

## Whether Sumerian language is related to Munda?

Alexander Akulov

independent scholar; Saint-Petersburg, Russia; e-mail: [aynu@inbox.ru](mailto:aynu@inbox.ru)

### Abstract

Having compared 34 lexical items of Sumerian with 34 lexical items of Mundari/Santali Diakonoff came to conclusion that Sumerian is a relative of Munda languages. The hypothesis seemed to be rather realistic, however, as far as comparison of randomly selected lexical items without due attention to structures can't say anything about languages relationship so the hypothesis was to be tested by a verified method. Verified method is Verbal Grammar Correlation Index (VGCI). According to VGCI methodology index of correlation of two languages belonging to the same stock is 0.4 or higher; index of correlation of two languages that aren't related is 0.3 or lower. VGCI of Sumerian and Mundari is 0.18 that evidently proves their unrelatedness. Due to transitivity of relatedness/unrelatedness Sumerian isn't a relative of Austroasiatic stock. Perspective directions of searching for potential relatives of Sumerian are the following: Basque, Kartvelian, Northeast Caucasian or Sino-Tibetan stocks.

**Key words:** Sumerian; Munda; Austroasiatic stock; comparative linguistics

### 1. Problem introduction

Hypothesis that Sumerian language could be distant relative of Munda<sup>1</sup> languages was spoken out by Diakonoff.

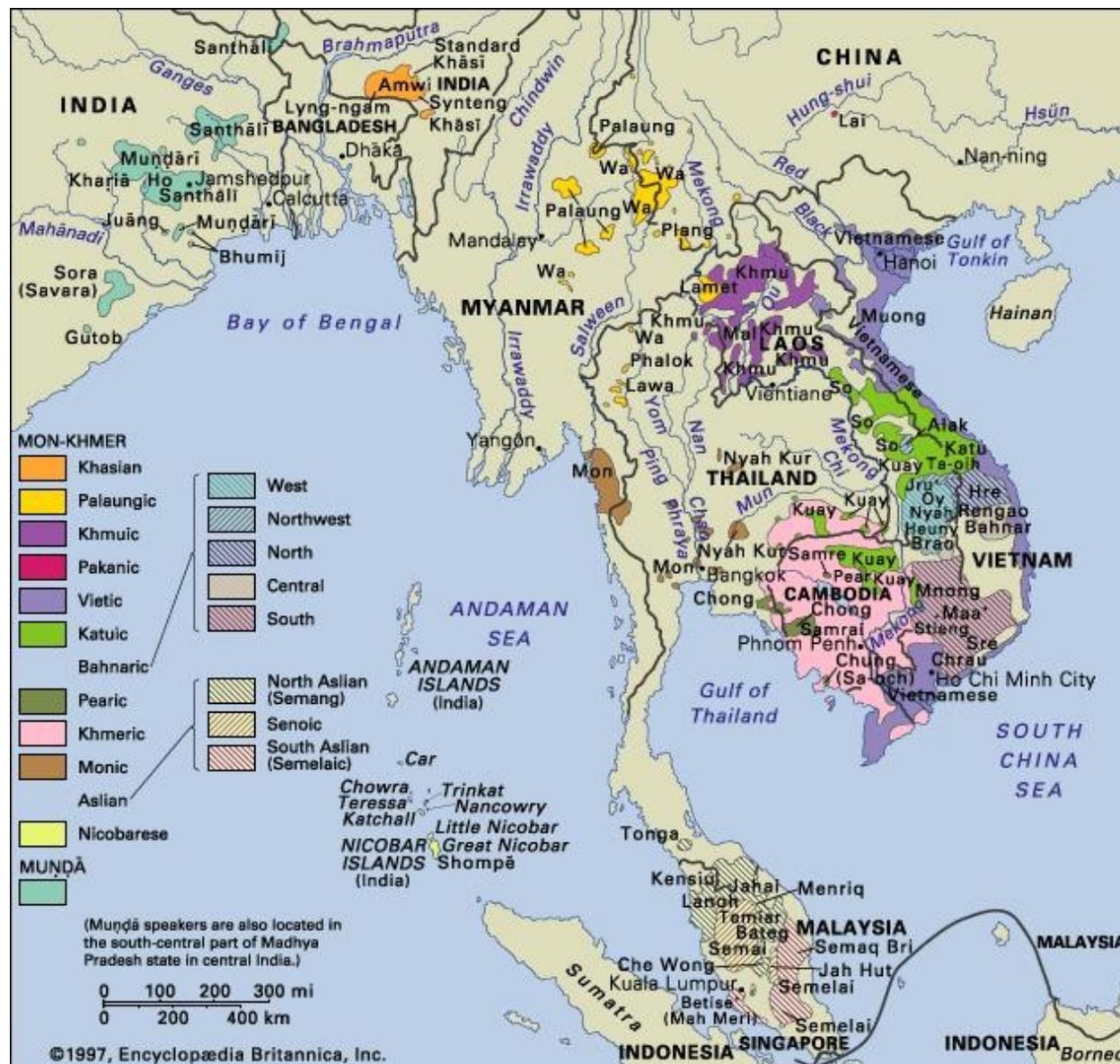


Pic. 1. Location of Munda languages and Sumer (made after Google map screenshot)

<sup>1</sup> A branch of Austroasiatic stock (pic. 1, pic. 2 show its location)

The hypothesis was based on the presupposition that Sumerian came from India that in those time was largely inhabited by people speaking in Munda languages.

Having compared 34 words and material components of certain grammatical morphemes of Sumerian with those of Kherwari<sup>2</sup> languages Diakonoff came to the conclusion that Sumerian was a relative of Munda languages (Diakonoff 1997).



Pic. 2. Map showing location of Austroasiatic stock in general and Munda and Kherwari languages (Mundari and Santali) in particular (source: Austroasiatic languages)

Methodology of detecting of relationship of certain languages based on comparison of lexis is outside of normal scientific methods since it allows completely different conclusions about the same language: conclusions are strongly influenced by character of used lexical materials and authors' presuppositions. This methodology is actually based on 'artist sees so' principle.

<sup>2</sup> A subgroup within Munda language, the group consists of two languages: Mundari and Santali (pic. 2 shows their location).

Lexical approach doesn't work even on the material of normally described languages and it's especially useless in the case of Sumerian since Sumerian phonology is not even a reconstruction but is just a very approximate imagination based on Sumerian-Akkadian dictionaries.

The case of Sumerian is this very issue where any comparison of material implementations should be completely eliminated and where any conclusions about possible relationship should be made on the analysis of pure structures only.

It is notorious fact that Diakonoff criticized Bengtson for inaccurate comparison of Sumerian lexical items with words of randomly selected languages (Diakonoff 1999; Bengtson 1997).

From a perfunctory point of view Diakonoff's approach seems to be more accurate than that of Bengtson. However, actually Diakonoff did the same procedure as Bengtson: they both made their conclusions on the base of comparison of certain randomly selected lexical items and both ignore that language is first of all system of grammatical relations, but not a heap of lexemes; both approaches are based on 'artist sees so' principle.

However, hypothesis of Diakonoff seems to be rather realistic, so in current paper it is tested by VGCI that is a precise verified method and can completely resolve questions of relationship of any languages.

## 2. Method

Verbal Grammar Correlation Index (VGCI) supposes direct comparison of really existing/existed languages.

The method is based on the idea that language is set of grammatical meanings and set of their positional distributions; and degrees of correlations of both sets can be calculated.

The method doesn't pay any attention to material exponents at all: has been proved that analysis of pure structures can completely solve questions of relationship of any languages.

VGCI is superposition (logical conjunction) of two indexes: grammatical meanings sets correlation index and index of correlation of positional distributions of common meanings. Scheme of calculation of VGCI can be represented as the following formula:

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

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.

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 higher then languages are related; if value of VGCI is about 0.3 or lower then languages are not related.

Error of measurement is about 2%.

Probability of coincidence of two languages is much less than  $10^{-19}$ .

(For more details see Akulov 2015.)

### 3. Comparison of Sumerian with Mundari

#### 3.1. List of Mundari forms

The list has been compiled after Anderson 2007; Osada 2008.

1. Action/gent oriented version -sfx
2. Agent: 11 prp / 11 -sfx
3. Benefactive: -sfx
4. Causative: prfx-
5. Dual: 4 prp / 8 -sfx
6. Exclusive: 2 prp / 4 -sfx
7. Future: zero marker (deleting indicative marker)
8. Inclusive: 2 prp / 4 -sfx
9. Imperative: zero marker
10. Indicative: -sfx
11. Iterative: -RR-
12. Negation: prp
13. Non-future: -sfx
14. Optative: -sfx
15. Passive: -sfx
16. Patient: 11 -sfx
17. Patient oriented version: -sfx
18. Perfect aspect 1 ~ Past: -sfx
19. Perfect aspect 2: -sfx
20. Perfect aspect 3: -sfx
21. Plural : 4 prp / 8 -sfx
22. Progressive aspect ~ Present continuous: -sfx
23. Progressive 2 -sfx
24. Progressive 3: -sfx
25. Prohibitive: prp
26. Reciprocity: -sfx
27. Reflexivity: -sfx
28. Singular: 3 prp / 6 -sfx
29. Subject: 11 prp / 11 -sfx

#### 3.2. List of Sumerian forms

The list has been compiled after Kaneva 2006.

1. Ablative dimensional marker: 2 prfx-
2. Agent: 4 prfx- / 7 -sfx / 3 crfx
3. Allative dimensional marker: 6 prfx-
4. Animate: 2 prfx- / 3 -sfx

5. Assertive mood: prfx-
6. Causative: zero marker
7. Centrifugal (non-topical version): prfx-
8. Centripetal version: prfx-
9. Centripetal version 2: prfx-
10. Comitative dimensional marker: 6 prfx-
11. Dative dimensional marker: 7 prfx-
12. Desiderative ~ Optative mood: prfx-
13. Future: -sfx
14. Hypothetical volition mood: prfx-
15. Imperative: R/zero marker
16. Imperfect aspect: -RR- / 3 suppletive forms / 56 prfx- \* -sfx / 56 prfx \* suppletive forms  
+ -sfx / 56 prfx- \* -RR- \* -sfx
17. Inanimate: 2 prfx- / 3 -sfx
18. Inanimate patient version: prfx-
19. Indicative: zero marker
20. Indirect inanimate object centrifugal version: prfx-
21. Indirect inanimate object centripetal version: prfx-
22. Intensity: -RR-
23. Iterative: -RR-
24. Locative dimensional marker: prfx-
25. Locative-allative dimensional marker: 3 prfx- / crfx
26. Negation: prfx-
27. Negation 2 ~ vetitive: prfx-
28. Non-future: zero marker
29. Patient: 8 prfx- / 7 sfx-
30. Perfect aspect: 28 prfx \* sfx / 28 prfx \* suppletive forms \* -sfx / 28 prfx- \* -RR- \* -sfx /  
21 crfx \* sfx / 21 crfx \* suppletive forms \* -sfx / 21 crfx \* -RR- \* -sfx
31. Plural: 3 -sfx / 3 -sfx \* -RR- / 30 -sfx \* suppletive forms / 12 prfx- \* -sfx / 12 prfx- \* -  
RR- \* -sfx / 32 prfx- \* suppletive forms \* -sfx / 9 -sfx \* crfx / 9 -sfx \* -RR- \* crfx /  
18 -sfx \* suppletive forms \* crfx
32. Prohibitive: prfx-
33. Reciprocity: -RR-
34. Request mood: prfx-
35. Singular: 4 -sfx / 16 prfx- \* -sfx / 32 prfx- \* suppletive forms \* -sfx / 16 prfx- \* -sfx /  
32 prfx- \* suppletive forms \* -sfx
36. Subject: 7 -sfx
37. Stative version: prfx-

### 3.3. VGCI of Mundari and Sumerian

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 M for Mundari; then schemes of expressions of this grammatical meaning in Mundari; then sign of correlation “~” or anti-correlation “≠”; then S for Sumerian; schemes of expressions of this grammatical meaning in Sumerian; and then number showing degree of correlation.

1. Agent: M.: 11 prp / 11 -sfx ~ S.: 4 prfx- / 7 -sfx / 3 crfx  $(11/22 + 11/14)/2 \approx 0.6$
2. Agent version: M: -sfx ≠ S.: centripetal version: 2 prfx- 0
3. Causative: M.: prfx ≠ S.: no marker 0

4. Future: M.: zero marker  $\neq$  S.: -sfx 0
5. Imperative: M.: zero marker  $\sim$  S.: zero marker 1
6. Indicative: M.: -sfx  $\neq$  S.: zero marker 0
7. Iterative: M.: -RR-  $\sim$  S.: -RR- 1
8. Negation: M.: prp  $\sim$  S.: prfx- 1
9. Non-future: M.: -sfx  $\neq$  S.: zero marker
10. Optative: M.: -sfx  $\neq$  S.: prfx- 0
11. Patient: M.: 11 -sfx  $\sim$  S.: 8 prfx- / 7 sfx-  $(7/11 + 7/15)/2 \approx 0.5$
12. Patient oriented version  $\sim$  inanimate patient version: M.: -sfx  $\neq$  S.: prfx- 0
13. Perfect aspect: M.: 3 -sfx  $\neq$  S.: 28 prfx \* sfx / 28 prfx \* suppletive forms \* -sfx / 28 prfx- \* -RR- \* -sfx / 21 crfx \* -sfx / 21 crfx \* suppletive forms \* -sfx / 21 crfx \* -RR- \* -sfx 0
14. Plural: M.: 4 prp / 8 -sfx  $\sim$  S.: 3 -sfx / 3 -sfx \* -RR- / 30 -sfx \* suppletive forms / 12 prfx- \* -sfx / 12 prfx- \* -RR- \* -sfx / 32 prfx- \* suppletive forms \* -sfx / 9 -sfx \* crfx / 9 -sfx \* -RR- \* crfx / 18 -sfx \* suppletive forms \* crfx  $(3/12 + 3/128)/2 \approx 0.13$
15. Progressive aspect: M.: 3 sfx  $\neq$  S.: -RR- / 3 suppletive forms / 56 prfx- \* -sfx / 56 prfx- \* suppletive forms \* -sfx / 56 prfx- \* -RR- \* -sfx 0
16. Prohibitive: M.: prp  $\sim$  S.: prfx- 1
17. Reciprocity: M.: -sfx  $\neq$  S.: -RR- 0
18. Singular: M.: 3 prp / 6 -sfx  $\sim$  S.: 4 -sfx / 16 prfx- \* -sfx / 32 prfx- \* suppletive forms \* -sfx / 16 prfx- \* -sfx / 32 prfx- \* suppletive forms \* -sfx  $(4/9 + 4/100)/2 \approx 0.5$
19. Subject: 11 prp / 11 -sfx  $\sim$  S.: 7 -sfx  $(7/22 + 1)/2 \approx 0.66$

Index of correlation of grammatical meanings sets is the following:

$$(19/29 + 19/37)/2 \approx 0.55;$$

index of correlation of positional distributions of common meanings is the following:

$$(0.6 + 4 + 2*0.5 + 0.13 + 0.66)/19 \approx 0.33;$$

And thus, VGCI of Mundari and Sumerian is  $0.55*0.33 \approx 0.18$ .

If value of VGCI is 0.3 or lower it means that languages are not related.

#### 4. Some preliminary conclusions

The fact that Sumerian and Munda aren't relatives means that Sumerian is not relative of whole Austroasiatic stock (due to transitivity of relatedness/unrelatedness).

Perspective directions of searching for potential relatives of Sumerian are the following: Basque stock, Kartvelian stock, Northeast Caucasian stock or Sino-Tibetan stock.

#### References

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

Anderson G. D. S. 2007. *The Munda Verb. Typological perspectives*. Mouton de Gruyter, Berlin – New York

Austroasiatic languages, Encyclopædia Britannica

<https://global.britannica.com/topic/Austroasiatic-languages> – accessed September 2016

Bengtson Jh. D. 1997. The Riddle of Sumerian: A Dene-Caucasic Language? *Mother Tongue*. Issue III, December 1997; pp.: 63 – 74

Diakonoff I.M. 1997. External Connections of the Sumerian Language. *Mother Tongue*. Issue III, December 1997; pp.: 54 – 62

Diakonoff I. M. 1999. More on Possible Linguistic Connection of the Sumerians. *Mother Tongue*. Issue V, December 1999; pp.: 141 – 144

Kaneva I.T. 2006. *Shumerskii yazyk (The Sumerian language)*. Peterburgskoe Vostokovedenie, St. Petersburg

Osada T. 2008. Mundari. In Anderson G. D. S. (eds.) *The Munda languages*. Routledge, London; pp.: 99 – 164 <https://www.routledgehandbooks.com/doi/10.4324/9781315822433.chThree> – accessed September 2016