

Origin of the Chukchee–Kamchatkan language family from the Paiwan language in Formosa: Evidenced by Swadesh basic vocabulary comparison, providing basic knowledge for understanding DNA haplotype distributions

Koji Ohnishi¹

¹Faculty of Science, Niigata University, Niigata 950-2181, Japan

(Tel: 81-25-262-5237, Fax: 81-25-262-5237)

¹ohnishi@bio.sc.niigata-u.ac.jp

Abstract: Chukchee-Kamchatkan (ChKm) basic body-part name words (BBPs) in Swadesh 200 basic word vocabulary. were compared with ca. 1,300 basic words of 80 different Austronesian (AN) languages, which has resulted in finding 13 cognates of ChKm BBPs. Out of the 13 ChKm BBP word-items, 5.3 items were found to have cognate in Formosan, and 3.5 items in Paiwan. Chi-squared statistical test concluded that ChKm have evolved from the Paiwan language or its close kin language. The results were compared with the results from similar analyses of Formosan-related non-AN languages and Gilyak (Nivkh). Results were also discussed from aspects of mitochondrial and Y-chromosomal DNA-haplotype distributions.

Keywords: 3-6 key words or phrases in alphabetical order, separated by commas.

1 INTRODUCTION

The Origin of Chukchee-Kamchatkan (ChKm) language family is unknown. ChKm is divided into Chukchee-Koryak branch (Chukchee, Koryak, Alutor) and Itelmen (Itel) branch, the latter consisting of Northern and Southern dialects (N.Itel, S.Itel) of Western Itelmen (W.Itel = Western Kamchadal) and distinct Eastern Itel and Southern Itel (Kurebito [1]). The aim of this paper is to elucidate phylogenetic position of ChKm language family by basic word (BW)-comparison method, and to provide basic knowledge for analyzing recently accumulating data of mitochondrial, Y-chromosomal and HLA haplotype distribution in Eurasia, Oceania, and Americas.

2 METHODS, RESULT and DISCUSSIONS

BBP comparison: Basic body-part name words (BBPs) (ca.32 word items for human and animals, See [2]), defined as body-part name words included in Swadesh 200 basic words (Sw200W), were taken from ChKm vocabulary in the references, Kure* (Kurebito [1]), Worth*, and Bogo*, listed in Table 1, and compared with ca. 1,300 Austronesian (AN) basic words listed in Tryon's "Comparative Austronesian Dictionary" (CAD* in Table 1), and with representative Eurasian languages. AN cognates of ChKm BBPs thus found are listed in Table 2, where BBP items are classified to 5 AN subgroups (FORM, W.MP, C.MP, SHWNG, and OC in Table 1) to which AN cognates most similar to the corresponding ChKm BBPs belong.

Cognates found in Niger-Kordofanian are also given in Table 2. Abbreviations used in Tables 2-6 are listed in Table 1.

Closest similarity analysis (CSA): CSA (Ohnishi [3]) was revised and employed to this study. Cognates most similar (= closest) to these ChKm BBPs were found in the frequencies as below. Cognates were found in 13 ChKm BBP items, out of which 11.2 BBP items are found in AN, and 1.8 items are in non-AN languages. The 11.2 items consists of (**Formosan 5.3 (Paiwan 3.5**, Tsou 1, Atayal 0.8), Western Malayo-Polynesian(=W.MP) 3 (Philippines 1 / Borneo 0 / Sundic 2/ Sulawesi 0), CentralMP= 0.5, Oceanic = 2), as shown in Table 3. These values, which we call here "closest similarity (CS)-scores", are calculated by the method shown detailedly in Table 3, using the cognate data in Table 2. The CS-scores, thus calculated for each of the 5 AN subgroups, tell us the subgroup to which ChKm would be most closely related. In the case of Table 3, Formosan seems to be most closely kin to ChKm.

Chi-squared test: In order to statistically evaluate the result of CSA in Table 3 and Table 4, χ^2 -test for the degree 2 of freedom was employed as below (Snedecor [4]).

$$\chi^2 \text{ is given by}$$
$$\chi_X^2 = \sum_i (Observed_i - Expected_i)^2 / (Expected_i)$$
$$= (O_X - E_X)^2 / E_X + (O_{non-X} - E_{non-X})^2 / E_{non-X} \quad [\text{Eq.1}]$$

By letting

Number of AN languages listed in CAD*: $N_{AN} = 80$.

Number of languages in a subgroup X of AN family: n_X

Number of languages in the remaining subgroup (non- X) of AN family: $n_{non-X} = N_{AN} - n_X$

Number of word items whose closest cognate is observed in subgroup X : O_X

Number of word items whose closest cognate is observed in subgroup $non-X$: O_{non-X}

Number of total number of word items whose closest cognates are found in AN: $O_{AN} = O_X + O_{non-X}$,

we have

$$E_X = O_{AN} (n_X / N_{AN}), E_{non-X} = O_{AN} (n_{non-X} / N_{AN}). \text{ [Eqs 2]}$$

From Eqs 2 and Eq.1, we obtain the value of χ^2 for subgroup X .

For $O_{FORM} = 5.3$ ($O_{AN} = 10.8$), and $O_{Paiwan} = 3.5$ ($O_{AN} = 10.8$), in Table 4, we can conclude $P < 0.05$ for both cases, as is shown in Table 4.

Thus we can conclude that ChKm is most kin to Formosan, especially to Paiwan language. It is very plausible that ChKm have evolved from an ancient Paiwan language or its kin language. Results of similar CSA and χ^2 -test for other possibly Formosan-related languages, Mayan, Eskimo, and Mongolic, taken from Ohnishi [3], are also given in Table 4, for the sake of comparison with ChKm. Further relationship among these possibly Formosan-related languages needs to be analyzed in future.

AN cognates of ChKm non-BBP vocabulary are also searched in CAD*, and listed in Table 5.

In order to know whether or not the Nivkh (= Gilyak) language is related or not related to ChKm, CSA and χ^2 -test were similarly done for analyzing AN cognates of Nivkh BBPs, and the result is given in Table 6 and Table 3. Nivkh is now concluded to be closely kin to Oceanic ($P < 0.005$), and is therefore unrelated to Nivkh.

Relationship to DNA haplotype distribution: Chukchee and Eskimo share some haplotypes such as Q and N3 haplotypes of Y chromosome, A2 and D2 haplotypes of mitochondrial DNA (Sakitani, 2008[5], p.13, p.p.31; Adachi *et al.*[6]). It seems to be interesting to answer the question whether or not these and/or some other haplotype might be shared by some different language groups of possible Formosan origins.

3 CONCLUSION

ChKm is a language derived from ancient Formosan, most plausibly from a Paiwan-like language. Eskimo seems to be considerably kin to ChKm, but seem to have been derived from different branches of Formosan or Formosan-related AN languages. Closest similarity analysis and chi-squared test concerning the distribution of basic body-part word cognates in AN subgroups was found to be an efficient method for finding phylogenical relationship to various subgroups of Austronesian.

Table 1: Abbreviations used in Tables 2 -5.

§ **General:** A < B, B > A: "A has/had been derived from B"; A <> B: "A is genetically (phylogenetically) related to B" (= "A cognates with B"); A << B, B >> A: "A has/had been borrowed from B". pX = protoX (e.g., pJp = proto-Japanese); dial. = dialect, N.=North(ern), S.=South(ern), E.= East(ern), W.=West(ern), C.= Central, wr.= Written, mid./Mid.= Middle, mod= Modern

§ **Language names:** **ChukKm** = Chukchee-Kamchatkan, Chuk.= Chukchee, Itel.= Itelmen | **Esk** = Eskimo | **AN** = Austronesian, FORM= Formosan, MP=Malayo-Polynesian, W.MP= Western MP {PHIL= Philippines, SND= Sundic, SLW=Sulawesi }, C.MP=Central MP, SHWNG= S.Halmahera and W.New Guinea), OC = Oceanic {ADM= Admiralties, W.OC = Western OC, S.E.Solom.= S.E.Solomons, ReOC= Remote OC {MicN= Micronesian, NCal= New Caledonian, Vanu.= Vanuatu, C.Pacif.= C.Pacific, PolyN= Polynesian} | **NiKord** = Niger-Kordofanian, Jxxx (J26, J145, etc.) = language number #xxx in John* (see below), where J1~ J226 belong to Bantu.

§ **References:** | **Bogo***= Bogoras, W.(1917) *Koryak Texts* (Publication of the American Ethnological Society, Vol.5), Leiden | **CAD***= Tryon, D.T.(1995) *Comparative Austronesian Dictionary Parts 1-4*, Mouton de Gruyter, Berlin/New York. | **CED***= Fortesque, M. *et al.* (1994) *Comparative Eskimo Dictionary*, Alaska Native Language Center, Fairbanks. | **John*** = Johnston, H.H.,Sir (1919, 1922) *A Comparative Study of the Bantu and Semi-Bantu Languages*, Volumes I, II, Clarendon Press, Oxford. | **Kure*** = Kurebito [1] | **MTFNT-App*** = Appendix in Ogawa (1935) *The Myths and Traditions of the Native Tribes (Texts and Notes)* (In Japanese), Inst. of Linguistics, Taihoku Imperial University, Taipei. (2nd printing in 1996) | **Savel'eva***= Savel'eva, V. N. and Taksami, Ch.M. (1970): *Nivkhsko-Russkii Slovar'*, Moskva. | **Tak***= Takahashi, M. (1942): *Karafuto Giryaku-go (Sakhalin Gilyak Language)* (In Japanese), Osaka Asahi-Shinbunsha, Osaka. | **Worth***= Worth, D.S. (1969) *Dictionary of Western Kamchadal* (Univ. of California Publications in Linguistics), Berkley and Los Angeles. | **YH*** = Yasumoto, B. and Honda, M. (1978) *Nihongo no Tan'jyou* (The Birth of the Japanese Language), Taishukan Shoten, Tokyo.

Table 2. Distribution of Austronesian and non-Austronesian cognates of Chukchee-Kamchatkhan basic body-part name words (BBPs), shown together with the method of closest similarity analysis (CSA) using CS-scores

§§ 1. Cognate distribution and CS-scores	§ Most similar to W.MP:
<p>§ Most similar to Formosan (FORM): 4 items (Paiwan) = [FORM (Paiwan) = 3, Tsou 1] (Tsou-like: 1 item)</p>	<p>3 items = [W.MP= 3(PHIL 1, SND 2)] # BACK(anatomical): N.Itel. <i>k'aac</i> /S.Itel. <i>k'awac</i> (< pItel. <i>*k'awac</i> < <i>*kawa-c</i>, where <i>*kawa-</i> < (pAN ?) <i>*kawa</i> < <i>*kuwa</i> < NiKord <i>*kua</i> "back") W.MP: (PHIL) Molbog <i>kaʔay likud</i> "behind" (< <i>*kaway likud</i> (Molbog <i>kaʔay</i> < <i>*kaway</i> < > pItel. <i>k'awac</i>), Cf. Aklanon, Palawan <i>likud</i> "back, behind") Note: Might cognate with; (FORM): Paiwan <i>qavan</i> "shoulder" (< <i>*kava-n</i> = <i>*kawa-n</i>) NiKord: J200, J220 <i>-k oan</i>, <i>-kuan</i> "back" (< <i>*kuan</i>) Note: Unrelated to: (W.MP: PHIL) Molbog <i>awak</i> //Esk (W.Greenlandic) <i>avaq</i> "back" # FAT: Chukchee <i>esʔən</i> /Alutor <i>asʔən</i> /Koryak <i>acʔən</i> "fat, grease" (< <i>*asʔən</i>) W.MP: (SND) Aceh <i>asəw</i> "flesh" # FLESH: N.Itel., S.Itel. <i>tʔaltʔal</i> "flesh" (< <i>*tʔal</i> < <i>*dgal</i> < <i>dagal</i>) W.MP: (SND) Batak Toba <i>jagal</i> "meat" (< <i>*Dagal</i> ~ <i>*dagal</i> < Nijer-Kordofanian <i>*dagala</i>) Nijer-Kordofanian: (Bantu) Lu-ganda <i>-dagala</i> "meat" (John*, vol.II, p.344)</p>
<p># EYE: N.Itel. <i>ʔaŋocʔ</i> (< <i>*ʔaŋocʔ</i>) FORM: Paiwan <i>ʔ-əm-əŋ-ʔəŋ</i> (= <i>ʔəŋ-ʔəŋ</i> + <i>-əm-</i>) (< <i>*ʔəŋ</i>)</p>	<p>§ Most similar to Central MP(C.MP)/Esk:</p>
<p># HEAD: N.Itel. <i>kəmtqol</i>, <i>k'əmtqol</i> "head" < <i>*kəmt-qol</i> "hair of head" < <i>*kəm-</i> "hair" + <i>*-t</i> "ʔ" + <i>*qol</i> "head", where <i>*qol</i> < FORM: <i>*qolo</i> "head" , S.Itel. <i>caqol</i> "head" (< <i>*ca-qol</i> < <i>*ca-</i> "ʔ" + <i>*qol</i> "head") FORM: Paiwan <i>qolo</i>, <i>ʔolo</i> (MTFNT-App*, p.2), <i>*qulu</i> (CAD*) "head" (< <i>*pAN</i> <i>*quluH</i> "head") Esk k: pEsk <i>*qulə-</i> "area above" NiKord: (Bantu) [J140] <i>-gulu</i> / [J153] li-gulu / [J155] li-kolo "above, up"</p>	<p>1 item = [C.MP= 0.5, Esk= 0.5]</p>
<p># TONGUE: Chk <i>jiliit</i> (< <i>*jiləjil</i>) /Koryak <i>jiləjil</i> "tongue, language, speech", Autor <i>jiləjil</i> "tongue" (< <i>*jilə-jilə</i> ~ <i>*jilə-jilə</i> < <i>*jilə</i> ~ <i>*jilə</i> < <i>*jila</i> ~ <i>*jila</i> < W.MP <i>*dila</i> "tongue" < <i>*d3ila</i> "to lick") FORM: Paiwan <i>j-il-ilaq</i> (= <i>jilaq</i> + <i>-il-</i>) (< <i>*dilaq</i>) W.MP: (SLW) Da'a <i>jila</i> /Uma <i>jilaʔ</i> // (SND) Madurese <i>ʔila</i> //Batak Toba <i>dila</i> /Aceh <i>dilah</i> // (Borneo) Murut <i>dilaʔ</i> // (P HIL) Tagalog, Aklanon <i>dilaʔ</i> "tongue" (< <i>*d3ila</i> "to lick") ; (SND) Batak Toba <i>dilat</i> /Minangkabau <i>jilat</i> /IndoN <i>mən-jilat</i> /Javanese <i>n-dilat</i> /Madurese <i>ʔilat</i> /Sasak <i>delaʔ</i> "to lick" (<i>*dila</i> ~ <i>*dela</i> < <i>*dela</i> < NiKord: <i>*deda</i> "tongue") NiKord: [J151 a] [J244] <i>*deda</i> / [J242] <i>lelā</i> "tongue" (< <i>*deda</i>)</p>	<p># NOSE: S.Itel. <i>qnasəŋ</i> "nose" (< <i>*qəŋə-səŋ</i> < <i>*qəŋə</i> "to smell ~ nose" + <i>*səŋ(u)</i> "to sniff, to smell") Esk: pEsk <i>*qəŋar</i> "nose" (< pAN <i>*qəŋ</i> "to smell") (< pAN <i>*qəŋ</i> "to smell") OC: Kwaio <i>g'əŋə(-na)</i> "nose" W.MP: (SND) Batak Toba <i>əŋgo</i> "to smell" (< <i>*əŋ-go</i> < <i>*əŋ-</i> < pAN <i>*qəŋ</i>) C.MP: Ngada <i>səŋu</i> "to smell, to smell" W.MP:(PHIL) Kagayanen <i>siŋŋu'tan</i> "to smell (v.t.)"</p>
<p># § Most similar to FORM/Oceanic(OC)/Turkic : 1 item = [FORM(Atayal) 0.3, OC 0.3, Turkic 0.3] # LIP: Alutor, Koryak <i>waməlkəŋən</i> "lip" (Kure*) (< <i>*waməl-kəŋən</i>, where <i>*-kəŋən</i> < <i>*kalŋə(-n)</i> < > OC: <i>kalŋə</i> "language") OC: (W.OC) Mbula <i>kalŋə-</i> "language" (< <i>*kal-ŋə</i> < <i>*kal-</i> "to speak"), (NCal) A'jje <i>kāŋāwā</i> "lip" (< <i>*kala-wā</i>) FORM: Atayal <i>k-um-at</i> "to speak" (< <i>*kal</i> + <i>-um-</i>) Turkic: Chu vash <i>kal-</i> "to speak"</p>	<p>§ Most similar to OC: 2 items</p>
<p># § Most similar to FORM/Eskimo(Esk): 1 item = [FORM(Paiwan)= 0.5, Esk= 0.5] # FAT: S.Itel. <i>qalk'</i> "fat, grease" (< <i>*qal-k'</i>, where <i>*qal-</i> < > FORM: <i>*qalum</i> "fat") FORM: Paiwan <i>qalum</i> (= <i>qalum</i>) "fat, grease" Eskimo: pEsk <i>*qalunəq</i> ~ <i>*qaluniq</i> "grease floating on stew", Gleenlandic Inuit <i>qaluniq</i> "fat, lard" (< <i>*qalu-niq</i>) Mongolic: Written Mongolian <i>qalim</i> "fat, skin cut off with remaining fat" NiKord:(?) [J193] <i>(a-)gali</i> "fat,oil"</p>	<p># HAIR: N.Itel. <i>k'əmkəm</i> "hair of head, hair", S.Itel. <i>k'mki'm</i> "hair of head" (< <i>*k'əmk'əm</i> ~ <i>*k'ink'im</i> < <i>*k'ink'im</i> < <i>*kumi</i> (i)kumi "beard") MP: (W.MP: PHIL) Salangani <i>Blaan kumi</i> // (C.MP) Ngada <i>kumi</i> //Sika <i>ʔumi-ŋ</i> "beard" OC: (W.OC) Roviana <i>yumi</i> // (North & Central Vanuatu) Raga <i>yumi</i> "i- // (Central OC) Rotuman <i>kumkumi</i> /Tahitian <i>ʔmi ʔmi</i> "beard" # HEAD: Alutor <i>lawət</i> /Koryak <i>lewət</i> /Chuk. <i>lewət</i> "head" (< proto- Chukchee-Koryak <i>*lawət</i> ~ <i>*lawət</i> "head" < <i>*law-(ət)</i> ~ <i>*law-(ət)</i> "feather ~ leaf (of head)" < <i>*llau</i> ~ <i>*lau</i> "feather ~ leaf" < <i>*lalau</i> "feather" ~ <i>*dau</i> "hair of head (= leaf (of head) ~ feather" < <i>*dau</i> < pMP <i>*d2ahun</i> "leaf") OC: (W.OC) Manam <i>lalau</i> // (N.C.Vanuatu) Raga <i>lalau</i> "feather" (< <i>*la-lau</i> < <i>*lau</i> "leaf" < <i>*dau</i> "leaf") ; (C.Pacif.) E.Fijian <i>d'au-ni-ulu</i> "hair-of-head" (< "leaf of head", Tongan <i>lau-ʔulu</i> "hair of head" (< <i>lau</i> "leaf" + <i>ʔulu</i> "head") ; Manam <i>dau</i> / Raga, Rotuman <i>rau</i> /Samoan <i>lau</i> "leaf" W.MP: Sasak <i>daun</i> /Aklanon <i>dahun</i> /Bangingi Sama <i>dahun</i> "leaf" Chinese: C antonese <i>t'au</i>²¹ "head" (< <i>*dau</i> "leaf")</p>
<p># § Most similar to FORM/Quechuan(Que.): 1 item = [FORM(Atayal)= 0.5, Que.= 0.5] # to BREATHE: Chuk, Koryak <i>wəjentok</i> /Alutor <i>wəjisʔtək</i> "to breathe" (< <i>*wəj-</i> < <i>*waj-</i>) Quechuan: wayra (= huayra) "to breathe" (< <i>*way-ra</i>) Atayal <i>bayhu</i> "wind" (< <i>*bay-</i>), <i>vāyu-</i> "wind" (Borrowed from IE: Sanskrit); Paiwan <i>vali</i> "air, wind" (< <i>*bāli</i> ~ <i>*bali</i> "wind"), <i>v-ən-ali</i> "to blow" (= <i>vali</i> + <i>-ən-</i>) (<i>vali</i> > <i>*vay</i> < > Koryak <i>wəj-</i> "to breathe") W.MP: (PHIL) Bangingi Sāma <i>baliyu</i> "wind, air", Kalinga Limos <i>bāli</i> "monsoon wind".</p>	<p># § Similar to many subgroups (Not counted in Table 3) : # SKIN: W.Itel. <i>qułxi</i> (Worth*)/Koryak <i>qıłh</i> "skin" (< <i>*gul(i)-xi</i> ~ <i>*kuli(-xi)</i>) OC: (W.OC) Maringe <i>guli</i> "skin" Note: Possibly further cognates with; (FORM) Puyuma <i>koljīt</i> "skin of fluit" //OC: (W.OC) Mbula <i>-kuli</i> // W.MP: (PHIL) Yami <i>ko lot</i> /Molbog <i>kulit</i> // (SND) Sundanese <i>kulit</i> // (SLW) Wolio <i>kuli</i> "skin" ; NiKord: (Bantu) [J25][J26] <i>-kuli</i> "skin". Note 2: S.Itel. <i>kilwiłx</i> "skin, hide" (Kure*) (< <i>*kil-wiłx</i> < <i>*kil-</i> "skin" < <i>*kul-</i> ?) OC: (MicN) Marshallese <i>kil</i> "skin" (< <i>*kul</i> ?)</p>

Table 3. Closest similarity analysis of the cognates found in Table 2. Method of calculating Closest similarity (CS)-scores for each language group or subgroup

§§ 1. CS-scores found in Table 2.	
§ Most similar to Formosan (FORM): 4 items (Paiwan) = [FORM (Paiwan) = 3, Tsou 1]	(i)AN= 11.2 (= 5.3 + 3 + 0.5 + 2.3)
§ Most similar to FORM/ OC/Turkic : 1 item = [FORM(Atayal) 0.3, OC 0.3, Turkic 0.3]	FORM = 4(Paiwan 3, Tsou 1) + 0.3 (Atayal) + 0.5(Paiwan) + 0.5(Atayal) = 5.3 (Paiwan 3.5, Tsou 1, Atayal 0.8)
§ Most similar to FORM/Eskimo(Esk): 1 item = [FORM(Paiwan)= 0.5, Esk= 0.5]	W.MP = 3 (PHIL 1, SND 2)
§ Most similar to FORM/Quechuan(Que.): 1 item = [FORM(Atayal)= 0.5, Que.= 0.5]	C.MP = 0.5
§ Most similar to W.MP: 3 items = [W.MP= 3(PHIL 1, SND 2)]	OC = 2 + 0.3 = 2.3
§ Most similar to Central MP(C.MP)/Esk: 1 item = [C.MP= 0.5, Esk= 0.5]	(ii) Non-AN: 1.8 (= 1 + 0.5 + 0.3)
§ Most similar to OC: 2 items	Esk = 0.5 + 0.5 = 1
	Quechuan = 0.5
	Turkic = 0.3
	Total sum = 13 items
§ Sum of scores:	
SC-scores for each language groups or subgroups obtained in §§ 1 above are summed up as follows;	§ Result :
	AN 11.2 = (FORM 5.3 (Paiwan 3.5 , Tsou 1, Atayal 0.8), W.MP 3(PHIL 1 /Borneo 0/SND 2/SLW 0), C.MP=0.5, OC=2)
	Non-AN 1.8 = (Esk 1, Que.= 0.5, Turkic 0.3)
	Total = 11.2 + 1.8 = 13 BBP items

Table 4. Occuring frequencies of Austronesian cognates of Chukchee Kamchatkan and Nivkh languages shown together with other possibly Formosan-derived language groups and with Nivkh.
(Results of χ^2 -test are indicated by: #, P<0.1, +, P< 0.05; ++, P< 0.025; *, P< 0.01; **, P< 0.005)

Subgroups of AN	FORM	W.MP	C.MP	SHWNG	OC	Non-AN languages	Number of word items whose closest cognates are found in AN
	(PHIL/Borneo/SND/SLW)				(ADM/W.OC/ SE.Sol/ReOC)		
Number of languages in CAD*	4	26 (11/ 2/ 9/ 5)	6	2	41 (1 /15/ 2/ 23)		
§ 1. Formosan (FORM)-like							
Chukchee-Kamchatkan (BBP)	** <u>5.3</u> (Paiwan:**3.5)	3 (1/ 0/ 2/ 0)	0.5	0	2	Esk.1, Quechuan 0.5, Turkic 0.3	10.8
Mayan (BBP)	** <u>7</u>	2(0/0/1/1)	0	0	3(0/2/0/1)	Mongolic 1?	12
§ 2. FORM-W.MP-like							
Eskimo (BBP)	** <u>4.5</u>	<u>7.5</u> (# <u>5</u> / 0 / <u>2.5</u> /0)	2	1	4.3(0/ 1.5/ 0.3/ 0.5)	Itelmen 0.3/Mong. 0.3	18.3
Mongolic (BBP)	** <u>3.5</u>	++ <u>8.5</u> (1/ 0 /** <u>5</u> / 0)	0	0	2 (0/1.5 /0/0.5)	Esk. 1.5, TRKc 2	14
§ 3. Oceanic-like							
Nivkh (=Gilyak)(BBP)	0	1	0.5	0	** <u>13.5</u> (0 /6.5 /1 /6)		15

REFERENCES

- [1] Kurebito M (ed.) (2001), *Comparative Basic Vocabulary of the Chukchee-Kamchatkan Language Family:1* (ELPR Publication Series A2-011).
[2] Ohnishi K (2012) Prehistorical multiple movements of modern humans from Old World to Americas: Evidence based on comparing basic body part-name words with Austronesians. In: *These Proceedings*.
[3] Ohnishi K (1999), Evolution of Mongoroid

- Languages. Shokado, Kyoto.
[4] Snedecor GW (1956), *Statistical Method* (Fifth ed.), Iowa State Univ. Press, Ames.
[5] Sakitani M(2008), A New History of Hokkaido, as viewed from Interdisciplinary Researches of the sciences of DNA, Archeology, and language. (In Japanese), Bensei Shuppan, Tokyo.
[6] Adachi N, et al.(2011), Mitochondrial DNA analysis of Hokkaido Jomon skeletons. *Am. J. Phys. Anthropol.* 146:346-360.

Table 5. AN and non-AN cognates of Chukchi-Kamchatkan vocabulary other than basic body-part names

§ Most similar to AN:
 # BODY: Chuk əwik /Alutor, Koryak uvik /S.Itel uwik (< *əwik ~ *əvik < W.MP: *awek ~ *avek) ||| W.MP: (SND) Sundanese, Balinese, Sasak awak /Javanese awa? (< *awak < *pMP *hawak “body”) || **OC: Paamese** avek ||| NiKord: J254 -aba “body”
 # COLD: N.Itel ləqləχ /S.Itel lqləχ (< pItel *ləqləχ) ||| **FORM: Paiwan** l'a-l'əkəl “cold” (< *l'əkəl)
 # to GO: Koryak jeləŋ təlek “to go” (jeləŋ < *jəlan < pAN *Zalan “road”) ; S.Itel təlekas “to go” (< *dalekas < *dalekas, where *dale- < *Zalan) ||| **FORM: Paiwan** jaŋan / Rukai ka-daŋan-anə “road”
 # GOOD: S.Itel melaχ ||| W.MP: (SND) Balinese mələh
 # GRASS: N.Itel, S.Itel isi “grass” (< *iʃi) ||| **FORM: Bunun** iʃimo:t “grass” (< *iʃi-mo:t), ??Paiwan *tsimil* “grass” (< *itsi-mul ?)
 # LEAF: N.Itel. pləpl /S.Itel. pələpəl (< *pələ-) ||| O C: (MicN) Marshallese p^wələk (< *pələk ~ bələk ?) || | Note: Cf. (FORM) Atayal *pai* “leaf”
 # LOUSE: N. Itel, S.Itel məlməl /Chuk. məmələ /Alutor mə mələ “louse” (< *mə(1)məl < *mələ “louse ~ flea”) ||| C.MP: Manggarai, Ngada mələ “flea”

ONE: Chuk., Koryak ənnen /Alutor ənnan “one” (< *ənnan < *hənnan < *hannan only” < > Indonesian haña “only”) ||| W.MP: (SND) Indonesian haña “only” (CAD*, #13.330), (SLW) Bugis ənnəŋ, inniŋ /Konjo annaŋ “six” (< *ənnəŋ-məlləŋin “6” < *ənnəŋ- “one” + * məlləŋin “5”, where *ənnəŋ- ~ *annaŋ- < > Indonesian haña “only”.) || C.MP: Manggarai *hanaŋ* “alone, only” ||| KOR: Middle Korean *həna(h)* “one” (< > Indonesian haña “only”.) ||| Note: pMP *h- > ChKm *Ø- (zero).

§ Most similar to non-AN

ELBOW: N.Itel. lot'əŋ /S.Itel. lot'əl “elbow” ||| Comecrudo lot “arm”
 # FOUR: Chuk ɲəraq /Alutor ɲəraqə /Koryak ɲəjaq “four” (< pChukchee *ɲəjaq ~ *ɲerjaq < > Finnish *neljä* “four”) ||| Korean: Mid.Korean nāi(h) “four” (< *nəri < *nəj < > pChukchee *ɲəjaq //Finnish *neljä* “four”) ||| Uralic: Finnish *neljä* “four” (< *nelj(a))
 # GRANDFATHER: S.Itel *mitix* “grandfather”, *mitix* “grand-mother” (<< Ainu ?) ||| Ainu *miti* “father” (<< ChKm ?)

Table 6. Distribution of Austronesian and non-Austronesian cognates of Nivkh (= Gilyak)* basic body-part name words (BBPs).

§ 1. W.MP-like: 1 item

SALIVA: (Amur) *lils* “saliva” ||| W.MP: (PHIL) Yami *lila* “tongue e”

§ 2. C.MP/W.OC(Yabem)-like: 1 item (0.5, 0.5)

LIP: əvləχ (< *ə^mveləχ < *ə^mveləř < *abelar ~ *i^mbela-) ||| **C.MP: Dobel** ɲabelar “tongue” || OC: (W.OC) Yabem *i^mbela* “his/her tongue”, *i^mbela-m* “your tongue” (CAD*)

§ 3. OC-like: 13 items = (0, 6, 1, 6)

§ 3.1. W.OC-like: 6 items

ARM: *tot* ||| OC: (W.OC) **Takia** to-n “arm”
 # BELLY: (Sakh.) *nauf*, ɲaur “belly” ||| OC: (W.OC) **Maringe** *naūafa* “heart” (< *naū- ~ *naūa- ?)
 # FEATHER: *tuɲr* (< *dup-r) ||| OC: (W.OC) **Adzera** *dzuf-* “to fly” (< *dup-), *dzə²-dzuf* “bird”, Tawala *lupa* “to fly” (< *dup-a ?) ||| Jpn: *tubasa* “wing” (< *duba-sa ?)

(Yabem-like) 3 items

EAR: (Amur) *nos* “ear” (< *no-s < *ɲo- “to hear”), (Sakh.) *noχo-* “to smell badly, to smell), *noχjai-* “to smell badly” ||| OC: **Yabem** -ɲò /Kiribati *oŋo* /Kwaio *loŋo-a* “to hear”, Tawala *nonoil* /Maringe *nomo* “to hear”
 # HEART/LIVER: (Amur) *ɲif* / (Sakh.) Trambaus dial., Tyk dial. *ɲif* (Murasaki*), *ɲif*, *nif* (Tak*) “liver” (< *ɲip ~ *nip) ||| OC: (W.OC) **Yabem** (*ɲa-nip-kaləp* “heart”
 # NECK: (Amur) *q'os* (YH*) // (Sakh.) *kōř* (Tak*) (< *ko-ř ?) ||| OC: (W.OC) **Yabem** *kəʔ-labeŋ* “throat” || ?W.MP:

(Sundic) Sasak *korok* “neck”. ||| Note: Cf. (OC:NCal) Nengone *ko den* “nape of neck” | (PHIL) Sarangani *Blaan kolloŋ* “neck”.

§ 3.2. S.E.Vanuatu-like: 1 item

TOOTH: (Amur) *ɲəys*, (Sakh.) Tyk dial. *ɲuɲzur*, Trambaus dial. *ɲəɲzuf* “tooth” (< *ɲəɲ-zur) ||| **OC: Kwaio** *nago* “gums”. Cf. Roviana *ɲaɲadolo* “gums”.

§ 3.3. ReOC-like: 6 items

BELLY: (Sakh.) *komk* (Tak*) // (Amur) *q'oχ* “belly” (Tak*) (< *kom-k) ||| **OC: Rapanui** *kokoma* “intestines, guts” (< *koma) ||| W.MP: (SLW) *ko^mpo* “stomach”
 # BLOOD: (Sakh.) *soχ*, *tfof*, *tfoχ* (Tak*) ||| OC: (MicN) **Woleaian** *cō*
 # EGG: (Amur) *ɲoiq* “egg” (< *ɲoi-q), (Sakh.) *ɲoi* “penis” (Tak*) (< *ɲoi “testis, egg”) ||| OC: **Port Sandwich** *ɲovüč* “yolk” (< *ɲo-vüč ?), Cf. Mbula *ɲonōno* “fruit” (< *ɲo-nōno) ||| C.MP: Buru *oŋo-n* “sphere, ball”
 # FOOT: ɲəcx ||| OC: **Kwamera** *nəsu-* “leg”, Marshallese *ne* “leg, foot”, Ponapean *nē* “leg” ||| W.MP: Sasak *nae* “leg, foot” ||| Jpn: Old Jpn *-ne* “leg” (in compounds)
 # LIVER: (Sakh.) *ɲulf* (< *ɲul-f < *ɲol(u)-) ||| OC: (MicN) *ɲoluŋol* (< *ɲolu-ɲol < *ɲol(u)-)
 # TAIL: (Amur) *ɲəki* // (Sakh.) *ɲaki*, *ɲakki* (Tak*) “tail” (< *ɲa-ki < Nivkh *ɲa* “beast, animal” + *-ki “tail”, where *-ki < > OC: Xaracuu *k^wi* “tail”.) ||| OC: (NCal) Xaracuu *k^wi* “tail”

* Nivkh vocabulary is obtained from YH* for Amur dialect (from Savel'eva*) and from Tak* for Sakhalin dialect.