

Российский государственный гуманитарный университет
Russian State University for the Humanities



Russian State University for the Humanities
Institute of Linguistics of the Russian Academy of Sciences

Journal of Language Relationship

International Scientific Periodical

N°18/3



Moscow 2020

Российский государственный гуманитарный университет
Институт языкознания Российской Академии наук

Вопросы языкового родства

Международный научный журнал

№18/3



Москва 2020

Advisory Board:

H. EICHNER (Vienna) / Chairman
W. BAXTER (Ann Arbor, Michigan)
V. BLAŽEK (Brno)
L. HYMAN (Berkeley)
F. KORTLANDT (Leiden)
A. LUBOTSKY (Leiden)
J. P. MALLORY (Belfast)
A. YU. MILITAREV (Moscow)
V. F. VYDRIN (Paris)

Editorial Staff:

V. A. DYBO (Editor-in-Chief)
G. S. STAROSTIN (Managing Editor)
T. A. MIKHAILOVA (Editorial Secretary)
A. V. DYBO
† S. V. KULLANDA
M. A. MOLINA
M. N. SAENKO
I. S. YAKUBOVICH

Founded by Kirill BABAEV

Редакционный совет:

Х. АЙХНЕР (Вена) / председатель

В. БЛАЖЕК (Брно)

У. БЭКСТЕР (Анн Арбор)

В. Ф. ВЫДРИН (Париж)

Ф. КОРГЛАНДТ (Лейден)

А. ЛУБОЦКИЙ (Лейден)

Дж. МЭЛЛОРИ (Белфаст)

А. Ю. МИЛИТАРЕВ (Москва)

Л. ХАЙМАН (Беркли)

Редакционная коллегия:

В. А. ДЫБО (главный редактор)

Г. С. СТАРОСТИН (заместитель главного редактора)

Т. А. МИХАЙЛОВА (ответственный секретарь)

А. В. ДЫБО

† С. В. КУЛЛАНДА

М. А. МОЛИНА

М. Н. САЕНКО

И. С. ЯКУБОВИЧ

Журнал основан К. В. БАБАЕВЫМ

Вопросы языкового родства: Международный научный журнал / Рос. гос. гуманитар. ун-т; Рос. акад. наук. Ин-т языкознания; под ред. В. А. Дыбо. — М., 2020. — № 18/3. — x + 91 с.

Journal of Language Relationship: International Scientific Periodical / Russian State University for the Humanities; Russian Academy of Sciences. Institute of Linguistics; Ed. by V. A. Dybo. — Moscow, 2020. — No. 18/3. — x + 91 p.

ISSN 2219-3820

<http://www.jolr.ru/>
gstarst1@gmail.com

Дополнительные знаки: С. Г. Болотов
Add-on symbols by S. G. Bolotov

Подписано в печать 10.07.2020. Формат 60×90/8.
Усл. печ. л. 12,8. Уч.-изд. л. 10,0. Заказ № 1125.
Тираж 1050 экз.

Издательский центр
Российского государственного гуманитарного университета
125993, Москва, Миусская пл., 6
www.rggu.ru
www.knigirggu.ru

Table of Contents / Содержание

Table of Contents / Содержание	vii
Contributors / Сведения об авторах	viii
Note for Contributors / Будущим авторам	ix

Articles / Статьи

<i>Ilya Egorov</i> . Basic vocabulary of closely related languages in contact: case study of Turkic languages on the Crimean Peninsula	170
[И. М. Егоров. Базисная лексика близкородственных контактирующих языков на примере крымскотатарских диалектов]	
<i>Alexander Militarev, Sergei Nikolaev</i> . Proto-Afrasian names of ungulates in light of the Proto-Afrasian homeland issue	199
[А. Ю. Милитарев, С. А. Николаев. Праафразийские названия копытных в свете проблемы афразийской прародины]	
<i>A. A. Trofimov</i> . Опыт составления аннотированного списка Сводеша для ведийского языка	227
[Artem Trofimov. An attempt at an annotated Swadesh wordlist for Vedic Sanskrit]	

Book reviews / Рецензии

<i>S. A. МЫЗНИКОВ</i> . Русский диалектный этимологический словарь. Лексика контактных регионов (<i>A. E. Аникин, И. И. Муллонен</i>)	249
[S. A. МЫЗНИКОВ. An Etymological Dictionary of Russian Dialects. Lexicon of Contact Areas. Moscow / Saint-Petersburg: Nestor-Istorija, 2019. 1064 p. (<i>Alexander Anikin, Irma Mullonen</i>)]	

Contributors

Alexander ANIKIN — doctor of sciences (Philology), principal researcher, Institute of the Russian Language of the Russian Academy of Sciences, alexandr_anikin@mail.ru

Илья EGOROV — researcher, School for Advanced Studies in the Humanities, RANEPA (Moscow), i.m.jegorow@gmail.com

Alexander MILITAREV — doctor of sciences (Philology), professor, Center of Ancient Eastern research, Institute for Oriental and Classical Studies, RSUH (Moscow), amilitarev@gmail.com

Irma MULLONEN — doctor of sciences (Philology), principal researcher, Institute of Linguistics, Literature and History of

Karelian Research Centre of the Russian Academy of Sciences (Petrozavodsk), mullonen@krc.karelia.ru

Sergei NIKOLAEV — doctor of sciences (Philology), lead researcher, Department of Slavic linguistics, Institute of Slavic studies, Russian Academy of Sciences (Moscow/Novosibirsk), sergenicko@mail.ru

Artem TROFIMOV — researcher, School for Advanced Studies in the Humanities, RANEPA (Moscow), artemii.trofimov@gmail.com

Сведения об авторах

АНИКИН, Александр Евгеньевич — доктор филол. наук, академик РАН, гл. науч. сотрудник Института русского языка РАН, alexandr_anikin@mail.ru

ЕГОРОВ, Илья Михайлович — научный сотрудник Лаборатории востоковедения и сравнительно-исторического языкознания ШГИ РАНХиГС (Москва), i.m.jegorow@gmail.com

МИЛИТАРЕВ, Александр Юрьевич — доктор филол. наук, проф. ИВКА РГГУ (Москва), amilitarev@gmail.com

МУЛЛОНЕН, Ирина Ивановна — доктор филол. наук, гл. науч. сотрудник Института языка, литературы и истории КарНЦ РАН (Петрозаводск), mullonen@krc.karelia.ru

НИКОЛАЕВ, Сергей Львович — доктор филол. наук, вед. науч. сотрудник Отдела славянского языкознания Института славяноведения РАН (Москва), sergenicko@mail.ru

ТРОФИМОВ, Артем Александрович — научный сотрудник Лаборатории востоковедения и сравнительно-исторического языкознания ШАГИ РАНХиГС (Москва), artemii.trofimov@gmail.com

Note for Contributors

Journal of Language Relationship welcomes submissions from everyone specializing in comparative-historical linguistics and related disciplines, in the form of original articles as well as reviews of recent publications. All such submissions should be sent to the managing editor:

G. Starostin
Institute for Oriental and Classical Studies
Russian State University for the Humanities
125267 Moscow, Russia
Miuskaya Square, 6
E-mail: gstarst1@gmail.com

Articles are published preferably in English or Russian, although publication of texts in other major European languages (French, German, etc.) is possible. Each article should be accompanied with an abstract (not exceeding 300 words) and keywords.

For more detailed guidelines on article submission and editorial policies, please see our website at: <http://www.jolr.ru> or address the editorial staff directly at gstarst1@gmail.com.

Будущим авторам

Журнал *Вопросы языкового родства* принимает заявки на публикацию оригинальных научных статей, а также рецензий от всех, кто специализируется в области сравнительно-исторического языкознания и смежных дисциплин. Рукописи можно высылать непосредственно заместителю главного редактора по адресу:

125267 Москва
Миусская площадь, д. 6
Российский государственный гуманитарный университет
Институт восточных культур и античности
Г. Старостину
E-mail: gstarst1@gmail.com

Предпочтительные языки публикации — английский или русский, хотя возможна также публикация статей на других европейских языках (французский, немецкий и т. п.). К каждой статье обязательно прикладывается резюме (не более 300 слов) и список ключевых слов.

Подробнее о требованиях к оформлению рукописи, редакционной политике журнала и т. п. вы можете узнать на нашем сайте по адресу: <http://www.jolr.ru> или же непосредственно, обратившись к редакции по электронной почте (gstarst1@gmail.com).

Basic vocabulary of closely related languages in contact: case study of Turkic languages on the Crimean Peninsula

The present paper provides two case studies of the basic vocabulary of the Turkic languages spoken on the Crimea Peninsula. Its aim is to illuminate the issues that a historical linguist, and in particular a phylogeneticist, faces when analyzing the basic vocabulary of closely related languages in a situation of intensive contact. The first case study is dedicated to the onomasiological reconstruction of the Proto-Karaim Swadesh list. The main problem here is detection of the West Oghuz loans and especially of contact-induced archaization (fake archaisms) in Crimean Karaim. The objective of the second case study is to identify the genealogical affiliation of the Crimean Tatar dialects. Both the manual analysis of the innovations in the basic vocabulary and the computational lexicostatistics (Bayesian approach, Neighbor-joining, Maximum Parsimony Analysis) confirm the traditional view that the Coastal dialect belongs to the Oghuz subgroup, the Orta dialect – to the West Kipchak subgroup, and the Steppe dialect – to the Nogai Kipchak subgroup. Such affiliations fully fit the documented ethnic history. The correct genealogical affiliation of the dialects in question became possible only after exclusion of all the loans, which has not been done in previous lexicostatistical studies of Crimean Tatar. Both cases show that careful elimination of areal influences is crucial for semantic (onomasiological) reconstruction and phylogenetic studies.

Keywords: phylogeny; semantic reconstruction; lexical borrowings; Karaim language; Crimean Tatar language; Turkic languages.

1. Introduction

The procedure of reconstruction in comparative-historical linguistics implies being able to distinguish between inherited and loaned items and patterns. This statement is true for phonologic, morphologic, and semantic reconstruction. Usually, when a word violates regular sound correspondences, it is treated as a borrowing unless it can be explained as the result of an analogical or another occasional change. Of course, extra sets of sound correspondences can appear between remotely related or unrelated languages as a result of phonological adaptation as well. However, as a rule, such borrowings can be revealed relatively simply, based on their distribution in the contacting subgroup. Various specific problems arise in the case of borrowings from a genetically related language, cf. for example the so-called “etymological nativisation”, described in detail by Ante Aikio (2007) for Finnish loans in the Northern Saami.

Problems caused by contacts between closely related languages are relevant not only for traditional historical-linguistic studies, but also for linguistic phylogeny. The issue of homoplasy and especially horizontal transfer has been redefined in the last decades, see Nakhleh, Ringe & Warnow 2005; Nelson-Sathi et al. 2011 and Kassian 2017. These works make linguists aware of the problem and propose methods to uncover and eliminate it. Early criticisms of Moris Swadesh’s lexicostatistic and glottochronological methods were caused mostly by incorrect interpretation of loans. One of the most known critical works is Knut Bergsland and Hans Vogt’s paper (1962), where it was argued that literary Norwegian (Riksmål) demonstrates a drastically longer distance from Old Norse than Icelandic. The problem was that both contact-

induced and autonomous replacements in the basic vocabulary were considered valid for measuring genealogical distance, whereas in reality the effect of the first group is highly dependent on the specific sociolinguistic situation. Revisiting this case, Sergei Starostin (2000: 230) has shown that 16 of 20 innovations in Riksmål are loans: 11 from Danish, 3 from Swedish and 2 from German. Hence, if they are excluded, the percentage of innovations more or less equals that in the other Scandinavian languages. Nowadays, the detection of loans when reconstructing phylogeny has become an obligatory requirement at least in the Moscow school of comparative linguistics.

However, such drawbacks still arise in more recent phylogenetic studies applying lexicostatistics. For instance, confounding true cognates and borrowings, Russell Gray and Quentin Atkinson (2003) and then Remco Bouckaert et al. (2012) have inferred such a structure for the Slavic group in which Polish forms one clade with Belarusian, Ukrainian, and Russian. This contradicts the existing consensus which assumes a trifurcation of Proto-Slavic into the following subgroups: [Polish, Czech, Slovak, Sorbian], [Slovenian, Serbian/Croatian/Bosnian, Bulgarian, Macedonian], [Belarusian, Ukrainian, Russian]. Such affiliation of Polish is caused by undetected Polish loans in the Belarusian wordlist used in the forenamed works (see the linguistic supplement in Kushniarevich et al. 2015 for more detailed criticism). In section 4.6, I address identical problems in a recent work on Turkic phylogeny.

In the present paper, I intend to discuss two cases which illustrate the problems with the basic vocabulary of the languages undergoing intensive influence on the part of their close relatives. In my investigation of the Turkic languages of Crimea, I attempt to show the challenges they pose to a historical linguist, when the new method of onomasiological reconstruction is applied to identify the genealogical affiliation of a language. The Turkic languages spoken on the Crimean Peninsula provide suitable material for discussion of these issues for the following reasons: (a) they are related to each other approximately at the same depth as Riksmål and its Scandinavian relatives from the canonical example cited above; (b) the tree structure and historical phonology of the Turkic family are known well enough for the purposes of our research; (c) the ethnic, sociolinguistic, and political history of Crimea is well documented.

The remainder of this paper is structured in the following way. Section 2 contains basic information on the Turkic languages of Crimea, their traditional genealogical affiliation, and the sociolinguistic situation in the region, along with a short annotated bibliography. Sections 3 and 4 deal with semantic (onomasiological) reconstruction of the Proto-Karaim wordlist and with revision of the genealogical affiliation of the Turkic varieties spoken in Crimea respectively. Each section contains its own introductory, methodological, analytical subsections and discussions of the results. Section 5 summarizes what can be learned from the considered cases.

2. Turkic languages of Crimea

In this section I provide the most important information on the sociolinguistic situation in the Crimean Peninsula, traditional genealogical affiliation of the languages, dictionaries, grammar and other sources used in the present study.

2.1. Crimean Tatar

The group of dialects traditionally referred to as the Crimean Tatar language actually represents a paraphyletic formation (Sevortyan 1966). It includes (1) Coastal dialect, which is genetically an Oghuz language most closely related to Turkish and Gagauz, (2) Orta (also called

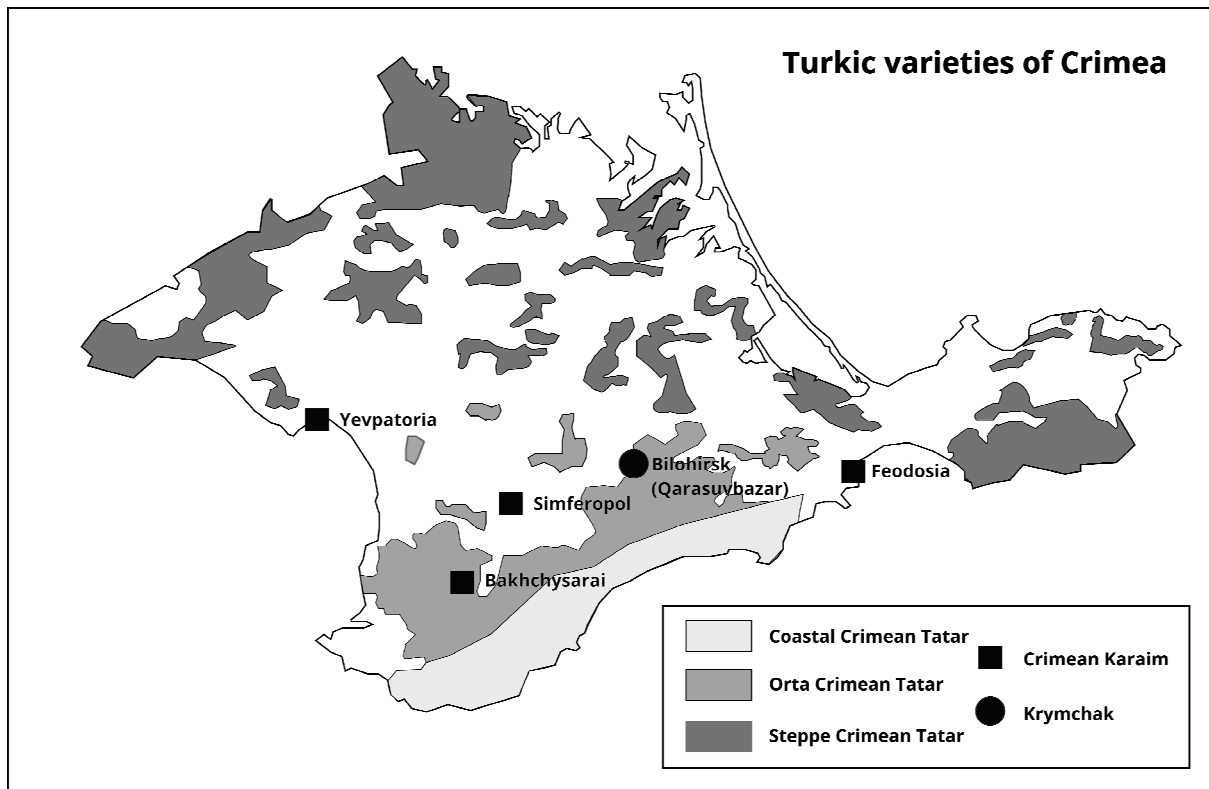


Figure 1. Turkic varieties of Crimea. The map has been drawn on the basis of the Soviet ethnographic map of Crimea 1926; Filonenko 1931 and Radloff 1896: xiv–xvi. The border between the Coastal and Orta dialects is somewhat arbitrary.

Central or Middle) and (3) Steppe dialects, both belonging to different Kipchak subgroups. The Coastal dialect is sometimes named Crimean Turkish; such term reflects its genealogical affiliation exactly. It became the dominant language in the Crimean Khanate, which was a vassal state of the Ottoman Empire. Modern literary Crimean Tatar is based on the Orta dialect.

Dictionaries: Useinov 2007 – dictionary of the Literary Crimean Tatar which is based on the Middle dialect.

Grammars: Sevortyan 1966; Izidinova 1996 – short grammar sketches; Kavitskaya 2010 – grammar based on the field notes from the early 2000s.

Other materials and studies: Polinsky 1992 – 100-wordlists for three Crimean Tatar dialects. I also use the wordlist recently collected after my own initiative, which can be found in Supplement 1.

2.2. Karaim

Karaim (also called Karaite) is a subgroup of Kipchak languages consisting of three dialects, sometimes treated as three separate languages. Only one of them was spoken by Jewish Karaite community in Crimea until recently. Two other dialects (Trakai and Halich) appeared as the result of the migration of Karaims from Crimean Khanate to the Grand Duchy of Lithuania. The migration to Trakai started in 1397, to Halich in 1407–1409 and continued up to the 15th and 16th centuries (Musaev 2010: 205–206). Karaim is traditionally classified together with Karachay-Balkar, Kumyk and with the Middle dialect of Crimean Tatar as a West Kipchak language (Johanson 1998: 82). Maria Polinsky treats Crimean Karaim as an “ethnolect” of Crimean Tatar belonging to the Oghuz subgroup.

Dictionaries: Baskakov, Szapszał & Zajączkowski 1974 remains the most reputable source on the lexicon of all dialects; Aqtay & Jankowski 2015 deals with Crimean Karaim, includes all Crimean materials from Baskakov, Szapszał & Zajączkowski 1974 and from other written sources.

Grammars: Musaev 1964 deals with the Trakai and Halich dialect; Musaev 2010 contains information on Crimean Karaim as well; Prik 1976 – grammar sketch of the Crimean Karaim.

Other materials and studies: Kocaoğlu 2006 – texts in the Trakai dialect with brief grammar sketch and vocabulary; Polinsky 1992 – 100-wordlist for the Crimean dialect; 110-wordlist for Trakai dialect, speaker’s self-recording made in 2019.

2.3. Krymchak

Krymchak is the language of the other Jewish community, which survived until the end of 20th century in the town of Qarasuvbazar (Ukr. *Bilohirsk*). In the late 20th and early 21st century, some attempts at revitalization were undertaken. However, now this language is extinct. In a number of works, Krymchak is treated as a Kipchak language. Polinsky names it (as well as Crimean Karaim) an “ethnolect” of Crimean Tatar, i.e. an Oghuz language.

Dictionaries: Rebi 2004 – the dictionary created by language activists; Ianbay 2016.

Other materials and studies: Polinsky 1992 – 100-item wordlist and short grammar sketch; Polinsky 1991 – text sample; Jankowski 2017 – overview of grammar and major sources.

3. Reconstructing the Swadesh wordlist for Proto-Karaim

3.1. Introductory remarks

Traditional reconstruction of lexical semantics remains extremely arbitrary. A typical meaning of a reconstructed root or even a lexeme is ‘a kind of tree’ or ‘to stack, to collect, to dump, to put in order, to build up’. Such definitions are the results of two wrong methods of semantic reconstruction: 1) reduction of all meanings attested in the daughter languages to a wide semantic component; 2) extrapolation of all attested meanings onto the proto-language. Sometimes this results in openly ridiculous situations: thus, according to Dybo 1996: 18, about 70% of the Proto-Indo-European verbal roots in Julius Pokorny’s dictionary (1959) mean ‘to bloat, to swell’ or ‘to bend’. For further criticism of the traditional semantic reconstructions see Burlak & Starostin 2005: 248. To solve this problem, the method of onomasiological reconstruction has been elaborated in the recent years. I will discuss it in the next section.

Why do we need reconstructed Swadesh lists? It seems reasonable to use reconstructed wordlists for commonly accepted low-level groups when investigating the tree structure of a deeper family. The principle of step-by-step reconstruction is a commonly accepted standard in comparative-historical studies. It is widely applied for phonological reconstruction. For instance, if one introduces a Germanic word into Indo-European comparison, methodologically it is more correct to use a reconstructed Proto-Germanic form instead of Gothic, Old High German, Old North and Old English, since each of them demonstrates innovations that are irrelevant to external comparison. Similarly, more correct is the use of a Proto-Germanic wordlist when reconstructing the Indo-European tree. This exact approach was recently used in Kassian et al. forthcoming. The reconstructed Proto-Karaim wordlist can be used when revising the topology of Turkic family and reconstructing the Proto-Turkic wordlist for further comparison.

3.2. Methods

A relatively strict method of onomasiological reconstruction was recently developed by the representatives of the Moscow School of comparative linguistics (see Kassian, Starostin & Zhivlov 2015: 304–306; Starostin 2016). It involves tracing a way from the meaning to its optimal exponent in the protolanguage, i.e., determining which word was used for a given concept in a protolanguage. Selection of the optimal candidate is guided by five principles, which are very similar to the ones used for detecting archaisms and innovations when reconstructing phonology. The basic principle is topological (1); others (2–5) are used in competitive situations, i.e. when tree topology allows no unambiguous judgment on the candidates. Here I only provide a brief synopsis; for strict definitions, further explanations and examples see the abovementioned works:

- 1) tree topology: the root attested in different branches is preferable;
- 2) external etymology: the root is preferable if its external cognates preserve the same Swadesh meaning;
- 3) internal derivability: the primary root (as opposed to polymorphemic derivatives) is a preferable candidate;
- 4) typology of semantic shifts: the typologically frequent direction of semantic shifts to be assumed when reconstructing scenario of semantic changes of the potential candidates;
- 5) areal effect exclusion: contact-induced innovations to be excluded.

It must be noted that these principles are different from the criteria for synchronic wordlists: terms for synchronic Swadesh wordlists are selected based on its frequency and stylistic neutrality.

Another advantage of onomasiological reconstruction is that it allows the elimination of uncertainties and mistakes in a synchronic list. Thus, if the quality of data on one language is not irreproachable, this can be compensated for with data on its relatives. In the process of reconstruction, an incorrectly selected term for one of the languages is likely to be seen as an innovation in this particular idiom. If a wordlist is overcrowded with inappropriate archaisms, the situation becomes more difficult. The solution to this problem is proposed in Section 3.3. Although the probability that one inaccurate list will influence the structure of the phylogenetic tree is lower and onomasiological reconstruction helps fix some defects in the data, the motto “garbage in – garbage out” remains fully true.

Next, I will concentrate on detecting contact innovations in Crimean Karaim, which make relatively shallow Proto-Karaim reconstruction difficult. Since the phonological inventory of Crimean Karaim and Coastal Crimean Tatar is very similar to each other, the phonological criterion is not particularly helpful in this case. Obviously, the reflexes of Proto-Turkic stems differ in these languages and sometimes these differences point to the Oghuz origin of the word. However, this criterion cannot be applied to every word. I will mostly use the distributional criterion, which can be described as follows. For instance, four languages (L1, L2, L3, L4) related to each other among which L1 is an outgroup and L2, L3, L4 form a separate clade are taken into consideration. The lexeme A is the basic term for the meaning ‘M’ in L1. In L2, lexemes A and B are synonyms; in L3, L4, ‘M’ is denoted only by the lexeme B (see Figure 2). If A is a primary root whereas B is a transparent derivative or semantic innovation, this constitutes strong evidence for reconstructing *A for the meaning ‘M’ in the Proto-L2–4. However, if L1 influences L2, A should be regarded as a borrowing. The coexistence of A and B in L2 is an additional argument for such a solution.

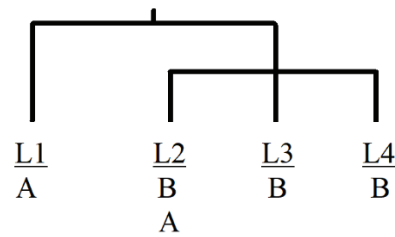


Figure 2. Tree structure and distribution of lexemes with meaning ‘M’ which hint at horizontal transfer of the stem A.

Extrapolating this scheme to the case in question, I will regard a word as an Oghuz borrowing in Karaim if it is widely spread in the Oghuz or at least in the West Oghuz languages and is uncommon or completely absent in Kipchak.

In addition to being subject to horizontal transfer (MAT-borrowings in Jeanette Sakel’s (2007: 15) terminology), some stems can also undergo contact-induced semantic shifts (PAT-borrowings in Sakel’s terminology). The last phenomenon is also known as loan meaning extensions (see Haspelmath 2009). In the case of closely related languages, this is driven by obvious, naïve logic: “if words sound similarly they must have similar meaning”.

3.3. Contact innovations in Crimean Karaim: clear cases

More evident borrowings will be considered at first. All of them are loaned from the Coastal dialect of Crimean Tatar. Some of these loans were in turn borrowed into West Oghuz languages from Persian and Arabic. Theoretically, one can assume that Oghuz-like lexemes in Crimean Karaim are inherited and Kipchak-like ones are borrowed. However, such assumption faces more difficulties, since it is hardly possible to find the source of potential Kipchak borrowings which occurred in the Trakai and Halich dialects and sporadically in the Crimean dialect as well.

In citing examples, I first give the number of the concept in the 110-item Swadesh list following Kassian et al. 2010, then list the language material with references. The abbreviation for the source used in Baskakov, Szapszał & Zajączkowski 1974 and Aqtay & Jankowski 2015 is given in brackets for Crimean Karaim forms. For the full reconstructed Proto-Karaim Swadesh wordlist see Supplement 1. For the transcription and transliteration of the examples, I use the Unified Transcription System applied in the Global Lexicostatistical Database (<https://starling.rinet.ru/new100/UTS.htm>).

5. **big** – CrKar. *balaban* (Sz) ‘big, huge’ (Baskakov, Szapszał & Zajączkowski 1974: 100). It is difficult to define whether *balaban* is appropriate even for the synchronic Crimean Karaim Swadesh list. There are two other candidates that will be considered below. CrKar. *balaban* is a clear Oghuz borrowing, cf. Tur. *balaban* ‘huge’, Gag. *balaban* ‘high’, CoCrTat. *balaban* ‘big’. The root is extremely rare beyond the Oghuz languages. Only two *comparanda* are mentioned in Dybo 2013: 128–131: Tat. dial. *balban* ‘fat, overweight, stout’, Kirg. *balpay-* ‘to seem big, bulky, clumsy’. Details of its etymology remain obscure (see the cited work for the review of existing hypotheses), however, the innovative nature of the meaning ‘big, huge’ is quite obvious.

CrKar., TrKar., HKar. *biyik* (Sz, Par. 101 v. 1) ‘big, high, great’ (Baskakov, Szapszał & Zajączkowski 1974: 115) is another stem which could be treated as a contact-induced innovation. This stem with the meaning ‘high’ is widespread across Turkic languages (see Dybo 2013: 123) whereas the more general meaning ‘big’ is limited to the Oghuz subgroup (Tur. *büyük*, Gag. *bü:k*, Az. *böyük*), Karakhanid Uyghur, Old Uyghur, and Sary Yugur. Such distribution theoretically can be an indication of the antiquity of the meaning ‘big’ (cf. Clauson 1972: 302 for an in-

terpretation), but **ulu* is the better candidate for both Proto-Karaim and Proto-Turkic ‘big’. All Karaim dialects demonstrate its reflexes: CrKar. *ulu* (Sz) ~ *uli* ‘great, big’, TrKar. *ullu* ‘big, great, important’ ~ *unlu* ‘great, big, elder’ ~ *ullux* ‘big, great’ (with additional suffix), HKar. *ullu* ‘big, great, important’ (Baskakov, Szapszał & Zajączkowski 1974: 577, 579). At least modern speakers of the Trakai dialect use *ullu* as the basic word for ‘big’, according to our data; it is also confirmed with materials published in Kocaoğlu 2006. The semantic shift ‘big’ > ‘great’ is typical for the world’s languages, the synchronic polysemy is also common in the Turkic family (Dybo 2013: 120–121) and cross-linguistically (Rzymiski et al. 2019). The direction of the shift ‘big (of a physical object)’ > ‘great (high status)’ is more probable than vice versa due to the common tendency of the development from concrete meanings to more abstract ones (Campbell 2013: 237). In sum, the old term for ‘big’ in the Proto-Karaim subgroup is **ullu*; the semantic shift ‘high’ > ‘big’ of the stem **biyik* should have been triggered by contact with the Oghuz dialect of Crimea before the start of the migration to the Grand Duchy of Lithuania. Note also the form *büyük* in the Crimean dialect, which looks as if it was recently borrowed from Turkish. Aqtay and Jankowski (2015: 88, 100) list the latter form with the gloss ‘great, big’ whereas *biyik* is glossed as ‘high’.

It is difficult to make a choice between three candidates with the meaning ‘big’ for synchronic Crimean Karaim based on existing sources, which partly contradict each other. Provisionally, I assume that *ulu* ~ *uli* has the more abstract meaning ‘great’ whereas *biyik* ~ *büyük* and *balaban* compete with each other in the basic meaning ‘big’. Both are loans, the meaning ‘big’ by the lexeme *biyik* is borrowed from Oghuz; *büyük* and *balaban* are MAT-borrowings. The choice is much simpler for Proto-Karaim. Thus, this case illustrates an important advantage of onomasiological reconstruction: uncertainty in the data on one of the languages does not influence the final list.

67. **red** – CrKar. *qirmizi* ~ *qirimzi* ‘red’. This is an Arabic loan common in Oghuz languages, cf. Tur., Gag. *qirmizi* ‘red’, Az., Turkm. *qirmizi* ‘red’. In the Trakai and Halich dialects, *kirmizi* denotes a specific shade ‘purple, magenta’ (Baskakov, Szapszał & Zajączkowski 1974: 381, 387). The archaic stem *kizil*, reflecting PT **Kir^vil* ‘red’, has been found with the meaning ‘red, orange’ for these dialects; it should be the basic term for ‘red’ (Baskakov, Szapszał & Zajączkowski 1974: 383). The reflexes of **Kir^vil* are not attested for Crimean Karaim in Baskakov, Szapszał & Zajączkowski 1974, but Aqtay and Jankowski (2015: 309) cite it with the gloss ‘red, ruddy’. The stem **qizil* can be reconstructed for Proto-Karaim ‘red’ with complete certainty. It is one of the most stable Turkic stems.

69. **root** – CrKar. *kök* ‘root’ (Baskakov, Szapszał & Zajączkowski 1974: 337). This stem is widely attested with the meaning ‘root’, however, its basic meaning in the majority of the Turkic languages is more abstract (‘basis’), it develops various metaphorical meanings as well. As the basic term for ‘root’ this stem is attested in the Oghuz languages, cf. Tur., Gag. *kök*, from which it has been borrowed into Crimean Karaim. The stem **damor* > **tamur* is a better candidate for Proto-Turkic and Proto-Karaim ‘root’. It is preserved all over the Turkic-speaking area including the languages in question and their numerous Kipchak relatives, cf. Karaim reflexes: CrKar., *tamur* ~ *tamar* ‘root, vein’, HKar. *tamar* ~ *tamur* ‘root, vein’, TrKar *tamur* ‘root, vein’ (Baskakov, Szapszał & Zajączkowski 1974: 509–510).

70. **round** – CrKar. *müdever* ~ *mudever* ‘round’, *yuvarlaq* ‘round, globular’, *tomalaq* ‘round, full, plump’, *yumalaq* ‘globular, round’ (Aqtay & Jankowski 2015: 247, 407, 472, 474), *tögerek* ‘round’ (Baskakov, Szapszał & Zajączkowski 1974: 541). The first term has an Arabic origin and is borrowed via Turkish, cf. Tur. *müdever* ‘circular, round’. The second one is an Oghuz borrowing as well, cf. Tur., Gag. *yuvarlaq* ‘round’. The stems *tomalaq* and *yumalaq* have been attested in other Turkic languages:

Uzb. *yumaloq* ‘round (sphere & circle)’, Karak. *žumalaq* ‘round (circle)’, for other derivatives with the Proto-Turkic bound root **yum-* see Dybo 2013: 441–442;

Gag. *tombarlaq* ‘round’, Uzb. *dumaloq* ‘round (sphere & cylinder)’, Uyg. *domlaq* ‘round (sphere)’, Bash. *tumalaq* ‘round (sphere & circle)’, Nog. *timalaq* ‘circle, sphere (n.)’, Kaz. *domalaq* ‘sphere (n.)’, Karak. *dumalaq* ‘round (sphere)’ – note that this set of phonetically similar forms demonstrates suspiciously irregular sound correspondences!

However, none of the Karaim stems listed above have been sufficiently confirmed by the Trakai and Halich data; the meaning ‘round’ is insufficiently documented in the existing sources. Only the stem *tögerek* has a Trakai cognate. According to the recently collected word-list for the Trakai dialect, either the collocation *galgan kibik* literally means ‘circle-like’ or *tägere* ~ *tegäräk* ‘round, circle’ is used as an adjective ‘round’. CrKar. *tögerek* and TrKar. *tägere* ~ *tegäräk* are treated as a Mongolian reborrowing and an inherited stem respectively in Dybo 2013: 238–239. However, I believe that it is reasonable to consider them true cognates and to reconstruct **tögerek* for Proto-Karaim with the meaning ‘round (circle)’.

Until there is a corpus for Karaim, the choice for the synchronic basic term is difficult both for the Crimean and Trakai dialects. The stem *tägere* ~ *tegäräk* is a single candidate for Trakai ‘round’, Crimean *tögerek* can apply at least for ‘round 2D’. Further details in synchronic dialects remain obscure.

79. **smoke** – CrKar., TrKar. *tüt-sü* (Sz) ‘smoke, incense’ (Baskakov, Szapszał & Zajączkowski 1974: 555). This is an old contact innovation shared by the Crimean and Trakai dialects. The substantives from the verb PT **tüt-* formed with the not especially productive suffix **-süg* have been found only in Oghuz languages, cf. *tütsü* Tur. ‘incense’, Az. *tüstü* ‘smoke’ with metathesis inside the consonant cluster, Turkm. *tüsse* ‘smoke’, Sal. *tissi* ‘smoke’, for this suffix see Räsänen 1957: 141. All other Turkic languages, including even Chuvash, demonstrate the suffix **-ün* (Dybo 2013: 479). So we treat CrKar. and TrKar. *tüt-sü* ‘smoke, incense’ as a borrowing which occurred before the migration from Crimea to the Grand Duchy of Lithuania. The inherited forms with **-ün* have been found in all Karaim dialects as well: CrKar. *tütün* ‘smoke, tobacco’, TrKar. *t^yut^yun^y* ‘smoke, tobacco’, HKar. *titin* ‘smoke’ (Baskakov, Szapszał & Zajączkowski 1974: 532, 555, 571). Hence, **tütün* must be reconstructed for Proto-Karaim ‘smoke’.

103. **near** – CrKar. *yaqın* is attested in the Crimean dialect (Baskakov, Szapszał & Zajączkowski 1974: 220) beside CrKar. *yurwuuq*, TrCar. *yurwux*, HCar. *yurwuk* ‘near’ (Baskakov, Szapszał & Zajączkowski 1974: 253–254). I consider the first stem an Oghuz borrowing: Tur. and Gag. *yaqın* ‘near’, Az. *yaχın* ‘near’, Turkm. *yaqı:n* ‘near’. Both stems are widely spread across Nuclear Turkic languages. However, the narrow distribution in the Karaim dialects allows us to treat **yaqın* as a borrowing. Its competitor, **yurwuuq*, which can be found in all Karaim dialects, is definitely the better candidate for Proto-Karaim ‘near’. The distribution of its external cognates points to the stem discussed above as to the main exponent of the meaning ‘near’ not only in Proto-Kipchak and even in Proto-Turkic, see Dybo 2013: 539–540.

3.4. Contact archaization in Crimean Karaim: fake archaisms

In this section, I consider the most curious cases. There are some stems which can seem archaic at first sight, but in reality turn out to be loanwords. For such cases, I suggest the term ‘fake archaisms’. Revealing this kind of borrowings is crucial for onomasiological reconstruction. Fake archaism must be suspected when principles of tree topology, external etymology, and internal derivability come in conflict with the principle of areal effect exclusion. The semantic plausibility principle, i.e. the typology of semantic shifts, theoretically, can also contradict the principle of areal effect exclusion but such cases have not been attested in our material.

Thus, fake archaisms can be successfully detected when areal distribution and the direction of influence are taken into account. In Crimean Karaim, four examples of fake archaisms have been found.

22. **to eat** – there are two candidates for filling this slot:

1) CrKar. *ye-* (Sz) ‘to eat’ (Baskakov, Szapszał & Zajączkowski 1974: 268);

2) CrKar., TrKar. *aš-a-* (Sz) ‘to eat’, HKar. *as-a-* ‘to eat’ (Baskakov, Szapszał & Zajączkowski 1974: 79, 91).

The root **ye-* should have an advantage due to the principle of external etymology. It is found not only in numerous non-Kipchak languages but even in Chuvash, whereas **aš-a-* is limited to the Nuclear Turkic languages. In many of them, it is often a marked polite term ≈ Rus. *kušat*’. The principle of internal derivability also speaks for the primary root **ye-*, since the verbal stem **aš-a-* can be analyzed as a synchronic derivative from **aš-* ‘food’. However, **ye-* is limited only to the Karaim dialect that was under intense influence on the part of Oghuz. The Oghuz languages preserve **ye-* as the basic exponent of ‘to eat’ (Tur., Az. *ye-*, Gag. *i-*, Turkm. *iy-*, Sal. *yí-*). Thus, one can simply consider CrKar. *ye-* a borrowing. A probable situation is that **aš-a-* already becomes the basic term for ‘to eat’ in Proto-Karaim, but archaic **ye-* as a marginal term still remains in Proto-Karaim. Under foreign influence **ye-* could become the basic term again, i.e. we deal with a semantic backformation.

83. **sun** – three words glossed in this way have been found in Crimean Karaim, and two of them can apply for the status of the basic term in Proto-Karaim.

1) CrKar. *kün* ‘sun, day’ (Sz, R) ~ *gun* ‘day’ (Par 77 v. 11), cf. TrKar. *k^yun^y* ‘day’, HKar. *kin* ‘day’ (Baskakov, Szapszał & Zajączkowski 1974: 167, 320, 353, 396);

2) CrKar. *küneš* ~ *güneš* ‘sun’ (Baskakov, Szapszał & Zajączkowski 1974: 354; Aqtay & Jankowski 2015: 169, 225);

3) CrKar. *quyaš*, TrKar. *kuyaš*, HKar. *kuyas* ‘sun’ (Baskakov, Szapszał & Zajączkowski 1974: 344, 372).

The first item attested mostly with the meaning ‘day’ demonstrates also the meaning ‘sun’ in Crimean Karaim. It is a reflex of the stable Proto-Turkic stem **gün* ‘day, sun’ which retained this meaning across the whole area of the Turkic languages. The second stem was derived from the first one with a not quite clear suffix. It occurs sporadically in various languages, cf. Tur., Kum. *güneš* ‘sun’, OT *küneš* ‘sun’. The Crimean Karaim form *küneš* ~ *güneš* is a transparent Western Oghuz loan due to the initial voiced consonant. The last stem, *quyaš*, is a result of the semantic shift ‘heat’ > ‘sun’, which should independently occur in a couple of Turkic languages. Thus, when one chooses between **kün* and **quyaš*, the external etymology principle strongly points to the first stem as the better candidate for filling the slot ‘sun’ in Proto-Karaim. However, in light of Tur., Gag. *gün* ‘sun, day’ (attested simultaneously with *güneš* in Turkish), it is reasonable to regard the meaning ‘sun’ of CrKar. *kün* as a result of backformation. Hence, the slot ‘sun’ must be filled by the stem **quyaš* in Proto-Karaim. The retention of the stem **kün* in Halich and Trakai collocations *k^yun^y batış* ‘sunset’ (lit. ‘sun diving’), *k^yun^y tuvuš* ‘sunrise’ (lit. ‘sun appearing’) can prove its antiquity in this meaning and, hence, the existence of **kün* ‘sun’ in Pre-Proto-Karaim, but it remains questionable whether this evidence is sufficient to reconstruct Proto-Karaim **kün* as the basic term for ‘sun’.

84. **to swim** – two candidates for this slot have been found:

1) CrKar. *yüz-* (Sz) ~ *üz-* (Sz) ‘to swim’ (Baskakov, Szapszał & Zajączkowski 1974: 261, 588);

2) CrKar. *čöm-* (Sz) ‘to swim, to dip’, TrKar. *čom-* ‘to swim, to dip’, HKar. *com-* ‘to swim, to flow’ (Baskakov, Szapszał & Zajączkowski 1974: 614, 632, 639).

The first one reflects the relatively stable Proto-Turkic stem **yür^y-* ‘to swim’, cf. Oghuz comparanda: Tur., Turkm. *yüz-* ‘to swim’, Gag., Az. *üz-* ‘to swim’, it is also common beyond the Oghuz subgroup (Dybo 2013: 490). This stem can be safely reconstructed for Proto-Turkic

‘to swim’. The second candidate is a transparent innovation. It reflects the semantic shift ‘to dive, to dip’ > ‘to swim’. The original meaning is confirmed by a number of languages:

Tuv. *šim-in-* (refl.) ‘to dip, to dive’, OUyg., KarakhUyg., Chag., Uyg. *čom-* ‘to dip, to dive’, Uzb. *čqm-* ‘to dip, to dive’, Tat. *čum-* ‘to dip, to dive’, Chuv. *čbm-* ‘to dip, to dive’ (Tat. borrowing?), Bash. *sumi-* ‘to dip, to dive’ (Dybo 2013: 491).

Note that the polysemy ‘to swim, to dive’ is attested in the Karaim dialects as well. Based on the external etymology principle, one could reconstruct **yüz-* for Proto-Karaim ‘to swim’. However, this stem must be regarded as a borrowing since it is limited to the one dialect in intimate contact with West Oghuz, while the archaic stem, on the contrary, is retained in Karaim.

86. **that** / 87. **this** – the system of demonstrative pronouns in the Crimean Karaim has been influenced on the part of Oghuz languages.

	Crimean	Trakai		Halich	
proximal	<i>bu</i> ‘this’	<i>bu</i> ‘this’	<i>ušpu</i> ‘this here’	<i>bu</i> ‘this’	<i>uspu</i> ‘this here’
medial	<i>šu</i> ‘this, that’	—	—	—	—
distal	<i>ol</i> ‘that’	<i>ol</i> ‘that’	<i>ošol</i> ‘that there’	<i>ol</i> ‘that’	<i>osol</i> ‘that there’

Table 1. The subsystems of the demonstrative pronouns in the Karaim languages.

The three-way deictic system, like in Crimean Karaim, can be potentially treated as archaic. Proto-Nuclear-Turkic **šu* functions as a medial deictic pronoun in several languages:

Gag. *šu* ‘this, that (medial deixis)’, Turkm. *šu* ‘this, that (medial deixis)’, Uzb. *šu* ‘this, that (medial deixis)’, Kum. *šu* ‘this, that (medial deixis)’, Kirg. *šu* ‘this, that (medial deixis)’.

Theoretically this could confirm the antiquity of CrKar. *šu*. The systems with bare **šu* as a medial deictic pronoun are common in the Oghuz languages (Tenishev & Dybo 2002: 145–156), but not typical for other Turkic subgroups. Outside Oghuz, the Proto-Turkic pronominal root **šu* is more frequently attested with various extensions:

Chuv. *šavz, šakz* ‘this’, *leš* ‘that’, Yak. *sol* ‘that’, Turkm. *šol* ‘that’, Bash. *ošo* ‘this’, *šul* ‘that’, Tat. *šul* ‘that’, Nog. *sosi* ‘this’, *sol* ‘that’, Kaz. *osi* ‘this’, *sol* ‘that’, Karak. *usi* ‘this’, *sol* ‘that’, Kir. *ušu* ‘that’.

The fact that Crimean Karaim, Kumyk, and Kirgiz feature simply **šu* sets them apart from other Kipchak languages. Therefore, it may be suspected that Crimean Karaim demonstrates another Oghuz loan. Thus, only **bu* ‘this’ and **ol* ‘that’ can be reconstructed for Proto-Karaim with certainty. In fact, **šu* must not be a deictic pronoun but rather a deictic particle, see Dybo 2013: 497–498.

3.5. Phonological variation in Crimean Karaim

Another result of strong Oghuz influence on Crimean Karaim is the presence of phonological doublets which reflect both Kipchak and Oghuz development of the same Proto-Turkic root. The Oghuz-like counterparts are borrowings. Due to the fact that Oghuz looks more archaic than Kipchak in some parameters, these cases, considered in Sections 3.4.1 and 3.4.2, can also be regarded as fake archaisms.

3.5.1. Reflexes of PT **g*

To the basic distinctions between Oghuz and Kipchak languages belong the reflexes of **g* after a low central vowel. The Oghuz languages demonstrate an uvular consonant whereas the west

majority of the Kipchak languages change the velar to a labial. A school-book example is the reflex of Proto-Turkic **da:g* ‘mountain’:

Oghuz: Tur. *da:* (dial. *daɣ*), Az. *daɣ*, Turkm. *da:g*, Sal. *da:ɣ*;

Kipchak: Kum. *taw*, K.-B. *taw*, Tat. *taw*, Bash. *tau*, Nog. *taw*, Kaz. *taw*.

Crimean Karaim demonstrates both *taw* ‘forest’ (Sz) and *taɣ* ‘mountain’ (Sz); the third variant is *daɣ* ‘mountain’ (ZR 45, 3). These forms contrast with TrKar., HKar. *taw* ‘mountain’ (Baskakov, Szapszał & Zajączkowski 1974: 168, 503, 505). The Crimean Karaim form *taɣ* must be treated as phonologically adopted. Voiced *d* was substituted with voiceless *t*, since only voiceless dentals are possible in word onset in the inherited vocabulary. The final velar does not undergo the adaptation since there is no general restriction on *ɣ* after vowels at least in the non-final position, cf. *alʒaɣim* ‘I will take’, *qartniŋ tayawɣi* ‘old man’s stick’. A simultaneous occurrence of adopted (to various degrees) and non-adopted items is typical for the situation of intensive influence, cf. Russian loans in Kazym Khanty and Finish loans in Northern Saami:

Khant. *áškola* ~ *škola* ‘school’ < Rus. *škola* ‘school’;

Khant. *wəntər* ~ *andrey* ~ *andrʲey* ‘a male personal name’ < Rus. *Andrey* ‘a male personal name’;

SaaN. *hirbmat* ~ *harbmat* ‘horrible’ < Fin. *hirmu* ‘horror’ (Aikio 2007: 28–29);

SaaN. *hapmu* ‘craving (for a particular food)’ ~ *hipmu* ‘lust, desire’ < Fin. *himo* ‘carving, desire’ (Aikio 2007: 28–29).

Another example for **-ag* in the final position found in the Karaim Swadesh list:

CrKar. *yaw* (Sz) ~ *yaɣ* (Sz) ‘fat’, cf. TrKar., HKar. *yaw* ‘fat’ (Baskakov, Szapszał & Zajączkowski 1974: 214–215).

Reflexes of the Proto-Turkic vocalic-consonantal cluster **-agi-* are a special case. In Kipchak, not only does **g* become a labial consonant, but **i* also becomes a rounded vowel. Oghuz demonstrates here an uvular consonant and an unrounded vowel.

CrKar. *awur* (Sz) ~ *awɣr* (Sz, R) ‘heavy’, cf. TrKar., HKar. *awur* ‘heavy’ (Baskakov, Szapszał & Zajączkowski 1974: 42, 44);

CrKar. *awuz* (Sz) ~ *awiz* (Sz) ~ *awɣz* (Par 84 v. 9) ‘mouth’, cf. TrKar., HKar. *awuz* ‘mouth’ (Baskakov, Szapszał & Zajączkowski 1974: 42, 44);

CrKar. *bawur* ‘liver’ (Sz) ~ *baɣr* (Sz) ‘chest, liver’, cf. TrKar., HKar. *bawur* ‘liver’, TrKar. *bawɣr* ‘liver’ (Baskakov, Szapszał & Zajączkowski 1974: 94, 96).

In Tenishev & Dybo 2006: 72–73, the double reflexes of **ag#* and **agi* have been postulated for Karaim, i.e. *aw* ~ *aɣ* and *awu* ~ *aɣi*. It seems more reasonable to regard the reflexes with *ɣ* as a result of Oghuz influence. If they are eliminated, Karaim will not differ from other Kipchak languages in its reflexes of **ag#* and **agi*. In the opposite case, the Karaim data would require reconstructing velar (or rather uvular) consonants for Proto-Kipchak in these clusters.

3.5.2. Initial voiced dental and velar consonants

Turkish and Gagauz reflect the Proto-Turkic distinction of initial voiced and voiceless dental and velar stops. For velars the opposition can be reconstructed only in roots with front vowels. The reconstruction of the initial Proto-Turkic voiced stops and some modifications which occurred in the Oghuz languages are described in all details in Tenishev & Dybo 2002: 68–83 (see also Dybo 2007 for further details and discussion). The majority of the Kipchak languages neutralize these oppositions in favor of the voiceless series. Crimean Tatar demonstrates contact-induced variation.

CrKar. *keča* ‘night’ (Par 83 v. 3) ~ *keče* ‘night’ (Sz, Man 3a, 8a) ~ *geže* ‘evening’ (Kž III–IV, 81) ~ *geče* ‘night’ (ZR 52, 26, Q 9) ~ *geča* ‘night’ (ZR 52, 20), cf. TrKar. *kʲečʲa* ‘night’, HKar. *kece*

‘night’ (Baskakov, Szapszał & Zajączkowski 1974: 159, 167, 311–312, 394–395; Aqtay & Jankowski 2015: 164, 202) < PT **ge:če*;

CrKar. *kel-* ‘to come’ (Sz, Cam, Dan 1:1, Man 2a) ~ *gel-* ‘to come’ (Man 3a, Q34), cf. TrKar., HKar. *kel-* ‘to come’ (Baskakov, Szapszał & Zajączkowski 1974: 301–302, 390; Aqtay & Jankowski 2015: 164, 204) < PT **gel-*;

CrKar. *köz* ‘eye’ (Sz, Cam, Psa 10:1, Man 5a) ~ *göz-* ‘eye’ (Par 82 v. 1, ZR 79, 15, ZR 95, 30, Man 5a, Q 4), cf. TrKar. *k^yoz^y-*, HKar. *kez-* ‘eye’ (Baskakov, Szapszał & Zajączkowski 1974: 161, 300, 312, 336; Aqtay & Jankowski 2015: 168, 221) < PT **gör^y*;

CrKar. *kör-* ‘to see’ (Par 83 v. 5, Man 1a, Q 38) ~ *gör* ‘to see’ (Q 36, 49), cf. TrKar. *k^yor-*, HKar. *ker-* ‘to see’ (Baskakov, Szapszał & Zajączkowski 1974: 306, 314, 339; Aqtay & Jankowski 2015: 167–168, 218) < PT **gör-*;

CrKar. *kün* ‘sun, day’ (Sz, R, Man 3a) ~ *gun* ‘day’ (Par 77 v. 11, Man 2b, Q 73), cf. TrKar. *k^yun^y* ‘day’, HKar. *kin* ‘day’ (Baskakov, Szapszał & Zajączkowski 1974: 167, 320, 353, 396; Aqtay & Jankowski 2015: 169, 224) < PT **gün*;

CrKar. *taš* ‘stone’ (Sz, Fil 7, 120, Q 81) ~ *daš* ‘stone’ (Par 83 v. 12, Q 21), cf. TrKar. *taš* ‘stone’, HKar. *tas* ‘stone’ (Baskakov, Szapszał & Zajączkowski 1974: 170, 516, 518; Aqtay & Jankowski 2015: 132, 386) < PT **dia:λ*;

CrKar. *taw* ~ *taκ* ~ *daκ* < PT **da:g* (details see above);

CrKar. *tamar* ‘vein, root’ (Sz) ~ *tamur* ‘vein, root’ (Sz, Q 431) ~ *damar* ‘vein’ (ZR 78, 18), cf. TrKar., HKar. *tamur* ‘vein, root’, HKar. *tamar* ‘vein, root’ (Baskakov, Szapszał & Zajączkowski 1974: 169, 509–510; Aqtay & Jankowski 2015: 131, 381) < PT **damor*;

CrKar. *terek* ‘tree’ (Man 10a) ~ *teraq* ‘tree’ (Par 83 v. 14) ~ *derek* ‘tree’ (Fil 8, 150, Q 58) ~ *direk* ‘tree’ (Sz, ZR 44, 30) ~ *diraq* ‘post, column’ (ZR 16, 21), cf. TrKar. *t^yer^yak* ‘fruit tree’, HKar. *terek* ‘id’ (Baskakov, Szapszał & Zajączkowski 1974: 178, 185, 522, 565, 567; Aqtay & Jankowski 2015: 136, 396) < PT **derek*;

CrKar. *tüz* ‘knee’ (Sz) ~ *diz* ‘knee’ (Q 628) ~ *düz* ‘knee’ (KM 61b), cf. TrKar. *tiz* ~ *tiz^y* ‘knee’, HKar. *tiz* ~ *kiz* ‘knee’ (Baskakov, Szapszał & Zajączkowski 1974: 317, 525–526; Aqtay & Jankowski 2015: 139, 144) < PT **dir^y*;

CrKar. *tolı* ‘full’ (Sz, Par 102 v. 13) ~ *tolu* ‘full’ (Sz, R) ~ *dolı* ‘full’ (Q 187), cf. TrKar., HKar. *tolu* ‘full’ (Baskakov, Szapszał & Zajączkowski 1974: 537; Aqtay & Jankowski 2015: 140, 407) < PT **do:l-*;

CrKar. *til* ‘tongue’ (Q 18, 49, Meq 60, 70) ~ *dil* ‘tongue’ (Q 223), cf. TrKar. *til^y*, HKar. *til* ~ *kil* ‘tongue’ (Baskakov, Szapszał & Zajączkowski 1974: 319, 528; Aqtay & Jankowski 2015: 403, 138) < PT **dıl* ~ **dil*;

CrKar. *tiš* ‘tooth’ (Sz) ~ *čiš* ‘tooth’ (Sz, Q 125) ~ *diš* ‘tooth’ (ZR 70, 12, Q 302); cf. TrKar. *tiš* ‘tooth’, HKar. *tis* ~ *kis* ‘tooth’ (Baskakov, Szapszał & Zajączkowski 1974: 178, 323, 531–532, 629; Aqtay & Jankowski 2015: 124, 132, 404) < PT **di:λ*;

CrKar. *tur-* ‘to stand’ (Sz, Man 5a) ~ *dur-* ‘to stand’ (Par 77 v. 12, Q 54), cf. TrKar., HKar. *tur-* ‘to stand’ (Baskakov, Szapszał & Zajączkowski 1974: 181, 547; Aqtay & Jankowski 2015: 142, 413) < PT **dur-*.

It should be noted that not all Proto-Turkic stems with initial **d* found in Swadesh list demonstrate voiced consonants in Crimean Karaim: PT **dırıŋa-k* > CrKar. *tırnaq* ‘fingernail, claw’, PT **di:λ-le-* > CrKar. *tišle-* ~ *čišle-* ‘to bite’, **dəri* > CrKar. *teri* ‘skin’, **dur^y* > CrKar. *tuz* ‘salt’. Such inconsistency indicates that in this case they are not regular reflexes but borrowings. The Proto-Turkic stems with initial **k* and **t* are found always with voiceless consonants: PT **kül* > CrKar. *kül* ‘ashes’, PT **kön-* > CrKar. *küy-* ‘to burn (intr.)’, PT **köp* > CrKar. *köp* ‘many’, PT **kiλi* > CrKar. *kiši* ‘man (person)’, PT **kičük* > CrKar. *kiči* ‘small’, PT **kem* > CrKar. *kim* ‘who’, PT **tük* > CrKar. *tük* ‘feather’, PT **tün* > CrKar. *tün* ‘night’, PT **tüt-ün* > CrKar. *tütün* ‘smoke’.

3.5.3. Other Oghuz loans

The initial consonant of PT **s(i)ač* ‘hair’ yields *č* or *š* in majority of Turkic languages. However, reconstruction of the initial **s* is proven by Yakut *as* (where **s-* > *0-* regularly) and Oghuz reflexes with retained *s-*. Crimean Karaim demonstrates doublets with Oghuz- and Kipchak-like reflexes. The first one should be a loan, since the Halich and Trakai dialects point to Proto-Karaim **č*. This case belongs to fake phonological archaisms.

CrKar. *sač* (Par 107 v. 13) ‘a hair (Rus. *volos* – hair[SG])’ ~ *seč* (Sz) ‘hair (Rus. *volosy* – hair-PL); tuft, crest’ ~ *čač* (R) ‘a hair, hair’, TrKar. *čač* ‘a hair, hair’, HKar. *cac* ‘hair, fur’ (Baskakov, Szapszał & Zajączkowski 1974: 470, 500, 613, 625).

Two more Oghuz loanwords in Crimean Karaim are *ver-* ‘to give’ and *var-* ‘to go’, which reflect an Oghuz innovation. Although these words are not fake archaisms, I include them here since they additionally confirm the direction of borrowings in the pair Crimean Karaim < Coastal Crimean Tatar / Turkish. These stems reflect the shift of initial **b* to *v* in monosyllabic stems with *r* in the coda.

CrKar. *ber-* (Sz, R) ~ *ver-* (Par 77 v. 19, ZR 32, 27) ‘to give’, cf. TrKar. *b^her-* ‘to give’, HKar. *ber-* ‘to give’ (Baskakov, Szapszał & Zajączkowski 1974: 112, 151, 158);

CrKar. *var-* (Q 4) ~ *bar-* (Sz) ‘to go’, TrKar., HKar. *bar-* ‘to go’ (Baskakov, Szapszał & Zajączkowski 1974: 102; Aqtay & Jankowski 2015: 436).

3.6. Preliminary conclusions

The onomasiological reconstruction of the Proto-Karaim Swadesh list is complicated mainly by the set of fake archaisms. Fake archaisms are a particular type of homoplastic development, namely MAT-borrowings and semantic back-formations from a sister subgroup which preserved more archaic (in the perspective of a whole family) items. In the Crimean Karaim case, Oghuz nature of the archaic-looking items is proven by the large amount of other Oghuz borrowings and by the history of the sociolinguistic situation.

Examination of sources used in Baskakov, Szapszał & Zajączkowski 1974 and in Aqtay & Jankowski 2015 shows that some of them are more “Oghuzized” than others. From our data it is clear that Par, ZR, and Q contain many more Oghuz forms than Sz. Apparently, they demonstrate language shift to Coastal Crimea Tatar. Aqtay and Jankowski’s more detailed study of the lexicon (2015: 9) confirms this statement. Data from these sources are inappropriate for phylogenetic studies.

Consistent detection of all borrowed elements allows mostly trivial reconstruction of the Swadesh list for the not particularly deep Proto-Karaim taxon.

4. Classifying languages of Crimea

In this section, I address the discussion of the genealogical affiliation of Crimean Karaim, each of three Crimean Tatar dialects and Krymchak. My goal within the scope of this section is not to build a complete phylogenetic tree but only to define the closest relatives of the idioms in question. Needless to say, disclosure of borrowings plays a crucial role in this procedure. Before comparing wordlists all loans, including inter-Turkic ones, must be excluded. Although this statement may seem trivial, in Section 4.6 it will be shown that even recent phylogenetic research still continues to be affected by undetected borrowings.

4.1. Previous research

Beginning with Radloff, the language of Crimean Karaims is fully identified with Crimean Tatar or seen as one of its dialects. This opinion is shared by Zajączkowski, Doerfer, and Polinsky. At the same time, Crimean Karaim (whatever the term means) is included in the Karaim-Russian-Polish dictionary (Baskakov, Szapszał & Zajączkowski 1974). In his earlier grammar (Musaev 1964: 36–37), Musaev maintains that the Crimean Karait's variety is not distinguishable from Crimean Tatar and must not be included in the notion *Karaim language*. Information on the Crimean dialect was later included in his sketch of Karaim dialectology (Musaev 2010) by the editors. The discussion is summarized in the work by Jankowski (2003: 109–112), who attempts to show that Crimean Karaim is different from Crimean Tatar, involving phonological, syntactical, onomastic and lexical arguments. Basic vocabulary remains beyond his interest. The Swadesh list was examined in Polinsky 1992. She comes to the conclusion that Crimean Karaim together with Krymchak language is very close to the Orta and Coastal Crimean Tatar dialects and, hence, belongs to the Oghuz subgroup. Polinsky does not distinguish borrowed and inherited vocabulary when calculating lexicostatistical matches, therefore her conclusions can be called into question. To be fair, it must be noted that, to the best of my knowledge, the requirement to exclude contact innovations was yet to be explicitly formulated in 1992.

Currently, no detailed descriptions of the Crimean Tatar dialects exist and they are unlikely to appear in the future. Commonly accepted is Ervand Sevortyan's (1966) dialectal classification, which distinguishes three dialects of Crimean Tatar (Steppe, Coastal and Central) highlighting their heterogeneous origin. According to Sevortyan's classifications, the Steppe dialect belongs to the Nogai Kipchak subgroup; Orta is Cuman Kipchak, i.e. West Kipchak; Coastal belongs to the Oghuz group. The original dialectal differentiation was violated as the result of Soviet deportation of Crimean Tatars to Uzbekistan in 1944. After the return to Crimea in the early 1990s, most families were not able to settle in their native villages. This provides further dialectal mixture. Already during Darya Kavitskaya's fieldwork in 2002–2003 and 2009, only older speakers had “clear dialect affiliation” (Kavitskaya 2010: 3). Dialectal mixture is quite visible both in my and Polinsky's data. See Normanskaya 2019 on dialectal mixture in literary Crimean Tatar.

4.2. Methods

To define the genealogical affiliation of the Turkic languages of Crimea, I first apply manual subgrouping based on lexical innovation and then compare the obtained results with the inference of the computational lexicostatistical algorithms. I use three approaches which are currently most widespread in linguistic phylogeny: Maximum Parsimony Analysis, MCMC Bayesian approach, Neighbor-joining algorithm.

An important advantage of the manual subgrouping applied in the present paper is that it fits the commonly accepted requirement to build genealogical classification based on innovations (Campbell 2013: 175). This requirement is ignored by the lexicostatistical framework, where every match, whether it is an innovation or a retention, has similar value. The principle of subgrouping sufficient for our purposes is drastically trivial. Languages A and B are regarded as specifically related to each other if this pair demonstrates the highest amount of shared non-contact-induced innovations. This method was used by Leonid Kogan (2015) for the classification of the Semitic languages. When reconstructing the phylogeny of a whole fam-

ily from scratch, this method leads to a vicious circle, since tree topology must be already known for most cases to distinguish between innovations and retentions. However, if the goal is merely to find the positions of newly involved taxa on a previously constructed tree, such a method is applicable.

The most important technical details on the applied computational lexicostatistical algorithms are summarized in Table 2.

Algorithm	Software	Basic settings
Maximum Parsimony	TNT v. 1.5 (Goloboff & Catalano 2016)	Implicit enumeration Collapse trees after search Outgroup: Yakut
Bayesian MCMC	MrBayes 3.2.7a x86_64 (Huelsenbeck & Ronquist 2001; Ronquist et al. 2012)	covariation F81 model; datatype = restriction; coding = noabsencesites; rates = gamma covariation = yes brlenspr = clock:fossilization clockvarpr = TK02
Neighbor-joining	Starling v. 2.7.0-42f0a13 (Starostin 2007a)	Method: Experimental Replacement rate: 4.88 (default value)

Table 2. Information on the software and basic settings applied for the lexicostatistical analyses.

Based on trees obtained as the result of Maximum Parsimony, a strict consensus tree was produced. The settings for Bayesian MCMC are adopted from Kassian et al. forthcoming. The full dataset and output files can be found in Supplement 2. Cognate encoding has been done within Starling software and then converted into the Nexus file with a binary matrix. The derivational drift free dataset has been used; on the principles of the derivational drift elimination see Kassian et al. forthcoming.

I have compared six wordlists of the Turkic idioms spoken until recently on the Crimean Peninsula (see Table 3) with Halich and Trakai Karaim, Turkish, Gagauz, Proto-Nogai, Proto-Kazakh-Karakalpak, Proto-Kumyk, Proto-Karachay-Balkar. Since Maximum Parsimony analysis requires an outgroup taxon, I included the Proto-Yakut list, which clearly belongs neither to the Oghuz nor to the Kipchak clade. Lists of the proto-languages are reconstructed (using methodology described in Section 3.1) by me in collaboration with Anna Dybo and Alexei Kassian within the framework of an ongoing project devoted to revision of the Turkic phylogenetic tree structure. All Karaim wordlists were collected from the sources mentioned in Section 2.1; Turkish and Gagauz ones are based on the dictionary sources as well (Parker 2008; Bogochanskaya & Torgashova 2009; Gaydarzhi et al. 1973; Sesli Sözlük; Rajki 2007). Collecting the lists, I was guided by the semantic specification proposed in Kassian et al. 2010.

Since some of Crimean Tatar wordlists do not meet all the modern requirements to Swadesh lists (see Section 4.3 for details), computational lexicostatistics can only play a secondary role in the present research. However, I believe that, despite somewhat faulty data, application of three different computational approaches still has some relevance.

4.3. Data

Table 3 shows which wordlists of idioms spoken in Crimea were used in the present paper.

Idiom	Comments and references
Crimean Karaim	The material of the list was collected from Baskakov, Szapszał & Zajączkowski 1974 and Aqtay & Jankowski 2015. All Oghuz loans discussed in section 3 have been excluded, so that the list reflects sources with minor Oghuz influence.
Coastal Crimean Tatar	The list published in Polinsky 1992.
Orta Crimean Tatar	The list published in Polinsky 1992.
Steppe Crimean Tatar	The list published in Polinsky 1992.
Crimean Tatar (dialect not defined)	The list was collected in 2020 from two speakers. I avoid labeling it with any dialectal affiliation due to the reasons described below.
Krymchak	The list published in Polinsky 1992, revisited and extended based on Ianbay 2016 and Rebi 2004.
Turkish	The material of the list was taken from Parker 2008; SesliSözlük; Bogochanskaya & Torgashova 2009.
Gagauz	The material of the list was taken from Gaydarzhi et al. 1973; Rajki 2007.
Proto-Kumyk	The list is reconstructed by the author in collaboration with Anna Dybo and Alexei Kassian on the basis of Bammatorov & Gadzhiakhmedov 2011 and recently collected dialectal data.
Proto-Karachay-Balkar	The list is reconstructed by the author in collaboration with Anna Dybo and Alexei Kassian on the basis of Gochiyayeva & Suyunchev 1989 and recently collected dialectal wordlists.
Proto-Nogai	The list is reconstructed by the author in collaboration with Anna Dybo and Alexei Kassian on the basis of Baskakov 1956 and recently collected dialectal wordlists.
Proto-Kazakh-Karakalpak	The list is reconstructed by the author in collaboration with Anna Dybo and Alexei Kassian on the basis of dictionary sources (Bektaev 1996; Bekturov & Bekturova 2001; Syzdykova & Khusaiyn 2008; Baskakov 1967) and recently collected dialectal wordlists.
Proto-Yakut	The list is reconstructed by the author in collaboration with Anna Dybo and Alexei Kassian on the basis of dictionary sources (Pekarskiy 1959; Pekarskiy 1916; Stachowski 1993; Stachowski 1998) and recently collected dialectal wordlists.

Table 3. 110-item Swadesh lists for the Turkic varieties of Crimea.

It should be noted that three Crimean Tatar wordlists used in this research had been collected long before the semantic specification of the Swadesh list was undertaken in Kassian et al. 2010. Therefore, they are not fully compatible with my data. Three important discrepancies have been found; ‘all (omnes)’, ‘to burn (intr.)’, ‘to go’ are used now instead of traditional ‘all (totus)’, ‘to burn (tr.)’, ‘to walk’. Incompatible items are marked as not attested in Polinsky’s data; moreover, the original Swadesh 100-wordlist has been extended with 10 items absent from the older record. To make my lists more compatible with older ones, I deviate from Kassian et al. 2010 on three points: taking ‘earth (ground)’ instead of ‘earth (soil)’, ‘round (2D)’ without ‘round 3D’ as a synonym and not accepting medial deictic pronouns.

I now avoid labeling my data on Crimean Tatar with any terms from the traditional classification. It has been collected from two informants, a couple, about 60 years old, who were interviewed independently. They were born in Uzbekistan, now they live in the Bakhchisaray district. The wife’s parents come from Duvanköy (Ukr. *Verxnjosadove*), the traditional territory of the Orta dialect; the husband’s parents come from the nearest suburbs of Gurzuf, the traditional territory of the Coastal dialect. The two obtained wordlists correspond with each other fully as far as lexical items are concerned; at the same time, they considerably differ in phonology. The woman’s idiolect lacks labial harmony and the distinction of voiced and voiceless stops in the word-onset; the voiceless uvular is a stop. So, it should be regarded as a Kipchak-

based dialect. Thus, she pronounces ‘fat’ with the uvular consonant as *yav*, but ‘forest’ with a labial as *taw*, cf. section 3.4.1. Her husband’s idiolect demonstrates consistent labial harmony, sporadic preservation of the PT initial **d*, and the fricativization of the voiceless uvular; these features are typical for the Coastal (Oghuz-based) dialect. The word *auz* ‘mouth’ has been found to have the typically Kipchak reflex of **agi* in the both idiolects, but *avir* ‘heavy’ with the Oghuz reflex. Polinsky’s lists contain some inconsistencies in phonology as well; the most glaring is a sporadic **y > ʃ* before *a* in the Coastal dialect, which is expected only in Steppe Crimean Tatar, and, at the same time, the lack of this shift in some lexemes from the Steppe dialect. The distribution of reflexes of PT **g* across dialects confounds expectations as well: cf. *aviz* in all dialects, CoCrTat. *tau*, *yav*, OrCrTat *yav*, *dav* (with *d* instead of expected *t!*). The phonology represented in Polinsky’s data is completely inconsistent with existing description of the Crimean Tatar dialects. As is shown below, basic vocabulary allows us to make some clearer conclusions about the original genealogical affiliation of the dialects documented in the wordlists under consideration and about the direction of their development.

4.4. Innovations in the basic vocabulary

Turkic languages of the Crimea demonstrate discrepancies in 25 slots out of the 110-item wordlist. Cases in which the variation is caused by Persian and Arabic borrowings are excluded, i.e. only potentially autonomous innovations are taken into account. Table 4 below presents the genetically relevant features. Archaisms, borrowings, and items innovative from the Common Turkic perspective but still not informative for the current question (i.e. innovations shared by both Kipchak and Oghuz languages) are underlined. The Oghuz loans revealed above and phonological variants have been excluded from the Crimean Karaim wordlist. Indexes in the superscript identify the subgroup in which cognates of the word are found, an exclamation mark labels singletons. These indexes are somewhat rough, since I ignore the fact that some of the considered words can occur sporadically in the other Turkic languages which cannot be applied to the closest relative of the studied Crimean varieties. These inaccuracies are partially clarified in the commentary immediately after the table. I use mostly *Kip* and *Ogh* meaning primarily West Kipchak and West Oghuz respectively; when possible, I refer to the low level subgroups instead of Kipchak. The full lists for the Turkic idioms of Crimea and for languages they have been compared with can be found in Supplement 2.

1. **all** – CrTat., Krym. *epsi*, cf. Tur., Gag. *hepsi* ‘all (omnes)’ – the word is limited to the mentioned languages; no fully acceptable Turkic or foreign etymology (Dybo 2013: 66), the root has a probable Persian origin (cf. Räsänen 1969: 158), but this hypothesis faces some phonological difficulties.

4. **belly** – CrKar., CrTat., Krym. *qursaq* – assuming the stem as a basic term is innovative as a result of the elevation ‘paunch’ > ‘belly’; widespread in the Kipchak subgroup (Dybo 2013: 106–108).

10. **bone** – CrTat., Krym. *kemik*, cf. Tur., Gag. *kemik* ‘bone’ – result of generalization ‘spongy bone’ > ‘bone’; as a basic term, limited to the mentioned languages (Dybo 2013: 173–174).

11. **breast** – Strictly speaking, both stems are archaic: CrTat., Krym. *göküs*, *köküs*, *kokus*, *koks* reflect PT **gökür*; CrKar. *körek*, CrTat. *körek* reflect the same root with the fossilized diminutive suffix **gökrek* (Dybo 2013: 178). Since the derivational connection between these stems has been erased a long time ago, I believe that their distribution is informative for genealogical classification. However, in agreement with the principles of derivational drift elimination, I mark the simplex and diminutive form with one index in the lexicostatistical dataset.

	Crimean Karaim	Crimean Tatar (author's data)	Steppe Crimean Tatar	Middle Crimean Tatar	Coastal Crimean Tatar	Krymchak
			(Polinsky 1992)			
all	<u>bari</u> <u>barča</u>	<i>epsi</i> ^{Ogh}	not attested			<i>epsi</i> ^{Ogh}
belly	<i>qursaq</i> ^{Kip}	<i>qursaq~χursax</i> ^{Kip}	<i>xursax</i> ^{Kip}	<i>qursaq</i> ^{Kip}	<u>qarın</u>	<i>qursaq</i> ^{Kip}
big	<i>biyik</i> ^{Ogh}	<i>balaban</i> ^{Ogh}	<i>bijk</i> ^{Ogh}	<i>buyuk</i> ^{Ogh}	<i>büyük</i> ^{Ogh}	<i>balaban</i> ^{Ogh} <i>buyuk</i> ^{Ogh}
bone	<u>süvek</u>	<i>kemik</i> ^{Ogh}	<i>kemik</i> ^{Ogh}	<u>süyek</u>	<u>süyek</u>	<i>kemik</i> ^{Ogh}
breast	<i>kökrek</i> ^{Kip}	<i>kokrek</i> ^{Kip}	<i>koks</i> ^{Ogh}	<i>kokus</i> ^{Ogh}	<i>göküs</i> ^{Ogh}	<i>kokus</i> ^{Ogh}
to burn	<i>küydür</i> - ^{Kar}	<u>yaq-</u>	not attested			
dog	<u>it</u>	<i>kopek</i> ^{Ogh}	<u>it</u>	<i>kopek</i> ^{Ogh}	<i>köpek</i> ^{Ogh}	<i>kopek</i> ^{Ogh}
dry	<u>quru</u>	<u>qurı</u>	<i>χati</i> ¹	<u>quru</u>	<u>quru</u>	<u>quru</u>
fat	<u>yaŵ</u>	<u>yaŵ</u>	<i>may</i> ^{Nog}	<u>yaŵ</u>	<u>yaŵ</u>	<u>yaŵ</u>
hand	<i>qol</i> ^{Kip}	<i>qol</i> ^{Kip}	<i>qol</i> ^{Kip}	<i>e</i> ^{Ogh}	<i>e</i> ^{Ogh}	<i>qol</i> ^{Kip}
feather	<i>yun</i> ^{Kar}	<u>qanat~χanat</u>	<i>qušin</i> ^{Nog}	<u>quš-qanat</u>	<i>lelek</i> ^{Ogh}	<u>puŵ</u>
man	<i>er</i> <u>erkak</u>	<i>aχay~aqay</i> ¹	<u>erkek</u>	<i>er</i>	<u>marda</u>	<u>erkek</u>
many	<u>köp</u>	<i>çoq</i> ^{Ogh}	<u>köp</u>	<i>çok</i> ^{Ogh}	<i>çok</i> ^{Ogh}	<i>çok</i> ^{Ogh}
mountain	<u>taw</u>	<i>bayır</i> ^{Ogh}	<i>bair</i> ^{Ogh}	<u>daŵ</u>	<u>tau</u>	<u>daŵ</u>
to sleep	<i>yuqla</i> - ^{Kip}	<i>yuqla~yuχla</i> - ^{Kip}	<i>žat</i> - ¹	<i>yuxla</i> - ^{Kip}	<i>yuxla</i> - ^{Kip}	<i>yuxla</i> - ^{Kip}
small	<u>kiči</u>	<i>ifaq</i> ^{Ogh}	<u>kişik</u>	<i>yufaq</i> ^{Ogh}	<i>yufaq</i> ^{Ogh}	<u>kičkene</u>
smoke	<u>tütün</u>		<i>duman</i> ^{Ogh}	<u>tutun</u>	<i>duman</i> ^{Ogh}	<u>tutun</u>
sun	<i>quyaš</i> ^{Kar}	<i>kuneš</i> ^{Ogh}	<i>kuneš</i> ^{Ogh}	<i>küneš</i> ^{Ogh}	<i>güneš</i> ^{Ogh}	<i>küneš</i> ^{Ogh}
to swim	<i>çom</i> - ^{Kar}	<i>yalta</i> - ^{Nog}	<i>žalda</i> - ^{Nog}	<i>yalda</i> - ^{Nog}	<u>yüz-</u>	<i>yalda</i> - ^{Nog}
tree	<u>awač</u>	<i>derek~terek</i> ^{Kip}	<i>derek</i> ^{Kip}	<i>terek</i> ^{Kip}	<u>awač</u>	<i>terek</i> ^{Kip}
		<i>teraq</i> ^{Kip}				
to go	<u>bar-</u>	<i>kit</i> - ^{Ogh}	not attested			<u>bar~var-</u> <i>kit</i> - ^{Ogh}
warm	<i>issi</i> <i>yilli</i>	<i>sižax</i> ^{Ogh}	<i>sižax</i> ^{Ogh}	<u>çilli</u>	<u>yilli</u>	<i>sižax</i> ^{Ogh}
woman	<u>qatın</u>	<i>apay</i> ¹	<i>χisayaxlı</i>	<u>qadın</u>	<u>qadın</u>	<u>qadın</u>
far	<u>yıraŵ</u>	<i>uzaq ~ uzaχ</i>	not attested			<i>uzaχ</i>
near	<u>yuwuŵ</u>	<i>yaqın ~ yaχın</i>	not attested			<i>yaqın</i>

Table 4. Innovations in the basic vocabulary of the Turkic varieties of Crimea.

12. **to burn** – CrKar. *küydür-* – causative from Proto-Turkic **kör-* ‘to burn’; cognates have been found in other Karaim dialects, Karachay-Balkar, Kazan Tatar and in some other languages, mainly in the Middle Asian area (Dybo 2013: 189).

18. **dog** – CrTat., Krym. *köpek*, Tur., Gag. *köpek* ‘dog’ – result of generalization ‘hound’ > ‘dog’; as a basic term, limited to the mentioned languages (Dybo 2013: 189).

20. **dry** – SCrTat. *χati* – result of the semantic shift ‘solid’ > ‘dry’ (Räsänen 1969: 241).

26. **fat** – SCrTat. *may* – result of the semantic shift ‘butter, suet’ > ‘fat’; as a basic term, common in Volga and Middle Asian Kipchak languages and in Nogai (Dybo 2013: 249–250).

28. **feather** – CrKar. *yun*, cf. HKar., TrKar., *yun* ‘feather / down’ – result of the semantic shift ‘fur / down’ > ‘feather / dawn’; with original meaning, common in the Kipchak languages (Dybo 2013: 259–260).

SCrTat. *quš'in* – etymology is somewhat obscure; the stem should be derived from *quš* ‘bird’, *-in* can be an old instrumental affix; exclusive isoglosses with the Nogai subgroup. The etymology proposed here is more probable than the hypothetical Persian loan mentioned in Dybo 2013: 261. Anna Dybo has as of now rejected the etymology involving Pers. *kuč* ‘fish scale’ (personal communication).

CoCrTat. *lelek*, cf. Tur. *yelek* ‘feather’ – derived from PT **ye:l* ‘mane’; attested in all Oghuz languages, including Salar, and in Khalaj (Dybo 2013: 259).

37. **hand** – CrKar., CrTat. *qol* – result of the semantic shift ‘arm’ > ‘hand’; extremely widespread across Nuclear Turkic languages, particularly in all languages belonging to the Kipchak subgroup. Oghuz and some other languages preserve reflexes of the stem **elg* ‘hand’ (Dybo 2013: 300–307).

51. **man** – CrTat. *aqay* ~ *aχay* – fossilized vocative form from **aqa* ‘uncle, older relative’ (Sevortyan et al. 1974: 121). The stem looks rather like a borrowing from baby-talk (cf. Russian baby-talk words *d'ad'a* ‘man’ < ‘uncle’ and *t'ot'a* ‘woman’ < ‘aunt’), so its relevance for the genealogical classification is questionable, cf. ‘woman’.

52. **many** – CrTat., Krym. *čoq*, cf. Tur., Gag. *čoq*, Az. *čoχ* ‘many’ – reconstruction of the original semantics is somewhat difficult but its innovative nature is obvious; as a basic term, attested only in the mentioned languages (Dybo 2013: 371).

55. **mountain** – CrTat. *bayır*, cf. Gag. *bayır* ‘mountain’ – result of the semantic shift ‘hill’ > ‘mountain’; attested as a basic term only in the mentioned languages (Dybo 2013: 380).

76. **to sleep** – CrKar., CrTat., Krym. *yuq-la-* – derived from **uyki* ‘sleep’; common in Kipchak languages (Dybo 2013: 473–474).

77. **small** – CrTat. *yufaq*, *ifaq*, cf. Tur. *ufaq* ‘small, little’ – result of the generalization ‘small, fine (of pebble, crumb, powder etc.)’ > ‘small, little’; attested in most languages with more specific meanings (Sevortyan et al. 1974: 560–561).

76. **smoke** – CrTat. *duman*, cf. Tur., Gag. *duman* ‘smoke’ – result of the semantic shift ‘fog’ > ‘smoke’; this meaning is found only in the mentioned languages (Dybo 2013: 481).

82. **sun** – CrTat., Krym. *kuneš*, *güneš*, *küneš* – formal innovation, root extension is not quite clear, apparently constitutes an analogical rhyme with *quyaš* ‘heat, blazing sun’; with this extension, this root is attested mainly in Oghuz languages but occurs beyond this subgroup as well (Dybo 2013: 488–489). Since no meaning shift has occurred, I mark reflexes **gün* and **güneš* with the same index in the lexicostatistical dataset.

83. **to swim** – CrTat., Krym. *yalta-*, *yalda-*, *žalda-* – derived from *yal* ‘horse mane’ with further semantic shift ‘to swim or to cross a river holding the horse’s mane’ > ‘to swim’; the verb with its original meaning occurs in Kipchak languages; as the basic term for ‘to swim’, only in Nogai (Dybo 2013: 492; Sevortyan et al. 1974: 93–94).

90. **tree** – CrKar., CrTat., Krym. *terek*, *derek* – result of the generalization ‘poplar’ > ‘tree’; attested as the basic term for ‘tree’ in West Kipchak and Sary Yugur (Dybo 2013: 510).

92. **to go** – CrTat., Krym. *kit-*, cf. Tur. Gag. *git-* – result of the semantic shift ‘to go away’ > ‘to go’; attested as the basic term for ‘to go’ in the Oghuz subgroup (Dybo 2013: 515).

93. **warm** – CrTat., Krym. *sižaχ*, cf. Gag. *sižaχ* ‘warm’ – a formal innovation, PT **isig* ‘warm, hot (?)’ extended with a diminutive suffix; found with this suffix only in the mentioned languages (Dybo 2013: 518; Räsänen 1969: 173–174).

99. **woman** – CrTat. *apay* – fossilized vocative form from **apa* ‘elder sister’ (Sevortyan et al. 1974: 159). The stem looks rather like a borrowing from baby-talk (cf. Russian baby-talk words

d'ad'a ‘man’ < ‘uncle’ and *t'ot'a* ‘woman’ < ‘aunt’), so its relevance for genealogical classification is questionable.

SCrTat. *χisayaχli* – the compound **qiz-ayal-ki* ‘girl-woman-nominal suffix’, metathesis of the cluster under the analogical influence of the more productive suffix *-li* (etymology proposed by Anna Dybo, p.c.). Due to the fact that the main component *-ayal-* is an Arabic loan, I mark the whole stem as a borrowing.

101. **far** – CrTat., Krym. *uzaq* ~ *uzaχ* – result of the shift ‘far (adj.)’ > ‘far (adv.)’; found in numerous languages. The isogloss is not particularly informative for affiliation of Crimean Tatar and Krymchak, however, it opposes Karaim to other Turkic varieties of Crimean (Dybo 2013: 534); the same is true for ‘near’.

103. **near** – CrTat., Krym. *yaqin* ~ *yaχin* – found in numerous languages, however, the distribution points to PT **yaguk* (Kar. *yuwuq*) as a more archaic term for ‘near (adv.)’ from the Common Turkic perspective (Dybo 2013: 539).

Table 5 shows the amount of the innovations shared by each of the investigated wordlists with other Turkic subgroups.

	Crimean Karaim	Crimean Tatar (author's data)	Steppe Crimean Tatar	Middle Crimean Tatar	Coastal Crimean Tatar	Krymchak
Oghuz	1	10	7	6	8	10
Kipchak	4	4	2	3	1	3
Nogai	0	1	3	1	0	1
Karaim	3	0	0	0	0	0

Table 5. The amount of the innovations (including both inherited items and inter-Turkic borrowings) shared with other Turkic subgroups.

Four innovations (‘to burn’, ‘feather’, ‘sun’, ‘to swim’) clearly connect Crimean Karaim with the Halich and Trakai dialect and oppose it to other studied languages. Since it demonstrates only one Kipchak-looking stem, Coastal Crimean Tatar must be classified as Oghuz language. The situation with other idioms is not so transparent. Oghuz-looking innovations are predominant in all wordlists, yet at the same time they demonstrate some typically Kipchak lexemes. Two possible interpretations of this situation may be offered: 1) the Oghuz innovations can be identified as inherited ones and Kipchak as substrate loans; 2) the Oghuz lexemes can be regarded as loans from the dominant language of the region and Kipchak innovations as inherited ones. Such ambiguity indicates that there has been a language shift. The range of borrowings can be easily identified in Crimean Tatar (Steppe, Orta and my data) and in Krymchak based on the same distributional criterion which has already been applied to the Crimean Karaim data. All lexemes with the meanings mentioned above are widespread in Oghuz languages but not in Kipchak. Taking into account the prestige status of Crimean Turkish, i.e. Coastal Crimean Tatar, extensive borrowing from it is very probable. Such interpretation is supported by the fact that the lexemes suspected to be loanwords concentrate predominantly in the less stable part of the Swadesh list: ‘small’ – 110; ‘mountain’ – 107; ‘many’ – 106; ‘big’ – 101; ‘far’ – 100; ‘near’ – 95; ‘warm’ – 90; ‘to go’ – 89; ‘all’ – 84; ‘to swim’ – 78; ‘breast’ – 49; ‘smoke’ – 40; ‘sun’ – 39; ‘bone’ – 38; ‘dog’ – 16. The number after the concept indicates its position in the stability ranking (the higher the number, the less stable the concept), according to Sergei Starostin’s estimations (2007b). Avoiding the discussion of every particular case, I would like to draw attention only to OrCrTat. *el* ‘hand’ which is a classic example of a fake archaism.

Other lexemes are more likely to be borrowed West Oghuz innovations. It must be also emphasized that at least in some cases we should deal not with MAT-borrowings (like *epsi* ‘all’; *küneš* ‘sun’; *sizax* ‘warm’; *kopek* ‘dog’ < ‘hound’) but rather with PAT-borrowings, i.e. with contact-induced semantic shifts (*kemik* ‘bone’ < ‘spongy bone’; *bayır* ‘mountain’ < ‘hill’; *ifaq* ‘small, little’ < ‘small, fine (of pebble, crumb, powder etc.)’; *kit-* ‘to go’ < ‘to go away’).

Considering all the Oghuz-looking lexemes as borrowings, hence, irrelevant for genealogical classification, I come to the conclusion that Crimean Karaim does indeed belong to the same subgroup with other Karaim dialects; Steppe Crimean Tatar – to the Nogai subgroup; Crimean Tatar (based on my data), Middle Crimean Tatar (Polinsky’s data) and Krymchak are close to West Kipchak. Any further conclusions about their proximity to a particular subgroup within Kipchak cannot be made with enough certainty.

The proximity of Steppe Crimean Tatar to the Nogai subgroup is proven by three non-trivial innovations (‘fat’, ‘feather’, ‘to swim’) which are not attested in any language from potential candidates for the closest relatives. Two of these innovations (*qušin* ‘feather’ and *yalda-* ‘to swim’) do not occur elsewhere in the Turkic languages. Orta Crimean Tatar, Crimean Tatar (based on my data), and Krymchak *yalda-* must be analyzed as a borrowing from the Steppe dialect. Despite the fact that the Steppe dialect was not a dominant idiom, some influence on its part cannot be excluded. Otherwise, such non-trivial (‘to swim or to cross a river holding horse’s mane’ > ‘to swim’) innovations can be regarded only as a signal of relatedness. This is less probable, since there are no other facts pointing to the specific proximity of Orta Crimean Tatar and of the dialect reflected in my data to the Nogai subgroup.

To sum up, I assume the following subgrouping based on innovations in the basic vocabulary: [Turkish, Gagauz, Coastal Crimean Tatar], [[Nogai, Steppe Crimean Tatar], [Halich Karaim, Trakai Karaim, Crimean Karaim], [Orta Crimean Tatar, Crimean Tatar (my data), Krymchak, Kumyk, Karachay-Balkar].

The alternative approach, i.e. consider Oghuz lexemes as inherited and Kipchak lexemes as borrowings, leads to difficulties. Must we mark as borrowings only lexemes looking similar to Kipchak innovations or must we regard typical Kipchak retentions as borrowings too, i.e. as fake archaisms? This question does not have a satisfactory answer. Had Kipchak innovations been considered borrowings, we would have to deal with a suspiciously archaic Oghuz idiom simultaneously overflowing with Kipchak loans. It should be noted that these fictitious Kipchak loans would have concentrated in the somewhat more stable part of the lists than the real loans considered above, cf. ‘belly’ – 109; ‘fat’ – 81; ‘to sleep’ – 73; ‘breast’ – 49; ‘tree’ – 42; ‘hand’ – 11. If Kipchak retentions such as *süyek* ‘bone’, *it* ‘dog’, *bar-* ‘to go’ etc. had been fake archaisms in the tentative Oghuz language, the mass of borrowings would have strongly contradicted the direction of influence proven by sociolinguistic factors. Thus, it seems reasonable to reject such a decision.

4.5 Results of lexicostatistical analysis

All expectations based on innovations in the basic vocabulary are confirmed by formal computational methods. Trees inferred by three applied algorithms differ only in some details. They all agree on the following points: (a) Coastal Crimean Tatar belongs to the same clade as Turkish and Gagauz; (b) other languages are included in the Kipchak subgroup; (c) Steppe Crimean Tatar is combined with Nogai and Kazakh-Karakalpak; (d) Crimean Karaim forms a clade with other Karaim dialects.

As for the internal structure of the Kipchak clade, the applied analyses are in minor disagreement with each other. Neighbor-joining and Bayesian MCMC suggests a first split into Nogai-Kazakh-Karakalpak-Steppe-Crimean-Tatar and remaining languages, followed by a split

into Karachay-Balkar-Kumyk and Karaim-Orta-Crimean-Tatar-Krymchak. Both algorithms have established that Krymchak is the closest to Crimean Tatar (according to my recently collected data) and the two idioms are related to the Orta dialect. The strict consensus tree build by Maximum Parsimony analysis shows multifurcation of the Kipchak clade into the following taxons: [Halich, Trakai, and Crimean Karaim], [Nogai, Steppe Crimean Tatar, Kazakh-Karakalpak], Orta Crimean Tatar, Crimean Tatar (my data), Krymchak, Kumyk, Karachay-Balkar. Such a structure for the tree fully fits all my assumptions made on the basis of common innovations (Section 4.4).

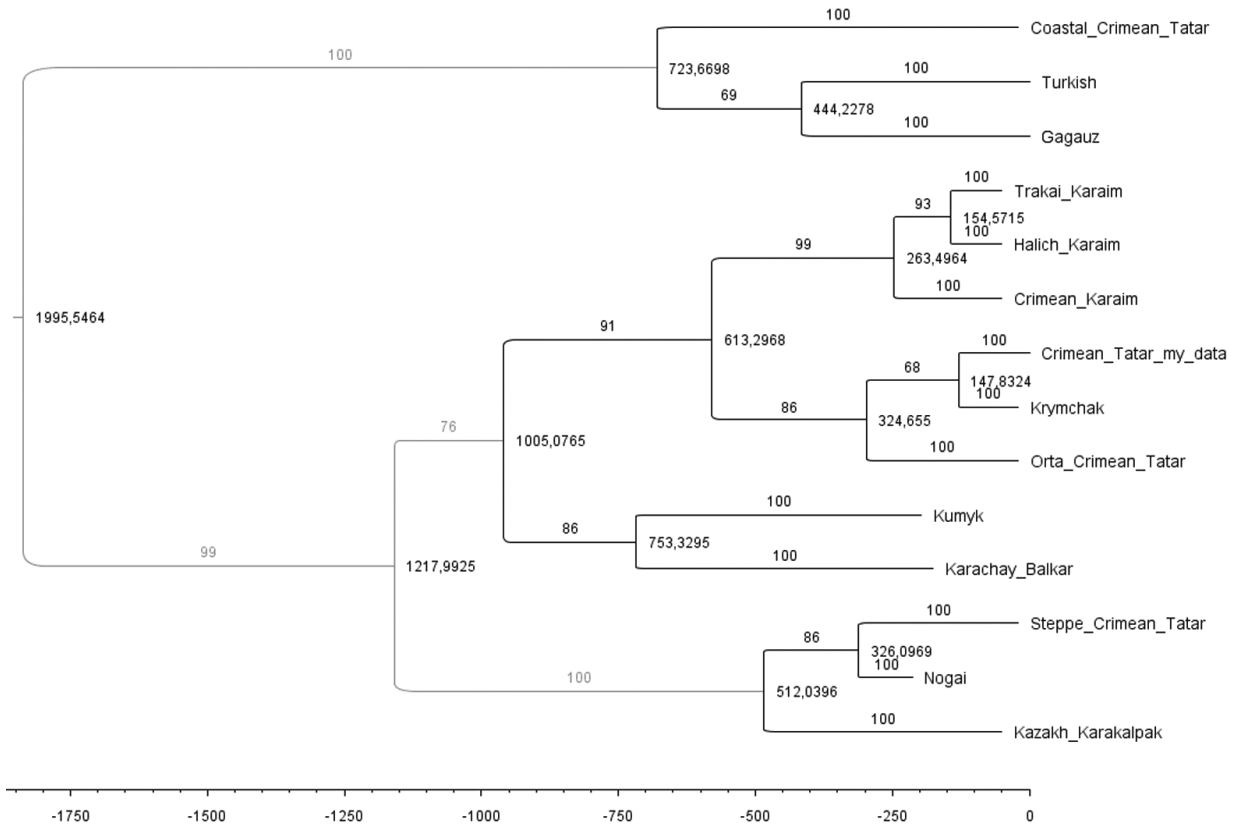


Figure 3. Tree constructed with Bayesian MCMC algorithm in MrBayes software visualized in FigTree software. Numbers near the nodes define mean age; numbers near branches define their probability in percent.

Node	Mean	Median	95% HPD
West Oghuz, i.e. [CoCrTat., [Tur., Gag.]]	723,6698	679,4849	[297,8038, 1238,5883]
[Tur., Gag]	444,2278	415,7284	[173,811, 774,879]
West Kipchak, i.e. [[CrKar., [TrKar., HKar.]], [[OCrTat. [CrTat.(my data), Krym.]], [Kum., KB]]]	1005,0765	959,0648	[581,0544, 1545,5003]
[[OCrTat. [CrTat.(my data), Krym.]], [CrKar., [TrKar., HKar.]]]	613,2968	578,6829	[281,4897, 1026,826]
[CrKar., [TrKar., HKar.]]	263,4964	246,5246	[118,5255, 446,7021]
[TrKar., HKar.]	154,5715	143,146	[66,1245, 269,8982]
[OCrTat. [CrTat.(my data), Krym.]]	324,655	297,9322	[100,4207, 605,93]
[CrTat.(my data), Krym.]	147,8324	129,064	[24,7837, 314,6661]
[Kum., KB]	753,3295	717,8587	[418,5504, 1164,1166]
South Kipchak, i.e. [Kaz-Karak., [Nog., SCrTat.]]	512,0396	484,1897	[242,3641, 844,1445]
[Nog., SCrTat.]	326,0969	311,1897	[149,6245, 521,9043]

Table 6. Mean age, median age and 95% HPD age of the law level nodes according to Bayesian MCMC.

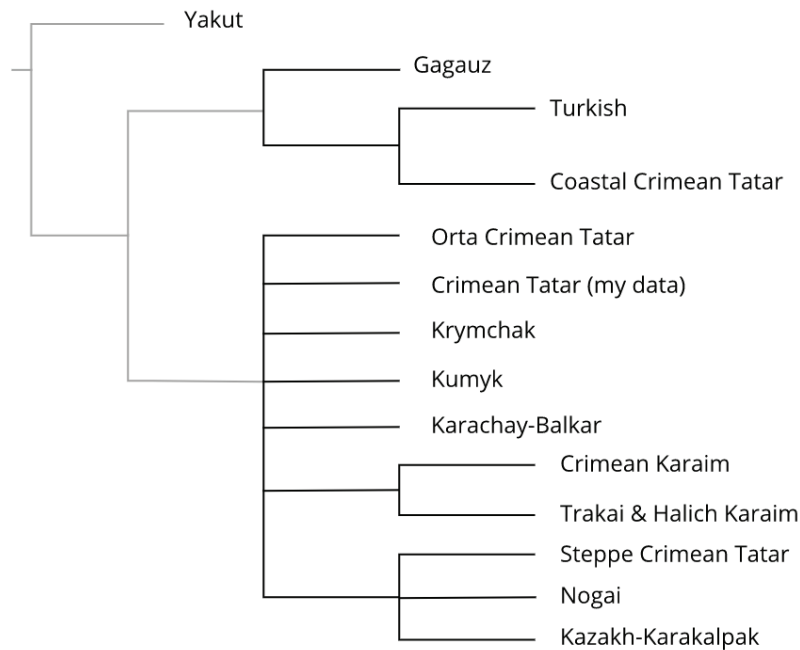


Figure 4. Manually redrawn strict consensus tree constructed with Maximum Parsimony algorithm.

The advantage of Bayesian MCMC and neighbor-joining is that they suggest a clade joining all the languages which had originated on the Crimea Peninsula. The ancestors of the Karaim and Orta Crimean Tatar speaking people came to this region when the West Kipchak languages should have already been (insignificantly) diversified. The invasion of Tatars into Crimea dates back to the 1220s (Fisher 1978: 2); around this time the adherents of Karaite Judaism migrated here from Byzantium (Jankowski 2017: 452–453) and adopted the local Kipchak language. Then the language common for Tatars and Karaims should have split due to the closeness of the Karaim community. The Krymchak speakers, groups of Rabbinic Jews heterogeneous in their origin, adopted Orta Crimean Tatar. Thus, Polinsky is right calling Krymchak an ethnolect of Crimean Tatar (see Polinsky 1992: 173–176). However, another of Polinsky's statements must be revisited. She classifies the Orta dialect together with Krymchak and even Crimean Karaim as Oghuz languages. But even if one admits it is methodologically tolerable not to exclude inter-Turkic borrowings, such an affiliation actually reflects a later language shift. Polinsky's data at least on Krymchak and Orta Crimean Tatars allows the reconstruction of their original genealogical affiliation. The identification of the borrowings plays here a crucial role. My study confirms Sevartyan's (1966) view on the Crimean Tatar dialects as three genealogically distinct items. The early separation of Nogai-Kazakh-Karakalpak-Steppe-Crimean-Tatar from remaining Kipchak languages in question corresponds to the traditional opinion that Nogai does not belong to the West Kipchak subgroup. The speakers of the Steppe dialect have massively settled in Crimea only in the beginning of the 17th century; this was the result of Nogai migration from Lower Volga Steppes which had started a century before (see Trepavlov 2014).

A recent attempt at another revision of the Turkic classification fell victim to undetected loans as well. Martine Robbeets and Alexander Savelyev (2020) include Crimean Tatar into the Oghuz subgroup. They discuss this contradiction with previous classifications and correctly explain it by a lot of Oghuz elements in the wordlist. However, the authors do not try to exclude them despite careful elimination of all non-Turkic borrowings. Robbeets and Savelyev's wordlist of Literary Crimean Tatar (based on the Orta dialect) is compiled on the basis of Useinov 2007. 13 lexemes from this list can be treated as Oghuz loans and contact-influenced semantic innovations: *el* 'hand', *qarın* 'belly', *kemik* 'bone', *koküs* 'breast', *köpek* 'dog', *uzaq* 'far',

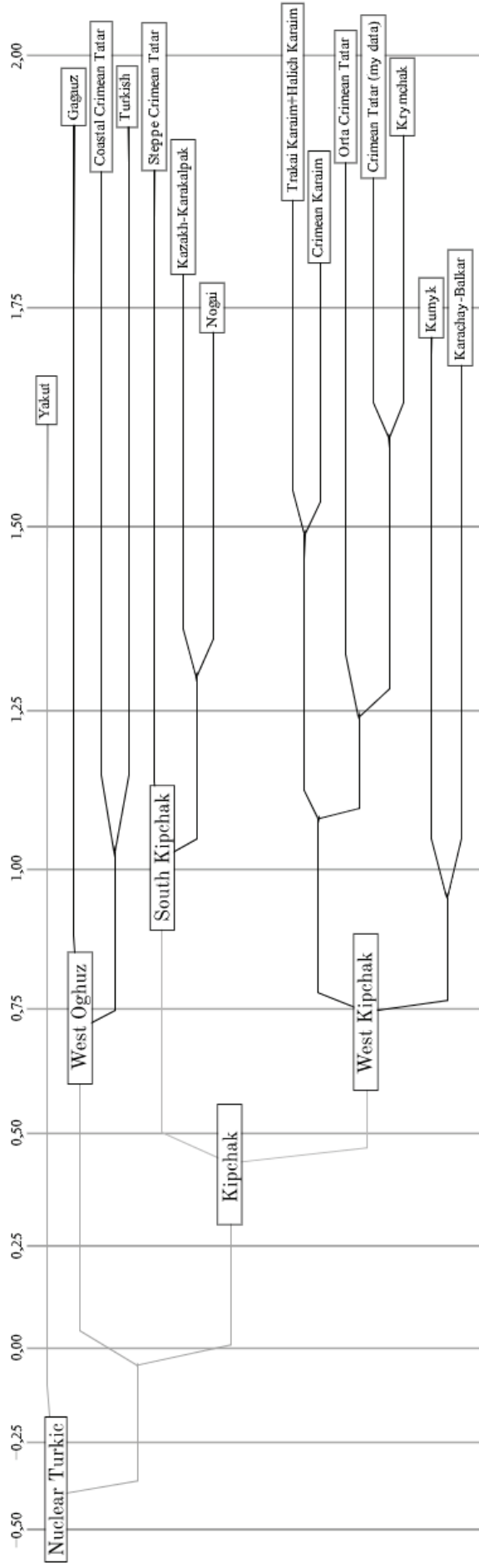


Figure 5. Tree constructed with neighbor-joining algorithm in Starling software.

ket- ‘to go’, *eyi* ‘good’, *čoq* ‘many’, *ufačiq* ‘small’, *küneš* ‘sun’, *sižaq* ‘warm’, *sivoalčan* ‘worm’. At the same time, the list contains the following lexemes typical for the Kipchak subgroup: *qursaq* ‘belly’, *aša-* ‘to eat’, *ayt-* ‘to say’, *yuqla-* ‘to sleep’, *terek* ‘tree’. My remark made on Polinsky’s results is true for this study as well. If an effort is made to exclude 13 items from the dataset, we should be able to identify the original genealogical affiliation of the Orta Crimean Tatar language. That a language should contain so many borrowings from one source in its basic vocabulary is not a frequent case, but hardly a unique exception either (the abovementioned case of Riksmål is another example of a language with a similar amount of borrowings in the basic vocabulary).

5. Conclusions

Detection of borrowings is a necessary procedure for the purposes of phylogenetic studies. A historical linguist should be attentive not only to external but also to intra-family loans. When closely related languages are in intimate contact with each other, the areal criterion becomes more important than the phonological one. The similar phonotactics, phonological inventories, and the minor differences in the phonological shape of words driven by historical phonological processes often make application of phonological criteria difficult or even completely impossible. Loans from a genetically related language can sometimes look like archaic items. They can be both MAT- and PAT-borrowings. These fake archaisms can be revealed based on the distributional criterion if the general direction of influence is known.

Borrowings from a closely related dominant language can strongly influence the basic vocabulary of less prestigious idioms and make the genealogical classification of the latter quite difficult. However, careful elimination of all borrowings makes it possible to identify the subgroup to which the language in question belongs. Both manual analysis of isoglosses and computational lexicostatistics give acceptable results if the dataset is free from borrowings.

Acknowledgments

I thank Anna Dybo and Alexei Kassian for their consultations at all stages of preparation of the present paper. I also acknowledge Maria Polinsky, Christopher Straughn, Mark Zimin, Eldar Idrisov, and Svetlana Egorova for making available to me the necessary materials and for their help in data collection. Last but not least, I would like to thank my friends Tatiana Kliuchnikova and Nikolay Kim for our informal discussions which helped me make much clearer the general idea and the actual text of the paper. All mistakes are, of course, mine.

All supplementary materials mentioned in the paper are archived online at: www.jolr.ru/jlr18/egorov.zip.

Abbreviations of sources used in Baskakov, Szapszał & Zajączkowski 1974 and Aqtay & Jankowski 2015

Cam – Cambridge manuscript of the whole Bible except Chronicles, vol i-iv, Cambridge University Library, classmark BSMS 288.

Fil – Filonenko, V. I. 1929. Atalar sozy: karaim idioms. Proceedings of the Tauride society of history, archeology and ethnography. Vol. 1. Simferopol.

Man – Manchester manuscript of some portions of the Bible. The Rylans Library, classmark Gaster H 170.

Meq – Meqabbeç, a prayer book printed in 1734.

R – Radloff 1896.

Par. – Karaim manuscript from French National Library, signature Hebr. 666.

Sz – Card files collected for Szapszał's dictionary of Crimean Karaim, manuscript.

ZR – Zeķer rav, published in 1831 in Istanbul by Joseph Shelomo Lucki, see edition Poznański 1913.

Q – Qılçı's Mejuma, edition Aqtay 2009.

Abbreviations for names of languages and dialects

Az. – Azerbaijani	Kaz. – Kazakh	SaaN. – Northern Saami
Bash. – Bashkir	Khant. – Khanty	Sal. – Salar
Chag. – Chagatai	Kip. – Kipchak	SCrTat. – Steppe Crimean Tatar
Chuv. – Chuvash	Kirg. – Kirgiz	Tat. – Tatar
CoCrTat. – Coastal Crimean Tatar	Krym. – Krymchak	TrKar. – Trakai Karaim
CrKar. – Crimean Karaim	Kum. – Kumyk	Tur. – Turkish
CrTat. – Crimean Tatar	Nog. – Nogai	Turkm. – Turkmen
dial. – dialectal	Ogh. – Oghuz	Turkm. – Turkmen
Fin. – Finnish	OrCrTat. – Orta Crimean Tatar	Tuv. – Tuvinian
HKar. – Halich Karaim	OT – Old Turkic	Ukr. – Ukrainian
K.-B. – Karachay-Balkar	OUyg. – Old Uyghur	Uyg. – Uyghur
Karak. – Karakalpak	PT – Proto-Turkic	Uzb. – Uzbek
KarakhUyg. – Karakhanid Uyghur	Rus. – Russian	Yak. – Yakut

References

- Aikio, Ante. 2007. Etymological nativization of loanwords: A case study of Saami and Finnish. In: Ida Toivonen, Diane Nelson (eds.). *Saami Linguistics* (Current Issues in Linguistic Theory 288): 17–52. Amsterdam, Philadelphia: John Benjamins Pub. Co.
- Aqtay, Gulayhan. 2009. *Eliyahu Ben Yosef Qılçı's Anthology of Crimean Karaim and Turkish Literature*. Istanbul: Yıldız Dil ve Edebiyat Dizisi 8.
- Aqtay, Gulayhan & Henryk Jankowski. 2015. *A Crimean Karaim-English dictionary* (Prace Karaimoznawcze 2). Poznań.
- Bammatov, Burgan G., Nurmagomed E. Gadzhiakhmedov. 2011. *Kumyksko-russkij slovar'* [Kumyk-Russian dictionary]. Makhachkala.
- Baskakov, Nikolay A. (ed.). 1956. *Russko-Nogajskij slovar'* [Russian-Nogai dictionary]. Moscow: Gosudarstvennoe izdatel'stvo innostrannyx i nacyonal'nyx slovarej.
- Baskakov, Nikolay A. 1967. *Russko-karakalpakskij slovar'* [Russian-karakalpak dictionary]. Moscow: Sovetskaya entsiklopediya.
- Baskakov, Nikolay A., Seraja Ben Mordechaj Szapszał, Ananiasz Zajączkowski. 1974. *Słownik karaimsko-rosyjsko-polski* [Karaim-Polish-Russian dictionary]. Moscow: Russkij Jazyk.
- Bektaev, Kaldybay. 1996. *Bol'šoj kazaxsko-russkij russko-kazaxskij slovar'* [Large Kazakh-Russian Russian-Kazakh dictionary]. Kazaxsij proekt razvitija gosudarstvennogo jazyka.
- Bekturov, Shabken & Ardak Bekturova. 2001. *Kazaxsko-russkij slovar'* [Kazakh-Russian dictionary]. Astana: Foliant.
- Bergsland, Knut, Hans Vogt. 1962. On the Validity of Glottochronology. *Current Anthropology* 3(2): 115–153.
- Bogochanskaya, Nina N., Anna S. Torgashova. 2009. *Bol'šoj russko-tureckij slovar'* [Large Russian-Turkish dictionary]. Moscow: Dom Slavjanskoj Knigi.
- Bouckaert, R., P. Lemey, M. Dunn, S. J. Greenhill, A. V. Alekseyenko, A. J. Drummond, R. D. Gray, M. A. Suchard, Q. D. Atkinson. 2012. Mapping the origins and expansion of the Indo-European language family. *Science* 337: 957–960. doi:10.1126/science.1219669.
- Burlak, Svetlana A., Sergei A. Starostin. 2005. *Sravnitel'no-istoričeskoe jazykoznanie* [Comparative-historical linguistics]. Moscow: Academia.
- Campbell, Lyle. 2013. *Historical linguistics: an introduction*. 3rd edn. Edinburgh University Press.

- Clauson, Gerard. 1972. *An etymological dictionary of Pre-Thirteenth-Century Turkish*. Oxford University Press.
- Dybo, Anna V. 1996. *Semantičeskaja rekonstrukcija v altajskoj étimologii: somatičeskie terminy (plečevoj pojas) [Semantic reconstruction in the Altaic etymology: somatic terms (shoulder girdle)]*. Moscow.
- Dybo, Anna V. 2007. *Lingvističeskie kontakty rannix tjurkov. Leksičeskij fond. Prattjurkskij period [Language contacts of the early Turkic peoples: vocabulary: Proto-Turkic epoch]*. Moscow: Vostochnaya Literatura.
- Dybo, Anna V. 2013. *Étimologičeskij slovar' tjurkskix jazykov [Etymological dictionary of the Turkic languages]*. Vol. 9: *Étimologičeskij slovar' bazisnoj leksiki tjurkskix jazykov [Etymological dictionary of the Turkic basic vocabulary]*. Astana: Prosper Print.
- Étnografičeskaja karta Kryma [Ethnographic map of Crimea]*. 1926. GosKrymIzdat.
- Filonenko, W. J. 1931. *Ethnografische Karte der Krimer autonomen sozialistischen Sowjetrepublik [Ethnographic map of the Crimean Autonomous Soviet Socialist Republic]*. Wien.
- Fisher, Alan W. 1978. *Crimean Tatars (Studies of Nationalities in the USSR)*. Stanford, Calif.: Hoover Institution Press.
- Gaydarzhi, Gavril A., Ludmila A. Pokrovskaya, Boris P. Tukan, Elena K. Koltsa. 1973. *Gagauzsko-russko-moldavskij slovar' [Gagauz-Russian-Moldovan dictionary]*. (Ed.) Nikolay A. Baskakov. Moscow: Sovetskaja entsiklopedija.
- Gochiyayeva, Sofya A., Kh. I. Suyunchev. 1989. *Karačaevo-balkarsko-russkij slovar' [Karachay-Balkar-Russian dictionary]*. Moscow: Russkij Jazyk.
- Goloboff, Pablo A., Santiago A. Catalano. 2016. TNT version 1.5, including a full implementation of phylogenetic morphometrics. *Cladistics* 32(3): 221–238. doi:10.1111/cla.12160.
- Gray, Russell D., Quentin D. Atkinson. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426: 435–439. doi:10.1038/nature02029.
- Haspelmath, Martin. 2009. Lexical borrowing: concepts and issues. In: Uri Tadmor, Martin Haspelmath (eds.). *Loanwords in the World's Languages: A Comparative Handbook*: 35–54. Berlin / New York: Walter de Gruyter. doi:10.1515/9783110218442.
- Huelsenbeck, J. P., F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.
- Ianbay, Iala. 2016. *Krimchak Dictionary*. Wiesbaden: Harrassowitz Verlag.
- Izidinova, Seveli R. 1996. Krymskotatarskij jazyk [Crimean Tatar language]. In: Edkhyam R. Tenishev (ed.). *Tjurkskie jazyki [Turkic languages] (Jazyki Mira [World's Languages])*: 298–309. Moscow.
- Jankowski, Henryk. 2003. On the Language Varieties of Karaims in the Crimea. *Studia Orientalia* 95: 109–130.
- Jankowski, Henryk. 2017. Karaim and Krymchak. In: Lily Kahn, Aaron D. Rubin (eds.). *Handbook of Jewish Languages*: 452–489. Leiden: Brill. doi:10.1163/9789004359543_015.
- Johanson, Lars. 1998. The history of Turkic. In: Lars Johanson, Éva Á. Csató (eds.). *The Turkic languages (Routledge Language Family Descriptions)*: 81–125. London: Routledge.
- Kassian, Alexei S., George Starostin, Anna Dybo, Vasily Chernov. 2010. The Swadesh wordlist. An attempt at semantic specification. *Journal of Language Relationship* 4: 46–89.
- Kassian, Alexei S., George S. Starostin, Mikhail A. Zhivlov. 2015. Proto-Indo-European-Uralic comparison from the probabilistic point of view. *Journal of Indo-European Studies* 43(3–4): 301–347.
- Kassian, Alexei S., Mikhail A. Zhivlov, George S. Starostin, Artem A. Trofimov, Petr A. Kocharov, Anna Kuritsyna, Mikhail N. Saenko. Forthcoming. *Rapid radiation of the Inner Indo-European languages: an advanced approach to Indo-European lexicostatistics*. <https://www.academia.edu/39903804/>.
- Kavitskaya, Darya. 2010. *Crimean Tatar (Languages of the World/Materials 477)*. München: LINCOM EUROPA.
- Kocaoğlu, Timur. 2006. *Karay: the Trakai dialect (Languages of the World/Materials 458)*. München: LINCOM EUROPA.
- Kogan, Leonid E. 2015. *Genealogical classification of Semitic: the lexical isoglosses*. Boston, Berlin: De Gruyter.
- Musaev, Kenesbai M. 1964. *Grammatika karaimskogo jazyka: fonetika i morfologija [Grammar of the Karaim language: phonetics and morphology]*. Moscow: Nauka.
- Musaev, Kenesbai M. 2010. Dialekty karaimskogo jazyka [Dialects of the Karaim language]. In: Anna V. Dybo (ed.). *Dialekty tjurkskix jazykov [Dialects of the Turkic languages]*: 205–235. Moscow: Vostochnaya literatura.
- Nakhleh, Luay, Donald A. Ringe, Tandy Warnow. 2005. Perfect Phylogenetic Networks: A New Methodology for Reconstructing the Evolutionary History of Natural Languages. *Language* 81(2): 382–420. doi:10.1353/lan.2005.0078.
- Parker, Philip M. 2008. *Webster's Turkish-English thesaurus dictionary*. San Diego: ICON Classics.
- Pekarskiy, Eduard K. 1916. *Kratkij russko-jakutskij slovar' [Concise Russian-Yakut dictionary]*. Saint Petersburg: Tipografija imperatorskoj akademii nauk.
- Pekarskiy, Eduard K. 1959. *Slovar' jakutskogo jazyka [Dictionary of the Yakut language]*. Leningrad: Izdatelstvo Akademii Nauk SSSR.

- Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch [Indo-European etymological dictionary]*. 2 vols. Bern / Munich: A. Francke.
- Polinsky, Maria. 1991. The Krymchaks: History and Texts. *Ural-Altai Yearbook* 63: 123–154.
- Polinsky, Maria. 1992. Crimean Tatar and Krymchak: classification and description. In: Howard I. Aronson (ed.). *The Non-Slavic languages of the USSR: linguistic studies*: 157–188. Chicago Linguistic Society, University of Chicago.
- Poznański, Samuel. 1913. Karäisch-tatarische Literatur. *KeletiSzemle* 13: 37–47.
- Radloff, Wasilij. 1896. *Der Volksliteratur der nördlichen türkischen Stämme [The folk literature of the northern Turkish tribes]*. Vol. 7: Die Mundarten der Krym. Saint Petersburg: Tipografija imperatorskoj akademii nauk.
- Rajki, András. 2007. *A concise Gagauz dictionary*. Budapest.
- Räsänen, Martti. 1957. *Materialien zur Morphologie der türkischen Sprachen (Studia Orientalia 21)*. Helsinki: Societas Orientalis Fennica.
- Räsänen, Martti. 1969. *Versuch eines etymologisches Wörterbuch der Türksprachen [Attempt at etymological dictionary of Turkic languages]* (Lexica Societatis Fenno-Ugricae). Helsinki: Suomalais-Ugrilainen Seura.
- Rebi, David. 2004. *Krymčakskij jazyk: krymčaksko-russkij slovar' [Krymchak language: Krymchak-Russian dictionary]*. Simferopol: Dolya.
- Ronquist, Fredrik, Maxim Teslenko, Paul van der Mark, Daniel L. Ayres, Aaron Darling, Sebastian Höhna, Bret Larget, Liang Liu, Marc A. Suchard, John P. Huelsenbeck. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61(3): 539–542. doi:10.1093/sysbio/sys029.
- Rzyski, Christoph, Tiago Tresoldi, Johann-Mattis List, Simon Greenhill, Robert Forkel. 2019. *The Database of Cross-Linguistic Colexifications, reproducible analysis of cross-linguistic polysemies*. <https://clics.clld.org/>.
- Sakel, Jeanette. 2007. Types of loan: matter and pattern. In: Jeanette Sakel & Yaron Matras (eds.). *Grammatical borrowing in cross-linguistic perspective (Empirical Approaches to Language Typology 38)*: 15–29. Berlin, New York: De Gruyter Mouton.
- Savelyev, Alexander, Martine I. Robbeets. 2020. Bayesian phylolinguistics infers the internal structure and the time-depth of the Turkic language family. *Journal of Language Evolution* 5(1): 39–53. doi:10.1093/jole/lzz010.
- SesliSözlük. *SesliSözlük. Online English-Turkish and Multilingual Dictionary*. <https://www.seslisozluk.net/> (29 July, 2020).
- Sevortyan, Ervand V. 1966. Krymsko-tatarskij jazyk [Crimean Tatar language]. In: Viktor V. Vinogradov (ed.). *Jazyki narodov SSSR, vol. 2: Tyurkskie jazyki [Turkic languages]*: 234–259. Moscow: Nauka.
- Sevortyan, Ervand V., Liya S. Levitskaya, Anna V. Dybo, Valentin I. Rassadin, Galina F. Blagova, Dmitriy M. Nasilov. 1974. *Ètimologičeskij slovar' tjurkskix jazykov [Etymological dictionary of the Turkic languages]*. 9 vols. Moscow: Nauka.
- Stachowski, Marek. 1993. *Dolganischer Wortschatz (Zeszyty naukowe Uniwersytetu Jagiellońskiego, Prace językoznawcze 1086. zes. 114)*. Kraków: Nakl. Uniewersytetu Jagiellońskiego.
- Stachowski, Marek. 1998. *Dolganischer Wortschatz: Supplementband*. Kraków: Księgarnia Akademicka.
- Starostin, George S. 2016. From wordlists to proto-wordlists: reconstruction as ‘optimal selection.’ *Faits de langues* 47(1): 177–200. doi:10.3726/432492_177.
- Starostin, Sergei A. 2000. Comparative-historical linguistics and lexicostatistics. In: Colin Renfrew, April McMahon & Larry Trask (eds.). *Time depth in historical linguistics, vol. 1*: 223–265. Cambridge, England: The McDonald Institute for Archaeological Research.
- Starostin, Sergei A. 2007a. Rabočaja sreda dlja lingvista [Linguist’s workspace]. In: Sergei A. Starostin. *Trudy po jazykoznaniju [Works on linguistics]*: 481–496. Moscow: Jazyki slavjanskix kul’tur.
- Starostin, Sergei A. 2007b. Opredelenie ustojčivosti bazisnoj leksiki [Defining the stability of basic lexicon]. In: Sergei A. Starostin. *Trudy po jazykoznaniju [Works on linguistics]*: 827–839. Moscow: Jazyki slavjanskix kul’tur.
- Syzdykova, Rabiga G., K. Sh. Khusaiyn. 2008. *Kazaxsko-russkij slovar' [Kazakh-Russian dictionary]*. Alma-Ata.
- Tenishev, Edkhyam R., Anna V. Dybo (eds.). 2002. *Sravnitel'no-istoričeskaja grammatika tjurkskix jazykov: regional'nye rekonstrukcii [Comparative grammar of Turkic languages: regional reconstructions]*. Moscow: Nauka.
- Tenishev, Edkhyam R., Anna V. Dybo (eds.). 2006. *Sravnitel'no-istoričeskaja grammatika tjurkskix jazykov: pratjurkskij jazyk-osnova. Kartina mira pratjurkskogo ètnosa po dannym jazyka [Comparative grammar of Turkic languages: the Proto-Turkic language. A world-view of the Proto-Turkic people according to language data]*. Moscow: Nauka.
- Trepavlov, Vadim V. 2014. Nogajskaja orda [Nogai Horde]. In: Rafael S. Khakimov (ed.). *Istorija tatar s drevnejšix vremen [History of Tatars from antiquity]*, vol. 4: 223–252. Kazan.
- Useinov, Seyran M. 2007. *Russko-krymskotatarskij, krymskotatarsko-russkij slovar' [Russian-Crimean Tatar, Crimean Tatar-Russian dictionary]*. Simferopol: Tezis.

И. М. Егоров. Базисная лексика близкородственных языков в ситуации языкового контакта: тюркские языки Крымского полуострова

Настоящая статья объединяет два разыскания в области базисной лексики тюркских языков Крымского полуострова. Ее цель — заострить внимание на проблемах, с которыми сталкиваются лингвисты при диахроническом — и в особенности филогенетическом — анализе интенсивно контактирующих друг с другом близкородственных языков. Первое исследование посвящено ономаσιологической реконструкции пракараимского списка Сводеша. Основная рассматриваемая здесь проблема — выявление западно-огузских заимствований и, в первую очередь, контактно обусловленных архаизмов (*fake archaisms*) в крымском диалекте караимского. Задача второго исследования — определение генеалогической принадлежности крымскотатарских диалектов. Ручной анализ инноваций в базисной лексике и алгоритмы вычислительной филогенетики (байесовский метод, метод ближайших соседей, метод максимальной бережливости) подтверждают традиционное мнение о том, что береговой диалект принадлежит к огузской группе, средний — к западно-кыпчакской, а восточный — к ногайско-кыпчакской группе. Такой результат полностью подтверждается данными по этнической истории. Установить правильную генеалогическую аффилиацию рассматриваемых диалектов удалось только после выявления всех заимствований, чего не делалось в предыдущих лексикостатистических исследованиях по крымско-татарскому языку. Оба изученных кейса показывают, что элиминация ареальных влияний принципиально важна и для семантической (ономаσιологической) реконструкции, и для филогенетических исследований.

Ключевые слова: филогенетика; семантическая реконструкция; заимствования; караимский язык; крымскотатарский язык; тюркские языки.

Proto-Afrasian names of ungulates in light of the Proto-Afrasian homeland issue¹

In this paper, we present the first section of a comprehensive thesaurus of Proto-Afrasian zoonyms, compiled and reconstructed by the authors. The list contains more than a hundred names of ungulates, including bovids, pigs, equids and camels, as well as large herbivores (elephants, hippopotamuses and rhinoceroses), allowing for a wholesome (at the current level of our knowledge) picture of these species of fauna in the period preceding the split of Proto-Afrasian (ca. 11th millennium BCE, according to glottochronology) within the area presumably inhabited by speakers of Proto-Afrasian. The results of the reconstruction may be interpreted in favor of two different points of view on the localization of the Proto-Afrasian homeland, namely, the Levant vs. East Africa; the paper presents arguments both in favor of and against each of these hypotheses.

Keywords: Proto-Afrasian homeland; Afrasian (Afro-Asiatic) languages; etymology of zoonyms; reconstruction of cultural lexicon.

Until recently, debates about the primary habitat of the speakers of Proto-Afrasian (Afroasiatic, Semito-Hamitic) have been based either upon general and extra-linguistic argumentation, sometimes accompanied with scarce, incomplete and unrepresentative lexical comparisons; or upon the reconstruction of Proto-Afrasian (PAA) terms interpreted as arguments in favor of intensive gathering of wild cereals and beans (Diakonoff 1981) or in favor of proto- or incipient agriculture (Militarev 2002).

It goes without saying that the *Urheimat* issue cannot be seriously discussed without a complete and detailed reconstruction of PAA animal and plant names, as well as general environmental and cultural vocabularies. Likewise, of vital importance here are the approximate dates of branching of PAA and its descendant proto-languages, obtained by glottochronology. A trustworthy solution of the problem can be achieved only through correlation of comparative linguistic data with the findings and outcomes of such related sciences as archaeology, paleoclimatology, paleontology, and human population genetics.

The genealogical classification of the Afrasian macrofamily as worked out by Alexander Militarev, based first and foremost on lexicostatistics, is as follows (the figures below refer to the approximate BCE dates of branching of proto-languages, obtained glottochronologically by applying Sergei Starostin's recalibrated method²):

¹ Alexander Militarev is responsible for all lexical material, reconstruction of zoonyms, and discussion of the hypothetical Levant homeland of Proto-Afrasians and their identification with Natufians. Sergei Nikolaev is responsible for the extra-linguistic evaluation of PAA zoonyms, their paleozoological interpretation, and discussion of the hypothetical African homeland of Proto-Afrasian. We would also like to extend our gratitude to Anastasia Kirillova for her help in compiling the lists of references, the alphabetic index of reconstructed roots, and the list of abbreviations.

Militarev's research is supported by The Russian Science Foundation (Project № 20-18-00159); the financing organization is The Institute of Linguistics, Russian Academy of Sciences.

² As described in, e.g., Starostin 2000.

Proto-Afrasian – ca. 10,500:

1. Proto-Semitic-Egyptian-Berber-Chadic (PSEBC) – ca. 9,000.
 - 1.1. Proto-Semitic – ca. 4,500.
 - 1.1.1. Proto-South Semitic (Proto-MSA).
 - 1.1.2. Proto-North Semitic – ca. 3,700.
 - 1.1.2.1. Akkadian.
 - 1.1.2.2. Proto-Central Semitic (Ethiopian-Arabic-Levantine) – ca. 3,000.
 - 1.2. Proto-Egyptian-Berber-Chadic – ca. 7,800.
 - 1.2.1. Egyptian.
 - 1.2.2. Proto-Berber-Chadic – ca. 6,500.
 - 1.2.2.1. Proto-Berber – ca. 1,000.
 - 1.2.2.2. Proto-Chadic – ca. 5,500.
2. Proto-Cushitic-OmotiC (PCO) – ca. 8,800.
 - 2.1. Proto-Cushitic – ca. 7,500.
 - 2.1.1. Proto-North-Central Cushitic – ca. 6,800.
 - 2.1.1.2. Proto-North Cushitic (Beja).
 - 2.1.1.3. Proto-Central Cushitic (Agaw) – ca. 2,000.
 - 2.1.2. Proto-East Cushitic (incl. Yaaku-Mogogodo) – ca. 6,000.
 - 2.1.3. Proto-South Cushitic (incl. Dahalo and Ma'a) – ca. 4,800.
 - 2.2. Proto-OmotiC – ca. 6,000.
 - 2.2.1. Proto-North OmotiC (incl. Dizi and Mao) – ca. 4,200.
 - 2.2.2. Proto-South OmotiC (incl. Ongota) – ca. 4,600.

Until now, any systematic paleozoological argumentation relying on a complete set of PAA animal names has been conspicuously lacking. Today, however, over a hundred names for animals reconstructed by Militarev (who has also drawn from the most reliable comparanda and reconstructions by Olga Stolbova, Aharon Dolgopolsky, Václav Blažek and Gabor Takács, as well as, naturally, those included in SED II) represent terms that can be labeled Proto-Afrasian and refer to species which may be relevant for addressing the problem of the PAA habitat from the point of view of both biocenosis and food acquisition (hunting, fishing, and domestication of animals).

The animal section which is by far the largest (some fifty items), the most economically vital for the Epipaleolithic man of Southwest Asian and North African areas, and one of the most essential for dealing with the issue of the original homeland, consists of PAA names of ungulates (bovids, camels, equids), conventionally labeled as “largest herbivores”.

Below we present the complete list³ of reconstructed forms⁴ on various levels, together with their reflexes in individual languages.

1. Bovids and pigs

1.1a. *ʔar(a)w/y-‘k. of larger bovid’

³ Regrettably, we did not have time to include data from the forthcoming Issue VI (Labials) of the Chadic Lexical Database by Olga Stolbova.

⁴ It must be stressed that, since historical vocalism in Afrasian (and even Semitic) comparative linguistics is still a weak point, our reconstruction of specific vowels in roots and stems on all levels is in many cases tentative and sometimes rather arbitrary.

Sem. **ʔi/arw-/*ʔawr-* ‘calf, bull’: Akk. (OAkk. on) *arwû*, *armû*, etc. ‘gazelle’ [*Gazella sp.*]⁵ || Syr. *ʔarw-ān-ā* ‘vitulus, vitula’ | Gur. **ʔar-* ‘cows’: Ez., Gt. *āray*, etc.⁶

(?) Egyp. (MK) *ir.t* ‘calf’, (Dem.) *iry.t* ‘milking cow’ (if < **ʔVr-* not **ʔVl-*).

Chad. **ʔarw-* ‘k. of bovid’ [*Tragelaphus?*]: C. **ʔarw-*: Baldiu *érèw-è* ‘antelope sp.’ || E. **ʔar-*: Migama *ʔār-ó* ‘duiker’ [*Tragelaphus sp.*].

Cush. **ʔaraw-* ‘male of large ungulates’: N. **-rayVw-* (**ʔ > 0* in Beja): Beja *o-reo* ‘bull, steer’ (*o-* is a masc. article) || E. **ʔawar-* (met.) ‘bull’: SA: Saho *awr*, Afar *awur* ‘bull’ | LEC: Somali *awr* ‘he-camel’, Rendille *or* ‘he-camel, bull’, Boni *oor* ‘male elephant’ [*L. africana*], Oromo *oor-oo* ‘burden camel’, Bayso *aar*, Arbore *ʔaar*, Dasenech *ar* ‘bull’ | HEC **ʔaray-*: Burji *array*, *arʔáy* ‘bull’, *aʔre* ‘calf’ | Yaaku *erer* ‘antelope sp.’ || S. **ʔarw-*: Ma’a *áro* ‘large herbivore’ | Dahalo *ʔáároole* (with *-l* root extension?) ‘eland’ [*Taurotragus sp.*].⁷

► Cf. SED II No.16; Blažek 2003b: 236, 269; Mil. 2009; AADB 2476.

1.1b. **ʔar(a)w/y-* ‘k. of lesser bovid’

Sem. **ʔarwiy-* ‘ibex’ [*Capra ibex*]: Sab. *ʔrwy-n* pl. ‘(female) mountain goat, ibex (?)’ [*Capra ibex*] | Arab. *ʔurw-iy-at-* ‘chèvre de montagne’ [*C. ibex*] || MSA: Mhr. *ʔar-īt* ‘goat’, Jib. *ʔarū* ‘she-goat’ (Nakano 1986), Mhr. *ḥā-ráw-n*, Hrs. *ḥe-werū-n* (met.), Jib. *ʔerú-n* ‘goats’, Soq. *ʔére-hon* ‘moutons, chèvres’.⁸

Berb.-Can.: Tenerife *ara* ‘goat’.

Chad. **ʔawr-* (met.) ‘small cattle’: W. **(ʔa)war-*: Hausa *ara-ara* (*-w-* lost in a reduplicated form?) ‘type of long-legged ram, goat, sheep’, Montol, Ankwe, Gerka *ur*, Kulere *wàr* ‘he-goat’ || E. **ʔawr-*: Lele *ōr-ē* ‘goats’.

Cush. **ʔaray-* ‘k. of lesser bovid’:⁹ E. **ʔaray-*: SA **ʔVray-*: Afar *eráy-taa* ‘goat’ | LEC **ʔar-*: Somali *ari, eri* ‘pecore e capre’, Bayso *arer* ‘antelope sp.’ | HEC **ʔaray-*: Burji *aráy* ‘sheep’ | Yaaku *erer* ‘antelope sp.’ || S. **ʔary-*: Iraqw *ari* ‘goat’, pl. *ar-a*, Alagwa, Burunge *ar-a* (pl.) ‘goats’ | Qwadza *ali-to* (*-l* < **-r-*) ‘goat’.

► Cf. SED II No. 18; AADB 2447.

1.2. **ʔayl-* ‘k. of lesser bovid’

Sem. **ʔayil-* ‘ibex [*Capra ibex*]’: (?) Akk. *al-u* (*el-u*) ‘a fine breed of sheep’ (reading arguable) || Ugr. *ʔal, ʔil* ‘ram of superior quality’ | Hbr. *ʔayil* ‘ram’ | ESA: Sab. *ʔyl* ‘mountain goat, ibex’ [*C. ibex*], Min. *ʔyl* ‘bélier, bouquetin’ [*C. ibex*] | Arab. *ʔiyyal-, ʔuyyal-* ‘bouquetin, bouc de montagnes’ [*C. ibex*].¹⁰

⁵ Presumable scientific zoological meanings are quoted in square brackets.

⁶ Also Tgr. *ʔarw-ät* (with f. suff. *-ät*) ‘female elephant’ [*Loxodonta africana*] quoted in SED II No. 17 **ʔarway-* ‘wild beast; lion’. For discussion on Gz. *ʔarwe* ‘animal, wild animal, beast’ with debatable etymological connections, see SED II: 27 (No. 18); on Amh. *awra* ‘male (animal), bull’ and its probable areal connections with E. Cush. **ʔaw(a)r-* ‘bull’ with uncertain direction of borrowing, see SED II: 23 (No. 16). Cf. also Tgr. *wuhər* ‘bull’.

⁷ Omot. N.: Oyda *arr* ‘cow’ is isolated, likely from E. Cush.

⁸ Cf. Sem. **ʔarn-* (likely related with a fossilized *-n*) ‘k. of antelope or mountain goat’: (?) Akk. (MB on) *arm-u* ‘buck of gazelle or mountain goat’ (**-n-* > *-m-* influenced by *armû?*); Aram.: Off. *ʔrn* ‘ram (?)’, Syr. *ʔarn-ā* ‘caper montanus’ [*C. ibex*]; Arab. *ʔirān-* ‘oryx mâle’ [*Oryx sp.*] (otherwise < **ʔaraw-*); Gz. *ʔorn-ā* ‘kind of antelope’, Gur.: Gog., Muh., *orrāññ-ā* ‘ram, small male sheep’. Probably related with Proto-Berb. **a-nHir* (< **-nVʔir*, met. < **ʔVrin* < **ʔVry-Vn* also with a fossilized *-n?*): Semlal *anīr* ‘antelope’, Zenaga *e-naʔrh* ‘gazelle-Dama’ [*Nangerdama*], Ghat *inir*, Ahaggar *enir*, Tadghaq *ti-nhir-t* ‘antelope mohor’ [*N. dama*].

⁹ Omot. N.: Koyra *aar-e* ‘flock of sheep’, Male *war-i* ‘goat’ are isolated terms, likely lws. from E. Cush.

¹⁰ Tna. *ʔil-ä* ‘kind of ram (from the Western Lowlands)’ isolated in Eth. is probably a Beja lw.; Jib. *ayyól* ‘Steinbock’ [*Capra ibex*] (Bittner 1917: 29, not in JJ) is probably an Arabism. Cf. also Proto-Sem. **ʔayyal-* ‘*Cervus elaphus*’ (SED II No. 25), similar to PIE **h₁el-* ‘deer’: Akk. *ayal-u* ‘stag, deer’, Ugr. *ʔayl* ‘deer’, Phoen. *ʔyl* ‘stag’ [*Cervus*].

Berb. **ti-halay* (< **-ʔayal*, met.?) ‘sheep’: Nefusa *t-ili* | Zenaga *t-iʒi* | Ahaggar *té-helé*, Ghat *či-hali*, Ayr *t-ele*, E. Wlm. *te-hale*, etc.¹¹

Cush. **ʔayl-* ‘k. of antelope (*Litocranius*, *Tragelaphus*, *Alcelaphus*): N. **ʔall-i*: Beja *all-i*, pl. *ill-i* ‘long-haired sheep’ || E.: SA **ʔill-*: Saho *ill-e*, Afar *ill-i* ‘small cattle’ | LEC **ʔel-* (< **ʔil-* or **ʔayl-*): Somali *eel-o* ‘tipo de gazella (antilopegiraffa)’ [*Litocranius walleri*], Arbore *ʔell-ém*, Elmolo *ʔél-em* ‘ram’ || S. **ʔayl-*: Gorowa *ele-te-mo* ‘bushbuck’ [*Tragelaphus sylvaticus*] | Ma’a *ʔal-é* ‘ram’, *ʔal-ú* (met. < **ʔayl-*?) ‘sheep’ | Dahalo *ʔèl-e* ‘hartebeest’ [*Alcelaphus buselaphus*].

▲ Egypt. (NE) *iyʔ* (if < **ʔVyl-*) ‘Widder’ (presumably > Copt.: Sahidic *oeile*, Bohairic *wili* ‘bélier’) is most probably a lw. from Hbr. *ʔayil* ‘ram’.

► Cf. SED II No. 24; Mil. 2009; AADB 2466.

1.3. **ʕafVw-* ‘k. of lesser bovid’

Sem. **ʕayp(V)w-* ‘foal, kid’: Arab. *ʕafw-*, *ʕifw-*, *ʕufw-* ‘ânon’, *ʕifw-at-*, *ʕufw-at-* ‘petit d’âne femelle’ || Soq. *ʕeyfif* (part. redupl.) ‘kid’.

Chad. **HafVw-* ‘goat’: W. **HVf-Vm-*: Ankwe *fùm* ‘he-goat’, Tangale *amf-o* (met.?) ‘ram’, Tsagu *úf-è* ‘goat’ || C. **HafVw-*: Hwona *wufi-rā* ‘she-goat’, Logone *ūfu*, Kotoko *hufu*, Mesme *afou* ‘goat’.

Cush. S. **ʕafaw-* ‘goat’: Qwadza *ufw-anko* ‘goats’ | Ma’a *afá* ‘goat’.

▲ The Chad. and S. Cush. forms testify to AA **f-*; Sem. **p* (> Arab. and MSA *f*) may originate from both AA **p* and **f*.¹² Cf. Canar.: Gomera *tufa* (< *t-ufa?*) ‘oveja’.

► Cf. SED II No. 49; Mil. 2009; AADB 2472.

1.4a. **ʕi(n)ʒ-* and **ʔi(n)ʒ-* ‘k. of smaller bovid’

Sem. **ʕa/inz-* and **ʔVʒ-* ‘she-goat’: Akk. (OB on) *enz-u* (*ezz-u*, *inz-u*, *azz-atu*) ‘she-goat, goat (gen. term)’ || Ugr., Phoen. *ʕz* | Hbr. *ʕēz* | Aram.: Off. *ʕnz*, *ʕz* ‘goat’, Bibl. *ʕizz-in* (pl.) ‘id.’, Syr. *ʕezz-* ‘capra’, etc. | Sab. *ʕnz* ‘goats’ (coll.) | Arab. *ʕanz(-at)-* ‘chèvre’ || MSA **ʔōz*: Mhr. *wōz*, Jib. *ʔoz* ‘she-goat’, Soq. *ʔoz* ‘chèvre, mouton’.

(?) Chad. W. **Haz-an-*: Bokkos *ʔazan* ‘ram’.

Cush. **ʕiʒ-* and **ʔiʒ-* ‘k. of lesser bovid (*Aepyceros melampus*, etc.)’: E. **ʕiʒ-* and **ʔiʒ-*: SA **ʕid* and **ʔid-*: Saho *ʕid-oo* (acc. to Conti Rossini 1913), *ayd-ó* (acc. to Reinisch 1890), Afar *ayd-á* ‘sheep’ | LEC **ʔiʒ-*: Somali *id-o*, Rendille *íy(-e)* ‘flock of sheep’, Arbore *ʔizz-e* ‘gazelle’, Elmolo *ʔéd-i* ‘goat’ || (?) S. **ʔand-*: Asa *ʔand-o* ‘impala; female impala’ [*Aepyceros melampus*].¹³

▲ Cf. Sum. *ùz* ‘goat, she-goat’ (EDS #2887) and *é.ze* ‘sheep’ (EDS #723).

► Cf. SED II No. 35; AADB 2471.

1.4b. **ma-ʕi(n)ʒ-* ‘k. of bovid’

Sem. **maʕiz-* or **miʕaz-* ‘goat’: Aram.: Jud. *mēʕazz-ē*, *mēʕazz-ayā*, *māʕizz-ē*, *māʕizz-ayyā* ‘goats’¹⁴ | ESA: Minean *mʕz-y* (du.) ‘chèvre’ | Arab. *miʕāz-*, *maʕīz-* ‘chèvre ou bouc’.

(?) Egypt. (Coffin texts) *myz-t* ‘horned animal (?)’¹⁵.

elaphus], Hbr. *ʔayyāl* ‘fallow deer’ [*Damadama*]; Aram.: Off. *ʔyl* ‘deer, stag’, etc.; Arab. *ʔiyyal-*, *ʔuyyal-* ‘cerf’ [*C. elaphus*].

¹¹ Cf. Ahaggar *élahei* ‘mouton à laine’: met.?

¹² Cf. what is perhaps a variant root with **p-*: Sem. ESA: Min. *fyʕ* ‘chevreau’, Arab. *faʕfaʕ-* id. (the meaning of both terms is questionable); Egypt. (20 Dyn.) *ʕpwy* (met.) ‘name of a holy ram’ [*Ammotragus leviola*]; Cush. E.: Dullay: Gawwata *peeʕ-atte*, Gobeze *piʕa-če* ‘she-goat’ (< **p/biʕ-*).

¹³ Acc. to Ehret 1980: 333, **ʕ* > Asa *ʔ*, **ʒ* (?) > **dy* > Asa *nd*; acc. to Takács 2011, **ʕ* > Asa *ʔ*, but **ʒ* > Asa *ʒ*.

¹⁴ Oddly translated in some Aram. dictionaries as ‘from goats, goats-hair, horn, etc.’.

¹⁵ Note lack of *-ʕ-* vs. the Sem. forms.

Chad. **maHa(n)z*- ‘k. of bovid (*Tragelaphus, Kobus*); male-goat’: W. **maHaz*-: Hausa *màaz-o* ‘harnessed antelope’ [*Tragelaphus scriptus*], Mburku *maaz-u* ‘antelope sp.’, Bade *máz-àrán* (compound?) ‘Ziegenbock’ || C. **ma/u(n)z*-: Fali-Bwagira *munza-ḥwun* ‘male goat’, Logone *máz-ā* ‘Antilopa hamariya’ [*Kobus kob*],¹⁶ Matakam *múz-àk* ‘he-goat’ || E.: Barein *múz-o* ‘ox’.

(?) Cush.: S.: Dahalo *máṣad-e* ‘female topi’ [*Damaliscus lunatus*].¹⁷

Omot. N. **miHi(n)z*- ‘cow, bull’: Wolaita *mízz-a*, Gofa *miz-a*, *minǰ-o*, Zala *mízz-ā* ‘cow’, *míz-a* ‘livestock, cattle’, Dorze *miiz*, Koyra *miz-a* ‘ox’, Chara *miy-ā*, Yemsa *miy-ā*, *miǰ-à* ‘cow’, Kafa *minǰ-o* ‘cattle’, Bworo *mii(n)z-à* ‘cow, cattle’, Dizi (Sheko) *mīnǰ-o* or *mīnz-a*.¹⁸

▲ Looks like a unique case of derivation with the prefix *m-* from **ʕi(n)z*- ‘k. of smaller bovid’ (1.4a.) as early as the PAA level. With all the problems in each of the proposed roots and/or their relationship, a rather rare combination of the radicals **ʕ* and **z* makes chance coincidence unlikely.

► SED II No. 148; EDE III 156–7; 796; AADB 2695.

1.5. **ba/i?*- ~ **baʔba?*- ‘k. of lesser bovid (*Tragelaphus silvaticus?*)’

Sem. **bib(b)*- (redupl.) ‘*Ovis orientalis, Tragelaphus sp.*’: Akk. (SB) *bibb-u* ‘wild sheep’ [*Ovis orientalis*] || Tgr. *bub-a* ‘koodoo (antelope)’ [*Tragelaphus sp.*].

Chad. W. **baw?* - and **baʔb(a?)*- (redupl.) ‘bushbuck (*Tragelaphus silvaticus*)’: Sura *pàab* ‘gazelle’, Angas *bwop* ‘he-goat’, Bolewa *ḥoḥ-oki*, Karekare *bàw-í*, Dera *bów-ò*, Tangale *bob-o*, Bele *bòḥ-ò*, Kirfi *bòʔʔ-ò*, Gera *bobb-i*, Galambu *bùb-à* ‘duiker’ [*Tragelaphus silvaticus*].

Cush. S. **baʔ*- and **baʔbaʔ*- (redupl.): Iraqw *buubuu-ti* (*-*a*- > -*u*- influenced by *b?*) ‘gazelle sp. (large and gentle)’ | Qwadza *baʔ-uko* ‘bush duiker’ [*T. silvaticus*] | Dahalo *ḥáḥaʔ-ááni* ‘bushbuck’ [*T. silvaticus*].

► SED II No. 54; cf. HSED #278.

1.6. **balay/w*- ‘k. of lesser bovid’

Sem. **w/yābil*- (met.) ‘ram’: Phoen. *ybl*, Hbr. *yōbēl* | Aram.: Dem. *ʔybl*, Jud. *yūbālā*-‘ram’ | Arab. *wābil-at*- ‘petits chameaux ou agneaux; petits d’un troupeau composé de chameaux ou de moutons’.¹⁹

Egypt. (OK, MK) *ibʔw* (very likely < **VbVl*- or **yVbVl*-, met.) ‘Mähnschaf (*Ovis tragelaphus*)’ [*Ammotragus lervia*].

Berb. N. **ballVy* ‘sheep, ram’: Gurara, Tuat, Tidikelt *belli* (pl.)²⁰

Chad. *(*m*-)*bil*- and **balaw/y*- ‘small cattle’: W. **m-bil*-: Geji *mbil-a* ‘sheep’ || C. **bVlaw*-: Matakam *báláw* ‘race de mouton sp.’ || E. **balVy*: Lele *bulóbuló* (redupl.) ‘he-goat’, Kabalai *bâl*, Migama *bòlly-ò*, Sokoro *bal* ‘goat’.

¹⁶ Cf. also Gudur *mangazaw* (< **manʕaz*-?) ‘reedbuck’ [*Redunca sp.*], Gisiga *mazawal* ‘male goat’.

¹⁷ *d* is considered to continue only **d*, but a few other cases of Dahalo *d* < **z*/**ǰ* including the present reflex in the triconsonantal root do not seem haphazard; an alternative comparison can be *mōʕo* (*mōdʕo*) ‘sp. large antelope’ commented in EDE III 158: “cannot belong here, being a late loan-word from SNil. (as suggested by EEN)”, which in the light of the AA parallels can be, on the contrary, a lw. from Dahalo.

¹⁸ Acc. to EDE III 157, «There seems to be a great confusion in the handling of the reflexes of NOM, **miz*- “Rind”...and NOM. **minǰ* “cattle”... usually mentioned together in Omotic etymologies...» The alternative approach suggested (ibid.): «NOM. **minǰ* “cattle” (palatalized from **ming*-?) is fairly close to ES: Amh. *mānga*, Gaf. *mānga* ‘flock, herd’» is hardly possible as the Omot. languages in question do not show palatalization of the velars (**g* and **k*) while the alleged source Eth. terms show *g*-, not *ǰ*-.

¹⁹ Cf. also Akk. *būlu* ‘herd of cattle, sheep or horses; wild animals (gen.)’ (unless from Sem. **bawl*- ‘mind, spirit, living creatures’).

²⁰ Infantile words: acc. to EDE II 24, “therefore cannot be considered” which is a debatable statement.

Cush. **(ʔa-m-)bVl-* ‘k. of bovid’ (*Tragelaphus sylvaticus*, etc.): E. **(ʔam-)bul-*: LEC **bulal-* (redupl.): Oromo *bulál-ē* ‘Lamm’ | HEC **ʔam-bul-*: Hadiya *ambul-a* ‘ram’, Kambatta *ambul-a* ‘ram, goat’ || S. **(ʔa)m-bal-*: Ma’a *mbāl-a* ‘Buschbock’ [*Tragelaphus sylvaticus*].

▲ Chad. **m-* and Cush **ʔa-m-* are fossilized nominal prefixes or “root-extensions” (see Militarev 2005).²¹

► Cf. HSED No. 2570; SED II No. 245; Mil. 2009; AADB 2484; EDE II 23–24.

1.7. **baray-* ~ **bayer-* ‘k. of lesser bovid’

Sem. **bayr-* ‘bull, ox’:²² Akk. (OB on) *būr-u* (*pūr-u*) ‘young calf’, (MB on) *bīr-u* ‘bull (for breeding); young cattle (up to three years)’ || Mand. *bir-a* ‘domestic cattle’, *bar-a* ‘lamb’ | Arab. (Yem.) *bār-ah* ‘cow’ | Tgr. *bar-a* ‘ox’ | Amh. *bare* (< **baray*) | Har. *bār-a* ‘ox, bull’ | Gur. **bawr-* ‘ox (for farming)’: Cha., Eža, Muh., Msq., Gog., Sod. *bor-a*, Gye. *bawr-a*, End. *bawr-ä*, Enn. *bawər-a* ‘ox (for farming)’.

Egypt. (OK) *bʔ* ‘heiliger Bock (*Ovis longipes palaeo aegyptiacus*)’ [*Ammotragus lervia*] | Copt.: Fayumic *barit*, Boheiric *barhit* ‘bouc’ (*-h-* is not clear).²³

Berb. S. **-baray* [*Addax* sp.]: Ahaggar *ē-berey* ‘faon d’antilope adax’ [*Addax nasomaculatus*], Ayr. E. Wlm. *e-bäräy* ‘faon d’oryx/adax’ [*Oryx* sp./ *Addax* sp.], W. Wlm. (Kel-Ataram) *bərwāy-a* ‘ram’.²⁴

Chad. **baray/w-* and **babar-* (redupl.) ‘k. of bovid (*Hippotragus*, *Cephalophus*, etc.): W. **baray/w-* and **(ba)bar-*: Hausa *bàrééw-áá*, pl. *bàrèey-íí* ‘gazelle’, Gera *bar-a* ‘buffalo’ [*Syncerus caffer*], Pa’a *bàbár* ‘roan antelope’ [*Hippotragus equinus*], Siri *babar-i* ‘gazella dama’ [*Nangerdama*], Tsagu *bur-āmən* ‘gazelle’, Polchi *mbār-àn* ‘antelope’, Zul *bar-námi*²⁵ ‘roan antelope’ [*H. equinus*], ‘buffalo’ [*S. caffer*] || C. **bVʔ/wr-*: Masa *boʔor-a* ‘k. of gazelle’, Banana *bbōr*²⁶ ‘cephalophe de Grime’ [*Cephalophus* sp.] || E. **ma-bur-*: Lele *mbur* ‘cobe onctueux (Cobe Defassa)’ [*Defassa waterbuck*], Kwang *mābār-ā* ‘duiker’ [*Tragelaphus* sp.].

Cush. **bayr-* (met.) ‘k. of bovid (*Oreotragus*, *Oryx* etc.)’:²⁷ LEC **bayr-*: Somali *bayr* ‘Antilope oder *Oreotragus saltator*, Klippspringer’ [*Oreotragus oreotragus*] (Reinisch 1902), *bārār* (pl.) ‘lambs’ || S. **baʔ/wr-*: Burunge *baʔur-u* ‘oryx’ [*Oryx* sp.].²⁸

²¹ There are also two phonetically tricky terms: Berb. E.: Ghadames *ta-bali*, Audjila *te-ḥel* ‘sheep’ which can originate either from **baHl-* or from **Hawl-* and Chad. C.: Boka *ḥwəl-ə* ‘goat’ where the voiced implosive is expected to originate from the combination **b(V)H-*; while no other AA zoonym conforms to such parameters, these forms may be reserved as Berb.-Chad. **bVHl-*. Cf. similar terms in non-Afras. African languages: C. Sud.: PMoru-Madi **bilɔ* ‘sheep’, Bagirmi *bal* ‘ram’; Ful *mbala*, *mbāl-u*, *bāli* ‘sheep’; W. Atlantic: Serer *mbal*, Mende, Kpele *bala* id.; Masai *em-balelo* ‘lamb’. Cf. also Chad. W.: **baHil-(Vm-)* ‘horn’: Montol *bulu*, Bolewa *ḥolum*, Galambu *ḥàli*, Maha *belem*, etc.; C. **ti-m-bil-*: Bura *ti-mbil*, Chibak *tə-mbəlá* ‘horn’ and Omot. S. **bal-*: Dime *bal-tu*, Galila *baal-i* ‘horn’.

²² Cf. also Proto-Sem. **baʕir-* ‘household animal; beast of burden’ probably derived from **bayr-* with a secondary *-ʕ-*: Ammonite *bʕr* ‘ox (beast of burden)’, Hbr. *bəʕir* ‘livestock, cattle’; Syr. *bəʕir-* ‘brutum, jumentum’; ESA: Sab. *bʕr* ‘cattle, head of cattle; camel’, Min. *bʕr* ‘chameau, tête de gros bétail’; Arab. *baʕir-* ‘chameau; toute bête de charge’; Gz. *bəʕar*, *bəʕ(ə)rā*, *bəʕ(ə)rāy*, *bəʕ(ə)rāwi* ‘ox, bull, horned cattle’, etc. and Muh. *bāre* ‘calves’.

²³ *-r-* in Fauymic points to **-r-* rather than **-l-*, though, acc. to EDE II, 23, it is “most probably, < **bl*”.

²⁴ Cf. also Ghadames *ābbur* ‘agneau’ (Motylinski 1904: 99), not in Lanfry 1973.

²⁵ *nami* ‘meat’.

²⁶ It is not clear if *bb* can convey *ḥ*.

²⁷ Also in N.: Beja *beʔrāy* ‘bull, cow’; C. **bir-* ‘ox, bull’: Bilin, Khamir, Kemant *bir-ā*, etc., E.: Afar *baʕer-ā* and Saho *beʕer* which must all be lws. from N. Eth. **bəʕray* ‘bull, ox’ (it is not so evident with Afar *abur* ‘bull, ox’), HEC **bāHōr-* ‘young bull’ (Sidamo *boor-to*, Hadiya *baara*, Kambatta *boora*) as well as N. Omot. Wolayta, Zala, Gofu, Kullo, Dache, Zaysse, Chara *bōra* ‘ox’, Gamu *bōra* ‘not castrated bull’, Kafa *bariy-ō* ‘calf’, Mocha *bariyō* ‘steer’, Bworo *berō*, Dizi (Sheko) *bariyō* ‘bull’ are more likely lws. from Amh. *bare* ‘ox, bull’.

► Cf. AADB 2498, 2487; HSED Nos. 183, 217, 341; ND 225a; EDE II: 24–25, 53 notes, 54–55; Blažek 2003b: 231, 238; SED II No. 53; Mil. 2009; CED #22.

1.8. *čawi?- ‘k. of bovid’

Sem. *šawa?- ‘sheep’: Akk. (OB on) *šuʔu* (*šū*) ‘sheep’, (OB) *šuātu* ‘ewe’²⁹ || Ugr. *š* ‘ram, sheep’ | Phoen *š* ‘sheep, one of a flock’, Hbr. *šā* ‘small livestock beast’³⁰ | Sab. *s₂h* ‘sheep’ | Arab. *šāʔ*- ‘brebis’, *šāt*- ‘brebis, mouton, espèce ovine; buffle (mâle)’.

(?) Eryp. (MK) *šw* ‘ass’ (with a meaning shift?).³¹

Chad. *ša-³² ‘cow; meat’ (with a meaning shift): W. *ša: Hausa *šá-nú*, *šá-nàye* (pl.) ‘ox’ | Jimi, Polchi *šáa* | Dwot, Buli, Zul, Ngizim *šáa*, etc. ‘cow’ || C. *ša ‘cow’: Bura, Margi, Gisiga *ša*, etc. || E. *sVwi: Kera *si* ‘cow’, Dormo, Nanchere, Kabalai *sī* ‘ox’, Ndam *swi* ‘zebu cow’, etc.³³

Omot. N. *šaʔ- ~ *ʔayš- (met.) ‘goat’: Bworo *eyšš-a*, Mao (Hozo) *šaa*, (Ganza) *saʔ-a*, Dizi *es-ku*.³⁴

► Cf. SED II No. 217; Mil. 2009; CED #831; AADB 229; HSED No. 517.

1.9a. *čaiw- ‘k. of lesser bovid’

Sem. *čaiw-: Ugr. *ʔ-at* (*ʔ-ut*) ‘sheep’ | Aram.: Anc. *šʔ-t*, *sʔw-n*, Sam. *šʔ-h* ‘sheep, ewe’, Mand. *ta-ta* ‘lamb, ewe, sheep’ | Arab. *čaiw-at*- ‘brebis maigre ou vielle’ || Mehri *čiw-īt*, Hrs. *čī-t*, Jib. *čī-t* ‘sheep’, Soq. *téʔ-e* ‘mouton, chèvre’.

Chad. *čVw- and *wVč- (met.) ‘small cattle’:³⁵ W.: Kirfi *wüč-i* ‘sheep’ || E. *čVw-: Kwang *čuw* ‘he-goat’.³⁶

(?) Cush. E.: LEC: *čawVʔ- (met.) ‘meat’: Somali *soʔ*, Oromo *fo-oni*, Baiso *soʔ-o*, Konso *sow-a*, etc.³⁷

Omot. *ʔač- ~ *ča(ʔ)-: N. *ʔač- (met.) ‘meat, flesh’: Koyra *ʔačč-o*, Zaysse *ʔač-ó*, Ganjule *ʔač-o*, Kachama *ʔač-ɔ* (the emphatic *č* is not justified – perhaps a record slip), Chara *ačč-aa*, Gimirra *ʔač* || S. *ča(ʔ)-: Ongota *ča-ta* ‘meat’.

► Cf. SED II No. 236; Mil. 2009; AADB 2652.

1.9b. *čawir- ‘male large bovid or herbivore’

Sem. *čawir- ‘*Alcelaphus buselaphus*; bull, ox’: Akk. *šūr-u* ‘bull (poet.)’ || Ugr. *tr* ‘bull’ | Hbr. *šōr* ‘one single beast, bovid’ | Aram.: Sam’al *šwr-h* ‘cow’, Bibl. *tōr* ‘ox, bull’, Syr. *tawr-ā* ‘bull’, Mand. *taur-a* ‘bull, ox’ | Sab. *čwr*, *čr* ‘bull’, Qat. *čwr* ‘steer, bull’, Min. *čwr* ‘bull’ | Arab. *čawr*- ‘taureau’ and ‘antelope bubalis’ [*Alcelaphus buselaphus*] (cf. Blažek 2003a: 8) | Gz. *sor*, *šor* ‘ox, bull’, *sawār* ‘fat ox’ | Tgr. *sor* ‘ox’ || Mhr. *čawr* | Hrs. *čawer* ‘bull’.

²⁸ Cf. also Cush. E.: HEC *hVbVr-: Qabenna *habur-ču* ‘goat’, Burji *ibār-i* ‘nanny goat’.

²⁹ Otherwise < *čaiw-.

³⁰ Jud. *šētā* ‘ewe’, *šē* ‘lamb’ and Sam. *šh* ‘sheep’ are most likely Hebraisms.

³¹ Cf. also *šʔy* ‘pig, sow’, if *ʔ* is not < *r* or *l*.

³² Perhaps *šaʔ- in view of such forms as *šáa*.

³³ In some of the languages, certain presumably related terms for ‘cow’ (not quoted here) show a voiced reflex (*b*), which Stolbova (CED #831) treats as originating “most probably... from *nḅ* <PCh *nʔ (note *n-* as a pref. of sg.)” (see in more detail in *šaʔ- ‘k. of bovid, herd’). Though in CED and other recent works by Stolbova, in accordance with most common usage in present-day Chadic studies, *t* and *b* are used instead of *š* (< AA *č) and *z* (< AA *š) accordingly, we continue here to follow the customary latter Afrasian notation. Cf. also the derivative *š(y)Vw- ‘meat’ (ibid.)

³⁴ Omotic reflexes of AA sibilants/sibilant affricates are not quite clear yet.

³⁵ Cf. W. *čič-: Kariya *čiči*, Gera *čača* ‘goat’, Warji *čiči-čá-ná*, Kariya *čiči* ‘he-goat’. Acc. to Stolbova’s observation, only Kariya *čiči* belongs here, while the other forms are < *čak^wi.

³⁶ Canarian (all islands) *chivato*, *chiva* ‘kid’ seems more likely < Spanish *chibo*.

³⁷ *s* can continue AA *s- and *č, while Oromo *f-* in *foon-i*, if it belongs here, points to *č.

Egyp. (OK) *sʷ* ‘taurillion’ (if < *čVr-) or (OK) šsʷw (if < *čVrw-) ‘antelope bubalis’ [*Alcelaphus buselaphus*].³⁸

Chad. *čiwār- (met.) ‘ox; elephant’: W.: (?) Dera *wór-é* ‘ox’³⁹ || C. *čiwār-: Bura *čiwār*, Chibak *isiwař*, Kilba *cuwar*, Nzangi *čūwar-e*, Kobochi *čūār-é* ‘elephant’ [*Loxodonta africana*].⁴⁰

Cush. *čawir- ‘bull; hippopotamus, waterbuck’: N. *sawsawr- (redupl.): Beja *sosar-i* ‘cattle’ || E.: LEC: Boni *šēr* ‘hippopotamus’ [*Hippopotamus amphibius*] || S. *čawr-: (?) Alagwa *cawr-u* ‘waterbuck’ [*Kobus ellipsiprymnus*]⁴¹ | Ma’a *čur-ú* ‘bull’.

Omot. N. *šawr-: Kachama *šor-o* ‘rhinoceros’ [*Dicerasbicornis*].

▲ This root is likely to have originated from #1.9a. *ča?iw- ‘k. of lesser bovid’ with suffixed *-r⁴² and loss of ? accompanied by an expected deletion of *-?- to size up to a triconsonantal root. The present reconstruction tilts the argument of Sem. *tawr- vs. IE *tauros towards the latter’s being borrowed from Sem.

► Cf. SED II No. 241; AADB 2525; Blažek 1994: 200 and Blažek 2003a.

1.10. *(?a-)dury- ‘k. of bovid’

Sem. *(?a-)dury-: Akk. *dudrū* ‘sorte de mouton’ (redupl.) || Arab. *diry-ān-at-* ‘espèce de taureau pourvu d’une bosse’ | Tgr. (pl.) *?addari-t* ‘dwarf-antelope’ [*Neotragini sp.*].

(?) Egyp. (OK) *idr* ‘Herde (von Vieh und Geflügel)’.

Berb. S. *-dVray- ‘oryx’ [*Oryx sp.*]: Ahaggar, Ayr, E.Wlm. *e-dəri*, pl. *i-dāra-n*.

Chad. *(?an-)dVry- ‘*Damaliscus sp.* etc.; small cattle’: W.*(?an-)dur-: Hausa *dari* ‘hartebeest’ [*Damaliscus lunatus*], Guruntum *dor-o*, *dōr-o* ‘goat’, Bokkos *?an-dūr*, Sha *n-dur* ‘ram’, etc. || C.*?u-dVr-: Musgu *u-deri* || E. *dar-: Dangla *dar-o* ‘gazelle sp.’

Cush. (?a-)dury- ‘*Tragelaphus sp.*, etc.’: E.: SA *wadar- (met.): Afar *wadār* ‘Ziegen, Kleinvieh’ | LEC *?a-dVry-: Somali *a-deri-o* ‘male kudu antelope’ [*Tragelaphus sp.*] || S. *(da)dury-: Burunge *dor-o* ‘zebra’ [*Hippotigrus sp.*] | Dahalo *dáđiiri* ‘Lesser Kudu’ [*Tragelaphus imberbis*].

Omot. *dury- ‘sheep’: N. *dur-: Koyra, Zaise *dor-ō*, Basketo, Doko *dori*, Oida *dur-o id.*, Gimira (She) *dor* ‘ram’ || S. *dayr- (met.): Ari *der-tí* ‘sheep’, Dime *der* ‘goat’.

▲ Cf. Sum. *dār*, *dāra* ‘ibex’ [*Capra ibex*] (EDS #439).

► AADB 2482. Cf. Takács 2011: 73.

1.11. *fa(?)n- ‘k. of large bovid’

Sem. *(w/ya-)pan- ‘young bull’: Ugr. *yp-t* ‘cow’ | Arab. *yafan-* ‘jeune taureau âgé de quatre ans’, *fan-āt-* ‘vache’ | Gz. *ta-yfan* ‘young bullock’, Tgr. *tā-fin* ‘giovenco, bue non ancora domato’, Amh. *wäyfän* (-äy- a diminutive infix?) ‘young bullock, calf of 2–3 years’, Gur. *wafan ‘young bull, steer’: Muher, Masqan, Goggot *wäfen*, Soddo, Wolane *m-ofän*, etc.

³⁸ Less likely Pyr. *šsr* ‘Schlachtrind’, which may be derived from *šsr* ‘Art des Schlachtens’.

³⁹*s, *č > Dera *w*, acc. to Stolbova 2009: 28; not in CED.

⁴⁰ Cf. C. *či/uw- ‘elephant’ (unless < Chad. *čVwar- with deleted *-r*): Bata *čūé*, Bana *čiwə̀*, Higi-Nkafa *čūwe*, Mwulien *čúwà*, etc. C. Chad. (Bura-Margi group) terms like *čVwVr- ‘elephant’ [*Loxodonta africana*] are usually considered by Chadacists to be < *čVwVn-; in Stolbova 2011 #433, the suggested reconstruction *ž/čVwVn- is compared with HEC *žanaVvy-; more likely these are two different roots since, acc. to CED: 28–29, reflexes of *č and *ž do not overlap, while forms like Bana *čiwə̀* are either to be treated as a separate entry (see 1.10a) or point to a deleted *-r* rather than *-n* in word-final position; acc. to CED 35, the former is generally less stable than the latter. At the same time, in CED: 44, **-n > -r* is postulated for C. Chad. groups 2, 3, 5 and 7, but not for group 4 including Nzangi-Kobochi in which, therefore, *-r < *-r* is regular which makes the whole idea of *čVwVr- < *čVwVn- less likely. Cf. also C.: Khamir *siriya* ‘Antilopensorte’ < *čir-?

⁴¹ But Burunge *cawadu* and Asa *šodo id.* probably point to *cawad- rather than *čawar-.

⁴² See the hypothesis on **-r* presumably conveying the meaning of a “larger dimension/quantity” in Militarev 2005: 49–52.

Chad. C. **fun-* ‘buffalo’ [*Syncerus caffer*]: Ga’anda *u-fan-a*, Higi-Dakwa *fun-u*, Higi-Kamale *n-fun*, Fali-Kirya *n-fu*, Banana *à-fún-tá*, etc.

Cush. S. **faʔan/m-* ‘k. of large herbivore’: Qwadza *faʔamo*⁴³ ‘buffalo’ [*S. caffer*], Asa *faʔan-ok* ‘elephant’ [*Loxodonta africana*].⁴⁴

▲ Cf. Bantu **-fun* ‘buffalo’ [*S. caffer*].

► SED II No. 250; AADB 2495.

1.12. **(ʔa-)fir-* ‘k. of bovid’

Sem. **parir-* (redupl.) and **paʔur-* (met.) ‘lamb; young of bovid’: Akk. *parr-u* || Hatra *pr-t?* ‘ewe or sheep’, Syr. *parr-ā*, Mand. *par-a* ‘lamb’ | Arab. *furār-*, *farūr-* ‘petit de brebis ou de chèvre’, *farīr-* ‘petit debuffle, de brebis, de chèvre’, *furfur-* (redupl.) ‘petit de brebis, agneau ou petit de chèvre, chevreau ou petit d’antilope’, (coll.) ‘agneaux ou chevreaux’, *fūr-* (pl.) ‘chamois’ [*Capra ibex*] || Hbt. *faʔoor* ‘male calf (from birth to 5 years)’.

Chad. **(ʔa-)fir-* ‘k. of larger bovid’ [*Tragelaphus*, *Hippotragus*, *Syncerus*]: W. **fir-*: Angas *fir* ‘roan antelope’ [*Hippotragus equinus*] || C. **(ʔa-)fir-ay-*: Hwona *ʔuf^oar-ā* ‘duiker’ [*Tragelaphus sp.*], Bura *fir*, Kilba *fur*, Margi *fur* ‘buffalo’ [*Syncerus caffer*], Mbara *fārāy* ‘cattle; dot (bride wealth)’.

Cush. E. **ʔa-fi/ur-* ‘small cattle’: Yaaku *apur* ‘sheep’ || S. **ʔa-fi/ur-* ‘he-goat’: Asa *ʔefer-et*, *ʔofor-ok*, Qwadza *aful-atu* (-l- < *r).

► Cf. HSED No. 1950; SED II No. 181; Mil. 2009; AADB 2494.

1.13a. **guday-* ‘k. of larger bovid’

Sem. **gaday-* ‘kid’: Ugr. *gdý* ‘kid’⁴⁵ | Phoen. *gd?* ‘goat’, Hbr. *gādī* ‘kid (of goat or sheep)’ | Aram.: Anc. *gdh* ‘goat’, Syr. *gady-ā*, Mand. *gadi-a* | Arab. *ʔady-* ‘kid’, *ʔadāy-at-* ‘gazelle; petit de gazelle’ [*Gazella gazella*].

(?) Berb. N. **gVnd-uz* ‘bull, calf’:⁴⁶ Rif *a-yenduz* ‘taureau’, Snus *a-yenduz*, Shenwa, Qabyle *a-genduz*, Sened *a-gendus*, etc. ‘calf’.

Chad. **ga/uday-*: W. **ga/ud-* ‘k. of larger bovid’: Hausa *gàd-áa* ‘antelope, duiker sp.’ [*Tragelaphus sp.*], Kariya *gud-am*, Miya *gud-an-zāku*, Pa’a *gud-an-cāka*⁴⁷ ‘Western kob’ [*Kobus sp.*], Ngizim *gád-ùwà* ‘duiker’ [*Tragelaphus sp.*] || C. **gVday-*: Zime-Batna *gódāy* ‘buck’.

Cush. **gad-* and **gund-* ‘k. of larger bovid’: E. **gad-am-*: LEC **gad-am-*: Oromo *gad-am-sa* ‘greater kudu’ [*Tragelaphus strepsiceros*],⁴⁸ Dirayta *gad-an-sa*, *gad-am-sa* ‘antelope’ | HEC **gud-*: Sidamo *god-a* ‘deer, gazelle’ || S. **gwand-*: Iraqw *gwand-a*, Alagwa *gwand-o* ‘ram’, Burunge *gond-i* ‘old ram’.

Omot. N. **gayd-* (met.): Zaisse *gaaid-é* ‘cattle’, *gaid-é-endo* ‘buffalo’ [*Syncerus caffer*].

► SED II No. 76; AADB 2490.

1.13b. **ga/ud-* ‘wild pig’

Chad. **gad-*: W.: Hausa *gàd-u* ‘warthog’ [*Phacochoerus africanus*].⁴⁹

Omot. N. **gud-in-*: Sheko *gudin-e*, Wolaita *gudun-ta*, Gamu *gudun-ci*, Dače *gudun-ç*, Gofa *gudun-ça*, Kafa *gudin-o*, Mao *kand* (k- < *g-; met.) ‘pig’.

⁴³ -m < *-n assimilated to f-?

⁴⁴ E. Cush.: Dullay: Harso *wayfan-akko*, Gollango *woyfanna* ‘bull’ are likely loans from Amh.

⁴⁵ Akk. (NB) *gadû* ‘male kid’ likely < W. Sem.

⁴⁶ It is not clear whether this unusual form could be a metathesis of the compound form **ged-n-uz*, -n- being *nota genitivi*; for *uz* see **ʔi(n)ʔ-* and **ʔi(n)ʔ-* ‘k. of smaller bovid’ 1.4a. (not attested in Berber).

⁴⁷ The two latter forms seem to represent word combinations of *gud* + *nota genitivi -n-* (see fn. 45).

⁴⁸ HEC: Burji *gadám-a* ‘greater kudu’ [*Tragelaphus strepsiceros*] may be a lw. from Oromo.

⁴⁹ Cf. Bolewa *gedd-o*, Mangas *gindir*, Boghom *kindi?*, Kiir *ʔindər* ‘fat’; Ngizim *gád-bàyák* ‘be fat’.

▲ Cf. Kanuri *godu*, Tubu *gadu*, Bantu **-gudu* ‘pig’, **gidi* ‘warthog’ [*Ph. africanus*].

► Cf. AADB 3758.

1.14a. **gawr-* ‘k. of bovid’

Sem. **(?a-)gurr-*: Akk. (MA) *gurr-atu*, *agurr-atu* ‘ewe’.⁵⁰

Egyp. (MK) *ḏr* ‘calf’ (if < **gVr*).

Berb. **-gur-* ‘small cattle’:⁵¹ Ghadames *a-žur* ‘bouc’ (Naït-Zerrad 2002: 859), Zenaga *a-grər*h ‘béliér’.

Chad. (a) **(?a-)garw/y-* ‘kind of bovid (*Redunca*, *Oryx*, *Ourebia*, etc.); animal, game’: W. **(?a-)gary-*: Hausa *āgārē* ‘a big, male red-fronted gazelle’ [*Eudorcas rufifrons*], Tsagu *gāre* ‘reedbuck’ [*Redunca sp.*], Mburku *gāri* ‘oryx’ [*Oryx sp.*], Ngizim *a-gare* ‘gazelle’ || C. *(?a-n-)gary-*: Logone *gari-a*, Makeri *ingərii* ‘antelope’, Buduma *ηgəri* ‘gazelle’ || E. **gVrw-* and **gVwVr-* (met.): Tumak *gəru*, Kwang *gowor-to* ‘antelope’, Ndam *gərù* ‘ourebi’ [*Ourebia ourebi*]; (b) **garaw-* ‘herd of cattle’: W. **gar(V)w-*: Hausa *gār-kei* ‘a herd, flock’, Bolewa *gaarùw-à* ‘pack ox’, Karekare *gaarùw-à* ‘bull’, Tangale *káarw-a* ‘cattle’ || C. **garaw/y-*: Bana *gàrəw-à* ‘troupeau de bœufs’, Gude *góra-nə* ‘herd of cattle’, *mà-gára* ‘shepherd, watchman’, Musgu *gari*, *gári* ‘Stier, Laststier’, Mandara *gári* ‘bull’, Hursa *goragor-a* (redupl.), Uzam *gwar-a*, Matakam *η-gwur* ‘ram’ || E. **garaw-*: Bidiya *gaaruw-o* ‘animal’, Migama *gáaráw* ‘bête, bétail’; (c) C. **g^war-* and **g^warg^war-* (redupl.) ‘ram’: Ouldem *g^wàr-à*, Mbuko *g^wàrg^wár-á*, Merey *g^warág^wàr-a*, Muyang *g^wòròg^wòr-à*, Mada *gurg^war-a* || E. **gagar-* (redupl.) ‘small cattle’: Sokoro *gaáger-o* ‘Schaf’, Mawa *gagar* ‘mouton’, Mubi *wegr-i* (met.) ‘chèvre’.

Cush. **garaw-* and **gwira?*- (met.) ‘k. of larger bovid (*Taurotragus*, *Alcelaphus*, etc.): N. **garuw-*: Beja *garuw-a* ‘male antelope; eland’ [*Taurotragus sp.*] || C. **gar-* ‘calf’: Bilin, Qwara, Kemant *gär*, Aungi *gara* || E.: HEC **gur(r)-*: Sidamo *gur-um?*-*iččo* ‘gazelle’, *garr-ančo* ‘Agazen antelope’ [*T. buxtoni*] || S. **g^wara?*?: Iraqw *gwará?-ai* ‘Hartebeest’ [*Alcelaphus caama*], Burunge *gera?*-*i* ‘Grant’s gazelle’ [*Nanger granti*].

Omot. N. **gaHar-*: Wolayta, Dawro *gaar-aa* ‘antelope dekula’ [*Tragelaphus decula*].⁵²

► AADB 2485, 3386, 4152; cf. CED #234, 235, 238; cf. SED II No. 85.

1.14b. **gur-* ‘wild pig’

Chad. E. **gawr-* or **gur-*: Jegu *gor-on* ‘Wildschwein’, Bidiya *gör-inya* ‘warthog’, *gerwa* (met.) ‘Wildschwein’, W. Dangla *bó-gór*⁵³ ‘pig doe’ (Skinner 1994).

Cush. C. **gVr-m-* ‘pig’: Aungi *gərm-i*, Kunfäl *gerim-i* (Appleyard 2006: 110).

Omot. N. **gur-it-* ‘pig’: Anfillo *gur-ico*, Bworo *gur-iccà*.

► AADB 2683.

1.15. **yawz-* ‘k. of smaller bovid’

Sem. **yuzāl-* ‘gazelle’ [*Gazella gazella*]: Akk. *huzāl-u*, *uzāl-u* ‘young gazelle’ || Aram.: Jud. *ʔwzl*, Syr. *ʔūzayl-ā* ‘antelope; hinnuleus’⁵⁴ | Arab. *yazāl-* ‘petit de gazelle; gazelle (mâle)’.⁵⁵

⁵⁰ Cf. Sab *grh* ‘domesticated animals’ (Biella 1982: 76; not in Beeston et al. 1982), Tna. *gahret* (-*h-* is non-etymological) ‘doe antelope’. Cf. Gz. *garħa* ‘tame, break a horse, mule’. Tgr. *gärwa* ‘kudu (antelope)’ [*Tragelaphus sp.*] is likely a lw. from Beja.

⁵¹ Cf. also Ahaggar *ă-gūr*, E. Wlm. *ă-gur* ‘animal castré’, Adghaq *a-žorh* ‘castrat (caprin)’ (Naït-Zerrad 2002: 878).

⁵² Lws. from Sidamo?

⁵³ Acc. to CED #234, *b-* is a prefix in E. Chad.

⁵⁴ Cf. *ʔwzl?* ‘porcus’.

⁵⁵ Cf. Gz. *kažal*, Tna. *kažal* ‘kind of antelope’ (< Arab., acc. to Leslau 1987: 457) and Chad. C.: Bura *kažal* ‘antelope sp.’, Fali-Bwagira *kažaran* ‘gazelle’. A variant root?

Berb. **zulay* (met. < **yuzal*) ‘buck’: Nefusa *zalay* ‘bélíer’, Sokna *zalay* ‘grand chevereau’, Siwa *zalaq*, Ghat *a-žulay*, Ahaggar *a-hûlay*, Ayr *ə-zolay* ‘bouc’, etc.

Chad. (no single form reconstructable): W.: Hausa *ganžar*⁵⁶ || C.: Gudur *maŋ-gazaw* ‘Redunca (reedbuck)’ [*Redunca sp.*]⁵⁷

Cush. C. **ʔa-gaz-in-* ‘agazen antelope’ [*Tragelaphus buxtoni*]: Qemant *agāzēn*, Aungi *agāziēn-ī*.

▲ Most likely, the root **ɣawɜ-* was joined with the suffix *-al-* (later fossilized) at the NAA level, forming Sem. **ɣuzāl-*, while the P.Berb. form underwent metathesis.

► Cf. AADB 2634; SED II No. 92.

1.16. **ħawr-* ‘k. of larger bovid’

Sem. **ħaw(a)r-* ‘bull, cow’: ESA: Sab. *ħwr-y*, Hdr. *ħwr*, Min. *ħwr* (probably) ‘cow’ | Arab. *ħawar-* ‘taureau’ (see SED II No. 106).⁵⁸

Berb. **-Hayr-* ‘calf’: Shilh *irey* (met.) ‘veau’, Ayr, E. Wlm. *e-ħər-i*, pl. *i-ħār-an* ‘jeune bœuf de 2 à 3 ans’.

(?) Chad. **ħar -*: E. **har-*: Somrai *hār-a* ‘elephant’ [*Loxodonta africana*].⁵⁹

Cush. E. **ħawr-*: LEC **ħawr-*: Oromo *hor-ii* ‘cattle’, Konso *hor-eeta* ‘wealth, livestock’⁶⁰ | HEC **hur-*: Burji *húr-i* ‘buffalo’ [*Syncerus caffer*].⁶¹

► AADB 4120, 380; cf. Stolbova 2019 #224.

1.17. **k^wayr-* ‘young (small) bovid’

Sem. **ka/irr-* ‘small cattle’: Akk. (Oakk) *kirr-u* (or *girr-u*) ‘a breed of sheep’ || Ugr. *kr* ‘ram’ | Hbr. *kar* ‘(young) ram’ | Aram.: Off. *kr* ‘sheep’ (not quite reliable).⁶²

Berb. **karw* and **-krar* (redupl.) ‘young small cattle’: Nefusa *a-krar* ‘billy-goat’, Semlal *i-kru* ‘chevreau’, *án-kkūr*, pl. *in-ukr-ān* ‘bouc’, Rif *i-karr-i*, Qabyle *i-kərr-i* ‘ram’, Ghat *a-kər-w-āt* ‘lamb’, Ahaggar *ē-krer*, pl. *a-krār-en* ‘bélíer’, *a-kerw-āt*, pl. *i-kerw-āten* ‘agneau’, etc.

Chad. **k^war-* ‘lesser bovid (*Redunca*, etc.); small cattle’: W. **k^walir-*: Kofyar *koor* ‘castrated goat’, Angas *kwēr*, Dira *kur-e* ‘reedbuck’ [*Redunca sp.*],⁶³ Gera *kwar-à-* ‘goat’,⁶⁴ Jimbin *kwar-iyā*, Mburku *kwaar-i* ‘meat, animal’, Zaar *kar-ò* ‘sheep’, Wangday *kér-ò* ‘ram’, Kulere *kiry-êt* ‘wild animal; meat’ || C. **kwarkwar-* (redupl.): Ouldem *kwàkwàr* ‘taureau’ || E. **ku/ar-*: Kera *kúrkí* ‘taureau’, Tobanga (N. Gabri) *kar-aŋ* ‘goat’.

Cush. E. **kawr-* ‘goat’: LEC: Arbore *kaar-íy* ‘heifer goat’, *kor-at* ‘male goat’ | Dullay **kawr-*: Dobase *kor-en-te* ‘female goat’ | Yaaku *kúrum-a* ‘young goat; lamb’.

► Cf. HSED No.1432; SED II No. 118; Mil. 2009; Stolbova 2011 #286; AADB 2491.

⁵⁶ Diri, Jimbin *ganžar* can be lws. from Hausa; cf. also Kanuri *nganžar* which may be the source of the Hausa term or, on the contrary, borrowed from it.

⁵⁷ Musgu *salak/g* ‘gazelle’ is a lw. from Berb.?

⁵⁸ Cf. also Gz. *ħarāwəyā* ‘pig, wild boar’, Tgr. *ħarāwəyā* ‘boar’, Har. *ħariyyā*, Amh. *ərəyā* ‘wild pig’, Selti *iräyā* ‘pig’, Gur. **ħəräyā* ‘porcupine’. Cush. N. **haraw-*: Beja *harawya*; E.: SA **ha(w)r-*: Afar *horay*, Saho *hara* ‘wild boar’ seem to be lws. from Eth. or *vice versa*.

⁵⁹ In CED #361, united with C.: Chibak *ħla* ‘Stier’, Kilba *hā-hál-dī* ‘bull’, Musgu *ħalúu*, *hollú*, Mbara *hùlúu* ‘buffalo’, where *-l-* may continue both **r* and **l*.

⁶⁰ Cf. Dirayta *hor* ‘multiply (of domestic animals)’.

⁶¹ Omot. N.: Malo *hāri*, Oyda (*h*)*arr* ‘cattle’ are from Oromo.

⁶² Cf. **k^(w)arm-*, probably related, with the suffix *-m-*: Amh. *kurm-a* ‘bull (not castrated)’, Har. *korm-a* ‘male (animal), brave’, Muh. *k^wärm-a* ‘the young male of a sheep bigger than *ḫābb^wät*, male (animal), bull’; cf. also Arab. *kawr-* ‘troupeau nombreux (de chameaux, de bœufs)’, ‘herd (of camels or gazelles, containing 150-200 animals)’.

⁶³ Also *kīr* ‘fattened ram’.

⁶⁴ Cf. Tangale *kārwa* ‘cattle’ and *kuri* ‘herd (of cattle)’, *kərne* ‘to protect, guard, keep, shelter, watch’, Dera *kárán* ‘herding’ < **kVr-* ‘to guard, to protect, to shelter’ in Stolbova 2011 #259.

1.18. *lawi?- ‘k. of large bovid’⁶⁵

Sem. *lawi?- and *lawli?-(redupl.) ‘k. of large bovid (*Bos*, *Syncerus*, *Connochaetes*, etc.): Akk. (OB on) *litt-u* (*līt-u*) ‘cow’, *lalû* (*lali?-u*, *lala?-u*) ‘kid’, *lulī-mu* ‘red deer, stag’ [*Cervus elaphus*] || Ebl. *lí-a-núm*, *lí-a-nu-um* ‘cow’ | Ugr. *ll?-u* ‘lamb, kid’ | Hbr. *lē?-ā* ‘wild cow’ [*Bos taurus primigenius*] (only as a pers. name) | Arab. *lā?aⁿ* ‘taureau sauvage, buffle’ [*Syncerus caffer*], *lu?lu?-* ‘antelope’ | Tgr. *lul-it*, pl. *läwäll-it* ‘(cow) with long, crooked horns’, Amh. *lil(lə)wil* ‘gnu’ [*Connochaetes sp.*] || Mhr. *lahay-tən* ‘cows’, Hbt. *leé*, pl. *lháyta*, Jib. *le?*, pl. *lhó-ti*, Soq. *?e-lh-eh* ‘cow’, *lúl-oh* ‘brébis’.⁶⁶

Egyp. (Pyr.) *iw?* ‘bull’ (if < *IVwV?-).

Berb. *w/yalaH- (met.): Izayan *ta-wala* ‘troupeau de bœufs, sangliers’, Ahaggar *élah-ei* ‘mouton à laine’.⁶⁷

Chad. *laway- ‘k. of larger bovid (*Tragelaphus*, etc.): W.: Dera *la-à* ‘cow’ || C.: Gude *la* ‘cow’ (Jungraithmayr & Ibriszimow 1994), Kapsiki, Higi-Nkafa *le*, Higi-Baza *lɔl-ε* (redupl.) ‘gazelle’, Higi-Ghye *le*, Bachama *lýey* ‘duiker’ [*Tragelaphus sp.*] (Stolbova 2005 #57), Masa *lúway* ‘troupeau, le betail’ (CED #549).

Cush. *lawi?- ‘cattle’: C. *luway-: Bilin *luwí*, Khamir *luwá*, Awngi *luwā* || E. *la?- and *la?la?- (redupl.): SA *la(?)-: Saho, Afar *lā* ‘cow, cattle’ | LEC *lo?(lo?)- ‘cows (coll.)’: Somali *lo?* ‘cows (coll.)’, Rendille (pl.) *loóly-o*, Oromo *loo-ni* ‘cows (coll.)’, *lal-eesa* ‘she-goat’, Konso *low-aa* ‘cows’, Dasenech *lal-o*, pl. *lal-i* ‘cow’, etc. | HEC *lal-: Sidamo *lal-o* ‘cows, cattle’, *laat-to* (f.) ‘young sheep, lamb’ | Dullay *lo?-, pl. *le?- ‘cow’: Harso, Dipina, Gollango *lól-o*, Tsamai *lō?-ō-*, pl. *lē?-ē*, etc. || S. *li?-: Iraqw *le?-i*, Gorowa *lee?-i* ‘goat’ | Qwadza *le?-amuko* ‘bull’.

► AADB 301, 2477. Cf. HSED Nos. 1632, 1647, 2595; SED II Nos. 142, 143; Mil. 2009; EDE I 86.

1.19. *mayn- ‘k. of large bovid’

Egyp. (MK) *mnmn.t* (redupl.) ‘herd, cattle’.

Chad. W. *mawan-: Boghom *mwòŋ* ‘ram, sheep’ || C.: Uldeme *maan-ša* ‘cow, bull’, Musgu *mwuŋ* ‘buffalo’.

Cush. *mayn-: C.: Qemant *miyän* ‘young cow’; E.: SA: Saho *móynoy* ‘bull’.

Omot. N. *mēn- ‘buffalo’: Wolaita *mien-ta*, Male *méni*, Kachama *mēn-o*, Ganjule *mēn-o*, Gimirra *men*, Dizi (Sheko) *mēn*, Mao (Bambeshi) *mènè* (cf. Kafa *mēnō* ‘meat’).

► EDE III 293–4: cf. E. Nilotic: Ongamo (pl.) *o-mónyí* ‘bull’, Maasai (pl.) *il-mónjí* ‘oxen’; AADB 4261.

1.20a. *mar(ay)- ‘k. of large bovid’

Sem. (a) *marVy- ‘bull’: Akk. (SB) *mīr-u* ‘young bull’, (OB) *mīr-tu* ‘cow’ || Hbr. *mārī(?)* ‘fatted steer’⁶⁸ | Arab. *māriyy-at-* ‘antelope blanche’.⁶⁹

Egyp. (Pyr.) *m?* ‘oryx-antelope’ [*Oryx sp.*] (if < *mVr).⁷⁰

Chad. *maluray- ‘*Kobus sp.*’: W. *maray: Hausa *mààrír-ì* ‘white oryx’ [*Oryx leucoryx*],⁷¹ *máàrày-áá* ‘Western kob (medium antelope)’ [*Kobus kob*], Saya *mááriy-á* ‘antelope’, Warji *mārai*, Ngizim *máàrày-á* ‘kob’ [*Kobus sp.*] || C. *mur-: Masa *murí* ‘antelope sp., prob. kob-défassa’ [*K. ellipsiprymnus*] || E.: Toram *múr-o* ‘gazelle’.

⁶⁵ Cf. Chad. *IVw(h)- ‘wild animal(s), meat’ (CED #549), likely related.

⁶⁶ Note non-etymological *-h-* in MSA.

⁶⁷ Cf. also *alu* ‘bull’, quoted in EDE I 86 as “Tamasheq”, dialect name and source not specified.

⁶⁸ Usually regarded as derived from the verbal root *mr? ‘to be fat’, which is less likely in view of the Akk. parallel.

⁶⁹ Unless ‘white, milk-colored’ is Arab. *miry-at-* ‘lait qui sort en filet abonant du pis de la femelle’ (BK II 1097-8).

⁷⁰ Cf. also (late) *mr.t* (if < *mVr-, not *mVl-) ‘cow’; cf. also (MK) *mr*, *mry* ‘Kampftier’ with the sign for bull/cow’.

⁷¹ Less likely < Arab. *māriyy-at-* ‘antelope blanche’ in view of other Chad. data; however, the similarity is somewhat striking.

Cush. E. **mar-*: LEC **mar-*: Rendille *máar* ‘male calf’, *maár* ‘female calf’, Arbore *máar* ‘calves’ (coll.) | HEC **mar-*: Hadiya *moor-a* ‘older calf’ | Dullay **mar-* ‘heifer’: Harso, Dihina *maar-akkó*, Tsamai *mār-e* || S. **ma(w)r-*: Ma’a *mor-o* ‘ox, steer’⁷² | Dahalo *morr-o* ‘duiker’ [*Tragelaphus sp.*].

Omot. N. **mar-*: Gamu *mar-a*, Dorze *mar*, Male *marr-o* ‘calf’ (probably lws. from E. Cush.)
 ► SED II No. 153; AADB 2331, 2460; EDE III 26, 393–4; HSED 1728; Mil. 2009.

1.20b. *(*ʔa-*)*mayr-* ‘k. of small bovid’

Sem. **ʔa/i-mmar-* ‘lamb, sheep’: Akk. (OA, OB on) *immer-u* ‘sheep; sheep and goats; ram’, (OB on) *immer-tu* ‘ewe, sheep (as a generic term)’ | Ugr. *ʔimr* (and *ʔamr*) ‘ram’ | Phoen. *ʔmr*, Hbr. *ʔimmēr* | Aram.: Anc., Off. *ʔmr* ‘lamb, sheep’, Bibl. *ʔimmēr* ‘lamb’, Sam. *ʔmr*, Syr. *ʔemr-ā* ‘lamb’, Mand. *ʔa/umbr-a* ‘sheep, lamb’ | Arab. *ʔa/immār-* ‘lamb’.

Chad. W. *(*ʔa-*)*mar-* ‘ram, goat’: Tangale *mar-a* ‘(castrated) goat’, Dirì *mar-ì* ‘goat’, Bokkos *maray*, Sha à*mar-á*, Fyer *ʔamar-a* ‘ram’.⁷³

Cush. E.: SA **mar-*: Saho, Afar *mar-ū* ‘ram’.

Omot. **mar(ay)-*: N.: Wolayta *mār-ā* ‘calf, lamb, kid’, Gamu *mar-a*, Dorze *mar*, Male *marr-o* ‘calf’, *màrày-i* ‘ram’, *marmar-o* (redupl.) ‘lamb’,⁷⁴ Koyra *mar-a* ‘ram’, Bworo *merēr-à-* ‘sheep’, etc.⁷⁵

► SED II No. 5; AADB 2591; EDE III 42, 394; HSED 1729; Mil. 2009.

1.21. **naʔiw-* ‘k. of bovid’

(?) Egyp. (NK) *wnwy* ‘Kalb (als Bezeichnung des Osiris)’.

Chad **naHaw-*: (a) W. **na* ‘cow’: Tangale *ta-ŋa* | Boghom, Tala, Jimi *naa* || C. **naw*: Lame *náo* | Mesme *nau*, etc. ‘cow’ | Mandara *náw-è* ‘goat’.⁷⁶

Cush. **naʔiw-*: N.: Beja *naaʔ* ‘sheep’ || C. **naw-* ‘calf’: Khamta, Khamir *níw* | Kailiña *nāw* | Aungi *nēw*, Kunfal *nāw-i*⁷⁷ || LEC **nVw-*: Boni *núì* ‘greater kudu’ [*Tragelaphus strepsiceros*].

► AADB 2025, 2530; HSED, 42; Blažek 2003b: 234; Stolbova 2005 #364, 364a; Mil. 2009; CED #585 (comp. with Hbr. and Arab.).

1.22. **saraw/y-* ‘k. of bovid’

Egyp. (OK) *sʔ* (if < **sVr-*) ‘Herde’.

Berb. S. **saraw/y-* ‘troop, herd’: Ahaggar *ǎ-sāra*, pl. *i-sārâ-tān* ‘troupe d’animals sauvages’, Taitoq *a-sera* ‘troupeau de moutons, de chèvres, d’ânes’, Tadghaq *ta-sāraw-t* ‘small herd of cows (up to about 30)’, *ta-sāray-t* ‘large herd (esp. of camels or cattle)’.

Chad. W. **sa(r)sar-* (redupl.): Hausa *sassar-o* ‘male red-fronted gazelle’ [*Eudorcas rufifrons*].

Cush. E. **sar(a)w/y-* and **sawr-*: SA **saraw-*: Saho *saraaw-ii* ‘einzwējähriker zur Arbeit noch nicht abgerichteter Stier’ | LEC **sary-* ‘Oryx, Madoqua’: Somali *sary-an* ‘adult male oryx’ [*Oryx sp.*] | Dullay **sawr-* ‘antelope dikdik’ [*Madoqua sp.*]: Gollango *sawr-ó*, *sayr-ó*, Tsamai *sawr-o*.⁷⁸

► AADB 299. Cf. Blažek 2003a: 9.

⁷² Acc. to Ehret 1980, < NS, but in light of other Cush. and AA parallels this seems less likely.

⁷³ And, perhaps, Tala *màar*, Buli *maro*, Polchi *mar* ‘goat’ unless < **maq-*. Cf. also W.: Dwot *maariiri* ‘rhinoceros’; C. **maray* ‘bull’: Mofu-Gudur *maray* ‘taureau engraisé à l’étable’, Matakam *mari* ‘bull’, Mafa *maray* ‘fête rituelle de taureau; taureau de case sacrifié lors la fête’.

⁷⁴ Diminutive reduplication?

⁷⁵ Ongota *mar-te* ‘she-calf’ < Tsamai.

⁷⁶ Cf. also Sem. **naw/yʔ-* ‘meat’: Akk. (YB) *unû* ‘k. of meat’; Hbr. *nāʔ* ‘raw (flesh)’; Arab. *nyʔ* ‘être cru, n’être pas cuit (viande)’; Chad. **nay(n)* ‘meat’: W. **nay(n)*: Mupun *nān* ‘flesh’, Galambu *nyàa* ‘meat’; C. **nay(n)*: Munjuk *nèŋ*, Musgu *nē*, *neŋ*, *nij* ‘meat’.

⁷⁷ Tgr. *nəway* ‘cow, possession/fortune’ is most likely from Agaw.

⁷⁸ Omot. S.: Ongota *saire* ‘antelope dikdik’ [*Madoqua sp.*] is from Tsamai.

1.23. *šaʔ- and *šaʔ- ‘k. of large bovid’

Berb.: Ahaggar *êsu* (< **Hišw*, met.) ‘bull’.⁷⁹

Chad. **žaH-* ‘large cattle, cow’: W. **žaHa:* Siri *žáà-ti*, Dott *žáá*, Buli, Dott *žáa* ‘cow’, etc. || C. **ža:* Tera *ža*, Gude *žà*, Buwal *žà*, Sukun *žə*.⁸⁰

Cush. **šaʔ-* or **šaʔ-*: N. **šaʔ-*: Beja *šaʔ* ‘cow’ || E. **saʔ-*: SA **saʔ-*: Saho *sáʔ-a* ‘cattle’, Afar *sáʔ* ‘small stock (animal)’ | LEC **saʔ-* ‘cow’: Somali *saʔ*, Jidda *sii*, Boni *sáʔ-a*, Rendille *saʔ*, Oromo *sáʔ-a*, Konso *sa-a*, Bayso *se*, Arbore, Dasenech *seʔ* | HEC **saʔ-*: Sidamo *saʔ-aa*, Darasa *sa-aya*, Kambatta *saʔ-a*, Hadiya *saa-yya*, Burji *saa* | S. **šaʔ-Vy-* ‘cow’: Iraqw, Alagwa, Burunge *še-e*⁸¹ | Asa *še-ok* | Qwadza *šae-ko*.

► Cf. HSED #517; AADB 381; CED #831; Mil. 2009.

1.24. *waʔil- ‘k. of lesser bovid’

Sem. **waʔil-* ‘mountain goat (*Capra ibex*, *Arabitragus jayakari*): Ugr. *yʃl* ‘mountain goat’ | Hbr. *yāʔēl* ‘ibex, mountain goat’ | Aram.: Syr. *yaʔl-ā* ‘ibex (?)’, Jud. *yaʔāl-ā*, *yāʔēl-ā* ‘mountain goat, wild goat’ [*Capra ibex*] | ESA: Sab., Qat. *wʃl* ‘mountain goat, ibex’ [*Arabitragus jayakari*], Hdr. *?-wʃl* (pl.) id. | Arab. *waʔl-*, *waʔil-*, *wuʔil-* ‘argali, bouc de montagnes’ [*Ovis ammon*]; bouc domestique’ | Gz. *wəʔəl-ā*, *waʔāl-ā*, *wāʔəl-ā* ‘kind of antelope, mountain goat’ [*C. ibex*], Tgr. *wəʔəl-a* ‘elk/élan’ (DRS 578) [*Cervus elaphus*], Amh. *waliy-a* ‘Walia ibex’ [*C. ibex*] || Mhr. *wēl*, pl. *wəʔōl*, *wəyʔōl*, Hrs. *wāl*, Jib. *ebʔóž* ‘Hemitragus Jaykari, mountain goat’ [*A. jayakari*], Hbt. *wəʔáal*, pl. *wəʔyoól* ‘wild goat, oryx’.

Berb. **wulli* ‘small cattle’: Qabyle *ulli* ‘sheep’, Ghadames *welli*, Ayr, E. Wlm. *wəlli*, Ahaggar, Tadghaq *ulli* ‘goat’.

Chad. **wayl-* ‘k. of lesser bovid’: W. **wal-*: Jimi *wal-aróo* ‘antelope’, Geji *wull-ì* ‘gazelle’ || E. **wil-*: Lele *òl*, Kabalai *yil-ə* ‘antelope’.

Cush. E. **wiʔl-* ~ **ʔawl-* (met.) ‘k. of lesser bovid’: LEC **ʔawl-*: Somali *ʔawl-kii*, Rendille *ʔol* (*ʔ < *ʔ*) ‘gazelle’ || HEC **wiʔl-*: Sidamo *wil-ii’l-icco*, pl. *wil-ii’le* ‘young (of sheep, goat)’.

► SED II No. 244; AADB 2470.

1.25. *wad- ‘k. of lesser bovid’⁸²

Sem. **wad-*: Arab. *wad-ān-* ‘mouflon’ [*Ovis orientalis*].

Berb. **a-wdad* (part. redupl.) ‘Barbary sheep (*Ammotragus lervia*): Nefusa *a-wdād*, Semlal, Izdeg, Ghadames, Ghat *udad*, Ahaggar *ūdad*.

Chad. **wad-* ~ **daw-* (met.) ‘k. of lesser bovid’: W. **wad-*: Hausa *ud-a* ‘k. of sheep’ || C. **daw/y-*: Mofu *dáw* ‘goat’, Zime-Batna *dèy* ‘duiker’ [*Tragelaphus sp.*].

Cush. E. **ʔid-* ‘sheep’: SA **ʔayd-* ‘sheep’: Saho *ayd-ó* ‘sheep (coll.)’, Afar *id-a* ‘ewe’ | LEC **ʔidad-* (redupl.): Bayso *idādo* ‘sheep’.⁸³

⁷⁹ Perhaps also P. Berb. **isa-n* (pl.) ‘meat’ which may alternatively originate from AA **sVʔ-* ‘meat’.

⁸⁰ In CED #831 (where *ž* is rendered as *ʃ*), the above forms are united with the forms in *š-* (rendered by *ʃ*) with the following comment: “The word for ‘cow’ shows irregular voiced reflex (*ʃ*) in a number of Ch Biu-Mandara languages. Most probably, *ʃ-* originates from *nʃ* < PCh **nʔ-* (note *n-* as a pref. of sing.);” this ingenious explanation “reconciling” the reflexes of two different AA laterals (Chad. *š/ʃ* < AA **č* and Chad. *ž/ʃ* < AA **š*) is opposed by an alternative solution proposed herein.

⁸¹ ? is < *ʔ, not < *ʔ in W. Rift languages.

⁸² Cf. also miscellaneous terms that are somewhat similar and probably eventually related: Sem.: Akk. (lex.) *dūdu* (< **dawd-*, partial reduplication?), (OA, SB) *aletūdu* ‘wild sheep (male)’; Off. Aram. (*hapax* in Frah) *ʔndwt* (**ʔV-n-daw-t?*) id. (cf. SED II No. 44) and Cush. **ʔad(d)Vy-* (?): E. **ʔaddi* ‘large cattle’: SA.: Afar *addi* ‘young heifer’, HEC: Hadiya *addi-cco*, Sidamo *adde*, Burji *addi* ‘heifer’; S. **deʔ-* (met.?): S.: Qwadza *deʔ-* ‘to herd’, Ma’a *dé* ‘cow, cattle’. Cf. Blažek 2003b: 244 (Beja *ḍa* ‘newly born calf’ and the E. and S. Cush. forms are probably not related, since Beja *ḍ* does not continue **d*).

⁸³ Difference in the first radicals and stem patterns makes cognation with the present root questionable.

▲ Cf. Sum. *udu* ‘sheep’ (EDS #2678).

► AADB 2531.

1.26. **(ʔa-)w/yan-* ‘k. of lesser bovid’

Chad. *(ʔa-)yan-* ‘goat’: W. **yan-*: Siri *yàan-i* || E. **ʔayn-* ‘goat’: Birgit *ʔàyn-éy*, Migama *ĩ:n-á*, Jegu *ʔàyn-éy (té-én-ó, pl. ʔéén* ‘he-goat’).⁸⁴

Cush. **(ʔa-)w/yan-*: N.: **ʔan-*: Beja *an-o* ‘sheep’ || LEC **wan-*: Somali *wan*, pl. *wanan* ‘ram’, Rendille *onó* ‘sheep’ | HEC **ʔayn-*: Burji *ayáan-e* ‘gazelle’.⁸⁵

Omot. S. **yVn-* ‘sheep’: Dime *iin*, (?) Ongota *hoona (h-* probably renders ?).

► AADB 2530.

1.27. **war-* ‘k. of larger bovid’

Egyp. (MK) *wr.t* ‘göttliche Kuh’, *wr* ‘Steer’, in dem Ortsnamen.

Berb. **wVrw-* (< **ʔarw*): Semlal *ta-uru-t* ‘troupeau (de bœufs)’, Azilal *ta-wru-tt* ‘troupeau’.

Chad. W. **war-*: Hausa *wariyy-a* ‘a k. of antelope’,⁸⁶ Dera *wor-e*, Kariya *wúr-ù*, Warji *ruwe* (met.) ‘cow’.

Cush. E.: HEC **war-*: Sidamo *war-aamo* ‘older calf’.

► HSED 2527; AADB 2497.

2. Camels and equids

2.1. **dawk-* ‘k. of equid’

Sem. **kawd-an-* (met.) ‘mule’: Akk. (Oakk. on) *kūdanu (kūdannu)* || Ebl. *gú-da-núm <kūdanum>* | Ugr. *kdn-t* | Aram.: Off. *kwdn*, Palm. *kwdn*, Syr. *kūdan-yā, kūdān-ā* | Arab. *kawdan-, kawdan-iy-*.⁸⁷

Chad.: W. **daw(a)k-* ‘horse’: Hausa *dók-ì*, Karekare *dókú*, Miya *duwak-a*, Burma *dok*, Ngizim *dük-à*, etc. || C. **(n-)dok*: Tera *doh* ‘stallion’, Pižimdi-Hina *n-dok* ‘pony’ || E. **dawk-*: Dangla *dewk-i* ‘antilope-cheval’ [*Hippotragus equinus*].

Cush. S. **dak-et-*: Iraqw *dak-eti*, Alagwa *dek-eti* ‘zebra’ [*Hippotigris sp.*].⁸⁸

► SED II No. 124; AADB 2639; cf. HSED 667.

2.2. **gayl-* ‘camel’

Sem. (a) **ʔa-gayl-* and **gilgal-* (redupl.) ‘k. of equid’: Akk. (Oakk. on) *agāl-u, agall-u* ‘an equid’ (CAD a1 141), ‘Reitesel’ (AHw: 15) || Ebl. *ag-lum* = ANSE.NITA.KUR (a donkey) | Tgr. *gālgāl* ‘foal (mule, horse)’, Tna *gālgāl* ‘puledro di cavallo o mulo non ancora domato’ (Bassano 1918: 815), Amh. *gālgāl* ‘the young of domestic animals (goat, sheep, horse, donkey)’; (b) **gimal-* (met. < **gayl-am-*)⁸⁹ ‘camel’: Hbr. *gāmāl* | Aram.: Off., Palm., Nab. *gml*, Syr. *gaml-ā* | Sab. *gml* | Arab. *ʔaml-, ʔamal-* | Gz. *gamal* || Jib. *gūl* (pl. *gmíhl*), Soq. *gimāl*.

⁸⁴ Cf. Can. **Han(y)-*: Tenerife *ana, haña, jana* ‘sheep’.

⁸⁵ Cf. also Afar *anaŕ-to* ‘lamb (female)’ (non-etymological -ŕ). Sem. Gur.: Cha. *onā*, Eža *onnā* ‘young male goat or sheep’ must be from Cush.

⁸⁶ Cf. also *warari* ‘a vicious bull’ (Skinner 1996: 283).

⁸⁷ Likely related is Tgr. *mākada*, name of the small Abyssinian horse.

⁸⁸ Cf. irregular variant forms (variant roots?): C.: Bilin, Khamir *dəh^wara*, pl. *daḵ^wal*, Aungi *dəy^wari*, etc. ‘donkey’; E.: LEC: Oromo *donḵoro* ‘stupid, fool’, (Bararetta dial.) *dongorra* ‘donkey’; S.: Iraqw, Gorowa, Alagwa, Burunge *daḵwai* ‘donkey’, Qwadza *dagwagwai-ko* (Dolgopolskiy 1973: 275), *daḵ^wak^way-iko* (Ehret 1980: 345). Cf. also Chad. C. **zVng^w-* ‘donkey’ (Stolbova 2009 #530).

⁸⁹ Cf. a similar process in **ḥimār-* < **ḥayr-ām-* (2.4).

Berb. **gilgal-* (redupl.): Ahaggar *ā-gelgāl-i* ‘chameau de charge commun et lourd’.⁹⁰

Cush. E. **gayl-* ‘camel’: SA **gāl-*: Saho, Afar *gāl-ā* || LEC **gayl-*: Somali *geel-a* (coll.), Boni *g’áál*, Rendille *gaal*, Oromo, Baiso *gaal-a*, Arbore *gaal*, Elmolo *gal-te*.⁹¹

► AADB 2508; cf. SED II No. 79.

2.3. **hi/ul(l)-* ‘k. of equid’

Sem. **hil(l)-ay-* ‘colt, foal’: Akk. (OB) *ilulay-u* ‘ein Kamel?’ || Tgr. *ḥel-e* ‘camel (poet.)’, Tna. *ḥawl-a* ‘donkey’s colt or foal of about two years’ || Hrs. *me-ḥeley-ōt* ‘she-camel with fully-grown young’.

Berb. S. **hulil* (redupl.) ‘wild ass, foal’: Ghat *a-hulil* ‘âne sauvage’ [*Equus asinus*], Ahaggar *ā-hāhul* ‘jeune chameau non castré’, Ayr, E. Wlm. *ə-hulel*, pl. *i-hule-tān* ‘poulain, ânon’.

Cush. E. **ʔi-ḥVII-* (redupl.) ‘k. of equid’: LEC *(*ʔi-*)*haw/yl-*: Rendille *é-ḥel*, Arbore *holl* | Dullay *(*H*)*all-*: Gollango *ooll-ó* ‘Pferd’ || S. **ḥVIIV?*- (met.?): Dahalo *ḥelleʔ-a* ‘zebra’ [*Hippotigris sp.*].⁹²

► AADB 2462.

2.4. **ha/ir-* ‘wild ass’

Sem. (a) **ḥayur-* ‘ass, camel’: Arab. *ḥuwār-*, *ḥiwār-* ‘petit de chameau récemment né, ou jusqu’à l’époque du sevrage’⁹³ || MSA **ḥayr-*, pl. *ḥayor*.⁹⁴ Mhr. *ḥayr* ‘male donkey’, *ḥīr-it* ‘female donkey’, Hrs. *ḥayr* ‘donkey’, *ḥayr-ēt* ‘she-donkey’, Hbt *ḥáyr*, pl. *ḥyeér* (m.), *ḥiiri-it*, pl. *ḥiyór-tə* (f.) ‘donkey, ass’, Jib. (Kathīri dial.) *a-ḥyér* ‘male donkey’; (b) **ḥimār-* ‘wild ass, onager, donkey’ (met. < **ḥayr-ām-*):⁹⁵ Akk. (Oakk. on) *imēr-u* ‘donkey, male donkey’ || Ugr. *ḥmr*; Hbr. *ḥāmōr* ‘he-ass’ | Aram.: Off. *ḥmr* ‘donkey, ass’, f. ‘she-ass’, Palm. *ḥmr* ‘donkey, ass’, Syr. *ḥmār-ā* ‘asinus’, *ḥmār-tā* ‘asina’, Mand. *hamar-a* ‘pack-animal, ass, donkey’ | Sab., Min. *ḥmr* ‘ass; wild ass, onager’ [*Equus asinus*, E. *hemionus*], ‘domesticated ass’ | Arab. *ḥimār-* ‘âne; âne sauvage, onagre’ [E. *asinus*, E. *hemionus*], *ḥimārat-* ‘ânesse’⁹⁶ | Gur. **əmar* ‘donkey’: Cha. Gog. *əmar*, Sod. *əmar*, Wol. *umar*, etc.⁹⁷

Cush.: N. **har-*: Beja *har-e* ‘camel’ (Blažek 2003b:13)⁹⁸ || E. **ḥa(y)r-* ‘wild ass, donkey’: SA **ḥayr-*: Saho *ḥer-a* ‘female donkey’ | LEC **harr-*: Oromo *harr-ee*, Konso *harr-eta*, Dirayta *harr-et* ‘donkey’, Dullay **ḥar(r)-*: Gawwada, Dobase, Golango *ḥarre*, Harso *ḥar-iččé* ‘ass’ [*Equus asinus*].⁹⁹

► SED II Nos. 98, 106; AADB 2463, 4120.

2.5. **k^war-* ‘wild ass or camel’¹⁰⁰

⁹⁰ Cf. Zenaga *a-žiž* (< **a-gil*) ‘âne’.

⁹¹ HEC **gāl-*: Sidamo, Darasa *gaala*, Burji *gaali*; Dullay **kāl-* (< **gāl-*): Gawwada *kaal-ḥó* (< **kāl-k-*), Harso, Dihina, Gollango *kaalá* probably from Oromo. Omot.: Dache, Zaysse *gaale*, Yemsa *gaala*, Kafa *gallo*, Dizi (Sheko) *gaale* ‘camel’ perhaps also from Oromo, probably via HEC.

⁹² Omot. N.: Zaisse *ollō*, *wollō*, Ganjule *ollo* ‘horse’ are probably lws. from Dullay.

⁹³ Tgr. *ḥəwar* ‘young (of camel or donkey)’ is likely an Arabic loan. Berb.: Ahaggar *awra*, pl. *iwerân* ‘chameau dans sa 1ère année’, Ayr, E. Wlm. *awāra*, pl. *iwāran* ‘chamelon de quelque mois’ are probably Arabisms.

⁹⁴ Erroneously related to Sem. **ḥayr-* ‘donkey’ in SED II No. 50, analyzed as forms with the definite article *ḥ-*. However, this article occurs only in “words with an initial radical ?...and certain...monosyllables with an initial w or y radical” (Johnstone 1987: 163-4). There seems to be no cases of the *ḥ-* article replacing initial etymological *ḥ-* in either Mhr. or Jib.

⁹⁵ Cf. a similar process in **gimal-* (< **gayl-am-* (2.2)).

⁹⁶ Cf. also *aḥmūr-*, *yaḥmūr* ‘onagre’ [*Equus hemionus*].

⁹⁷ Cf. Mhr. *ḥəmūr* ‘to tame, break (a horse)’, Hrs. *ḥemōr* ‘to tame, break (a riding beast)’, Jib. *ḥōr* ‘to tame, break, domesticate (an animal)’.

⁹⁸ Cf. *harri* ‘mount (anything that can be ridden)’.

⁹⁹ Wolamo *hariya*, Gofa *haare*, Gamo, Dorze, Cancha, Kullo *hare*, Malo, Dache, Zayse *hare*, Kachama *haarre*, Ari *harra* ‘donkey’ – likely lws. from Oromo.

¹⁰⁰ Cf. NAA **kawr-* ‘load, carry a load’: Sem.: (a) Arab. *kwr* ‘porter qqch. sur son dos’, II ‘empaqueteur (des bagages)’ (BK II942), (b) ‘pack saddle, bat’: Hbr. *kar* ‘bât (de chameau)’, Arab. *kūr-* ‘selle de chameau’, Tna, Tgr. *kor*,

Sem. **k^wVrr-* and **k^wVrk^wVr-* (redupl.) ‘she-camel’: Hbr. *kirkär-ät* ‘she-camel’ (redupl.) | Arab. *kurr-* ‘ânon, poulain’ (DRS: 1283)¹⁰¹ | Amh. *k^wark^war*, Wolane *kurkur* ‘interjection one uses when calling a donkey’ (redupl.) || Hobyot *kóokar* (redupl.) ‘female camel (from 2 to 3 years)’.¹⁰²

Berb. S. **kirkar* (redupl.): Ayr, E. Wlm. *i-karkar* (pl.) ‘camels’.

Chad. **kwar-*: W. **kwar-* ‘wild ass, donkey’: Chip *kor-o*, Bolewa *kor-o*, Ngizim *kwár-á*, etc. || C. **kwar-* ‘ass’: Tera *kór-ó*, Hwona, Margi, Gude *kwar-a*, Banana *kwár-a*¹⁰³ || E. **kur-* ‘ass’: Ndam *kùr-ō*, Kwang *kúr-ā*, etc.

Cush. **kular-*: N. **kar-*: Beja *kár-ee-see* ‘Kamellaus’ (*see* ‘louse’) || E.: LEC *(?*Vn-*)*kur-*: Somali *kór-on* ‘castrated camel; barren she-camel’, Rendille *in-kuraarr-ú* ‘donkey colt, young donkey’.

Omot. N. **kur-* ‘donkey’: Chara *kur-ā*, Kafa, Mocha, Anfillo *kur-ō*, Gimirra, Dizi (Sheko) *kur*.

► AADB 2468.

2.6. **pirā?* ‘k. of equid’¹⁰⁴

Sem. **para?* ‘onager, wild ass’ [*Equus asinus*, *E. hemionus*]: Akk. (OB on) *parû* ‘Onager, Maultier’ || Hbr. *pärä(?)* | Sab. *fr?* | Arab. *fara?*-.

(?) Egyp. (18 Dyn.) *rnp* (met. < **pr-n?*) ‘young horse’.

Chad. W. **pirr-* ‘k. of equid’: Gerka *pirr-i* ‘mare’ || C. **pur-*: Bura *pur-u* ‘white horse’.¹⁰⁵

Cush. **filar-* ‘k. of equid’: N **firr-*: Beja *firr-at* ‘she-camel just foaled’ || E. **far-*: SA **farar-* (redupl.): Saho, Afar *farar* ‘horse’ | LEC **faraw*: Somali, Arbore *faraw* ‘zebra’ [*Hippotigris sp.*].¹⁰⁶

▲ Seems like a genuine Proto-Afrasian term for a kind of equid.

► SED II No. 176; AADB 2503.

3. Largest herbivores¹⁰⁷

3.1. **da(n)g^w(Vr)-* ‘elephant’ [*L. africana*]

Chad. E. **dugur-*: Gabri *dūgūr-ù*, Sokoro *dogol* (-*l* < **-r?*).¹⁰⁸

Amh. *kor-āča*, Arg. *kor-āčča* ‘selle’ (DRS 1202), Amh. *kārkar* ‘louage des bêtes de somme’ (DRS 1276); Chad. (a) **kVr-*, *kVrkVr-* ‘load, carry a load’ (CED #446): W.: Bolewa, Karekare *karài* ‘load’, Ngizim *kákárâ* ‘load’, *kàkàrú* ‘carry a load’, Bade *karày* ‘goods’, Duwai *kàrkàrà* ‘load carried on the head’, C.: Bura *kari* ‘goods, possessions’, Fali-Kiryá *k^wàr* ‘to unload’, E.: Tumak *gàr* ‘carry; a load’ (*g-* < **k-* is regular), (b) W.: Hausa *ta-karkari* ‘pack ox’.

¹⁰¹ Cf. *kawr-* ‘troupeau nombreux (de chameaux, de bœufs)’.

¹⁰² Note a striking similarity both in form (reduplication) and meaning (she-camel) in Hbr. and Hbt. Cf. also Amh. *kārkar* ‘louage des bêtes de somme’.

¹⁰³ Cf. Nzangi *kəra*, Bata-Zumo *kara* ‘horse’.

¹⁰⁴ There are several (five, to be exact) somewhat similar triconsonantal roots in which it is not easy to distinguish between chance similarity and borrowing from suppletive stems such as Somali *faras* and pl. *fárdó*, Bilin *fardā* and pl. *fáriž*, *bādrā* and pl. *bázir*, Khamir *bázrā* and pl. *baris*. Two of them are well represented in Sem. (**pird-* and **paraš-*), nearly all in E. Cush and all in C. Cush while a few Chad. and Omot. parallels are most likely loanwords; notably, they seem to have no parallels in Beja, S. Cush., Egyp. and Berb. Surprisingly, all of them mean ‘horse’, which apparently excludes their PAA status. This entangled situation is worth a separate piece of research, which the authors are planning to publish soon.

¹⁰⁵ Hardly named after its color, since there seems to be no corresponding color name in the available Bura lexicon.

¹⁰⁶ Omot. N. **far-* ‘horse’: Wolayta, Zala, Dawro, Gofa, Gamu, Chara *far-a* < SA?

¹⁰⁷ Some examples from individual languages are quoted after Blažek 1994; however, distribution of the forms among different entries, based on regular consonantal correspondences, is in most cases quite different from this highly valuable collection of data.

¹⁰⁸ Also *dúger* or *dúkur* ‘rhinoceros’. Cf. Dangla *dugulo*, Migama *dijillu* ‘trunk’ and W. Saharan: Daza *dugugul* ‘trunk of elephant’.

Cush. S. **dang*^w-. Iraqw *dangw*-.

Omot. **dang*-*Vr*-: N.: Zaise *dongor*, Kafa *dangiy-ō* || S.: Ari *dangór*, Hamar *donger*.

► AADB 2519. Cf. Blažek 1994: 199.

3.2. **duhr*- ‘elephant’

Egypt. (late) *dnhr* (met. < **dhr-n?*) ‘elephant’ [*Loxodonta africana*].

Chad. W. **dor-in*- ‘hippopotamus’ [*Hippopotamus amphibius*]: Hausa *dòrìn-a*, Mupun *dòrin-a*, Miya *dòrin-á*, Tsagu *dòrán-à* (the last three words can be lws. from Hausa).

Omot. **duHur*- ‘elephant’ [*L. africana*]: N.: Dizi (Nao, Maji) *dōr*, (Sheko) *door* || S.: Dime *duur-o*.

► AADB 2524; cf. Blažek 1994: 199.

3.3. **fil*- ‘a proboscidean mammal’

Sem. **pīl*- and **pVlpVl*- (redupl.): Akk. (OB) *pīl-u* || pB. Hbr. *pīl* | Aram.: Syr. *pīl-ā*, Mnd. *pīl(-a)* | Arab. *fil*- ‘elephant’ [*Loxodonta africana*] | Gz. *falfal* ‘water buffalo; elephant’ [*Syncerus caffer*; *L. africana*].

Chad. **pīl*-: C. **pīl*-: Bura *pīl-ngir*¹⁰⁹ ‘aardvark’ [*Orycteropus afer*].¹¹⁰

Cush. S. **fil*-: Iraqw *fil-a*, Alagwa *fil-i*, Burunge *fil-u* ‘aardvark’ [*O. afer*] | Qwadza *fil-imbayo* ‘honey-badger’ [*Mellivora capensis*].

► Cf. SED II No. 173; AADB 2528. Cf. Blažek 1994: 196.

3.4. **gumar*-~ **garyam*- ‘hippopotamus’

Chad.: C. **garyam*-¹¹¹ ‘hippopotamus’ [*Hippopotamus amphibius*]: Musgu *géryam*, Vulum, Mbara *gàriyàm*, Masa *gáryam*, Banana *garyam-ba* || E. **girim*-: Tumak *girim*, Birgit *girim-tí*.

Cush. **gumar*-¹¹² ‘hippopotamus’ [*H. amphibius*].¹¹³ C. **gumar-ī*: Khamir *gumār-i*, Khamta *gumār*, Qwara *gumār-ē*, *gumār-ī*, Kemant *gumār-ī*, Aungi *gumār-í* || E. **gumar*-: Saho *gūmār-ē*, Afar *gūmār-ī* | HEC: Qabenna *gomarr-a*, Tembaro *gumorr-a*.¹¹⁴

▲ Cf. NS **ngwír*: ‘large horned or tusk herbivore’ (Ehr. #514) and **ngwí:rént* ‘hippopotamus’ (ibid. #515).

► AADB 2526; cf. Blažek 1994: 202–3.

3.5. **lab?*- ‘k. of largest herbivore’

Egypt. *īrb?* (< **?V-lbV?*?)¹¹⁵ ‘rhinoceros (?)’ [*Diceros bicornis*].

Chad. W. **lab*-: Tangale *lab-ata* || E.: Mokilko *?ēlb-í* (met.?) ‘elephant’ [*Loxodonta africana*].¹¹⁶

¹⁰⁹ Blench, 2009 :141 (draft printout); *ngir* ideof. describes large number of people or birds sitting in one place.

¹¹⁰ Cf. also W. **pīl*: Gwandara *pīl* ‘spear grass’ [*Heteropogon contortus*], Angas *pīl* ‘elephant grass’ [*Pennisetum purpureum*]. It is an important element of forage for elephants in Africa, hence its name “elephant grass”.

¹¹¹ The metathesis is probably due to folk etymology, cf. Chad. C. **gar*- ‘bull’: Musgu *gari* ‘(big) bull’, Muskum *gèrré*, Vulum *gàri* ‘bull’ and PAA **yVm*- ‘water’. Cf. Blažek 1994: 204.

¹¹² Blažek also refers to Militarev’s Ms. from 1987, in which Berb. **a-gumār* ‘horse’ is compared to this Cush. term, referring to a well-known connection between horse and hippopotamus. However, the expected semantic development is reference to a hippo as “(river) horse” metaphorically; the reverse meaning shift is hard to imagine, which makes the Chad-Cush. comparison, already problematic because of metathesis, even less probable.

¹¹³ Borrowed in Eth. as **g^wamāri* ‘hippopotamus’ (note also Gz. *gamar*, Amh. *gumarr-e* ‘large monkey’). Blažek (1994: 202) quotes Gordon’s interpretation of Ugr. *gmr* as ‘a kind of animal capable of fighting ferociously’ (Gordon, 1965: 380), calling it “a promising Semitic cognate”. In Olmo Lete and Sanmartín 2003, however, *gmr* is glossed as ‘champion, fighter’ with Sem. cognates.

¹¹⁴ Cf. Omot. N.: Kullo *gomār-a* ‘hippopotamus’, isolated, likely a lw. from HEC.

¹¹⁵ Otherwise < **?V-rbV?*- comparable to E. Cush. **?arb*- ‘elephant’.

¹¹⁶ Cf. Mokilko *lòllib-ò* ‘trunk’.

Cush. HEC **laʔab-*(met.?): Hadiya, Kambatta *loob-a* ‘hippopotamus’ [*Hippopotamus amphibius*].
 ► AADB 2513. Cf. HSED #1662; Blažek 1994: 203.

3.6. **pir-* ‘k. of largest herbivore’

Sem.: Akk. (OB on) *pīr-u*, *pēr-u* ‘elephant’ [*Loxodonta africana*].¹¹⁷

Chad. **pir-*: W.: Hausa *fyár-mà* ‘a young female elephant’ || C.: Margi *pir*, Sukur *n-ver-i* (< **n-per-i*) ‘elephant’.¹¹⁸

Cush. E. **pury-*: Yaaku *puriaʔ*, pl. *puri-ain* ‘rhinoceros’ [*Diceros bicornis*].

► AADB 4259.

3.7. **warš-* ‘rhinoceros’¹¹⁹

Sem.: Eth. **hawriš-* ‘rhinoceros’ [*Diceros bicornis*]:¹²⁰ Gz. *hariš*,¹²¹ Tna., Tgr. *hariš*, Amh. *haris*, *awraris*.¹²²

Chad. C. **warž-*: Mbara *wí(r)ž-ĩ*, (?) Gidar *walš-ya* (-l < **-r* and **ž* > *š* influenced by l?) ‘bull’.

Cush. E. **worš-* ‘rhinoceros’ [*Diceros bicornis*]: LEC **worš-*: Oromo *worse-sa*,¹²³ Konso *orš-ayta* (also ‘hippo’ [*Hippopotamus amphibius*] acc. to Black), Gidole *orš-ayt*, Dirayta *orš-ayt* (also ‘gnu’ [*Connochaetes sp.*]) | HEC **worš-aʔ-*: Hadiya *oršaʔ-aḏo*, Burji *wórš-a* | Dullay **wVrs-Vʔ-*: Gollango, Harso *oršaʔ-aḏo* | Yaaku *írseʔ*.

▲ The resemblance of all these forms can hardly be accidental, though the origin of Eth. **h-* is far from clear. It could have been explained by **h-* > *Ø-* in some of the non-Semitic languages quoted, but **h* is expected to be preserved as *h* or *h* in Hadiya, Gollango and Yaaku. It could be suggested that the Eth. term may go back to the compound form **hayāw* ‘alive, living’¹²⁴ + **wariš-* paralleling *ʔarwehoras* ‘animal / beast + rhinoceros’ > *ʔawrāhars* > *ʔawrāris* (above). Cf. especially Proto-North Caucasian **wHārχwə* ‘pig, piglet’.

► AADB 2529; cf. Blažek 2003b: 270, Blažek 1994: 201.

3.8. **žag^w-* ‘k. of largest herbivore’¹²⁵

Egyp. (MK) *zʔg.t* (met. < **ʔzg* or *ʔ* renders *-a-*?) ‘ein Fabeltier’.

Chad. W. **(n-)žungw-*(met. < **žag^w-n-*): Diri, Pa’a *žungw-a*, Siri *žinw-a* ‘hippopotamus’ [*Hippopotamus amphibius*], Dwot *n-dzugh-u* ‘elephant’ [*Loxodonta africana*].

Cush. E. **zag-am-*: Yaaku *sogóm-èi* ‘elephant’ [*L. africana*].¹²⁶

Omot. N. **ʔa-zag-ay* ‘hippopotamus’ [*H. amphibius*]: Kačama *azāgē*, Koyra *azzāgē*, *azzagi*, Ganjule *azagé*.

¹¹⁷ Usually regarded as a by-form of *pīl-u* (see in 3.3) due to its sound Sem. etymology, indisputable in light of Gz. *falfal* (cf. discussion in SED II: 228).

¹¹⁸ Cf. W.: Kofyar *feer*, Warji *farai*, *párai*, Karya *pár*, Miya *à-pár* ‘horn’; C.: Daba *fālām* (-l can be < **-r*), Buwal *fár-ám*, etc. ‘horn’.

¹¹⁹ The reconstruction of **-š* (not **-č*) is based on the Mbara form alone.

¹²⁰ There is also Syr. *hars-ūm-ā* ‘proboscis; labia bovis’ which matches the Eth. noun phonetically, but semantically seems not close enough to be related.

¹²¹ Also *haris*, *horas*, *ʔarwehoras*, *ʔawrāhars*, *ʔawrāris*.

¹²² Arab. *hariš-* is a lw. < Gz.; E. Cush.: Saho *hariš* < Mod. Eth.

¹²³ < **worse-ta*; cf. Bayso *wōrsēsa*, borrowed from Oromo.

¹²⁴ < Sem. **hyw/y* ‘to live’ (> Hbr. *hayyā* ‘animals, beasts’, Syr. *hayyūt-*, Hbt. *haywuún* ‘animal’).

¹²⁵ Cf. the similar SAA root **žakan-* ‘elephant’ [*L. africana*]: Cush. E.: Saho, Afar *dakaan-o* (< **dakan-* or **žakān-*), LEC: Somali *dagon* (< **dakVn-* or **žakVn-*), Dahalo *dokom-i*, *ḏokom-i* (< **dVkVm-* or, perhaps, **žVkVm-*; the Dahalo reflex of AA **ž* is not clear — see Takács 2011: 119); Omot. N. **zak(kV)-nV*: Kačama, Koyra *zākk-ā*, Gofa *zākk-i*, Ganjule *zākk-a*, Yemsa *zakn-ō* (the Omot. terms may otherwise have been borrowed from Bantu **nzok* ‘elephant’).

¹²⁶ s in Yaaku may reflect **z* < **ž* / *š*.

► AADB 2521. Cf. HSED 2658; Blažek 1994: 204.

3.9. **žunaʔ*- ‘elephant’

Chad. **žunaʔ*- ‘elephant’ [*Loxodonta africana*]: W. **žawon*-: Ngizim *žàun-àk* || C. **žuwān*-: Tera *žùwàn*, Sukur *žuwān* || E.: **žunaʔ*: Nanchere *ženaʔ*, Gabri *žén-u*, Kabalai, Dormo *žun-o*.

Cush. **žunaʔ*- ‘elephant’ [*L. africana*]: C. **žaHun*- (met.): Bilin, Qwara, etc. *žān-ā*, Khamir *zohón*, Aungi *ziyon-i* || E.: HEC **žanaʔ*-: Darasa *dānaʔ-è*, Kambatta *zan-ó*, etc. || S.: Dahalo *dannab-a*¹²⁷ ‘female elephant’.

► AADB 2522. Cf. HSED #2658; Blažek 1994: 199.

Discussion

Given the above genealogical classification of Afrasian, only those roots whose reflexes are attested in both SEBC and CO branches can be labeled PAA; attestation in Semitic makes this status incontestable. In case of faunistic terms, though, this principle cannot be strictly followed, since the retention of PAA animal names depended on availability of the *denotata* in the area where the respective language speakers had migrated.

There are two opposing views on the ancestral home of the Afrasians. Alexander Militarev’s Levantine hypothesis is supported by Blažek (2013), and a similar view is held by the archaeologist Peter Bellwood (2013). The idea of an African homeland, originally put forward by Igor Diakonoff (1981; later, however, he would lean more towards the Levantine hypothesis, cf. Diakonoff 1998) is promoted in various works by Christopher Ehret (e.g., Ehret 2011) and Roger Blench (e.g., Blench 2006) and, to a certain degree, shared by George Starostin (Starostin 2017).

The hypothesis according to which Proto-Afrasians may be associated with the highly developed Natufian and Post-Natufian Epipaleolithic culture of hunters and gatherers (12,500–9,500 BCE) is based on several main arguments. First, some of the multiple reconstructed terms¹²⁸ point to incipient agriculture rather than foraging.¹²⁹ For such an early period, both are typical of West Asia, the South Levant in particular, rather than of East or North Africa. Second, the whole picture of PAA culture reconstructed from the PAA lexicon (referring to social and economic differentiation, some terms relating to trade, weaponry, building and dwelling, etc.) features a society highly developed for that period which, again, implies the Levant rather than East or North Africa. Third, there are at least several dozens of hypothetical cultural isoglosses between the non-Semitic Afrasian daughter proto-languages (Berber, Chadic, Cushitic) and, on one hand, Sumerian (Militarev 1995) and, on the other, North Caucasian (Militarev & Starostin 2007), speaking in favor of ethnic and cultural contacts between

¹²⁷ If analyzed as *danna-ba*, this would be another case of Dahalo *d* (usually < **d*) as a reflex of **ž/ʒ*, besides *máʕade* ‘in **maʕi(n)ʒ*- ‘k. of ungulate’ (1.5b.)

¹²⁸ Ca. 30 in Militarev 2002; the number has more than doubled with a lot of newly processed data, especially Chadic — publication of a new list is forthcoming.

¹²⁹ Since the 11th millennium is regarded by modern scholarship as too early a period for incipient farming, the following scenario could explain the inventory of terms referring to farming and formally qualifying for Proto-Afrasian status: PSA speakers moved southward from the Levant to the north and, later, south Arabia (from where they moved to East Africa supposedly by or in the 8th mill.), while PNA speakers stayed in the Levant and created the Pre-Pottery Neolithic A culture which set farming in motion in the 10th mill.; some of the farming practices and terms could have passed from them to Proto-South Afrasians who were still located in Arabia, creating the illusion of Proto-Afrasian farming terminology when, in fact, it is representing only the PNA (PSEBC) stratum.

the corresponding speaking groups which could take place in West Asia in the period between the split of PAA and migration of most of its daughter proto-languages to Africa.¹³⁰

Counter-arguments rely on the repertoire of the PAA zoonyms denoting numerous African wild bovids (including most “antelopes”), elephants, rhinoceroses, hippopotami, crocodiles and odd-toed ungulates (except for onager) and large carnivores (lion, leopard, wolf, jackal, etc.). African animals did not inhabit the Levant during the warmer Preboreal phase of Holocene, which includes the Pre-Pottery Neolithic period (ca. 9,500–8,500 BCE, see Table 1), though the names of these animals are technically reconstructed on the PAA level. Regular consonantal correspondences between the reflexes of these roots point to their PAA origin, though the primary meaning of PAA terms is hard to specify considering the length of time that has passed since the PAA split.

Climate of the Levant in the era of the Natufian culture (ca. 12,500–9,500 BCE) was unacceptable for tropical mammals. The Natufians hunted only the following ungulates: gazelles [*Gazella gazella*, *G. subgutturosa*], deer [*Cervus elaphus*, *Damadama*], ibex [*Capra ibex*], aurochs [*Bos primigenius*], wild boar [*Sus scrofa*] and onager [*Equus hemionus*].¹³¹ A corresponding set of wild animal names has been reconstructed for Proto-Semitic (see SED II) whose Urheimat was most likely situated in West Asia, including the Levant, whereas for PAA as many as ca. 50 terms for wild ungulates have been reconstructed, most of them presumably representing various species. Such a variety of ungulate species is typical of the North East African sub-equatorial fauna, which has essentially remained unchanged since the end of Pleistocene.

Many terms denoting tropical African “antelopes”¹³² and largest animals such as elephant, hippopotamus, and rhinoceros, are to be formally considered PAA as well. These PAA roots are usually well reflected in Chadic (less so in Berber and Egyptian), Cushitic, and/or Omotic, but poorly attested in Semitic, whose original habitat being West Asia, the terms for “largest herbivores” sometimes refer to mythical animals (‘monster’, ‘fabulous bull’ and the like), while those for other ungulates usually denote various species of cattle, sheep and goats, asses,¹³³ horses and camels.

Another group of PAA zoonyms relevant for the Urheimat problem is represented by predators. There are 8 reconstructed terms for canines including ‘dog’ (**kawit-*, **k^wihan-*, **KV3/žim-*), ‘k. of wild canine’ (**ʔaway-*, **ʔaw/ys-*, **ba^wiḥ-*, **kur(-ay-)*, **wanš-*), 4 undifferentiated terms ‘k. of canine or hyena’ (**ʔalusk-ay-*, **baʔVš-*, **gVd-^wahr-*), 9 terms for felines including ‘lion’ (**labiʔ-*), ‘leopard or lion’ (**baʔy-*), ‘k. of (wild) feline’ (**ʔariw-*, **baʔis-*, **da(ʔ/y)m-*, **giwar-*, **layč-*, **mary-*, **sawr-*), and one undifferentiated term ‘k. of feline, hyena or viverra’ (**ʔažur-*).¹³⁴ These considerations should rather stimulate searching for the PAA homeland in the East African area, a hypothesis already voiced by several scholars, including Igor M. Diakonoff and others.

¹³⁰ As for Sumerian-Afrasian (non-Semitic) suggested isoglosses, all of them have a PAA origin and, on chronological grounds, can only be loanwords into Sumerian from various intermediate proto-languages of different Afrasian branches. It is possible that another (extinct and unattested) branch of AA in Mesopotamia could function as a substratum under Sumerian; this assumption by Militarev (1995) was accepted as probable by Diakonov (1996).

¹³¹ See Mithen 2003; Lev et al. 2020.

¹³² “Antelopes” comprise a miscellaneous group within bovids, encompassing Old World (mostly African) ruminants that are not bovines, sheep, or goats.

¹³³ In this connection it is worth mentioning that the only species of odd-toed ungulates attested in the Levant is the onager, whereas Africa was home to several species of wild equids.

¹³⁴ The detailed elaboration of these terms is prepared for publication by the same authors as well as reconstruction of other PAA zoonyms including monkeys (4 terms – none in Semitic), other mammals (8 terms, a selection), birds (11, a selection), reptiles (9, a selection), water reptiles and amphibia (4, a selection), fish (5, a selection).

Linguistic families	Disintegration of language families (year BCE) ¹³⁵	Palaeoclimatic periods (Europe, Western Asia)	Palaeoclimatic periods: approximate dates (year BCE)	Significant archaeological cultures in Levant	Archaeological cultures: approximate dates (year BCE)
Pleistocene					
		Oldest Dryas (cold)	12,000–11,600	Kebaran (Epipaleolithic, wildcrafting and hunting)	16,000–10,500
		Bølling-Allerød interstadial (warmer/colder/warmer)	12,600–10,900	Early Natufian (Epipaleolithic, wildcrafting and hunting)	12,000–10,800
Proto-Afrasian	10,500	Younger Dryas ¹³⁶ (colder)	10,900–9,640	Late Natufian ¹³⁷ (Epipaleolithic, developed wildcrafting and hunting)	10,800–9,500

¹³⁵ Approximate glottochronological dates.

¹³⁶ See Platt et al. 2017.

¹³⁷ See Grosman 2013: 622–627.

Linguistic families	Disintegration of language families (year BCE) ¹³⁸	Palaeoclimatic periods (Europe, Western Asia)	Palaeoclimatic periods: approximate dates (year BCE)	Significant archaeological cultures in Levant	Archaeological cultures: approximate dates (year BCE)
Holocene					
Proto-Sem.-Eg.-Berb.-Chadic	9,000	Preboreal (warmer)	9,640–8,500	Pre-Pottery Neolithic A ¹³⁹ (local origin, beginning of agriculture)	9,000–8,000
Proto-Cush.-Omotic	8,800				
Proto-Eg.-Berb.-Chadic	7,800	Boreal ¹⁴¹ (cool, dry, rising temperature)	8,500–6,900	Pre-Pottery Neolithic B ¹⁴⁰ (Anatolian origin, agriculture and beginning of cattle breeding)	9,000–7,000
Proto-Cushitic	7,500				
Proto-North-Central Cushitic	6,800				
Proto-Berb.-Chadic	6,500				
Proto-East Cushitic ¹⁴⁴	6,000				
Proto-Omotic	6,000	Atlantic ¹⁴² (warm, moist, maximum temperature)	6,900–3,700	Pottery Neolithic ¹⁴³ (agriculture and cattle breeding) and several subsequent neolithic cultures (developed agriculture and cattle breeding)	7,000–5,000
Proto-Chadic	5,500				
Proto-South Cushitic ¹⁴⁵	4,800				
Proto-South Omotic ¹⁴⁶	4,600				
Proto-Semitic	4,600				
Proto-North Omotic ¹⁴⁷	4,200				

Table 1. Comparative dates of Proto-Afrasian and Afroasiatic language families, palaeoclimatic periods and synchronous archaeological cultures in Levant

¹³⁸ Approximate glottochronological dates.

¹³⁹ See Chazan 2017: 197, Kuijt & Finlayson 2009: 10966–10970.

¹⁴⁰ See Chazan 2017: 197.

¹⁴¹ See Dansgaard et al. 1989: 532–534.

¹⁴² See Rasmussen et al. 2005.

¹⁴³ See Goring-Morris & Belfer-Cohen 2013.

¹⁴⁴ Including Yaaku.

¹⁴⁵ Including Ma'a and Dahalo.

¹⁴⁶ Including Ongota.

¹⁴⁷ Including Dizi and Mao.

Solid retention of PAA roots denoting representatives of African fauna in Proto-Chadic contrasts with weaker integrity in Egyptian and Proto-Berber and largely relict reflexation in Proto-Semitic. This may be explained by the latitudinal arrangement of the proto-languages descending from PSEBC dialects. The aggregate area of the latter probably stretched from Northeast Africa up to the Levant. The Cushitic and Omotic homelands would be situated in East Africa further to the south than the PSEBC homeland.

Appendix: Alphabetic index of reconstructed roots

- **ʔar(a)w/y-* ‘k. of larger bovid’ 1.1a.
 **ʔar(a)w/y-* ‘k. of lesser bovid’ 1.1b.
 **ʔayl-* ‘k. of lesser bovid’ 1.2.
 **ʔafVw-* ‘k. of lesser bovid’ 1.3.
 **ʔi(n)ʔ-* and **ʔi(n)ʔ-* ‘k. of smaller bovid’ 1.4a.
 **ba/iʔ-* ~ **baʔbaʔ-* ‘k. of lesser bovid (*Tragelaphus silvaticus?*)’ 1.5.
 **balay/w-* ‘k. of lesser bovid’ 1.6.
 **baray-* ~ **bayar-* ‘k. of lesser bovid’ 1.7.
 **čawiʔ-* ‘k. of bovid’ 1.8.
 **čaiw-* ‘k. of lesser bovid’ 1.9a.
 **čawir-* ‘male large bovid or herbivore’ 1.9b.
 **da(n)g^w(Vr)-* ‘elephant’ 3.1.
 **dawk-* ‘k. of equid’ 2.1.
 **duhr-* ‘elephant’ 3.2.
 **(ʔa-)dury-* ‘k. of bovid’ 1.10.
 **fa(?)n-* ‘k. of large bovid’ 1.11.
 **fil-* ‘a proboscidean mammal’ 3.2.
 **(ʔa-)fir-* ‘k. of bovid’ 1.12.
 **ga/ud-* ‘wild pig’ 1.13b.
 **gawr-* ‘k. of bovid’ 1.14a.
 **gayl-* ‘camel’ 2.2.
 **guday-* ‘k. of larger bovid’ 1.13a.
 **gumar-* ~ **garyam-* ‘hippopotamus’ 3.4.
 **gur-* ‘wild pig’ 1.14b.
 **ɣawʔ-* ‘k. of smaller bovid’ 1.15.
 **ħa/ir-* ‘wild ass’ 2.4.
 **ħawr-* ‘k. of larger bovid’ 1.16.
 **ħi/u(l)l-* ‘k. of equid’ 2.3.
 **k^wayr-* ‘young (small) bovid’ 1.17.
 **k^war-* ‘wild ass or camel’ 2.5.
 **labʔ-* ‘k. of largest herbivore’ 3.5.
 **lawiʔ-* ‘k. of large bovid’ 1.18.
 **ma-ʔi(n)ʔ-* ‘k. of bovid’ 1.4b.
 **mayn-* ‘k. of large bovid’ 1.19.
 **mar(ay)-* ‘k. of large bovid’ 1.20a
 **(ʔa-)mayr-* ‘k. of small bovid’ 1.20b.
 **naʔiw-* ‘k. of bovid’ 1.21.
 **pir-* ‘k. of largest herbivore’ 3.6.
 **piraʔ-* ‘k. of equid’ 2.6.
 **saraw/y-* ‘k. of bovid’ 1.22.
 **šaʔ-* ~ **šaʔ-* ‘k. of large bovid’ 1.23.
 **waʔil-* ‘k. of lesser bovid’ 1.24.
 **wad-* ‘k. of lesser bovid’ 1.25.
 **(ʔa-)w/yan-* ‘k. of lesser bovid’ 1.26.
 **war-* ‘k. of larger bovid’ 1.27.
 **warš-* ‘rhinoceros’ 3.7.
 **ʒag^w-* ‘k. of largest herbivore’ 3.8.
 **ʒunaʔ-* ‘elephant’ 3.9.

Abbreviations

Names of languages

AA	Afrasian = Afro-Asiatic	Copt.	Coptic	ESud	East Sudanic
Akk.	Akkadian	Cush.	Cushitic	Eth.	Ethiopic
Amh.	Amharic	Dem.	Demotic	Ez.	Eza
Arab.	Arabic	E. Wlm	East Tawllemmet	Gog.	Gogot
Aram.	Aramaic	Ebl.	Ebla	Gt.	Gyeto
Berb.	Berber	Egyp.	Egyptian	Gur.	Gurage
Bibl.	Biblical Aramaic	End.	Endegen	Gz.	Ge'ez
Ch.	Chaha	Enn.	Ennemor	Har.	Harari
Chad.	Chadic	ESA	Epigraphic South Arabian	Hbr.	Hebrew

Hbt	Hobyot	NE	New Kingdom	Sab.	Sabaic
HEC	Highland East Cushitic	NS	Nilo-Saharan	Sam.	Samaritan
Hrs.	Harsusi	OAKk	Old Akkadian	Sel.	Selti
Jib.	Jibbali	OB	Old Babylonian	Sem.	Semitic
Jud.	Judaic Aramaic	Off.	Official Aramaic	SNil	South Nilotic
LEC	Lowland East Cushitic	OK	Old Kingdom	Soq.	Soqotri
Mand.	Mandaic	PAA	Proto-Afrasian	Sum.	Sumerian
MB	Middle Babylonian	PCO	Proto-Cushitic-Omotiic	Syr.	Syriac
Mhr.	Mehri	Phoen.	Phoenician	Tgr.	Tigre
Min.	Minaean	PSEBC	Proto-Semitic-Egyptian-Berber-Chadic	Tna.	Tigrinya (Tigray)
MK	Middle Kingdom			Ugr.	Ugaritic
Ms.	Masqan	Qat.	Qatabanian	Wol.	Wol.
MSA	Modern South Arabian	S.Cush.	Southern Cushitic	YB	Young Babylonian
Muh.	Muher	SA	Saho-Afar		

Terminology

coll.	collective	k.	kind	redupl.	reduplication
dial.	dialect	lw.	loanword	m.	masculine
dissim.	dissimilation	met.	metathesis	sg.	singular
f.	feminine	pers.	personal	sp.	species
gen.	generic (term)	pl.	plural	subst.	substantive
id.	idem	pref.	prefix	suff.	suffix

Other notational symbols

- || separates branches within a language family
- | separates subbranches or groups
- denotes morphemic boundaries¹⁴⁸

Abbreviations of sources

AADB = Militarev & Stolbova 2020	DRS = Cohen et al. 1970–1993
AHw = von Soden 1965–1981	EDE = Takács 1999, 2001, 2008
BK = Biberstein-Kazimirski 1860	EDS = Parpola 2015
CAD = Brinkman et al. 1956–2010	HSED = Orel & Stolbova 1995
CED = Stolbova 2016	Mil. 2009 = Militarev 2009
	SED II = Militarev & Kogan 2005

References

- Appleyard, David. 2006. *A Comparative Dictionary of the Agaw Language*. Köln: Rüdiger Köppe.
- Bassano, Francesco da. 1918. *Vocabolario tigray-italiano e repertorio italiano-tigray*. Roma: Casa editrice italiana & C. de Luigi.
- Beeston, Alfred F. L., Mahmud Ali Ghul, Walter W. Müller, Jacques Ryckmans. 1982. *Sabaic Dictionary*. Louvain-la-Neuve, Éditions Peeters; Beyrouth, Librairie du Liban.
- Bellwood, Peter. 2013. *First Migrants. Ancient Migration in Global Perspective*. Chichester: Wiley – Blackwell.
- Biberstein-Kazimirski, Albert de. 1860. *Dictionnaire arabe-français*. Vol.1–2. Paris: Maison-Neuve et Co.
- Biella, Joan C. 1982. *Dictionary of Old South Arabic. Sabaean Dialect*. Harvard Semitic Studies, 25. Chico (CA): Scholars Press.

¹⁴⁸ In some cases, especially in Chadic lexemes, boundaries drawn between the stem and affixes and between the affixes may not be exact, since ideal morphemic analysis requires more detailed knowledge of each individual language's morphology and morphophonemics than possessed by the authors.

- Bittner, Maximilian. 1917. *Studien zur Shauri-Sprache in den Bergen von Dofar am Persischen Meerbusen*. Wien: Hölder.
- Blažek, Václav. 1994. Elephant, Hippopotamus and Others: On Some Ecological Aspects of the Afroasiatic Homeland. *Asian and African Studies* 3/2: 196–212.
- Blažek, Václav. 2003a. Semitic *ṭáwar*- 'Bull' and Its Relatives. In: M. L. Bender, D. Appleyard, G. Takács (eds). *Selected Comparative-Historical Afrasian Linguistics Studies in Memory of Igor M. Diakonoff*: 7–12. München: Lincom Europa.
- Blažek, Václav. 2003b. Fauna in Beja Lexicon. In: Ilya Smirnov (ed.). *Studia Semitica: Fs. for Alexander Militarev*: 230–294. Moscow: Russian State University for the Humanities.
- Blažek, Václav. 2013. Levant and North Africa: Afroasiatic linguistic history. In: Peter Bellwood (ed.). *The Encyclopedia of Global Human Migration. Vol. 1. Prehistory*: 125–132. London: Blackwell.
- Blench, Roger. 2006. *Archaeology, Language, and the African Past*. Lanham: Altamira Press.
- Blench, Roger. 2009. *A Dictionary of Bura*. Available at: www.rogerblench.info/Language/Afroasiatic/Chadic/Central/Bura/Bura%20dictionary%202010.pdf [Accessed 28.12.2020].
- Brinkman, John A., Miguel Civil, Ignace J. Gelb, A. Leo Oppenheim, Eruca Reiner (eds.). 1956–2010. *The Assyrian Dictionary of the Oriental Institute, the University of Chicago*. Chicago, Illinois, U.S.A.: Oriental Institute; Glückstadt, Germany: J.J. Augustin Verlagsbuchhandlung.
- Chazan, Michael. 2017. *World Prehistory and Archaeology: Pathways Through Time. 4th edition*. London: Routledge.
- Cohen, David, et al. 1970–1993. *Dictionnaire des racines sémitiques ou attestées dans les langues sémitiques*. Fasc. 1–2. Paris: Mouton. Fasc. 3–5. Leuven: Peeters.
- Conti Rossini, Carlo. 1913. *Schizzo del dialetto Saho dell'Alta Assaorta in Eritrea*. Roma, Tip. della R. Accademia dei Lincei. V. XXII, Fasc. 5.
- Dansgaard, Willi, James W. C. White, Sigfús J. Johnsen. 1989. The abrupt termination of the Younger Dryas climate event. *Nature* 339: 532–534.
- Diakonoff, Igor M. 1981. Earliest Semites in Asia: agriculture and animal husbandry according to linguistic data (VIIIth–IVth millennia). *Altorientalische Forschungen* 8: 23–74.
- Diakonov, Igor M. 1996. Shumery i afrazijscy glazami istorika [Sumerians and Afrasians through the eyes of an historian]. *Vestnik Drevnej istorii* 4: 81–88.
- Diakonoff, Igor M. 1998. The earliest Semitic society. Linguistic data. *Journal of Semitic Studies* 43 (2): 209–219.
- Dolgopolskiy, Aaron B. 1973. *Sravnitel'no-istoričeskaja fonetika kušitskix jazykov* [Comparative-Historical Phonetics of the Cushitic Languages]. Moskva: Nauka.
- Ehret, Christopher. 1980. *The Historical Reconstruction of Southern Cushitic Phonology and Vocabulary*. Berlin: Dietrich Reimer Verlag.
- Ehret, Christopher. 2001. *A Historical-Comparative Reconstruction of Nilo-Saharan*. Köln: Rüdiger Köppe.
- Ehret, Christopher. 2011. *History and the Testimony of Language*. Berkeley: University of California Press.
- Erman, Adolf, Hermann Grapow. 1957–1971. *Wörterbuch der ägyptischen Sprache. I–V*. Berlin: Akademie-Verlag
- Grosman, Leore. 2013. The Natufian Chronological Scheme – New Insights and their Implications. In: Ofer Bar-Yosef, François R. Valla (eds.). *Natufian Foragers in the Levant: Terminal Pleistocene Social Changes in Western Asia*: 622–637. New York: Berghahn Books.
- Johnstone, Thomas M. 1987. *Mehri Lexicon and English-Mehri Word List*. London / New York: Routledge.
- Jungraithmayr, Herrmann, Dymitr Ibrizimow. 1994. *Chadic Lexical Roots. Vols. I–II*. Berlin: Dietrich Reimer Verlag.
- Goring-Morris, A. Nigel, Anna Belfer-Cohen. 2013. The Southern Levant (Cisjordan) During the Neolithic Period. In: Ann E. Killebrew, Margreet Steiner (eds.). *The Oxford Handbook of the Archaeology of the Levant: c. 8000–332 BCE*: 141–163. Oxford University Press.
- Kuijt, Ian, Bill Finlayson. 2009. Evidence for food storage and predomestication granaries 11,000 years ago in the Jordan Valley. In: *Proceedings of the National Academy of Sciences of the United States of America* 106 (27): 10966–10970.
- Lanfry, Jacques. 1973. *Ghadamès. Vol. II. Glossaire (Parler des Ayt Waziten)*. Algérie: Fort-National, Fichier de documentation berbère.
- Leslau, Wolf. 1987. *Comparative Dictionary of Ge'ez (Classical Ethiopic)*. Wiesbaden: Otto Harrassowitz.
- Lev, Ma'ayan, Mina Weinstein-Evron, Reuven Yeshurun. 2020. Squamate bone taphonomy: A new experimental framework and its application to the Natufian zooarchaeological record. *Nature Scientific Reports* 10. Available at: <https://www.nature.com/articles/s41598-020-66301-5> [Accessed 28.12.2020].
- Militarev, Alexander. 1995. Sumerians and Afrasians. *Journal of Ancient History* 2: 113–127.

- Militarev, Alexander. 2002. The prehistory of a dispersal: Proto-Afrasian (Afroasiatic) farming lexicon. In: Peter Bellwood, Colin Renfrew (eds.). *Examining the Farming/Language Dispersal Hypothesis*: 135–150. Cambridge: McDonald Institute for Archaeological Research.
- Militarev, Alexander. 2005. Root extension and root formation in Semitic and Afrasian. *Aula Orientalis* 23(1): 83–130.
- Militarev, Alexander. 2009. Proto-Afrasian Lexicon Confirming West Asian Homeland: Pastoralism. *Journal of Language Relationship* 1: 95–106.
- Militarev, Alexander, Leonid Kogan. 2005. *Semitic Etymological Dictionary*. Vol. II: Animal Names. Münster: Ugarit-Verlag.
- Militarev, Alexander, Olga Stolbova. 2020. *Afroasiatic Etymology Database*. Available at <http://starling.rinet.ru> [Accessed 28.12.2020].
- Mithen, Steven. 2003. *After the Ice: A Global Human History, 20,000–5000 BC*. Cambridge, MA: Harvard University Press.
- Motyliniski, Adolphe de Calassanti. 1904. *La dialecte berbère de R'edamès*. Paris: E. Leroux.
- Naït-Zerrad, Kamal. 1998, 1999, 2002. *Dictionnaire des racines berbères*. Paris-Louvain: Peeters Publishers.
- Nakano, Akio. 1986. *Comparative Vocabulary of Southern Arabic: Mahri, Gibbali and Soqotri*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Olmo Lete, Gregorio del, Joaquín Sanmartín. 2003. *A Dictionary of the Ugaritic Language in the Alphabetic Tradition*. Leiden-Boston: Brill.
- Orel, Vladimir, Olga Stolbova. 1995. *Hamito-Semitic Etymological Dictionary*. Leiden / New York / Köln: Brill.
- Parpola, Simo. 2015. *Etymological Dictionary of the Sumerian Language. Part 1. Lexical Evidence*. Winona Lake, Indiana: Eisenbrauns.
- Platt, Daniel E., Marc Haber, Magda Bou Dagher-Kharrat, Bouchra Douaihy, Georges Khazen, Maziar Ashrafian Bonab, Angélique Salloum, Francis Mouzaya, Donata Luiselli, Chris Tyler-Smith, Colin Renfrew, Elizabeth Matisoo-Smith, Pierre A. Zalloua. 2017. Mapping Post-Glacial expansions: The Peopling of Southwest Asia. *Nature Scientific Reports* 7: 40338. Available at: <https://www.nature.com/articles/srep40338> [Accessed 28.12.2020].
- Rasmussen, Sune O., Bo M. Vinther, Henrik B. Clausen, Katrine K. Andersen. 2005. The Greenland Ice Core Chronology 2005 (GICC05). Early Holocene section. *IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series # 2006–119*. NOAA/NCDC Paleoclimatology Program, Boulder CO, USA.
- Reinisch, Leo. 1890. *Wörterbuch der Saho-Sprache*. Wien: Alfred Hölder.
- Reinisch, Leo. 1902. *Die Somali-Sprache. Wörterbuch*. Wien: Alfred Hölder.
- Skinner, A. Neil. 1996. *Hausa Comparative Dictionary*. Köln: Rüdiger Köppe Verlag.
- Soden, Wolfram von. 1965–1981. *Akkadisches Handwörterbuch*. Wiesbaden: Otto Harrassowitz.
- Starostin, Sergei A. 2000. Comparative-historical linguistics and lexicostatistics. In: Colin Renfrew, April M. S. McMahon, Robert Lawrence Trask (eds.). *Time Depth in Historical Linguistics*. Vol. 1: 233–259. Cambridge: The McDonald Institute for Archaeological Research.
- Starostin, George. 2017. Macrofamilies and agricultural lexicon. Problems and perspectives. In: M. Robbeets, A. Savelyev (eds.). *Language Dispersal Beyond Farming*: 215–233. Amsterdam / Philadelphia: John Benjamins.
- Stolbova, Olga V. 2005. *Chadic lexical database. Issue I: Letters L, N, NY, R*. Moscow: Institute of Oriental Studies, Russian Academy of Sciences.
- Stolbova, Olga V. 2009. *Leksicheskaja baza dannyx po chadskim jazykam. V. 3. Sibil'anty i sibil'antnyje affrikaty*. [Chadic Lexical Database. Issue III. Sibilants and Sibilant Affricates.] Moscow: Institute of Oriental Studies, Russian Academy of Sciences.
- Stolbova, Olga V. 2011. *Leksicheskaja baza dannyx po chadskim jazykam. V. 5. Vel'arnyje*. [Chadic Lexical Database. Issue IV. Velars.] Moscow: Institute of Oriental Studies, Russian Academy of Sciences.
- Stolbova, Olga V. 2016. *Chadic Etymological Dictionary*. Moscow: Institute of Oriental Studies of the Russian Academy of Sciences.
- Stolbova, Olga V. 2019. *Leksicheskaja baza dannyx po chadskim jazykam. V. 5*. [Chadic lexical database. Issue V.] Moscow: Institute of Oriental Studies, Russian Academy of Sciences.
- Takács, Gábor. 1999. *Etymological Dictionary of Egyptian. Vol. 1: A Phonological Introduction*. Leiden / Boston / Köln: Brill.
- Takács, Gábor. 2001. *Etymological Dictionary of Egyptian. Vol. 2: B-, P-, F-*. Leiden / Boston / Köln: Brill.
- Takács, Gábor. 2008. *Etymological Dictionary of Egyptian. Vol. 3: M*. Leiden / Boston: Brill.
- Takács, Gábor. 2011. *Studies in Afro-Asiatic Comparative Phonology: Consonantism*. Berlin: Dietrich Reimer Verlag.

А. Ю. Милитарев, С. Л. Николаев. Праафразийские названия копытных в свете проблемы афразийской прародины

В статье представлена первая порция полного тезауруса праафразийских зоонимов, подготовленного авторами: более ста названий копытных, включая полорогих, свиней, эквидов и верблюдов, а также крупных травоядных (слонов, гиппопотамов и носорогов), отражающие максимально полную (на современном уровне наших знаний) картину этих видов фауны в период накануне распада праафразийского языка (11 тыс. до н.э., по глоттохронологии) и в зоне обитания праафразийцев, две различные точки зрения на локализацию которой – Левант и Восточная Африка – также обсуждаются авторами.

Ключевые слова: афразийская прародина; афразийские (афро-азиатские) языки; этимология зоонимов; реконструкция культурной лексики.

Опыт составления аннотированного списка Сводеша для ведийского языка

Статья посвящена проблемам, возникающим при составлении 110-словного списка Сводеша для ведийского языка. Качественный список такого рода имеет исключительную важность для лексикостатистической работы с индоиранским и индоевропейским материалом, но составление его осложняется стилистической и хронологической неоднородностью ведийских текстов. В качестве основного идиома был выбран язык Атхарваеды, второй из Вед, по той причине, что в нем меньше лаун и в текстах употребляется больше базисной лексики, в том числе соматической. Была применена семантическая спецификация, используемая в ГЛБД (Глобальной лексикостатистической базе данных). При этом текст Атхарваеды сравнивался с Ригведой для получения более надежного результата и исследования языковой эволюции на ранних этапах развития ведийского языка. В результате исследования выяснилось, что (а) 87 элементов списка определяются без проблем; (б) различными для Атхарваеды и Ригvedы являются 6 основ; (в) морфологические различия между текстами Ригvedы и Атхарваеды засвидетельствованы в 3 случаях; (г) наконец, в 10 случаях было несколько весомых кандидатов на заполнение слота, но работа с текстами позволила сделать уверенный выбор базовых терминов. В тексте Атхарваеды всего 4 лауны. В пяти случаях в основной слот вынесено две основы по причине супплетивизма, еще в пяти — по причине синонимии. Проведенное исследование показывает, что ведийский список может быть успешно составлен при опоре на тексты и на строгий семантический стандарт.

Ключевые слова: лексикостатистика; список Сводеша; ведийский язык; Атхарваеда; Ригведа; синонимия; супплетивизм.

Составление аннотированного 110-словного списка Сводеша для ведийского языка представляется актуальной задачей, поскольку ведийский является одним из древнейших индоевропейских языков и подобный список может представлять повышенный интерес для глоттохронологического и лексикостатистического анализа индоевропейской семьи. Наличие такого рода списков для древних индоевропейских языков представляется особенно важным, потому что их использование для классификации представляется приоритетным, хотя и может быть сопряжено с рядом трудностей филологического характера, с проблемой выбора синонимов в поэтических текстах и т. д. Как показывает опыт таких работ, как Kassian 2011, посвященной составлению 50-словного списка Сводеша для древнегреческого языка Геродота, и Starostin 2019, в которой разбираются и сравниваются относительные скорости замены базисной лексики (по 100-словному списку Сводеша) на протяжении языковой эволюции китайского языка, от раннедревнекитайского к классическому древнекитайскому, среднекитайскому и современному китайскому, задача является в целом успешно осуществимой.

Ведийский язык представляет собой наиболее древнюю разновидность древнеиндийского языка. Его памятники состоят, с одной стороны, из метрической части с текстами, записанными стихотворным размером; она представлена четырьмя Ведами (Ригведа, Атхарваеда, Самаведа и — частично — Яджурведа). С другой стороны, в ведийском корпусе есть прозаическая часть, в которую входят самхиты и брахманы (Wacker-

pagel 1896: XI–XXXIV). Наиболее древнее состояние отражено в языке Ригведы, причем I-я и X-я мандалы позднее по составлению. Следующий по архаичности текст — Атхарваведа, далее следует Яджурведа и язык прозы. Поскольку для древней Индии при установлении хронологии существует проблема с абсолютными датами, доступны только приблизительные оценки датировки ведийских текстов, основанные в значительной степени на лингвистических данных текстов в сопоставлении с данными других древних индоиранских и, шире, индоевропейских языков и на сопоставлении исторических данных о древней Индии с реальностью, описываемой в ведийских текстах. Считается, что Ригведа была составлена в середине 2-го тысячелетия до н. э. (Elizarenkova 1989: 435–436), а Атхарваведа — несколько позже; так, например, М. Витцель датирует ядро Атхарваведы временем *post quem* 1200 до н. э. (Witzel 1997: 281). Для ведийского списка как эталон текста мы выбираем Атхарваведу, представленную двумя основными редакциями, Пайппалада (Paippalada) и Шаунака (Śaunaka).

Атхарваведа — второй по древности ведийский текст после Ригведы, датируемый приблизительно рубежом 2-го и 1-го тысячелетий до н. э. Тексты Атхарваведы предпочтительнее, поскольку содержат заклинательные гимны и лучше отражают реальное употребление лексики, в том числе соматической. Практика сбора списка также показала, что в тексте Атхарваведы меньше лакун. Вместе с тем при составлении ведийского списка данные Ригведы невозможно игнорировать, и сравнение наиболее древнего состояния ведийского языка с более поздним, представленным в Атхарваведе, неизбежно становится одной из задач при составлении 110-слового списка Сводеша для ведийского языка.

Основные лексикографические источники, которые были использованы при составлении списка — Большой Петербургский словарь санскрита (BR 1855–1875), словарь языка Ригведы Г. Грассмана (Grassmann 1873) и незаконченный лексикон языка Ригведы (RIVELEX). Для нахождения и подсчета словоформ в текстах Ригведы и Атхарваведы были привлечены конкордансы Lubotsky 1997, Kim 2020, Whitney 1881. Для ведийской литературы, датирующейся временем после РВ и АВ¹, использованы конкордансы Bloomfield 1906 и Vishva Bandhu 1935–1965. Для указания форм глаголов в списке использованы работы Bloomfield 2008, Whitney 1885, также для презентных форм — Gotō 1987 и для перфектных форм — Kümmel 2000.

При этом составление списка для ведийского языка наталкивается на ряд проблем, возможность успешного разрешения которых и хотелось бы продемонстрировать в рамках этой работы. В ряде случаев присутствует разветвленная синонимия, и решение о вынесении того или иного слова в список должно приниматься на основании тщательного филологического анализа. В силу ограниченности имеющегося корпуса неизбежны лакуны в списке и немногочисленность контекстов для ряда лексем.

Применяемая семантическая спецификация основывается на работе (Kassian et al. 2010), в которой для каждого слота 110-слового списка Сводеша приводится ряд диагностических контекстов. В статье для сомнительных случаев будут указаны контексты и приведены дополнительные критерии выбора из имеющихся возможных кандидатов на заполнение основного слота списка.

Принципы Глобальной лексикостатистической базы данных (ГЛБД) описаны в Starostin 1998–2005; также можно отметить описание в работе Kassian 2011: 65–66.

В работе будут указаны слова, которые бесспорно являются основными как в Атхарваведе, так и в Ригведе (или только в Атхарваведе, что значительно реже), при этом для

¹ Она привлекается только в тех случаях, когда соответствующей лексемы или основы нет в собственно четырех Ведах (самхитах).

списка этих случаев будут даны только необходимые краткие комментарии. В основном используется редакция Шаунака (Śaunaka), поскольку существует полное издание текстов (Roth-Whitney 1856–1857; Orlandi 1992), соответственно, примеры из текстов приводятся по этой редакции. Языковые данные Пайппалады используются частично; прежде всего используется 5-я мандала по изданию Lubotsky 2002. Встречаемость исследуемых лексем в корпусе устанавливается по конкордансу Kim 2020.

При описании глаголов указываются основы презенса, аориста и перфекта с опорой на (Gotō 1987²; Kümmel 2000³; Whitney 1885), при супплетивизме в основной слот выносятся все дополняющие друг друга основы. Переводы текстов Атхарваеды приведены по изданию Elizarenkova 2005–2010, Ригведы — по Elizarenkova 1989–1999. В отдельных случаях были использованы для сверки значений классический перевод Гельднера (Geldner 1951), современные переводы Джемисон-Бреретон (Jamison-Brereton 2014) и Витцеля и др. (Witzel et al. 2007).

Отдельно будут подробно разобраны сомнительные случаи (в которых, например, приходилось делать выбор из ряда синонимов), а также все те случаи, когда для Ригведы основное обозначение в том или ином слоте отличается от устанавливаемого для Атхарваеды. Также особняком будут указаны позиции, в которых тексты Ригведы и Атхарваеды демонстрируют только морфологическую разницу.

Фонетический облик для выносимых в слот основ указывается по стандарту ГЛБД (GLD) и по традиционному стандарту записи древнеиндийских текстов (IAST — International Alphabet of Sanskrit Transliteration); в комментарии для простоты дана только традиционная запись. Взаимные соответствия транскрипций даются ниже, в таблице 1.

IAST	GLD	IAST	GLD	IAST	GLD
a	a	k	k	th	t ^h
ā	a:	kh	k ^h	d	d
i	i	g	g	dh	d ^h
ī	i:	gh	g ^h	n	n
u	u	ñ	ŋ	p	p
ū	u:	c	č	ph	p ^h
ṛ	ṛ	ch	č ^h	b	b
ṝ	ṛ:	j	ǰ	bh	b ^h
ḷ	ḷ	jh	ǰ ^h	m	m
ḹ	ḷ:	ñ	ɲ	y	y
e	e:	ṭ	ṭ	r	r
ai	a:y	ṭh	ṭ ^h	l	l
o	o:	ḍ	ḍ	v	w
au	a:w	ḍh	ḍ ^h	ś	ʃ
ṁ	Ṃ	ṇ	ṇ	ṣ	ʂ
ḥ	h	t	t	h	ɦ

Таблица 1. Соответствия между транскрипциями IAST и GLD

² Материалы данной работы используются для аннотации глаголов, имеющих презенс I класса или таких, которые не являются презенсами I класса, но считались таковыми по традиции.

³ С опорой на эту монографию указываются перфектные основы.

Гласные могут иметь слабый носовой призвук, что отмечается в традиционной записи знаком *m̐* (анусвара), а в ГЛБД — \tilde{V}^A .

Однозначные случаи

1. all / весь: *w'if-wa- {vísva-}*

Бесспорно основное слово, имеющее значение ‘весь’, как для Ригведы, так и для Атхарваведы, четко отличающееся от *sárva-* ‘целый (entire, whole)’. Основа *vísva-* обладает полисемией ‘весь / всякий, каждый’, ср. RV X, 60, 6 «Ты наступил на Пани, на **всех** скупцов, о царь», англ. «You have trodden down the niggards, trodden on **all** the ungenerous ones, o king», а значение ‘всякий, каждый’ можно проследить, например, в RV I, 100, 7 «Он один управляет *каждым* действием», англ. «he alone is lord of **every** undertaking» (Jamison-Brereton).

2. ashes / пепел, зола: *'a:s-a- {ása-}*

Некоторая проблема состоит в том, что данное слово употреблено только один раз в Атхарваведе (Шаунака) в контексте AVŚ IX, 8, 10 «пусть пеплом станет баласа, мочой болезнетворной!» и два раза в Пайппаладе: AVP XVI, 30, 8; AVP XVI, 74, 10. Несмотря на неясное контекстуальное значение, можно утверждать, что слово должно было значить ‘зола; пепел’. В пользу этого говорят этимологические связи данного слова: оно производно от корня **h₂eh₁s-* ‘жечь / сушить’, как и, например, др.-в.-нем. *asca* ‘зола’; также в ряде дардских и нуристанских языков засвидетельствованы слова с подходящим значением, произведенные от данной основы, ср. кати *asə* ‘пепел’ (EWAia I: 182–183). Более поздние употребления в ведийском также подкрепляют эти доводы, ср. ŚBr. XII, 4, 1, 4. *āsāḥ pāṃsavah* ‘пеплы (золы)-пески’.

3. bark / кора: в Ведах (самхитах) не засвидетельствовано. В более поздних памятниках ср. *tvác-* с полисемией ‘кожа / шкура / кора’ (BR 1861: 463–464). Также стоит отметить термин *bāhya-* ‘внешняя кора дерева’, употребленный в таком значении в ŚBr. III, 7, 1, 8; первое значение — собственно ‘внешний’ (BR-5 1868: 84).

5. big / большой: *maḥ-'ant- {mahánt-}*

Также более архаичные варианты *māh-* и *mahí-* < **megh₂-*. Сочетается с максимальным количеством объектов.

7. bite / кусать: *dāḥ- {daṃś-}*

Традиционно относится к I классу, но содержит нулевую ступень и относится к так называемым аористным презентам (Aorist-Präsens) (Gotō 1987: 74; Hill 2007: 172–173). Финитные формы этого глагола плохо представлены в Ведах (самхитах), но причастные употребления безусловно доказывают, что это основное обозначение. Так, в Ригведе засвидетельствованы презентное причастие *dásant-* (RV I, 189, 5) и причастие интенсива *dándaśāna-* (RV X, 95, 9); и в Ригведе, и в Атхарваведе представлены перфектное активное причастие *dadaśvāṃs-* (AVP III, 16, 6; RV IV, 38, 6) и пассивное перфектное причастие *daśtá-* (AVP III, 16, 6; AVP XX, 14, 9). Кроме того, в более поздних ведийских и послеведийских текстах встречаются финитные формы этого глагола, например, 3 л. мн. ч. *dásanti*

⁴ В данном случае сознательно игнорируется фонетическая разница между анусварой и анунастикой (последняя означает нализацию предшествующей гласной) в связи с тем, что анунастика в рассматриваемых формах не встречается.

(ŚBr. VII, 4, 1, 29). Всего в Атхарваведе Пайппаладе 9 употреблений форм этого глагола, а в Шаунаке — 2 (Kim 2020: 329), в Ригведе — 4 употребления (Lubotsky 1997: 655).

8. black / черный: *kṛṣ-ṛ'a-* {*kṛṣṇá-*}

9. blood / кровь: *'as-ṛ-ṣ-* {*ásrj-*}

10. bone / кость: *'ast^h-i-* {*ást^hi-*}

12. burn tr. / жечь перех.: *dañ-* {*dah-*}

Презенс (I) 3 ед. *dáhati*, корневой атематический аорист 3 ед. *adhāk* (Gotō 1987: 174–176) Корень *kṣā-*, родственный на праязыковом уровне *dah⁻⁵*, встречается только один раз в Ригведе (редуплицированный аористный инфинитив), 4 раза в Пайппаладе и 2 — в Шаунаке в сходных контекстах (Kim 2020: 235).

13. nail / ноготь: *nak^h-a-* {*nakhá-*}

14. cloud / облако: *n'ab^h-as-* {*nábhas-*}

15. cold / холодный: *ṣi:-t'a-* {*śítá-*}

16. come / приходить: *gam-* {*gam-*}

Презенс (с суффиксом *-cha*) *gáchati*, корневой тематический аорист 1 ед. *ágatam*, 3 ед. *ágan*, перфект 3 ед. *jagáma*, 3 мн. ч. *jagmús* (Gotō 1987: 73; Whitney 1885: 34; Kümmel 2000: 155–160). Относится традиционной индийской грамматикой к I классу (Whitney 1885: 34), но такая классификация смешивает собственно презенсы с полной или продленной ступенью корня с другими образованиями (Gotō 1987: 73).

17. die / умирать: *mar-* {*mar-*}

Презенс (IV) 3 ед. *mriyáte* (с вторичным ударением как у пассива, возможно, вызванным неудобством прямого отражения последовательности *CṛiV-, см. работу Kulikov 1997); корневой атематический аорист 3 ед. *amṛta*; перфект 3 ед. ч. *tamára*, 3 мн. ч. *tamrús* (Gotō 1987: 82; Kümmel 2000: 370–372). Зафиксирована форма *máрати*, которая в литературе считалась I классом презенса (Whitney 1885: 124), однако П. Тедеско показал, что основа *mára-*, встречающаяся в Ригведе, функционально соответствует футуральной основе *marīṣya-*, соответственно, представляет собой субъюнктив корневого аориста, употребляемый в значении будущего времени (Tedesco 1944: 212–213).

18. dog / собака: *ṣwan-* {*śván-*}

19. drink / пить: *pa:-* {*pā-*}

Презенс 3 ед. *pípati* I←III / I^{III} (в нотации Kulikov 2008, 2017) с тематизацией изначально атематического редуплицированного презенса (Gotō 1987: 72), (II) 3 мн. *pānti* (RV II, 11, 14); аорист 3 ед. *apāt*, перфект 3 ед. *parāi*, 3 мн. *parús* (Whitney 1885: 95–96; Kümmel 2000: 308–309).

20. dry / сухой: *ṣuṣ-ka-* {*śúṣka-*}

21. ear / ухо: *k'arṇ-a-* {*kárṇa-*}

23. eat / есть: *ad-* {*ad-*} и *aṣ-* {*aś-*}

Невозможно уверенно сделать выбор между этими двумя корнями с презентными основами: (II) 3 ед. *átti*, 3 мн. *adánti* и (IX) 3 ед. *ásnāti*, 3 мн. *ásnánti* соответственно.

⁵ Корень *dah-* отражает праформу **d^heg^{wh}-*, в то время как *kṣā-* выводится из **d^hg^{wh}-eh₁* (LIV²: 133–134; EWAia I: 712–713).

Авторы словаря *Rigveda-Lexikon*, следуя традиционному взгляду, считают, что здесь налицо следующий супплетивизм: корень *ghas-* ‘поглощать’ служит аористом, а деноминативная основа *āvaya-* образует имперфект (RIVELEX 1: 117). Тем не менее, формы этих глаголов редки в Атхарваведе и встречаются в основном в местах, имеющих параллели в Ригведе (Whitney 1881: 51, 103). Поэтому я предпочитаю не включать их как отдельные корни.

24. egg / яйцо: *a:ṇḍá-* {*āṇḍá-*}

25. eye / глаз: *akṣ-* {*akṣ-*}

Две основы, *ákṣi-* и *akṣán-*.

26. fat n. / жир: *p'i:-was-* {*p'ívas-*}

27. feather / перо: *par-ṇ'a-* {*parṇá-*}

28. fire / огонь: *agn-i-* {*agní-*}

29. fish / рыба: *m'atsy-a-* {*mátsya-*}

30. fly / летать: *pat-* {*pat-*}

Презенс (I) 3 ед. *pátati*, 3 мн. *pátanti*, аорист 3 ед. *ápaptat*, перфект 3 ед. *parpāta*; также *áya-*основа *patáyati*. Как показал Л. Куликов, образования *pátati* и *patáyati* противопоставлены как предельный и неопредельный глаголы, соответственно, они параллельны русской паре ‘лететь’ — ‘летать’ (Kulikov 2008). В применяемой спецификации при наличии противопоставления значений ‘летать’ и ‘лететь’ предлагается брать соответствующие слова как синонимы (Kassian 2010: 60).

31. foot / нога: *pad-* {*pád-*}

32. full / полный: *pu:r-ṇ'a-* {*pūrṇá-*}

33. give / давать: *da-* {*dā-*}

Презенс (III) 3 ед. *dádāti*, аорист 3 ед. *ádāt*, перфект 3 ед. *dadaú* (Whitney 1885: 71–72).

34. good / хороший: *w'as-u-* {*vásu-*}

Синоним *bhadrá-* имеет первое значение ‘благой, счастливый’ и часто упоминается по отношению к божествам (Lubotsky 1997: 975–976; Kim 2020: 491–492).

35. green / зеленый: *h'ar-i-ta-* {*hárīta-*}

36. hair / волосы: *k'e:f-a-* {*kéśa-*}

37. hand / рука: *h'as-ta-* {*hástā-*}

39. hear / слышать: *fraw-* {*śrav-*}

Презенс (V) 3 ед. *śṛṇóti*, *śṛṇuté*, корневой атематический аорист 1 ед. *áśravam*, перфект 3 ед. *śúśrāva* (Whitney 1885: 179–180). Форма 2 ед. *śrosi* представляет собой изолированный *-si-*императив к сигматическому аористу (Cardona 1965; Vammesberger 1982/3; Shields 1997).

40. heart / сердце: *h'rd-aya-* {*hṛdaya-*} ~ *hṛd-* {*hṛd-*}

41. horn / рог: *śṛṅga-* {*śṛṅga-*}

42. I / я: *ah'am-* {*ahám-*} ~ *ma* {*ma*}

Основа косвенных падежей *ma-* {*ma-*}.

Авторы словаря *Rigveda-Lexikon*, следуя традиционному взгляду, считают, что здесь налицо следующий супплетивизм: корень *ghas-* ‘поглощать’ служит аористом, а деноминативная основа *āvaya-* образует имперфект (RIVELEX 1: 117). Тем не менее, формы этих глаголов редки в Атхарваведе и встречаются в основном в местах, имеющих параллели в Ригведе (Whitney 1881: 51, 103). Поэтому я предпочитаю не включать их как отдельные корни.

24. egg / яйцо: *a:ṇḍ'-a-* {*āṇḍá-*}

25. eye / глаз: *akṣ-* {*akṣ-*}

Две основы, *ákṣi-* и *akṣán-*.

26. fat n. / жир: *p'i:-was-* {*p'ívās-*}

27. feather / перо: *par-ṇ'a-* {*parṇá-*}

28. fire / огонь: *agn-'i-* {*agní-*}

29. fish / рыба: *m'atsy-a-* {*mátsya-*}

30. fly / летать: *pat-* {*pat-*}

Презенс (I) 3 ед. *pátati*, 3 мн. *pátanti*, аорист 3 ед. *ápaptat*, перфект 3 ед. *rapāta*; также *áya-*основа *patáyati*. Как показал Л. Куликов, образования *pátati* и *patáyati* противопоставлены как предельный и неопредельный глаголы, соответственно, они параллельны русской паре ‘лететь’ — ‘летать’ (Kulikov 2008). В применяемой спецификации при наличии противопоставления значений ‘летать’ и ‘лететь’ предлагается брать соответствующие слова как синонимы (Kassian et al. 2010: 60).

31. foot / нога: *pad-* {*pád-*}

32. full / полный: *pu:r-ṇ'a-* {*pūrṇá-*}

33. give / давать: *da-* {*dā-*}

Презенс (III) 3 ед. *dádāti*, аорист 3 ед. *ádāt*, перфект 3 ед. *dadaú* (Whitney 1885: 71–72).

34. good / хороший: *w'as-u-* {*vásu-*}

Синоним *bhadrá-* имеет первое значение ‘благой, счастливый’ и часто упоминается по отношению к божествам (Lubotsky 1997: 975–976; Kim 2020: 491–492).

35. green / зеленый: *h'ar-i-ta-* {*hárīta-*}

36. hair / волосы: *k'e:f-a-* {*kéśa-*}

37. hand / рука: *h'as-ta-* {*hástā-*}

39. hear / слышать: *fraw-* {*śrav-*}

Презенс (V) 3 ед. *śṛṇóti*, *śṛṇuté*, корневой атематический аорист 1 ед. *áśravam*, перфект 3 ед. *śúśrāva* (Whitney 1885: 179–180). Форма 2 ед. *śrosi* представляет собой изолированный *-si-*императив к сигматическому аористу (Cardona 1965; Vammesberger 1982/3; Shields 1997).

40. heart / сердце: *h'rd-aya-* {*hṛdaya-*} ~ *hṛd-* {*hṛd-*}

41. horn / рог: *śṛṅga-* {*śṛṅga-*}

42. I / я: *ah'am-* {*ahám-*} ~ *ma* {*ma*}

Основа косвенных падежей *ma-* {*ma-*}.

‘красный’. Она переводится как ‘красный, кровавый’, а в позднейшем языке начинает использоваться для обозначения ‘крови’ (BR-6 1871: 386–387; EWAia II: 453). Она употребляется 5 раз в Пайппаладе (из них 3 раза в одном гимне) и 1 раз в Шаунаке (Kim 2020: 597).

67. road / дорога: *pat^h-* {*path-*}

Основы: *pánthā-* / *pathí-* / *path-*. Им. ед. ч. *pánthās*, лок. ед. ч. *pathí*, дат.-твор. мн. ч. *pathíbhī-*.

68. root / корень: *m^hu:l-a-* {*mūla-*}

69. round / круглый: не засвидетельствовано.

70. sand / песок: *pā:s-u-* {*pāmsú-*}

и *s^hik-a-ta-* {*síkatā-*}

Данные слова употребляются как синонимы, которые сложно различить из-за малочисленности контекстов: *pāmsú-* употребляется по 2 раза в Пайппаладе и Шаунаке (Kim 2020: 431), а *síkatā-* по 3 раза соответственно (Kim 2020: 758). В контексте AVŚ VII, 109, 2 они употребляются в одном стихе, но на его основании нельзя провести четкое разграничение: *g^hṛtām apsarāb^hyo vaha tvām agne pāmsūn akṣéb^hyaḥ síkatā aprás ca* ‘О Агни, отвези ты жир для (нимф) апсарас, Пыль для игральных костей, песок и воду!’

71. say / сказать: *braw-* {*brav-*}

Презенс (II) 3 ед. *brāvōti*, 3 мн. *bruvānti*, имперфект 3 ед. *ābravīt* (Whitney 1885: 107).
wač- {*vac-*}

Редуплицированный тематический аорист 3 ед. *avocat* с гласной удвоения *-a-*. Также от данного глагола употребляется редуплицированный презенс (III) *vivakti* и перфект 3 ед. *vavāsa* и *ivāsa*, 3 мн. *ucús* (Kümmel 2000: 441–444).

Отличается от *vad-* ‘говорить; сказать; повествовать; восхвалять’ (BR-6 1871: 649–657; Grassmann 1873: 1199–1201), встречающегося в основном в первом и третьем значениях и часто обозначающего приподнятую, торжественную речь.

72. see / видеть: *paḥ-* {*paś-*}

Презенс (IV) 3 ед. *páśyati*.

darḥ- {*darś-*}

Перфект *dadárśa*, сигматический аорист *adrkṣ-* (Whitney 1885: 78; Grassmann 1873: 62; Kümmel 2000: 231–233).

73. seed / семя: *bⁱ:ž-a-* {*bíja-*}

74. sit / сидеть: *a:s-* {*ās-*}

Презенс (II) *áste*. Корень *sad-* имеет значение ‘сесть, садиться’, соотносится с *ās-* как инхоативный. На английский последовательно переводится как ‘sit down’.

75. skin / кожа: *twač-* {*tvāc-*}

78. smoke / дым: *d^hu:-m^a-* {*dhūmá-*}

79. stand / стоять: *st^ha-* {*sthā-*}

Презенс 3 ед. *tīṣṭhati* I←III / III^h (в нотации Kulikov 2008, 2017) с тематизацией изначально атематического редуплицированного презенса (Gotō 1987: 72), корневой атематический аорист 3 ед. *ásthāt*, перфект 3 ед. *taṣṭāú* (Whitney 1885: 194–195; Kümmel 2000: 580–582).

81. stone / камень: *‘af-man-* {*ásman-*}

82. sun / солнце: *s^hu:r-ya-* {*sūrya-*}

83. swim / плавать: *plaw-* {*plav-*}

Презенс (I) 3 ед. *pl'aw-a-te* (Gotō 1987: 214–215).

85. that / тот: *sa-* {*sá-*} / *ta-* {*tá-*}

Данное супплетивное местоимение является самым употребительным дейктическим местоимением в ведийских текстах и имеет значение ‘тот; он, она, оно’ (Gotō 2013: 67–69). Обычно «отсылает к чему-то уже известному как к только что упомянутому или к общеизвестному» (Macdonell 1916: 293). Для более конкретного дистального дейксиса служит местоимение *asáu-*.

as-'a:w {*asáu-*}

Имеет значение ‘вот тот’. У этого местоимения три формы: *asáu-*, *adás* и *amú-*; две последние используются в косвенных падежах (Gotō 2013: 71–73). По мнению исследователей, все эти основы так или иначе восходят к праязыковому дейктическому местоимению **u* ‘там’ в полной ступени **ew-* (возможно, новой) с различными наращениями, подробный сценарий и литература могут быть найдены в работах (Wackernagel 1930: 530; EWAia I: 98; Gotō 2013: 71–73; LIPP 2: 191, 821).

86. this / этот: *a-* {*a-*} ~ *i-* {*i-*}, *ay-* {*ay-*}

Местоимение значит ‘вот этот’ (в непосредственной близости к говорящему) и противопоставлено *asáu-* (Gotō 2013: 70–71).

87. thou (you) / ты: *tuw-'am* {*tuváam*} ~ *tw-am* {*tvám*}

88. tongue / язык: *žih-w-'a-* {*jihvá-*}

91. two / два: *dwa-* {*dvá-*} ~ *duw'a-* {*duvá-*}

92. walk (go) / идти: *i-* {*i-*}

Презенс *éti*, 3 мн. *yánti*, перфект 3 ед. *iyāya* (Kümmel 2000: 98–100).

ga- {*gā-*}

От этого корня образуется корневой атематический аорист 3 sg. *agāt* и более редкий редуцированный презенс (III) *jígāti* (Whitney 1885: 35).

93. warm (hot) / теплый (горячий): *uṣ-ṇ'a-* {*uṣṇá-*}

Данное слово обладает полисемией ‘теплый / горячий’.

95. we / мы: *way-'am* {*vayám*}

Косвенная основа *as-m'a-* {*asmá-*}.

96. what / что: *ka-d* {*kád*}

97. white / белый: *ṣwe:t-'a-* {*śvetá-*}

В Атхарваведе есть также основа *śvitrá-* ‘белый / светлый’, производная от того же корня (BR-7 1875: 422; Grassmann 1873: 1435; Kim 2020: 698), но она не выступает в качестве базового слова ‘белый’.

98. who / кто: *ka-* {*ká-*}

100. yellow / желтый: *ñ'ar-i-ta-* {*háríta-*}

101. far / далеко: *du:r-'e-* {*dūré-*} и *a:r-'e-* {*āré-*}

Невозможно уверенно сделать выбор между этими двумя словами. Второе слово встречается только в Ригведе и Атхарваведе и в постведийский период практически не представлено (BR-1 1855: 686, 689).

102. heavy / тяжелый: *gur-'u-* {*gurú-*}

103. near / близко: *'an-ti-* {*ánti-*}

104. salt / соль: *law-a-ṅa-* {*lavaṅá-*}

(только AV, употребляется дважды в Пайппаладе и один раз в Шаунаке (Kim 2020: 605)).

105. short / короткий: не засвидетельствовано. В Ведах есть основа *kr̥dhú-* ‘укороченный / недостаточный’, она один раз употреблена уже в Ригведе (Lubotsky 1997: 447) и один раз в Атхарваведе (Kim 2020: 227); позднее появляется слово *hrasvá-* ‘короткий; небольшой’, оно употребляется, например, в контекстах ŚBr. I, 4, 4, 7; III, 7, 2, 7; TS II, 6, 11, 5; VI, 6, 4, 2.

106. snake / змея: *'aṅi-i-* {*áhi-*}

107. thin / тонкий: *tan-'u-* {*tanú-*}

Данная основа засвидетельствована впервые только в брахманах (EWAia I: 620), но других кандидатов в принципе нет, а его безупречная индоевропейская этимология < *t_ṅH-u- (EWAia I: 620–621) указывает на древность образования.

108. wind / ветер: *w'a:-ta-* {*vāta-*}

109. worm / червь: *kṛ-m-i-* {*kṛmi-*}

(надежно только AV)

110. year / год: *sā=vat-sa-r'a-* {*saṃvatsará-*}

Таким образом, по тексту Атхарваеды для 110-словного списка Сводеша представляется возможным установить 87 основ, что можно считать солидным показателем. При этом во всем списке есть только четыре лакуны: 3. bark / кора; 50. louse / вошь; 69. round / круглый; 105. short / короткий. Они, очевидно, объясняются ограниченностью корпуса ведийских текстов. Наконец, только для одного слота, 107. thin / тонкий, была взята основа, встречающаяся в более поздних ведийских текстах, но такой выбор оправдан отсутствием любых других кандидатов.

На следующем этапе целесообразно рассмотреть три типа случаев:

- 1) разные лексемы в Ригведе и Атхарваведе;
- 2) морфологические расхождения между Ригведой и Атхарваведой;
- 3) случаи, когда было сложно сделать выбор между синонимами.

1) Расхождения между Ригведой и Атхарваведой

6. bird / птица: RV *wi-* {*ví-*} / *we:-* {*vé-*}, *way-* {*váy-*} — AV *pakṣ-'in-* {*pakṣín-*}

В Ригведе основным обозначением служит основа *ví-* / *vé-*, *váy-* ‘птица’. Она имеет надежную индоевропейскую этимологию, родственна др.-арм. *haw* ‘птица; петух; курица’, греч. гом. *αἰετός* ‘орел’, лат. *avis* ‘птица’, ср.-ирл. *ai* ‘лебедь’, валл. *hwyaad* ‘утка’ и т. д. и восходит к праформе *h₂wei- (EWAia II: 507–58).

ví- / *vé-*, *váy-* встречается также и в Атхарваведе, но как архаизм: все контексты за исключением одного параллельны контекстам Ригведы (Kim 2020: 615; Whitney 1881: 268). Независимое от Ригведы употребление содержится в следующем месте: AVŚ VI, 2, 2 *á yám viśántíndavo váyō ná vṛkṣám ánd^hasaḥ* «(он тот), в кого входят капли Сомы, как птицы (устраиваются) на дереве».

Для Атхарваеды же основным определенно является слово *pakṣín-* с полисемией ‘пернатый / птица’. Внутренняя форма данного слова прозрачна — это субстантивация прилагательного с продуктивным суффиксом *-in*, образованное от *pakṣá-* ‘крыло / плечо,

бок (человека)' (BR-4 1865: 344–346; Grassmann 1873: 757), не имеющего надежной этимологии за пределами индоиранской группы языков (EWAia II: 62).

В индоарийских языках послеведийского периода преобладают рефлексы именно этой основы, ср. (Turner 1962–1985: 7636) с большим количеством примеров.

В данном случае речь идет о замене слова с надежной индоевропейской этимологией на эпитет. Подобный же процесс происходил в иранских языках: рефлексы родственной праиранской основы **wi-* / **way-* 'птица' отражаются в авест. *vi-*, *vay-* 'птица' и в среднеперсидском *way* 'птица' (MacKenzie 1986: 89), но в большинстве среднеиранских и современных иранских языков данная основа вытеснена другими, в том числе **mrga-* (ESIJ 5: 379–383).

11. breast (грудь) RV *w'akṣ-as-* {*vákṣas*} — AV *'ur-as-* {*úras-*}

В Ригведе основным словом для обозначения 'груди', как мужской, так и женской (но не в смысле 'грудь, которой кормят детей', ср. *stána-* 'женская грудь'), следует признать *vákṣas-*. Оно встречается в тексте 9 раз (Lubotsky 1997: 1215), в то время как *úras-* только два раза, причем в поздних мандалах (RV I, 158, 5; RV X, 155, 4). Можно привести следующие примеры, доказывающие обоснованность именно такого выбора: RV I, 64, 4 «на **груди** они (Маруты) носят золотые пластинки для блеска»; RV I, 92, 4 «она обнажает **грудь**, словно корова — вымя».

Основа *úras-* же явно становится основным обозначением груди в Атхарваведе, как мужской, так и женской (Kim 2020: 180). При этом основа *vákṣas-* засвидетельствована уже только в связанном виде в составе двух композитов: *rukṁávákṣas-* 'золотогрудый, несущий на груди золотые украшения (эпитет Марутов)' и синонимичного *hiraṇyavákṣas-* 'золотогрудый' (Kim 2020: 596, 808; Whitney 1881: 255; MW 1899: 882, 1300). Ср. пример из Атхарваеды для слова *úras-*: AVŚ X, 3, 11 «этот варана на **груди** у меня, царь».

Происхождение слова *vákṣas-* не получило однозначного убедительного объяснения. М. Майрхофер полагает, что сравнения с осетинским иронск. *wæxsk* (*yæxck*) / диг. *usqæ*, *wæskæ* (*usqæ*, *yæskæ*) 'плечо' (ср. Abaev 4: 100–101) или с др.-инд. *vakṣāṇā-* 'утроба' неубедительны; при этом исследователь не исключает возможности происхождения от глагольного корня *vakṣ-* 'расти' < и.-ир. **(H)uakš-* 'расти' (EWAia II: 487–488).

При этом *úras-* является прозрачным производным от прилагательного *urú-* 'широкий' (EWAia I: 226). Таким образом, несмотря на то, что оно имеет более надежную этимологию, чем *vákṣas-*, встречающееся в Ригведе, в данном случае опять можно было бы думать о вытеснении основного обозначения эпитетом.

Однако если учитывать характер более поздних данных, приходится признать, что это достаточно спорный случай. Достаточно заметить, что обе основы употреблялись в послеведийскую эпоху в древнеиндийском, а их производные имеют продолжение в среднеиндийских и новоиндийских языках.

Рефлексы *vákṣas-*, засвидетельствованные в индоарийских языках, в большинстве своем начинают обозначать 'бок', а также 'легкое' и 'вымя', хотя праkritские основы *vakkha-*, *vaccha-* еще имеют значение 'грудь' (Turner 1962–1985: 11188).

Соответствия *úras-* и прилагательного *urasá-* 'с сильной грудью' также нашли продолжение в индоарийских языках (Turner 1962–1985: 2350).

Соответственно, в данном случае нельзя исключать и диалектных или стилистических различий между текстом Ригведы и Атхарваеды.

54. moon / луна: RV *ma:s-* {*mās-*} — AV *čand-r'a-* {*candrā-*}

В Ригведе сохраняется древнее имя *mās-* 'луна / месяц' (BR-5 1868: 758; Grassmann 1873: 1036). Оно имеет безупречную индоевропейскую этимологию: тох. А *mañ*, тох. В *meñe* 'луна / месяц', авест. *mā* [maṇah], греч. *μήν* 'луна', готск. *mena* 'луна', лат. *mēnsis* 'месяц' и т. д. < **meh₁ns-* (EWAia II: 352–353).

Основным обозначением луны уже в Атхарваведе становится основа *candrá-* (фонетический вариант *ścandrá-*). Данное слово образовано от корня (*ś*)*cand-* < *(s)*kend-* ‘блестеть’ (EWAia I: 528–529; LIV²: 554) и представляет собой очевидный эпитет, так что в данном случае опять речь идет о замене старого обозначения эпитетом, при этом *mās-* в Атхарваведе удерживает свое второе значение ‘месяц’, как и, например, лат. *mēnsis*.

Любопытно, что засвидетельствован также поэтический композит *candrámmās-*, имеющий значение ‘луна, бог луны’ (буквально ‘блестящая луна’) (BR-2 1858: 943; Grassmann 1873: 436). Он мог представлять собой промежуточное звено при переходе от *mās-* к *candrá-*.

76. sleep / спать: RV *sas-* {*sas-*} — AV *svap-* {*svap-*}

В Ригведе в значении ‘спать’ употребляются корни *sas-* и *svap-*. Детали их распределения вызвали определенную научную полемику, о которой будет сказано ниже.

В Атхарваведе корень *sas-* уже употребляется всего один раз (именительный падеж действительного причастия настоящего времени *sasán*); соответственно, для ведийского списка, ориентированного на Атхарваведу, остается реконструировать только *svap-*.

Распределение *sas-* и *svap-* в тексте Ригведы в общих чертах может быть описано следующим образом: именно *sas-* употребляется в презенсе, данный глагол вообще представляет собой *praesens tantum*. От *svap-* засвидетельствована только одна форма императива 2 лица *svapa* ‘спи’ (RV VII, 55, 2) и дважды действительное причастие *svapánt-* (RV V, 44, 13; RV X, 164, 3). При этом от *svap-* образованы формы перфекта и каузатива (в том числе аористные) (Grassman 1873: 1625–1626). Анализ С. Джемисон, который представляется убедительным, позволяет говорить скорее о супплетивизме этих двух корней в ведийском, нежели о семантическом различии, вопреки мнению, высказанному в (Maurohofer 1965). Как показывает исследовательница, непереходные презенсы от корня *svap-* необходимо признать в целом вторичными; реальные унаследованные из праязыка презентные формы представлены только вед. *sásti* ‘спит’ и т. д. (напрямую соответствует хеттскому *šešzi* ‘спит; проводит ночь; отдыхает, покоится’). Ч. Р. Бартон присоединяется к данному мнению и устанавливает, что для и.-е. **ses-* необходимо реконструировать не предельное (*atelic*) значение ‘спать’ и праязыковой презенс **sésti* в противоположность предельному (*telic*) значению **swep-* ‘уснуть, заснуть’ с праязыковым корневым аористом **swép-t* (Barton 1985). Можно говорить о противопоставлении презентного корня **ses-* и аористного **swep-*.

Примечательное распределение корней *sas-* и *svap-* в Ригведе в определенной степени соответствует хеттской ситуации. В хеттском также засвидетельствованы корни *šeš-* / *šaš-*, ср. *šešzi* ‘спит; проводит ночь; отдыхает, покоится’ (CHD Š: 439–445) и *šupp-* / *šuppa-* / *šuppiya-* (act. *šuppiyezi* / med. *šuppari*, *šuptari* ‘засыпает’) (CHD Š: 607–608). Второй глагол вообще более редок, а значение ‘засыпать’ постепенно переходит к имперфективной основе *šeške-* (от *šeš-*), перед которой в предложении стоит клитика *-za* (CHD Š: 445).

Такую трактовку поддерживает П. Кочаров, рассматривая праиндоевропейские глаголы, имеющие значение ‘спать’, и их смысловые и грамматические различия (Kocharov 2016: 81–86).

Если данное предположение верно, для Ригведы основным глаголом ‘спать’ безусловно служит *sas-*, что соответствует лексической спецификации для данного глагола в (Kassian et al. 2010: 76):

1. *People usually sleep in the night and keep awake by day.*
2. *Do not wake him up, he is sleeping.*

Static verb. Should be kept apart from inchoatives like ‘to fall asleep’.

1. *Обычно ночью люди спят, а днем бодрствуют.*
2. *Не буди его, он спит.*

Длительное состояние. Отличать от инхоативов вроде ‘заснуть’.

При этом *svap-* в Ригведе будет служить именно инхоативом.

На данном примере можно видеть, как от Ригведы к Атхарваведе произошло следующее развитие в лексиконе: корень *sas-* ‘спать’ исчез и был вытеснен близким по семантике корнем *svap-* ‘уснуть’.

84. tail / хвост: AV *p'uč^h-a- {púcha-}* — RV *śe:p-a- {śépa-}* и *śif-n'a- {śísna-}*

Основа *púcha-* не встречается в Ригведе, но в Атхарваведе однозначно обозначает ‘хвост’ (змеи AVŚ VII, 56, 6, быка AVŚ IX, 4, 13 и коровы AVŚ X, 9, 22).

В Ригведе по одному разу в прямом значении ‘хвост’ засвидетельствованы основы *śépa-* и *śísna-*: ср. соответственно RV X, 105, 2 «двоих скакунов — хвосты (их) как у птицы» и RV I, 105, 8 «как крысы вгрызаются в (свои) хвосты». Второе значение обеих основ — ‘penis’, и именно в этом значении данные слова встречаются в Атхарваведе.

В индоарийских языках более позднего периода в значении ‘хвост’ преобладают рефлексы *púcha-* (Turner 1962–1985: 8249). Менее многочисленные продолжения имеет производная суффиксальная основа **śēp-ya-* (Turner 1962–1985: 12606–12607).

Слово *púcha-* не имеет ясной этимологии: М. Майрхофер полагает оптимальным вариантом сравнение с др.-в.-нем. *fuhs* ‘лиса’ и т. д., ср. торвали *pūš* ‘лиса’ < **pucchin* ‘хвостатый, имеющий хвост’ (EWAia II: 140). Я следую мнению, высказанному в Lubotsky 2001: 312, и рассматриваю *púcha-* как заимствование из неизвестного источника; в этом случае логично предполагать, что оно постепенно вытеснило индоарийские по происхождению основы.

94. water / вода: RV *wa:r- {vār-} / ud-'an- {udán-}* — AV *ud-a-k'a- {udaká-}*

Это особенный случай. Для Ригведы необходимо постулировать супплетивную парадигму *vār- / udán-* (род. п. *udnás* и т. д.) (Lubotsky 1997: 352, 1257; Lubotsky 2013), а в Атхарваведе и в более поздних текстах уже нормально употребляется *udaká-* (Kim 2020: 173; Bloomfield 1906: 257).

Существует мнение, что основа *vār-* восходит к **wed-r* при признании так называемого «эффекта Кортландта», перехода **d > *h₁* в консонантных кластерах, и родственна, таким образом, основе косвенных падежей *udán-* (Lubotsky 2013; Garnier 2014: 148–149; de Vaan 2008: 644).

В своей работе Trofimov 2017 я показываю, что с учетом семантики и распределения когнатов вед. *vār-* в других индоевропейских языках необходимо возводить данную основу к отдельной праформе **weh₁r-* ‘мокрота, сырость’ и отделять ее от *udán-* и *udaká-* < **wed-* ‘вода’, что дает для Ригведы дополнительный слот с основой другого происхождения.

2) Морфологические отличия Ригведы по сравнению с Атхарваведой.

В ряде случаев в Ригведе слово, оказывающееся главным кандидатом для списка, имеет тот же корень, что и в Атхарваведе, но при этом представлено другой основой (например, корневой против суффиксальной или с другим суффиксом).

65. rain / дождь: RV *wṛṣ-t'i- {vṛṣṭi-}* — AV *warṣ-a- {varṣá-}*

В Атхарваведе именно слово *varṣá-* является основным (Kim 2020: 624), в то время как синонимичная этимологически родственная основа с другим суффиксом, *vṛṣṭi-*, менее употребительна и встречается менее 10 раз как в Пайппаладе, так и в Шаунаке (Kim 2020: 659–660). В Ригведе, напротив, *varṣá-* встречается только дважды (RV V, 58, 7; RV V, 83, 10), при этом более частотное *vṛṣṭi-* употреблено более 50 раз, а также засвидетельствовано в составе многих композитов (Lubotsky 1997: 1361–1362; BR-6 1871: 1346; Grassmann 1873: 1351–1352). Можно привести следующие примеры: RV I, 38, 8 «когда пролился их **дождь**»;

RV II, 6, 5 «(дай) нам **дождь** с неба!»; RV II, 27, 15 «для него оба обращенных друг к другу (мира) набухают от **дождя** с неба» и т. д.

В данном случае *ti*-основа заменяется тематической.

77. small / маленький: RV 'arb^h-a- {ár**b**ha-} — AV 'arb^h-a-ka- {ár**b**haka-}

В Ригведе основа без суффикса употребляется 12 раз против 6 с суффиксом (Lubotsky 1997: 139), а в Атхарваведе она засвидетельствована уже реже, чем *ár**b**haka-* (Kim 2020: 81; Whitney 1881: 36).

В Атхарваведе наряду с основами *ár**b**ha-* и *ár**b**haka-* употребляется основа *á**l**pa-*, трижды в редакции Пайппалада и дважды в редакции Шаунака (Kim 2020: 83), ср. AVŚ IV, 16, 3: «и два океана — две стороны живота Варуны, и в этой **малой** капле воды он сокрыт»; AVŚ XI, 3, 23–24: «Тот, кто будет знать величие рисовой каши, Не сможет сказать ни «(Оно) малое», ни «(Оно) без добавлений», ни «Оно какое ни на есть». Также в составе композитов *á**l**pa-**p**aśu-* ‘тот, у кого мало скота’, *á**l**pa-śayu-* ‘название мелкого насекомого’; кроме того, есть лексема *alpikā-* (в гимнах Кунтапы AV 20.136.3) — по-видимому, слово, которое можно отнести к «вульгарно-сексуальному регистру» и должно передаваться русскими эквивалентами вроде *малышка*, *крошка*, *детка* (Kulikov 2011: 318–319; Kulikov 2012: 492).

80. star / звезда: RV *stṛ-* {stṛ-} ~ t'ar-a-k-a- {tā**r**akā-} — AV t'ar-a-k-a- {tā**r**akā-}

В Ригведе 8 раз засвидетельствован вариант с начальным *s-* (а именно инструменталис *stṛ**b**his*) (Lubotsky 1997: 1587; Grassman 1873: 1589), уже не встречающийся в более поздние эпохи развития древнеиндийского языка.

В Атхарваведе, как и позднее, употребляется только основа *tā**r**akā-* (по 5 раз в тексте как Пайппалады, так и Шаунаки) (Kim 2020: 303; Whitney 1881: 128).

Первая, засвидетельствованная в Ригведе, основа сохраняет начальный *s-*, содержащийся в праформе **h₂ster-* ‘звезда’ (EWAia II: 755–756), соответственно, представляет собой более архаичный вариант.

3) Сложные случаи.

4. belly / живот: *ud'ar-a-* {udá**r**a-}

Лучшим кандидатом на заполнение слота 4. *belly* / *живот* представляется основа *udá**r**a-* (BR-1 1855: 915; Grassmann 1873: 253; RIVELEX 2: 320). У данного слова наблюдается полисемия ‘живот / утроба’ (в том числе в переносном смысле). Ряд контекстов из Атхарваеды показывают, что оно имеет интересующее нас значение ‘живот’: AVŚ IV, 16, 7 «пусть сидит негодяй с обвислым **животом**»; AVŚ IV, 34, 1 «Брахман — ее голова, брихат — ее спина, (напев) Вамадевы — **живот** рисовой каши» и т. д.

Стоит отметить, что слово встречается только в I, VIII и X мандалах Ригvedы в следующих контекстах: RV I, 25, 15 «(как и) в **утробах** наших»; RV I, 30, 3 «когда эти (соки) с(ливаются) для мощного опьянения у него в **животе**»; I, 42, 9 «наполни живот!» (о Пушане); RV I, 162, 10 «(та) оставшаяся в животе пища, которая пахнет» (о коне); RV VIII, 1, 23 «Этот выжатый напиток, о благой, — пей, переполнив **живот**!» (об Индре) и т. д.

Близким синонимом традиционно считается *kukṣi-*, переводимое как ‘живот, брюхо’ (Kim 2020: 224; BR-2 1858: 306–307; Grassmann 1873: 328). Однако тот факт, что лексема *kukṣi-* в основном употребляется в двойственном числе, а в контексте AVŚ II, 33, 4 («Из внутренностей, из твоих кишок, из прямой кишки, из **чрева** (*udá**r**a-*), якшму из двух сторон **живота** (*kukṣi-*), из потрохов(?), из пупа я вырываю у тебя!») *kukṣi-* и *udá**r**a-* употреблены друг за другом, говорит в пользу того, что они обозначали разные части тела, особенно если учитывать анатомическую точность текстов Атхарваеды. Опираясь в том числе и на это обстоятельство, С. Джэмисон показывает, что первоначальное значение

kukṣí- — не ‘живот’, а ‘щека’ (Jamison 1987). Она реконструирует семантический переход ‘(две) щеки’ > ‘ягодицы’ ~ ‘живот беременной’ > ‘живот’ (Jamison 1987: 80). Приведенные соображения позволяют считать именно *udāra-* основным рефлексом.

Таким образом, рассматриваемая синонимия развилась ввиду близости первоначальных значений, но это две разные основы с различной исходной семантикой.

Следует также отделять основу *jaṭhāra-* с полисемией ‘живот / желудок / утроба / внутренности’ (Kim 2020: 276; BR-3 1861: 12–13; Grassmann 1873: 464–465). В своем первом значении эта основа в первую очередь обозначает специфически живот, наполненный сомой, особенно Индры и других божеств (ср. RV V, 34, 2; RV VI, 67, 7 и т. д.), соответственно, ее нельзя считать стилистически нейтральным синонимом.

Наконец, есть более редкий синоним *vakṣānā-*, который часто употребляется в значении ‘нутро, недра, (материнская) утроба’ (Kim 2020: 608; BR-6 1871: 616; Grassmann 1873: 1190), ср. RV I, 32, 1 «недра гор»; RV V, 42, 13 «в чреве дочери» и т. п. Вообще его значение и этимология неясны (EWAia II: 487).

22. earth / земля: *b^hu:-mi-* {*bhūmi-*}

Здесь три кандидата: основы *kṣám-*, *pr̥thivī-* и *bhūmi-*. Наиболее солидной этимологией среди перечисленных обладает *kṣám-*: она восходит к и.-е. **d^hǵ^hem-* ‘земля’ (EWAia I: 424–425). Тем не менее, основа *kṣám-* уже в Ригведе употребляется только в определенных контекстах в противоположность *dív-* ‘небо’ (ср. RV I, 103, 1 «на **земле** здесь одно, на **небе** другое от нее») и адвербиально (ср. RV III, 8, 7 «(те,) что срубленные, (лежат) на **земле** (*kṣámi*)»). В Атхарваведе же эта основа — уже очевидный архаизм, употребляющийся 6 раз в Пайппаладе и 9 раз в Шаунаке в подобных контекстах, причем многие оказываются зависимы от Ригvedы, ср. AVŚ V, 1, 5 «двое движутся вместе к **земле** (*kṣám*)» (Kim 2020: 233; Whitney 1881: 92).

Намного более употребительна, особенно в Атхарваведе, основа *pr̥thivī-* ‘земля’ (Kim 2020: 447–449; Whitney 1881: 188–189), которая при этом также обозначает прежде всего ‘землю в противоположность небу’ и землю как обожествленную сущность, божество (Grassmann 1873: 855–857).

В любом случае, и *kṣám-*, и *pr̥thivī-* никогда не имеют значения ‘земля как почва’, предполагаемого используемой семантической классификацией (Kassian et al. 2010: 57), а значат только ‘земля как поверхность, ground’.

Соответственно, для базового термина ‘земля’ в ведийском подходит только слово *bhūmi-* с параллельной формой *bhūmī-* и полисемией ‘земля (ground) / земля (почва) / страна’ (Kim 2020: 506–507; BR-5 1868: 348–349; Grassmann 1873: 951). Слово *bhūmi-* также часто используется по отношению к обожествленной и персонифицированной Земле, как и упомянутые синонимы. Примеры: AVŚ III, 17, 5 «на счастье пусть взрезают **землю** хорошие лемехи!», AVŚ V, 11, 6 «пусть уползут дасы вниз в **землю!**» и т. д.

38. head / голова: *fir-as-* {*śiras-*} ~ *fir-ṣ-an-* {*śirṣán-*} (косв.)

Несмотря на то, что значение ‘голова’ для основы *śiras-* устанавливается однозначно как в Атхарваведе, так и в Ригведе, проблема состоит в наличии достаточно близкого синонима *mūrdhán-* (Kim 2020: 539–540; Lubotsky 1997: 1080; BR-5 1868: 856–858; Grassmann 1873: 1053; MW 1899: 826).

Сама основа *śiras-* обыкновенно служит для обозначения ‘головы’ людей и антропоморфных существ, богов и животных: AVŚ IV, 14, 7 «**голове** козла положи в восточной стороне»; AVŚ IX, 8, 21 «из ног, из колен, из ягодиц, из промежности, острые боли спинного хребта, из **головы** — я заставил исчезнуть болезнь».

Основа *mūrdhán-* же неоднократно засвидетельствована в значении ‘глава’ или ‘вершина’, ср. RV I, 24, 5 «к **вершине** богатства». В связи с этим в Большом Петербургском

словаре санскрита специально отмечается, что в древности это слово редко имеет прямое значение и в основном употребляется в переносных (BR-5 1868: 856). Также стоит отметить, что контекстов с основой *mūrdhān-* в Атхарваведе меньше (Kim 2020: 539–540; Whitney 1881: 228), чем в Ригведе (Lubotsky 1997: 1080; Grassman 1873: 1058), а *śiras-* и *śiṛṣān-* по-прежнему частотны (Kim 2020: 686; Whitney 1881: 291–292).

Поскольку в семантической спецификации для аннотированных списков Сводеша, которой я пользуюсь, предлагается отличать слово ‘голова’ от «а) более узких терминов — ‘затылок’, ‘макушка’ и т. п.», «в) метафорического переноса ‘голова’ > ‘верхушка’, ‘начальник’ и т. п.» (Kassian et al. 2010: 63), *mūrdhān-* следует считать второстепенным синонимом.

45. know / знать: *vid-* {*vid-*}

На первый взгляд близкие по значению глагольных корней *vid-* ~ *jñā-* образуют супpletивную пару, поскольку первый образует прежде всего древний перфект *véda* со значением ‘знать’ и реже — корневой презент *vetti* (Kümmel 2000: 501–504), а от второго в основном образуются презентные формы (Kim 2020: 290–291).

Однако искомый, согласно спецификации, «глагол, обозначающий уже имеющее место знание относительно какой-л. ситуации» (Kassian et al. 2010: 66) — это именно *vid-*, например, RV X, 11, 1 *viśvaṃ sā veda* ‘он знает все’, ср. нем. ‘der weiß alles’ (Geldner).

При этом корень *jñā-* как в Ригведе, так и в Атхарваведе имеет более специфическое значение ‘узнавать, понимать’ и ‘иметь знакомство, знать (дорогу, человека), быть знакомым’, которое как раз необходимо отличать от базового глагола ‘знать’, как он описан в Kassian et al. 2010: 66.

Это доказывается многочисленными контекстами с *jñā-*, например, AVŚ XV, 12, 9: «он не знает ни пути, исхоженного отцами, ни (пути), исхоженного богами»; RV 10, 34, 4: «Отец, мать, братья говорят о нем: мы не знаем (его!) Уведите его связанным!»; RV 10, 114, 2: «Три Гибели сидят при этом, чтоб указывать, — ведь возницы (жертвы), которых далеко слышно, **знают** их хорошо» и т. д.

Соответственно, я склоняюсь к тому, что именно корень *vid-* должен быть вынесен как элемент 110-словного списка в значении ‘знать’. Тем не менее, необходимо дальнейшее исследование по разграничению значений этих двух глаголов, тем более что иногда контексты с глаголом *vid-* переводятся при помощи нем. *kennen*, ср. RV II, 14, 10 *véda ahám asya níbhṛtam me etát* «Ich **kenne** ihn; das steht für mich fest» (Geldner).

51. man / мужчина: *nar-* {*nár-*} и *p'umā:s-* {*p'umāṃs-*} ~ *pūs-* {*p'um̐s-*}.

В Ригведе и Атхарваведе встречается несколько синонимов, служащих для обозначения ‘мужчины’: это основы *nár-*, *vīrá-* и *p'umāṃs-* (в косвенных падежах *p'um̐s-*). Дополнительные синонимы *vīṣan-* ‘мужественный / сильный / мужчина / самец / самец любого животного и т. д.’ (BR-6 1871: 1338–1340; Grassmann 1873: 1345–1348; MW 1899: 1012) и *márya-* / *máriya-* ‘молодой человек в расцвете сил’ очевидно имеют более специфическое значение и не годятся в качестве кандидатов.

Интересно, что *nár-* и *vīrá-* имеют надежную индоиранскую или индоевропейскую этимологию с соответствиями, которые указывают в родственных языках в том числе и на значение ‘мужчина’: ср. греч. *ἀνήρ* ‘мужчина; человек’, арм. *ayr* (род. п. *aṛn*) ‘мужчина’ и т. д. (EWAia II: 19–20), второе — лат. *uir* ‘муж, мужчина / супруг’, гот. *wair* ‘муж / мужчина / человек’ и т. д. (EWAia II: 569–570).

Для ведийского списка, тем не менее, я выбираю только два из этих синонимов, *nár-* и *p'umāṃs-* ~ *p'um̐s-* на основании анализа употребления рассматриваемых основ в тексте.

Первый, часто употребляющийся и в Атхарваведе, и в Ригведе, имеет точный антоним ‘женщина’, *nári-* ~ *nāri-*, и иногда употребляется в этой оппозиции: RV I, 43, 36; RV VIII, 77, 8. Примером употребления в своем основном значении может служить кон-

текст AVŚ XIV, 2, 9 «сейчас, о **мужи**, услышьте же от меня, каким благословением супружеская чета достигнет (всего) ценного».

Подобным образом и слово *pūtāms-* ~ *pūms-* имеет антоним *strī-* ‘женщина / самка’ и также употребляется в контекстах противопоставления мужчины женщине: AVŚ I, 8, 1 «кто это совершил, женщина (или) **мужчина**, пусть этот человек здесь объявит себя!»; AVŚ V, 14, 6 «если женщина или если **мужчина** во зло сотворили колдовство».

К сожалению, однозначных примеров, где для мужчин и женщин описывались бы именно присущие им занятия (вроде охоты / плетения корзин), в ведийских текстах обнаружить не удается.

В то же время *vīrā-* чаще всего имеет значение ‘герой’ и постоянно употребляется по отношению к божествам, ср. RV VI, 21, 6 «мы воспеваем тебя (Индра), о **герой**, привлекаемый молитвой»; AVŚ III, 23, 2 «да родится тут **герой**, сын у тебя, десятимесячный!».

60. night / ночь: *r'a-tr-i-* {*rātrī-*} ~ *r'a-tr-i-* {*rātri-*}

В ведийском есть три основы, которые могли бы теоретически претендовать на заполнение данного слота.

Более древние по происхождению основы *kṣāp-* (BR-2 1858: 530; Grassmann 1873: 362) и *nākt-* (BR-4 1865: 6; Grassmann 1873: 703) уже в Ригведе употребляются как наречия или в связанных контекстах в устоявшихся поэтических выражениях (Lubotsky 1997: 458–459, 785). Для *kṣāp-* стандартны употребления адвербиального типа, ср. RV IX, 99, 2 «и вот, приготовленный ночью, он (Сома) ныряет вперед к наградам») и контексты с основой *uṣās-* ‘заря’ во множественном числе (ср. RV VI, 52, 15 «Чтобы (мы) процветали (и) **ночами**, (и) и на утренних зорях!»; RV VII, 15 «свети (Агни) в **ночах** и на утренних зорях!»). В Атхарваведе эта основа уже не употребляется (Kim 2020: 235; Whitney 1881: 92), позднее в памятниках древнеиндийского языка встречается только производное *kṣapā-* (BR-2 1858: 531). Соответственно, это слово никак не может претендовать на заполнение слота.

Похожим образом употребляется *nākt-*: это либо обожествленная сущность, либо слово, употребляемое вместе с *uṣās-* ‘заря’, либо наречие *nāktam* ‘ночью’, произведенное от винительного падежа единственного числа соответствующей основы, часто вместе с противоположным по значению *dīvā* ‘днем’, ср. AVŚ V, 7, 3 «пусть будет в порядке днем и **ночью** наше награждение, созданное богами!», RV I, 24, 12 «О том **ночью**, о том днем мне твердят».

Поэтому базовым обозначением ‘ночи’ в ведийском языке можно считать основу *rātrī-* (Kim 2020: 594–595). Ср. AVŚ II, 8, 2 «Прочь пусть засветит эта **ночь** (кшетрию)». Она также может выступать как обожествленная персонифицированная сущность, ср. RV I, 35, 1 и гимн к Ночи AVŚ XIX, 49, подробно с параллелями в ведийской литературе см. Kulikov 2010.

64. person / человек: *ṣan-a-* {*jāna-*}

Для слота ‘человек’ есть целых три кандидата, *mānuṣa-* (RV, в AV только субстантивированное прилагательное *mānuṣyā-* ~ *mānuṣīya-*), *pūruṣa-* ~ *pūruṣa-* и *jāna-*. Их употребление в тексте Ригведы и Атхарваведы показывает примечательное распределение: *mānuṣa-* и *mānuṣyā-* ~ *mānuṣīya-* обозначают человека в противоположность божествам, *pūruṣa-* ~ *pūruṣa-* человека в оппозиции к (другим) животным, *jāna-* человека вообще (Trofimov 2019). Для последнего в связи с этим частотное употребление во множественном числе в значении ‘люди’ (например, RV II, 12, 1 «тот, о люди, Индра!») и далее неоднократно в тексте того же гимна).

В связи с этим необходимо выбрать *jāna-* как самый нейтральный синоним.

Примечательно, что в различных индоарийских языках все четыре основы имеют рефлекс с базовым значением ‘человек’ (Turner 1962–1985: 5098; 8289; 9827–9828), причем преобладают потомки производного *mānuṣyā-* ~ *mānuṣīya-*, на втором месте — продолжения *jāna-*, в достаточно небольшом количестве языков — продолжения *pūruṣa-* и *mānuṣa-*.

89. tooth / зуб: *dant-* {*dánt-*}

В данном случае в Ригведе и Атхарваведе встречаются два слова, *dánt-* и *jámbha-*, оба имеют хорошую индоевропейскую этимологию и связаны с когнатами в других языках со значением ‘зуб’: ср, например, лат. *dens, dentis* ‘зуб’ (EWAia I: 695–696) и тох. А *kat*, тох. В *kete* ‘зуб’, праслав. **zobъ* ‘зуб’ (EWAia I: 572–573). Очевидно, что определение основного понятия в ведийском является дополнительным аргументом в пользу реконструкции **dent-* или **ǵombʰo-* на прауровень в качестве основного обозначения ‘зуба’.

Несмотря на то, что синоним *jámbha-* чаще употребляется (особенно в Ригведе), он имеет специфическую семантику. В основном это слово применяется в Ведах в метафорическом смысле, когда речь идет о ‘зубах огня’, то есть бога Агни (ср. RV I, 143, 5 «(этот) Агни ест-пожирает острыми зубами»). Также имеет значение ‘челюсть’, ‘ряд зубов’, часто отраженное в словарях и переводах, ср. RV VIII, 60, 14 «Не устоять перед твоими **челюстями**, о Агни-бык» и т. д.

Кроме того, необходимо отметить, что *jámbha-* является производным по отношению к глаголу *jabh-* ~ *jamh-* ‘хватать; защелкивать(ся)’, а это хорошо согласуется со значением ‘ряд зубов / челюсти’.

История и внутренние связи рассматриваемого ведийского слова прослежены в работах (Narten 1965; Mumm 1999). Оба исследователя приходят к выводу, что слово *jámbha-* имело в ведийском значение ‘ряд зубов’, ‘челюсть’. В связи с этим очевидно, что основное обозначение ‘зуба’ в Ведах — именно *dánt-*. Это хорошо согласуется с такими парами когнатов, как греч. *γόμφος* ‘клин, гвоздь’ при *ὀδούς*, род. п. *ὀδόντος* ‘зуб’, др.-исл. *komb* ‘гребень’ при *tonn* ‘зуб’ и т. д.

90. tree / дерево: *wrkṣ-a-* {*wrkṣá-*}

Значение ‘дерево’ подтверждается для основы *wrkṣá-* достаточным количеством контекстов: AVŚ X, 7, 38 «как ветви, разошедшиеся вокруг ствола **дерева**»; RV III, 45, 4 «стряхни, как спелый плод крюком с **дерева**»; RV IV, 20, 5 «спелое **дерево** (с плодами)»; RV VI, 24, 3 «ветви **дерева**».

Близкий синоним, основа *vána-*, имеет значение ‘дерево, лесное дерево’ большей частью в Ригведе (Lubotsky 1997: 1228–1229; BR-6 1871: 666–667; Grassmann 1873: 1206–1207; MW 1899: 917–918), в Атхарваведе это уже в основном ‘лес’, как и в более поздних памятниках, ср. AVŚ VI, 33, 1 «чтобы подталкивать людей, **лес**, небо»; AVŚ VIII, 6, 11 «кто поднимает шум в **лесу**».

Часто употребляемая основа *vanaspáti-* ‘(лесное) дерево’ буквально переводится как ‘повелитель леса’ (BR-6 1871: 673–674; Grassmann 1873: 1208–1209) и употребляется в стилистически окрашенных контекстах, например, AVŚ IV, 3, 1 «божественное **дерево**», AVŚ V, 27, 11 «о **дерево**, отпусти (жертвенное животное), даря (его)!»

Также не подходит в качестве кандидата на основное обозначение *dāru-* ‘деревянное бревно, балка / дрова’ (Kim 2020: 339; BR-3 1861: 595; Grassmann 1873: 595–596), восходящее к и.-е. **dor-u-* / **der-u-* / **dr-ew-* ‘дерево’.

99. woman / женщина: *n'a:r-i-* {*nári-*} ~ *n'a:r-i-* {*nārī*} и *str-i-* {*strī-*}

Как и в случае со слотом ‘мужчина’, засвидетельствован большой ряд синонимов, имеющих определенные коннотации или указывающих на возрастную семантику. Среди них *gná-* ‘божественная женщина, женское божество’ (BR-2 1858: 830; Grassmann 1873: 415; MW 1899: 370); *jáni-* ‘жена’ (BR-3 1861: 29; Grassmann 1873: 475), *yóṣañā-* ‘девочка, девушка’ с вариантами основы *yóṣā-* и *yóṣit-* (BR-6 1871: 200; Grassmann 1873: 1128), *ménā-* ‘самка любого животного / женщина’ (BR-5 1868: 903–904; Grassmann 1873: 1064).

В качестве элементов списка Сводеша необходимо выбрать *nāri-* ~ *nārī-* (Kim 2020: 395; BR-4 1865: 117; Grassmann 1873: 725) и *strī-* (Kim 2020: 777; BR-7 1875: 1275–1276; Grassmann 1873: 1596), прямые антонимы к *nār-* и *ṛítām̐s-* ~ *ṛít̐s-* ‘мужчина’ соответственно.

Проведенное исследование показывает, что ведийский список может быть в целом успешно составлен при должном привлечении данных текстов и с опорой на определенный семантический стандарт. Общее число случаев расхождения между Атхарваведой и Ригведой — всего 6 основ. Морфологические различия между текстами Ригведы и Атхарваведы засвидетельствованы в 3 случаях. В 10 случаях было несколько весомых кандидатов на заполнение слота, но работа с текстом позволила сделать уверенный выбор базовых терминов. В тексте Атхарваведы 4 лакуны: 3. bark / кора; 50. louse / вошь; 69. round / круглый; 105. short / короткий. По две основы в главный слот вынесено 5 раз из-за супплетивизма (6 раз для Ригведы, где дополнительно наблюдается супплетивизм *vár-* / *udán-* ‘вода’) и 5 раз из-за неснимаемой синонимии.

References

- Abaev, V. I. 1958–1995. *Istoriko-ètimologičeskij slovar' osetinskogo jazyka. V 4-x tomax* [Historical-Etymological Dictionary of Ossetic. In 4 volumes]. Moskva-Leningrad: Izdatel'stvo AN SSSR.
- Bammesberger, Alfred. 1982/83. Zur Entstehung der vedischen Imperative auf *-(s)i*. *Zeitschrift für vergleichende Sprachforschung* 96: 1–5.
- Barton, Charles. 1985. PIE **suep-* and **ses-*. *Die Sprache* 31: 17–39.
- Bloomfield, Maurice. 1906. *A Vedic concordance, being an alphabetic index to every line of every stanza of the published Vedic literature and to the liturgical formulas thereof, that is an index to the Vedix Mantras, together with an account of their variations in the different Vedic books*. Cambridge, MA: Harvard University Press.
- Bloomfield, Maurice. 2008. *Maurice Bloomfield's A Vedic Concordance Enhanced with New Material Taken from Seven Vedic Texts. Edited by Marco Franceschini. Harvard Oriental series 66*. Cambridge, MA: Harvard University Press.
- BR = Böhtlingk, Otto, Rudolph Roth. 1855–1875. *Sanskrit Wörterbuch, herausgegeben von der kaiserlichen Akademie der Wissenschaften, bearbeitet von Otto Böhtlingk und Rudolph Roth. 7 Bde*. St-Petersburg: Eggers.
- Cardona, George. 1965. The Vedic imperatives in *-si*. *Language* 41: 1–18.
- CHD = P. M. Goedegebuure, H. G. Güterbock, H. A. Hoffner, T. P. J. van den Hout (eds.) 1980–. *The Hittite Dictionary of the Oriental Institute of the University of Chicago*. Chicago: The Oriental Institute of the University of Chicago.
- Elizarenkova, T. J. (translator). 2005–2010. *Atharvaveda. V 3-x tomax* [Atharvaveda. In 3 volumes]. Moskva: Vostočnaja literatura RAN.
- Elizarenkova, T. J. (translator). 1989–1999. *Rigveda. V 3-x tomax* [Rigveda. In 3 volumes]. Moskva: Nauka.
- ESIJ = V. S. Rastorgueva, D. I. Edel'man. 2000–. *Ètimologičeskij slovar' iranskix jazykov* [Etymological dictionary of the Iranian languages]. Moskva: Vostočnaya literatura RAN.
- EWAia I–II = Manfred Mayrhofer. *Etymologisches Wörterbuch des Altindoarischen. I–II Bde*. Heidelberg: Universitätsverlag Winter.
- Garnier, Romain. 2014. Nouvelles considérations sur l'effet Kortlandt. *Glotta* 90: 139–159.
- Geldner, Karl F. 1951. *Der Rig-Veda aus dem Sanskrit ins Deutsche übersetzt. 3 Bde*. London and Wiesbaden: Verlag Otto Harrassowitz.
- Gotō, Toshifumi. 1987. *Die “I. Präsensklasse” im Vedischen*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Gotō, Toshifumi. 2013. *Old Indo-Aryan Morphology and its Indo-Iranian Background (Veröffentlichungen zur Iranistik): In cooperation with Jared S. Klein und Velizar Sadoovski*. Wien: Austrian Academy of Sciences Press.
- Grassmann, Hermann G. 1873. *Wörterbuch zum Rig-Veda*. Leipzig: F.A. Brockhaus.
- Jamison, Stephanie W. 1982. “Sleep” in Vedic and Indo-European. *Zeitschrift für vergleichende Sprachforschung* 96: 6–16.
- Jamison, Stephanie W. 1987. Linguistic and philological remarks on some Vedic body parts. In: Calvert Watkins (ed.). *Studies in Memory of Warren Cowgill (1929–1985). Papers from the Fourth East Coast Indo-European Conference. Cornell University, June 6–9, 1985: 66–91*. Berlin / New York: Walter de Gruyter.

- Jamison, Stephanie W., Joel P. Brereton. 2014. *The Rigveda: the earliest religious poetry of India*. Oxford University Press.
- Hill, Eugen. 2007. *Die Aorist-Präsentien des Indoiranischen. Untersuchungen zur Morphologie und Semantik einer Präsensklasse*. Bremen: Hempen Verlag.
- Kassian, Alexey, George Starostin, Anna Dybo, Vasily Chernov. 2010. The Swadesh wordlist. An attempt at semantic specification. *Journal of Language Relationship* 4: 46–89.
- Kassian, Alexey. 2011. Annotated 50-item wordlist of the basic lexicon of the Ancient Greek languages (the idiolect of Herodotus). *Journal of Language Relationship* 6: 65–100.
- Kim, Jeong Soo. 2020. *Index Verborum der Paippalādasamhitā und der Śaunakasamhitā des Atharvaveda (eine vorläufige Version)*. Ms. (in print).
- Kocharov, Peter. 2016. Proto-Indo-European lexical aspect and stem patterns. *Faits de langues* 47: 75–88.
- Kulikov, Leonid. 1997. Vedic *mriyāte* and other pseudo-passives: notes on an accent shift. In: Irén Hegedus, Peter A. Michalove, Alexis Manaster Ramer (eds.). *Indo-European, Nostratic, and Beyond: Festschrift for Vitalij V. Shevoroshkin (JIES monograph series; 22)*: 198–205. Washington DC: Institute for the Study of Man.
- Kulikov, Leonid. 2008. The Vedic type *patayati* revisited: Semantic oppositions, paradigmatic relationships and historical connections. In: Alexander Lubotsky, Jos Schaeken, Jeroen Wiedenhof (eds.). *Evidence and Counter-evidence: Essays in honour of Frederik Kortlandt. Vol. I. Balto-Slavic and Indo-European Linguistics*. (Studies in Slavic and General Linguistics 32): 245–261. Amsterdam: Rodopi.
- Kulikov, Leonid. 2010. Atharvaveda-Śaunakīya 19.49.1 = Atharvaveda-Paippalāda 14.8.1: an etymological note on Vedic *rātri-* ‘night’. In: T. M. Nikolaeva, P. M. Arkad’ev, A. A. Zaliznjak, N. N. Kazanskij (eds.) *Ю Issledovanija po lingvistike i semiotike: Sbornik statej k jubileju Vjač. Vs. Ivanova*: 174–179. Moskva: Jazyki slavjanskix kul’tur.
- Kulikov, Leonid. 2011. Passiv, antikauzativ, “intransitiv”? K interpretaciji trex vedijskix prezensov s suffiksom *-ya-* (*jāyate, mriyāte, yabhyate*) [Passive, anticausative, “intransitive”? Notes on interpretation of three Vedic present formations with the suffix *-ya-* (*jāyate, mriyāte, yabhyate*)]. *Indoeuropejskoje jazykoznanie i klassicheskaja filologija. Materialy chtenij* 14: 310–322.
- Kulikov, Leonid. 2012. *The Vedic -ya-presents: Passives and intransitivity in Old Indo-Aryan*. Amsterdam: Rodopi.
- Kulikov, Leonid. 2017. Ch. 6. Indo-Aryan. In: Mate Kapović (ed.). *The Indo-European Languages*: 214–262. London: Routledge.
- Kümmel, Martin Joachim. 1996. *Statio und Passivaorist im Indoiranischen. (Historische Sprachforschung, Ergänzungsheft 39)*. Göttingen: Vandenhoeck & Ruprecht.
- Kümmel, Martin Joachim. 2000. *Das Perfekt im Indoiranischen. Eine Untersuchung der Form and Funktion einer ererbten Kategorie des Verbums und ihrer Weiterentwicklung in den altindoiranischen Sprachen*. Wiesbaden: Reichert Verlag.
- LIPP = Dunkel, George E. 2014. *Lexikon der indogermanischen Partikeln und Pronominalstämme. Band 1: Einleitung, Terminologie, Lautgesetze, Adverbialendungen, Nominalsuffixe, Anhänge und Indices. Band 2: Lexikon*. Heidelberg: Winter.
- LIV² = Rix, Helmut (ed.). 2001. *Lexikon der indogermanischen Verben. Die Wurzeln und ihre Primärstammbildungen*. Wiesbaden: Dr. L. Reichert Verlag.
- Lubotsky, Alexander. 1997. *A Rgvedic Word Concordance*. 2 vols. New Haven: American Oriental Society.
- Lubotsky, Alexander. 2001. The Indo-Iranian substratum. In: Christian Carpelan, Asko Parpola, Petteri Koskikallio (eds.). *Early contacts between Uralic and Indo-European: linguistic and archaeological considerations*: 301–317. Helsinki: Suomalais-Ugrilainen Seura.
- Lubotsky, Alexander. 2002. *Atharvaveda-Paippalada, Kanda Five: Text, Translation, Commentary*. Cambridge, MA: Harvard University Press.
- Lubotsky, Alexander. 2013. The Vedic paradigm for ‘water’. In: A. I. Cooper, J. Rau, M. Weiss (eds.). *Multi Nominis Grammaticus. Studies in classical and Indo-European linguistics in honor of Alan J. Nussbaum on the occasion of his sixty-fifth birthday*: 159–164. Ann Arbor / New York: Beech Stave Press.
- MacKenzie, David N. 1986. *A concise Pahlavi dictionary*. London / New York / Toronto: Oxford University Press.
- Mayrhofer, Manfred. 1965. Hethitisches und arisches Lexikon. *Indogermanische Forschungen* 70: 245–257.
- MW 1899 = Monier-Williams, M. A. *Sanskrit-English dictionary: etymologically and philologically arranged with special reference to cognate Indo-European languages*. Oxford: Clarendon Press.
- Mumm, Peter A. 1999. Deutsch *Kamm, Kimme* und die Bedeutung von idg. **ǵomb^h-o-*, **ǵemb^h-*. In: Wolfgang Schindler, Jürgen Untermann (eds.). *Grippe, Kamm und Eulenspiegel. Festschrift für Elmar Seebold zum 65. Geburtstag*: 295–312. Berlin / New York: De Gruyter.

- Narten, Johanna. 1965. *Ai. jámbha-*, gr. γόμφος und Verwandtes. *Zeitschrift für vergleichende Sprachforschung* 79: 255–264.
- Orlandi, Chatia. 1992. *Gli inni dell'Atharvaveda: Śaunaka. Traslitterazione a cura di Chatia Orlandi e Saverio Sani*. Pisa: Giardini.
- RIVELEX 1–2 = Krisch, Th. (ed.). 2006-. RIVELEX. *Rigveda-Lexikon. Bd. 1. Words beginning with "a". Bd. 2. Words beginning with other vowels than "a"*. Graz: Leykam.
- Roth, Rudolf von, William D. Whitney. 1856–1857. *Atharvaveda Atharva Veda Sanhita, herausgegeben von R. Roth und W.D. Whitney*. Berlin: F. Dümmler.
- Shields, Kenneth. 1997. The origin of the Vedic 2nd sing. Imperative in «-si». *Indo-Iranian Journal* 40: 149–155.
- Starostin, George. 2019. Chinese basic lexicon from a diachronic perspective: implications for lexicostatistics and glottochronology. *Journal of Language Relationship* 17: 153–176.
- Starostin, Sergei (ed.). 1998–2005. *The Tower of Babel. An etymological database project*. Available at: <http://starling.rinet.ru/> [accessed 12.11.2020].
- Tedesco, Paul M. 1944. The Supposed Rigvedic Present *márate*. *Language* 20: 212–222.
- Trofimov, A. A. 2017. Vedijskie osnovy *vār-*, *udán-* i *udaká-* 'voda': ix paradigmatičeskie i ètimologičeskie otnošeniya [Vedic stems *vār-*, *udán-*, *udaká-* 'water' and their paradigmatic and etymological interrelations]. *Journal of Language Relationship* 15: 181–192.
- Trofimov, A. A. 2019. Sootnošenie slov *jána-*, *púruša-* i *mánuša-* 'čelovek' v Rigvede i Atharvavede [The correlation between the words *jána-*, *púruša-* and *mánuša-* 'person' in Rigveda i Atharvaveda]. In: *Sovremennoe sostojanie indoevropejskogo jazykoznanija i sanskritologii. Sbornik naučnoj konferencii, posvyščennoj pamjati zaslužennogo professora Moskovskogo universiteta imeni M. V. Lomonosova Very Aleksandrovny Kočerginoj (1924–2018)*: 310–315. Moskva: MCNMO.
- Turner, Ralph L. 1962–1985. *A comparative dictionary of Indo-Aryan languages*. London: Oxford University Press.
- Vaan, Michiel de. 2008. *Etymological dictionary of Latin and the other Italic languages*. Leiden/Boston: Brill.
- Vishva Bandhu, Bhim Dev, S. Bhaskaran Nair (eds.). 1935–1965. *Vaidika-Padānukrama-Koṣa: A Vedic Word-Concordance*. Hoshiarpur: Vishveshvaranand Vedic Research Institute.
- Wackernagel, Jakob. 1896. *Altindische Grammatik. Bd. I. Lautlehre*. Göttingen: Vandenhoeck and Ruprecht.
- Wackernagel, Jakob. 1930. *Altindische Grammatik. Bd. III. Nominalflexion — Zahlwort — Pronomen*. Göttingen: Vandenhoeck and Ruprecht.
- Whitney, William D. 1881. Index verborum to the published text of the Atharva-Veda. *Journal of the American Oriental Society* 12: 1–383.
- Whitney, William D. 1885. *The roots, verb-forms, and primary derivatives of the Sanskrit language: a supplement to his Sanskrit grammar*. Leipzig: Breitkopf and Hartel.
- Witzel, Michael. 1997. The Development of the Vedic Canon and its Schools: The Social and Political Milieu. In: Michael Witzel (ed.). *Inside the Texts, Beyond the Texts: New Approaches to the Study of the Vedas, Harvard Oriental Series, Opera Minora; vol. 2*: 257–348. Cambridge: Harvard University Press.
- Witzel, Michael, Gotō Toshifumi, Mislav Ježić, Eijiro Dōyama. 2007-. *Rig-Veda — Das heilige Wissen. Erster und zweiter Liederkreis. Dritter bis fünfter Liederkreis*. Frankfurt am Main und Leipzig, Berlin: Verlag der Weltreligionen.

Artem Trofimov. An attempt at an annotated Swadesh wordlist for Vedic Sanskrit

The goal of the present paper is to illustrate certain philological and semantic issues that arise during the compilation of the 110-item Swadesh wordlist for the Vedic language, which is itself a necessary procedure for any satisfactory lexicostatistical analysis of Indo-Iranian and Indo-European data. The language of the Atharva Veda, the second of the Vedas, was chosen as the main idiom, since it does not have quite as many lacunae and is more stylistically neutral than the other texts. For the purposes of the wordlist, I have applied the semantic specifications currently used in the GLBD (Global Lexicostatistical Database). The text of the Atharva Veda was subsequently compared with the Rig Veda in order to obtain more reliable and comprehensive results and to study linguistic evolution at the early stages of development of the Vedic language. Research shows that 87 items of the Vedic list are more or

less uncontroversial. Six meanings correspond to different stems in the Atharva Veda and the Rig Veda; there are also three cases of consistent morphological differences between the two texts. In 10 cases, it was necessary to choose the most appropriate candidate from a variety of synonyms. Only four lacunae remain in the text of Atharva Veda. Finally, there are five cases of suppletive stems and five cases of synonymy. The study shows that the Swadesh list for Vedic can be successfully compiled based on textual data and with the application of a strict semantic standart.

Keywords: lexicostatistics; Swadesh list; Vedic language; Atharva Veda; Rig Veda; synonymy; suppletivism.

А. Е. Аникин^{†*}, И. И. Муллонен^{‡**}

[†] Институт филологии Сибирского отделения РАН; Институт русского языка им. В. В. Виноградова РАН; alexandr_anikin@mail.ru

[‡] Институт языка, литературы и истории КарНЦ РАН; mullonen@krc.karelia.ru

С. А. Мызников

Русский диалектный этимологический словарь. Лексика контактных регионов.

М.; СПб: Нестор-История, 2019. —1064 стр.

Опубликован новый важный справочник по русской этимологии — огромный труд С. А. Мызникова «Русский диалектный этимологический словарь. Лексика контактных регионов» (далее РДЭС).

Книга содержит довольно пространное Введение (с. 4—23), в котором излагаются общие сведения о словаре, состав включенной в него лексики, дается характеристика словника и источников, описание структуры словарных статей и прочие «технические» данные. Имеется большой указатель русских слов (с. 1006—1062) и внушительный список литературы (с. 966—1005), список сокращений (с. 24—30). Отдельные разделы Введения посвящены обсуждению проблемы достоверности диалектных данных и теоретических вопросов, касающихся терминов «заимствование» и «субстрат», характеристик некоторых этимологических исследований диалектной лексики неисконного происхождения (Я. Калима, М. Фасмер, А. К. Матвеев и др.), диалектных этимологических словарей и др.

К крупной форме в области «чистой» этимологической лексикографии С. А. Мызников прибегнул впервые. Как одному из основных авторов и (в последние годы) руководителю Словаря русских народных говоров (СРНГ) ему прекрасно знакома работа в области описательной диалектной лексикографии. РДЭС можно расценить как обобщение и развитие его предшествующих фундаментальных трудов (прежде всего монографий и атласов¹,

не говоря о «*kleine Schriften*»), посвященных главным образом ареально-лингвогеографическому и этимологическому аспектам изучения лексики прибалтийско-финского, но также пермского, марийского и мордовского происхождения в русских говорах. Как и в предшествующих трудах Мызникова, в РДЭС затрагиваются и иные заимствования (балтизмы, тюркизмы, англицизмы и др.), а также некоторые исконные слова — в тех случаях, когда они представляют особый этимологический интерес. В словаре дается относительно большое количество топонимов, которые привлекаются по большей части как вспомогательный материал при рассмотрении апеллятивов. Некоторые статьи РДЭС посвящены анализу русских диалектизм, вызывающих сомнения в их достоверности.

Словарь опирается на богатейшие данные по русским говорам, извлеченные из опубликованных и рукописных источников и картотек, но также на чрезвычайно обширные полевые материалы самого С. А. Мызникова, собиравшиеся им в течение более чем 30 лет. Представлены материалы, записанные от носителей русских диалектов на территориях Северо-Запада и Центра России, Южного Урала и др. (всего около 20 областей и республик РФ), а также Казахстана. Речь идет о тех местах, где имеют место более или менее интенсивные и давние контакты русских и соседних диалектов. Словарь построен по алфавитно-гнездовому принципу, когда лексика объединяется вокруг одного и того же этимона или нескольких сходных этимонов, но доминирует алфавитный порядок расположения слов. В этимологическом плане ставка делается на детальное (особенно в плане лингвогеографии) представление непосредственного этимона, с меньшим вниманием к его праязыковой истории, иноязычным связям и соответствующим реконструкциям.

* Публикация подготовлена по проекту ИФЛ РАН № 0328-2016-0003 «Язык и фольклор русских в Сибири как отражение культурных взаимодействий» и проекту ИРЯ РАН № АААА-А18-118021690163-2 «Реконструкция истории славянской лексики от праславянского периода до современного состояния и этимология лексического состава русского языка».

** Публикация подготовлена в рамках выполнения государственного задания КарНЦ РАН.

¹ (Мызников 2003; Мызников 2004; Мызников 2005; Мызников 2007; Мызников 2010). При активном участии С. А. Мызникова в Институте лингвистических исследований РАН (СПб) ведется работа над крупным Всероссийским коллективным

проектом, объединяющим многие ВУЗы и научные институты — Лексическим атласом русских народных говоров.

Весьма ценно, что РДЭС содержит словник от *a-* до *я-*, т. е. является единым целым, не разделяющимся на части и выпуски. При этом он содержит материал, значительная часть которого еще не подвергалась этимологической обработке. Если оценивать РДЭС с позиций «общерусского» этимологического словаря, каковым у нас является прежде всего известный труд М. Фасмера, то следует констатировать, что по существу любая статья РДЭС (не считая отсылочных) может служить дополнением для словаря Фасмера и для будущих этимологических словарей русского языка. В этой связи, кстати, уместно отметить наличие многих этимологических решений, совпадающих в РДЭС и «Русском этимологическом словаре» (РЭС). Примером может служить севернорусское *дóра* ‘небольшая плоскодонная лодка с широкой кормой’ (СРНГ 8: 130), ошибочно объясняемое в «Православном словаре» как рефлекс праслав. **dora*, далее к *драть* (SP 4: 111). В действительности, *дóра*, как справедливо сказано в РДЭС (с. 178), происходит из англ. (США) *dory* (*dorey*) ‘морская плоскодонная гребная лодка’, неясного происхождения (см. также (РЭС 14: 220, 382))².

Вместе с тем словарь С. А. Мызникова в очередной раз демонстрирует перспективность и/или реалистичность создания диалектных (и иных «частных») этимологических словарей русского языка. Большой «общерусский» словарь в духе словаря Фасмера является трудно реализуемым, поскольку требует для своего завершения много времени даже в случае привлечения коллектива специалистов. Однако диалектный этимологический словарь является и «научно необходимым и самодостаточным научным проектом» (Варбот 1997: 193), способным дать выигрышную точку зрения на лексику языка.

Несомненно, что такой важный труд, как РДЭС, не является чем-то застывшим. В перспективе желательны переиздания этого труда с учетом новых научных достижений и т. д. Предлагаемые в данной рецензии соображения по поводу РДЭС рассчитаны именно на будущее этого словаря.

Переходя к оценке РДЭС³, следует прежде всего заметить, что, несмотря на преимущественное внимание к заимствованной лексике, С. А. Мызников сделал некоторые наблюдения, весьма важные для изучения исконнорусской лексики.

² Опубликованные выпуски РЭС (1–14) охватывают отрезок русского словника от *a-* до *д-*. В данной рецензии эти выпуски для краткости используются как правило без ссылки.

³ Используемые в РДЭС многочисленные диалектные пометы в данной рецензии ради краткости как правило опускаются.

Рус. диал. *бот* ‘длинный шест с утолщением (чтобы бить по воде и пугать рыбу), ботало’ в этимологической литературе обычно понимается как продолжение праслав. **botъ* от **botati* > рус. диал. *бóтать* ‘качать, болтать, колебать, бить, пахтать масло; бить по воде, загоняя рыбу в сеть’ (Фасмер 1: 200; SP 1: 341; ЭССЯ 1: 224). Иначе РДЭС (с. 91) — *бот* ‘ботало’ из *бóйт* < *болт* от *боўтáть* < *болтáть*. Аналогично объясняется *гобѣц* ‘деревянная пристройка к русской печи’ < *гоўбѣц* < *голбѣц* (РДЭС: 164). Подобные примеры, как выяснилось, являются реликтами древненовгородской фонетики, что подтверждается многочисленными примерами типа диал. *вóглый* — *вóлглый*, *коч* — *колч* и т. п. (Зализняк, Гиппиус 2018: 18–19) (со ссылкой на (Мызников 2004: 45)).

Но в центре внимания С. А. Мызникова как этимолога и лингвгеографа, разумеется, финно-угорские и особенно прибалтийско-финские истоки русской диалектной лексики. На этом материале в РДЭС, как и в предшествующих монографиях и статьях Мызникова, им предложены сотни авторских этимологий, значительно пополнивших список ставших уже классическими интерпретаций Я. Калимы, И. Микколы, А. Л. Погодина, М. Фасмера⁴. Кроме того, многие из этих интерпретаций осмыслены в словаре Мызникова по-новому, с учетом полученных на протяжении XX столетия новых полевых и архивных материалов. Он успешно использует наработки своих коллег-этимологов, работающих с финно-угорским материалом. Им выявлены некоторые принципиально важные особенности фонетической адаптации финно-угорских этимонов в русской языковой среде, ареальная дистрибуция русских диалектных лексем финно-угорского происхождения, их лексико-семантические группы.

Среди новых авторских этимологий изящная интерпретация загадочного для предшественников термина из былинных текстов *ольга* ‘болото, топкая низина’ как производительного варианта слова *орга*, широко бытующего в говорах Северо-Запада с той же семантикой и имеющего очевидные прибалтийско-финские истоки (с. 555–556). Безупречной представляется вепсская этимология для *ра́хта* ‘жидкость, остающаяся при сбивании масла’, не попадавшего ранее в поле зрения этимологов (из вепс. *rahtod* ‘вытопки от масла’). При этом автором предложены более убедительные истоки для самого прибалтийско-финского слова, чем это сделано в классическом для прибалтийско-финских

⁴ Е. А. Хелимский в шутку и всерьез называл Мызникова «новым Калимой» и «новым Жильероном».

языков этимологическом словаре SKES (с. 656—657). Достоверность большого ряда этимологических этюдов основана не только на собственно лингвистических параметрах, но и на привлечении экстралингвистических данных. Показательный пример — термин *нярега*, называющий полоску из жести, закрепляющую косу на рукоятке. Как показано С. А. Мызниковым, этимология явно хранит память о том, что традиционно для этой цели использовали не жечь, а древесный материал из молодой ели, еловой поросли, карел. *näre, näreh* (с. 544). Список подобных примеров может быть многократно расширен.

Материалы РДЭС представляют значительную ценность для прибалтийско-финского и, шире, финно-угорского языкознания кроме прочего в том плане, что в русских говорах сохранились многочисленные реликты языковых единиц, утраченные в языках, послуживших источниками заимствований или субстратных включений. Так, русские говоры Обонежья бесспорно важны для истории вепсского языка, который, как известно, относительно поздно оказался в поле зрения собирателей и исследователей, когда произошло уже практически полное обрусение территории исторического вепсского Присвирья и южного Обонежья. Показателен в этом смысле богатый ряд дескриптивных глаголов, зафиксированный в русских говорах в зоне вепсского субстрата (например, *лявндать, лямзать, гуньяндать, кабандать, нюлайдатель, нявандать*). При этом сам вепсский оригинал не отражен в доступных источниках по вепсской лексике, что вынудило автора РДЭС оперировать данными родственных прибалтийско-финских языков. Однако исторический и географический контекст позволяет реконструировать в конечном счете именно вепсские этимоны и тем самым обогатить исторический фонд вепсской лексики.

Таким образом, русские диалекты могут сохранять исходный вепсский этимон, утраченный в самом вепском. Убедительный пример этого рода — рус. диал. *кйкла* ‘наклонная жердь, подпирающая изгородь’ на фоне вепс. *t’ilg* ‘клин’. В результате двух фонетических процессов (развитие *ki > ti* и метатеза в конечном кластере) вепсский этимон значительно изменил свой облик, отчего С. А. Мызников даже засомневался в связи между вепским и русским словом (с. 297). Однако она бесспорно подтверждается прибалтийско-финскими параллелями, в частности, фин. диал. *kiklata* ‘подпирает (дверь)’ (SMS).

Можно обратить внимание на потенциал РДЭС и для анализа вепсского словообразования, которое

исследовано лишь фрагментарно. Ограничимся примером из РДЭС, дополняющим ряд вепсских отглагольных имен с суффиксом *-im*. С. А. Мызников совершенно прав, соотнося бытующие в южном Обонежье рус. *кйтма* ‘веревка, на которой приведена для продажи корова’ и *кйтма* ‘веревка, на которой водят корову’ с карел. *kytkie* ‘привязывать’, вепс. *kütte* ‘привязь для рогатого скота’ (с. 305, 395). Однако остается вопрос о происхождении *-м*-ового элемента. Привлечение финских диалектных данных дает ответ: соответствующая лексическая единица представлена в них в виде *kytkein* или *kytkin* (SMS), где конечное *-in* (вепс. *-im*) — суффикс, образующий названия орудий действия от глагольных основ. Исходя из этого, *кйтма, кйтма* можно возвести к вепс. **kytkim*, которое в полном соответствии с закономерным для вепсской фонетики выпадением гласного второго слога при первом закрытом переходило в **kyt(k)m*. В русских диалектах реально сохранилась, таким образом, производная лексема, утраченная вепскими говорами.

Отмечая исключительный потенциал РДЭС для исторического финно-угроведения, нельзя пройти мимо тех возможностей, которые он предоставляет для реконструкции фрагментов утраченных финно-угорских языков и праязыков на территории Русского Севера и Верхневолжья. Местные русские говоры содержат богатейший финно-угорский субстрат, постоянно привлекающий внимание этимологов. Это чрезвычайно ценный материал, хотя и «неудобный» для этимологии, поскольку его не удастся возвести непосредственно к какому-либо конкретному живому финно-угорскому языку. Материалы РДЭС позволяют, к примеру, увидеть важные фонетические признаки субстратных финно-угорских языков на территории Русского Севера. Это, в частности, заднерядный характер звука *e*, отражающегося в русских заимствованных словах в виде *a* (*вахта* ‘трава на болоте’, *вагмас* ‘болото у реки, лес на болоте’, *шялга* ‘часть леса, лес’, см. РДЭС: 108, 120, 903), что неоднократно привлекало внимание исследователей, а также менее известный факт существования заднерядного местного прибалтийско-финского *i*, который в ряде случаев отразился в русских говорах архангельской области в виде *ы*: *выранда* (также *веранда*) ‘куча сучьев на подсеке’, *нылас* (и *нилас*) ‘намерзающий у берегов лед’, *вйргут* ‘знак-ориентир возле убитого зверя во время морской охоты’. Последнее слово оставлено в словаре без этимологии (РДЭС: 146), хотя и семантически (‘знак-ориентир...’) и фонетически убедительно соотносится с приб.-фин. *virg* ‘метка, зарубка на дереве’ (Saarikivi 2009).

Во Введении С. А. Мызников оговаривает, как уже сказано выше, ограничения, накладываемые на сведения об этимологических связях этимона, его праязыковом облике, что представляется в большинстве случаев оправданным. Но в ряде случаев такие ограничения ведут к неиспользованию важных сведений. Понятно, что при исключительно широком исследовательском поле и огромном объеме охваченного материала некоторые новые этимологические работы и их результаты оказались в РДЭС неучтенными. Стоило бы особенно обратить внимание на изыскания нового поколения финских этимологов (Ante Aikio, Petri Kallio, Janne Saarikivi, Santeri Junttila, Jaakko Häkkinen, Sampsa Holopainen), активно включившихся в исследовательскую работу уже после выхода на рубеже тысячелетий этапного для финно-угроведения словаря Suomen sanojen alkuperä (SSA), и в некоторых отношениях отошедших от представлений, сформировавших идеологию SSA. Эти ученые предложили новые этимологические интерпретации, новый взгляд на этапы формирования финно-угорского языкового сообщества, их хронологию и ареал.

Хотя соответствующие исследования, естественно, не всегда имеют прямое отношение к этимологии русских диалектных финно-угризмов, они актуальны для определения финно-угорских этимонов в целом ряде статей РДЭС. Одновременно они дают несомненный импульс и непосредственно для анализа русской диалектной лексики, позволяя не привязывать этимологию к конкретному современному финно-угорскому языку, а исходить из праязыковых реконструкций, которые в плане фонетического облика и семантического наполнения нередко оказываются более адекватными для территорий Русского Севера и Верхневолжья, где не сохранились живые финно-угорские языки. Так, Я. Саарикиви предложил новую этимологию для многозначного русского диалектного финнизма *вóрга* (в значении ‘лесная дорога, тропа’, ‘охотничья тропа с ловушками вдоль нее’, ‘путь, отмечаемый затесками на деревьях’), которая вводит это слово в один этимологический ряд с олон. *вэрги* ‘зарубки на деревьях’. Последнее считается уже более поздним, карельским заимствованием: из карел. *virka* ‘затеска на дереве’. Особый вокализм *ворга* вызван спецификой адаптации исторически заднерядного слова в русской языковой среде⁵.

Другой пример — глагол *вóймовать* ‘слушать, понимать, воспринимать, знать’: в новых работах

⁵ Подробнее об отношении рус. диал. *вэрги*, *выргут* и *ворга* см. (Saarikivi 2009).

по ряду причин фонетического и семантического свойства предлагается отказаться от ранее предложенных интерпретаций (на которые опирается и РДЭС: 143) в пользу западноуральского **wajma* ‘дух, душа’ с последующей субституцией *a > o* (Saarikivi 2009: 127–131)⁶.

Попутно следует заметить, что в РДЭС хотелось бы видеть больше обсуждений этимологических дилемм между финно-угорскими и исконнославянскими решениями. Так, прослеживающееся в ряде русских говоров развитие **-dl- > -гл-* нередко усматривают в рус. *негла, нёгла* ‘лиственница’, предполагая исходное **ne-edla* (Tesnière 1933: 59; Николаев 1989: 192–193; SP 6: 30). Однако русский апеллятив, известный также в вариантах *мегла́, мигла́, мегли́на* и др., имеет альтернативное и более вероятное объяснение (Я. Калима, А. К. Матвеев, С. А. Мызников и др.) как заимствование из источника типа фин. *neula, niekla*, карел. *niekla*, вепс. *ñegl* и др. ‘игла, хвоя’ (РДЭС, с. 525). Выбирая более перспективное решение, С. А. Мызников не упоминает об альтернативе, что представляется упущением.

Судьбы диалектных слов в контактных ареалах очень прихотливы и запутанны, при их рассмотрении в словаре практически неизбежно оказываются учтенными не все возможные интерпретации. Далее следуют несколько примеров таких неиспользованных возможностей. Масштабы РДЭС таковы, что, обращаясь к русскому материалу прибалтийско-финского происхождения, в рамках настоящей рецензии пришлось остановиться лишь на небольшой его части, а именно, на отрезке словника на *к*-:

ка́бач ‘пучок коры’, для объяснения которого в РДЭС привлекается вепс. *käčak* ‘пучок, клубок, моток (лыка)’ (с. 213), целесообразнее связывать с прибалтийско-финской основой **kapa-*, отразившейся в фин. диал. *kapa* и его суффиксальным производным *kapakka* ‘сильно пересохшая кожа, одежда и т. д.’ (SMS), карел. *kapakka* ‘затвердевший, пересохший (напр., о земле, сене, коже)’ (KKS). Финаль *-ч* в русском слове указывает на возможный вепсский этимон (**kabač*), поскольку восходит к характерному вепсскому суффиксу *-čči*, семантически равнозначному карел., фин. *-kka*;

ка́гача, качага ‘сойка’, для которого в РДЭС приводится не вполне ясно изложенная интерпрета-

⁶ При рассмотрении обсуждаемого лексического материала (и целого ряда других слов, рассматриваемых в РДЭС), следует считаться с еще одним важным и оригинальным этимологическим справочником — (МСФУСЗ) и с другими этимологическими трудами екатеринбургско-свердловских авторов, начиная с А. К. Матвеева. Эта литература, как правило, находит отражение в РДЭС.

ция, исходящая из примарности варианта *качага* (с. 218), заманчиво возводить к прибалтийско-финскому источнику вепсского типа, родственному фин., карел. *kahakka* ‘ссора, крик’, ‘хриплый, осипший, скрипучий (о голосе)’ (SMS, KKS), см. выше об отношении вепс. -*čši* ~ фин. -*kka*. Этимология мотивирована особенностями голоса птицы; *кэ́кка* ‘грибной нарост на березе’ в тверских говорах, для которого словарь не предлагает этимологии («неясно», с. 276), восходит, конечно, к карел. *kiekko* ‘круглая плитка, колесо, диск’ (KKS) с дескриптивными прибалтийско-финскими истоками (SSA 1: 353);

кёрка ‘деревянный кружок, надеваемый на конец шеста, которым при ловле ударяют по воде’ (с. 285) логично возводить к точному карельскому соответствию *kerkkä* ‘деревянный кружок в основании ботала’ (KKS). Стоит отметить, что слово *kerkkä*, будучи многозначным, в восточно-финских говорах бытует в значении ‘шишка (еловая или сосновая)’ (SMS). Это наводит на мысль о том, что приводимое в РДЭС *кэ́чка* ‘еловая или сосновая шишка’ (с. 289) может являть пример искажения лексики при записи или публикации (подобные примеры затрагиваются во Введении к РДЭС, с. 8–11);

кёртик ‘небольшой невод, которым ловят в небольших травянистых озерах’ нет необходимости сопоставлять с саамскими данными (с. 287), особенно с учетом олонечского ареала русского слова. Слово входит в один ряд с рус. *кёрдега* ‘рыболовная сеть’ (с. 282), у которого надежные прибалтийско-финские истоки, ср. фин. диал. *kierre* в Приладожье, а также карел. ливв. *kierre* ‘рыболовная сеть’ (SMS, KKS), исторически **kērdek*. Карельская и финская лексемы здесь приведены намеренно, поскольку они отсутствуют в словарной статье РДЭС *кёрдега*, которая оперирует только предложенной в свое время Я. Калимой вепсской реконструкцией **kerdeg* (с. 282);

кúкка ‘пирог с начинкой’ широко представлено в зоне карельских контактов и вполне логично возводится в РДЭС к карел. *kukko* ‘о выпечных изделиях с начинкой’. Однако отождествление последнего вслед за Калимой и Фасмером с приб.-фин. *kukko* ‘петух’ (с. 364) неправомерно, что вполне убедительно показал М. Бентлин, предложив исходить из нижненемецкого варианта нем. *Kuchen* ‘пирог’ (Bentlin 2008) и вернуться тем самым к интерпретации, предложенной еще в конце XIX в. Э. Леннротом, но позднее забытой исследователями.

Приведенные выборочные примеры этого типа свидетельствуют о важности для исследования истоков русской диалектной лексики северо-запад-

ного ареала такого источника, как Словарь финских диалектов (SMS), оставшегося, к сожалению, не учтенным в РДЭС⁷. Из других перспективных электронных источников, которые стоило бы использовать в РДЭС, надо упомянуть этимологическую базу данных саамских языков⁸.

В РДЭС включено некоторое количество топонимического материала — в основном в качестве своеобразных иллюстраций к бытованию апеллятивной лексики, но имеются и самостоятельные топонимические статьи. Представлен целый ряд ценных фиксаций и убедительных этимологических объяснений географических названий (среди них *Кезанки*, *Кеды*, *Кебялка*, *Кевка*, *Кохта*, *Кугаматка*, *Куткостров* и др.).

Определенный упрек, который можно высказать по поводу этого материала в РДЭС, состоит в отсутствии специальной пометы, отличающей топоним от апеллятива. Лишь в некоторых случаях отмечается, что речь идет о топониме (например, ‘название покоса’), чаще же после заголовочного слова (топоним) в толковании дается непосредственная характеристика названного географического объекта: *Конда* ‘возвышенность’, *Саймига* ‘кладбище’, *Сайкора* ‘возвышенное место’, что создает неверное представление о присутствии в русских говорах соответствующих апеллятивов. Кроме того, некоторые из топонимов и анализируются в РДЭС как апеллятивы, ср. «лексема *саймига*» (с. 709). Не вполне учитывается типология топонимов, не позволяющая, к примеру, реконструировать вепс. *märg* ‘мокрый’ в качестве второго элемента в составе названия лесного участка *Кономарга* (с. 333). Детерминант в подобных композитах должен быть представлен термином, называющим разновидность объекта, в данном случае это, очевидно, вепс. *orgo* ‘лес’.

Некоторые из предложенных в словаре этимологий топонимов (*Кондопога*, *Кивинит*, *Киверно*, *Кулатка*), по-видимому, нуждаются в пересмотре. Так, вряд ли верно связывать первый элемент сложного по структуре топонима *Кондопога* с приб.-фин. *kon'd'i*, *kondie* ‘медведь’ (с. 332). Людиковский оригинал *Kompohd* < **Kondpohj* явно апеллирует к карельскому поселенческому термину с твердой основой *kondu*, *kond* ‘однодворная деревня, состоящая из земельного надела и жилого дома’. Второй элемент топонима *-пога* из карел. *pohj* ‘конец залива’.

⁷ Первые восемь выпусков SMS опубликованы в традиционном бумажном виде, а начиная с 2012 г. словарь пополняется в электронном виде (<http://kaino.kotus.fi/sms/>).

⁸ Etymological database of the Saami languages (<http://kaino.kotus.fi/alg/>).

В РДЭС предложены также этимологии некоторых этнонимов, — возможно, самых сложных для этимологического анализа слов. При этом, кажется, каждое новое поколение исследователей вновь возвращается к этому материалу, иногда к старым интерпретациям. Например, современная феннистика отказалась от саамской этимологии этнонима *чудь*, приводимой в SKES (на нее опирается и РДЭС, с. 887), предпочтя вернуться в определенных позициях к традиционному германскому этимону, но не в его готском облике *þiuda* ‘народ’ (Фасмер 4: 378), а в восточно-норвежской реконструкции **þjūd* ‘люди’ — исходя из того, что среди всех германцев именно варяги были в самом близком контакте с чудью — населением Новгородской земли. Согласно новейшей интерпретации, заимствованный из языка новгородской чуди апеллятив **suuti* ‘люди’ был усвоен в язык славянского населения Новгородской земли как этноним *чудь* и впоследствии по народной этимологии сближен с лексемами *чудо*, *чужой*, что способствовало закреплению этнонима в фольклоре (Kallio 2015: 91). В свою очередь, при объяснении *корелы* в РДЭС (с. 341—342) не нашла отражения получившая наибольшее признание в феннистике германская этимология, возводящая этноним к прагерм. **harja* ‘группа людей’, адекватная с точки зрения типологии этнонимов, хотя и вызывающая некоторые фонетические вопросы (SSA 1: 313; Grünthal 1997: 75—78).

При этом толкования этнонимов и иллюстрации в этнонимических статьях РДЭС содержат замечательные редкие материалы, ценность которых бесспорна даже при том, что предлагаемые этимологии в ряде случаев дискуссионны.

Далее следуют выборочные дополнения и/или альтернативы к отдельным этимологическим решениям, предлагаемым в разных статьях РДЭС, содержание которого, как уже сказано выше, посвящено отнюдь не одному только прибалтийско-финскому материалу. Можно заметить, что в ряде случаев в РДЭС вердикт «неясно» был несколько поспешным, т. к. этимон лежит на поверхности.

бáло ‘орудие для гнутья санных полозьев’. С. А. Мызников присоединяется к мнению о заимствовании коми *бала*, *бава* ‘приспособление для гнутья санных полозьев’ из русского (с. 71), что справедливо. Не совсем точно однако считать, что рус. *бáло* из *гибало* (< **gyba-dlo*). Исходным для *бало* может быть только **gъba-dlo* (от **gъbati* ‘гнуть’, см. (ЭССЯ 7: 187)), откуда также рус. диал. *гбáло*, *гáло*, *мгáла* ‘то же’ (СРНГ 2: 65, 83, 90; 6: 116; 18: 63; Матвеев 1964: 287). Не совсем точное возведение к *гибало* дается в РДЭС и при специальном обсуждении упомянутых *гáло*, *мгáла* (с. 155, 492);

бамбáк ‘пряжа для вязки сетей’. Приводимое в РДЭС сравнение с тур. *bükte* ‘крученые, сученые нитки’ (с. 73) нецелесообразно. Слово тождественно известному в русских говорах на Дону *бамбáк*, *бамбúк* и т. п. ‘нить для вязки сетей’, *банбáк* ‘заграничная нитка для сетей’, которые, видимо, из укр. диал. *бомбáк* ‘фабричная пряжа’, далее из польск. *bombak* ‘вид полотна’ или из молд. *бумбáк*, рум. *bumbác* (ЕСУМ 1: 228);

бзні́ка ‘ягоды паслена, при употреблении вызывают рвоту и понос’, также *бздні́ка*, *бздник*, *бзду́ніка* и др. (СРНГ 2: 287—288). С. А. Мызников справедливо пишет о проблематичности сближения с коми *баз*, *паз* ‘вороника’ (с. 87), но для подтверждения стоило бы указать этимон русского фитонима: от *бздеть*, на что указано у Фасмера (1: 163);

бúга ‘затопляемый половодьем низменный береговой лес и кустарник’. Цитирование этого слова завершается вердиктом «неясно» (с. 94). Но имеется неплохая и в целом принятая специалистами этимология: из праслав. **buga* ‘низкий берег реки, заливаемый водой’, к и.-е. **bheug(h)-* ‘изгибать, сгибать’, ср. к семантике рус. *лукá* ‘изгиб, кривизна’, ‘поворот реки’, диал. ‘поемный луг’ и под., родственное рус.-цслав. *лащи* ‘сгибать’, праслав. **lęk-/lǫk-* (ЭССЯ 3: 77; 15: 62; 16: 141—142; SP 1: 435);

говня́к ‘передовая часть саней’, *говéнник* ‘поперечное крепление в передней части саней’ понимаются как прозрачные по своей структуре, что, видимо, подразумевает связь с *говно́* (с. 164), поскольку речь идет о той части саней, которая ближе всего к лошади (собственно, к ее хвосту). Эту связь С. А. Мызников подтверждает финно-угорскими параллелями вроде вепс. *sitked* ‘гибкий’ — *sit* ‘кал, помет’, но это явно гетерогенные образования. При рассмотрении русского слова следует учесть его варианты, которые не объясняют сближение с *говно́*. Ср. диал. *говéдник*, *говéльник* ‘крепление (в передней части саней)’ (СРНГ 6: 252—253), *говя́дник*, *говя́нник* ‘то же’ волог. (СГРС 3: 60). Вся эта лексика, возможно, из диал. *головéнник* ‘передняя изогнутая часть полозьев саней’, далее к *головá*^о. Не исключено развитие *-оло-* > *-owo-* и выпадение *w*. Следует считаться и с возможностью гаплогогии (или синкопы предударного слога) исходного *головéнник*. Развитие ‘голова’ > ‘передняя часть саней’ повторяется, например, в лтш. (*ragavi galvas*, букв. ‘головы саней, дровней’);

гон ‘удилище’ едва ли как-то связано с вепс. *oŋg* ‘удочка’ (с. 165), но является продолжением праслав. **gōpъ*, имени действия от **gъpnati*, **ženo* ‘гнать’ (ЭССЯ 7: 27; SP 8: 89—91). В свою очередь, рус. диал. *гóнка* ‘лес, сплавляемый в плотях’ не следует сопо-

ставлять, как справедливо пишет С. А. Мызников, с фин. *honka* ‘сосна’ (РДЭС: 165). Можно добавить, что русское слово связано с *гон* ‘плот’ (ср. выражение *гонять плоты*), далее к тому же праслав. **gonь*;

гук ‘мыс, выдающийся узкой полосой в море, озеро, реку’ объясняется (с. 169) как заимствование из коми *гуг* ‘пространство, находящиеся за чем-л., задняя часть чего-л.’. Правдоподобнее голландский этимон: *hoek* ‘угол’, ‘сторона’, ‘мыс’, откуда также нем. (у моряков) *Huk* ‘Landspitze’ (Kluge 1911: 382). Кстати, ошибочно включение русского слова в число рефлексов праслав. **gukь* (ЭССЯ 7: 169). Сама эта лексема расценена как ненадежная (SP 8: 301);

едун ‘песчаная почва’, ‘сухая неплодородная почва’ расценено как неясное (с. 184), но это слово скорее всего тождественно *едун* ‘обжора’, ‘аппетит’ (ср. *едун напал* ‘о прожорливости’), что подтверждается толкованием: *почва пдун* ‘песчаная сухая, съедаящая назем’ (Даль 4: 680);

ёдра ‘корова красной масти’, *ёдрый* ‘желтый с красным, бурым оттенком’. Отмечается сходство с коми *гёрд* ‘красный’ (с. 184), что едва ли можно признать убедительным. Рус. слова тождественны диал. *рёдра* и *рёдра* ‘пестрая белоголовая корова’, ‘кличка рыжей коровы’, *рёдрый*, *рёдрый* ‘рыжий (о рогатом скоте)’ < прасл. **rьdrь(jь)* ‘рыжий, красно-желтый (о скоте и др.)’ (Фасмер 3: 459). Начальное *р-* в *ёдрый* исчезло вследствие диссимилиации *р...р > j...р*;

колеть ‘умирать’ вслед за О. Б. Ткаченко (искавшим в русской лексике меряньское наследие) сопоставляется (с. 324) с вепсским и эстонским материалом, хотя гораздо более реалистично объяснение на рус. почве: тождество с рус. *о-колеть*, укр. *колѣти*, блр. *калѣць* ‘цепенеть’ (Фасмер 3: 129);

коняга ‘выдолбленное из дерева корыто’, *коняги* ‘две лодки, соединенные вместе’ понимаются как переосмысленное по народной этимологии *комляга* (*комляга*) ‘лодка, выдолбленная из дерева’ (с. 333). Видимо, имеется в виду близость к зоониму *коняга*, деривату с суффиксом *-яга* от *конь*. Словам *комляга*, *коняга*, *комляга* в РДЭС уделено немало места (с. 329–330, 333). К сказанному там можно добавить, что рус. *коняга*, *коняги* в значении ‘корыто’, ‘две лодки’, по всей вероятности, обязаны своим обликом не тому, что на них повлияли зоонимы *конь* и *коняга*, но тому, что отражают реликт фонетики древненовгородского диалекта: развитие *мл’ > н’* как в случаях *крень* — *кремль*, *на зени* — *на земли* и проч. (Зализняк 1988: 66);

крут ‘сыр из коровьего или овечьего молока’ расценивается (со знаком «?») как возможное заимствование из «казахского или башкирского»

(с. 353). Можно уточнить, что слово не отделимо от рус. диал. *курут* ‘сыр у алтайцев’ < алт., тел. и др. *qurut* ‘сыр’, ср. др.-тюрк. *qurut* ‘высушенный творог’ (Räsänen: 304);

лбса ‘спокойная поверхность воды’. Упоминаются малоудачные сравнения с иноязычным материалом (с. 451), но не учтена очевидная этимология на основании данных ЭССЯ (16: 91): *лбса* субстантивированная форма ж. р. адъектива **losь(jь)* ‘гладкий, блестящий (о поверхности воды)’. В ЭССЯ (Там же) есть также ссылка на публикацию Я. Калимы, где специально отмечается исконное (нефинское) происхождение русских слов;

лыч ‘картофельная ботва, ботва корнеплодов’. Не без основания принимается влияние коми *лыч* ‘ботва корнеплодов’ (с. 464), но не исключено, что это слово само может быть из русского. Стоило бы учесть этимологию ЭССЯ (17: 18–19), где рус. *лыч* ‘ботва’ толкуется как продолжение праслав. (вост.) **lyčьjь*, букв. ‘относящийся к лыку’ от **lyko*. Правда, это объяснение недостаточно убедительно. Интересен приводимый в РДЭС (с. 464) русский диалектизм *хлыч* ‘ботва овощей’, но отношение *лыч* — *хлыч* пока неясно;

лю́стра ‘зеркало’ дается без комментария (с. 465). Это то же самое слово, что блр. *лю́стра* ‘зеркало’ < польск. *lustro* < итал. *lustrò*;

митусить ‘напряженно всматриваться’, ‘рябить в глазах’, ‘моросить’ (с. 499). При обсуждении этого глагола были бы полезны данные ЭССЯ (19: 60), включающие помимо великорусских другие славянские материалы. Трудно согласиться с суждением С. А. Мызникова об отсутствии родства между *митусить* ‘моросить’ и наречием *митусь* ‘попеременно’. Объединяющей ‘дождь’ и ‘попеременность’ является идея пестроты (Невская 1984: 131);

музя, *муся* ‘рот, губы’ (в Петербурге) — вслед за Я. Калимой предполагается заимствование из эст. *тизи* ‘рот, поцелуй’ через посредство нянь-эстонок (с. 505). Стоит обратить внимание на сходство этих слов с польск. *buzia* ‘рот, лицо’, *dać buzi* ‘поцеловать’, укр. *бузя* ‘лицо, ротик’, нем. *Buss* ‘поцелуй’ и подобным ономастическим материалом (Boгуś 2005: 48);

му́лить ‘мутить воду’, *муль* ‘муть’ — сравнения с прибалтийско-финскими данными справедливо отклоняются (с. 505, 506). Это решение было бы более очевидным при использовании данных, приводимых в ЭССЯ (20: 181–185) s. v. **muliti*;

наракуй ‘гололедица, наледь’ — дается неконкретизированное указание на заимствование из ненецкого (с. 522), что нетрудно уточнить: из нен.

narako 'весенний наст, время наста, весна до ледохода' (Хелимский 2000: 356);

обурд́ать 'есть сырое мясо или рыбу': дается неубедительное сопоставление с саамскими данными (с. 551). В действительности глагол *обурд́ать* является приставочным производным от *бурд́ать* из более раннего *айбурд́ать* (русское слово, упомянутое в ненецком словаре Лехтисало), который связан с рус. диал. *айбарча, айбарча́* 'свежая оленьея кровь, которую пьют олениводы', 'кушанье из оленьея крови с мелко нарезанными кусочками мяса' арх. (СГРС 1: 13–14). Это заимствование из ненецкого, точнее, из нен. зап. (с утратой начального *η*- и возможной заменой *-ць* на *-ч*) *ajbarč* = *айбарць, аябарць* 'есть мясо или рыбу в сыром виде' (Матвеев 1996: 226). Возможно посредство коми (< нен.) *айбарчи* 'струганное мерзлое оленье мясо или рыба (как особое блюдо), строганина' (МСФУСЗ 1: 24–25);

оведь, оудь 'яровая рожь'. Дается сравнение с вепс. *овед* 'яровая рожь' (с. 552), но это слово явный русизм. Связи рус. *оведь, оудь* в общем ясны: к рус. *овень, овьнь* и др. 'то же' (Попов 1974: 54), др.-новг. *овьдь* XIV или XV в. (Зализняк 2004: 638), диал. *юдить* 'зреть', *юдное* (зерно) 'зрелое' (Фасмер 1: 368–369, 149). Праслав. **obudy* в ЭССЯ (30: 239), по-видимому, правильно, но не учитывает др.-новг. *овьдь*;

оп́ока 'ровная подводная мель' (с. 558). Остались неучтенными обширные лексические и иные материалы ЭССЯ (28: 245–248) с изложением разных этимологических объяснений праслав. **obroka*;

пахору́кий 'о неловком, неуклюжем человеке', 'с одной здоровой рукой'. Первая часть сложения сравнивается с вепс. *ра́на* 'плохой' (с. 386). Однако сложение (с которым связана фамилия *Пахорук*) входит в ряд исконных *криворукий, сухорукий, однорукий*. Исходной может быть форма *накору́кий* 'с изуродованной рукой' (СРНГ 25: 159, 293), где первая часть связана с *наки́а* 'кисть руки', 'безрукий (прозвище)', *наки́ла* 'левая рука', *наку́ля* 'сухорукий человек', *наку́ша* 'левая рука' волог. (РЭС 2: 106, s. v. *бакорукий*). Возможно, инлаутное *-х*- появилось под влиянием *пах*;

паэ 'каменистые гряды, образованные выходами коренной породы' справедливо объясняется как заимствование из нен. *пэ* (с. 589). Стоило бы учесть записанное Р. Джемсом в Холмогорах рус. *ро́и* (т. е. *пои*) 'камень' (Stira 1981: 58), а также топоним *Пай-Хой*, ненецкое по происхождению название горного хребта к Северу от Полярного Урала (Матвеев 1987: 134). В русских говорах слово *ро́и* не получило распространения из-за полного доминирования исконного *камень*, также в качестве обозначения

Уральского хребта (в этом случае *Камень* калька с ненецкого, коми и мансийского);

писяга 'тонкое одеяло' дается как возможное саамское заимствование (со знаком «?»), без указания источника (с. 611). Между тем, указывался ненецкий источник: *pī́* (*pīs-* в косвенных падежах и сложениях) 'покрывало из шкур, сукна для женских нарт' с добавлением суф. *-яга* (Хелимский 2000: 356);

под 'прорубь', *пода* 'то же', *запускная пода* в говорах Томской обл. Согласно РДЭС (с. 615), не заимствования из селькупского (как считал Е. А. Хелимский), а исконные слова, связанные с рус. диал. *подда́ча, поддавальная прорубь* и *под*. 'прорубь, в которую опускают невод'. С. А. Мызников пишет, что место селькупской лексемы на самодийской почве неясно, отсюда вывод, что если заимствование в русский и было, то только обратное. Непонятно, однако, каким образом *поддача* может превратиться в *под(а)*. Речь идет скорее всего о случайном сходстве, не более того. У селькупской лексемы как будто не найдены соответствия в других самодийских языках, но таких лексем немало. Этимон рус. *пода* на уровне синхронно-фонологической реконструкции выглядит как *pōta* (Хелимский 2000: 372), его отношение с засвидетельствованными селькупскими данными (*pōd*), а также распространение рус. *под(а)* (только в зоне распространения селькупского языка, в том числе в качестве субстратного) свидетельствует о заимствовании в русские говоры от (ассимилированных) селькупов;

пóтага 'кнут' дается с указанием на возможное латышское происхождение (с. 621). Действительно, речь идет об обратном заимствовании из лтш. *pātaga* 'кнут', которое наряду с лит. *botāgas* происходит из др.-рус. *батогъ*;

потма́р 'пространство под русской печью (для ухватов и т. п.)', также *батма́р* и др. Приводятся правильные сопоставления с волжско-финским и др. материалом: морд. (эрзя) *потмар* 'поставец, шкаф для посуды', (мокша) *потмар* 'ящик под лавкой для посуды' (с. 622). Однако эти слова все-таки из русского, что следует из наличия надежного русского этимона: *под* + *нары*, ср. диал. *подма́ры* 'нары для спанья', *подна́ры* 'место под нарами' (Добродомов 2010: 105–110);

пра́за 'земля, сданная в аренду' (с. 624–625). Было бы полезно учесть данные древнененовгородского диалекта и комментарий А. А. Зализняка: *празка* < **празд-ък-а* от *празда* или (?) *праздь* 'арендная плата' (см. подробнее (Зализняк 2004: 620–622));

пúрищик ‘пекарь’, *пúрня* ‘хлебопекарня’, дается с вопросом «из кавказских языков?» (с. 637). Действительно, речь идет о производных от заимствования из груз. *piri* ‘хлеб’, греческого происхождения (Климов, Халилов 2003: 189);

пэндель ‘кисть для побелки известью’, определено как неясное (с. 644). Из польск. *peźdel* ‘кисть’;

пэнкнуть ‘лопнуть, треснуть’, определено как неясное (с. 644). Из польск. *peknąć* ‘то же’;

пюки ‘шкура вместо постели’. Дается с неконкретизированной ссылкой на самодийское происхождение (с. 644). Речь идет, несомненно, об искаженном диалектизме (весьма известном на севере Западной Сибири, на Русском Севере) *нюки*, мн. от *нюк* ‘полотнище из нескольких сшитых оленьих шкур’ < хант. (Казым) *ńyki* и др. (Steinitz 1960: 507; Фасмер 3: 92);

сян ‘повод в оленьей упряжи’, ‘костяшки и деревянные в оленьей упряжи’ — дается сопоставление с рус. *сса* ‘узкий длинный ремень в оленьей упряжи’ (с. 757). В действительности *сян* из нен. *śan* ‘одна из двух плоских костяшек, скрепленных ремнем на узде оленьей упряжи’ (Хелимский 2000: 357);

тремонтан ‘горный ветер’, *тремонтано* ‘северный ветер’. Дается сравнение с рум. *tare* ‘сильный, крепкий’ и *montan* ‘горный’ (с. 799). Более точное объяснение в том, что русские слова — варианты названия северного ветра *трамонтана* из итал. *tramontana*, в числе прочих черноморских отголосков средиземноморской лексики (особенно связанной с мореплаванием), занесенной на Черное море генуэзцами и венецианцами (Богородский 1979: 101, 103, 106);

халеть ‘умирать’ (с. 841). К приводимым в РДЭС соображениям стоило бы добавить данные ЭССЯ (8: 14), в особенности сопоставление со словац. диал. *hal’es* ‘чахнуть, усыхать’ и оговорку о возможном проникновении из языка словацких разносчиков-«венгерцев» в русские жаргоны;

хапорúкий ‘неопытный, неумелый’ рассматривается как композит, где *хапо-* сравнивается с ливв. *hapattiñ* ‘гнилой’ и под. (с. 843). Но возможна метатеза *хапорúкий* < *нахорúкий* (см. выше);

хаутуры ‘похороны’. Ставится вопрос, можно ли сопоставлять с фин. *hauta* ‘могила’ (с. 845). Ответ может быть лишь отрицательным уже ввиду географии русского слова: смоленское, западное, по существу, белорусское, ср. блр. *хаўтуры* ‘погребение’, *хаўту́рны* ‘погребальный’. Далее тождественно рус. *халту́ра*, диал. ‘поминки’, о котором см. (Фасмер 4: 218) и (Хелимский 1990: 35—38)⁹;

хиб ‘шея, загривок у животного’. Справедливо сравнивается с блр. *хи́б* ‘щетина у свиньи’ (с. 846), но явно не хватает ссылки на ЭССЯ (8: 154), где дана и этимология слова (к праслав. **xybati*). Упущение еще более заметно при рассмотрении в РДЭС (с. 846) апеллятивов *хибен*, *хибены* ‘горные возвышенности на Кольском полуострове’ и названия гор *Хибины* в Мурманской обл., которые рассмотрены В. Л. Васильевым (2018: 78), сделавшим вывод о славянских истоках этих русских слов и окончательно отклонившим обсуждаемую в РДЭС (с. 847) возможность их финского происхождения;

холь ‘чистота как результат ухода, заботы’, *хольный* ‘чистоплотный’ сравниваются с карел. *huoli* ‘забота’ и под. (с. 851), но не учтены данные ЭССЯ (8: 61), которые значительно усиливают предпочтительность исконной этимологии слова *холь*, деноминатива от *хóлить*;

хорéи ‘шест, которым погоняют оленей’ — дается правильное сравнение с коми *хареи* (с. 851), но не учитывается более точный этимон, установленный К. Редее и Е. А. Хелимским: нен. *χarei* (в современном ненецком не сохранилось), откуда и коми слово. Показательный пример сохранения утраченного источника заимствования в заимствованном слове (в данном случае в русском и коми);

хúста ‘кусочек холста’ сравнивается с фин. *hursti* ‘дерюга, мешковина’ и под. (с. 854), но *хúста* явный полонизм, ср. польск. *chusta* ‘платок, перевязь’ (о последнем см. (Boryś 2005: 72));

червя́к и *червя́к* ‘серп’ (с. 857) можно расценить как ценное дополнение к праславянскому диалектизму **čьrvь* ‘серп’, который реконструируется только на основе рус. диал. *черв* (*черп*) ‘серп’ и сблизается с лит. *kīvois* топор секира и под. (ЭССЯ 4: 171; SP 2: 271). Слово *черп* ‘серп’ сравнивается в РДЭС с вепс. *čirp* ‘серп’, допускается «финно-угорское воздействие» (с. 869). Скорее всего, в вепсском русизм;

челпа́н ‘круглый хлеб’, ‘округлый курган’. Имеющиеся объяснения противоречат друг другу (с. 867—868). Стоило бы указать, что в пользу исконной этимологии (к праслав. **čьlрь*) названия кургана свидетельствуют древненовгородские топонимы *Челпино*, *Челпаново*, *Челпново* (Зализняк 2004: 427), см. также (ЭССЯ 4: 144). Для *челпа́н* ‘круглый хлеб’ имеется коми этимология (коми *чóвтон*, см. (КЭСК: 310)), но можно думать о развитии ‘округлая вершина’ > ‘округлый предмет, хлеб’ у исконного по происхождению русского слова;

⁹ Предлагаемая в этой статье монгольская этимология рус. *халту́ра* не была включена ее автором в книгу (Хелим-

ский 2000), что, видимо, свидетельствует об отказе от данной этимологии.

чичер 'снег с дождем' дается как неясное (с. 882). Стоило бы по крайней мере упомянуть этимологию в ЭССЯ (4: 110) — к праслав. *čičerъ;

шиши́бары 'комья застывшей грязи', шишобáрчик 'неумытый непричесанный ребенок'. Упоминается сопоставление (А. К. Матвеев) с коми шишы́бар 'вид растения', 'репей' с оговоркой о возможном русском происхождении коми слова (с. 929). Стоило бы учесть русский диалектизм шишобар 'чертополох', сложение шиши 'черт' и -бар от борóть(ся) (Меркулова 1967: 96);

ягáрма 'бранчливая баба' и под. Принимается заимствование из коми ягарма 'бойкая баба' (с. 956). Стоило бы учесть возможность исконной этимологии, предлагаемой в «Праславянском словаре», где рус. диал. ягáрма, игáрма, егáрма, ягáрьма сближается с ягáть 'кричать, шуметь' (Даль 4: 693) < праслав. *egati (SP 6: 96).

В заключение уместно пожелать автору РДЭС при дальнейшей работе над словарем по возможности повысить требовательность к предлагаемым этимонам и устранить имеющиеся в тексте шероховатости в изложении.

* * *

Этимологические словари всегда дают основания для многочисленных замечаний, поправок и уточнений. Не является исключением и РДЭС, что ни в коей мере не отменяет общей сугубо положительной оценки этого фундаментального труда, выдающегося вклада в русское и финно-угорское языкознание. С.А. Мызникова следует поздравить с большим и заслуженным успехом.

Литература

- Богородский, Б. Л. 1979. Загадочное ли слово горбин? *Диалектная лексика 1977*: 100–106.
- Варбот, Ж. Ж. 1997. Рец. на кн.: Boryś W., Popowska-Taborska H. *Słownik etymologiczny kaszubszczyzny*. Т. 1. А–Ї. Warszawa, 1994. *Этимология 1994–1996*: 189–193.
- Васильев, В. Л. 2018. О древненовгородских языковых следах в регионах Восточно-Европейского Севера и Сибири. *Вопросы языкознания 3*: 72–88.
- Даль, В. И. 1955. *Толковый словарь живого великорусского языка*. 2-е изд. Тома 1–4. Москва: Государственное издательство иностранных и национальных словарей.
- Добродомов, И. Г. 2010. Из лексики Волго-Камского языкового союза. *Урало-алтайские исследования 2 (3)*: 105–111.
- ЕСУМ = Мельничук, О. С. (ред.). 1982–2012. *Этимологічний словник української мови. Томи 1–6*. Київ: Наукова думка.
- Зализняк, А. А. 1988. Древненовгородское койне. *Балто-славянские исследования 1986*: 60–78.
- Зализняк, А. А., А. А. Гишпиус. 2018. Берестяные грамоты из раскопок 2017 г. в Великом Новгороде и Старой Руссе. *Вопросы языкознания 4*: 7–24.

- Климов Г. А., М. Ш. Халилов. 2003. *Словарь кавказских языков. Сопоставление основной лексики*. Москва: Восточная литература.
- КЭСК = Лыткин, В. И., Е. С. Гуляев. 1970. *Краткий этимологический словарь коми языка*. Москва: Наука.
- Матвеев, А. К. 1964. Заимствования из пермских языков в русских говорах Северного и Среднего Урала. *Acta linguistica Academiae Scientiarum Hungaricae XIV (3–4)*: 285–315.
- Матвеев, А. К. 1996. Новые данные о ненецких заимствованиях в севернорусских говорах. *Этимологические исследования 6*: 72–79.
- Меркулова, В. А. 1967. *Очерки по русской народной номенклатуре растений (Травы. Грибы. Ягоды)*. Москва: Наука.
- Мызников, С. А. 2003. *Русские говоры Обонежья: ареально-этимологическое исследование лексики прибалтийско-финского происхождения*. Санкт-Петербург: Наука.
- Мызников, С. А. 2004. *Лексика финно-угорского происхождения в русских говорах Северо-Запада. Этимологический и лингвогеографический анализ*. Санкт-Петербург: Наука.
- Мызников, С. А. 2005. *Русские говоры Среднего Поволжья. Чувашская республика. Республика Марий Эл*. Санкт-Петербург: Наука.
- Мызников, С. А. 2007. *Атлас субстратной и заимствованной лексики русских говоров Северо-Запада*. 2-е изд., испр. и доп. Санкт-Петербург: Наука.
- Мызников, С.А. 2010. *Русские говоры Беломорья. Рабочие материалы*. Санкт-Петербург: Институт лингвистических исследований РАН.
- МСФУСЗ 1 = Матвеев, А. К. (ред.). 2004. *Материалы для словаря финно-угро-самодийских заимствований в говорах Русского Севера. Выпуск 1 (А–И)*. Екатеринбург: Издательство Уральского университета.
- Невская, Л. Г. 1984. Лит. *targas* (семантические связи постоянного эпитета). *Славянское и балканское языкознание: Язык в этнокультурном аспекте*: 130–136.
- Николаев, С. Л. 1989. Следы особенностей восточнославянских племенных диалектов в современных великорусских говорах. I. Кривичи (окончание). *Балто-славянские исследования 1987*: 187–224.
- Попов, А. И. 1974. Из истории слов пассивного лексического пласта одного северно-русского говора. *Диалектная лексика 1973*: 51–62.
- РЭС = Аникин, А. Е. 2007–2020–. *Русский этимологический словарь. Выпуски 1–14–*. Москва: Рукописные памятники Древней Руси; Знак; Институт русского языка РАН; Нестор-История.
- СГРС = Матвеев, А. К., М. Э. Рут (ред.). 2001–2014–. *Словарь говоров Русского Севера*. Выпуски 1–6–. Под ред. Екатеринбург: Издательство Уральского университета.
- СРГК = Герд, А. С. (ред.). 1994–2002. *Словарь русских говоров Карелии и сопредельных областей. Тома 1–6*. Санкт-Петербург: Издательство Санкт-Петербургского университета.
- СРНГ = Филин, Ф. П., Ф. П. Сороколетов, С. А. Мызников (ред.). 1965–2019–. *Словарь русских народных говоров. Выпуски 1–51–*. Москва / Ленинград / Санкт-Петербург: Наука.
- Фасмер, Макс. 1986–1987. *Этимологический словарь русского языка*. Тома 1–4. 2-е изд. Москва: Прогресс.
- Хелимский, Е. А. 1990. Этимологические заметки. *Исследования по исторической грамматике и лексикологии*: 30–41.
- Хелимский, Е. А. 2000. *Компаративистика. Уралистика. Лекции и статьи*. Москва: Языки славянской культуры.
- ЭССЯ = Трубочев, О. Н., А. Ф. Журавлев, Ж. Ж. Варбот (ред.). 1974–2018–. *Этимологический словарь славянских языков. Праславянский лексический фонд. Выпуски 1–41–*. Москва: Наука.

References

- Bentlin, Mikko. 2008. *Niederdeutsch-finnische Sprachkontakte. Der lexikalische Einfluss des Niederdeutschen auf die finnische Sprache während des Mittelalters und der frühen Neuzeit* (Suomalais-Ugrilaisen Seuran Toimituksia 256). Helsinki: Suomalais-Ugrilainen Seura.
- Bogorodskij, Boris L. 1979. Zagadochnoe li slovo gorbin? *Dialekt-naja leksika* 1977: 100–106.
- Boryś, Wiesław. 2005. *Słownik etymologiczny języka polskiego*. Warszawa: Wydawnictwo literackie.
- Dal', Vladimir I. 1955. *Tolkovyj slovar' zhivogo velikorusskogo jazyka*. 2-e izd. Toma 1–4. Moskva: Gosudarstvennoe izdatel'stvo inostrannykh i natsional'nykh slovarj.
- Dobrodomov, Igor' G. 2010. Iz leksiki Volgo-Kamskogo jazykovogo sojuza. *Uralo-altajskie issledovanija* 2 (3): 105–111.
- ESSJa = O. N. Trubachev, A. F. Zhuravlev, Zh. Zh. Varbot (eds.). 1974–2018. *Etimologičeskij slovar' slav'anskih jazykov*. Vypuski 1–41. Moskva: Nauka.
- ESUM = O. S. Mel'nychuk (ed.). 1982–2012. *Etymologičnyj slovník ukrajins'koji mowy*. Tomy 1–6. Kyjiv: Naukova dumka.
- Grünthal, Riho. 1997. *Livvistä liiviin. Itämerensuomalaiset etnonyymit* (Castrenianumin toimitteita 51). Helsinki: Suomalais-Ugrilainen Seura.
- Kallio, Petri. 2015. The Language Contact Situation in Prehistoric Northeastern Europe. In: Robert Mailhammer, Theo Vennemann gen. Nierfeld, Brigit Anette Olsen (eds.). *The Linguistic Roots of Europe: Origin and Development of European Languages*. (Copenhagen Studies in Indo-European 6): 77–102. Copenhagen: Museum Tusulanum.
- KESK = Lytkin, Vasilij I., Yevgenij S. Guljajev. 1970. *Kratkij etimologičeskij slovar' komi jazyka*. Moskva: Nauka.
- Khelimskij, Yevgenij A. 1990. Etimologičeskie zametki. *Issledovanija po istoričeskoj grammatike i leksokologii*: 30–41.
- Khelimskij, Yevgenij A. 2000. *Komparativistika*. *Uralistika*. *Lektsii i statji*. Moskva: Jazyki slavjanskoj kul'tury.
- KKS = *Karjalan kielen sanakirja*. 1968–2005. 1–6. (Kotimaisten kielten tutkimuskeskuksen verkkojulkaisuja 18). Helsinki: Suomalais-Ugrilainen Seura.
- Klimov, Georgij A., Madzhid Sh. Khalilov. 2003. *Slovar' kavkazskikh jazykov*. *Sopostavlenie osnovnoj leksiki*. Moskva: Vostochnaja literatura.
- Kluge, Friedrich. 1911. *Seemannssprache. Wortgeschichtliches Handbuch deutscher Schifferausdrücke älterer und neuerer Zeit*. Halle a. d. Saale: Verl. d. Buchh. d. Waisenhauses.
- Matveev, Aleksandr K. 1964. Zaimstvovanija iz permskikh jazykov v russkikh govorakh Severnogo i Srednego Urala. *Acta linguistica Academiae Scientiarum Hungaricae* XIV (3–4): 285–315.
- Matveev, Aleksandr K. 1996. Novye dannye o nenetskikh zaimstvovanijakh v severnorusskikh govorakh. *Etimologičeskie issledovanija* 6: 72–79.
- Merkulova, Valentina A. 1967. *Očerki po russkoj narodnoj nomenklature rastenij (Travy. Griby. Jagody)*. Moskva: Nauka.
- Myznikov, Sergej A. 2003. *Russkiye govory Obonezhja: areal'no-etimologičeskoe issledovanie leksiki pribaltijsko-finskogo proiskhozhenija*. Sankt-Peterburg: Nauka.
- Myznikov, Sergej A. 2004. *Leksika finno-ugorskogo proiskhozhenija v russkikh govorakh Severo-Zapada. Etimologičeskij i lingvo-geografičeskij analiz*. Sankt-Peterburg: Nauka.
- Myznikov, Sergej A. 2005. *Russkije govory Srednego Povolzhja. Chuvashskaja respublika. Respublika Marij El*. Sankt-Peterburg: Nauka.
- Myznikov, Sergej A. 2007. *Atlas substratnoj i zaimstvovannoj leksiki russkikh govorov Severo-Zapada*. 2nd edition. Sankt-Peterburg: Nauka.
- Myznikov, Sergej A. 2010. *Russkije govory Belomorja. Rabochie materialy*. Sankt-Peterburg: Institut lingvističeskikh issledovanij RAN.
- MSFUSZ 1 = A. K. Matveev (ed.). 2004. *Materialy dlja slovarja finno-ugro-samodijskikh zaimstvovanij v govorakh Russkogo Severa*. Vypusk. 1 (A–I). Ekaterinburg: Izdatel'stvo Ural'skogo universiteta.
- Nevskaja, Lidija G. 1984. Lit. margas (semantičeskie svjazi postojannogo epiteta). In: E. I. Zelenina, V. V. Usacheva, T. V. Tsvijan (eds.). *Slavjanskoje i balkanskoje jazykoznanije: Jazyk v etnokul'turnom aspekte*: 130–136. Moskva: Nauka.
- Nikolaev, Sergej L. 1989. Sledy osobennostej vostočnoslavjanskikh plemennykh dialektov v sovremennykh velikorusskikh govorakh. I. Krivichi (okončanie). *Balto-slavjanskije issledovanija* 1987: 187–224.
- Popov, Aleksandr I. 1974. Iz istorii slov passivnogo leksičeskogo plasta odnogo severno-russkogo govora. *Dialekt-naja leksika* 1973: 51–62.
- Räsänen, Martti. 1969. Versuch eines etymologischen Wörterbuchs der Türksprachen. *Lexica Societatis Fenno-Ugricae* XVII (1). Helsinki: Suomalais-Ugrilainen Seura.
- RES = Anikin, Aleksandr E. 2007–2020–. *Russkij etimologičeskij slovar'*. Vypuski 1–14–. Moskva: Rukopisnye pamjatniki Drevnej Rusi; Znak; Institut russkogo jazyka RAN; Nestor-Istorija.
- Saarikivi, Janne. 2008. Itämerensuomalais-slaavilaisten kontaktien nykytilasta. In: Jussi Ylikoski (ed.). *The Quasiquicentennial of the Finno-Ugrian Society*. (Suomalais-Ugrilaisen Seuran Toimituksia 258): 109–160. Helsinki: Suomalais-Ugrilainen Seura.
- SGRS = Matveev, A. K., M. E. Rut (eds.). 2011–2014–. *Slovar' govorov Russkogo Severa*. Vypuski 1–6–. Ekaterinburg: Izdatel'stvo Ural'skogo universiteta.
- SKES = Y. H. Toivonen, Erkki Itkonen (eds.). 1958–1981. *Suomen kielen etymologinen sanakirja*. 1–7. (Lexica Societatis Fenno-Ugricae XII). Helsinki: Suomalaisen Kirjallisuuden Seura.
- SMS = Anonymous. 2011–2021. *Suomen murteiden sanakirja* (Kotimaisten kielten keskuksen verkkojulkaisuja 30). Available online at: <http://kaino.kotus.fi/sms/>.
- SP = Słowski, Franciszek (ed.). 1974–2001. *Słownik prastowiański*. T. 1–8. Wrocław: Zakład im. Ossolińskich.
- SRGK = Gerd, S. A. (ed.). 1994–2005. *Slovar' russkikh govorov Karelii i sopredel'nykh oblastej*. 1994–2005. Sankt-Peterburg: Izdatel'stvo Sankt-Peterburgskogo universiteta.
- SRNG = Filin, F. P., F. P. Sorokoletov, S. A. Myznikov (eds.). 1965–2019–. *Slovar' russkikh narodnykh govorov*. Vypuski 1–2. Moskva / Leningrad / Sankt-Peterburg: Nauka.
- SSA = Erkki Itkonen, Ulla-Maija Kulonen (eds.). 1992–2000. *Suomen sanojen alkuperä. Etymologinen sanakirja*. 1–3. (Suomalaisen Kirjallisuuden Seuran Toimituksia 556, Kotimaisten Kielten Tutkimuskeskuksen julkaisuja 62). Helsinki: Suomalaisen Kirjallisuuden Seura.
- Steinitz, Wolfgang. 1960. Ostjakische Lehnwörter im Russischen. *Zeitschrift für Slawistik*. Bd. 5. H. 4: 483–493.
- Stipa, Günter J. 1981. Vom Ostseefinnischen bis zum Samoje-dischen reichender Lehnwortschatz im "Russian Vocabulary" (1618–1619) des Richard James. *Journal de la Société finno-ougrienne* 77: 41–66.
- Tesnière, Lucien. 1933. Les diphtonges tl, dl en slave: essai de géolinguistique. *Revue des études slaves* 13 (1-2): 51–100.

- Varbot, Zhanna Zh. 1997. Recenzija na knigu: Boryś W., Popowska-Taborska H. Słownik etymologiczny kaszubszczyzny. T. 1. A-Ć. Warszawa, 1994. *Etimologija 1994–1996*: 189–193.
- Vasil'jev, Valerij L. 2018 O drevnenovgorodskikh jazykovykh sledakh v regionakh Vostochno-Evropejskogo Severa i Sibiri. *Voprosy jazykoznanija* 3: 72–88.
- Vasmer, Max. 1986–1987. *Etimologičeskij slovar' russkogo jazyka*. V. 1–4. 2-e izd. Moskva: Progress.
- Zaloznjak, Andrej A. 1988. Drevnenovgorodskoe kojne. *Baltoslavjanske issledovanija 1986*: 60–78.
- Zaloznjak, Andrej A., Aleksej A. Gippius. 2018. Berestjanye gramoty iz raskopok 2017 g. v Velikom Novgorode i Staroj Russe. *Voprosy jazykoznanija* 4: 7–24.