

The Ainu-Minoan (Sino-Caucasian) substratum in Indo-Iranian languages

Tresi Nonno

independent scholar; Chiba, Japan; e-mail: tresi_nonno@hotmail.com

Abstract

In Sanskrit and in Iranian languages there are about 55 words that have no reliable Indo-European etymology, but are supposed to have originated from the language of the Bactria-Margiana Archaeological Complex (BMAC). The BMAC language seems to be quite close to Burushaski and Yeniseian. Burushaski and Yeniseian languages are quite close relatives, and both belong to the Ainu-Minoan stock / macrofamily. At least some of these 55 substratum words can be correlated with Burushaski and Yeniseian words. The most notorious examples are: the Sanskrit word *atharvan* – a legendary priest who, according to Hinduism, performed the first yajna – originally was the following word-form: *a-th-a-r-u-an* – “the one who performs rite”; the Sanskrit theonym *Indra* correlates with the Proto-Yeniseian form **tingVr* “high”; the Sanskrit word *nagnáhu* “sourdough” and Persian *nān* “bread” correlate with Common Burushaski **nayé* “dung” (sourdough sometimes does not smell particularly pleasant and is similar in consistency to the dung).

Keywords: substratum in Sanskrit; Burushaski language; Yeniseian languages; Bactria-Margiana Archaeological Complex; Ainu-Minoan stock; Sino-Caucasian stock

1. Introduction to the problem

In Sanskrit and in Iranian languages there are about 55 words which have no reliable Indo-European etymology (Lubotsky 2001, Wikipedia 2024c).

These words relate to cereal-growing and bread making (bread, ploughshare, seed, sheaf, yeast), waterworks (canal, well), architecture (brick, house, pillar, wooden peg), tools or weapons (axe, club), textiles and garments (cloak, cloth, coarse garment, hem, needle) and plants (hemp, mustard, soma plant) (Witzel 2005).

Lubotsky pointed out that these Sanskrit words have morphological and phonological counterparts in Iranian languages, and this fact indicates that there was a common substratum. He supposed that these words could have been borrowed from the language of the Bactria-Margiana Archaeological Complex (BMAC) which was supposedly connected with the Harappan language – the language of the Indus Valley civilization (Lubotsky 2001).

It is noteworthy that the BMAC and the Indus Valley civilization coexisted for a certain period and that their territories were quite close (Fig. 1, Fig. 2).

The external links of the BMAC language are quite unclear, but there is a supposition that this language was quite close to Burushaski and Yeniseian.

Burushaski and Yeniseian are quite close relatives (Clarke 1870, van Driem 2001, Bengtson 2010), and both Yeniseian and Burushaski belong to the Ainu-Minoan (Sino-Caucasian) macrofamily (Bengtson 2008, Akulov 2019).

At least some words of these 55 substratum words can be correlated with Yeniseian and Burushaski words.



Fig. 1. The region where the Bactria-Margiana Archaeological Complex (BMAC) existed in about 2250 – 1700 BCE (image source – Wikipedia 2024a)

2. Substratum words of Indo-Iranian languages that can be correlated with Burushaski or Yeniseian words

2.1. Words with more reliable correlations

2.1.1 Sanskrit *atharvan* and Avestian *āθrauan* “priest”

It is supposed that Sanskrit *átharvan* “priest”, Avestian *āθrauan*/*aθaurun*- “priest”, Pehlevi *āsrōn*, Tocharian A *atār*, Tocharian B *etre* “hero” from BMAC **atʰr* (Wikipedia 2024c).

I seriously doubt that it is correct to derive these Indo-Iranian forms from BMAC **atʰr*, I suppose that in the BMAC language there was a form of the following view: **atharuan*.

According to Hinduism Atharvan is a legendary Vedic sage (rishi) of Hinduism who, along with Angiras, is supposed to have authored (“heard”) the Atharvaveda. He is also said to have first instituted the fire-sacrifice or yajna, and thus can be described as an actor.

The BMAC form **atharuan* looks like a verbal form that can be decoded through Burushaski and Yeniseian grammars.

For the sake of convenience, I rewrite the BMAC word-form of *atharuan* in the following view:
a-th-a-r-u-an.



Fig. 2. The territory occupied by the Indus Valley civilization (image source – Wikipedia 2024b)

The initial *a-* correlates with a marker of patient of Burushaski. In Burushaski the *a-* prefix expresses 1sg.pt./sb., also in Burushaski there are three prefixes *o-* expressing 3pl.pt. of male, female and inanimate nouns classes.

Also, the initial *a* can be correlated with a marker of patient/object existing in the Kott language: *ā* expressing 3sgpt/sb¹

Taking into account the significant time distance between the modern Burushaski and modern Yeniseian languages on the one hand and the language of BMAC on the other hand it is completely natural and logical to suppose a certain shift of meanings and to suppose that in the language of BMAC there could have existed a marker of 3sg.pt. that had the view of *a-*.

The following *-th-* correlates well with the Common Burushaski root **-t-* “to do”, “to make”, “to set up”². And also correlates with the Kott causative marker *-t^h-* that is infixed into a verb

¹ Werner 1998: 90.

² Starostin 2005b.

stem close to its end³, in all languages which have formal causative constructions causative is expressed by verb “to do” / “to make” or by its derivatives.

The component *-a-* can be correlated with the Kott marker of present tense *-a-* that is infix in a verb stem⁴.

The component *-r-* correlates with the Kott marker of imperfective aspect *-r-/-l-* that is infix into a verb stem⁵. The marker of aspect is placed to the right of the tense marker⁶.

The next component *-u-* looks like an epenthetic vowel.

The ending *-an* correlates with markers of 3pl.ag. of male and female noun classes existing in the modern Burushaski.

Also, the final *-an* can be correlated with a marker of agent existing in the Kott language: *-an* expressing 1sgag⁷.

Taking into the account the significant time distance between the modern Burushaski and the modern Yeniseian on the one hand and the language of BMAC on the other hand it is completely natural and logical to suppose a certain shift of meanings and the possibility that in the language of BMAC there could have existed a marker that had the view of *-an* and expressed 3sg.ag. of male and female noun classes.

Thus, we have the following:

a-th-a-r-u-an.

3sg.pt.thing-make/perform-PRES-IMPRF-epent.vow.-3sg.ag.m.

In other words, the word *atharvan* / *atharuan* originally meant simply “the one who performs a rite”.

2.1.2. Sanskrit theonym *Indra*

The Sanskrit theonym *Indra* has no reliable Indo-European etymology, and it is supposed that the theonym derived from BMAC theonym **Indra* (Wikipedia 2024c).

I suppose that the word *Indra* can be correlated with Proto-Yeniseian form **tingVr* meaning “high”⁸.

It is noteworthy that the Proto-Yeniseian **tingVr* / **tingir* also seems to have been the source for Turkic theonym *Tengri*, i.e.: the Turkic word ultimately originates as a loanword from Proto-Yeniseian **tingir* “high” (Georg 2001).

Also, it is noteworthy that the Proto-Yeniseian form **tingVr* seems to be a cognate with Sumerian word *diĝir* “god” / “goddess”, this word seems to have been pronounced as [tiŋir] that is very close to the Proto-Yeniseian form **tingVr*. The Sumerian language belongs to the same branch within the Aino-Minoan stock, as well as Yeniseian languages (Akulov 2022).

³ Werner 1998: 115

⁴ Werner 1998: 103.

⁵ Werner 1998: 108.

⁶ Werner 1998: 84 – 85.

⁷ Werner 1998: 90.

⁸ Starostin 2005f.

In the case of *Indra* and **tṛṅgVr* took place the following processes:

- 1) falling away of the initial /t/,
- 2) softening of the /g/, i.e.: /g/ → /dʒ/.

2.1.3. Sanskrit *nagnáhu* “sourdough” / “yeast” and Persian *nān* “bread”

The Sanskrit word *nagnáhu* “sourdough” / “yeast” and related Persian *nān* “bread” are supposed to have been derived from BMAC **nagna* meaning “sourdough” “yeast” (Wikipedia 2024c). This BMAC form **nagna* correlates with the Common Burushaski form **nayé* “dung”⁹. At first sight, it is quite strange that the words bread and manure turn out to have derived from the same root. However, we should keep in mind the fact that sourdough sometimes does not smell particularly pleasant and is similar in consistency to the dung.

2.2. Words with less reliable correlations

2.2.1. Sanskrit *śaṇa*, Middle Persian *šan* “hemp”

Sanskrit *śaṇa* and its cognate Middle Persian *šan* meaning “hemp” are supposed to have been derived from BMAC **k/śan-* “hemp” (Wikipedia 2024c). However, I suppose that it is more correct to collate the Sanskrit and Middle Persian forms with Common Burushaski **sar* “thread”¹⁰. Hemp was and still is one of the main materials for producing thread/textile.

2.2.2. Sanskrit *śépa* “tail”

Sanskrit *śépa*, Prākṛit *cheppā*; Late Avestian *xšuaēpā* “tail” are supposed to have been derived from BMAC **šwaipa* “tail” (Wikipedia 2024c). This BMAC form can be collated with Common Burushaski **sum* “tail”; “sprout”, “shoot”; “snout (of a vessel)”¹¹.

2.2.3. Sanskrit *sūcī* and Late Avestian *sūkā* “needle”

Sanskrit *sūcī*, Late Avestian *sūkā* meaning “needle” are supposed to have originated from BMAC **sūcī* “needle” (Wikipedia 2024c). I suppose that the Common Burushaski form **sel* “needle”¹² can be collated with the BMAC form **sūcī*.

2.2.4. Sanskrit *yavīyā* and Old Persian *yauwiyā* “canal”

Sanskrit *yavīyā* and Old Persian *yauwiyā* “canal”, “irrigation channel” are supposed to have originated from BMAC **jawījā* “canal” (Wikipedia 2024c).

I suppose that BMAC **jawījā* can be correlated with Common Burushaski **dalá* “large irrigation canal”¹³.

⁹ Starostin 2005c.

¹⁰ Wkitionary 2024.

¹¹ Starostin 2005e.

¹² Starostin 2005d.

¹³ Starostin 2005a.

References

Akulov A. 2022. Sumerian and the Ainu-Minoan stock. *Cultural Anthropology and Ethnosemiotics*, Vol. 8, № 1; p.: 6 – 11

Akulov A. 2019. Yeniseian languages and Ainu-Minoan stock. *Cultural Anthropology and Ethnosemiotics*, Vol. 5, № 3; pp: 12 – 17

Bengtson Jh. D. 2010. Burushaski and Yeniseian and the Karasuk Culture. *14th Harvard Round Table on the Ethnogenesis of South and Central Asia*. Harvard University <https://www.academia.edu/10091472> – accessed June 2024

Bengtson Jh. D. 2008. Materials for a Comparative Grammar of the Dene–Caucasian (Sino-Caucasian) Languages. *Aspects of Comparative Linguistics*, vol. 3, RSUH Publishers, Moscow; pp.: 45 – 118

Clarke H. 1870. Response to Leitner’s verbal presentation (November 23rd, 1869). *Journal of the Ethnological Society of London*, New Series II; pp.: 32 – 34

Georg S. 2001. Türkisch/Mongolisch tengri “Himmel/Gott” und seine Herkunft. *Studia Etymologica Cracoviensia* 6; pp.: 83 – 100

Lubotsky A. 2001. The Indo-Iranian Substratum. In Carpelan, C.; Parpola, A.; Koskikallio, P. (eds.). *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations*. Suomalais-Ugrilainen Seura, Helsinki; pp.: 301 – 317

Starostin S. A. 2005a. Burushaski etymology database. Meaning “canal” https://starlingdb.org/cgi-bin/response.cgi?single=1&basename=%2fdata%2fsinocauc%2fburuet&text_number=371&root=config – accessed June 2024

Starostin S. A. 2005b. Burushaski etymology database. Meaning “to do” https://starlingdb.org/cgi-bin/response.cgi?root=config&morpho=0&basename=%5Cdata%5Csinocauc%5Cburuet&first=1&off=&text_proto=&method_proto=substring&ic_proto=on&text_meaning=to+do&method_meaning=substring&ic_meaning=on&text_yas=&method_yas=substring&ic_yas=on&text_hun=&method_hun=substring&ic_hun=on&text_ngr=&method_ngr=substring&ic_ngr=on&text_notes=&method_notes=substring&ic_notes=on&text_any=&method_any=substring&ic_any=on&sort=proto – accessed June 2024

Starostin S. A. 2005c. Burushaski etymology database. Meaning “dung” https://starlingdb.org/cgi-bin/response.cgi?single=1&basename=%2fdata%2fsinocauc%2fburuet&text_number=298&root=config – accessed June 2024

Starostin S. A. 2005d. Burushaski etymology database. Meaning “needle”

https://starlingdb.org/cgi-bin/response.cgi?root=config&morpho=0&basename=%5Cdata%5Csinocauc%5Cburuet&first=1&off=&text_proto=&method_proto=substring&ic_proto=on&text_meaning=needle&method_meaning=substring&ic_meaning=on&text_yas=&method_yas=substring&ic_yas=on&text_hun=&method_hun=substring&ic_hun=on&text_ngr=&method_ngr=substring&ic_ngr=on&text_notes=&method_notes=substring&ic_notes=on&text_any=&method_any=substring&ic_any=on&sort=proto – accessed June 2024

Starostin S. A. 2005e. Burushaski etymology database. Meaning “tail”

https://starlingdb.org/cgi-bin/response.cgi?root=config&morpho=0&basename=%5Cdata%5Csinocauc%5Cburuet&first=1&off=&text_proto=&method_proto=substring&ic_proto=on&text_meaning=tail&method_meaning=substring&ic_meaning=on&text_yas=&method_yas=substring&ic_yas=on&text_hun=&method_hun=substring&ic_hun=on&text_ngr=&method_ngr=substring&ic_ngr=on&text_notes=&method_notes=substring&ic_notes=on&text_any=&method_any=substring&ic_any=on&sort=proto – accessed June 2024

Starostin S. A. 2005f. Yeniseian etymology database. Meaning “high”

https://starlingdb.org/cgi-bin/response.cgi?single=1&basename=%2fdata%2fyenisey%2fyenet&text_number=710&root=config – accessed June 2024

van Driem G. 2001. *Languages of the Himalayas (2 vols) An Ethnolinguistic Handbook of the Greater Himalayan Region containing an Introduction to the Symbiotic Theory of Language*. Leiden, Brill

Werner H. 1998. *Abriß der kottischen Grammatik*. Harrassowitz, Wiesbaden

Wikipedia 2024a. Bactria–Margiana Archaeological Complex –
https://en.wikipedia.org/wiki/Bactria%E2%80%93Margiana_Archaeological_Complex
 accessed June 2024

Wikipedia 2024b. Indus Valley Civilisation
https://en.wikipedia.org/wiki/Indus_Valley_Civilisation – accessed June 2024

Wikipedia 2024c. Substratum in Vedic Sanskrit
https://en.wikipedia.org/wiki/Substratum_in_Vedic_Sanskrit – accessed June 2024

Wiktionary 2024. Burushaski word list
https://en.wiktionary.org/wiki/Appendix:Burushaski_comparative_vocabulary_list – accessed June 2024

Witzel M. 2005. Central Asian Roots and Acculturation in South Asia. Linguistic and Archaeological Evidence from Western Central Asia, the Hindukush and Northwestern South Asia for Early Indo-Aryan Language and Religion". In Osada, T. (ed.). *Linguistics, Archaeology and the Human Past. Indus Project*. Kyoto, Research Institute for Humanity and Nature; pp.: 87 – 211