© G.S. Starostin

# THE DIFFICULTIES OF RECONSTRUCTING THE CULTURAL LEXICON FOR A MACROFAMILY-LEVEL PROTOLANGUAGE (BASED ON THE AFRASIAN EXAMPLE)

The paper offers a critical analysis of several Afrasian etymologies with presumably "military" semantics, put forward by Alexander Militarev. The conclusion is that these etymologies typically suffer from multiple problems, such as lack of proper attention to the historical typology of semantic shifts and insufficient consideration for the distribution of potential reflexes in daughter languages. Because of this, the reconstructability of a large Proto-Afrasian lexical layer of specifically "military" terms remains questionable – a t least not until such a reconstruction has been diligently conducted on each of the chronological levels preceding Proto-Afrasian (particularly on the various intermediate levels of the Cushitic family, since only a secure reconstruction of any select etymon on the ProtoCushitic level can in turn properly guarantee its Proto-Afrasian status).

The paper of Alexander Yu. Militarev, one of the largest Russian specialists in historical semitology and Afrasistics, continues his long-term research in the field of reconstruction of not only the phonetic appearance and vocabulary of the Proto-Afroasiatic language, but also the cultural world of the ancient Afrasians. I would like to take this opportunity, by no means belittling the relevance and significance of the author's etymological research, to nevertheless express a few long-pending considerations concerning some fundamental methodological problems that have not been completely overcome either in Afroasiatic linguistics, or in general in studies that somehow deal with "distant" (as opposed to "near") linguistic kinship.

Professional historical linguists are well aware that there is, although not very clearly defined, but still a real demarcation line between the "basic" comparative material, on the basis of which the fact of linguistic kinship *is proved* (or at least justified), and a kind of additional layer, which by itself, for various reasons, cannot prove the kinship, but can be used for further linguistic reconstruction, provided that the kinship has already been proven. The "basic" material primarily includes elements of grammar and so called basic vocabulary (more or less universal lexical concepts that are rarely borrowed from language to language and have a high level of historical stability); the additional layer is usually defined as «cultural lexicon», which includes more complex, less universal concepts that are more prone to shifting and borrowing.

**George Starostin** | https://orcid.org/0000-0002-4024-6625 | gstarst1@gmail.com | National Research University "Higher School of Economics" (20 Myasnitskaya Str., Moscow, 101000, Russia)

In modern comparative studies, the hypothesis of Afroasiatic relationship, from the point of view of the "basic" material, is in a much more favorable position than, for example, the Nostratic hypothesis of Vladislav M. Illich-Svitych or the Sino-Caucasian hypothesis of Sergei A. Starostin: comparative material confirming it from the field of grammatical paradigmatics and basic vocabulary, although not numerous, is still usually considered sufficient for the world linguistic community to accept as a given the descent of the Semitic, Berber, Chadic, Egyptian and Cushitic languages from a common linguistic ancestor (some doubts have recently been expressed only regarding the Omotic languages).

However, reconstruction of the Proto-Afroasiatic cultural vocabulary and, in general, Afroasiatic etymology as such, are in a much more difficult situation. None of the two Afroasiatic etymological dictionaries published to this date (by Christopher Ehret and by Vladimir Orel and Olga Stolbova) enjoys significant authority, and although in general the dictionary of Orel and Stolbova is significantly stronger than the works of Ch. Ehret from the point of view of the quality of phonetic and semantic comparisons, most of the etymologies presented in it can be considered "raw material" for building a full-fledged dictionary rather than the final product. Work on the preparation of this kind of publication has long been carried out as part of the etymological base by Alexander Militarev and Olga Stolbova (from which the material for the concerned article is mainly borrowed), however, in my opinion, its successful completion is still hindered by a number of unresolved problems, which are quite clearly manifested, including in the data given in the publication of Alexander Militarev.

Above all, it is necessary to emphasize once again the historical specifics of the *cultural* vocabulary (as opposed to the *basic* one). First, the cultural vocabulary as a whole is less *stable*, i.e. the words that do not belong to the basic layer, as a rule, have a shorter "life cycle" and are replaced by lexical innovations much faster. Accordingly, for example, the a priori probability that the same word with the meaning "war" or "club" will remain unchanged (except for the phonetic appearance) for five or even more than ten thousand years, even in one language (not to mention more), should be much lower than for the word "hand" or the pronouns "I" and "you". At best, various semantic shifts should be expected for lexical elements belonging to the cultural layer, at worst − complete disappearance. The exceptions, of course, are possible, but in general, only an incurable optimist can count on a successful reconstruction of a full−fledged corpus of cultural vocabulary at the level of, say, 10−12 thousand years BC.

Secondly, in the presence of even minimal language contacts, the cultural vocabulary is easily subject to *borrowings*, which creates numerous problems in cases where originally related languages are in contact with each other; this requires the establishment of fairly clear criteria for distinguishing genetically related and borrowed vocabulary, which is actually not always possible, especially in situations where there are ambiguities with historical phonetics.

Surprisingly, in situations of distant relationship between several language families, each of which is represented by dozens (Semitic, Berber, Cushitic) or even hundreds (Chadic) of languages, some researchers may find these problems not so significant or even negligible, although they often turn out to be significant for relatively small families. This is primarily due to the fact that in the presence of a huge linguistic material from which one can more or less arbitrarily select formally "suitable" etymological parallels, the possibility of taking random coincidences in form and meaning (as a rule, partial, but sometimes complete) for such parallels increases significantly.

Thus, in order to consider this or that etymology at the level of deep relationship convincing (we will define as "deep" the kind of relationship whose glottochronological age exceeds the approximate time depth of such a reference family for comparative studies as Indo-European, i.e. about 5-6 thousand years), it is necessary that not only its individual members be formally compatible with each other in terms of phonetics and semantics, but

also that their *reconstructibility* at *intermediate* levels of comparison could be monitored. Thus, in Indo-European linguistics, it is traditionally not very common to trust comparisons based on data from modern languages that do not have parallels in ancient languages, i.e. an etymology consisting, for example, of modern Persian, Irish and Albanian words will be a priori doubtful if their ancestors are not found in ancient Iranian and old Celtic corpora. For most Afroasiatic branches (except, of course, Semitic and Egyptian), where there are no ancient corpus data, this kind of verification function is played by the representation of the same etymon in different languages of different sub-branches of the family, which allows it to be reliably projected to the Proto-Cushitic, Proto-Chadic and similar levels and thereby significantly increase confidence in the etymologization of the upper (Afroasiatic) level.

It seems to me that, despite, as one would think, the rather complex and well-thought-out system of assessing the quality of etymology developed by Alexander Militarev and briefly described in the concerned article, it still does not solve the problem of reconstructibility at intermediate levels well enough. To clearly confirm this concern, we will analyze, in order to avoid subjectivity, the first three etymologies put forward by the author:

1. *ma/iṭw- ~ *may/wṭ- "club, cudgel": this root, apparently, is well represented in Semitic languages and is confidently brought to the proto-Semitic level. The Egyptian parallel, although not perfect in terms of semantics (the Egyptian *mdw* staff was probably more a sacral than military tool), is well traced from ancient times in the Egyptian corpus and can be considered convincing.

The situation is much worse with the Chadic part of the etymology presented in one (Hausa) of several hundred languages; in order for this parallel to be acceptable, it is necessary to at least show its reconstructibility to the Proto-Chadic or at least to the Proto-West Chadic level. However, it is also marked as unreliable in the article of Alexander Militarev.

The most significant for etymologization, both in this and in all other cases, is the comparison of Northern Afrasian (SEBCH) and Cushitic (or Cushitic-Omotic) materials, since only a successful comparison at this level, in fact, allows us to give the etymology a pan-Afrasian status. And here some problems are immediately detected. The Cushitic part of the etymology consists of two languages - Oromo (*Lowland East Cushitic*), where we have *muṭuṭḗ* "club", and Sidamo (*Highland East Cushitic*), where the form *amāṭṭ-o* (either "spear" or "arrow", depending on the source) is used in the comparison. First, it is clear that we are talking at best about Eastern Cushitic, and not about the Common Cushitic level. Secondly, even within Eastern Cushitic comparison is not perfect: "club" and "arrow" are objects that are semantically far from identical, while the root forms are so different from each other (what is initial *a-* in Sidamo? is it really possible that such vocalization models in Oromo and in Sidamo can regularly correspond to each other?), that a specific phonetic reconstruction is not even proposed for them on the (relatively shallow!) Proto-East Cushitic level.

Similarly, the Omotic parallel is actually reconstructed at most at the level of one of the sub-branches (Gimojan) of one of the two main branches of the Northern subfamily, and also represented in the area in contact with Eastern Cushitic languages, which raises the question of possible areal diffusion (it is honestly noted in the etymology that the Omotic word can be borrowed from Sidamo).

Conclusion: although at this level of analysis it is impossible to clearly *reject* a hypothesis of kinship between Oromo *muṭuṭḗ* "club" and the Semitic-Egyptian etymology (for this it is necessary to present concrete arguments in favor of a different origin of the Cushitic word), it also cannot be considered convincing at least until this word in the Oromo receives a reliable Common Cushitic etymology.

2. *ḳVs- "bow": at the Proto-Semitic level the root *ḳaš-t- "bow" is reliably reconstructible (it is not very clear, however, how it is related to the Arabic *ḳaws-*), but further comparisons are challenging. It is said about the Egyptian word *ḳꜣs* that its original

meaning was most likely "bowstring"; even if this is so, the semantic connection between "bow" and "bowstring" is not as unambiguous as it may seem (synchronous polysemies of this kind are unknown to me), but, most importantly, already in the oldest texts this word has common meanings "to bind, tie" (verb) and "tie, fetters" (nominal), a special extension of which is sometimes "to attach the bowstring on the bow" etc. The argument that the original meaning was "bowstring" here is clearly based on an external comparison with Semitic material and, thus, turns out to be circular. It is quite possible, since the phonetic correspondences are regular and the semantics are generally compatible, that the Egyptian word does indeed go back to a common ancestor with Semitic; but it is much more likely in this case that the original meaning was still the broad "to attach, tie" rather than the narrow "bowstring", let alone "bow".

The situation is even worse with the Chadic parallel *k̲Vs-* "arrow, bow". The dictionary of Olga Stolbova referred to by the author refers in turn to the dictionary of Hermann Jungraithmayr and Dymitr Ibriszimow, where a number of *kɛsɛ*-type forms with appropriate meanings, found in a number of Eastern and Central Chadic languages, is defined as the result of borrowing from the Central Sudanic language Bagirmi, which is well known to have been a donor language for many languages of the Chadic family. Olga V. Stolbova generally rejects this idea, pointing out that the word in Bagirmi means "bow" and therefore cannot be the source of those forms which in Chadic languages mean "arrow" (*Stolbova* 2016: 240). This is a weak argument because the word *kɛsɛ* in Bagirmi means exactly "arrow" (*Stevenson* 1969: 13) as well as in the vast majority of Chadic languages (the meaning "bow" is most often found in compound formations, which confirms its secondary nature). If it is indeed a borrowing from Bagirmi - the arguments for such a scenario are quite strong — the comparison should be rejected.

Let us move on to the South Afrasian (Cushitic) part. Everything seems to be going well here at first sight: reflexes are found in three branches of the Cushitic family, which formally allows to project the stem onto the Proto-Cushitic level. However, it is again no coincidence that there is no Proto-Cushitic reconstruction in the dictionary entry, since it proves difficult to reduce all the reflexes to one another:

a) in the case of Bilin *k̲ist-*, there is every reason to suspect a borrowing from an Ethiosemitic source (cf. above: Amharic *k̲äst* etc.); the same opinion, by the way, is shared by Olga Stolbova in her Chadic Dictionary (*Stolbova* 2016: 240);

b) in Somali and Boni it is not very clear how the forms relate to each other (where does the word-medial -*n*- come from in Somali?), nor whether they can be safely reconstructed for the Proto-Eastern Cushitic level (thus, this root is absent in Paul D. Black's seminal study [*Black* 1975] on the reconstruction of the *Lowland East Cushitic* vocabulary);

c) it is not at all clear how Gollango *kaaš-ankó* "shield" is connected with these words (perhaps a descriptive interpretation of this word as something like "protection against arrows" (?) is intended, but in that case one would need a careful morphological analysis, preferably supported by data from other East Cushitic languages);

d) the South Cushitic parallel consists of an isolated form from the extinct language Kwadza *k̲asamato* "bow", which is arbitrarily divided into components *k̲asa-* and -*mato*. Taking into account the fact that the component -*to* in Kwadza is a frequent noun suffix, a segmentation into *k̲asama-to* would be more acceptable, in which case this "bow" is hardly separable from the Proto-Western Rift paradigm *\*kasaːma*, pl. *\*kasamu* "arrow" (*Kießling, Mous* 2003: 173). The only problem is that the initial consonant is voiceless (not ejective); perhaps it means that *k̲-* was written in Kwadza incorrectly (this concerns a single extinct language as opposed to an entire group of living Western Rift languages — Iraqw, Alagwa, Burunge), but in this case the common South Cushite reconstruction should be *\*kasama-*, and in this case this word cannot be cognate with other Afro-Asiatic forms with initial *k̲-*.

To sum up, the etymology, which the author himself rates with 'five stars', is in fact problematic from all sides. Only its Semitic part is beyond doubt, while all the others raise numerous questions in terms of semantics, phonetics, cognate distribution and areal relations.

3. *dVg- "bow and arrow". The very meaning in the title of this etymological article raises doubts (how frequent and natural is the existence of a separate lexical root with a "common" meaning "bow and arrow"? − in all the provided materials such polysemy, apparently, is stated only for the Sidamo language and requires verification at the text level).

At the Proto-Semitic level nothing is really reconstructible here; there is a good Ethiosemitic root *dVgVn- "bow", for which at best an isolated Arabic parallel is found, problematic both phonetically and semantically. A purely Ethiosemitic root without parallels in Middle Eastern Semitic languages is of course a priori suspicious as a possible African substrate element.

At the Proto-Berber level, a verbal root like *dVg- with the semantic variation "to stab, pierce, fight" may be recovered, although it is a de facto isogloss between two closely related Tuareg languages and one North Berber language (Wargla), hardly deserving a "four-star" rating.

For the Chadic languages, there are a number of parallels whose connections to one another remain unclear. One of the cited roots is *dVŋg-, "bow", allegedly the result of metathesis from *dVgVn-, as stated in Olga Stolbova's Dictionary (*Stolbova* 2016: 97). However, firstly, the only argument in favor of such metathesis is an external comparison with Ethiosemitic data (i.e. circular logic again). Secondly, the dictionary itself shows that this root most often has the verbal meaning "to bow, to bend", whereas the nominal meaning "bow" is often interspersed with the semantics of "angle" or "boomerang", which confirms the original verbal meaning "to bend".

The forms in the first group (Ngizim, Logone; it is unclear whether Tobanga *doge* "to throw" also belongs here) deserve a little more attention, but whether it is possible to really reconstruct Common Chadic "arrow" from such scarce data is unclear; Jungraithmayr and Ibriszimow, for example, suspect here a substrate element from the Niger-Congo languages (*Jungraithmayr, Ibriszimow* 1994: 1).

The South Afrasian parallels turn out to be quite weak - with one isolated form each in Eastern Cushitic (Sidamo) and in South Omotic (Aari) respectively, and it is not yet clear how they are related to each other phonetically (here the author's "one star" assessment is quite fair).

In the end we have an isogloss between the Ethiosemitic nominal root "bow", the (Narrow) Berber verb "to pierce" and a suspicious Chadic root "arrow". Can this be considered a convincing argument in favour of a Common Afrasian (or even Common North Afrasian) root "bow and arrow"? Even if all three parts of this etymology could be confirmed by additional data, showing the reconstructibility of these stems at the Proto-Semitic, Proto-Berber and Proto-Chadic levels respectively, the presence among the mentioned meanings of the verb "to pierce" shows that the direction of development could well be from verb to name and, therefore, it would not be appropriate to claim that the Proto-South Afrasians used *dVg- "bow" or "arrow" (let alone both simultaneously).

The list of claims about the proposed etymologies could, of course, be expanded, but in this case the commentary on the article would far exceed the size of the article itself. However, it seems that some conclusions can already be drawn from the conducted analysis, namely:

The above criticisms by no means detract from the significance of the comparative studies carried out by the author of the article under discussion (as well as by his colleagues, in particular Olga V. Stolbova), and none of the presented etymologies *in general* can be unambiguously "rejected" (except those components for which more

reasonable etymological solutions can already be proposed, e.g. borrowing from Bagirmi for Chadian *kɛsɛ* etc.).

At the moment, however, most of these etymologies should rather be taken as "food for thought", and a truly serious discussion of the degree of militancy of the Proto-Afroasiatic speakers should probably be postponed until the military terminology systems have been more or less thoroughly reconstructed at the level of individual families or groups. Thus, it looks strange that due to the absence or incompleteness of intermediate reconstructions we have (allegedly) a better understanding of the military terminology at the level of Proto-Afroasiatic language (dating back to the 11th millennium BC!), than, for example, at the level of the Proto-Cushitic language (dating back to the 8th millennium BC). This is clearly an abnormal situation, unequivocally suggesting that a large number of macro-level comparisons may turn out to be accidental coincidences.

Particular care must be taken in any attempt at "cultural reconstruction" with regard to semantics. In modern research whose aim is to truly advance historical science, we should be striving to move away from the aura of "semantic vagueness" that has so far haunted comparative linguistics, towards greater discreteness, manifested at least in a reduction in the level of proto-language synonymy. Thus, in the final list, the author cites no fewer than five allegedly Proto-Afroasiatic terms, the meaning of which includes the component "arrow". What remains unclear is how their semantics differed not only at the Proto-Afroasiatic, but often even at less profound levels. Were these different types of arrows? Or did only one of these words actually mean "arrow", while the others developed this meaning secondarily in one or another of their daughter branches? Of course, there is no guarantee that this issue will ever be thoroughly understood, but at least a more rigorous study of the distribution of these terms in descendant languages, and of the semantic oppositions in their meanings in those languages in which several of these words are represented at once, etc., can be undertaken. Until this research is carried out to the extent to which it is at all possible, talk about the Proto-Afroasiatic semantics of these reconstructions - and, accordingly, talk about those realities of material culture that were known to Proto-Afroasiatic speakers − will remain at the level of pure speculation.

However, to end the commentary on a positive note, it is worth clarifying that some (alas, rather few) of the comparisons cited by Alexander Militarev look like fairly reliable candidates for Proto-Afrasian status. Thus, for example, the root *ʕalw/y- "to attack" seems to be reconstructible both at the Proto-Semitic level (Ugaritic, Sabean, Arabic) and, at least, at the Proto-Eastern Cushitic level (as *ʕawl- "dispute, war"; for semantic reasons the connection with Dahalo ʕáála "bow" is inconclusive). Lack of representation of the root in the Ethiosemitic languages excludes the possibility of borrowing, and its wide occurrence in the Eastern Cushitic group hints at a possible archaic origin (although the listed Omotic parallels are most likely the result of borrowing from a Cushitic source). Whether this term, reconstructible for the Proto-Afroasiatic stage, can be interpreted in a "military" sense is a separate question, requiring a detailed semantic study of all its reflexes in the daughter languages; but at least there are no fundamental complaints about the etymology itself here, either in terms of phonetics, semantics or distribution. Hopefully, in time, all the collected *comparanda* so far, potentially related to the cultural world of the Proto-Afrasians, will be adequately tested for both reconstructibility at the intermediate levels, and for the possibility to construct a fairly clear scenario of their semantic evolution from the original protolanguage to their descendant languages.

### References

Black, P.D. 1975. Lowland East Cushitic: Subgroupingand Reconstruction. PhD diss., Yale University.
Jungraithmayr, H., and D. Ibriszimow. 1994. *Chadic Lexical Roots*. Vol. I, *Tentative Reconstruction, Grading, Distribution, and Comments*. Vol. II, *Documentation*. Berlin: Dietrich Reimer Verlag.

Kießling, R., and M. Mous. 2003. *The Lexical Reconstruction of West-Rift Southern Cushitic*. Koln: Rudiger Koppe Verlag.

Orel, V.E., and O. V. Stolbova. 1995. *Hamito-Semitic Etymological Dictionary: Materials for a Reconstruction*. Leiden: Brill.

Stevenson, R.C. 1969. *Bagirmi Grammar*. Khartoum: University of Khartoum.

Stolbova, O.V. 2016. *Etimologicheskii slovar' chadskikh yazykov* [Chadic Etymological Dictionary]. Moscow: Institut vostokovedeniia RAN.