Ainu is a relative of Sino-Tibetan stock (preliminary notes)

Alexander Akulov

PhD in history, independent scholar; St. Petersburg, Russia; e-mail: aynu@inbox.ru

Abstract

Verbal Grammar Correlation Index (VGCI) can completely answer the question: whether languages are related. VGCI is logical conjunction of two indexes: grammatical meanings sets correlation index and index of correlation of positional distributions of common meanings. If value of VGCI is about 0.4 or more then languages are related; if value of VGCI is about 0.3 or less then languages are not related. VGCI of Ainu and Qiang is 0.41. Due to transitivity of relatedness Ainu is a relative of whole Sino-Tibetan stock. Evidences of linguistics correlate well with those of genetics: Ainu, Qiang, Tibetan have different subclades of Y haplogroup D.

Key words: Ainu language; Qiang language; Sino-Tibetan stock; Ainu relatives; typology; comparative linguistics

1. Some general notes on method

Verbal Grammar Correlation Index (VGCI) is a precise typological method of comparative linguistics. The method supposes direct comparison of really existing/existed languages, and it allows us to see whether languages are related.

The method is based on the idea that language is determined by set of grammatical meanings and by their positional distributions; and degrees of correlations of both issues can be calculated. The method doesn't pay any attention to material exponents at all.

And thus VGCI is superposition (logical conjunction) of two indexes: grammatical meanings sets correlation index and index of correlation of positional distributions of common meanings. Formula for calculation of VGCI is the following:

$$VGCI = \left(\frac{N_{e(A \cap B)}}{N_{e(A)}} + \frac{N_{e(A \cap B)}}{N_{e(B)}} \right) \frac{1}{2} \times \left(\frac{i_1 + i_2 + \dots + i_n}{N_{e(A \cap B)}} \right)$$

where:
A is set of grammatical meanings of one language;
B is set of grammatical meanings of another language;
N_e means amount of elements;
i means index of positional correlation.

More closely related languages demonstrate higher values of VGCI (since they have more alike sets of grammatical meanings, so intersection ratio to each set of grammatical meanings is higher; and due to the fact that common grammatical meanings are distributed in more alike positions); languages that are more distant relatives demonstrate lower values of VGCI and those that are not related demonstrate much lower values of VGCI.

Tests of the method on the material of firmly assembled stocks (Indo-European; Sino-Tibetan; Austronesian) show the following: if value of VGCI is about 0.4 or more then languages are related; if value of VGCI is about 0.3 or less then languages are not related.

Error of measurement is about 2%.

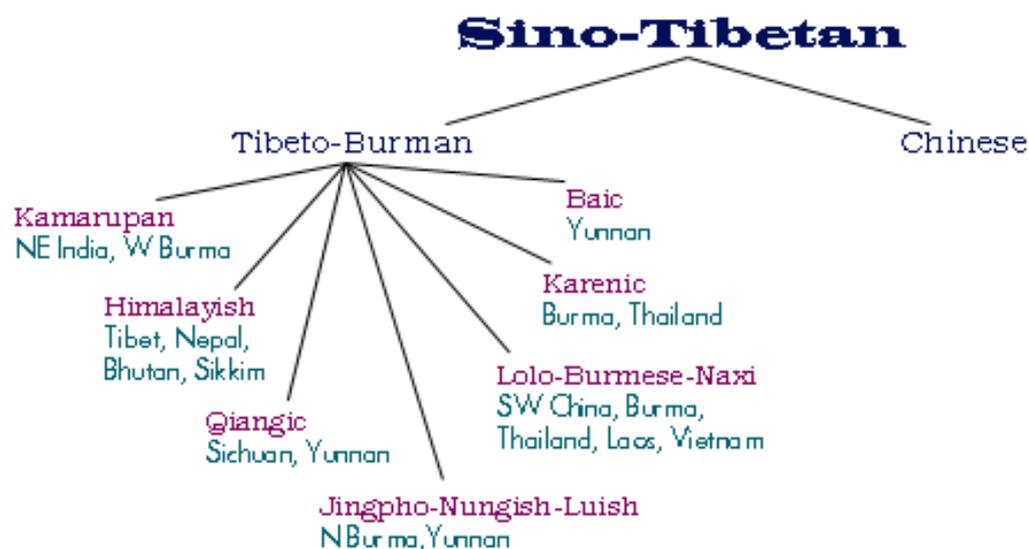For a more detailed description of VGCI method see Akulov 2015b.

2. Comparison of Ainu with Qiang

2.1. List of Qiang forms

List of Qiang forms has been compiled after LaPolla, Huang 1996.

Pic. 1. Scheme representing geographic location of Qiangic languages (grey) among other Sino-Tibetan languages (source The Language Gulper)

Pic. 2. Position of Qiang among Sino-Tibetan languages (source STEDT)

Each grammatical meaning is followed by scheme/schemes representing positional implementations.

Notations of positional implementations are the following:

prp- – preposition;
prfx- – prefix;
-RR- – reduplication;
suppletion;
-sfx – suffix;
-pp – postposition;
zero marker – there is no special marker for this meaning.

If there are some different forms of same position (i.e.: forms used in different contexts) they are numbered in the following way: prp_1-/prp_2-/prp_3- and distinguished by slash; elements that are components of the same implementation are recorded in the following way: prp- + -sfx.

If certain element can optionally be omitted then it is written in square brackets: [prp-].

1. Actor: 6 -sfx
2. Ability 1 learned: -sfx
3. Ability 2 natural: $-sfx_1$/$-sfx_2$
4. Ability 3: -sfx
5. Causative: -sfx
6. Centrifugal version: -pp
7. Centripetal version: -pp
8. Change of state aspect: -sfx
9. Continuative aspect ~ Progressive: prfx-/-sfx
10. Declarative: -pp
11. Deontic modality: -sfx + -pp

12. Directional: 8 prfx-
13. "To dare" to do: -sfx
14. Evidence hearsay: -sfx
15. Evidence mirative: -sfx
16. Evidence see: -sfx
17. Experiential aspect: -sfx
18. Habitual aspect: $-sfx_1/-sfx_2$
19. Hortative: $-pp_1/-pp_2 /-pp_3/-pp_4$
20. Imperative: zero marker
21. Imperfective aspect: zero marker
22. Indicative: zero marker
23. Intentional: -sfx + -pp
24. Interrogative: $-pp_1/-pp_2/-pp_3/-pp_4/-pp_5$
25. Negation: prfx-
26. Optative: $-pp_1/-pp_2 /-pp_3$
27. Patient/non-actor/benefative: 6 -sfx
28. Perfective aspect – resultative: 8 prfx-/ prfx- + -pp
29. Permissive: -sfx
30. Plural: 9 -sfx
31. Prohibitive: prfx-/$-pp_1$/$-pp_2$
32. Reciprocity: -RR-
33. Repetition: -sfx
34. Resultative: prfx- + -pp
35. Simultaneity: -pp
36. Singular: 9 -sfx
37. Subject (of intransitive): 6 -sfx
38. Volition: -sfx

2.2. List of Ainu forms

List of Ainu forms has been compiled after Tamura 2000.

1. Agent: [prp-] +7 prfx-
2. Causative $-sfx_1/-sfx_2$
3. Deontic modality: -pp
4. Desiderative mood: -pp
5. Dubitative mood ~ evidence "may be": -pp
6. Evidence "people say": -pp
7. Evidence: "speaker has seen it" -pp
8. Evidence: "something can be heard/smelt" -pp
9. Evidence "something can be seen" -pp
10. Exclusive $prfx-_1/prfx-_2/-sfx$
11. Frequency: -RR-/suppletion/$-sfx_1/-sfx_2$
12. Future conditional: $-pp_1 /-pp_2$
13. Future tense: -pp
14. Hortative: -pp
15. "If" mood: -pp
16. Imperative (plain): R
17. Imperative (polite): -pp
18. Impossibility mood: -pp

19. Inclusive: prfx-$_1$/prfx-$_2$/-sfx
20. Indicative mood: zero marker
21. Intensiveness: -RR- / -sfx$_1$/-sfx$_2$/suppletion
22. Interrogative: -pp$_1$/-pp$_2$
23. Negation: prp-
24. Patient: [prp-] +7 prefixes
25. Past simple: -pp
26. Past perfect: -pp
27. Plural number: [prp-] +10 prfx- / [prp-] +2 -sfx/ [prp-] +10 prfx- + -sfx / prp- + -sfx +2 -sfx / [prp-] +10 prfx- + suppletion/ [prp-] + suppletion +2 -sfx/ [prp-] +10 prfx- + -RR-/ [prp-] + -RR- +2-sfx
28. Possibility: -pp
29. Present continuous tense: -pp$_1$ / -pp$_2$
30. Present-future (general tense) zero marker
31. Present perfect: -pp
32. Prohibitive mood: prp-
33. Reason/intention: -pp
34. Reciprocity: prfx-
35. Reflexivity: prfx-
36. Singular number: [prp-] +9 prfx-/ [prp-] +9prfx- + suppletion
37. Subject: [prp-] + prfx-$_1$/prfx-$_2$/prfx-$_3$/prfx-$_4$/prfx-$_5$/-sfx$_1$/-sfx$_2$
38. Transitivator: prfx-
39. Unidentified agent: prfx-$_1$/prfx-$_2$/prfx-$_3$
40. Unidentified agent's causative -sfx

2.3. VGCI of Ainu and Qiang

System of recording is the following: first is name of a grammatical meaning that is common for both of compared languages (or meanings that are correlated); then A for Ainu; then scheme of expression/schemes of expressions of this grammatical meaning in Ainu; then sign of correlation "~" or anti-correlation "≠"; then Q for Qiang; then scheme of expression/schemes of expressions of this grammatical meaning in Qiang; and then number showing degree of correlation.

1. Actor: A: [prp-] +7 prfx- ≠ Q: 6 -sfx 0
2. Causative: A: -sfx ~ Q: -sfx 1
3. Deontic: A: -pp ~ Q: -sfx + -pp 1
4. Desiderative: A: -pp ~ Q -sfx 1
5. Durative: A: -pp$_1$/-pp$_2$ ~ Q: prfx-/-sfx 0.5
6. Evidence "may be" ~ mirative: A: -pp ~ Q -sfx 1
7. Evidence "people say" ~ hearsay: A: -pp ~ Q -sfx 1
8. Evidence: "speaker has seen it" ~ see: A: -pp ~ Q: -sfx 1
9. Experiential aspect ~ Past simple: A: -pp ~ Q -sfx 1
10. Hortative: A: -pp ~ Q: -pp$_1$/-pp$_2$ /-pp$_3$/-pp$_4$ (1 + ¼)/2 = 0.625
11. Imperative: A: zero marker ~ Q: zero marker 1
12. Imperfective ~ General tense: A: zero marker ~ Q zero marker 1
13. Indicative: A: zero marker ~ Q: zero marker 1
14. Intention: A: -pp ~ Q -sfx + pp 1
15. Interrogative: A: -pp1/-pp2 ~ Q: -pp$_1$/-pp$_2$ /-pp$_3$/-pp$_4$/-pp$_5$ (1 + 2/5)/2 = 0.7
16. Negation: A: prp- ~ Q: prp- 1

17. Patient: A: [prp-] +7 prefixes ≠ Q: 6 -sfx 0
18. Perfect: A: -pp ≠ Q: 8 prfx-/ prfx- + -pp 0
19. Plural: A: [prp-] +10 prfx- / [prp-] +2 -sfx/ [prp-] +10 prfx- + -sfx / prp- + -sfx +2 -sfx / [prp-] +10 prfx- + suppletion/ [prp-] + suppletion +2 -sfx/ [prp-] +10 prfx- + -RR-/ [prp-] + -RR- +2-sfx ~ Q: 9 -sfx (6/48 + 6/9)/2 = 0.39
20. Potential: A: -pp ~ Q: $-pp_1/-pp_2/-pp_3/-pp_4$ (1 + ¼)/2 = 0.625
21. Prohibitive: A: prp- ~ Q: $prfx-/-pp_1/-pp_2$ (1+1/3)/2 = 0.66
22. Reciprocity: A: prfx- ≠ Q: -RR- 0
23. Repetition: A: $-RR-/suppletion/-sfx_1/-sfx_2$ ~ Q: -sfx (1 + ¼)/2 = 0.625
24. Singular: [prp-] +9 prfx-/ [prp-] +9prfx- + suppletion ≠ Q: 9 -sfx 0
25. Subject: A: + $prfx-_1/prfx-_2/prfx-_3/prfx-_4/prfx-_5/-sfx_1/-sfx_2$ ~ Q: 9 -sfx (2/7+2/9)/2= 0.25

(25/38 + 25/40)/2*(12 + 0.5 + 0.7 + 3*0.625 + 0.39 + 0.66 + 0.25)/25 = 0.41;

VGCI (Ainu; Qiang) = 0.41.

It means that Ainu and Qiang are relatives.

Due to transitivity of relatedness it is possible to state that Ainu is relative of whole Sino-Tibetan stock.

3. Conclusions

3.1. On origin of Ainu

Above described facts are firm evidences of southern origin of Ainu.

3.2. Ainu-Andamanese-Sino-Tibetan stock

Previously was demonstrated Ainu relatedness with Great Andamanese (Akulov 2015a), so now it is possible to speak about Ainu-Andamanese-Sino-Tibetan stock.

Main structural features of the stock are the following: voices are badly elaborated; often can be seen personal markers; also category of evidence is elaborated rather well.

Qiang is closer to Ainu than to Mandarin.

Mandarin language is marginal idiom inside the stock since it has well elaborated category of voice and has no category of evidence.
It seems that Mandarin underwent certain influence of Austric languages since Austric languages have well elaborated category of voice (system of trigger/focuses).

3.3. Evidences of genetics

Different subclades of Y haplogroup D are rather wisely spread among people of Southwest China (Xue et al. 2006). The same Y haplogroup can be seen among Ainu people (Tajima et al. 2004) and among Andamanese (Chandrasekar et al. 2007).

It's notable fact that nowadays geographic spreading of languages preserving old features of Ainu-Andamanese-Sino-Tibetan stock correlates with that of Y haplogroup D (pic. 3).

Pic. 3 Map representing spreading of Ainu-Andamanese-Sino-Tibetan stock; genetic evidences and perspective territories for looking for other relatives of the stock (map drawn after Google maps screenshot)

It seems that initially large spaces of Southern and East Asia were occupied by people with Y haplogroup D and they were the first speaker of the idiom that was proto-language of Ainu-Andamanese-Sino-Tibetan stock.

As it has been noted in 3.1 it seems that initially Mandarin also was a language with set of typical features of the stock, but had been undergone serious influence of certain Austric languages; and this hypothesis also correlate well with data of genetics, since there is almost no Y haplogroup D in the plains of China. It seems that people of Y haplgroup D mixed with newcomers who spoke Austric languages and bore Y haplogroup O upon the plains, or were ousted to the mountain where they could save their languages.

3.4. Location of proto-language of Ainu-Andamanese-Sino-Tibetan stock

Proto-language of Ainu-Andamanese-Sino-Tibetan stock probably located somewhere in South Asia.

3.5. Other potential relatives

Potential relatives can probably be found among West Papuan and Halmahera languages (pic. 3).

References

Akulov A. 2015a. Ainu and Great Andamanese are relatives (proved by Prefixation Ability Index and Verbal Grammar Correlation Index). *CAES*, Special Issue, October 2015; pp.: 1 – 24

Akulov A. 2015b. Verbal Grammar Correlation Index (VGCI) method: a detailed description. *CAES*, Vol.1, № 4; pp.: 19 – 42

Chandrasekar A., Saheb S.Y, Gangopadyaya P., Gangopadyaya S., Mukherjee A. Basu D., Lakshmi G.R., Sahani A.K., Das B., Battacharya S., Kumar S., Xaviour D., Sun D., Rao V.R. 2007. YAP insertion signature in South Asia. *Annals of Human Biology* 34 (5); pp.: 582 – 586

The Language Gulper http://www.languagesgulper.com/eng/Tibeto.html – accessed May 2016

LaPolla R.J, Huang Ch. 2003. *A Grammar of Qiang with annotated texts and glossary*. Mouton de Gruyter, Berlin – New York

STEDT – The Sino-Tibetan Etymological Dictionary and Thesaurus http://stedt.berkeley.edu/about-st – accessed May 2016

Tajima A., Hayami M., Tokunaga K., Juji T., Masafumi M., Sangkot M., Omoto K., Horai S. 2004. Genetic origins of the Ainu inferred from combined DNA analyses of maternal and paternal lineages, *Journal of Human Genetics* 49; pp.: 187 – 193

Tamura S. 2000. *The Ainu language*. Sanseidō, Tokyo

Xue Y., Zerjal T., Bao W., Zhu S., Shu Q., Xu J., Du R., Fu S., Li P., Hurles M.E., Yang H., Tyler-Smith C. 2006. Male Demography in East Asia: A North-South Contrast in Human Population Expansion Times. *Genetics* 172 (4); pp.: 2431 – 2439