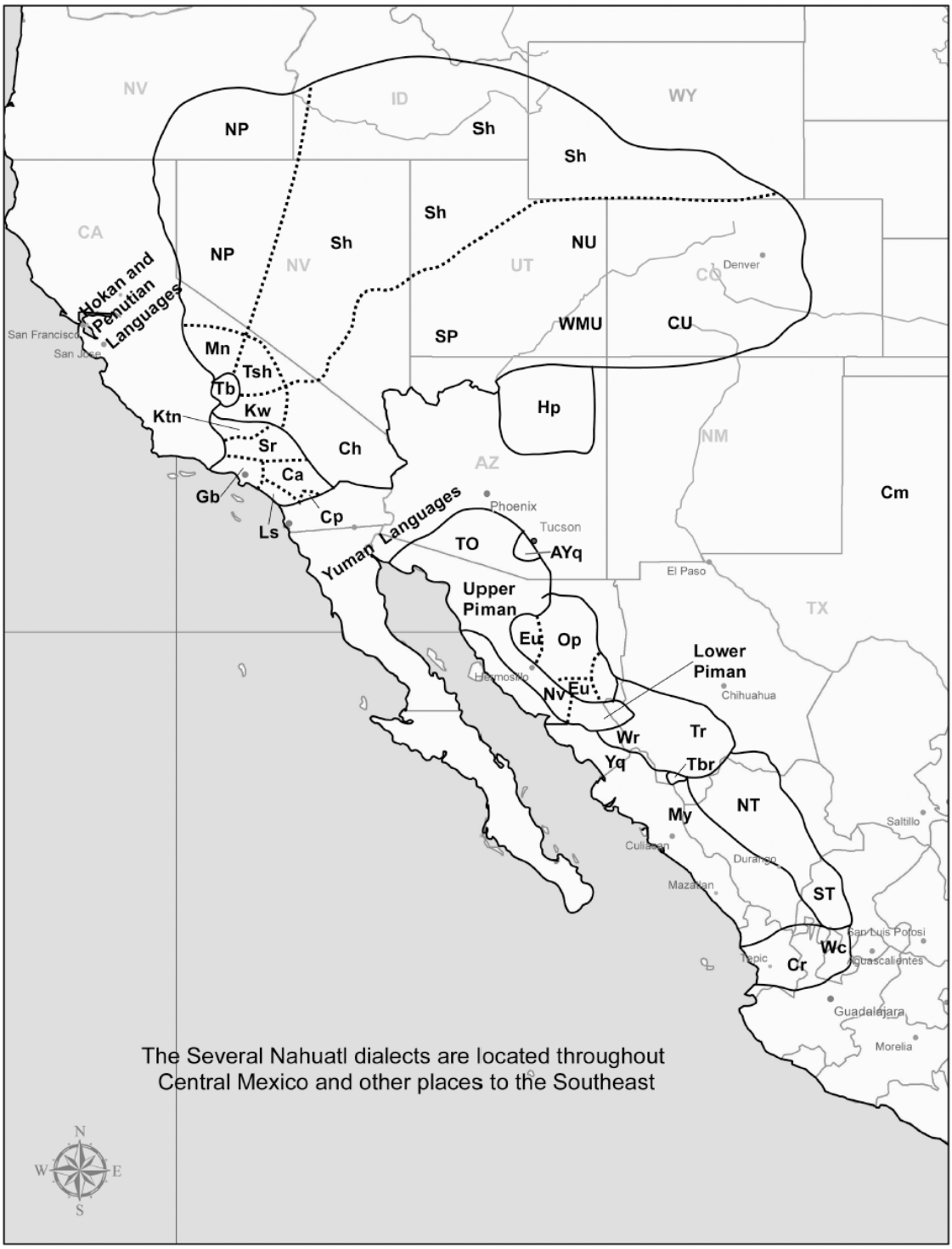
	Mn													
NP	77	NP												
TSh	59	58	TSh											
Sh	58	58	87	Sh										
Cm	57	58	79	88	Cm									
Kw	52	56	54	55	49	Kw								
Ch	50	55	61	58	54	75	Ch							
SP	53	58	62	62	59	79	86	SP						
CU	52	57	59	61	59	76	78	87	CU					
Tb	39	42	37	38	35	39	42	39	40	Tb				
Gb	26	26	26	26	23	24	27	26	27	40	Gb			
Sr	26	24	24	24	21	26	28	27	27	35	45	Sr		
Ca	29	27	27	27	24	27	31	31	29	38	42	50	Ca	
Cp	28	27	24	24	23	26	30	31	28	37	34	38	65	Cp
Ls	26	27	25	24	22	24	27	27	26	34	38	35	50	48 Ls
Hp	33	32	27	23	22	31	33	31	32	38	29	29	31	31 26 Hp
TO	23	26	25	25	23	26	28	28	30	35	25	27	31	28 25 32 TO
LP	24	26	24	24	23	24	26	26	27	35	24	27	30	27 24 35 85 LP
NT	25	28	26	26	23	27	28	30	29	37	26	30	32	29 26 33 79 79 NT
ST	22	24	23	23	21	24	24	26	27	33	26	28	31	28 25 30 73 75 82 ST
Wr	26	29	23	23	24	24	24	25	28	36	29	34	34	29 28 32 44 47 47 48 Wr
Tr	23	27	21	21	21	22	22	23	26	32	28	34	33	26 28 28 41 42 42 43 83 Tr
Op	26	29	21	20	20	20	26	24	23	33	26	31	33	29 24 33 40 44 40 39 55 54 Op
Eu	28	27	23	23	22	26	24	26	27	35	26	30	34	29 25 35 45 47 45 43 59 52 73 Eu
My	27	28	25	26	24	27	25	27	28	35	29	33	36	26 28 34 43 45 49 49 58 51 53 61 My
Yq	29	30	26	26	24	29	26	29	30	35	28	32	35	26 28 36 45 47 49 49 58 51 55 62 93 Yq
Tbr	28	27	27	28	27	28	27	30	31	33	24	28	29	26 23 30 40 41 46 43 48 44 42 51 51 53 Tbr
Wc	25	24	23	23	21	23	23	24	25	32	24	28	34	26 27 28 41 43 42 41 51 48 48 49 48 51 41 Wc
Cr	25	22	22	23	21	22	21	22	23	30	19	21	24	23 22 26 34 34 35 35 42 38 35 42 45 46 39 58 Cr
CN	18	18	16	16	14	16	15	16	16	24	20	22	23	19 19 24 29 29 30 29 32 33 39 40 38 39 36 39 37 CN
Te	19	18	16	16	14	17	15	16	17	25	20	22	24	20 19 24 30 30 30 29 32 34 38 40 38 39 35 37 35 85 Te
Za	17	17	15	15	13	16	16	17	18	26	21	20	24	20 19 24 31 31 32 31 29 33 35 39 37 38 35 35 33 80 85 Za
Pl	16	15	14	14	12	16	15	16	17	24	21	19	23	20 18 24 30 30 29 29 33 34 38 40 39 39 37 37 35 79 81 77

Many students of UA see a primary split between Northern Uto-Aztecan (NUA) and Southern Uto-Aztecan (SUA)(Heath 1977:27; Heath 1978:222; Langacker 1977:5; Langacker 1978:197, 269; Fowler 1983:234, Cortina-Borja and Valiñas 1989), yet a few reject NUA and Manaster Ramer (p.c.) rejects SUA. Jane Hill (2001a and b, 2010) also cites evidence for NUA vs. a lack of such for SUA. NUA does exhibit phonological innovations *-l- > n, *-c- > -y- (Manaster Ramer 1992b) and some morphological innovations (Heath 1977:1978), while SUA may exhibit a slightly closer lexical unity. (See discussion in Miller 1983, Goddard 1996, Cortina-Borja and Valiñas 1989.) But until comprehensive morphological studies clarify matters, objecting to the objectors of either half of UA may be premature. Accordingly, NUA has traditionally consisted of Numic, Takic, and two single-language branches: Tübatülabal and Hopi. SUA branches include Tepiman, Opatan, Tarahumaran, Cahitan, Tubar, Corachol, and Aztecan.

Numic (Num) has three subbranches. From southern California, Western Numic (WNum) spread northward along the California-Nevada border into Oregon and Idaho. Central Numic (CNum) spread northeastward through central Nevada, northwestern Utah, into Idaho, Wyoming, and onto the plains. Southern Numic (SNum) spread eastward into southern Nevada, northern Arizona, most of Utah, and the mountainous west half of Colorado. Western Numic includes Mono (Mn) and Northern Paiute (NP). To Central Numic belong Tumpisha Shoshoni (TSh), Shoshoni (Sh), and Comanche (Cm). Southern Numic

includes Kawaiisu (Kw), Chemehuevi (Ch), Southern Paiute (SP), Northern or Uintah Ute (NU), White Mesa Ute (WMU), and Colorado Ute (CU).

Map of the Uto-Aztecan Languages



The term Colorado Ute here replaces Southern Ute, since northern vs. southern is not a language division, but relocation options for the many dialects: e.g., the Uncompahgre Utes from southern Colorado went north to the Uintah-Ouray reserve, though their dialect and ties are closer to southern Colorado Ute; and White Mesa Ute (Stubbs 2011, 6-10), often labeled Southern Ute (because it in the south), retains features in NU and California SNum, but lost in Ignacio's Colorado Ute; and none of the three so-called Northern Ute dialects (two from Colorado) is recorded. So the northern-southern distinction is recent-geographic, not linguistic, and of at least five dialects, only Ignacio's is left in Colorado, thus, the term Colorado Ute.

The tabulations above show high correlations within each branch of Num (76-88), but less between the Num languages of different branches (49-62). Lamb (1958) and others have explained the Num languages' spread from the NUA homeland in southern California out into the Great Basin. The data show the inner-most language of each branch to be more closely related to the outer-most language of the same branch than to the closer neighboring Num languages of different branches. This pattern shows more diversity in Southern California between languages of differing branches only a few miles away vs. closer ties to tongues of the same branch 1,000 miles away. For example, TSh in Southern California is linguistically much closer to Sh (87) in Wyoming and Cm (79) on the plains, all three of Central Numic (CNum), than TSh is to nearby Mn (59), of Western Numic (WNum) and also in Southern California, or to nearby Kw (54), of Southern Numic (SNum) and also in Southern California. This greater diversity in the geographically limited Numic (and NUA) homeland speaks convincingly for a three-way Numic split in Southern California before spreading north, northeast, and eastward into the Great Basin. Shaul (2014) presents many details about the Numic spread, suggesting SNum spread first and WNum last.

Takic (Tak) has traditionally included the UA languages of Southern California, less Tübatülabal (Tb) and Numic languages. Within Tak is a tighter **Cupan** (Cup) group—Luiseño (Ls), Cahuilla (Ca), and Cupeño (Cp)—though the numbers above show Sr as close to Ca as Ls is to Ca. Serrano (Sr), Gabrielino (Gb), Kitanemuk (Ktn) and other now extinct languages together with Cupan constitute the Tak branch. Tak shows a much greater diversity than Numic. The numbers between the Tak pairs range from 35 to 50 (except for Ca-Cp 65) vs. Numic's numbers (49-88). Matters relating to that diversity have periodically caused the unity or exclusivity of the Tak branch to be questioned. **Californian** (Alexis Manaster Ramer 1992a; Kenneth Hill 1998) has been a contemplated union of Tb with Tak. Numbers as low as 34 between Gb and Cp, and 35 between Sr and Ls approximate several other 34's between Tak and non-Takic languages (Wr, Tr, Eu, Tb, Wc). Those inter-Tak numbers are no larger than the 35 through 40 that Tb shares with four Tak languages (Gb, Sr, Ca, Cp). Thus, the union of Tb and Tak into a Californian branch of NUA is reasonable enough in view of the above data, and questioning the traditional Tak unity merits consideration. Nevertheless, the author sees support for Tb's separation from Tak (see discussion under Tb), though hardly overwhelming. Kenneth Hill (2010, 1) also notes Tb's lack of initial ŋ and allowing ŋ only after vowels to be like the Numic languages and unlike the Tak languages' initial ŋ, and sees Tb's lenited absolutive suffix's (*-t > -l) similarity to the Cupan languages as likely coincidental.

Tübatulabal's (Tb) numbers with Num range from 35 to 42, with Tak they range from 34 to 40, and the Tb-Hp number is 38. The differences are so slight and the ranges so overlapping that Tb appears to be about equidistant lexically to other branches of NUA; thus, Tb seems to hold an especially central place in NUA. Yet viewing matters from the other directions, we see that Num is closer to Tb (35-42) than Num is to Tak (21-31) or to Hp (22-33), and that Hp is closer to Tb (38) than Hp to Tak (26-31) or Hp to Num (22-33). Furthermore, Cortina-Borja and Valiñas (1989, 235) see Tb to be slightly more closely associated with Hp and Num than with Tak. So it may be useful to retain Tb as a NUA branch for now. In any case, Tb and Hp both hold especially central positions, not only in NUA, but in UA generally: the Tb and Hp numbers with SUA branches are higher than other NUA languages with SUA languages, though Ca and Sr are not far off.

Hopi (Hp), presently spoken in northern Arizona, holds a unique position in UA—unique as a single-language branch of NUA and as the only known UA tribe to participate in the Ancient Pueblo tradition, along with three other language families (Kiowa-Tanoan, Keresan, and Zuni). Some measures put Hp closer to Tak (Cortina-Borja and Valiñas 1989, 228), while the numbers above show the closest Hp correlate to be Tb (38). Interestingly, however, Hp's next highest numbers are shared with Yq (36), Eu (35), LP (35), and My (34), all of SUA, after which several low 30's (30-33) are shared with some Tak and Numic languages, but also with some other Tepiman and Taracahitan languages. This fairly equal distancing with so many SUA and NUA languages further confirms Hp's unique place in UA.

Southern Uto-Aztecan (SUA) has consisted of Tepiman (Tep), Taracahitan (TrC), Corachol (CrC), and Aztecan (Azt), mostly from Arizona to Mexico City. Miller (1984) included Tep, TrC, and CrC in **Sonoran**; however, Tep and CrC in many respects differ more from TrC phonologically and grammatically than any two NUA branches; and below TrC is further divided. In contrast to earlier leanings toward a UA homeland in NUA areas, hints of greater diversity in SUA areas surface regularly, bringing Manaster Ramer, Jane Hill, and myself to deem SUA areas as more likely prospects for the UA homeland. One such hint is the close proximity of all UA reflexes for PUA *kw in the heart of SUA. Within miles of each other are Tep b, Cahitan bw, Tbr kw, and Tr w/b/ko (Stubbs 1995), while NUA reflects a nearly unanimous kw.

Tepiman (Tep) is so unique phonologically (*kw > b, *c > s, *s > h, *y > d, *w > g) among UA languages that it may merit distinction strictly on phonological grounds and grammar, regardless of word counts. Yet even word counts show a tight Tep entity with numbers from 73-85 between Tep languages, while 34-49 are the numbers between other Sonoran languages and the Tep languages, about the same as between NUA branches. That fact and the unique Tep phonology both recommend a separate Tep branch, here represented by Tohono O'odham (TO) in Arizona and Nevome (Nv) in Mexico, both of Upper Pima, while Lower Pima/Pima Bajo (LP) includes Pima de Yepachec (PYp) and Pima de Yécora (PYc). The Tepehuan languages include Northern Tepehuan (NT) and Southeastern Tepehuan (ST) in western Mexico.

Taracahitan (TrC) has been a term for the core Sonoran languages, i.e., Miller's Sonoran minus Tepiman and Corachol. However, Shaul's (2014) work shows a lack of evidence for a Taracahitan node and recommends four finer divisions for the geographic collection of languages in northwest Mexico between Tepiman and Corachol:

Opatan (Opn) is the closely related pair of Eudeve (Eu) and Opata (Op) or Tewima/Tegwima (Shaul, p.c.).
Tarahumaran (TrWr) includes the dialects of Tarahumara (Tr) and the dialects of Guarijio (Wr).
Cahitan (Cah) has Yaqui (Yq), Arizona Yaqui (AYq), and Mayo (My), with Yq and My sharing 93 items.
Tubar (Tbr) is its own branch. These four branches diverge nicely in reflecting Proto-Uto-Aztecan *kw: PUA *kw > Eu/Op *b, > Cahitan bw, > Tr/Wr *w, and > Tbr kw. Miller (1984) has called **Sonoran** a mesh of languages, which indeed it is with its overlapping and multi-directional influences, and with its intertwining phonological and lexical complexities. For example, **Tubar**, as a unique language in the center of the "Sonoran mesh/mess," is a difficult classification for two reasons: one, the lexical data are limited; two, the limited data, obtained shortly before extinction, show numerous loans and influences upon this small language surrounded by other larger UA languages. It is apparent that Tbr is in part a product of phonological influences from Tep and lexical loans from Cahitan and Tarahumaran, yet it is a kw-language, isolated geographically from the only other kw-languages of SUA: i.e., the Corachol and Aztecan branches. Classification by word counts may be misleading, due to lexical influences upon the small Tbr-speaking population surrounded by larger numbers of Tep (NT) and Tr, Wr, My, and Yq speakers. Phonological influences from neighboring Tep languages upon Tbr include some *s > h, some *w > g, and initial *p > w (Stubbs 2000b). Tbr's lexical position may be more due to loans and meshing movements than to genetic position. Thus, I previously hesitated to call Tbr a single-language branch—because, unlike Hopi's clear distinctions and massive database, Tbr has neither—yet I must concede that the meagerly documented Tbr hardly fits elsewhere and so should be its own branch. However, the work of rewriting and dividing all the TrC notations will happen in a future edition.

Corachol (CrC) consists of Cora (Cr) and Huichol (Wc), showing a closer lexical relationship to each other (58) than to any other UA languages, but phonologically they form a pair and align better with Aztecan than with the old Sonoran grouping. They share an innovation with Aztecan of *p > h/ø and a retention of *kw, neither of which is prevalent in Tep or TrC.

The **Aztecan** (Azt) branch consists of the many dialects related to Classical Nahuatl. Cortina-Borja and Valiñas (1989) include nine in their classification study. Suarez' (1986) admirable comparative study of Nahua dialects merits more use. Of interest is that Azt yields numbers of 30-40 with other SUA languages, but only teens to 20 with NUA languages, except with Tb, Hp, and Ca, with which languages the Aztecan numbers are 23-26.

1.41 Primary Sources for the Lexical Data (in addition to the cognate collections):

Mono (Mn):

Bethel, Rosalie, Paul V. Kroskrity, Christopher Loether, and Gregory A. Reinhardt. *A Dictionary of Western Mono*. 2nd ed. 1993. — Mn

Northern Paiute (NP):

Bednark, James. Project director. *Paiute-English, English-Paiute Dictionary*. A publication of the Yerington Paiute Tribe, 1987. — NP

Liljeblad, Sven, Catherine S. Fowler, Glenda Powell, and Harold Able. *The Northern Paiute-Bannock Language: A Dictionary*. Salt Lake City: University of Utah Press, 2011. — NP(B)

Tümpisa Shoshone / Panamint (TSh):

Dayley, Jon P. *Tümpisa (Panamint) Shoshone Dictionary*. Berkeley: U of C Press, 1989. — TSh

Shoshone (Sh):

Miller, Wick R. *Newe Natekwinappeh: Shoshone Stories and Dictionary*. University of Utah Anthropological Papers, number 94. Jesse D. Jennings, ed. Salt Lake City: University of Utah Press, 1972. — Sh

Crapo, Richley H. *Big Smokey Valley Shoshoni*. Desert Research Insitute Publications in the Social Sciences, number 10. Don D. Fowler, ed. 1976. — Sh(C)

Gould, Drusilla, and Christopher Loether. *An Introduction to the Shoshoni Language*. Salt Lake City: University of Utah Press, 2002. — Sh(GL)

Crum, Beverly, and Jon Dayley. 1993. *Western Shoshoni Grammar*. Boise State University: Occasional Papers and Monographs in Cultural Anthropology and Linguistics, no. 1. — WSh

Comanche (Cm):

Robinson, Lila Wistrand, and James Armagost. *Comanche Dictionary and Grammar*. Summer Institute of Linguistics and the University of Texas at Arlington Publications in linguistics, number 92. 1990. — Cm

Kawaiisu (Kw):

Zigmund, Maurice L., Curtis G. Booth, and Pamela Munro. *Kawaiisu: A Grammar and Dictionary with Texts*. Berkeley: University of California Press, 1991. — Kw

Chemehuevi (Ch):

Press, Margaret L. *A Grammar and Lexicon*. Berkeley: University of California Press, 1979. — Ch

Laird, Carobeth. *The Chemehuevis*. Banning: Malki Museum Press, 1976. — Ch(L)

Southern Paiute (SP):

Sapir, Edward. 1931. Southern Paiute Dictionary. Proceedings of the American Academy of Arts and Sciences 65:537-730. — SP

White Mesa Ute (WMU):

Stubbs, Brian, Mary Jane Yazzie, Aldean Ketchum, Loretta Posey. 2011. *White Mesa Ute: A Dictionary and Lessons*. Preliminary edition. — WMU

Colorado Ute (CU):

Givon, Talmy, ed., and Southern Ute Tribe. 1979. *Ute Dictionary*. Ignacio, CO: Ute Press. — CU

Jean O. Charney. 1996. *Ute Dictionary*. Ignacio, Colorado: Southern Ute Indian Tribe. — CU(Ch)

Hopi (Hp):

Hill, Kenneth C., and the Hopi Dictionary Project. 1998. *Hopi Dictionary*: *Hopìikwa Laváytutuveni*. Tucson: The University of Arizona Press. — Hp

Voegelin, C. F., and F. M. Voegelin. 1957. *Hopi Domains: A Lexical Approach to the Problem of Selection.* IJAL Memoir 14. — Hp(V)

Seaman, P. David. 1985. *Hopi Dictionary*. NAU Anthropological Paper, no. 2. Flagstaff: NAU. — Hp(S)

Tübatulabal (Tb):

Voegelin, Charles F. 1958, "A Working Dictionary of Tübatulabal," *IJAL* 24:221-28. — Tb

Munro, Pamela, and William E. Mace. 1995. *A New Tübatulabal Dictionary*. (revised preliminary version) UCLA. — Tb(M)

Hill, Kenneth C. 2010. Tübatulabal Dictionary. Draft manuscript. — Tb(H)

Serrano (Sr):

Hill, Kenneth C. *Serrano Dictionary*. In preparation, 2001 edition. — Sr

Kitanemuk (Ktn)

Anderton, Alice Jeanne. 1988. *The Language of the Kitanemuks of California*. Ph.D. dissertation, UCLA. — Ktn

Cahuilla (Ca):

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

Cupeño (Cp):

Hill, Jane H., and Rosinda Nolasquez. 1973. *Mulu'wetam: the First People: Cupeño Oral History and Language.* Banning: Malki Museum Press. — Cp

Hill, Jane H. 2005. *A Grammar of Cupeño*. UCPL 136. Berkeley and Los Angeles: University of California Press.

Luiseño (Ls):

Bright, William. 1968. *A Luiseño Dictionary*. UCPL51. Berkeley: University of California Press. — Ls

Elliott, Eric B. 1999. *Dictionary of Rincon Luiseño*. Ph.D. Dissertation. San Diego: University of California at San Diego. — Ls(E)

Tohono O'odham (TO):

Saxton, Dean, Lucille Saxton, and Susie Enos. *Dictionary: Tohono O'odham/Pima to English, English to Tohono O'odham/Pima,* 2nd ed. 1983. R.L. Cherry, ed. Tucson: University of Arizon Press. — TO

Saxton, Dean and Lucille. *Dictionary: O'odham Milgaan, English Papago/Pima*. 1969. Tucson: University of Arizona Press.

Mathiot, Madeleine. *A Dictionary of Papago Usage*. 1976. Tucson: University of Arizona Press. — TO(M)

Upper Pima (UP) and **Nevome (Nv)**:

Bascom, Burton W. *Proto Tepiman*. Ann Arbor: University Microfilms, Inc., 1965. — UP

Munro, Pamela, et al. Akimel O'odham. In Preparation. — UP(M)

Pennington, Campbell W. ed. *Vocabulario en la Lengua Nevome: The Pima Bajo of Central Sonora, Mexico*. vol. 2. Salt Lake City: University of Utah Press, 1979. — Nv

Pima Bajo (PB) or **Lower Pima (LP):**

Bascom, Burton W. *Proto Tepiman*. 1965. PhD Dissertation: University of Washington. Printed at Ann Arbor: University Microfilms, Inc. — LP(B)

Escalante, Roberto, and Zarina Estrada Fernandez. 1993. *Textos y Gramática del Pima Bajo*. Hermosillo, Mexico: Departamento de Letras y Lingüistica, Universidad de Sonora. — LP(EF)

Shaul, David Leedam. "A Sketch of the Structure of Oob No'ok (Mountain Pima)" in *Anthropological Linguistics*. vol. 36, number 3, Fall 1994. — PYp

Barragan, Luis M. fieldnotes on the Pima Bajo dialect of Yécora, Sonora, Mexico. — PYc

Northern Tepehuan (NT):

Bascom, Burton. *Northern Tepehuan Dictionary*. In preparation. — NT

Bascom, Burton. *Proto Tepiman*. Ann Arbor: UMI, 1965.

Southern Tepehuan (ST):

Willett, Thomas. *Southeastern Tepehuan Dictionary*. 2005. Preliminary edition. — ST

Bascom, Burton. *Proto Tepiman*. 1965. Ann Arbor: UMI.

Eudeve (Eu):

Anonymous. 1981. *Arte y Vocabulario de la Lengua Dohema, Heve, o Eudeva*. Ed.Campbell W. Pennington. Mexico City: Mexico, Instituto de Investigaciones Filológicas, Universidad Nacional Autónoma de Mexico. — Eu

Lionnet, Andrés. 1986. *Un Idioma Extinto de Sonora: El Eudeve*. Mexico City: Universidad Nacional Autónoma de México. — Eu(L)

Tubar (Tbr):

Lionnet, Andrés. 1978. *El Idioma Tubar y Los Tubares*. Universidad Iberoamericana. — Tbr

Yaqui (Yq):

Estrada Fernandez, Zarina, et al. 2004. *Diccionario Yaqui-Español y Textos.* — Yq

Johnson, Jean B. *El Idioma Yaqui*. 1962. — Yq(J)

Arizona Yaqui (AYq):

Molina, Felipe S., and David Leedom Shaul. 1993. *A Concise Yoeme and English Dictionary*. Tucson: Tucson Unified School District. — AYq

Mayo (My):

Collard, Howard, and Elisabeth Scott Collard. 1984. *Vocabulario Mayo*. Serie de vocabularios indígenas, no. 6. Mexico, D.F.: Instituto Lingüistico de Verano. — My

Guarijio (Wr):

Miller, Wick R. *Guarijio: Gramática, Textos, y Vocabulario*. 1996. — Wr

Medina Murillo, Ana Aurora. 2012. *Diccionario Léxico-Morfológico del Guarijío*. — Wr(M)

Tarahumara (Tr):

Brambila, David. *Diccionario Raramuri-Castellano*. 1976. La Obra Nacional de la Buena Prensa. — Tr

Hilton, K. Simón. *Diccionario Tarahumara de Samachique*. 2nd ed. Serie de vocabularios indígenas, no. 101. Tuscon: Instituto Lingüistico de Verano, 1993. — Tr(H)

Lionnet, Andrés. 1972. *Los Elementos de la lengua Tarahumar.* México: UNAM. — Tr(L)

Cora (Cr):
McMahon, Ambrosio, and Maria Aiton de McMahon. *Cora y Español*. 1959. Cr
Serie de Vocabularios Indigenas, no. 2. Mexico City: Instituto Lingüistic de Verano.
Huichol (Wc):
Grimes, José E., Pedro de la Cruz Avila, José Carrillo Vicente, Filiberto Díaz, Román Díaz, Wc
Antonio de la Rosa, and Toribio Rentería. *El Huichol: Apuntes Sobre el Léxico*. 1981.
New York: Cornell University.
Classical Nahuatl (CN) and other **Aztecan Dialects**
Campbell, R. Joe. *Draft Lexicon of Molina and Florentine Codex Vocabulary*, 2006. CN(RJC)
Karttunen, Frances. 1983. *An Analytical Dictionary of Nahuatl*. Austin: University of Texas Press. CN
Simeon, Remi. *Diccionario de la Lengua Nahuatl or Mexicana*. Mexico City: CN(S)
Siglo Veintiuno, 1977. First published in French in Paris: la imprimerie nationale, 1885.
Campbell, Lyle. 1985. *The Pipil Language of El Salvador*. New York: Mouton Publishers. Pl
Campbell, Lyle, and Ronald W. Langacker. 1978. Proto-Aztecan Vowels. CL.Azt
IJAL 44: 85-102, 197-210, 262-279.

1.42 Sound Correspondences and Comparative Phonology of Uto-Aztecan

Some Proto-Uto-Aztecan (PUA) consonants attract debate—PUA *l / *r, and *ŋ vs. *n—while the more secure PUA consonants include *p, *t, *k, *kw, *', *h, *s, *c, *m, *n, *l, *w, and *y. Exceptions for *kw before round vowels (*kwo, *kwu) are discussed in Stubbs 1995. Some PUA *t palatalized to c/č in time to participate in the Tepiman sound change *c > s, and are thus mistaken for PUA *c (Stubbs 2000a). The PUA vowels are *i, *a, *u, *o, and *ï. An oversimplified portrayal of the consonant correspondences follows (per Sapir 1913-14, VVH 1962, Miller 1967, 5, Steele 1979, Manaster Ramer 1992b, Stubbs 2003):

Table 5: Consonant Sound Correspondences (mostly initial position)

PUA	*p	*t	*k	*kw	*m	*n	*c	*s	*w	*y	*'	*h
Num	p	t	k	kw	m,ŋw, w	n	c,-y-	s	w	y	'	h
Hp	p	t	k,q	kw	m	n	c,-y-	s	w,l	y	'	h
Tb	p	t	h,k	w	m	n	c,-y-	š	w	y	'	h
Sr	p	t	k,q	kw	m	n	c,-y-	ş,h	w	y	'	h
Ca	p	t	k,q	kw,w	m	n	c,-y-	s	w	y	'	h
Ls	p	t	k,q	kw	m	n	c,-y-	s,š	w	y	'	h
Tep	w,v,-p-	t,c	k	b	m	n,ñ	s, š	h,ø	g	d̰,j	ø,'	',h
Eu	b,p	t	k	b	m	n	c, č	s	w	d	ø,'	h
Tr,Wr	b,p	t	k	w,-'w-	m	n	c, č	s	w	y	ø,',h	h
Yq,My	b,p	t	k	bw	m	n	c, č	s	w	y	'	h
Tbr	w,-p-	t	k	kw	m	n	c, č	s,h	mw, ñ	y,ñ	ø,h	h
Cr	h	t	k,č	kw,čw	m,mw	n	c, č	s	w	y	'	'
Wc	h	t	k	kw	m	n	c, č	s,z	w	y	ø	ø
CN	ø,p	t	k	kw	m	n	c, č	s,š	w	y	ø, ',h	ø

Table 6: UA Vowel Correspondences and medial *l (Sapir 1913-14, VVH 1962, Miller 1967, Bright and Hill 1967, Langacker 1970, Munro 1990, Stubbs 2003):

PUA	*i	*a	*u	*o	*ï	*l	
Num	i	a	u	o	ï	n	
Hp	i	a	o	ö	ï	n,l	
Tb	i	a	u	o	ï	n	
Sr	i	a	u	ö	ï	n,r	
Ca	i	a	u	i	e	n,l	
Cp	i	a	u	i	ɛ/ə	n,l	
Ls	i	a	u	e(i)	o(u)	n,l	
Gb	i,e	a	u,o	e,o	o	n	(Kenneth Hill, p.c. 2002)
Tep	i	a	u	o	ï	l,d̰,r	
Tr,Wr	i	a	u,o	o	e,i	l,r	
TrC	i	a	u	o	e	l,r	
CrC	i	a	ï	u	e	l,r	
CN	i	a	i	o	e	l	

1.43 Consonant Clusters in Proto-Uto-Aztecan Stems

The traditionally accepted form for UA stems has been CVCV (C = consonant; V = vowel). While many stems undoubtedly align with CVCV, evidence is emerging to suggest that many Proto-Uto-Aztecan (PUA) stems contained consonant clusters not previously recognized: CVCCV and others. First of all, Manaster Ramer and Blight (1993b) and Manaster Ramer (1997) have noted evidence for reconstructing clusters for several etyma, such as *kapsi 'thigh' vs. *kasi (Miller 1967). Sometimes those clusters survive in only one language. Second, we see frequent evidence in UA that vowel syncopation (the deletion of an internal vowel as a common phenomenon in UA) creates additional clusters, and that even those later clusters are reduced quite quickly (CVCVCV > CVCCV > CVCV), suggesting that most UA languages do not maintain consonant clusters well. Third, the difficulties found in the correspondences of the medial consonants in UA are likely the result of reductions of previous clusters. In Miller (1967, 5), one can see in table 5 above that the initial consonant correspondences are fairly clear and consistent, while the medial consonant correspondences are more varied and less consistent. Yet many medial consonants being reduced consonant clusters may explain some of the variety and difficulty, if not most of it. If UA had 13 proto-consonants (also debatable), then 169 possible combinations (13 x 13) exist. Perhaps some of those clusters reduced to the velar nasal (ŋ) in some languages, others to a glottal stop (') in some languages, etc. A certain cluster might reduce five different ways among the branches of UA. Complications of clusters may underlie the medial consonant difficulties, which Uto-Aztecanists have only begun to unravel. The UA medial consonant correspondences as listed in Miller (1967, 5) illustrate the confusion:

Table 7: Some of the Medial Consonant Correspondences depicted in Miller (1967, 5)

	*-p-	*-t-	*-k-	*-k^w-	*-s-	*-m-	*-n-	*-w-	*-y-	*-'-	*-h-	*-l-
SP	v,hp,mp	r,ht,c	x,hk,ŋk,k^w	k^w,hk^w,ŋk^w	s,š,ø	ŋw,m	n,hn,ŋ	----	y	----	ø,h	n
Tb	p, b, hp	l, t, d	h,g,hk	----	š	w, m	n, ŋ	w	y	'	', ø	n
Ca	v, p	l, t, š	x, k, q	----	s, š	w, m	n, ŋ	w	y	'	h	l, n
Sr	v, p	r, t, ç	k, q	----	h, ş	m	n, ñ, ŋ	----	y	'	h, ø	r, n
Hp	v, p	r, l, t	k, q	k^w	s	m	n, hn, ŋ	w, l	y	ø	ø	n
TO	v, p	ḍ, t, c	k	b	h	m	n, ñ	g	d	'	', ø	l, ḍ
Tr	b, p,'w	r, l, t	k	w	s	m	n	w	y	h,'	----	l, r
My	b, p	t	k	b^w	s	m	n	w,b	y	'	h	l, r
CN	p, hp	t	k	k^w	s, š	m,-n	n	w	---	---	ø	l

Other medials not listed above include some Num m : NUA ŋ : SUA n (see salt 280, lung 281, husband 284). For those 3 and other cognate sets, PUA *ŋ > SUA n (some say) and PUA *n > SUA l, and that PUA had no liquids; others see the change in the other direction: PUA *n > NUA ŋ and PUA *l > NUA n. The medial liquid(s) (l/r) await explanation, but see 7.9. On the positive side, some progress has been made since Miller 1967: AMR (1992a) clarified PUA non-initial *-c- > *-y- in NUA and other medial matters cited in coming pages. This work also clarifies matters for Tr initial t vs. r (6.1), TrC b vs. p (6.2) both previously thought from PUA *p, the Tb k vs. h < PUA *k dichotomy (6.3), and Hopi l vs. w before low vowels (6.5). Semitic explains Takic *qa vs. *ka syllables and other matters may suggest additional PUA consonants. Of interest is a general lenition shift of consonants in Tep: *t > c (before high Vs), *c > s, *s > h, *h > ', *' > ø.

PUA	*p	*w	*y	*t	*c	*s	*h	*'
Tepiman	w	g	d	t/c	s	h	'	ø

Phonemic Frequencies in Uto-Aztecan

The phonological frequencies of initial syllables in Miller 1988 (M88) were calculated. The exact numbers of initial syllables among UA cognate sets are subject to adjustment, yet those in M88 are reasonably proportionate and available for quick inspection, until this work's sets settle sufficiently for counting. The first column is the number of sets beginning with glottal stop-vowel or initial vowel. (Enough UA languages require glottal stop before otherwise initial vowels that Miller (M88), Ken Hill (KH/M06), and others deem the same for PUA.) The other columns are sets beginning with the specified CV combination. Totals of the lines (vowel totals) are to the right; and totals of the columns (consonant totals) are below. The total number of sets in M88 is 1185, the total both of the rows and of the columns.

Table 8: Initial Syllable Frequencies

	’	c	h	k	kw	m	n	p	s	t	w	y	totals
a	39	18	17	43	15	43	38	64	29	48	27	28	409
i	11	23	7	10	16	6	2	28	18	1	18	--	140
ï	19	15	9	17	6	11	15	17	22	54	12	19	216
o	27	20	8	38	--	11	12	26	15	26	14	10	207
u	9	20	21	37	--	23	5	23	21	24	2	28	213
	105	96	62	145	37	94	72	158	105	153	73	85	1185

Some observations of interest and relevant to the phonological discussions include:
(1) The vowel ***a*** is about twice as frequent as other vowels.
(2) The syllables kwo, kwu, and yi are absent. Yet there are 38 ko and 37 ku syllables, respectively, vs. 10 ki and 17 kï. The ko/ku are nearly as many as the 43 ka, which vowel, across the board, is normally twice what others are. The increase in ko/ku syllables is probably related to the absence of kwo/kwu syllables, though the same cannot be said for an increase in ***i*** in absence of yi.
(3) Among all tV syllables, only one ***ti*** syllable (M88-ti1 ‘man’) existed until Ken Hill redistributed it (to KH/M06-ci24, tu10, tï9), so now no ***ti*** syllables exist (in KH/M06) vs. 48 ta, 54 tï, 26 to, and 24 tu. In contrast, the number of ***ci*** syllables (23) is larger than other ***cV*** syllables (18, 15, 20, 20) in spite of the fact that ***i*** is the least frequent vowel: i.e., 140 ***i*** vs. 409 for ***a*** and vs. 200-plus for the other three vowels. All this suggests that many apparent ***ci*** may be from an earlier ****ti***.

Final Features as Evidence of Earlier Consonant Clusters

Final features suggest the presence or absence of internal consonant clusters. Final features have been discussed by several (Sapir 1914, 451-2; Sapir 1930, 62-65; Irving Miller 1982; Wick Miller 1983; Manaster Ramer 1992b, 2004) and involve the presence or absence of underlying final consonants, whose presence causes consonant cluster behavior at morpheme boundaries. These final features are found in much of NUA, most notably and clearly in Num, but also in Tak and Tb. Sapir (1930) found that Num stems had one of three final features: gemination (-”) or (-C) causes a doubling of the next consonant (> -CC-); nasalization (-N) adds a nasal dimension to precede the next consonant (> -NC-); or spirantization appears to be a lack of a final underlying consonant, such that the next morpheme’s initial consonant appears as it typically does between vowels (*-k- > -x-/-ġ-, *-t- > -r-/-l-/-d-, *-p- > -v-/-b-). Miller, Elzinga, and McLaughlin (2005) provide some TSh examples with the post-position -pa’a ‘on’ after spirantization (*naka-pa’a > naġa-va’a ‘bighorn sheep-on’), gemination (*tuaC-pa’a > tuappa’a ‘son-on’), and nasalization (*pïyïN-pa’a > pïyïmba’a ‘duck-on’). The variety of absolutive suffixes (*-ta > -t(a), -l(a), etcetera) mostly in NUA, also leaves hints of the existence and type of final consonant (Sapir 1914, 451; Manaster Ramer 1992b; 2004). For example, in Tak and Tb, an absolutive suffix -l means the stem ended with a vowel and *V-ta became V-la between vowels (*V-ta > V-la > V-l), whereas absolutive suffix -t suggests the noun stem had an underlying final consonant no longer obvious (*VC-ta > V-t). The peculiar Ls -la is treated at 6.4.

Intervocalic *-t- vs. *-tt-/*-Ct- Clusters, and Many NUA -c- < *-tt-/*-Ct-

Intervocalic *-t- usually goes to -r- or -d- in Num and to -l- in Cupan and Tb (Sapir 1914, 451; Manaster-Ramer 1992b). So when we see intervocalic -t- in those languages, it is usually due to an underlying geminated *-tt- or to a cluster approximating *-Ct- that behaves much like *-tt-. Sapir (1914, 452) also noticed that Num geminated -tt- corresponds to Tak and Tb -t-. Later, Alexis Manaster Ramer (1992a) demonstrated PUA medial *-c- > -y- in NUA, and accordingly suggests the various NUA medial -c- are from other sources than PUA *-c-, unless *-cc- is geminated or clustered. Thus, the source of NUA -c- is often a palatalized *-tt- or *-Ct-, especially adjacent to high vowels. (See 534, 969, 1445.) In fact, Sapir (1914, 445) noted that many UA *c* may be from syncopated *ti. I would add that many, if not more, are also from non-syncopated *ti / *-tti or *tï / *-ttï. In the data below, note the frequency of *t-/*-tt-/*-Ct- > c/-c-, often adjacent to high vowels, but not always.

1368 UA *attip-na 'good': CU ’atti 'good'; SP ’attïN ‘good’; Cp á’či’a 'good'; Ca áča’e 'good, fine, well, very’; Hp -’civa 'accord with’, Hp a’civa ‘behave as expected, do what one can with one’s personal resources and limitations’; Hp àacipna/a’cipna 'do as expected’. Note that Hp a’cipna and Cp á’či’a are identical in five segments (a’ci . . . a) except for a consonant cluster in Hp that aligns with a glottal stop in Cp, and both align with SNum (CU, SP) *’atti, suggesting *-tti- > -ci-. [Syriac ’aṭib / ’aṭ(’)ib ‘do good, treat well’ (causative of ṭ’b; Hebrew haṭṭiib ‘do well’]
UACV-124 ***paCti’a** ‘bat’ > *paci, *pali, etc. NP pidahana'a ‘bat’ actually shows -t-. See discussion at ‘bat’.
534 UA *paCtï ‘daughter’ > Num *pattï ‘daughter’, but pacï in SP and CU. [Hebrew **batt** ‘daughter’ (< *bant / bint)]
1227 UA *patta/*patti ‘flat’ > *paci.

More Examples of Proto-Uto-Aztecan *t/*tt > c and in time for *c > s in Tepiman

We not only see *t or *-tt- > -c-, but sometimes that change was early enough to undergo the Tepiman sound change of *c > s, such that **some PUA *t / -Ct- > c > Tep s**:
437 UA *matta > *maca/i ‘tick’: NP madabi (< *matapi); Kw muu’maa-ci; CU mata-ci (< *matta-ci); Ch mata-vi (<*matta-pi); Cp máči-ly; Ca máči-l; Ls ’amáča; Sr maca-c; Hp màaca; TO maamş; Wr macá; Tr mačá; Wc mate. Takic, Hp, and TrC show -c- (in both NUA and SUA), but Num and Wc show -t-/-tt- (again in both NUA and SUA), yet TO has ş (< c < *-tt-). [Egyptian **mht** ‘an insect’]
1464 UA *takola/*takula 'round, (en)circle': Eu takóris 'circle'; AYq tekolai ‘round’; My tékolai ‘redondo’; Sr ta’kï’q ‘be round, circular’. From the first vowel *a* (Eu, Sr), note some raised vowels (AYq, My). If raised a little more, then:
1464 UA *tikola > *cikola (> Tep *sikola/i) '(a)round': TO sikod ‘round, circumscribed’; TO sikol ‘circular, round’; NT šikóra; NT šikóóraka; ST šikar. Ken Hill adds Cahita číkola ‘alrededor’ exactly the link theorized.
638 NUA *tïkïya 'deer’ is found in most Numic languages and Tb, yet compare
638 SUA *ciki 'white-tailed deer' (Tep *siki < *ciki < *tiki): TO siiki 'white-tailed deer'; PYp siiki 'white-tailed deer’
UACV-108 *paNtuC > *paicu’ 'badger': ST vaisïly ‘tejón’; Cr haihcï(-te) ‘tejón(es)’; and Wc háisï ‘tejón’ all match *paicV (*p > ST v; *p > CrC h). CN peeso’-tli ‘badger’ also parallels ST vaisïly and Wc háisï, all pointing to s.th. near *paicu, though CN s should be c and CN has p while Cr and Wc have h, so CN may be from an early loan. Most forms suggest an originally round final vowel, but puzzles remain. Wr pincúri ‘tejón’ and Tr batúwi ‘tejón’ must be included and may be key to the cluster. Wr pincúri shows *-nc-, a nasal-alveolar cluster, and the dipthong **ai* > *i* instead of > *e*, like CN. ST *s* agrees nicely with the *c* of CrC and Wr. In light of many PUA *t > c adjacent to high vowels and in light of Tr’s t and in light of Cr, Wr, Tr showing PUA *u after the t/c, something like *paNtu may explain all forms, especially since other examples of UA vowels before alveolars tending toward *i* would explain *paicu (< *pantu). In addition, Wr’s nasal in the cluster may explain such a cluster > -c- in most languages, for this may have been a different kind of cluster than in ‘bat’, resulting in Cr -c- vs. Cr -hc- for ‘badger’. This is a 4th example of *t > c > Tep s.
UACV-124 ***paCti’a** ‘bat’ note the -pisa of PYp ho’opisa (Tepiman) and pida- of NP pidahana’a ‘bat’ among the dozen-plus reflexes. Because of NUA -c-, the reconstruction must include *-Ct-/*-t- and NP actually has -t- among many Num -c-, yet in a Tep language (PYp) we find -s-, the usual reflex of *c, but ultimately from *t or *-Ct-.
*paCti’a > Ca pali, > *paci’a > *paca’a (Tb, Kw, Ch, SP, CU), > *pita- (NP pitahana'a ‘bat’),
> *paci’i > háci’i (Cr)
> *paci > *so’-peci (TrC: Tr, Wr, Eu) > *soci (Yq, My); *paCti > *paci > *so’o-pica > Tepiman ho’o-pisa (PYp)
UACV-935 *natipa (> ***nacipa** > ***nacpa** > Tep *naspa) 'fold': ST naspa’ ‘doblar, torcerse’; Eu nátpa ‘doblar’; Nv nasa ‘plegar una cosa’. Eu -t- aligns with Tep -s-, suggesting palatalization before c > s in Tep.
210 UA *tuti > *cuci > Tep *susi(-ka) > Tep susaka ‘sandals’: TO šuušk; LP šuušak; NT súúsaka; ST suusak. In light of Tep’s frequent aniticipatory V assimilation (*V-a > a-a), an original *tuti would have high vowels following both consonants (*tuti > *cuci > Tep *susi), then suffixed -ka would later encourage *susi-ka > susaka. As we often see Tep s < c < *t (i.e., Tep *susa < *susi < *tuti) and since Hp o < *u, then Hp tooci (< *tuti) ‘shoe, moccasin’ agrees with Tep entirely. [Egyptian **twt** ‘sandal’]
620 UA *tapputi / *tïpputi ‘flea’: TO čïïpš; PYp teepas; NT tapïïši; ST tapïïš; Eu tepú’u / tepú; Yq téput, tepučim (pl); My tépput; Wr tehpucí; Tr ŕipučí; Tbr tipú-t; Wc teepïï; Cr tepï-, tepï-ci (pl.). We see a 3rd consonant -t- in Yq, My, and Tbr, and even if the -t- was originally part of a suffix, it understandably palatalized in Tr, Wr, and the Yq pl, and that palatalization (c) is likely the source of Tep s, that is, the 3rd consonant in several Tep forms. The first vowel may well be ***a***; for NT and ST both show *a*, not *ï*, and if ï (a high V) were original, then results similar to *t > c > s as in ‘deer’ and ‘sandals’ for the first consonant would have resulted, but that did not happen, and perhaps because an original initial *ta syllable, which only later became tï, prevented it. [Semitic *đabbot ‘flies’]
809 UA *’ati / *ata / *aCti ‘laugh’: Wr a’ci 'estar riendose'; Tr ačí 'reirse'; My aače 'reírse'; AYq aače; Cr ra-’á’ace 'he is laughing at him'; TO a’as; LP ’a’aši; PYp a’asi; NT ááši-/ásyi; ST ’aas/ašia. Miller includes probable Ca ’ála’ 'mock, echo s.o., vt’. Because Ca ’ála’ has l, the Cupan reflex for intervocalic *-t-, it again may suggest a medial *-t- or cluster *-Ct- originally, which again did the cycle *t > c > s in Tepiman *asi. Ca ’ála’ is a transitive verb, perhaps preserving the final vowel -a, of the alternation -a ‘transitive, active’ vs. -i intransitive, stative’. [Semitic ***-hattil** ‘to mock’]
UACV-2205 *tïyuna ‘keep’: Mn tïyuna ‘store, v’; NP notïïna 'keep s.th.'; Ca téyan 'preserve, carry on (custom, rite)’; NT šiid^{y}úñd^{y}i ‘retacar, guardar, llenar mucho’. In *t- > *c > Tep *s by high vowels, Mn and NT agree well in * tïyuna.

Medial -p- (vs. -v-) from a Previous or Underlying Consonant Cluster

Many UA languages yield intervocalic -v- < *-p-, as the first set suggests. So when those same languages show -p-, it is from gemination or a cluster, perhaps even in Tep, as several sets suggest.
188 UA ***nopi** / ***nohopi** 'hand, arm': TO nowi ‘hand, arm’, pl: noonhoi; PYp novi, pl nonovi; Nv novi, pl: nonovi; NT novi; ST nov. TO pl shows h but no v. [Egyptian **nḫbt** ‘nape of the neck]
221 UA *wïr-pa’a ‘tall, long, great-height/length’: Hp wïïpa ‘tall, long’; Cp wevášša ‘long’; Cp wevášiš ‘tall’. Miller (M67-229) astutely sees Hp wïïpa 'tall, long' as a compound of *wïr-pa’a ‘big-height/length’. Intervocalic -p- in Hp instead of -v- supports Miller’s observation, though Cp -v- in Cp means it was sooner perceived as clusterless or non-geminated in Tak. [Egyptian wr ‘great’]
1070, 1071 UA *naNkapï ‘leaf’: Kw naga-vï; Ch nanká-va; SP maavï-naŋqa-vï ‘leaf’ (vs. SP naŋqava ‘ear’); CU nïká-’a-vi (vs. CU nïká-vi ‘ear’); Tb naŋhabïï-l; Hp nàapi / nahpi. Hp lost intervocalic -ŋk-, collapsing -ŋkap- > -ŋkp- > -p- in Hp nàapi / nahpi showing -p- instead of -v-, due to a previous cluster. [Semitic *na-qšab ‘be perked up’]
UACV-1547 *mukpiC ‘nose’: While Num *muvi lost all signs of a medial cluster, Sr and Ktn *mukpi agree with Hp mòope(q) ‘in front’ in showing evidence of the cluster.
UACV-1550 *sïCpowa / ***sïk-powa** 'numb': CN sepoowa 'be numb (of body part, from cold or lack of circulation)'; Eu zopóre ‘encogerse’. The first element of the CN term is suggested to be CN sek-tli ‘snow, ice’. Eu normally has intervocalic -v- for *-p-, so Eu -p- (vs. -v-) suggests a cluster in Eu as well.

Reduplication Created Clusters That Later Separated

Some sets show the base form (non-reduplicated) in NUA, while SUA shows the reduplicated form. Another consistency in both sets is that the second consonant is a liquid (-l- or -r-), and it appears that the reduplication first created a cluster, which caused the liquid to change to glottal stop, which was later separated from the other consonant by an echo vowel: *-VLC- > -V’C- > -V’VC-.
221 *wïr, reduplicated ***wïrwïru** > ***wï’wïru** > ***wï’ïwïru** 'big' or Tep gï’ïgïru: among the several UA forms, the reduplicated form is usually the plural form of *wïr. [Egyptian wr / wrw ‘great’]
630 *koli, reduplicated ***kolkoli > *ko’koli** > ***ko’okoli)** 'hurt, be sick, chili pepper': many SUA forms show *ko’okoli, while Cupan shows the non-reduplicated form with its vowel change *koli > *qoli > qili: Cp qilyíqa-t ‘hot, spicy, strong’; Cp qilyíqtu’ni ‘hurt, sting, vt’; Ca qélya ‘feel sore, v’; Ca qélyak ‘peppery, pungent, creating a burning sensation’. In SUA: TO s-ko’ok 'be painful'; TO ko’okol ‘chile pepper’; TO ko’okod ‘hurt, give pain to, vt’; NT kóóko ‘be sick’; NT kóókoli ‘chile’; ST -ka’ook ‘be sick’; ST ko’okoly ‘chile’; Eu kókoe- ‘doler’; Wr ko’koré- ‘dolerse’; Wr ko’kóri ‘chile’; My kó’okori ‘chile’; My kó’okore ‘enfermo’. [Hebrew **xolɛ** ‘be sick, hurting’]

1.44 The Labial Labyrinth in Uto-Aztecan

The labiovelar spectrum in UA is fraught with intrigue. The syllabic frequencies show a complete lack of *kwo and *kwu among UA initial syllables paralleled by a marked abundance of about twice as many ko and ku syllables as k with other vowels: 38 ko and 37 ku syllables vs. 10 ki and 17 kï, and nearly as many as the 43 ka, though across the board, *a*-syllables are normally twice what others are. Lack of *kwo/kwu syllables alongside about double the usual vocalic ratio for *ko/ku syllables may suggest that many *kwo/kwu became ko/ku, or that bo/bu > ko/ku, but ba, bi, bï before other vowels.

A count of TO’s initial syllables provides an even greater discrepancy. Considering that TO b corresponds to PUA *kw, notice that a rough count from Saxton’s (1983) dictionary yields the following:

	a	ï	i	o	u
b (< *kw)	ba(40)	bï(5)	bi(28)	bo(0)	bu(0)
k	ka(48)	kï(20)	ki(13)	ko(70)	ku(88)

Again in TO, a complete lack of bo/bu syllables contrasts with about triple the expected number of ko/ku syllables, as if in Tep languages *kwo/kwu > ko/ku. Note the TO variants of a plant (Mathiot 1976, 362):
UA bihul / hikul ‘a plant’. These alternate forms switch first and second consonants, except that PUA *kw is b before i, but *kw is kw before u. In PUA terms, *kwisuL > TO bihul, and *sikwuL > TO hikul.

If we take each language’s initial correspondences for *kw and place them before o and u, the likely results are *bwo/bwu > bo/bu in Cah (Yq, My), *wo/wu > o/u in Tr/Wr, *kwo/kwu > ko/ku in the kw-languages and in Tep as well, and *kwu > kwi in CN. Interestingly, some semantically plausible sets show that very array of correspondences.
UACV-1896 *kwuhV ‘scrape off, degrain (corn)’: Yq buh-te ‘espigar [take grain from ear]’; My búh-tuk ‘se espigó’; My búh-te ‘está espigando’; Tr ohó ‘desgranar [remove grain from ears]’; CN kwi’kwi ‘chip off (wood or stone), clean up a surface, take s.th. away, get ready, be prepared’. As Miller points out that Tr sometimes shows o as well as u for PUA *u, these four languages show PUA *kwuh ‘scraping off s.th.’: *kwu > Cah bwu > bu; > Tr oh; > CN kwih/kwi’.

UACV-1974 *kwuya (> *kwoya) ‘growl, scold’: Eu búde/nevúde/nepúde ‘growl, bark’ (Eu d < *y); My buuye ‘snarl, growl, bark, scold’; Hp qö’öqöya ‘scold, vt’; Hp(S) qöyqöya ‘he’s scolding’; Tr oyo ‘become angry’; TO kodog ‘rumble, gurgle’; and perhaps CN kwikwinaka ‘make a low sound in the throat; for a dog, to growl; for a person, to hum’ since CN i < *u. But TO kodog with d is usually < PUA *l/r rather than *y.
18 UA *sakwo > *sikwo/sikwi 'witch, bewitch': My sisibo 'hechizar'; My sibori 'hechizado’. Cp sekwíte 'curse, whip' (Cp i < *o) suggests a semantic tie such that the set under *sakwi 'whip, v’ (at whip) may be related: M88-sa27; KH.NUA: Cp sekwíte 'curse, whip'; Cp sekwítxe-l 'whip, n'; Sr şakwit(kin) 'whip, swat, vt sg obj’ (borrowed from Cup?); Gb sakwít 'castigar'; Ls şíqwi 'to punish, whip’ (vowel is wrong, Miller notes), but Miller speaks of the first vowel, often putting too much emphasis on the unstable, unaccented vowels; Tr siku- 'hechizar'; Tbr sigu-l 'hechicero’. Ls -qw-, rather than -kw-, suggests a non-high second vowel, i.e., a second vowel of *o instead of *i originally (Langacker 1970), which agrees with SUA TrC (Tr, My). As for the first V, it appears that *_a_ went to the schwa options—i and ï—suggesting it may have been unstressed previously, with Sr and Gb maintaining the original *a*. And note My -bo- (< *bwo) with Tak *-kwo-. Tr ku < *kwu may be the medial reflex vs. the initial.

We also often see what we might call **kw-reduction—*kwVC > kuC/koC**—where the vowel between *kw and the next C becomes short enough that the rounding of *kw overrides it, and the result is k + round V + C: e.g., 15 Tr kusá at *kwasa ‘eagle’; 44 Ca kuş at *kwïsi ‘grasp, take’; 24 Tr oke/weke at *kwïkï ‘weep’; etc. Perhaps kw-reduction is more likely between two bilabials, as below:
36 *kwawa/i ‘invite, call’: Cp kwawe ‘call, invite’; Tr o’wí ‘invite’; Wr oí ‘invite to work’; Eu bowá ‘invite’; perhaps the baa- of TO baamuđ ‘plead, invite’ (lack of TO g < *w is frequent enough). These forms show kw-reduction in some (TrC), which brought the kwo-phenomenon into play in Eu, Tr, Wr, while Cp may come nearest the original *kwawV. [Hebrew baʕaa ‘enquire, search’]
8 UA *cakwa / *cakwo / *cakwi 'catch, grasp, close, lock': Ls čáqwi 'seize, catch'; Cp čáqwe 'catch, grab, cling to'; TO šaakum 'catch, grasp’; NT saakómi 'handful'; ST saakum 'handful’; CN cakwa 'close, enclose, lock up'; CN cakwi 'close, get closed, vi’; Pl cakwa (pret cak) ‘close, shut, cover’; Mn cakwiti’i 'close, lock, bolt’. Here kw-reduction in Tep between two labials (*kw and m) triggers Tep ku < *kwu, instead of bu < *kwu. [Semitic *ḍabba / ṣabba ‘grasp, lock’]

Infrequently mentioned is the fact that Tr often lends itself to Tepiman-like phonology in the labial realm or has variants with Tep correspondences in addition to the usual Tr correspondences. The widely publicized sound correspondence for *kw in Tr is w initially and for *w is also Tr w. While those two are most frequent, Tr has dozens of variant pairs, in which one variant indeed shows the touted w < *kw or w < *w or b < *p, but one variant resembles Tepiman phonology: *kw > w/b or *w > w/g/k or *p > w/b:
<u>*kw > b</u>
Tr wasi-/basi-bura ‘loincloth’ (< *kwasi ‘tail, penis’) 5
Tr wasu/basu ‘cook in water’ (< *kwasV ‘boil’) 4
Tr we-móri/be-móri ‘dust’ (< *kwiya- ‘earth’) 19
Tr wa’wé/ba’wé ‘eagle’ (< *kwa’awV > TO ba’ag; Eu páwe)
<u>*kw > gu/go</u>
Tr witá/guté ‘feces’ (< *kwita ‘feces’)
Tr ciwá/cigó ‘rob’ (< *ïcikwa ‘steal’)
<u>*w > g/k</u>
Tr oná/koná ‘salt’ (< *oŋa/*omCa; Wr woná) 280
Tr oona/koona ‘corncob (Wr wo’ná)
<u>*p > w/b</u>
Tr wici-/bici- ‘believe’ (< *piti)
Tr wíso/bíso ‘infect(ion)’ (Wr pehsóni; PUA *pisVk ‘rot, infection’)
Tr bo’o / ko’o ‘del otro lado [of/from the other side]

Other Tr forms show similar and considerable phonological variety: Tr uusabi / kuusabi / guusabi ‘Prunus Capuli’; Tr utuburi / tutuguri / ŧutuburi ‘type of dance’ (note b-g alternation medially)

121 Most intriguing is the pair—Tr bineri ‘alone, only, sg’ and Tr a’wineri ‘alone, only, pl’—as if *p > kw when geminated medially, since -’w- is a reflex of medial *-kw- in Tr, perhaps also in *kap(p)a ‘egg’ below.
UACV-803 *kap(p)a ‘egg’: Eu akabo-ra; Yq kaba; My kabba; Tr ka’wa, among others.
UACV-995 Note medial *-p- > -kw- exists in Num: ***yïpana** 'autumn': Mn yïba, yïbano ‘be autumn’; NP yïbano; TSh yïpani; Sh yïpani; Kw yïvana; Ch(L) yïvana; SP yïvannaC / yïvwanna; CU yuvwa-na(-ttï) / yugwa-na(-ttï).

Note that when the labiovelar glide -w- develops in SP -vw-, then the labiovelar -kw- is the next step in the next language east (CU). Similarly, I have heard native speakers of Yaqui pronounce intervocalic -w- with some velar contact: -gw- (< *-w-), and Shaul and Yetman (2007) suspect Op gw was an intermediate step from *w > gw > g. At *hupa (> *howa ‘back’), the Tbr variants (ova/owa/ogo) show another instance of velarizations of labials preceding

round vowels. Larry Hagberg (p.c.) informed me that in My also PUA *wo is usually pronounced *wo*, but occasionally *go*, but not *gwo*; but with other vowels, *wa, for example, is never pronounced *gwa* only *wa*. Also at 613 Tr gohi < Tep wohi 'bear'. So round vowels can trigger velarization in labials. In contrast, Monzón and Seneff (1984) note *kw > w, bw, b in various Nahuatl dialects.

Manaster Ramer's (1993a) suggestion of *-tw- > -kw- finds support in the My reflex of *ïcikwa/*ït(i)kwa 'steal'. Among the TrC reflexes (Eu écba'a-n, Tbr icikwa, Yq 'étbwa) is My ekbwa, which essentially does the change that Manaster Ramer proposed, changing non-velar t/c to a velar -k- adjacent to the labio-velar *kw/bw.

1.45 Nasals of Uto-Aztecan

Uto-Aztecanists have long held to the correspondences of NUA ŋ: SUA n and NUA n: SUA L (L = either liquid, l or r). David Shaul (1985) and Jane Hill (2007b) summarize the history of the matter well, stating that Miller (in Miller and Silver 1997, 285) viewed the matter as PUA *ŋ > SUA n and PUA *n > SUA *L (l/r). Others, VVH (1962), Campbell and Langacker (1978), Manaster Ramer (1993), and Dakin (2001), have argued for the opposite direction of change: *L > NUA n, and *n > NUA ŋ. Sapir (1915, 475), on the other hand, considered *ŋ > SUA n more probable, but also considered PUA *L and *n to have merged in NUA, or *L > NUA n (Sapir 1915, 477), and that *n remained n in both NUA and SUA, though disappearing in SP when not geminated (Sapir 1915, 473-4). Sapir's view comes nearest the author's. I see PUA as having at least one liquid, if not both *r and *l, in addition to both *n and *ŋ.

The correspondence of NUA n : SUA n is much more frequent than NUA ŋ : SUA n. In Miller 1988 we see n:n in both NUA and SUA in na-1*naka 'ear'; na-2 *naki 'want'; na-5 *napu 'prickly pear'; na-7 *na'i 'fire'; na-29 *naka 'meat'; ni-1 *nioki 'say'; nï-2 *nïma 'liver'; nï-9 *nïmi 'walk around' (126); nï-11 *nïpaR 'snow'; 266, 274, etc.) So if *n > ŋ in NUA, then why did so many more *n remain n in NUA instead of doing the sound change *n > ŋ, like the other one-third of them did? The correspondence NUA ŋ : SUA n is much less frequent and may be limited to medial positions, as we do see ŋ:n in *laŋi 'tongue' (698), *omwa 'salt' (280), *kumwa 'husband' (281), *somwo 'lung' (283). However, the candidates for ŋ:n in initial position may not be valid, that is, may have different stems in NUA and SUA respectively: na-6 ŋa 'root' and na-10 ŋa 'cry'. The set of no-2 ŋo/no 'return, bend' has the best chance for viability, but even they may be different NUA and SUA sets (931).

NUA ŋ is often the reduced result of a consonant cluster, one of which is often a nasal. Because many ŋ are from cluster reductions (though not all), it seems less reasonable that *n became ŋ and then ŋ blossomed into an array of consonant clusters, but rather that *-NC-/-CN- > *ŋ > SUA n. For example, *kumCa 'husband' (below) > *kuŋa (NUA) > *kuna (SUA) seems more likely than *kuna > *kuŋa > *kumwa. The parallel corollary of such a change would be PUA *n > SUA l, and is sometimes the case, yet again I agree with Sapir, that in other cases PUA *L > NUA n. The *n-*L complex remains mysterious in part, though something like a merger of *n and *L to n in NUA, which Sapir (1915, 477) also suggested, and *l and some *n merging to SUA l may hold some potential, though groups of exceptions litter the aspired neatness and await insightful explanation. The next six sets exemplify NUA n : SUA n.

1070: UACV-752a ***nakka / *naNkapa (< *na(N)kasapa)** 'ear' [Semitic *na-qšab 'be perked up (to hear)]

Mn	náqa	Hp	naqvï	Eu	nakát 'oreja'
NP	naka	Hp	naaqa 'ear pendant'	Eu	kéisiven 'oido'
		Tb	naŋha-l 'ear, leaf'	Tbr	naká-r
TSh	naŋki	Sr	qävaač 'ear, leaf'	Yq	náka
Sh	nainki	Ca	náq-al	My	nákka-m
Cm	naki	Ls	náq-la	Wr	nahká
Kw	naga-vi-vi	Cp	náq'a	Tr	naká
Ch	naŋkávï	TO	naak	Cr	našaíh
SP	naŋkava-vi	PYp	naaka	Wc	naaká
SP	naŋka 'hear, v'	NT	naáka	CN	nakas-tli
CU	nïká-vi	ST	naak/nak	Pl	nakas

UACV-1366 ***nïmaC / *nïmaN** 'liver':

Mn	nïwï	Hp	nïïma	Eu	hemát
NP	nïma	Tb	nïïma-l	Tbr	yamá-t
TSh	nïmï(cci)	Sr	nïmïič	Yq	héemam
Sh	nïmïn; nïwïn	Ca	ném'a	My	heémam
Cm	nïïma	Ls	nóóma	Wr	emá
Kw	nïwï-bi	Cp	néma; pípiviska	Tr	imará; emará
Ch	nïwïmpi	TO	nemaj; nem 'a liver'	Cr	neemwa
SP	nïŋwï-n, nïŋwï-mpi	Nv	nïmadi	Wc	néma
WMU	nú<u>u</u>-ppü-n 'my liver'	PYp	nemar; LP hïm	CN	eel-li
CU	núu-pï-n 'my liver'	NT	nïma(dï)/númai	ST	lumaad

126 UACV-1012 *nïmi ‘walk around, live’: **NUA:** NP nïmmi ‘walk’; TSh nïmi ‘one moves’; Sh nïmi ‘live’; Cm nïmi ‘move about, walk, sg’; Ca ném ‘walk around’; Ca némi ‘chase, follow tradition’; Sr nïm/nïmï- ‘walk, walk around, walk along’; Ktn nïm ‘walk, vi, walk on, vt’; Hp -nïma ‘go around doing s.th.;
SUA: CN nemi ‘live’; HN nemi’ ‘walk’; Pipil nemi ‘be, exist’. [Egyptian **nmi** ‘travel, traverse, go’]
885 UACV-878 ***na’ay** ‘fire’; ***na’aya** ‘build/light a fire’:
SUA: Wr na’í ‘flame’ and Wr na’yá-ni / na’i-ma ‘make a fire’; Tr na’í / na’y- ‘fire’ and Tr na’yá- ‘make a fire’; My na’- ‘burn, v’ and My náyya ‘hacer lumbre’; AYq naya’i ‘fire’; TO naada ‘fire, n’ (TO d < *y); ST naada’ ‘make fire’; NT naadá; Nv nadda; Cr á-úu-na’ara ‘go build a fire’;
NUA: Mn ani ‘burn, vi’; NP nai ‘fire, burn vi’; NP na’i’yu ‘burn, vi’; Kw ne’e ‘burn’; SP na’ai ‘burn’; CU na’ay-ttï ‘fire, light’; Ca ná’ ‘burn’; Ls ná’ ‘burn’. [Arabic naar ‘fire’ but written na’r / na’ar]
720 UACV-7a *no’pal / *napu ‘prickly pear cactus/fruit’: **NUA:** NP nabu; TSh napumpï; Sh nabombï (Fowler83); Kw navu-bï; Ch navumpï; SP nabumpï; Hp naavï; Sr naavt; Ktn navïh-t; Ca návet; Cp návet; Ls náávu-t;
SUA: TO naw/nawï; Nv nubo(nïvo); LP(B) nav; NT návoi; ST nav; Eu navúc; Wr napó; Tr napó; Yq naabo; My naabo; CN no’pal-li. [Semitic nbl / Syriac n’bl ‘skin-bottle’]
1407 UACV-2085 ***mo’ona(C) / *monna / *moCna** ‘son-in-law, in-law’: **NUA:** Sh monappï; Kw mono; SP munna / mona-ci; Hp mö’önaŋw 'male in-law';
SUA: Eu mónwa; My mó’one; Yq mó’one; Wr mo’né; Tr mo’né-ra; Wc muune; Cr -mu’un 'yerno'; CN moon-tli 'son-in-law’. [Hebrew maɦ^{a}ne < *maɦne ‘camp, people of the camp’; as in-laws become family]

Medial *-’m- and Other Consonant Clusters with Nasals Underlie Some Medial -ŋ-

UACV-1221 ***sï’moci** 'hummingbird': Wr se’móci ‘hummingbird’; Tr semučí / simučí ‘hummingbird’; NP soŋoi’i 'hummingbird’. NP aligns with *sï’muci in that NP's 2nd and 3rd vowels agree with Tr and Wr, and if the 1st assimilated to the 2nd (*ï-o-i > NP o-o-i), and PUA *-c- > -y- (or i or ’), then *sï’moci > *so’moyi/*so’mo’i > *soŋoi’i has NP being a decent match with Tr/Wr, and glottal stop plus m (-’m-) aligning with -ŋ-. The next three sets show the -’m- cluster in SUA, and -ŋ- in NUA.
771 UA *cu’mi ‘suck, sip’: Kw čohmi ‘suck, v’; Cp čúŋe ‘kiss,vt’; Cp čúmum 'suck obj, as venom’; Cp čúme ‘suck, vt’; Ca čúŋ suck, vt’; Ls čúúŋi 'suck (breast)'; Ls čúŋi 'kiss'; Sr čuuŋ 'suck, vt'; Wr cu’mi 'suck or slurp food'; Tr cu’mi 'kiss, sip'; My čuune; AYq čuune; Hp coocona 'kiss, suck’; CN (paal)čičiina ‘soak up, suck in, smoke, vt' and CN ilčiina 'suck up, consume'; HN čičiina / čičiini’. Nv tup’suma ‘suck, vt’; NT višúúsumai 'suck’. These forms suggest *cu’ma. Six languages show medial -m- or -Cm- aligning with the frequent NUA ŋ and SUA n. [Hebrew **ṭʕm** ‘taste, eat’; plural prtcpl **ṭoʕmiim > *cu’mV > *cuŋV** 'suck, sip, kiss’]
1144 UA ***o’mana** ‘sad, suffering’: CN a’mana ‘be upset, disturbed’; Tr o’moná / o’móna- ‘be afflicted, saddened’; Tr o’móna-ri ‘sadness, affliction’; in Sr the -uŋani- portion of Sr ahauŋanik ‘sad, miserable’; Sr hahauŋan ‘be poor, pathetic, miserable’; Sr hauŋanič ‘poor one, orphan’ (u often pronounced o); and Ktn haoŋa ‘poor’. Words as long as the Sr forms are certainly compounds, so -uŋani- likely aligns with CN and Tr. Here the cluster -’m- appears in SUA (CN and Tr) and as ŋ in Sr and Ktn, as in 771 cu’mi in Tr/Wr and ŋ in NUA; in addition, the Tr and CN forms agree perfectly in the consonants -’m-n-, but disagree in the vowels: a-a-a vs. o-o-a. However, the vowels of Sr and Ktn are between the two, agreeing fairly well with both, perhaps:

PUA *o’mana > CN a’mana
> Tr o’mona
> Sr -uŋani- / Ktn -oŋa [Hebrew ’almaanaa ‘widow’; Arabic ’alima ‘to experience grief’]

856 UA ***yu’mi** > **yuŋi** 'warm’: NP yuwi; NP yui; Sh yuai ‘warm’; Cm yu’a ‘warm (of weather)’; SP yuuttui ‘be warm’; SP yu’mi ‘warm (of water)’, yu’ata (of weather); Hp yoŋi 'be warm’. Even if SP yu’mi and Hp yoŋi have an extra morpheme than the others, Hp (-ŋ-) and SP (-’m-) still suggest a medial cluster. The fact that 9 sets (in UACV) show m in some languages and ŋ in others suggests that medial -m-, when clustered (-Cm-/-mC-), reduces to -ŋ-. [Hebrew **yɦm** ‘be in heat’ (alternate form of ɦmm ‘feel warm, get warm’]
1114 UA *sïk-mukki ‘numb’ < ‘ice/cold-dead’: Hp sűmokiw|ta ‘be numb, vi’; NP ta/ma-sïsïŋi 'foot/hand goes to sleep'; Cm sïsï’nitï ‘numb, feel numb, asleep’; WMU sï’<u>ú</u>ú 'be numb’. The first morpheme could well be a cognate of CN sek-tli ‘ice/cold’. Hp lost the velar stop, but preserved the vowel pattern best. In NP, Cm, and WMU are cluster reductions, showing residual features of both consonants, in which the velar + nasal cluster -km- went various directions: *-km- > ŋ (NP); -’n- (Cm); and ’<u>u</u> (WM; underlined V = nasal V), for all show signs of a velar (velar nasal or glottal stop) and a nasal; a nasalized vowel shows the nasalization in WMU. [Hebrew šɛlɛg ‘snow’ + Hebrew mukkɛ ‘smitten’]

After five examples of -'m- aligning with -ŋ- , consider three well known examples of NUA ŋ aligning with SUA n, but with several seldom-highlighted m's among the NUA reflexes as well.

HUSBAND; MARIDO

Mn	kúwa	Hp	kooŋya	Eu	kúnwa
NP	guma	Tb	kuuŋa	Tbr	--
Tsh	kuhma(cci)	Sr	--	AYq	kuuna
Sh	kuhma/kuha	Ca	--	My	kuuna
Cm	kumahpï'	Ls	kúúŋ; tó'ma-vu	Wr	kuná
Kw	kuhma	Cp	kúŋ	Tr	kuná(ra)/guná(ra)
Ch	kumá	TO	kun	Cr	kïïn (2nd V stressed)
SP	kumma	LP	kun	Wc	kïna
WM	piwá	NT	kúna	CN	--
CU	piwá	ST	kun		

284 UA *kumCa / *kuCma 'husband': this set is one of few whose reflexes appear in 25 or more UA languages. Note Hp, Tb, and Tak ŋ aligns with SUA n, while 9 Num languages show -m(m)- / -Cm-. WMU and CU have piwá 'husband', but kumma 'male' also, in a slight semantic shift on SNum's east end:

SP kumma 'male, husband' SP piŋwá 'wife, spouse'

CU kumáa-vi 'male animal, stud, macho' CU piwá 'spouse, husband, wife'

The fact that nearly all UA languages show a form agreeing with *kuNa, but only vary in the type of nasal, three different nasals, no less—bilabial in Num; velar in Hp, Tb, Tak; alveolar in SUA—suggests that we are dealing with a single proto-form whose medial consonant is likely a reduced cluster, probably involving *m* and something else. Reflexes of 'lung' and 'salt' do similarly. [Egyptian qm']

LUNG(S); PULMÓN(ES)

Mn	sóno	Hp	halayna; mïma	Eu	abokadaga-di
NP	soŋo/sono	Tb	mošooha-t	Tbr	wopaN-s; sorá komwa-lí-t
Tsh	somo/soŋwo/soŋo	Sr	--	Yq	saré'ečia
Sh	sonko/sonno	Ktn	šoŋa-č	AYq	hemaha'ačim
Cm	soomo	Ca	yávayva	My	sáre'ečiam
Kw	soo-vï	Ls	şavá-şva-š	Wr	so'locá
Ch	soo-vi	Cp	qíqilye	Tr	sonorá
SP	soo-vi	TO	hahaw	Cr	šáïñi-mee; ta'atime
CU	sö'ö-vï	PYp	hakadaga; pl: havdaga	Wc	šaaka
		ST	havkal	CN(RJC)	mimiyawayo-tl

291 UA *somCo / *suNCa 'lungs': Mn; NP; TSh; Sh; Cm; Kw; Ch; SP; CU; Tb; Sr; Ktn; Gb sár; Tbr; Cr; and HN sooneewa' 'to swell up (of vipers)'; Tr sonorá. Tr has the expected SUA n for NUA ŋ, but we see NUA -m- (Tsh, Cm) and -ŋ- as well as SUA -n-. [Egyptian sm']

SALT; SAL

Mn	omábi; omaa- 'to salt'	Hp	öŋa	Eu	onát/ónta
NP	oŋabi	Tb	uŋaal	Tbr	oná-t
Tsh	oŋwapi(cci) / omapi-	Gb	'oŋó-r	Yq	'óna
Sh	oŋa- / onka-/ ona-pin	Ca	'íŋ-il	My	oona
Cm	ona-/onaabi/ona'aitï	Ls	'éŋ-la	Wr	woná
Kw	'owa-vi	Cp	íŋeyu 'to salt'	Tr	oná/koná/noná
Ch	--	TO	on	Cr	unáh
SP	oa	PYp	ona	Wc	'únaa
WMU	'öá-vi	NT	ónai		
CU	'öá-vi	ST	'on	CN	--

280 UA *omCa/*oNCa > *oŋa (> SUA ***ona**) 'salt': Reflexes are in all branches except Azt, and medial consonants (n, ŋ, m, ø) again show a pattern similar to 'lung' and 'husband' with Mn and TSh showing m. [Egyptian ḫm'(t) 'salt']

1246 Canaanite ***ha-sim'al** 'left' > Tb **aašiŋan** 'left'

1012 Hebrew **šiqma(t)** 'sycamore' > **UA *sïŋŋa(C)** 'cottonwood and/or aspen tree'

807 *sïm 'laugh': Cp šeme; Ca sém; Od hïhïm; ST h(ï)mpa, h(ï)mia; Nv 'ï'ïmï 'smile'; Ca sém-yaw 'smile'; Ca séŋi 'smile' may involve the same stem as Ca sém-yaw, but with a differing suffix, then ŋ becoming a cluster reduction. [Hebrew **śimḫ** 'be happy'; Hebrew **śimḫaa / śimḫat** 'joy, gladness']

Above are 11 sets having medial clusters of m plus something else corresponding to some NUA ŋ and SUA n. Below are other cluster combinations corresponding to NUA ŋ and SUA n.

1418 UA ***taŋa** 'bag, sack, contain(er)': Sr taŋat 'sack'; Gb taŋár 'sack'; Hp taŋa 'contained things';
Hp patŋa 'squash' (with pa- prefixed); Tbr tanaté 'zurrón, mochila de cuero en que se acarrea a la espalda el ineral'; -ta'ni of Mn kusatá'ni 'sack' (kusa 'sack'); CN taana'-tli 'basket with a handle'; and Yq 'ía-tana 'this shore/side' (a shore as that which contains/encloses water). *taŋa compounded with *pa- 'water' produces *pa-taŋa 'squash, pumpkin, gourd, i.e., liquid-container' (Stubbs 2003:4 and KH/M03-pa66 'squash'): Ch paráŋar(a) 'pumpkin'; SP paráŋwaraN 'pumpkin'; and Hp patŋa 'squash, pumpkin'. Note that the only NUA language not showing ŋ (Mn) does show a cluster of glottal stop plus n (-'n-), which suggests a cluster. [Semitic *ta-ʕra' > UA taŋa']
1066 UA ***corowa / *corwa** 'be hungry': Wr coloá-ni 'be hungry'; (Wr co'-cóla-ni 'be hungry, pl');
Hp cöŋö-w(ï)- 'hunger'; Hp cöŋ-moki 'die of starvation'. Wr coloá- and Hp cöŋö- match well, since Hp ö < *o, and if -owa- > -oa- in Wr, then syncope causing a cluster of *-lw- > -ŋ- in Hp is natural, for w is a labio-velar and SUA liquids often become NUA nasals, so the nasal and velar dimensions' becoming the velar nasal is reasonable. Note Tr čiriwísa 'tener hambre', which has the same three consonants (c, r, w). In light of alveolar consonants causing V > i in Tr, as also in Tr bikiyá 'three'< *pakay. [Arabic ḍrʕ > UA *cor(V)wV]
628 UA *ca'ro 'chin, jaw': Tr ča'ró 'chin'; Wr caló 'chin, jaw'; CN teen-čal-li 'chin'; CN kama-čal-li 'jaw'; Yq čao 'barba'; My čaro hímsim 'bigote'; My čaro wá'asa'ari 'quijada'; Hp càŋw-ti 'open the mouth'. The medial *-'ro- of SUA likely corresponds to Hp -ŋw- much like we saw in *corowa 'hungry' above. These sets (*corowa, *ca'ro, and UACV-326 *yïLCa) with Hp ŋ aligning with SUA liquid plus round vowel suggest two things: (1) they suggest a liquid > NUA nasal, since *ŋ > l/r is hardly likely in the other direction; (2) and they show Hp ŋ aligning with likely clusters of a nasalizing element (*l/r > N in NUA) plus w or round vowel. [Hebrew ***đaqn-o** 'chin-his']
681 UA *wïl 'grow': Ca wél 'to grow, rise up high'; Cp wéle 'to grow'; Ls wola/i 'grow (of plants or anim subj)'; and Hp wïŋwa 'grow, grow up' (< ʕVlwa). [Hebrew **ʕlw** / ʕly / **ʕalaa** 'ascend, go up, grow']

One among many examples of a medial NUA -ŋ- corresponding to SUA -n-, but not from a cluster:

952 UA *poŋa / ***poŋo** 'hit, pound': Cp píŋe 'knock on, knock around'; Ls péŋa/i 'throw, be thrown'; Sr pööŋ 'pound'; Ktn poŋ 'hit with the fist'; Hp pöŋöŋöta 'be making a knocking or rapping sounds'; AYq poona 'knock'; Yq pónne 'pound, crush'; My póona 'hit, touch'; and My popona 'hit/pound with a hammer'. [Hebrew **pgʕ** 'meet, attack']

The prominent UA cognate for 'tongue' is in 7 of 8 branches, in every branch except Numic, and it is yet another example of NUA -ŋ- corresponding to SUA -n- medially as above. Hp and Tb begin with l- and all other UA languages begin with n-, so the Uto-Aztecanists figure that *n- is the initial consonant and that Hp and Tb disassimilated. However, the opposite direction of assimilation is more likely, as explained below:

698 UA *laŋi / *laŋu 'tongue': Hp leŋyi / leŋi 'tongue'; Cp naŋ; Ca náŋ-il^{y}; Sr naŋ|ač; Ktn nïŋi-č; Gb -nóŋin (poss'd); Tb lalan-t / lalun-t; Eu nenét; Tbr niní-r; Yq níni; My ninni; Wr yení; Tr inará/inirá; TO neeni; LP nïnni; PYp neeni; NT nïïni; ST nïïn; Cr nanuri; Wc neení; CN nene-pil-li 'tongue'; CN nene-tl 'female genitals'; Pl nenepil 'tongue'. Sapir suggests that Hp and Tb dissimilated *neŋi > leŋi, then Tb assimilated again > l-l. The reverse is more likely (*laŋa > naŋi), the liquid assimilating to the following nasal, as anticipatory consonant harmony is common in UA. And Tb does preservative V assimilation, so perhaps in this case preservative C harmony also. Initial *l is not common in UA, so assimilation to the usual (*l- > n-) seems more likely than dissimilation to the unusual (*n- > l-). Note also that initial l happens in Hopi (695, 698, 700). Sapir also notes the voweling *a-u in Cr and Tb. Since none of the languages show *e-u, but rather all with u show first vowel a, then the voweling *ï-i could be the 1st assimilating to the 2nd, such that the original 1st vowel was likely *a*, as it appears in Tb, Sr, Ca, and Cr. The 2nd may have more likely been u (which aligns with Hebrew pl), and final V > i is common, but anything else > u is not. So the reconstruction *laŋu serves best. [Arabic *lahgat 'tongue', the Hebrew voweling for an unattested plural would be *lahgoot]

Four decades ago Munro (1973) demonstrated that a half dozen sets show Ls ŋ < PUA *w. The forty years since that time have turned up a few more examples but not an explanation. In fact, some rather sporadic ŋ < *w in some other languages (mostly Takic) seem to complicate more than clarify. The matter is mostly clarified in 6.7, but not entirely.

757 UACV-2575a ***siwa < *si(ŋ)wa / *siwNa** 'female, sister, daughter': Sapir; M67-470; Munro 1973: Hp siwa 'sister of a man'; CN siwaa-tl / sowa-tl 'woman, wife'; Pl siwaa-t 'woman, wife'; Ls ṣawáa-may 'daughter'. Miller and Bright's observation that Ls ṣawáá-may 'daughter' is the diminuitive of Ls ṣuŋáá-l 'woman' is very relevant to the nasal clustered with -w-. CN may show a vowel assimilation to w (*siwa > *sowa) that occurred in other languages also, probably in Tak *suŋa, TrC *sona 'wife' and Tep *hooniga 'wife'.
UACV-2575b ***sï'a** 'girl': I.Num195 *sï'a (young) girl; M88-sï11 'young girl'; KH/M03-sï11: Mn sï'a; NP sïa'a / cïa'a. The WNum forms likely tie to *siwa/siwŋwa, but until an explanation emerges, a separate letter is good.
UACV-2575c ***suŋa** 'man's daughter, wife': M88-su21; KH.NUA; KH/M03-su21: Cp ṣuŋáma 'man's daughter';
Ca súŋama 'man's dau'; Ls ṣuŋáá-l 'woman, wife'; Gb ásoŋ 'wife'; Sr ṣuuŋ 'man's dau'. Add Ktn huŋ 'descendant' and Ktn nïmihuŋ 'wife', pl: nïmihuŋam (< *nïmi-suŋa 'man's-girl/woman').

UACV-2575d ***sona < *suŋa** < ***si(ŋ)wa** 'woman, wife': B.Tep73 *hooniga 'wife'; B.Tep72 *hoonita/hoonata 'to take a wife'; L.Son256 *sona ‘esposa’; BH.Cup ṣuŋáma’ daughter of man (diminuitive of woman); M88-so8; KH/M03-so8: Tb so’yiil ‘wife’ (cognate?); Tbr soná-r ‘esposa’. [Hebrew **šipḫaa** ‘maid, maid-servant, concubine’]
1059 UA *tï(N)wa / *tïnwa (AMR) ‘name’: Hp tïŋwa ‘name, refer to, vt’; Tb ’ïndïŋwa-l ‘name’; Cp téw’a ‘name (n. poss’d)’; Ca téwal; Ls túŋ-la; Sr tïwan(č) ‘name, n’; Ktn tïw; TO čïïg ‘(1) find, (2) call by name’; PYp teegi ‘name’; Eu tewát; Tbr temwa-ra; Yq tea; My tééwam; Wr tewá; Tr ŕewá; Wc téévää; Cr an-tyawaa ‘he is named X’. Munro suggests -ŋw- may explain *o > u in Ls. Note ŋ with w in Hp and Tb. [Arabic dʕw / daʕaa ‘to call, name’]
332 UA *koNwa ‘snake’ reflects a medial -rḫ- cluster (< *qVrḫat). This widespread cognate is in 6 of 8 branches, and while Joe Campell (1976) cites a Nahuatl dialect showing *koŋwa, most show *kowa, except Takic, which has Tak medial -ŋ-: Cp qeqiŋi-ly ‘king snake’ and Ls qiqeŋ-la ‘ring snake’ < Tak *koŋo.
[Egyptian **qrḫt** ‘serpent (sometimes bird determinative instead of serpent), friend/partner’]
Four more instances of pharyngeal ḫ reflecting Ls ŋ follow:
270 UACV-70 ***tïpiwa / *tïpiN** 'ask': Mn tïbiyu; Mn tïpiwï (M88); Mn tïtïwï- ‘ask for (objects)’; NP tïbiŋa; TSh tipiŋa; Sh tïpinka (= tïpiŋa) ‘ask for’; Kw tïvina; Ch tïviŋi; SP tïvi / tïvi-ŋu ‘to ask’; CU tïvïyuy; Hp tïïviŋ-ta 'ask, inquire of, ask for'; Ls tuvyuŋi 'ask a question'; Cp túvyuŋ 'ask'. [Egyptian **dbḫ** ‘ask for]
411 UA *hoŋ ‘body’; remember Tepiman n corresponds to NUA ŋ: TO hon 'body'; Nv hona; PYp hona; Ls heŋča-wu-t ‘cheerful, contented’ is key: Ls e < *o, and Ls ŋ corresponds to pharyngeals and to UA *w also in woman, name (Munro 1973) and to SUA n; and Egyptian ḫʕ unites the meanings ‘happy’ and ‘body’.
[Egyptian ḫʕ / ḫʕw ‘body’, Egyptian ḫʕwt ‘joy, rejoicing’]
412 Ls heŋča-wu-t ‘cheerful, contented’. [Egyptian ḫʕ / ḫʕw ‘body’, Egyptian ḫʕwt ‘joy, rejoicing’]
413 Ls hiŋé’-ma-l / hiŋéé-ma-l ‘boy’. Ls even shows the 3rd consonant glottal stop [Egyptian ḫʕ’ ‘child, boy’], besides the first 2 consonants matching in the last 3 sets: Egyptian ḫʕ > Ls hVŋ.

1.46 NUA Liquids Corresponding to SUA Liquids

In contrast to PUA *l > NUA n or *n > SUA l (as Uto-Aztecanists have seen matters heretofore), several sets show liquids for both NUA and SUA:
6 UA *kwïlu 'swallow': Hp kwelo(-k) ‘sample by tasting'; Eu béru’u 'swallow'; Tb weleeh 'swallow’. Hp and Eu correspond perfectly through 4 segments, since Hp o < *u and Eu b < *kw. And Tb’s w (< *kw) agrees through 3, the last V assimilating to the first, yet all NUA and SUA forms show a liquid.
630 NUA *koli, SUA reduplicated ***kolkoli > *ko’okoLi.** Again, all SUA and NUA forms show liquids.
88 UA ***walaka** ‘snail’: CN wilaka ‘caracol de monte’; Tr warákoara ‘caracol’; Ls muvílaqa 'snail' (Ls múúvi-l ‘nose’); Wr alágaloci ‘snail’; Wr nalágeloci ‘snail’; Tr narákuri ‘snail’; another example of a NUA liquid (Ls) and SUA liquids, though some languages added prefixes that eliminated initial w(V)-. [Hebrew **ʕaluqaa** ‘leech’; Arabic **ʕalaqat** ‘leech’]
381 UA ***wirhukuN** 'buzzard, turkey vulture, zopilte’ (in 7 of 8 branches, missing only in Tep):

Mn wiho; NP wi’ho/wiho	[WNum]
Tsh wihnumpi(cci) / wihumpiccih / wiyombic; Sh wikkumpiccïh	[CNum]
Kw wiku-mahaa-zi; SP wikkuN; CU wəkúci-ge-tï	[SNum]
Hp wisoko; Tb wišokombiš-t 'song of the turkey buzzard'; Sr wirukt	[other 3 branches of NUA]
Yq wiiru; My wiiru; Tr wirú; Tbr wilú	[TrC]
Wc wirïkï; Cr viskï	[CrC]
CN wiiloo-tl, pl: wiiloo-me’ 'dove'	[Azt]

Besides a general NUA liquid and SUA liquid correspondence, we see the liquid > -s- in three languages (Hp, Tb, Cr), and being clustered with a voiceless spirant best explains the devoicing of *-r/l- > -s-. Wc (SUA) and Sr (NUA) show all 3 syllables of *wirhukuN, while the rest are reductions. [Egyptian wr ḫq’w ‘buzzard’]

1.47 Some Uto-Aztecan *-k- > NUA -h-, > SUA -k-, and > ø in Hp, Tb, Eu, Op

TWO; DOS

Mn	wahá-i/tu	Hp	lööyöm	Eu	wodí(m) (gen. woke; acc. wok)
NP	waha(’yu)	Tb	woo/wooh; wooyo ‘both’	Op	gode
			woo’ami ‘twice’	Tbr	nyohór
TSh	waha	Sr	wöh	Yq	wói
Sh	waha/waa-ttïn	Ca	wíh	My	wooyi
Cm	waha	Ls	wéh	Wr	woká
Kw	waha	Cp	wíh	Tr	okwá
Ch	wahá	Od	gook	Cr	wá’apua
SP	waa	Nv	gok	Wc	húuta ‘pair, double’
WM	wáyIni	NT	goóka		’útïmana ‘second (place)’
CU	wáy-ini	ST	gok	CN	oome

570 NUA *wakay ‘two’: Mn; NP; TSh; Sh wahattïwïh; WSh wahattïn; Cm; Kw wahayu; Ch; SP; WM; CU; Sr waah- / wah- ‘twice’; Gb wahá ‘other, companion’; Ktn wah- / weh- ‘twice’; Cr wá’apua.
SUA *wokay / *wokoy: Sr wöh; Ls wéh; Ca wíh; Cp wíh; Gb wehé’; Hp; Tb; Eu wodí(m)/wok; Tbr n^{y}ohór; Yq wói; My wooyi; Wr woká; Tr okwá. Note liquids in Yq and My wo’olim 'twins’ and Tbr in contrast to -y- in Hp, Eu, Op, and Num. While *wakay and *wokay are likely variants of an original unity, UAnists often separate them according to first vowel, which is fine for the sake of tidiness. Both Num and Cr show initial *wa, while the rest of UA rounded the vowel adjacent to w: *wakay > wokay. [Semitic ’axar]

THREE; TRES

Mn	pahí-i/tu	Hp	paayom	Eu	veidúm
NP	pahi’yu	Tb	paai	Op	vaide
TSh	pahi/pai	Sr	paahi’	Tbr	vayí-r
Sh	paih-	Ca	páh / páx	My	bahi
Cm	pahihtï	Ls	pááhay	Yq	báhi
Kw	pehe/peheyu	Cp	páh	Tr	bikiyá
Ch	pahí	Od	waik	Wr	paiká
SP	pai	Nv	vaiko	Cr	waihka
WM	páyIni	NT	váika	Wc	háika; hairíeka ‘third’
CU	pay-ni	ST	vaik	CN	eei

UACV-2623 *pakay 'three': a form of *pakay is in every language above, plus WSh pahaittïn; Ktn pahi’; Gb páhe’; and note Kw peheyu. Note the k syllable in Wr, Tr, CrC, and Tep, in three branches. Note also Ca páh / páx, with an alternate form suggesting *-k- > -x-/-h-. In nearly the same languages as in *wakay ‘two’ above, here also *k > k in Tr, Wr, Tep; *k > h in most of Num, Tak; *k > ø in Hp, Tb, SP, CU, Eu, Op. The -k- is clear in Tr, Wr, CrC, and Tep.
1071 UA *naNkapï ‘leaf’: Kw naga-vï; Ch nanká-va; SP maavï-naŋqa-vï ‘leaf’ (vs. SP naŋqava ‘ear’);
CU nïká-’a-vi (vs. CU nïká-vi ‘ear’); Tb naŋhabïï-l; Hp nàapi/nahpi ‘leaf’. The last three sets show Hp losing intervocalic -k-/-ŋk-, but Hp nàapi/nahpi shows -p- instead of -v-, as evidence of a previous cluster.
170 UA ***tïku** 'drunk': Wr tekú 'be drunk'; Tr ŕiku 'become drunk, sick, faint'; Tr téguri/tékuri 'ebrios, borrachos, pl’. [Egyptian(F) **txw** ‘drunkard’]
170 UA ***tïhu** 'angry': Mn tïhuyee 'be angry'; Sh tuhu" 'angry'; TSh tuupïkkan 'be angry’. In light of other examples of a correspondence between Tr/Wr k and h in Num and other languages (agave, two, three, deer), a relationship between Num *tïhu ‘angry’ and TrC *tïku ‘drunk’ is reasonable. [Egyptian(F) **txw** ‘drunkard’]
638 UA ***tïkïya** 'deer': Mn tïhïta 'deer'; Mn tïhïya 'old buck'; NP tïhïdda; TSh tïhïya(n); Sh tïhïyan; Cm tïhïya 'horse'; Kw tïhïya; Ch tïhíya; SP tïġia ‘deer’; SP tï- 'deer, game'; CU tïíyï. Though the first vowel is problematic, Tb tohii-l 'deer' is likely related, since the other three of the first four segments agree. From Sapir on, some have mixed these with *tïnnV 'antelope’ (< *tïmïna), which is another example of syllable reduction causing a cluster: *tïmïna (Ktn) > tïmna > *tïnna.) For ‘deer’ the SP form shows *-k-, while the other Num forms show -h- or nothing. So again, *k > h in most of Num. [Hebrew ***raxel** ‘ewe, sheep’]
638 UA ***ciki** 'white-tailed deer': Od siiki 'white-tailed deer'; PYp siiki 'white-tailed deer’. In light of the frequency of *ti > ci, this Tep stem (*ciki > Tep *siki) likely ties to NUA *tïkïya 'deer’. The Tep k with Num h (< *k) is consistent with the above terms (two, three, drunk/angry) as well. [Hebrew ***raxel** ‘ewe’]

1.48 Consonant Harmony and Consonant Anticipation

Instances of consonant harmony in UA seem to be consistently regressive or anticipatory: that is, the earlier consonant harmonizes with the next consonant:
1100 UA *tanapiko ‘heel’: among others are My témpe’erim and Yq pémpe’im, Yq’s first consonant harmonizing with the second.
96 UA *tïpa > *pïpa ‘throw, v’: Yq and all of TrC show *pïpa while other branches show *tïpa.
1028 UA *yoli 'live, alive, bear, be born': most reflexes align with *yoli, and so does Cr ruúrikame 'alma, vida’ (Cr u < *o) except that the first consonant harmonized to the second.
665 UA *huCkuN- ‘dust’: while 7 languages show *hukkuNpV, CU kukupï (< *kukkuppï) has consonant harmony.
UACV-2233 *pacay ‘shine’: TO wađađ-k ‘be shiny, bald’; PYp vasad ‘shine, vi’. Consonant harmony in TO.
UACV-1851 *pakwa ‘pus’: Tr bawana/wawana ‘erupcion purulenta, sarna’; Ls ’apáákwaya ‘rotten wood, punk’. Medial *-kw- > Tr -w-, so outside of a preceding vowel that Tr lost or Ls gained, both match *pakwa. However, note the consonant harmony in one of the two Tr variants: wawana.
UACV-1943 *turipa / *tVrV ‘shake’: whether the final *-pa in CrC is a suffix or not, notice that Cr harmonized the second consonant to the third: Wc tïtïriva ‘estar temblando’; Cr rubibéh ‘tiembla’;
Eu turiré nomíkdaa ‘shake, stir’; Hp tïrïrï ‘be shivering, trembling, shaking’.

Anticipatory consonant harmony and consonant anticipation (being moved to the preceding syllable) have in common a consonant being moved forward or repeated forward. Uto-Aztecan does both.
UACV-160 *ku(C/N)ta(N)(pa) 'bee': Cp kutáŋva-l 'bumblebee'; Ls kúúkunta-la 'bumblebee'; My kuta kúmera 'bee that lives in wood'; Nv kuarhagi mumuva 'abejas grandes que hacen panales'; WMU kučávi 'bee'. Ls anticipates the nasalization a syllable earlier than is apparent in Cp, while the SUA languages (My, Nv) do their typical lack of clustered nasalization. WMU -č- (vs. -r-) and Cp -t- (vs. -l-) signify a cluster.
UACV-1194 *(na-)patï(N)kï(N) 'fight, v': Mn pidïkï 'fight'; Mn nanna-pidïkï 'fight one another'; TSh napitïŋkïn / napitiŋkïn 'fight'; Sh napitinka" 'to fight'; Cm nabitïkïrï 'war, battle'; Tb paandïgït 'fight'. WNum and CNum *na-pitïNkï and Tb *paNtïkï show Tb anticipating the nasalization a syllable before Numic's nasal feature, and even Num *pitïNkï may be anticipating nasalization from *pVtïkïN.
UACV-390 *pina 'bring, gather, acquire': Tb pin ~ 'imbin 'bring it'; Sr pinai 'bring, bring back'; Wc piini 'be the property of'; Nv vino'o 'for river to carry s.th.'; Tr bi'ni/be'ná 'recoger uno a uno, pepenar'. Note nasalization anticipation in Tb above and below:

Without nasal anticipation	With nasal anticipation
Tb kiig ~ 'ikik 'to sack, store, load'	Tb kam'-(ut) ~ 'aŋgam' 'it fits'
Tb kita ~ 'ikita 'it is locked'	Tb kin-(at) ~ 'iŋgin 'he brings it'
Tb kuuŋut ~ 'uuguuŋu 'she married'	Tb kumaawa'(ït) ~ 'uŋgumaawa' 'it is shady'
Tb kamiiž ~ 'akamiič 'to catch it'	Tb paam ~ 'ambam 'make into a ball'
Tb paabï ~ 'aabaabï 'be tired'	Tb pin ~ 'imbin 'bring it'
Tb pacaa'in ~ 'apacaa'in 'he caches'	Tb paan ~ 'amban 'to close it'
Tb tomocka ~ 'otomocka 'to stumble'	Tb tana ~ 'andana 'to get down'
Tb tuluumiin ~ 'utuluumiin 'to roll his blanket'	Tb taŋ ~ 'andaŋ 'it is raining'
Tb tulu'uma ~ 'utulu'uma 'it rolls'	

The Tb telic (perfective) form generally reduplicates the first vowel. If the second consonant is a nasal, sometimes that nasalization is anticipated with the prefixed vowel, but not always. The cognate languages show no inherent nasalization in front of the verb stem, so it must come from anticipating the nasalization two consonants away. This phenomenon may explain Tb's nasalization in other places.

Besides nasals being anticipated, glottal stops frequently jump to the preceding syllable, and liquids on occasion. This glottal stop hop or anticipation occurs often in TrC, especially in Tr and Wr, and Sapir (1930, 59) noticed the glottal stop's mobility in SP. I have also noticed it in WMU.
8 and UACV-400c Note the glottal stop hop at 'carry' in Tr ca'pi 'coger' vs. Tr na'cabi 'coger pl obj's.
UACV-153 *ci'ma / *(C)a'cima 'beautiful': Tr či'má in Tr či'má(k)ame 'precioso, primoroso, bello'; Tr či'má-re-ma 'ser bello, primoroso, precioso'Cp á'čimal 'pretty, nice'; PYp la'sima 'beautiful'. With additional prefixes in Cp and PYp, the glottal stop hops, as all agree in five segments otherwise—(')ci(')ma—and PYp s < *c.
724 While other forms point to ***paro'osi** 'jackrabbit' at 'rabbit' (such as My paaros, pl. paró'osim), Wr pa'loísi and Tr ba'loísi anticipated or transposed the glottal stop a syllable forward.
UACV-210 Among forms of ***curaka'i** 'bird, woodpecker' is Wr cu'rukí 'bird' with the ' moved two syllables forward.

1.49 Vowel Behavior (or Misbehavior) in Uto-Aztecan

Early on, Sapir (1913, 402) noticed that "most UA languages seem to assimilate vowels of successive syllables to each other to some extent, though in varying manner." He also noted the frequency of vowel syncope and that the existence of many consonant clusters was due to it (Sapir 1913, 415). In fact, Sapir (1913, 417) goes so far as to say, "In Nahuatl (as presumably in UA generally) there were no consonant clusters to begin with. All present clusters have been brought about by the disappearance of short vowels." I vary from that view only slightly: even if many present clusters were brought about by vowel syncope, there were also original clusters, even if many are largely now lost, but sometimes perceptible in the reduction of the old cluster to a single consonant, whether the components of the cluster are retrievable or not.

The UA vowel correspondences are fairly straightforward and obvious by inspection of table 6 (page 46). Hopi shifted them one direction (*u > o; *o > ö), while the Corachol languages shifted them the other (*u > ï; *o > u). CN continued the CrC shift one step further: *u > ï > i. The Tak languages offer less obvious scenarios, treated by Langacker (1970), who also explains PUA *k > Cup q/_o, which q remained even after *o became high front vowels in Cupan: Tak *ko > *qo > qe (Ls) / > qi (Cp, Ca). Examples are at *kuta 'neck'; *koloka 'beads'; and elsewhere.

Vowels > i/ï/e in Unstressed Syllables

Vowel centralization is common in language change. Sapir (1913, 416) noticed that many vowels appear to change to *i* in shortened/aspirated syllables and that a 'dulling' to ə is common in SP in unaccented syllables (Sapir

1930, 8). This is similar to the schwa-phenomenon in English, wherein short unaccented vowels of longer words become ə. The UA schwa-equivalents are i and ï/e.
UACV-504 *(pa)-hawa 'fog, steam': Yq báhe(wa) 'fog'; AYq haawa 'vapor, steam, n'; AYq vahewa 'mist, fog'; AYq vaiweče 'fog, mist'; My baihwo 'neblina, brisa'; My háawa 'vapor'; Eu baúua (baúwa) 'rocío, neblina'; Eu beiwat 'neblina'; Ca háway 'be foggy, vi'; Ca háway-š 'mist, fog'. The diachronic fragility of h results in a dipthong and the loss or near loss of the middle syllable after the prefix *pa-. Also of interest is the fact that all forms without the prefix *pa- show *hawa (Ca, My, and one AYq form) because the first syllable was likely stressed, while all forms with prefix *pa- show a higher vowel after pa-, i.e., pa-(h)ïwa/(h)iwa with second syllable reductions, because pa- was stressed and thus not the first syllable of *hawa. Furthermore, those high vowels are the UA schwas, and, like the English schwa, sometimes result from lack of stress in unaccented syllables, not from PUA *ï or *i.
UACV-2601 *hatawa 'yawn, v': Mn na'ïdawï 'yawn, vi'; NP ïdamuwïnï 'yawning, vi'; TSh hïtawa 'yawn, vi'; Cm ïhtamakï'atï 'yawn, vi'; Kw 'atawa 'yawn'; Eu hátawa (prêt: hátauhri) 'yawn'; My ten háha'awa 'is yawning'; Yq háawe 'yawn'; Cr ha'ateewa 'he yawns'. Note a glottal stop in Cah corresponding to *t in the other UA languages: *t > l/r > ' in Cah. Interestingly, in TrC where the first vowel is stressed, the *a is retained while second and third vowels sometimes change, but in Num where the second vowel is more often stressed, the first vowel goes to ï, the UA schwa, in all Num forms except Kw.
UACV-1067 *ata(N)kaC 'grasshopper', note the 2nd vowel is consistently *a* in TSh aattaŋki(cci); Sh aattenkih; Cm aatakíí'; Kw 'aataka-piži; SP aataŋkaC, aataŋka-ppici except for some CU variants: CU 'áa-rïká-ci / 'áa-raká-ci / 'aa-taká-ci. In the one CU variant, the unaccented a > ï between two accented syllables. In CU the third vowel is also *a*, so only unaccented schwa-like behavior can explain *a > ï in one of the CU variants.
UACV-1850 *ayakwi 'pus': Cp áyexwi-š / áyaxwi-š 'pus'; Ls 'iyáxwi-š 'pus'. Ls and one Cp form both show an unaccented *a > i/ï*, while accented á remains in all cases.
UACV-1286a *yaCV 'laugh': Mn yawi; TSh yahi/yahe; Sh yahnaiC; Cm yahneetï 'laugh, v sg' vs. Cm na'yïnetï 'laugh, v pl'. The two Cm forms are quite identical except that when the prefix *na- is added, the first vowel a becomes the second, and in the unaccented position becomes *ï*.
676 UA *pakuwa 'mushroom, fungus': Mn paagú 'type of pink mushroom'; PYp vikoga 'mushroom(s); Wr wehkoári 'fungus'; Tr wikubékuri 'large white edible mushroom'; Tr wekogí 'mushroom'; Tr wehorí 'type of edible mushroom'; Tr čohowékuwi 'large white edible mushroom'. The phonological variety in Tr is typical (-weku-, wiku-, béku, weko, weho-) and some forms suggest Tep influence. The Mn, PYp, and one Tr form (-beku-) suggest initial *p, whose reflexes in Tep (v/w) are the loan source of some Tr/Wr forms. The first vowel is probably *a* on the strength of the Mn form, which *a* easily assimilates or centralizes to *ï/e/i* when a greater stress is later in the word.
269 *taka 'fruit' are 11 languages with reflexes of *taka, but Kw tïkïpiya 'fruit' shows *a > ï/_i.
1120 *yuhu 'fat, grease': among several Num *yuhu forms with stress usually on the second syllable, we find Kw yïhuu/yuhuu-vï and CU yïú-vi 'fat, oil, grease, lard' which changed *u > ï when unstressed.
UA *pašwel 'young man': Ca pašwél-iš 'young man'; Cp pišwéliš 'young man'.
93 UA *toci 'head': among other SNum *tocí- forms, all accented on the second syllable, is CU tïcí-vi.
UACV-2614 *pana 'yucca whipplei': Ls panáá-l; Cp pəná-l; Ca pána-l. Note Cp ə < *a in the unstressed syllable.

Additional examples of schwa-like behavior (V > ï/i), usually in unaccented syllables, can be found in the UACV at *malkocowa 'hug'; *paca 'long, thin, stretch'; *patto- 'swell'; and *sakwo > *sikwo/sikwi 'bewitch, whip'; etc.

Uto-Aztecan Vowel Assimilations Anticipating Following Consonants

Uto-Aztecan vowels also move toward the point of articulation of the following consonant, anticipating its place of articulation, though again, more often in unaccented syllables, that is, V > o/u before labials and V > i before alveolar consonants: e.g., Semitic baraq 'lightning' > UA beroq 'lightning' raises and fronts -a- > -e- before -r- and backs -a- > -o- anticipating uvular -q.

Some vowels round before labials: e.g., UA *sa'maC 'spread': Kw sa'ma 'spread out (as blanket)'; Kw sa'ma-pï 'blanket, mat'; SP sa'ma / sam'a 'spread out (a blanket)'; SP sa'mappï 'spread out, ptc, cover on which s.th. is laid'; Ch som'á 'spread a blanket'. Note Ch's assimilation of *a > o/_m. Other examples exist dot the data.

Vowels > i before alveolar consonants, especially in unstressed syllables. Note how often vowels become high-front when preceding an alveolar or when anticipating what might be considered a "high front" consonant:
UACV-108 *paNtu' > *paicu' 'badger'.
UACV-358 *packo'or 'prickly pear sp.': PYp pasko'or 'type of prickly pear'; Tr péčuri 'nopal species'.
1066 UA *corowa 'hungry': Tr ciriwísa exemplifies the raising influence of three of four consonants being alveolar, with perhaps help from assimilation toward the third accented -í-.
UACV-2623 *pakay 'three', Tr bikiyá shows the anticipatory influence of -y-.
308 UA From *pa-surV / sura 'sweat' the last two syllables of Wc kwaašiiya 'sweat, n' assimilate the V toward y, while Cr táisï'e 'sweat, vi' or Cr -sï'e (< *surV) agrees well with all the other *pa-surV/sura forms, mostly of Tep.

Kenneth C. Hill notes that Spanish *frazada* is the source of Hp pösaala, and is the likely source of other UA words for blanket: Ca sáala'a, Tbr pirisál, Yq piisam. Comparing Tbr and Yq, note Yq's quick loss of r since European arrival. Also note the tendency of alveolars to raise and front preceding vowels (a > i/_ before r/l/s/t) in Tbr, Yq.
Hp kapiira is from Spanish cabra. To separate the Spanish consonant cluster, i emerged, perhaps partially due to its schwa properties, though having become a long vowel hardly has it schwa-like any more, so perhaps more likely is the influence or anticipation of r.

Vowels' effects on consonants: besides the palatalizing effect of high vowels (*t > c) discussed above, low vowels (PUA *a and *o) often caused *k > q. *k > q/_a is common in Num, Tak, and Hp, but Tak changed *ko > *qo, then kept q even after the subsequent Cupan vowel changes of *o > i (Ca, Cp) and > e (Ls), which then yield Ls qe and Ca/Cp qi < *ko (Langacker 1970). Examples include 1014 *kuta 'neck'; 630 *koli 'hurt, be sick, chili pepper'; 594 *ko'ci 'older sister'; UACV-1637 *koyni 'plow' at 'plant, v'; and others.

Vowels assimilate to other vowels, anticipating the following vowel or preserving the preceding vowel. Relevant to Sapir's (1913, 402) generalization that "most UA languages seem to assimilate vowels of successive syllables to each other … in varying manner" are *u-a > o-a, *i-a > e-a, vowel leveling *a-i or i-a > e-e, Tübatülabal's preservative vowel assimilation, and Nahuatl's anticipatory vowel assimilations, and Tepiman's anticipatory vowel assimilations, each treated below:

The Partial Anticipatory Assimilation *u-a > o-a

UACV-69c *kuC-taC-pï 'ashes': TSh kuccappïh; Kw kuca-pï; SP kuččaC 'ashes, light gray'; CU kuca-pï; Ls koškuyat 'soot' (vowel is wrong, Miller notes); Hp qöcvi (vowel is wrong, Miller notes). The two vowels that Miller notes as wrong (Ls and Hp) are likely due to *u-a > o-a, because three other forms show *u-a, and *u-a > o-a is natural and explains Ls o; otherwise, Ls o < *ï, which would not work here.
UACV-1734 *hupa 'pull out': Kw hovo 'pull out (hair, grass, seeds), v'; Ch hová 'pull out, v'; Nv 'upana 'arrancar'. The semantics are identical, as are the correspondences nearly, since Nv ' < *h. The only difference is *u-a > o-a in NUA, then Kw further assimilated the second vowel to the first.
UACV-1128 *yula 'hang': Ca yúlaa 'to hang'; Ls yóóra 'to swing, hang in the air'. Ls and Ca are similar except for the explainable vowel assimilation in Ls. That assimilation was later than the one in P175 below, wherein the change was before the Ls vowel shift of o > Ls e: that is, *suka > *soka > Ls *sexa. For note that all of SUA and even Sr in Tak show *suka while Ls has *seka.
1260 UA *LukV 'stoop': Ca lúku 'bend the body forward'; Ls lóóqa 'stoop'. The fact that Ls has final -a allows *u-a > o-a to explain Ls o, as in the next set also and others.
UACV-525 *suka 'to heat, be hot (weather)': Ls šéexa 'to simmer, of water when it is about to boil'; Ls šéx-la 'to warm water'; Eu sukáe-n 'caliente'; Op sukkara; My súkka 'está caliente'; AYq suka/sukkai 'warm';
Tr sukáre 'calentarse'; Wc šïkáa 'caliente'; Cr šïká 'sun'; Cr wa-šïka 'be hot (weather)'; Nv 'ukadida 'calentar, vt'; Nv 'ukagï 'calentarse a la lumbre'; NT uukádyi; ST huukad; TO huukaji. Ls e < *o suggests *u-a > o-a as an intermediate step: *suka > *soka > Ls *sexa.
UACV-354 *yuŋa 'cactus fruit': Hp yöŋö 'prickly pear cactus'; Wc yïna; TO juni 'dried saguaro cactus fruit'. Both Wc and TO agree with *u, and *u-a > o-a likely preceded o > Hp ö, as in P169 and P175 also.
UACV-1289 *uŋa > *oŋa '(feel/be) lazy': Hp ööna 'not feeling like doing'; Hp naa'öna 'lazy'; Sr 'ööŋa' 'lazy'; Cp íŋi-š, pl. í'iŋčam 'lazy'; Cp íŋiču 'be unmoving'; Cr wá-'ïna-ase 'he feels lazy, dragged out'. Note Hp n vs. Tak ŋ as in 'suck'. Also note Cr ï < *u, and *u > NUA *o is easily feasible before a following a.
683 UA *'uma 'be cloudy': Hp oomaw 'cloud'; Tr na'oma 'become cloudy, erased'; Tbr homé-k 'be cloudy'. A reconstruction of the first vowel as *u instead of *o is preferred, as we would expect Hp ö < *o, and Tr sometimes shows o where u is expected anyway, and even if that were not the case, a vowel assimilation or lowering *uma > *oma, a common phenomenon in UA, also explains the Tr and Tbr forms.
UACV-847 *muwa 'father': Kw muwa; Ch móa; SP moa; WMU muuwá-; CU múa; *u-a > o-a in Ch and SP.

The Cupan languages show a vowel assimilation from *kuta > *qola (Proto-Cupan) 'neck' 1014, as well as *yuŋa 'cactus fruit'; *uŋa 'lazy'; *uma 'cloud'; *hupa 'pull out'; *suka 'heat'; and *kuta 'neck'; that is, seven show NUA lowering the round vowel in assimilating *u-a > o-a, while SUA languages do not as much.

Subbranches may do so: WNum does *u-a > o-a in WNum *toka (NP, Mn) at *tuka 'black, night, fire go out';
UA *tuCcaC / *tuCCaC 'dirt(y)': Mn tocábi 'dirty one'; NP tocaggïti 'dirty clothes, v'; TSh tuccaappï 'dirt, dirty';
Ch tucá-vi 'dirt'.
UACV-536 *mura 'ear of grain': *mura > Cah mo'a > mo(w)a): Yq móa 'espiga'; My mówwa 'espigar', while the rest of SUA is consistent with *muLa: TO muda 'tassel'; Nv murhadaga 'espiga'; Eu murát 'espiga'; Wr mulá 'espiga'; Tr murá 'espiga'; Cr mwée-yu 'spike/espiga'; NT muurádadï 'la espiga'.

The Partial Aniticipatory Assimilation *i-a > ï/e-a

Similar to *u-a > o-a, so is *i-a > e-a (or > ï-a) as common in UA.
UACV-742 *kisa 'chicken hawk': Tak and Hp show *kisa (Cp kísi-ly; Ca kísily 'chicken hawk'; Ls páákiš-la 'chicken hawk'; Gb pakísar 'chicken hawk'; Sr paakiha-ţ 'chicken hawk'; Hp kiisa 'chicken hawk'). But SNum assimilated the first vowel to the second or *i-a > ï-a (Kw kïsa-vi 'chicken hawk'; Ch(L) kïsavu 'hawk species').
225 UA *witta > wïtta 'wrap' shows SNum *witta, but *wïtta in CNum and WNum.
UACV-614 *sika / *siki 'cut (hair), mow', Tr has two stems: Tr sikí and a secondary stem Tr seká. Other forms (at 'cut') with 2nd vowel *a* also show the change (> ï-a); yet other forms level the vowels (> ï-ï).
UACV-2028 *huppa 'skunk': among many *huppa forms is CN epa-tl 'skunk' which likely acquired its vowel thus—**uppa > *ipa > CN epa—the last step being i-a > e-a.
UACV-1338 *wina > *wïna 'limp, be lame': Cm wihnai mi'arï 'walk lamely, limp'; Ls wóna 'limp, be lame'. Note the identity of three of four segments (*wVna), with *i-a > ï-a, and ï > Ls o.
630 UA *koli (*kolkoli > *ko'okoli) 'hurt, be sick, chili pepper': While many SUA forms show the reduplication *ko'okoli, Ca and Cp show *koli > *qoli > qili. Then after acquiring final -a, Ca lowers *i-a > e-a: cf. Cp qilyíqa-t 's.th. hot, spicy, strong'; Cp qilyíqatu'nine 'hurt, sting, v'; Ca qélya 'feel sore, v'; Ca qélyak 'peppery, pungent, creating a burning sensation'.

Vowel Leveling

Hopi *e* is the only Hp vowel not aligning with PUA's five vowels, but vowel leveling of *i-a* or *a-i* is often the source of Hp *e*. Ken Hill (p.c.) also mentions reductions of *ai* as a source of *e*, which is another form of vowel leveling:
1457 UA *cikwa 'rain, v': TO siibani 'drizzle, sprinkle' and Hp cekwekwe-ta 'be raining big drops as at the outset of heavy shower' (cekwe- 'soak') suggest *cikwa with vowel leveling in Hp.
UACV-109 *kwila / *kwita 'badger / tejón': Ca wílyaly 'badger'; Tbr kwelé-t/keré 'tejón'.
19 UA *kwiya 'earth, land': most vowels reflect *kwiya, but Tr, Wr, and Cr leveled the vowels *i-a > e'e.
1105 UA *kali 'kidney': SP qaniN-, qanimpi 'kidney' and the k^{y}ele- portion of Hp k^{y}elevosna 'kidney'.
640 UA *piska 'rot, pus, infection' and Hp peek^{y}e 'pus, pus-filled infection'. (*piska is more fully elaborated below under phonological reductions.)
UACV-234 *ciya 'bitter': CN čičiya 'bitter, sour' and Tb ceeyee'ït / 'eceeyeeu 'be bitter' show *i-a > e-e.
890 UA *kani 'house': In SUA: Wr karí; CN kal-li; Tbr kalí-n 'pueblo'. In NUA: NP kani; TSh kahni; Sh kahni; Cm kahni; Kw kahni; WMU kaní; CU káni; Tb hanii-l; and Hp qeni 'place, room, space'. Note how many of the vowel leveling exampes involve Hp.
1095 UA *pisa 'pound': NT viaáhai 'remoler'; Hp pïsïsï-ta 'be a continuous drumming or pounding sound'. With vowel leveling, these agree.
135 UA *mana/mani 'stumble, roll (over), fall over/off/down': Cp máne 'roll, fall off, stumble'; Ca mána/i 'fall down (rolling), roll, stumble over'; Cp manániŋiyqal 'he fell over'; Ls máána/i 'stumble and fall, roll down (a hill) vi, vt'; Sr manamk 'fall down'. Note Hp mïnï(k) 'stumble and fall, fall down' the leveled vowels: *mani > mïnï.
UACV-1391 *laya 'lie with legs/feet spread/pointing outward': The specific semantic identity of Hp lèesi-kiw-ta 'lie with feet pointed outward' and of Ls láya 'lie with legs spread apart' makes this match probable, when we consider that Hp e is usually from vowel leveling, such as a-i / i-a > e-e, or as we have here: aia/aya > ee, as in Ls laya and Hp lèesi, if -si is of another morpheme.
UACV-2358 *ta'ika 'tomorrow': Ch ta'íka 'tomorrow'; Kw te'eka-su 'tomorrow'. Kw again levels the vowels.
1043 UA *mama'u 'woman': While other languages show *mama'u, Kw levels the vowels to Kw momo'o:
Kw momo'o 'woman'; Ch mamá'u 'woman'; Ch(L) mamau'u 'woman'; SP mamma'u-ci 'woman, young woman'; WMU mamá-či 'woman'; CU mamá-ci 'woman'.
2580 UA *pami 'girl': My beeme 'girl'; Yq béeme; AYq veeme; Tr bamirá. Tr probably shows the more original vowels with vowel leveling occurring in Cah: *a-i > e-e.
162 UA *siwa(N) 'sand': While Num shows *siwaN, the TrC terms level the vowels of 'sand' similarly: *siwa > se'e.

Tübatülabal's Frequent Preservative Assimilation of Second Vowel to the First

UACV-1587 *huna 'out(side)': NP hunaggwa 'outside'; Sh hunankwa 'outside'; Cm hunakï 'outside';
Tb 'oonooban 'the outside'. Probably *u-a > o-a > o-o.
6 UA *kwïlu 'swallow': Hp kwelo(-k) 'sample by tasting'; Eu béru'u 'swallow'; Tb weleeh 'swallow'.
Hp and Eu correspond perfectly through 4 segments, since Hp o < *u and Eu b < *kw. With Tb w (< *kw), Tb agrees as well, considering that the second vowel assimilated to the first.
UACV-137 *mo'olV 'bear': Kw mo'orii-ži 'brown or black bear'; Tb mo'olohy 'brown bear'.
206 UA *tuwaC / *tu'aC 'to bear, son, child': among many forms approximating *tuwa'/tu'a, we have
Tb tu'mul 'baby, offspring' which even assimilated the vowel of the suffix *-maL 'small, young'.

829 UA *pit-kanas ‘loincloth, rear-cover’: Hp pitkïna 'kilt, breechclout' and Tb pigiiniš-t 'shirt'; the latter portion likely relates to *kïna ‘cover’ and the *kanas of Cr (at clothing) with preservative vowel assimilation in Tb.
742 UA *comi /*comya ‘hair’: CN comi-, Hp -cmi, Tb comoo-, with preservative vowel assimilation in Tb.
UACV-234 *ciya 'bitter': Tb ceeyee’ït~’eceeyeeu 'be bitter’; CN čičiya 'bitter, sour’; likely *i-a > e-a > e-e.
UA *hu-ma’sa ‘(arrow-)feather’: Hp homasa ‘wing feather’; Tb ’umuša-t ‘arrow feathers’.
677 UA *wakol > *wikol ‘round’: Tep gakod; NP wïkono’o ‘ring, circle’; Mn wigo’onogi ‘crooked’; but Tb(M) wiiginat ~ iwiigin ‘stir, v’.
826 UA *mulawi 'dance, v': Tb muuluwat 'dance, v’; TO mualig '(of a person) to spin or dance’.

Nahuatl’s Anticipatory Assimilation of First Vowel to Second Vowel

162 UA *siwaN ‘sand’: Most of Numic suggests *siwa(N), while most of SUA lost -w- and some leveled vowels, such as My see’e. However, some SUA forms kept the original vowels: Nv hia, TO -hia,
Tbr siha-t, and Wc šie.káari almost. However, CN šaal-li again anticipated the second vowel (iwa > aa), though š is evidence for the original first vowel (AMR 1996d).
UACV-1685 *wiwa ‘amaranth, pigweed’: Hp wiiwa ‘amaranth (pig weed)’; CN waaw-tli ‘amaranth’. Another example of CN’s propensity for assimilating 1st V to 2nd: *wiwa > *wawa > waw.
692 UA *cako ‘small’: Hp cay, pausal acc: càako ‘small, little’; CN coko ‘s.th. very small’. Comparing Hp’s pausal accusative form, CN’s first vowel anticipated or assimilated to the second.
UACV-1739 *(ta)tacowa ‘push’: CN totočoaa ‘to push, shove someone or something to the front’; Tr na’tačo 'push each other'; Cr raa-tátahči ‘lo empuja’; Yq táhta 'bump’. CN assimilated *a-o > o-o.
UACV-1746b *to’asa 'throw': Wc túaša ‘tirar’; Cr tiú’utu’asah ‘tira (piedra)’; CN tlaasa 'throw s.o. down’.
597 UA *taputi 'cottontail rabbit': Sixteen languages match perfectly the four segments *tapu, which consistency is rare in UA. For CN tooč-tli, we have both loss of intervocalic *-p- and a change of first vowel to second: *taputi > *tapoč(i) > *taoč- > CN tooč-. CrC kept the first vowel, but also lost intervocalic *-p-: *tapoci > *tapci > CrC *taciu 'rabbit’ in Wc táciu; Cr táciu’u.
1245 UA *su’i / *suwi ‘hare’: while all of Tak, Hp, and Tb show *suwi/*su’i ‘jackrabbit’, CN si’-tli shows anticipation in *su’i > si’i, then loss of final vowel; though *u > CN i also, no palatalizing s > š.
98 SUA *tïkpa-wa (< *tukum-pa-wa) 'up, above, sky, on': Tr ŕe’pá; Tr ŕe’paní 'sky, up'; Eu téva(n)/tewa;
Tep *tïvagi (< *tïpawi) aligns with *tïkpa-wa (cf. Hp tokpela, Hp l < *w); CN tlakpa-k 'above, on top’. Note that while all others (and others not repeated here) show ï-a, CN has a-a. See ‘sky’ for details on other forms.
1144 UA *o’mana ‘sad, suffering’: CN a’mana ‘sad, troubled’; Tr o’moná-/o’móna- ‘be afflicted, saddened’; Tr o’móna-ri ‘sadness, affliction’. Tr and CN agree in the cosonants -’m-n-, but disagree in vowels: a-a-a vs. o-o-a. Note CN again has earlier vowels anticipating following vowels *o-V-a > CN a-a-a.
UACV-1042 *tapusa > tïposa > tïposi 'gopher': TO jewho/čïwho; PYp tïvua; NT tïvóóhi; ST tïvua; Eu tïvósi;
Yq tébos; Wr te’pósi; Tr repósi. For CrC and Azt, ***tapusa > tausa > tusa** > **tosa**: CN tosan ‘gopher’; Cr tauhsa ‘tuza’. At both *tapusa ‘gopher’ and *taputi ‘rabbit’, CrC kept the first vowel (a), but CN assimilated the first vowel toward the second (a-u > o-o).

Anticipatory Vowel Assimilation in Tepiman: *u-a > ua-a, and *i-a > ia-a

Nevome’s vowel anticipates the vowel on the other side of the consonant in the other languages.
UACV-160 *ku(N)ta(N)(pa) 'bee': Cp kutáŋva-l 'bumblebee'; Ls kúúkunta-la 'bumblebee'; My kuta kúmera 'bee that lives in wood'; **Nv kuarhagi** mumuva 'abejas grandes que hacen panales’; WMU kučávi 'bee’.
1102 UA *suma 'hungry': Eu hisúmrava 'hambre, n'; Eu hisúme 'haber hambre'; Eu hisúm-ce 'tener hambre';
ST uama 'die of hunger’. From *suma > Tep (h)uma > ST uama, as ST anticipates the following vowel.
826 UA *mulawi 'dance, v': TO mualig '(of a person) to spin or dance'; Tb muuluwat 'dance, v';
Tb muuluwii-l 'dance, n’. This pair shows three consonants in agreement. It is plausible that the Tb vowels assimilated between the initial syllable's u and the third C w, or second assimilating to first as above, then with the frequent Tep vowel anticipation, TO's vowels reflect the original, though shifted a syllable forward: *muLawi > mualig.
297 UA *masiwa ‘centipede’: Eu másiwa; Yq masíwe; My masia; TO maihogi; PYp maihig; Nv maiokka (< *mahioga < *masiwa). Wr ma’yáka, Tr maagá/ma’agá, and Tr mahará may derive from Tep loans: *masiwa > Tep *mahiga > mahaga (Tr) and > ma’yaka (Wr). Vocalically TO behaves much like in *muLawi above, anticipating the 2nd vowel, but with rounding toward -w-, a form of anticipation: *masiwa > *maisowV > maihogi.
739 UA *si’a > Tep hi’a ‘urinate, v’: TO hi’a; Nv i’a/’i’a; PYp hia’a. PYp aniticipates the following vowel.
1095 UA *pisa 'pound': NT viaáhai 'remoler'; Hp pïsïsï-ta 'be a continuous drumming or pounding sound’. Note NT anticipatory assimilation and Hp’s vowel leveling.
210 UA *tuti-ka > *cuci-ka > *susi-ka > susa-ka also shows Tep anticipatory vowel assimilation.

Vowel Transposition or Vowel-Line Shift

Another phenomenon frequent in TaraCahitan and sometimes in Tep is what might be called vowel-line shift, transposition, or leapfrog; that is, a sequence of vowels shifts in position relative to the consonants, similar to TO: *mulawi > TO mualig.

UACV-1171 At 'heel' Tr ŕanikura and Eu tenuka have matching consonants (*t-n-k) and the two forms have a similar string of vowels (i/e-u-a), but the vowels have shifted one slot relative to the consonants.
264 At 'rainbow' is another vowel-line shift in these four forms: though the feeble -h- dropped out in Tr/Wr, the vowel pattern persisted, thus shifting the remaining consonants: NT kiihónali 'rainbow'; TO gihonalï;
Wr kenolá; Tr ginorá. Note:

'rainbow' *kihonali (TO, NT) 'heel' Tr ŕanikura
*kinola (Wr, Tr) Eu tenuka

88 among CN wilaka 'caracol de monte'; Ls muvílaqa 'snail'; Tr warákoara 'caracol'; Wr alágaloci 'snail'; Wr nalágeloci 'snail'; Tr narákuri 'snail', note another example of vowel transposition:
Wr a-a-a(l)o-i
Tr a-a-u(r)i

Often *u > ï in Numic

1368 UA *tu'a- 'good': CU tïï'ay 'be good/well'; CU tïï'a-tï 'good'; WMU tïï'a-; Yq tú'i 'bueno, está bueno'; My tu'uri 'be good/well'.
UACV-2069 *suku 'snake, lizard': TSh pa-suku 'water snake'; Mn pasúgu 'water snake'; Tb pišuugat 'red racer snake'; Yq/AYq sikkuča'a 'coral snake'; Ch sïgïpici 'lizard'; CU sïgï-naġóy-či 'lizard'; Kw čigïpi-ži 'lizard' (*s > c?).
622 UA *cukka/*cukki 'crowded, mixed': CN ciciika 'stuff s.th. tight'; SP cïkki 'be mixed with'; CU cïku'mi 'narrow, constricted'; Cm cïhki-/cïkk- 'crowded'. Since *u > i in CN and *u > ï in Num is frequent enough, Num and CN agree through *cuk, and the final vowels (-a vs. -i) are the active/transitive in CN and stative in Num (except CU).
UACV-2300 *hu'uC 'thorn': Kw hu'u-pi-vï 'boxthorn, desert thorn'; Sh hï'ï- 'stickers'.
754 UA *puni 'turn, look, see': Mn puni/poni; NP puni; TSh puniC 'see, look at, study'; Sh puniC/puiC 'see'; Cm puni-tï; Ch puunii 'see, look'; SP pïnni 'see'; CU pïni-kya 'see, vt'; CU pïni-'ni 'look at'.
Hp poni-ni-ykï 'start moving, wake up' is cognate with Num *puni 'see/look', as would the more basic stem Hp poni- 'turn, bend' be also, as in Hp poni-l-a 'turn, make turn, steer' as well as the Tak forms *puni 'turn'. 'He turned to look' and 'he turned' and 'he looked' can all apply to the same instantaneous event. Note that the eastern end of the SNum line (SP, CU) changed *u > ï.
UACV-166 *hupi 'bumblebee': Mn hïbíwu 'bumblebee'; NP huupi nodda 'bumblebee'; Sh hïpi-muih 'bumblebee'.
81 UA *hupi (*huppi?) 'woman, wife': While other UA languages show forms consistent with *hupi, the Num languages show *hïpi/*hïppi (< *hup(p)i): Mn hïïpí'; TSh hïppicci(cci); Sh hïpi; Cm hïbi, though occasional gemination remains to be clarified.
UACV-353 *muCta 'cholla cactus': Cp múta-l; Ca múta-l; Ls múúta-l; Sr muutu|ţ; Sh(C) mïca 'cactus'. While Tak shows u, the Num form has ï, as well as -c- < *-Ct- or *-tt-.
UACV-2319 *yuna/i 'pour, put': Mn tïyuna 'pour into'; Cm payunitï 'pour water on, water, vt'; Ch yuná 'put pl obj's'; CU yunáy 'scatter, put pl obj's'; Kw yïna/yuna 'pour'. Note a Kw form showing yïna < *yuna.

Pima de Yepáchic (PYp) Vowel Metatheses

PYp occasionally metathesizes its first two vowels from a pattern of PUA *a-i > i-a, or *a-u > u-a:
UACV-124 *paCti'a 'bat' several languages illustrate *paCti'a > *paci/*paca, but PYp -pisa < *pica.
UACV-1697 *yalipá 'poison': Mn (y)enipá' 'poison, n'; Mn enipa'a 'poison, v'; Wr yeloá 'poison, n'; Wr yeloé-na 'poison, vt'; PYp dirav 'poison for fish'. PYp fits well, because Tep d < *y and v < *p, and it shows the same metathesis as in 'bat': i-a < *a-i. TrC (Wr) often shows intervocalic -p- > -w- late in a word.
597 From *taputi 'cottontail rabbit' note the vowel metathesis in PYp tuuva 'cottontail'.

Compensatory Vowel Lengthening with Consonant Cluster Reduction

Other examples exist, but the following instroduce the phenomenon of compensatory vowel lengthening in conjunction with consonant cluster reductions: CVCCV > CVVCV. Examples in Tb include Tb(V) paanïnt 'ant' vs. Tb(M) pa'nïnt 'ant'; and Tb(M) polo'mat ~ 'opoloom 'bend, vi'.

Ls also provides examples. At UACV-2386 'touch' are Cp ŋášxa 'be rough'; Cp ŋašxaŋášxa'a-š 'rough, adj'; and Ls ŋááxa/i 'scratch, scrape, vi, scratch, brush against, vt'. These show a cluster in Cp being reduced in Ls with compensatory lengthening of the vowel. In contrast to most Tak terms for 'sky' having no long vowels (Ca túkva-š,

Cp túkva'a-š, Sr tukuhpt̨), we see the long vowel in Ls túúpa-š, which again reduced the cluster. Ls *p remaining a stop (vs. -v-) is evidence of the previous -kp- cluster (*tukupa > *tukpa > *tuupa) with a long vowel in Ls.

Hopi's long vowel with falling tone in some dialects (àa), aspiration in others (ah), usually signifies a previous consonant cluster reduced to one consonant with compensatory vowel lengthening, for -àa- at least and for -ah- if -h- is considered a voiceless vowel continuation of the preceding vowel.

1071 *naNkapV 'leaf': Kw naga-vï; Ch nanká-va; SP maavï-naŋqa-vï 'leaf'; SP naŋqava 'ear';
Tb naŋhabïï-l; Hp nàapi / nahpi 'leaf'. Note that Hp lost -k- / -ŋk- and that Hp nàapi / nahpi shows -p- (instead of -v-) usually due to a previous cluster, and with the reduced cluster, Hp has a long vowel.

221 UA *wïr-pa'a 'tall, long, great-height/length': Hp wïïpa 'tall, long' is a compound of *wïr-pa'a 'big-height/length'. Hp -p- (vs. -v-) means a cluster, yet the first morpheme does not inherently have a long vowel. So the long vowel in the compound is due to a cluster's reduction with compensatory lengthening.

274 UA TO toon-k 'hill'; SP tonnoqqi / tunnuqqi 'a hill rises'. The long vowel in TO appears to be long due to the cluster reduced in TO, but still apparent in SP.

1407 UA *mo'na / *mo'ona > monna / moona 'son-in-law': Sh monappï; Kw mono; SP monna; Hp mö'önaŋw 'male in-law'; Eu mónwa; Wr mo'né; Tr mo'né-ra; My mó'one; Yq mó'one; Tbr moa-saká-r; Wc muune; Cr mú'u 'affinal relative'; mu'un 'yerno'; CN moon-tli 'son-in-law'; Pl muunti; Ca míŋkiw'a. The long vowels in CN, Pl, and Wc are obviously not original, as a dozen other UA forms show short vowels with an intervocalic glottal stop or a cluster (-'n- / -nn-), so the long vowels in the three are secondary and appear to be due to reduced consonant clusters.

With *yu'ma 'tired, worn out' we see clusters in Tb yu'mat~'uuyu'm 'worn out' and Ch yum'á 'tired, suffer, drunk, dead, pl', but without the cluster, we see a longer vowel in Yq yúume 'cansarse' and My yuúme 'se está cansando'. These examples suffice to introduce the fact that consonant cluster reduction with compensatory vowel lengthening is a feature of UA comparative phonology.

The Vowel Changes from Semitic and Egyptian to Uto-Aztecan are treated in section 7.1.

Pattern of Presentation of the Uto-Aztecan and Semitic Data

First is listed the relevant Semitic / Egyptian forms; the most relevant are in bold calibri font. Then is cited the UA reconstruction(s) and the relevant UA set from the reference work *Uto-Aztecan Comparative Vocabulary* (UACV). The UA data are listed thus: UACV-the set number in UACV: then a reconstruction and definition: then the preceding UA cognate collections citing that set: then are listed the UA cognates from the various UA languages, followed by discussion. Some later data and detail, perhaps of interest only to Uto-Aztecan specialists, may be in small print. Then follow a bracket of searchable code for phonological detail, and a bracket of the branches represented by that UA set. Times New Roman is the font for most of the book, but Times New Roman when bolded is less clear, so Calibri font is often used for the primary bolded forms to be compared.

Sections 2 through 5 focus mainly on consonant correspondences of the 1500+ parallels, with occasional comment on vowel correspondences; however, section 7.1 more properly or thoroughly addresses vowel correspondences; section 7.2 shows the medial consonant cluster results in UA; and section 7.3 treats the Near-East grammatical and morphological parallels in UA. Those three normally comprise the comparative method. Yet in addition to those, section 6 shows how these language ties explain seven puzzles of UA previously unexplained. Section 8 reviews the Aramaic leaning of the Semitic-p contribution in UA.