Online Appendix 2

Cognate Sets

The majority of the cognate sets presented here have been proposed in other studies of Uto-Aztecan historical linguistics. The principal published compilations of Uto-Aztecan cognate sets are Voegelin, et al. (1962), Miller (1967), and Stubbs (2011). Stubbs (2011) is the most detailed compilation, including numerous cognate sets not found in the other sources, as well as additional cognates or resemblants for sets previously identified. The author also provides analyses of the regular and irregular correspondences attested, cross-references to other relevant sets, and an extensive bibliography of previous studies.

The cognate sets presented here are intended only to document regular correspondences, in particular the three correspondences that are the focus of this essay: 1) NUA -*r/l*- :: SUA -*r/l*-; 2) NUA -*n*- :: SUA -*r/l*-; and 3) NUA -ŋ- :: SUA /*n*/. I do not include reconstructions of the etyma that the cognates reflect nor, in the case of some sets, all of the cognates attested in the UA languages. In most instances, the cognates that I leave out can be found in Stubbs (2011). I provide, in parentheses following the header gloss of each set, the number of the corresponding set in Stubbs (2011).

The words included in each cognate set show the regular correspondences expected of cognates in the initial syllable or, in the case of disyllabic morphemes, in the initial syllable plus the inital segment of the second syllable. Major deviations from the expected are noted in the "Comments." Glosses are given for individual cognates only when they deviate from the header term or terms that I have assigned to the set. When a cognate has multiple referents and one corresponds to that of the header, the other referents usually are not indicated. The principal

source of data for each language is the first source listed in appendix 1. Data from other sources are included when cognates are not attested in the principal sources or when the data differ between the principal and secondary sources.

My orthography corresponds to that of the Americanist phonetic notation, in which <c> represents the voiceless alveolar affricate, <č> the voiceless post-alveolar and alveopalatal affricates, and <š> the voiceless alveopalatal fricative. I retain the modern technical orthographies developed for each of the languages considered with a few exceptions. I use <ş> to represent a voiceless retroflex sibilant and <ï> instead of <ɨ> as the grapheme for a high, central or back unrounded vowel. I have adopted <ṛ> as the grapheme for the Névome sound represented in the original source with the digraph <rh>. In his study of Tubar, Lionnet (1978) uses <o> as the symbol for an allophone of the phoneme /u/ and <ǫ> (o-comma overstrike) to represent the phoneme /o/. Here I represent the Tubar phoneme /o/ as <o>; there are no attestations of the Tubar allophone of /u/ in these sets. I retain Hill and Nolasquez's use of <e> to represent Cupeño /ə/.

For the most part, I use the citation forms of the original sources, which in the case of nouns often include non-possessed ("absolutive") noun suffixes. Here these suffixes are not separated from the stems to which they are attached. Stress usually is indicated if it is marked in the original sources or is relevant to the analysis. The grave accent in Third Mesa Hopi words represents falling tone. Identical vowel sequences in Yaqui, Mayo, and the Tepiman languages are retained as such: -VV-. Vowel length in Tubatulabal, Hopi, Luiseño, Nahuatl, and Pipil is denoted by a colon following the lengthened vowel: <-V:->. I eliminate initial glottal stops preceding vowels except in the case of Tubatulabal verbs.

2

The cognates in each set are organized by subfamilies in the following order: 1) Numic (NUM), 2) Tubatulabal (TUB), 3) Hopi (HOP), 4) Takic (TAK), 5) Tepiman (TEP), 6) Taracahitan (TRC), 7) Tubar (TBR), 8) Corachol (CRC), and (9) Aztecan (AZT). Most sets lack cognates from several subfamilies, but the order and numbering are retained. The abbreviations for all of the languages and subfamilies are found in appendix 1.

ASPEN (S-559). 1) NUM: NP *kaibasïŋabi* ~ *kaibassïŋaabi* 'quaking aspen' [*kaiba* 'mountain' + *sïŋaabi* 'willow']. NP-Y *sïŋŋabi* 'cottonwood'. EMn *suŋáava* 'cottonwood'. TSh *sïŋapin*. WSh-B *sïnkappin* ~ *sínkappïh*. WSh-D *sïnnapin*. WSh-G *sïnkapin* ~ *sïnnapin*. NSh *sïnaaʔbi* ~ *sanaaʔbi*. ESh *sïna* (Shimkin 1949:208). SP *šïávï*. SUt *sïavi*

BIRD (S-210). 1) Num: WMn *coocoonáʔ* 'western bluebird'. 2) TUB: *culušt* 'woodpecker'. Tb *čïlust* 'type of woodpecker'. 3) HOP: *co:ro* 'bluebird, mountain bluebird'. 6) TRC: Wr *cuʔrú* 'kind of blue bird with a long tail, perhaps a kind of jay'. Wr *cuʔrukí* 'bird (generic)'. Rr *čurugí* 'bird (generic)'. **Comments**: a) Based on the Hopi and Warihó words, PUA **cuʔru could be reconstructed as the label for some kind of 'blue bird'. Hp /o/ :: Wr /u/ is a regular correspondence, and while Warihó usually retains -ʔ- as the reflex of an antecedent morpheme-internal, preconsonantal glottal stop, Hopi seldom does. b) This etymon is the best example I have found for concluding that at least some PUA labels for 'birds' could be onomatopoetic in origin. Western Mono /o/ reflects PNUA *o, so WMn *coocoonáʔ* cannot be identified as a cognate in this set. However, it is analyzed by Bethel, et al. (1993:261) as combining a sound-imitative stem, *coocoo-*, and the suffix -*naʔ*, which denotes 'it says X', "X" being the sound imitated in the stem. In Western Mono, this construction is not restricted to labels for birds. It is attested, for example, in *mááʔ-naʔ* 'mosquito' and in two different names for Bigfoot:

3

qasiqasi-náʔ 'white-coated Bigfoot' and *qauqau-náʔ* 'brown-coated Bigfoot'. According to Bethel, et al. (1993:164-166), "*qasiqasi*" is the sound made by Bigfoot's feet walking on snow and "*qauqau*" when the creature walks on dry leaves. c) It is impossible to determine if the second syllable of the other words in this set could be reflexes of a PUA suffix **-ru with a morphological function comparable to WMn -naʔ. The final syllable in PUA **wiruku 'turkey vulture' [see (4)] definitely reconstructs as **-ru, but **-ri appears to be the required reconstruction for the final syllables of **tukori 'owl' [(3)] and **kïri 'a kind of hawk' [(5)]. Perhaps *-ru-* in reflexes of **cuʔru and **wiruku is the result of vowel harmonization, and a PUA onomatopoetic suffix should be reconstructed as **-ri or **-ti. It is suggestive that /t/ is the initial consonant of words meaning 'to say', 'to talk', or 'mouth' in languages in all UA subfamilies. In addition, intervocalic *-t-* (as the initial /t/ of a suffix would be) has shifted to lenited *-r/l-* in the ancestral languages of Tubatulabal and the Cupan subdivision of the Takic subfamilies, and *-t-* continues to lenite in the Numic languages (Stubbs 2011: 15, passim, #1876a - #1878, #1882; Bright and Hill 1968: 360; Voegelin, et al. 1962: 41, 63; McLaughlin 1992:172).

CHEST (OF THE BODY) (S-427). 1) NUM: NP *niŋaabï*. EMn *núŋava*. TSh *nïŋappïh* 'rib cage area just below the breasts'. WSh-B *nenkappïh.* WSh-D *nïnkappïh ~ nïnnappïh*. WSh-G *nïnkappïh ~ yïnkappïh*. NSh *nïnapï*. ESh *nïnápï*. Cm *nïnapï*. Ch *nïŋapï*. SUt. *nï̃avï*.

CHILD (S-140). 3) HOP: *ma:na*. 5) TEP: To *maḏ(ï)*. PYp *mar*. NT *már(a)*. 6) TRC: Yq *maára* 'daughter (male speaker)'. Yq *maála* 'mother'. Yq-Az *maala* 'daughter (male speaker), mother'. My *maála* 'daughter (male speaker)'. Ed *márwa* 'daughter (male speaker)'. Wr *malá* 'daughter'. Rr *mará* 'daughter of a man'.

CONTAIN, CONTAINER (S-111). 1) NUM: SP *taŋi* 'to put into'. 3) HOP: *taŋata* 'put into a rigid and/or enclosed container or structure'. 4) TAK: Kt *taŋatat* 'bag, box, quiver, tobacco bag'. 9) AZT: Na-Cl *ta:naʔtli* 'basket with a handle, woven of palm'. Pp *ta:nah* 'traveling bag, bundle, pack'

FLY (S-43; S-915). 1) NUM: NP *aŋibi* 'gnats'. WMn *anípi* 'big mosquito found in the mountains'. TSh *aŋipi ~ aŋimmuih*. WSh-B *anamuih ~ ankamuih*. ESh *ániwuʔi* (Shimkin 1949:203). Kw *anivi* 'gnat, fruitfly'. SP *aŋívi* 'mosquito'. Ch *aŋivi* 'gnat'.

FOOT, HOOF (S-937). 2) TUB: *tanaʔpit ~ tanaʔpil* 'little baby heel, knuckle'. 3) HOP: *tana*. 5) TEP: To *taḍ*. Nv *taṛa*. PYp *tar*. NT *tára*. 6) TRC: Ed *tarát*. Wr *talá* 'sole of the foot'. Rr *ŕará* 'foot, sole of the foot, footprint'.

GENTLE (S-2083). 1) NUM: WSh-D *yuun-*. WSh-G *yuun*. WSh-G *yuni* 'soft, spongy'. 8) CRC: Hc *yiiríme* 'soft, tender, young'

HUSBAND (S-1240). 1) NUM: NP *kuma*. EMn: *gúmaʔa* 'son's wife's father, daughter's husband's father (female speaker)'. WMn *kúwa*. TSh *kuhma*. WSh-D *kuhma ~ kuha* 'male, gelding'. WSh-D *kuhmappï ~ kuhappï* 'husband'. WSh-G *kuhma ~ kuha* 'husband, male animal'. WSh-G *kuhmahpi* 'husband'. Cm *kuhma* 'male, man'. Cm *kumahpiʔ ~ kumahpiʔ*. Kw *kuhma*. SP *kummá*. Ch *kuma*. SUt *kumáavi* 'macho'. 2) TUB: *ku:ŋan*. 3) HOP: *ko:ŋya*. 4) TAK: Ca *kúŋlu* 'to propose to marry (of a woman); to have intercourse'. Cp *kuŋ*. Ls *kú:ŋ*. 5) TEP: To *kun*. Nv *kuna*. PYp *kun*. NT *kúna*. ST *kun*. 6) TRC: Yq *kuna*. AYq *kuuna*. My *kuúna*. Eu *kúngʷa*. Wr *kuná*. Rr *kuná*. 7) TBR: *kuna* 'to marry (of a woman)'. 8) CRC: Cr *kiïn*. Hc *kïna*.

Comment: The Numic cognates reflect Proto-Numic *kuma, while all the other cognates reflect PUA **kuŋa. Kaufman (1981:118, 237) interprets the *-m- in PNumic etymon as the

result of assimilation to the round vowel *u in the first syllable. Based on research in progress, I reconstruct Proto-Central Numic(PCNumic) *kuhma and attribute the *h to a shift from antecedent vowel length to preaspiration (*kú:mi > *kuhmi) that occurred in Proto-Eastern Numic, the ancestral language that gave rise to Proto-Central Numic and Proto-Southern Numic (Freeze and Iannucci 1979; Babel, et al., n.d.). Central Numic -*hC*- clusters regularly correspond with Southern Paiute lengthened (or "geminated") consonants, including when these clusters derive from vowel loss. An example of the latter is seen in (20): PNUA * káhani > PENumic *kahni > PCNumic *kahni > SP *kanní*.

KNEE (S-941). 1) NUM: EMn *taŋadovo*. WMn *tanabódo ~ tanobódo ~ tonobódo*. TSh *taŋappïh*. WSh-G *tankappïh ~ tannappïh*. NSh *dannapï*. ESh *tankapï*. Cm *tana*. Kw *tanavï*. SP *taŋávi*. Ch *taŋa*. SUt *táavi*. 2) TUB: *toŋo:l*. 5) TEP: To *toon*. Nv *tona*. PYp *toni*. NT *toona*. ST *toon*. 6) TRC: Yq *tonom*. My *tónnom*. Ed *tonót*. Wr *tonó* 'foot'. Rr *ŕonó* 'foot, leg'. 7) TBR: *tonór*. 8) CRC: Cr *tunú*. Hc *tunú*.

LOOK FOR (S-1898). 4) TAK: Ca *hál*. Cp *hále*. Ls *há:l*. 6) TRC: Yq *haríwa*. AYq *hariwa*. AYq *haliwaka* 'to look for tracks'. My *haría*.

LUNG(S) (S-1409). 1) NUM: NP *soŋo*. EMn *jŭŋ ~ jóno*. WMn *sóno*. TSh(M) *soŋo*. WSh-G *sonko ~ sonno*. NSh *sooʔwoN*. ESh *soonko*. Cm *soomo*. Kw *soovï*. SP *soóvi*. Ch *soovi*. SUt *sőövï*. 4) TAK: Kt *šoŋač*. 6) Taracahitan: Rr *sono*.

NECKLACE (S-1505). 4) TAK: Ca *qénxa*(t). Cp *qínxat*. Ls *qénxat*. Kt *konakat*. 6) TRC: Yq *koókam*. My *koókam*. Ed *kórka*. Wr *koloká*. Rr *gorogá*. 9) AZT: Na-Cl *ko:skatl*. Pp *ku:skat*. **Comments**: a) The Yaqui and Mayo cognates show the loss of -r-. b) The occurrence of -*s*- instead of -*l*- in the Aztecan terms for 'necklace' is an expected secondary development. In

Proto-Aztecan or an earlier post-PSUA ancestral language, the vowel following -r- was lost, resulting in the juxtapositioning of -r- and -k- and the subsequent replacement of -r- by -s-. The same vowel loss is encountered in the Eudeve cognate *kórka* but without the shift of -rk- to -sk-.

PITCH, GUM, GLUE, SAP (S-1634). 1) NUM: NP *sanapi*. WMn *sanápi*. TSh *sanappin*. WSh-D *sanappin*. WSh-G *sanappin*. Cm *sana* 'sticky'. Kw *sanapï*. SP *sanáppi ~ sannáppi*. Ch *sanapi*. SUt *sanápi*. 2) TUB: *sa:nat*. 3) HOP: *sa:na*. 4) TAK: Ca *sáanat*. Cp *sáanat*. Ls *şá:nat*. Kt *hanat* 'tar'. 9) AZT: Na-Cl *sa:loaa* 'to glue'. Pp *sa:luaa* 'to glue'.

POUR (S-2319). 1) NUM: Cm *payunitï* 'pour water on'. Kw *yuniʔi ~ yïniʔi*. SUt *yunáy* 'to scatter, put (plural objects). 4) TAK: Ls *yuníʔi* 'to sprinkle'. 6) TRC: Ed *dúridaan* 'to be emptied'. Wr *yuʔripúna* 'to empty, throw out a liquid'. **Comment**: The initial syllable *pa-* in Comanche *payunitï* is the reflex of the PUA etymon for 'water'.

PULVERIZE (S-1080). 3) HOP: *pi:ŋya*. 4) TAK: Ca *píŋ* 'to get pulverized, to grind'. Kt *piŋan* 'to crumble (vi)'. 9) AZT: Na-Cl *pinolli* 'flour, something ground'. Pp *pinu(u)l* 'powder or flour for eating or drinking'

RING, RATTLE. 2) TUB: *halala:ʔït ~ ʔahalala:* 'to rattle'. 3) HOP: *qalalata* 'to be ringing, clanking, clacking, clinking, tinkling'. 4) TAK: Ls *ká:ri* 'to ring, croak, belch' 9) AZT: Na-Cl *kala:ni* 'to jingle, rattle'. Na-Cl *kakalaka* 'to rattle'

SALT (S-1865). 1) NUM: NP *oŋabi*. EMn *oŋávi*. WMn *omábi*. TSh *oŋwapi*. WSh-G *onapin*. NSh *onaaʔbi*. ESh *ónaabi*. Cm *onaabi*. Kw *owavi*. SP *oávi*. SUt *öávi*. 2) TUB: *uŋa:l*. 3) HOP: *ö:ŋa*. 4) TAK: Ca *íŋilʸ*. Cp *íŋeyu* 'to salt'. Ls *éŋla*. 5) TEP: To *on*. Nv *ona*. PYp *ona*. NT *onai*. ST *on*. 6) TRC: Yq *oóna*. My *oóna*. Ed *onát*. Wr *woná*. Rr *oná ~ koná ~ noná*. 7) TBR: *onát*. 8)

CRC: Cr *unáh*. Hc *úna*.

SCOOP. 1) NUM: NP-Y *yuŋaʔhu* 'to scoop fish'. TSh *yuŋwah* 'to scoop or dip up'. WSh-D *yunnah* 'to scoop or dip a liquid, to ladle'. WSh-G *yunnah* 'to scoop or dip up a liquid'. Cm *yunarï* 'to skim from the surface'.

SHAKE, SWING (S-1928). 2) TUB: *wilikoʔyat ~ ʔiwilikooy* 'to swing'. 3) HOP: *wi:wila* 'to shake, swing or wave around'. 6) TRC: Tbr *wimwirá* 'to shake, tremble'.

SMALL (S-1356). 4) TAK: Ca *ínišilʸ ~ ínišmal* 'small one'. Ls *alúʔmal*. Kt *anuciʔ* 'small; baby, little finger'. 5) TEP: To *al* 'little'. To *ali* 'baby, child'. PYp *laʔali ~ lahali* 'boy, youth'. NT *áli*. ST *ályii*. 6) TRC: Yq *ili*. My *iliči ~ iliʔiči*. **Comment**: The -*l*- in the Luiseño term probably is a secondary development in which -*n*- was denasalized to -*l*-. A similar denasalization of -*n*- is attested in the Luiseño term *čóri* 'to roll', which is derived from PCupan *čïnï. The other Cupan cognates are Cahuilla *čénen* and Cupeño *čéne*.

SWALLOW (S-785). 2) TUB: *welehat ~ ʔeweleh.* 3) HOP: *kwelok-* 'sample by tasting'. 6) TRC: Ed *béruʔu*.

THINK (S-2284). 3) HOP: *iʔna* 'to recall, remember'. 4) TAK: Ca *eʔnan* 'to know, recognize, learn, find out'. Ls *óʔna* 'to know, recognize, be acquainted with'. Kt *ïn* 'to know, know how, understand'. 5) TEP: To *ili∂* 'to think (about), decide, conclude, wish'. NT *ilíídïi* 'to think, believe, want'. ST *ilʸiid* 'to think'. 6) TRC: Yq *ea* 'to think, reflect, want'. My *eiya* 'to think'. Ed *erám* 'to think, feel, want'. Wr *eʔláni ~ elamá* 'to think about, think to be so, be concerned about, be considerate'. 8) CRC: Hc *érie* 'to think, believe, feel'. 9) AZT: Na-Cl *ilna:miki* 'to remember, reflect on something'. Pp *elna:miki* 'to remember'. **Comments**: a) The Yaqui cognate shows the loss of -*r*-. b) In the Mayo cognate, the /*i*/ is unexpected while the /*y*/ is

inserted to break up a three-vowel cluster.

TURKEY VULTURE (S-343). 1) Num: NP *wiho*. EMn *wihópi*. WMn *wího*. Mn *wiho*. TSh *wihnumpicci*. TSh(M) *wihumpiccih*. Kw *wikumahaazi*. SP *wikkumpïci*. SUt *wẖkúcigetï*. 2) TuB: Tb(V) *wišokombišt* 'song of the turkey buzzard'. 3) Hop: *wisoko*. 4) Tak: Kt *wirukuht*. 6) TrC: Yq *wiíru*. My *wiíru*. Rr *wirú*. 7) Tbr: *wilú*. 8) Crc: Cr *viskï*. Hc *wirïkï*. 9) Azt: Na-Cl *wi:lo:tl* 'dove'. Pp *wi:lut* 'bird, dove'. **Comments**: a) PUA **wiruku can be reconstructed based on the Kitanemuk and Huichol cognates; Huichol /ï/ reflects PUA **u. b) Except for the medial -*s*-, the Hopi term corresponds precisely to the Kitanemuk and Huichol terms (Hopi /o/ is the reflex of PUA **u), but a shift of medial *-r- to Hopi -s- is unusual, suggesting that it is not cognate. c) The distinct referents of the Aztecan terms also raise doubts about their cognacy, despite their phonological similarities to **wiruku (Jane Hill, personal communication 2011). d) The Numic terms show a variety of secondary developments, including the shift of the intervocalic *-r- to other consonants and the loss of the final or medial syllables. A medial -*n*-, postulated as the regular Numic correspondence of -*r*- or -*l*- in the other UA languages, is attested only in Timbisha Shoshone.