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

№ 22, issue 1—2

Moscow 2024

Российский государственный гуманитарный университет
Институт языкознания Российской Академии наук

Вопросы языкового родства

Международный научный журнал

№ 22, часть 1—2

Москва 2024

Advisory Board:
Heiner EICHNER (Vienna) / Chairman
William H. BAXTER (Ann Arbor, Michigan)
Václav BLAŽEK (Brno)
Larry M. HYMAN (Berkeley)
Frits KORTLANDT (Leiden)
Alexander LUBOTSKY (Leiden)
James P. MALLORY (Belfast)
Alexander MILITAREV (Moscow)
Valentin VYDRIN (Paris)

Editorial Staff:
George STAROSTIN (Editor-in-Chief)
Anna DYBO (Managing Editor)
Tatyana MIKHAILOVA
Mikhail SAENKO
Ilya YAKUBOVICH
Anastasia Garnet ROMANTSEVSKY (Editorial Secretary)

Founded by Kirill BABAEV

© Russian State University for the Humanities, 2024

Редакционный совет:
Х. Айхнер (Вена) / председатель
В. Блажек (Брно)
У. Бэкстер (Анн Арбор)
В. Ф. Выдрин (Париж)
Ф. Кортландт (Лейден)
А. Лубоцкий (Лейден)
А. Ю. Милитарев (Москва)
Дж. Мэллори (Белфаст)
Л. Хайман (Беркли)

Редакционная коллегия:
Г. С. Старостин (главный редактор)
А. В. Дыбо (заместитель главного редактора)
Т. А. Михайлова
М. Н. Саенко
И. С. Якубович
А. Г. Романцевски (ответственный секретарь)

Журнал основан К. В. Бабаевым

© Российский государственный гуманитарный университет, 2024

Вопросы языкового родства: Международный научный журнал / Рос. гос. гуманитар. ун-т; Рос. акад. наук. Ин-т языкознания; под ред. Г. С. Старостина. — М., 2024. — № 22/1—2. — x + 188 с.

Journal of Language Relationship: International Scientific Periodical / Russian State University for the Humanities; Russian Academy of Sciences. Institute of Linguistics; Ed. by George Starostin. — Moscow, 2024. — № 22/1—2. — x + 188 p.

ISSN 2219-3820

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

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

Учредитель и издатель
Российский государственный гуманитарный университет
125047, Москва, Миусская пл., 6

Свидетельство о регистрации СМИ
ПИ № ФС77-73409 от 03 августа 2018 г.
Периодическое печатное издание (4 номера в год)

Подписано в печать 03.10.2024. Выход в свет 11.10.2024
Формат 60×90/8. Уч.-изд. л. 14,0. Усл. печ. л. 24,8
Тираж 1050 экз. Свободная цена
Заказ № 2053

Отпечатано в типографии Издательского центра
Российского государственного гуманитарного университета
125047, Москва, Миусская пл., 6
www.rsuh.ru

Table of Contents / Содержание

Contributors / Сведения об авторах viii
Note for Contributors / Будущим авторам x

Issue 1 / Часть 1

Articles / Статьи

Septhia Irnanda. Acehnese: The dating of Into-Sumatra migration. 3
[*Септиа Ирнанда*. К вопросу о датировке миграции ачех на Суматру]

Anton Kogan. On the etymological stratification of borrowed
Indo-Iranian vocabulary in Burushaski 21
[*А. И. Коган*. Об этимологической стратификации
индоиранских заимствований в языке бурушаски]

Е. Е. Арманд. Рефлексы древнеиранских превербов в шугнанском языке 43
[*Elena Armand*. Reflexes of Old Iranian preverbs in the Shughni language]

А. О. Бадеев. Указательные местоимения в древнеиранских языках:
основные типы употребления и реконструкция праиранской системы 75
[*Artyom Badeev*. Demonstrative pronouns in Old Iranian languages: main types of usage
and the reconstruction of the Proto-Iranian system]

Issue 2 / Часть 2

Articles / Статьи

В. Н. Харьков, А. Ю. Урманчиева, Н. А. Колесников, А. А. Зарубин, Л. В. Валихова, И. Ю. Хитринская, И. А. Грунтов, И. М. Егоров, В. М. Лемская, А. В. Дыбо, В. А. Степанов. Популяционная генетика и лингвистическая компаративистика:
филогения и филогеография гаплогруппы N1a2b
и лингвистическая предыстория самодийцев 109
[*Vladimir N. Kharkov, Anna Y. Urmanchieva, Nikita A. Kolesnikov, Aleksey A. Zarubin, Larisa V. Valikhova, Irina Yu. Khitrinskaya, Ilya A. Gruntov, Ilya M. Egorov, Valeriya M. Lemskaya, Anna V. Dybo, Vadim A. Stepanov*.
Population genetics and comparative linguistics: phylogeny and phylogeography
of haplogroup N1a2b and linguistic prehistory of the Samoyed peoples]

Н. Р. Сумбатова. Соответствия постувулярных согласных в даргинских языках 131
[*Nina Sumbatova*. Correspondences of post-uvular consonants in Dargwa languages]

Mikhail Zhivlov. Hokan I: A review of comparative studies 162
[*М. А. Живлов*. Хока I: обзор исследований по сравнению хоканских языков]

Contributors

Elena Armand — candidate of sciences (Philology), faculty member, Institute of Oriental and Classical Studies of the Higher School of Economics (Moscow), armandlena@yandex.ru

Artyom Badeev — researcher, School of Linguistics of the Higher School of Economics; assistant researcher of the Institute of Linguistics of the Russian Academy of Sciences (Moscow), badeev@iling-ran.ru

Anna Dybo – corresp. member of the Russian Academy of Sciences, doctor of sciences (Philology), professor, head of the Department of Ural-Altaic Languages, Institute of Linguistics of the Russian Academy of Sciences (Moscow), adybo@mail.ru

Ilya Egorov — post-graduate student, Institute for Finno-Ugristics/Uralistics, Ludwig Maximilian University (Munich), i.m.jegorow@gmail.com

Ilya Gruntov – candidate of sciences (Philology), senior researcher, Research Institute for Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences (Tomsk); senior researcher, Department of Ural-Altaic Languages, Institute of Linguistics of the Russian Academy of Sciences (Moscow), grountov@mail.ru

Septhia Irnanda — Ph.D., faculty member, Universitas Serambi Mekkah (Banda Aceh), septhia.irnanda@serambimekkah.ac.id

Vladimir Kharkov – doctor of sciences (Biology), senior researcher, Laboratory of Genomic Identification and Laboratory of Evolutionary Genetics, Research Institute for Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences (Tomsk), vladimir-kharkov@medgenetics.ru

Irina Khitrinskaya – candidate of sciences (Biology), deputy director for science & organization work and external relations, academic secretary, Research Institute for Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences (Tomsk), khitrinskaya@tnimc.ru

Anton Kogan — doctor of sciences (Philology), head of the Department of Asian and African languages, Institute of Oriental Studies of the Russian Academy of Sciences (Moscow), kogan_anton@yahoo.com

Nikita Kolesnikov – candidate of sciences (Biology), junior researcher, Laboratory of Genomic Identification and Laboratory of Evolutionary Genetics , Research Institute for Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences (Tomsk), nikita.kolesnikov@medgenetics.ru

Valeriya Lemskaya – candidate of sciences (Philology), deputy director, Institute of Foreign Languages and International Cooperation, Tomsk State Pedagogical University (Tomsk), lemskaya@tspu.edu.ru

Vadim Stepanov – member of the Russian Academy of Sciences, doctor of sciences (Biology); professor, director, head of the Laboratory of Evolutionary Genetics, Research Institute for Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences (Tomsk), vadim.stepanov@medgenetics.ru

Nina Sumbatova – doctor of sciences (Philology), senior researcher, Department of Caucasian languages, Institute of Linguistics of the Russian Academy of Sciences (Moscow), nina.sumbatova@gmail.com

Anna Urmanchieva – doctor of sciences (Philology), deputy director for research, head of the department of the Russian Federation Peoples' Languages, Institute for Linguistic Studies of the Russian Academy of Sciences (Saint Petersburg), urmanna@yandex.ru

Larisa Valikhova – post-graduate student, junior researcher, Laboratory of Genomic Identification, Research Institute for Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences (Tomsk), larisa.valikhova@medgenetics.ru

Aleksey Zarubin – candidate of sciences (Medicine), junior researcher, Laboratory of Population Genetics, Research Institute for Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences (Tomsk), aleksei.zarubin@medgenetics.ru

Mikhail Zhivlov — candidate of sciences (Philology), lecturer, Institute for Oriental and Classical Studies, Russian State University for the Humanities; senior researcher, Institute for Oriental and Classical Studies, HSE (Moscow), zhivlov@gmail.com

Сведения об авторах

Арманд, Елена Евгеньевна — кандидат филол. наук, ? преподаватель Института классического Востока и античности НИУ ВШЭ (Москва), armandlena@yandex.ru

Бадеев, Артем Олегович — сотрудник Лингвистической лаборатории по корпусным исследованиям НИУ ВШЭ; лаборант-исследователь Института языкознания РАН (Москва), badeev@iling-ran.ru

Валихова, Лариса Витальевна – аспирант, младший научный сотрудник лаборатории геномной идентификации Научно-исследовательского института медицинской генетики Томского национального исследовательского медицинского центра РАН (Томск), larisa.valikhova@medgenetics.ru

Грунтов, Илья Александрович – канд. филол. наук, старший научный сотрудник Научно-исследовательского института медицинской генетики Томского национального исследовательского медицинского центра РАН (Томск); старший научный сотрудник отдела урало-алтайских языков Института языкознания РАН (Москва) grountov@mail.ru

Дыбо, Анна Владимировна – чл.-корр. РАН, доктор филол. наук; профессор, заведующая отделом урало-алтайских языков Института языкознания Российской академии наук (Москва), adybo@mail.ru

Егоров, Илья Михайлович – аспирант Института финно-угроведения/уралистики Университета Людвига Максимилиана (Мюнхен), i.m.jegorow@gmail.com

Живлов, Михаил Александрович — канд. филол. наук, доцент Сектора компаративистики Института восточных культур и античности РГГУ, старший научный сотрудник Института классического Востока и античности ВШЭ (Москва), zhivlov@gmail.com

Зарубин, Алексей Андреевич – канд. мед. наук, младший научный сотрудник лаборатории популяционной генетики Научно-исследовательского института медицинской генетики

Томского национального исследовательского медицинского центра РАН (Томск), aleksei.zarubin@medgenetics.ru

Коган, Антон Ильич — доктор филол. наук, заведующий Отделом языков Азии и Африки Института востоковедения РАН (Москва), kogan_anton@yahoo.com

Колесников, Никита Александрович – канд. биол. наук, младший научный сотрудник лаборатории геномной идентификации и лаборатории эволюционной генетики Научно-исследовательского института медицинской генетики Томского национального исследовательского медицинского центра РАН (Томск), nikita.kolesnikov@medgenetics.ru

Ирнанда, Септиа — Ph.D., преподаватель университета Серамби Мекка (Банда-Ачех), septhia.irnanda@serambimekkah.ac.id

Лемская, Валерия Михайловна – канд. филол. наук, заместитель директора института иностранных языков и международного сотрудничества ФГБОУ ВО «Томский государственный педагогический университет» (Томск), lemskaya@tspu.edu.ru

Степанов, Вадим Анатольевич – академик РАН, доктор биол. наук; профессор, директор, руководитель лаборатории эволюционной генетики Научно-исследовательского института медицинской генетики Томского национального исследовательского медицинского центра Российской академии наук (Томск), vadim.stepanov@medgenetics.ru

Сумбатова, Нина Романовна – доктор филол. наук, старший научный сотрудник отдела кавказских языков Института языкознания РАН, nina.sumbatova@gmail.com

Урманчиева, Анна Юрьевна – доктор филол. наук, заместитель директора по научной работе, зав. отделом языков народов Российской Федерации Института лингвистических исследований РАН (Санкт-Петербург), urmanna@yandex.ru

Харьков, Владимир Николаевич – доктор биол. наук, ведущий научный сотрудник лаборатории геномной идентификации и лаборатории эволюционной генетики Научно-исследовательского института медицинской генетики Томского национального исследовательского медицинского центра РАН (Томск), vladimir-kharkov@medgenetics.ru

Хитринская, Ирина Юрьевна – канд. биол. наук, заместитель директора по научно-организационной работе и внешним связям, ученый секретарь Томского национального исследовательского медицинского центра РАН (Томск), khitrinskaya@tnimc.ru

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. Articles are published preferably in English or Russian, although publication of texts in other major European languages (such as French or German) may be considered by the Editorial Board.

All submissions should be uploaded electronically in MS Word and PDF format, using the online *Manuscript Submission Form* at the official website of the Journal (http://jolr.ru). Each article should be accompanied with information about the author(s) (names, affiliations, contact information), an abstract (not exceeding 300 words) and relevant keywords.

For more detailed guidelines on article submission and editorial policies, please see our website (http://jolr.ru) or address the Editorial Board directly at gstarst1@gmail.com. Inquiries may also be sent by regular mail to the official address of the Journal:

Journal of Language Relationship
Institute for Oriental and Classical Studies
Russian State University for the Humanities
125047 Moscow, Russia
Miusskaya Square, 6

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. Предпочтительные языки публикации — английский или русский, хотя по согласованию с редакционной коллегией возможна также публикация статей на других крупных европейских языках (французский, немецкий и т. п.).

Подача материалов для публикации (в форматах MS Word и PDF) осуществляется через *Электронную форму подачи рукописей* на официальном сайте журнала (http://jolr.ru). К каждой статье обязательно прикладываются сведения об авторах (имена, аффилиации, контактная информация), краткое резюме (не более 300 слов) и список подходящих 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 непосредственно от редакционной коллегии по электронной почте (gstarst1@gmail.com). По различным вопросам с редакцией журнала можно также связаться по обычной почте:

125047 Москва
Миусская площадь, д. 6
Российский государственный гуманитарный университет
Институт восточных культур и античности
В редакцию журнала «Вопросы языкового родства»

Issue 1 / Часть 1

Articles / Статьи

Septhia Irnanda
Universitas Serambi Mekkah, Banda Aceh; septhia.irnanda@serambimekkah.ac.id

Acehnese: The dating of Into-Sumatra migration

The debate on the dating of the Acehnese-Chamic subgrouping has been inactive for nearly twenty years. However, new data on the Acehnese sub-lects and the interior mainland Chamic languages, and the historical studies of the Champa kingdom provide more clues about the Acehnese ancestors' time of migration into Sumatra. Using an etymological analysis approach, this study shows that despite being the first to leave the Chamic groups, Acehnese shares more innovation with the mainland Chamic languages than previously thought. This suggests that Acehnese might have left the group within a close time frame to the highland Chamic languages, which subgrouped in the 10th century CE. Based on the 8th century CE Javanese raids to the southern area of the Southeast Asia mainland, recorded in a Chamic inscription, the disappearance of the Linyi name from the Chinese records around the same date, the appearance of the Lamuri kingdom with its Austroasiatic toponymy, and the historical findings on Srivijaya power across Southeast Asia, all suggests that the Acehnese's Indochina-Sumatra migration is likely to have occurred around the middle of the 8th century CE. The final location of this first Chamic migration to Sumatra indicates that it was an intentional relocation rather than an emergency refuge. Furthermore, post-Proto-Chamic-breakup (PC) innovations and borrowings on the mainland are also found in higher numbers than previously reported. This finding proves the post-PC-breakup Chamic dispersal into Aceh-Sumatra, as Thurgood once claimed, to occur in the 15th century AD. Yet, the variants across dialects for these post-PC-breakup words reported in the present study open a probability of an earlier date of this second wave of Chamic-speaking migrants' arrival in Sumatra[1].

Keywords: Acehnese language; Chamic languages; languages of Southeast Asia; language change.

Introduction

It is now widely accepted that the Acehnese are indeed immigrants from the Indochina of today's Vietnam (Blust 1994, 2000; Thurgood 1999, 2007b). Durie's (1990) Proto-Acehnese and Thurgood's (1999) Proto-Chamic works that included the Acehnese in the reconstruction have provided strong evidence that the Acehnese and mainland Chamic languages were once a single community sharing several innovations. Yet, a significantly low number of Austroasiatic loans in Acehnese when compared with the other Chamic languages serves as a strong indicator that Acehnese missed many centuries of linguistic innovations in the Chamic language (Sidwell 2005, 2006), and is believed to be the first to leave the Chamic language family. Thurgood (1999: 29) argued that the Acehnese ancestors were the most northerly Chamic speakers who populated the northern part of the Champa kingdom with its capital Indrapura. According to Thurgood, other than Acehnese, the Tsat and Northern Roglai speakers are also the descendants of that northern Champa area who ran away from the Vietnamese raid in 982.

Meanwhile, historical records suggest that Acehnese might have existed in Sumatra in the 9th CE. The existence of what seems to be a Hindu-Buddha civilisation called the Lamuri king-

[1] I am deeply indebted to Paul Sidwell for his useful comments and guidance and also for answering all my inquiries related to this topic.

dom, located in today's Banda Aceh and Great Aceh (Daly et al. 2019; McKinnon 1988) areas was reported by some Arabic travellers as early as 9th CE. An Arab geographer Ibn Khurdadhbih mentioned the place as *Rām(n)ī* (رامني), *Lawrī*, *Lāmurī* and other variants (Cowan 1933).

The *lam-* as in Lamuri seems to come from Proto South-Bahnaric's **glam* or Proto North-Bahnaric **lam*, which means 'inside'. The distribution of this *lam-*, along with the morpheme *cot-* meaning 'hill', and *krueng* 'river' in many other place names within Aceh, particularly within the Banda Aceh and Great Aceh areas strongly indicates the presence of a Chamic-speaking community in this northern tip of Sumatra as early as the 9th CE. This historical record of Lamuri thus predates the Champa's northern capital fall (in the 10th CE) that triggered the subgroupings of Tsat and Highland Chamic of Roglai and Rade. This leads me to conclude that Acehnese left Champa at least one century before the event.

Durie (1990) and Sidwell (2005, 2006) suggest that the Acehnese-Chamic split may have occurred much earlier. Sidwell (2005) proposed the split date a few centuries before the common era, citing the percentage of AA loans in Acehnese and their unspecific AA language source(s). However, he later changed his stance to a more recent date, around the 5th century CE, suggesting the Chinese raids to subjugate Champa as a historical event (2006). Durie (cited in Sidwell 2005) also believes that Acehnese left at a much earlier date than the split of Tsat proposed by Thurgood (1999). Durie also once proposed that Acehnese was a sister of Proto-Chamic (PC) spoken geographically closer to the Malay peninsula and was probably part of Funan power in the first millennium, as cited in Sidwell (2005). Despite my agreement with Sidwell's (2006) statement about Champa instability within the first part of the first millennium, his 5th century date hypothesis is merely a statistical linguistic argument, thus lacking objective support. Considering the strategic position of North Sumatra in between the India and China trade route, more consistent records about Lamuri from much earlier centuries should also have been attested. Yet, none of the records about this Lamuri toponymy older than from the 9th CE are available. Furthermore, in the first millennium, from the Chinese report of Linyi (or read as Lâm Ấp in Vietnamese) until their first recognition of Champa in 859 AD, the Vietnam coasts were always a battlefield (Schweyer 2010). All the raids thus presented opportunities for population migration, but which ones most probably drove off the Acehnese ancestors from the mainland so that they ended up in Sumatra is a question the answer to which should include proper historical support.

A recent study of the ancient Chinese court documents by Suzuki (2019) led to the speculation that Linyi in the north was attacked by the Srivijaya Group[2] sometime in the 8th century CE, and the Linyi fleet was destroyed, preventing them from sending tribute to China. Therefore, Suzuki argues, from 757 until 859, i.e. for about 100 years, the Linyi name disappeared from Chinese records due to the impact of this sacking. The Acehnese ancestors' migration to Sumatra could occur around this time. This timing is a more logical proposal because the attackers' origin was the SEA islands (Sumatra, Java and the Malay peninsula). This timing explains the situation with the presence of Mainland-origin Chamic speakers in Sumatra (Modern Acehnese speakers), about which modern linguists have been wondering. At the very least, this timing provides a logical explanation for why the Chamic people should have chosen such a direction.

A possible scenario might be as follows: some Chamic people who believed in Mahayana Buddhism might have decided to submit to the Srivijaya group. Mahayana Buddhism was the

[2] Suzuki believed that Srivijaya was not merely a centralized kingdom in Palembang (Southern Sumatra), but rather a group of kingdoms across the mainland and islands around the peninsula. The Sailendra dynasty of Java, Funan and Chela were believed to be parts of this group.

Srivijaya's official religion at that time, while Linyi was Hindu and praised Siva. The archaeological evidence confirms this religion-rivalry hypothesis as the Shiva statue in Linyi's temple was thrown into the sea by the attackers during the raid (Majumdar 1927). It is unknown which Chamic kingdom hosted the Acehnese ancestors in the Linyi era, but I speculate that they lived in the one in the southern part close to Chenla (Funan, as Durie suggested in Sidwell 2005). From there, the journey might have been taken through the trans-peninsular of Thai-Malay as these areas were protected under the influence of the Srivijaya group around that period (Kelley 2023)[3]. From 9 AD onward, Chamic small kingdoms united into Champa, with a more unified local consciousness and a single Cham ethnicity (Schweyer 2010), whilst in Sumatra, Acehnese ancestors founded Lamuri as one of Srivijaya trade ports, perhaps to accommodate Barus, the camphor producer, in the west coast of Sumatra. Two heads of the *avalokiteśvara* were discovered in Banda Aceh and Great Aceh areas, dated between the 8th and 10th centuries. The fortresses of Indrapatra and Indrapurwa, as well as the old mosque of Indrapuri which was built atop a ruined Hindu-Buddha temple are all found dated between the 8th and 12th centuries. Therefore, I propose the 8th century CE as the correct timing of the Acehnese ancestors' arrival in Sumatra.

Other than this Acehnese subgroup dating from the Chamic family, the debate on the Acehnese language's historical place is made more complex by the open possibilities of other Chamic migrations into the island driven by continuous war on the mainland. Thurgood proposes 15 AD as the logical time and argues that the rise of the Aceh kingdom in the early 16th century CE and the information in the Malay Annals about the Champa prince who fled Champa after the sack of Vijaya in 1470 are two strong pieces of evidence for this second Chamic migration into the SEA islands. Unfortunately, Thurgood does not provide any strong linguistic evidence for this dating proposal.

Consequently, the present article offers linguistic arguments for two important proposals; (a) that Acehnese split from Chamic around the 8th century AD, and (b) that sometime after the break-up of PC on the mainland, another Chamic-speaking group entered the Aceh region and mixed with the already present Chamic-speaking inhabitants.

The data

The present study consults a set of databases both from websites and printed publications as listed in Table 1. In addition to these, a set of Acehnese data was gained from informants speaking various Acehnese lects. They are:

1. A 70-year-old male from Tantuha, Great Aceh
2. A 43-year-old female from Seulimeum, Great Aceh
3. A 67-year-old female from Blang Galang, Pidie
4. A 62-year-old female from Meureudu, Pidie Jaya
5. A 40-year-old female from Seunuddon, North Aceh
6. A 32-year-old female from Langsa, East Aceh
7. A 51-year-old female from Lamno, Aceh Jaya

All the informants above were contacted individually by the author via Whatsapp text, call and/or voice note. Each informant was given a specific gloss (in Bahasa Indonesia) to be translated into Acehnese according to their respective dialect.

[3] Kelley (2023) even suggested that "Shilifoshi" (a Chinese toponym that many historians believe to be a transcription of Srivijaya) was located around today Trang, Thailand during the 8th century.

Table 1. Language Databases Used in the Present Study

Language	Database	Source
Austronesian I	The Austronesian Comparative Dictionary Online (ACD)	https://acd.clld.org/ (Blust, Trussel & Smith 2023)
Austronesian II	Austronesian Basic Vocabulary Database	https://lpan.eva.mpg.de/ (Greenhill, Blust & Gray 2008)
Austroasiatic	The SEALANG Mon-Khmer Languages Project	http://sealang.net/monkhmer/index.htm
Proto-Chamic	Thurgood's PC Reconstruction	Thurgood 1999
Proto-Malayic	Adelaar's PM Reconstruction	Adelaar 1992
Acehnese I	Acehnese-Indonesian-English Thesaurus	Daud & Durie 1999
Acehnese II	Wordlist on the sub-lects of Great Aceh, Pidie, Pidie Jaya, North Aceh and Aceh Jaya	Author's own data
Malayic Dialects	Pusat Rujukan Persuratan Melayu (PRPM)	Dewan Bahasa dan Pustaka https://prpm.dbp.gov.my/
Javanese Dialects	SEAlANG Library Javanese Dictionary Corpus	http://sealang.net/java/dictionary.htm

Results & discussion

1. The Birth of Chamic Language

The Proto-Chamic (PC) language — as reconstructed by Thurgood (1999) — was presumed to have developed a few centuries before the common era in the area of the Mekong River of modern-day southern Cambodia. PC is the result of a mix between a Western Malayo-Polynesian (WMP) language with an Austroasiatic language of mainland southeast Asia, probably a proto-variety of Bahnaric (Grant 2005; Thurgood 1999). The exact geographical situation in which the language was born is unknown, nor is the precise period. Yet, the emergence of what was believed a Chamic kingdom around the 2nd century CE in central Vietnam could be the marking point of the Chamic-speaking community's first political recognition. Before that, they may have simply existed as a scattered minority group of Austronesian, being relatives of Proto-Malayic-speaking seafarers who, just like them, were also fighting for dominance in the ancient trade routes of Southeast Asia.

Blust (1981) proposed that Chamic was the grand-niece of the Moken/Moklen language, the language of the sea nomad minority which has existed since at least 2000 years ago, and in the past was believed to have mediated the trade of goods between the trans-national traders (e.g. Indian, Arab and Roman) and the inlanders[4] around the Isthmus of Kra (Bellina, Blench & Galipaud 2021). Moken/Moklen left the family of WMP before Chamic-Malayic was formed as a single language community. Larish (1999), who reconstructed Proto-Moken/Moklenic, suggested that Chamic speakers were similar to the Moken/Moklenic before settling in. Further research is needed to prove the sea-nomadic life of early Chamic speakers as claimed by Larish, yet I suppose that their activities as mediators between the inland ethnic groups and transnational traders were highly probable and could explain the source of small dialectal variations that existed before the break-up of PC as envisioned by Thurgood (1999), Durie (1990) or Brunelle (2019).

[4] Aslian people (thus in Bellina, Blench & Galipaud 2021); however, for the mainland Chamic inhabitants, the inlanders must have been the Bahnaric groups of Austroasiatic speakers.

Although the Linyi kingdom of the 2nd century AD mentioned in the Chinese source seems indeed to be a Chamic kingdom, no evidence proves that the kingdom represented all the Chamic speakers living in that period. Thus, parallel to the suggestion by Schweyer (2010) and Suzuki (2019), I also regard Linyi as merely a fraction of the overall Chamic-speaking group, probably the first group that adopted Hinduism and flourished into a recognised kingdom. So, during the Linyi era, or before the 9th century CE, the mainland Chamic communities were scattered along the coast of Vietnam from the south to the north, speaking a mutually intelligible variant of PC with few linguistic differences.

2.The First Dispersal of Acehnese

Although the date is debated, all linguists have agreed that Acehnese left the Chamic family first, as demonstrated by the significantly minimal presence of AA loanwords, the abundance of Malay loanwords, distinct phonological and lexical innovations not shared with the other Chamic, and some other arguments. Thurgood (1999) implies that the date of split of Acehnese should be around the same period as that of Tsat in the 10th century. He states:

> "The Chamic dialect chain extended north at least as far as Quang-Tri, where they undoubtedly encountered Katuic speakers….. Under pressure from the north, these Chamic speakers left, probably to become the modern Acehnese of northern Sumatra."
>
> (Thurgood 1999: 20)

Pressure from the north did not start before Vietnamese independence from China, which occurred in 938 AD (Fall 1971). Therefore, the quotation from Thurgood might suggest that the Acehnese left the mainland between 9th and 10th centuries AD.

In another part where he discusses the pre-syllabic clusters of post-PC, he states that:

> "The under-representation in Acehnese and Tsat along with the strongest representation in the highlands languages combines to suggest that many of these forms were borrowed after the Acehnese had left the mainland on the journey that would eventually take them to Sumatra and after the Tsat had broken from the Roglai and gone to Hainan, fleeing the Vietnamese push to the south."
>
> (Thurgood 1999: 98)

In this second statement, Thurgood proposes that Acehnese and Tsat left at two different points in time. Since he believes that the Tsat left in the 10th century, the Acehnese must have left earlier than that. Considering Thurgood's view that the Vietnamese's "push to the south" (which started in the 9th century) was the motive, Acehnese should have left sometime around that same period. In short, the linguistic evidence provided by Thurgood leads him to the conclusion that Acehnese has been in Sumatra for at least 1000 years, i.e. since the 10th century or earlier. Given that Thurgood does not restrict the relevant factors to Champa's conflict with the Vietnamese, a date of the 8th century AD would still fit his linguistic arguments for the date of the separation of the Acehnese subgroup.

Regardless of the reasons, for being the first to leave the family, Acehnese was thought to have missed many linguistic changes in the Chamic family. In terms of shared innovations across the Chamic language, Acehnese was thought to share none with any Chamic languages, thus it could not be subgrouped with any of them. For example, the Acehnese final accretion of velar nasal *-ŋ* in such words as 'five', 'flower', and 'woman' was thought to be an innovation developed in SEA islands, until Lee (1998) finally demonstrated that an understudied Chamic language also had this feature. The Cat Gia Roglai (CGR; the first word is sometimes written as Cac) is a Chamic language spoken in Cat Hai, north of Phan Rang city, in the southern part of Vietnam (Lee 1998). Lee successfully shows two interesting phonological innova-

tions shared between Acehnese and CGR — the accretion of final nasals and the raising of PC's central nasal vowels. We should now discuss the two shared innovations of the Acehnese and the mainland Chamic of CGR according to Lee in more detail.

2.1. The raising of PC's main syllable vowels

Both Acehnese and CGR raised the low back unrounded vowel of PC to central high unrounded in CGR and central or back rounded and unrounded vowels in Acehnese, as shown in Table 2.

Table 2. The main syllable vowels of PC in Acehnese and CGR

Gloss	PAN	PC (Thurgood 1999)	Ac (Daud & Durie 1999)	CGR (Lee 1998)	NR (Thurgood 1999; Lee 1998)	PRC (Thurgood 1999)	Malay
five	*lima*	**lima*	*limʌŋ*	*ləmɨŋ*	*lumã*	*lamɨ*	*lima*
betelnut	–	**pina:ŋ*	*pinɯŋ*	*pinɨŋ*	*pinãŋ*	*panɨŋ*	*pinaŋ*
gold	–	**ʔama:s*	*mɯh*	*mɨh*	*mãh*	*mɨh*	*mas*
to nurse	–	–	*mɔm*	*mɨm*	*mãm*	–	–
land	**tanəq*	**tanah*	*tanɔh*	*tanɨh*	*tanãh*	*tanɨ̆h*	*tanah*
enter	**tama*	**tama*	*tamʌŋ*	*tamɨŋ*	*tamã*	*tamɨ*	**tamaʔ*

To illustrate, the PC low back unrounded long vowel **-a:-* which is preceded by a nasal consonant (as in 'betel nut' and 'gold') became the high back unrounded vowel *-ɯ-* in Acehnese, and a high central unrounded vowel *-ɨ-* in Cac Gia Roglai (CGR) and Phan Rang Cham (PRC). Meanwhile, despite its geographical closeness to CGR, Northern Roglai (NR) retained the low back unrounded vowel, but nasalised it into *-ã-*. This leaves NR as the only Chamic language in the table to retain the low unrounded long vowel of PC. CGR, Acehnese, and PRC also treat similarly the PC final short vowel **-a(h)*, either raising it to /ɨ/ in CGR and PRC, or to /ʌ/ or /ɔ/ in Acehnese (see glosses 'five', 'land' and 'enter' in Table 2).

The abovementioned shared innovations give us the impression that Acehnese, PRC, and CGR were close in the Chamic dialectal chain before the Vietnamese 'push to the south' or before the 9th century, in which Acehnese and CGR formed one group while PRC belonged to another. Meanwhile, as Thurgood once suggested, NR was a lect spoken in northern Champa and grouped with Tsat. Thus, Acehnese-CGR and PRC were the southern lects of the Chamic dialectal chain in the first millennium. This fits the hypothesis proposed in the present study: the Chamic cities attacked by the Srivijaya in the 8th century were Kauthara and Panduranga, both located in the southern coastal area of modern Vietnam. When Srivijaya attacked Kauthara and Panduranga, the Acehnese-CGR population which dwelled in the area split up: the Acehnese left for Sumatra where they further developed the innovations they once shared only with CGR, whereas CGR joined the more northern Chamic speakers which were pushed to the south, and stopped resembling the Acehnese relatives who had fled the mainland.

2.2. The accretion of final nasals

Acehnese and CGR share the accretion of final nasal velar *-ŋ* which suggests a common stage of development for both languages for a few centuries after PC. This innovation in both languages cannot be a coincidence because they occur in a similar environment and within the

same basic vocabularies. Not only do they both develop a final *-ŋ* in the PC open syllable ending in **-a*, but Acehnese and CGR also raise the PC short vowel that preceded the sound; to /ɨ/ in CGR and PRC, and /ʌ/ or /ɔ/ or /ɯ/ in Acehnese.

The nasal velar coda *-ŋ* after the PMP final vowel *-a* or *-au* occurs when the final vowel is preceded by the nasal consonant **-m[n/ŋ]*. Similar change patterns took place in some other WMP languages, such as Malayic and Old Javanese, or languages nearby, such as Batak and Sundaness (see Table 3), indicating that the Acehnese velar nasal accretion did not originate already after the Acehnese migration to the SEA islands; however, this innovation seems to have further spread on Sumatra, postdating the exodus of the ancestors of Acehnese. The presence of similar forms in Javanese can be taken as evidence of contacts between Acehnese-CGR with an unknown dominant language which carried this phonological characteristic to the mainland. Studying the distribution of this PMP velar nasal accretion within Chamic and some other WMP languages, which occurs only when the main initial consonant of the syllable is a nasal sound, I suspect that this pattern was borrowed from an unknown AN language of Borneo during the early first millennium. This language lent out the PMP-descended words for 'women', 'flower', and 'tiger', all of them with a final *-ŋ,* to the ancestral states of Javanese, Sundanese, and Malayic in the islands, as well as to the Acehnese-CGR in southern Vietnam. Specifically, in Chamic the contact with this AN language must have occurred sometime between the 1st and 7th centuries AD in the southern area of Vietnam or during the rise of the Funan kingdom in today's Cambodia.

Let us take a look at PAN **buŋa* 'flower' (Table 3, gloss 6), which is also reconstructible as PMP **buŋaŋ* in Blust's ACD. Both Acehnese and CGR treat the etymon similarly by raising the vowel that preceded it (Acehnese *buŋɔŋ* vs. CGR *puwɨŋ*). Notice that in other WMP languages the vowel that precedes *-ŋ* remains low. PRC also shows a change that is consistent with Acehnese and CGR, raising the main syllable vowel so that the form becomes *piŋu*, except it does not get affected by the accretion of velar *ŋ*, indicating that the shared raising of the vowel in this Acehnese-CGR-PRC cluster predates the velar accretion in Acehnese-CGR. Furthermore, the same innovation which occurred in both Acehnese and CGR items 'five', 'ricefield' and 'enter' suggests that it took place on the mainland when Acehnese and CGR were still a single group.

The PMP etymon for 'tiger', **rimau* (Table 3, gloss 2), is also found with final nasal velar reflexes in all modern varieties of Chamic and in some WMP languages of Java, Bali and Borneo. In Borneo, both variants, with and without the velar nasal final and the diphthong *-au*, are found in the languages of interior hunter-gatherers, pointing to the island of Borneo as the origin of this word. However, no historical records have ever reported the existence of tigers on Borneo island; probably, the Borneo PWMP speakers use the word to refer to the smaller-than-tiger leopard species that lives on their home island. This may be why, in Old Javanese and Old Balinese, the word *mong* specifically refers to a baby tiger, not an adult one. For adult tigers, which in the past were found in abundance in Java and Bali islands, people have their own term *macan*. A phonologically similar pattern is even found in the Khmer-speaking community inland[5]; Thurgood lists this word as a post-PC-breakup borrowing from an unclear source (1999: 361). However, its *-au* diphthong retained in Written Cham *rimauŋ* confirms its AN origin. The diphthong is also retained in Acehnese *rimuəŋ* which shows that the Acehnese left when the borrowing had just occurred or when the monophthongisation to /o/ or /ɔ/ + *-ŋ* in Rade, Jarai and Chru had not yet taken place in this word. Therefore, as a post-PC-breakup

[5] The form *taa mouŋ* តាមោង is a word for 'tiger' in certain Montagnard languages of Cambodia (Khmer) (Headley et al. 1997).

item, this word should have been borrowed into Chamic sometime after the common era. I suggest that this word was brought onto the mainland by the Shailendra's SEA-island colonists/traders who invaded Chenla. They probably spoke a WMP language which could be closer to either Old Javanese or Proto-Malayic. The connection between the mainland and SEA island power (Java-Malay Srivijaya) is proven by the Sdok Kak Thom inscription which mentioned that Jayavarman II, the king of Chenla, declared its independence from Java (Coedes 1975).

On the other hand, in Javanese the final velar nasal accretion seems to be morphological rather than phonological. Old Javanese has *lima* for 'five', *telu* for 'three', but *lima-ŋ puluh* for 'fifty' and *telu-ŋ puluh* for 'thirty'. Meanwhile, Acehnese and CGR do not change any other numbers except for 'five'; thus, even if the final *-ŋ* in these languages was borrowed from the ancestors of an AN-speaking community who settled in the mainland in the pre-Chenla era, the influence was phonological rather than morphological. Perhaps this contact occurred only through trading.

Table 3. The Acehnese accretion of final velar /*-ŋ*/ before PMP final /*-a*/ and /*-au*/

No	Acehnese I	Proto-Chamic (and other Chamic)	Proto-Malay	PMP	Other
1	*inɔŋ* 'female'	**ʔina*	**inaŋ* 'mother, mother's sister' (ACD)	**inaŋ; *ina* 'mother, mother's sister'	*ina-ŋ* 'mother, elderly woman' (Old Javanese)
2	*rimuŋ, rimuəŋ* 'tiger'	**rimɔ:ŋ*	**harimau*	**rimau*	*moŋ* (Old Javanese)
3	*limʌŋ, limɔŋ* 'five'	**lima* *ləmɨŋ* (CGR)	**lima*	**lima*	*lima-ŋ puluh* 'fifty' (Old Javanese)
4	*tamʌŋ, tam tamɔŋ* 'get in, enter'	**tama* *tamɨŋ* (CGR)	–	**tama*	*ka-tama-n* (Old Javanese) '*ka-tama-n*' (Tae' — North Sulawesi)
5	*umʌŋ, umɔŋ* 'ricefield'	**huma* *hmɨŋ* (CGR)	*huma* (Malay)	**quma*	
6	*buŋʌŋ, bunɔŋ* 'flower	**buŋa* *puwɨŋ* (Cac Gia Roglai) *piŋu* (PRC)	**buŋa*	**buŋa*[1] **buŋaŋ*[2] 'flower' (ACD)	
7	*taɲʌŋ, taɲɔŋ* 'ask'	–	**taɲa* 'ask'	**kutaña*	*kutaʔnaŋ* 'ask for or about' (Makassarese)
8	*ɲawʌŋ, ɲawɔŋ* 'soul'	**jawa*	**ɲawa*	**NiSawa* (PAN)	
9	*lɯŋʌŋ, lɯŋɔŋ, lɯŋɔ* 'sesame'	**laŋa*		**leŋa*	
10	*laŋɔŋ, laŋʌŋ* 'insect: green bottle fly'	–		**beRŋaw*	
11	*kɯnɔŋ, kɯnʌŋ* 'touch, meet, hit'	–	*kena* (Malay) **meŋena* (PWMP)	**keNa* (PAN)	

Thus, Acehnese and CGR were formerly a single Chamic lect in southern Vietnam that some time between 2nd and 8th centuries underwent contact with a seafarer community speaking an Austronesian language that was characterised by the accretion of final velar /-ŋ/ in PMP final open syllables of the type *n[m/ŋ/w]a[au].* This innovation took place in the southernmost lects of the Chamic dialectal chain and did not have time to spread to the more northern ones, since most of those southern speakers were displaced by the Shailendras of Srivijaya to their region in Sumatra. The number of Chamic immigrants that arrived in Sumatra after Srivijaya's raid on Kauthara in the 8th century must have been significantly high, considering the Shailendra dynasty's famous strength and dominance. With such a high number of people, it must have been quick and easy to significantly reduce the presence of the local language(s) of Sumatra, while the Sumatran Chamic speakers that became the Acehnese, in contrast, multiplied over a relatively short time.

3. Post-PC-breakup Phonological Innovation & Borrowing: Another into-Sumatra Chamic Migration?

Based on the rise of the Aceh kingdom in the 16th century, the story of the runaway prince from Champa told in the Malay Annals, and the Chinese bell kept in Aceh museum that is dated around the same period as the historical and archaeological evidence, Thurgood (1999) proposes that another migration must have occurred around the end of the 15th century CE as a result of the Vietnamese attack on Vijaya, the southern capital of Champa. In a book review, Brunelle (2007) also presents an overview of Thurgood's stand on the split dates of Acehnese. In a later paper, Thurgood further clarifies his stance on Acehnese's out-of-the-mainland dates of migration: the precise date for the first journey remains unknown, while the second journey must have taken place in the 15th century — the only date that Thurgood is certain of (Thurgood 2007a). He also re-emphasizes the important role played by the Aceh Kingdom that flourished between the 16th and 19th centuries in shifting the other languages in northern Sumatra to Acehnese, which explains the Acehnese's relatively great number of speakers compared to the other Chamic languages.

In this section of the paper, I conduct an analysis of the post-PC break-up wordlist from Thurgood (1999), presenting linguistic evidence for a different Chamic-speaking arrival in Aceh. However, instead of agreeing with the 15th century hypothesis, I suspect an earlier date. The reasons are discussed below.

The phonological innovations and lexical borrowings that occurred after the breakup of PC are believed to be missing or severely limited in Acehnese, thus confirming its relatively early split from PC. Surprisingly, however, several post-PC breakup innovations were spotted in certain understudied sublects of Acehnese, such as those of the Seulimuem and Tantuha villages of Great Aceh; Blang Galang and Garot of Pidie; Meureudu of Pidie Jaya; Seuneddon of North Aceh; and Langsa of East Aceh. Figure 1 shows the Acehnese districts of the old kingdom which once existed around those areas. Furthermore, some omission of sounds in the initial position observed in certain lects of Pidie and Great Aceh could have happened as the result of a situation where the newly arrived Chamic speakers tried to curb their 'odd' dialectal characteristics as the impact of this language mix. Furthermore, many words borrowed after the breakup of PC, which were missing in most Acehnese lects (Thurgood 1999: 346–364), were also found within these specific lects. All of this could be evidence of another wave of Chamic-speaking people migrating to Sumatra after this innovation had taken place on the mainland.

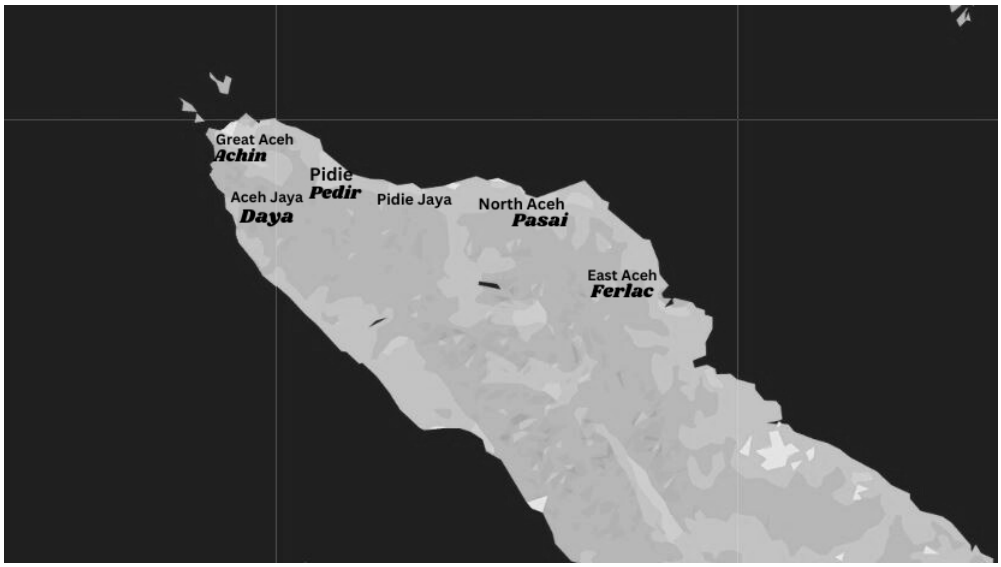


Figure 1. Modern Districts and Old Kingdoms in Aceh

3.1. The post-PC change of /*c-/ to /t-/

A mainland Chamic linguistic innovation proposed by Thurgood (1999: 79), namely, the shift of /**c-*/ to /*t-*/ in word-initial position if the main (second) syllable shares the same initial, turns out to be present in some of the Acehnese sublects. Thurgood (1999: 27) shows that, since this pattern is found only in Jarai, Northern Roglai, Phan Rang Cham and Written Cham, this should reflect the effects of post-PC borrowing. He states:

> "…the apparent failure of Acehnese to participate in this change, with the exception of the word 'lizard', suggests that Acehnese left before much of this occurred" (Thurgood 1999: 78).

Due to Thurgood's limited data, he had only one relevant word ('lizard'), and concluded that the Acehnese left before the pattern of borrowing could become extensive. Yet, in addition to 'lizard', my data from the Acehnese speakers around Seulimuem and Indrapuri of Great Aceh, as well as some villages in Pidie and Pidie Jaya, demonstrate at least three more words that follow this change. With these new data, I revise the table of the relevant changes presented in Thurgood 1999 (see Table 4).

Table 4. The change of PC /*c-/ to Acehnese /t-/ in the initial syllable

Gloss	Malay	Proto-Chamic (Thurgood 1999)	Ac. (Daud & Durie)	Ac. (Seulimuem & some Pidie)	Jarai (Thurgood 1999)	NR (Thurgood 1999)	PR Cham (Thurgood 1999)	Wr. Cham (Thurgood 1999)
grandchild	*cucu*	**cucɔ*	*cucɔ*	*tucɔ*	*təco*	*tico*	*taco*	*tičauv*
chop	–	**cacɔh*	*cicah*	*ticah*	*tĕcɔh*	*ticɔh*	*tacoh*	*tičauḥ*
greatgrandchild	*cicit*	**cicɛ̆t*	*cot*	*cot*	*tacɛ̆ʔ*	*ticɛ̆ʔ*	*tacɛ̆ʔ*	*tičaiʔ*
home lizard	*cicak*	**cicak*	*cicaʔ*	*ticaʔ*	*kăŋkăʔ*	*tacăʔ*	*kacăʔ*	–
bird	–	**cim*	*cicem*	*ticem*	*cĭm*	*cip*	*cĭm*	*cim*
spoon	*camca*	–	*canca*	*tanca*	–	–	–	–
rinse	*cucur*	–	*cuco*	*tuco*	–	–	–	–
to taste	*cicip*	–	*cicap*	*ticap*	–	–	–	–

In Seulimeum, this change of /*c-/ to /t-/ occurs in the same environment as in the other Chamic languages; initial *c-* only shifts to *t-* when the second syllable of the word starts with the same initial. The Acehnese example of 'spoon', which was borrowed from Hindi *camcā* or Persian *čamča*, with a bilabial nasal sound in the middle, also follows this change, indicating that it was still in effect when this word was borrowed (probably between the 12th and 15th centuries AD), or after the Chola raid, which was followed by the occupation of the area by Tamil Nadu traders who later intermarried with the royal natives and founded smaller independent kingdoms along the northern coastal area. Meanwhile, most other Acehnese lects retain the older form with /*c-*/.

3.2. PC /*c-/ to Acehnese /j-/

In Thurgood 1999: 70, Durie explains that PC initial /*c-/ is voiced in Acehnese only when the main syllable also contains a voiceless stop. He demonstrates this with two examples: PC **cupat* 'to squeeze', which becomes *jupat* in Acehnese, and PC **champa* 'Champa', which becomes *juᴍpa* in Acehnese. Indeed, in some Acehnese lects, certain PC words with initial /*c-/ changed it to /j-/; however, in the majority of lects the older form with /*c-/ is still retained. This is particularly true for the gloss 'to squeeze', for which the variant *jupat* was preferred by speakers in the Great Aceh area, while *cupat* characterized other areas from Pidie to East Aceh. The gloss 'Champa' is indeed found only as *juᴍpa* in all attested Acehnese lects.

However, not all Acehnese words which satisfy the condition proposed by Durie have the initial consonant voiced to /j-/. Some items, such as *cukɛh* 'to poke', *cutɔʔ* 'to peck', and *cutīət* 'to pinch' have no variants with /j-/ in any lects. Hence, this *c- > j-* sound change is likely an unstable phenomenon that lasted over a relatively short time in the past and thus affected only a small region.

It is perhaps useful to analyse the two identity-related terms to find further clues, namely, the Champa and Cham. We know that the previous Chamic migration was hinted at by the name of its speakers as a mutation from the toponyms 'Champa' or 'Cham'. In particular, the name of 'Utsat' or 'Tsat' in Hainan comes from the word 'Cham', while the 'Kampong Cepa' in Kelantan comes from the word 'Champa'. Given that the name 'Champa' (as a kingdom) does not appear until the 9th century AD (Schweyer 2010), the existence of the word 'Jeumpa', which probably represents the Champa kingdom, could indicate another Chamic group of migrants in Sumatra after the 9th century. Judging by the incomplete spread of this *c- > j-* pattern, this second Chamic group in Aceh must have been small in numbers and was quickly absorbed by the more dominant Chamic language.

Furthermore, the name *Jeumpa /juᴍpa/* has long been known as the name of Aceh's iconic flower, the *Magnolia champaca*, as well as place names in Bireuen and Pidie areas, before historians proposed that the word could mean a place or kingdom located in today's Vietnam. In Malaysia, *cempa* could mean two things as well: a flower species and a place[6]. This suggests that the word was once a shared term between Acehnese and Malaysian Malay for this sacred plant (compare it to the terms used in Indonesian Malay — *campaka* or *kamboja* 'magnolia flower'). The *cempa* which became *juᴍpa* in Acehnese could also refer to a place — not the Champa kingdom in Vietnam, but merely any small town nearby where many Chamic immigrants used to reside. From this we may conclude that the Acehnese word *Jeumpa* is a term

[6] The word is given as *cempaka*, 'Magnolia flower' in Kamus Bahasa Melayu, Malaysia. The other definition for *cempa* in the same Malaysian dictionary is a place or town, supported by examples from poems. However, it is not defined where the actual location of Cempa is situated.

related to Champa kingdom immigrants and their famous iconic flower. Considering its shared form and meaning with Malaysian Malay, I hypothesize that its presence was also dominant in the Malay peninsula until it was finally absorbed into Malay.

Meanwhile, the other term, *Cham*, seems to evolve into the modern word *Aceh*, the name of the region and its people who speak Acehnese. Unlike the word *Champa* which had its initial sound changed to *j-* due to its plosive initial in the main syllable, the word *Cham* in Acehnese retains its initial consonant, transformed into *Achim, Achin* and finally *Aceh.* The form *Achim*, still retaining the nasal consonant *-m*, is attested in the early 16th century Portuguese documents referring to the area of today's Banda Aceh. Although it appeared for the first time in the 16th century, this does not mean that the term did not exist before that or only existed in that time period. Was the Aceh toponym introduced by Chamic newcomers who migrated there as latecomers? Or could it have existed for centuries before or since the first Chamic dispersal to Sumatra? Without more historical and archaeological evidence, the issue remains problematic. However, in general, the ethnic name *Cham* is older than the kingdom name *Champa*; the name of *Aceh* or *Achin* or *Achim* could have been present in Sumatra as early as the first Chamic migration in the 8th century[7].

On the other hand, the younger word *Champa*, or *Jeumpa* in Acehnese, appears much later. The two places named after this word could give additional clues: the Jeumpa subdistrict in the Bireuen area was an area between the kingdoms of Pasai and Pedir, and the Jeumpa town in the Pidie area was only 6 km into the interior of Batee, the old port of Pedir. These two locations, up to the eastern part of Banda Aceh, are also areas where the *j-* variants are common, cf. such examples as *cupaʔ* > *jupaʔ* 'measurement tool', *cupat* > *jupat* 'to squeeze', *ciŋɯʔ* > *jiŋɯʔ* 'to peep, to keep an eye' and *ciŋklɛt* > *jiŋklɛt* 'to stand on one's toes', which all have corresponding forms in Malaysian Malay (*cupaʔ, jepit, jenguk* and *jengket*, respectively meaning 'measurement tool', 'to clasp', 'visit' and 'tiptoe'). A fuller study of Pedir as an old kingdom in Aceh, and what seems to be its close connection with the Malay peninsula, should be encouraged to find more clues about this second group of Chamic immigrants.

3.3. The voicing of PC /*k-/ to /g-/ in Acehnese

The voicing of the initial presyllabic velar /*k-*/ to /*g-*/ in Acehnese occurs when the initial consonant of the main (second) syllable is plosive (Durie in Thurgood 1999). This change seems to have started early because the voiced variant is universally found in almost all lects across Aceh, while the older unvoiced variant is not found in any lects. Therefore, I suggest that the shift **k-* > *g-* is one of the first independent innovations of the Acehnese as a Chamic language. Additionally, I also observe that this specific voicing phenomenon extends to *k-* in the initial consonant cluster *kl-*, e.g. in PC **kliŋa* 'ear' which became *gliɲuŋ* in Acehnese, and in Dutch *klas* 'class' which became Acehnese *glah* (see Table 5).

Some assumptions may be drawn from this sound change pattern and its distribution across Acehnese lects that follow the change. First, due to being widely attested in both the interior and coastal lects, the change must have started early, or before the move inland had started. This period must have been as early as the 8th to 11th centuries, during which the changes in such basic words like 'foot' and 'ear' took place, and continued until as late as the 15th–17th centuries, as seen from such PM words as **kancĩŋ, *kapur*; the Javanese word for 'pota-

[7] The alternate names for Pasai found in Portuguese documents include *Pachem* and *Pasem*, which might also be connected to the name *Cham* (see Pires & Rodrigues 1944).

Table 5. The voicing of PC initial velar plosive /**k*/ to /*g*/

Gloss	PAN/PMP/PWMP (ACD Website + ABVD Website)	Proto-Chamic (Thurgood 1999)	Proto-Malayic (Adelaar 1992; ACD Website + ABVD Website)	Acehnese
'foot/leg'	**qaqay*	**kakay*	**kaki*	*gaki*
'head louse'		**kutɔw*	**kutu*	*gutɛ*
'nail'		**kukɔw*	**kuku*	*gukɛ*
'shirt button'			**kancĩŋ*	*ganceŋ*
'chalk/limestone'			**kapur*	*gapu*
'cotton'	**kapuk*	**kapa:s*	**kapoʔ*	*gapuh*
'fart'	**qutut*	**kəntut*	**kəntut*	*gɯntot*
'tree sp.'			**ketapaŋ*	*gɯtapaŋ*
'we (incl.)'			**kita*	*gəta +nyoe*
'massage'			**kusuk*	*gusuʔ*
'mountain name'			**kratau*	*gɯrɯtɛ*
'ear'		**kliŋa*	**teliŋa*	*gliŋuŋ*

toes', *kənthaŋ*, which entered Southeast Asia in the 16th century; and the Portuguese word *classe* (or Dutch *klas*) which was borrowed around the 16th–17th centuries. I also suggest that the name of a mountain which separates the Great Aceh and Aceh Jaya areas, Geureutè /*gɯrɯtɛ*/, was also the product of this change from the Gayoness word *krato* and its Malayic equivalent *kratau*, denoting a species of mulberry plant whose leaves are used to feed the silkworm that grows abundantly in the high mountainous area[8] (see Table 5). The spread of this change to both basic terms and words with specifically narrow semantics suggests that Acehnese speakers rely on this special phonological characteristic of the Acehnese language to make it distinct from Malay, their other dominant language.

3.4. The loss or glottalization of the Acehnese plosive /*g-*/, /*b-*/ and affricate /*j-*/ initials

Within Pidie and Great Aceh, there are some Acehnese lects which drop the initial consonant in items with PC **k-*, or Acehnese *g-* as discussed in section 3. In these lects, the loss does not only extend to the Acehnese initial voiced velar /*g-*/, but also to two other initial consonants: the voiced plosive bilabial /*b-*/ and voiced affricative palatal /*j-*/ (see Table 7). The loss mainly occurs in words with reduplication, such as PC **kakay* and PAN **baqbaq*; however, it also takes place, quite exceptionally, in the word *jarɔə* 'finger, hand'. I suppose that it may have been due to its close semantic relationship with 'foot', with the word altered to contain the same initial sound for purposes of harmony (semantic attraction).

Given that another group of Chamic immigrants arrived in the north coastal area of Aceh (today's Pidie), the loss of initial voiced sounds in some words could result from the interaction between two Chamic dialects. The Chamic speakers which arrived later may have adjusted the respective forms in their language to blend with the Chamic speakers already present in the area.

[8] Silkworm farming flourished in North Sumatra and silk was recorded to be one of the trade commodities of the place in the 15th–16th centuries (see Pires & Rodrigues 1944).

Table 6. The loss of initial plosives /k/, /g/ and /b/

Gloss	PAN/PMP/PWMP (ACD Website + ABVD Website)	Proto-Chamic (Thurgood 1999)	Proto-Malayic (Adelaar 1992; ACD Website + ABVD Website)	Acehnese (Daud & Durie, 1999)	Acehnese (Field data from Pidie & Great Aceh)
'foot/leg'	**qaqay*	**kakay*	**kaki*	*gaki*	*aki*
'nail'		**kukɔw*	**kuku*	*gukɛə*	*ukɛə*
'tooth'		**gigɛy*	**gigi/gigoy*	*gigɔə*	*igɔə*
'mouth'	**baqbaq* 'opening, gate'	**mabah* 'mouth'	–	*babah*	*abah*
'lips'	**biRbiR*	–	*bibir*	*bibi*	*ibi*
'hand, finger'			*jari*	*jarɔə*	*arɔə*

4. The post-PC-breakup borrowings

Given Thurgood's implicit assumption that the Acehnese left around the 9th-10th centuries, the PC words listed in Thurgood 1999 as borrowings after the split should be found in small to moderate numbers in Acehnese. Indeed, Thurgood only lists 36 Acehnese words out of 201 as borrowings that took place after the break-up of PC. The list of Acehnese borrowings includes 9 (out of total 20) words of Indian origin, 2 (out of total 3) of Arabic origin, and 25 (out of total 179) of unknown MK origin.

Thurgood's list was rechecked in the present study with the help of Acehnese-speaking informants from Tantuha, Seulimeum & Indrapuri (Great Aceh district), Simpang Tiga (Pidie district), and Meureudu (Pidie Jaya district). The result shows that Acehnese contains more post-PC lexical borrowings than reported by Thurgood. Compared to Thurgood's 36 Acehnese words (out of 201), the present study found 25 more (see Table 7), bringing the total up to 61/201. This means that Acehnese has a total of 30.3%, not 12.4% as previously thought, of post-PC-breakup lexical borrowings in its vocabulary.

Furthermore, the presence of more than one variant for some of these words in Acehnese suggests that they had existed long enough on the island to form a dialect chain. For instance, in the first pattern, where the PC main syllable ending in **-ŋ/n* corresponds to Ac *-ŋ/Ø* with a back high vowel *-u-* in the main syllable, the three Acehnese variants stemming from the same PC word hint at a subgrouping of Acehnese in the past. There are three sets of words with this pattern found in the list: PC **bamɔ:ŋ* 'banana stalk', PC **jhuəŋ* 'long-legged', and PC **kun* 'bundle'. The distribution of the first and second variants for these words in Acehnese dialects is complicated. The first variant for the first item (*mu*) is distributed widely across almost all the lects of Aceh. The second one (*muŋ*) has so far been found in two minority lects — one in the Tantuha village of Great Aceh, recorded from a 70-year old informant, and the other one in the Aceh Jaya sublect of the Lamno area.

Meanwhile, the variants of PC **jhuəŋ* are observed as distributed from Pidie to the east coast of Langsa, forming a dialect chain as *jhuəŋ* and *ɲhuəŋ* within the Pidie and Pidie Jaya area, and *ɲhuə* in North Aceh and East Aceh area. Surprisingly, the variants *ɲhuə* and *ɲhuəŋ* are also found in Tantuha of Great Aceh, indicating a significant merge between the North Aceh and Great Aceh communities sometime in the past. This could result from contact between the Pasai and the Aceh Sultanate in 1524 AD, when the latter conquered the first and made it into a military base to attack Aru and Johor in the Malay Peninsula. Therefore, considering this Pasai-Aceh union in 1524 AD, the arrival of post-PC-breakup *bamɔ:ŋ, jhuəŋ,* and *kun* lexical items into the island must predate the event. I hypothesize that they first started to

Table 7. Additional post-PC borrowings into Acehnese (after Thurgood 1999).

Gloss	PC (Thurgood 1999)	Acehnese (Field data from Pidie & Great Aceh; Daud & Duri 1999)
'man; person'	**manus*	*manuih*
'left (side)'	**iãw*	*wiə*
'morning; dawn'	**ʔəm-aguãh*	*bəŋʌ(ɔ)h*
'banana blossom, stalk'	**bamɔ:ŋ*	*mu:ŋ; mu*
'scoop up; ladle out'	**chɔʔ*	*cɔʔ*
'mark; draw line'	**crɛh*	*creh*
'hide'	**dəp*	*padʌʔ*
'small'	**dɒt*	*dit*
'lie prone'	**gruãʔ*	*cruəp, ʃrup*
'long-legged'	**jhuəŋ*	*jhuəŋ; ɲhuəŋ, ɲuə*
'scratch; claw'	**kuac*	*kuət*
'fold, bundle'	**kun*	*kuə*
'firm; solid'	**kəjap*	*jap*
'clay'	**lən +*kliat*	*tanɔh kliət*
'to swallow'	**luən*	*taluəm*
'rich'	**madar*	*mada*
'to squirt, to spit out, to blow'	**prush/h*	*pruh*
'a comb; hand of bananas'	**sisi(r)*	*isi*
'pull out'	**suac*	*suət*
'shake, tremble'	**tatuh*	*tʌt.tʌt*
'barrel'	**thuŋ*	*thoŋ*
'mynah bird'	**tioŋ*	*tioŋ*
'horn; antler'	**tuki*	*luŋkɛ*
'turn aside, turn'	**wɛh*	*wɛh, wet, wɛ̃t*
'descend'	**yu:ʔ*	*yuʔ*

appear between the 12th and 13th centuries. One other option could be the year 1470, right after the fall of Vijaya on the mainland, as suggested by Thurgood (1999, 2007b); however, earlier dates corresponding to other historical events seem more logical.

The third Acehnese item, *kuə* 'bundle', the presence of which was reported only for the Pidie and Pidie Jaya speakers (without any traces of the older variant *kun*), increase the probability that the Pidie area is the true origin of post-PC borrowings in Acehnese, being the precise location where the second wave of Chamic immigrants had first landed. Additionally, the predominant usage of the Acehnese post-PC words such as *dit* 'small, little', *cruəp, ʃrup* 'lie prone', *tʌt.tʌt* 'tremble, shake' and *yuʔ* 'descend; reduce' by the Acehnese speakers in Pidie and Pidie Jaya area once again strongly confirms them as the initial destinations for this second wave. This is strong evidence in favor of the second wave of Chamic migration to Aceh as suggested by Thurgood (1999); however, precise dating of this event needs further investigation.

In *Suma Oriental*, a Portuguese traveller, Tomé Pires, who visited the area around 1515 AD, wrote interesting remarks about Pedir, the kingdom between Pasai and Achin (the Aceh Kingdom centralised in the most northern tip of Sumatra). He states that:

> Pedir in the island of Sumatra, before 1510, used to be important and rich. It once held the mouth of the channel, had all the trade and was visited by traders more than Pasai. The king of Pedir was always at war with the king of Achin. The king of Pirada (which borders Pasai) was related to Pedir. One of the Pedir King's captains now reigns the kingdom and throws out one of the princes who then runs away to Pasai (Pires & Rodrigues 1944: 139).

This information indicates that in the 16th century, Pedir was an important port that was at war with Aceh which it bordered on the west, but had a more harmonious relationship with kingdoms bordering it from the east (Pirada, Pasai). Therefore, I assume that Pedir and Pasai should have been in a state of intense language contact during this time, which explains why the dialect chain for the lexical item *jhuəŋ ~ ɲuəŋ ~ ɲuə* extends from Pidie to Pidie Jaya to North Aceh respectively. Later (from 1524 and onward), the Pasai-Pedir-Achin union resulted in a dialectal mix which blurred out the linguistic evidence of this second Chamic arrival to Sumatra.

In the meantime, phonological studies on the sublects of Great Aceh, Aceh Jaya and West Aceh[9], with the strong resemblance of their final diphthongs to the PC state, should be further encouraged, as the Acehnese lects spoken in these areas could be the oldest Chamic dialects in Sumatra.

Conclusion

The proposed historical explanation for the evidence discussed above is as follows. Around the mid-8th century, a group of Chamic speakers, skilled in warfare and navigation but without any traces of literacy, sailed from Vietnam to northern Sumatra. There they became subjects to one of the Srivijaya kingdoms centralised in the Palembang area and became known as the Modern Acehnese. The Acehnese left the Chamic community when the southern mainland kingdoms came under attack from the Shailendra dynasty of Southeast Asian islands. The ancestral Acehnese lect during that time was probably spoken in the south of Vietnam, away from Chinese influence and close to the Funan (probably bilingual in Old Javanese and Malay) of today Cambodia. The study suggests that Acehnese, Chamic and Old Javanese relationships should be studied to investigate whether the contact took place on the mainland. The combination of evidence from Chinese and Arabic records, the oldest Hindu-Buddhist archaeological remnants found in Aceh, and the toponymy of Lamuri and other place names around the Great Aceh and Pidie districts support the 8th century hypothesis. Additionally, in line with Thurgood's proposal, this study also suggests a second Chamic dispersal to Sumatra. Evidence for this comes in the form of word variants with characteristics that are similar to post-PC breakup innovations in various sublects within the districts of Great Aceh and Pidie. Statistically, when data from several understudied dialects of Acehnese were included, a much higher number of post-PC-breakup borrowings were found in Acehnese than previously thought. This linguistic insight opens a discussion on the dating of the second Chamic migration into Sumatra, which probably started from the Pidie area. A more refined chronology of contacts between old kingdoms and ports in the area should be produced to support the critical analysis of linguistic variation in Acehnese and ultimately uncover the dynamic of human migration and contacts in the ancient islands and mainland of Southeast Asia.

[9] The reconstruction of the Proto-Acehnese dialect by Durie (1990) was carried out on the basis of only a small fraction of the Acehnese lects, namely, those that spread from the northern coast to the tip of the land, or the Banda Aceh area. The reconstruction thus missed the linguistic variations that exist in the interior lects of the western coast. Recently, some of these understudied lects spoken in the Great Aceh, Aceh Jaya, Nagan Raya and West Aceh districts were investigated anew (Yusuf et al. 2022; Yusuf, Fata, & Karwinda 2021; Masykar et al. 2021; Masykar, Nurrahmi, & Al Mulhim 2021), providing data on Acehnese vowels with a stronger resemblance to PC.

Bibliography

Adelaar, K. Alexander. 1992. *Proto Malayic: The Reconstruction of Its Phonology and Parts of Its Lexicon and Morphology. (Pacific Linguistics: Series C, 119).* Canberra: Research School of Pacific and Asian Studies, Australian National University.

Bellina, Bérénice, Roger Blench, Jean-Christophe Galipaud (eds.). 2021. *Sea Nomads of Southeast Asia: From the Past to the Present*. Singapore: NUS Press.

Blust, Robert. 1981. The reconstruction of Proto-Malayo-Javanic: An appreciation. *Bijdragen Tot de Taal-, Land- En Volkenkunde / Journal of the Humanities and Social Sciences of Southeast Asia* 137: 456–469.

Blust, Robert. 1994. The Austronesian Settlement of Mainland Southeast Asia. In: Karen Adams, Thomas Hudak (eds.). *Papers from the Second Annual Meeting of the Southeast Asian Linguistics Society 1992*: 25–83. Tempe Arizona: Arizona State University.

Blust, Robert. 2000. From Ancient Cham to Modern Dialects: Two Thousand Years of Language Contact and Change (Review). *Oceanic Linguistics* 39(2): 435. doi: 10.2307/3623433

Blust, Robert, Stephen Trussel, Alexander D. Smith. 2023. *CLDF Dataset Derived from Blust's 'Austronesian Comparative Dictionary*.' Available online at: github.com/lexibank/acd. doi: 10.5281/zenodo.7741197

Brunelle, Marc. 2019. Revisiting the Expansion of the Chamic Language Family: Acehnese and Tsat. In: Andrew Hardy, Geoff Wade (eds.). *Champa: Territories and Networks of a Southeast Asian Kingdom*: 287–302. Paris: École française d'Extrême-Orient.

Coedes, George. 1975. *The Indianized States of Southeast Asia*. Honolulu: University of Hawai'i Press.

Cowan, H. K. J. 1933. Lamuri — Lambri — Lawri — Ram(n)i — Lan-Li — Lan-Wu-Li — Nan-Po-Li. *Bijdragen Tot de Taal-, Land- En Volkenkunde / Journal of the Humanities and Social Sciences of Southeast Asia* 90(1): 422–424. doi: 10.1163/22134379-90001421

Daly, Patrick, Edmund Edwards McKinnon, R. Michael Feener, Tai Yew Seng, Ardiansyah, Andrew Parnell, Nizamuddin, Nazli Ismail, Kerry Sieh, Jedrzej Majewski. 2019. The Historic Trading Port of Lamri on the North Sumatran Coast. *Bulletin de l'École Française d'Extrême-Orient* 105: 115–144.

Daud, Bukhari, Mark Durie. 1999. *Kamus Basa Acèh= Kamus Bahasa Aceh = Acehnese-Indonesian-English Thesaurus. (Pacific Linguistics, Series C, 151)*. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Durie, Mark. 1990. Proto-Chamic and Acehnese Mid Vowels: Towards Proto-Aceh-Chamic. *Bulletin of the School of Oriental and African Studies* 53(1): 100–114.

Fall, Bernard B. 1971. The History and Culture of Vietnam. *Naval War College Review* 23(6): 48–54.

Grant, Anthony P. 2005. The Effects of Intimate Multidirectional Linguistic Contact in Chamic. In: Anthony Grant, Paul Sidwell (eds.). *Chamic and beyond: Studies in Mainland Austronesian Languages*: 37–104. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Greenhill, Simon, Robert Blust, Russell D. Gray. 2008. The Austronesian Basic Vocabulary Database: From Bioinformatics to Lexomics. *Evolutionary Bioinformatics* 4: 271–283.

Headley, Robert K., Rath Chim, Ok Soeum. 1997. *Modern Cambodian-English dictionary*. Kensington: Dunwoody Press.

Kelley, Liam. 2023. *Revisiting the Chinese Sources on Early Southeast Asian History. (Universiti Brunei Darussalam Working Paper No. 73)*. Gadong: Institute of Asian Studies, Universiti Brunei Darussalam.

Larish, Michael David. 1999. *The Position of Moken and Moklen Within the Austronesian Language Family (Thailand)*. Doctoral dissertation. Honolulu: University of Hawai'i.

Lee, Ernest W. 1998. The Contribution of Cat Gia Roglai to Chamic. In: David Thomas (ed.). *Papers in Southeast Asian Linguistics No. 15: Chamic Studies*: 31–54. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Majumdar, R. C. 1927. *Champa. History and Culture of an Indian Colonial Kingdom in the Far East 2nd–16th century A.D.* New Delhi: Gyan Publishing House.

Masykar, Tanzir, Roni Agusmaniza, Tabsyir Masykar, Huang Shan, Febri Nurrahmi. 2021. Variation of Acehnese Monophthong /ʌ/ in Western Acehnese Dialect. *Journal of English Language Teaching and Linguistics* 6(3): 695. doi: 10.21462/jeltl.v6i3.655.

Masykar, Tanzir, Febri Nurrahmi, Abdullah Al Mulhim. 2021. Diphthongized Monophthongs of Acehnese Oral Vowels in Samatiga Dialect. *Journal of Language and Literature* 21 (2): 418–428. doi: 10.24071/joll.v21i2.3379

McKinnon, E. Edwards. 1988. Beyond Serandib: A Note on Lambri at the Northern Tip of Aceh. *Indonesia* 46: 103–121.

Pires, Tomé, Francisco Rodrigues. 1944. *The Suma Oriental of Tomé Pires and the Book of Francisco Rodrigues*. Ed. by Armando Cortesao. London: Hakluyt Society.

Schweyer, Anne-Valérie. 2010. *The Birth of Champa. Crossing borders in southeast asian archaeology, Sep 2010, Berlin, Germany*. Ms., available online at: shs.hal.science/halshs-00828812.

Sidwell, Paul. 2005. Acehnese and the Aceh-Chamic Language Family. In: Anthony Grant, Paul Sidwell (eds.). *Chamic and beyond: Studies in Mainland Austronesian Languages*: 211–245. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Sidwell, Paul. 2006. Dating the separation of Acehnese and Chamic by etymological analysis of the Aceh-Chamic lexicon. *The Mon-Khmer Studies Journal* 39: 105–122.

Suzuki, Takashi. 2019. *The History of Srivijaya, Angkor and Champa*. Tokyo: Mekong Publishing.

Thurgood, Graham. 1999. *From Ancient Cham to Modern Dialects: Two Thousands Years of Language Contact and Change*. (*Oceanic Linguistics Special Publications*. *Vol. 28*.) University of Hawai'i Press.

Thurgood, Graham. 2007a. *The Acehnese: Genetic Affiliations, Diversity and Absorption*. Paper presented at The First International Conference on Aceh and Indian Ocean Studies. Banda Aceh: Universitas Syiah Kuala.

Thurgood, Graham. 2007b. The Historical Place of Acehnese: The Known and the Unknown (A Working Paper). In: *First International Conference of Aceh and Indian Ocean Studies*. Banda Aceh. Ms., online at: www.sabrizain.org/malaya/library/acehnese.pdf.

Yusuf, Yunisrina Qismullah, Ika Apriani Fata, Sinta Karwinda. 2021. Oral Monophthong Vowel Qualities of the Jamee Language in Aceh. *Indonesian Journal of Applied Linguistics* 10(3): 794–803. doi: 10.17509/ijal.v10i3.31767

Yusuf, Yunisrina Qismullah, Zulfadli A. Aziz, Faisal Mustafa, Husnul Khatimah. 2022. The unique accent features of the stigmatized Greater Aceh subdialect in Sibreh, Aceh, Indonesia. *International Journal of Language Studies* 16(2): 143–164.

Септиа Ирнанда. К вопросу о датировке миграции ачех на Суматру

Дискуссия относительно датировки истории ачех-чамской подгруппы не возобновлялась уже почти двадцать лет. Однако новые данные, собранные по диалектам ачех и материковым чамским языкам, а также исторические исследования королевства Чампа дают возможность сформулировать новые гипотезы о времени миграции предков ачех на Суматру. Этимологический анализ материала показывает, что, несмотря на то, что ачех были первой ветвью, отделившейся от чамской группы, они разделяют больше общих инноваций с материковыми чамскими языками, чем считалось ранее. Это говорит о том, что отделение ачех могло произойти примерно в одно и то же время с отделением горно-чамских языков (X век н.э.). Учитывая набеги яванцев в VIII веке н. э. на южное побережье Юго-Восточной Азии, зафиксированные в чамской надписи, исчезновение названия Линьи из китайских летописей примерно в то же время, появление царства Ламури с его австроазиатской топонимикой и исторические находки, относящиеся к государству Шривиджая, логично предположить, что миграция ачех из Индокитая на Суматру скорее всего имела место в середине VIII века н. э. Географическое расположение этих первых чамских мигрантов на Суматре свидетельствует скорее в пользу преднамеренного расчета, чем экстренного бегства. Помимо этого, на материковых территориях обнаруживается намного больше инноваций и заимствований, относящихся к периоду после распада чамского единства, чем считалось ранее. Это согласуется с гипотезой о том, что основное расселение чамского этноса в Ачех-Суматре, как было ранее предположено Г. Тергудом, относится к XV в. Тем не менее, диалектные варианты этих слов, появившихся после распада чамской общности, заново ставят вопрос о возможности более ранней даты релокации на Суматру этой второй волны мигрантов, говоривших на чамском языке.

Ключевые слова: язык ачех; чамские языки; языки Юго-Восточной Азии; языковые изменения.

Anton Kogan
Institute of Oriental Studies, Russian Academy of Sciences, Moscow; kogan_anton@yahoo.com

On the etymological stratification of borrowed Indo-Iranian vocabulary in Burushaski

In this paper, I attempt to refine the generally accepted view on the origin of Indo-Iranian loan vocabulary in Burushaski. Having analyzed extensive lexical material, I come to the conclusion that a very large part of this vocabulary was borrowed before the split of the Proto-Burushaski state, and cannot be derived from any of the traditionally identified donor languages (i.e. Shina, Khowar, Urdu, Persian and East Iranian Pamir languages). It is also demonstrated that this "pan-Burushaski" segment of the Indo-Iranian lexical stratum is etymologically heterogeneous, with some Dardic lect as one of the possible sources.

Keywords: Burushaski language; Indo-Iranian languages; Dardic languages; language contact; lexical borrowing; etymology.

Introduction

Burushaski has for centuries been and still is in close contact with a number of Indo-European languages belonging to the Indo-Iranian subfamily. The Burushaski-speaking area is almost completely surrounded by the speaking areas of several Dardic and East Iranian lects.[1] In addition, certain Indo-Iranian languages were and are used by the Burusho people for official and literary purposes. The natural corollary of this situation is the presence in the Burushaski vocabulary of numerous Indo-Iranian loanwords borrowed from different sources. Among these sources the following are identified by most scholars: Classical Persian (the official and literary language of the region in the Middle Ages and later till the 20th century), Urdu (the official and literary language of the region nowadays), East Iranian languages (Wakhi and Ishkashimi), Dardic languages (Shina and Khowar). Khowar and Shina are believed to have exerted the strongest influence on the Burushaski lexicon.

Although the sources of Aryan loans are often reduced to the aforementioned four, there are good reasons to consider the above list incomplete. As early as the 1940s, Georg Morgenstierne demonstrated the possible existence of certain Sanskrit lexical elements in Burushaski (Morgenstierne 1945).[2] The presence of such elements seems to be quite natural if we bear in mind that Buddhism was most probably spread in the area under study in pre-Islamic

[1] It is only in the southeast where the immediate neighbor of Burushaski is not an Aryan language. Here the Burushaski-speaking area borders on the Tibetan-speaking one.

[2] The presence of early Middle Indian loanwords is also possible. They may include, e.g. the Burushaski names of the days of the week: *adít* 'Sunday' (cf. OIA *ādiya-* 'sun', *ādiyavāra-* 'Sunday'), *ċandúra* 'Monday' (cf. OIA *candravāra-* id.), *aṅáro* 'Tuesday' (cf. OIA *aṅgāraka-* 'the planet Mars', Sindhi *aṅāro* 'Tuesday'), *bódo* 'Wednesday' (cf. OIA *budha-* 'the planet Mercury', *baudha-*, *budhavāra-* 'Wednesday'), *biréspat* 'Thursday' (cf. OIA *bṛhaspati-* 'name of a deity, the planet Jupiter', *bṛhaspativāra-* 'Thursday'), *śúkuro* 'Friday' (cf. OIA *śukra-* 'the planet Venus', *śukravāra-* Friday'), *śimśér* 'Saturday' (cf. OIA *śanaiścara-* 'the planet Saturn', *śanaiścaravāra-* 'Saturday'). The Indo-Aryan origin of these words being obvious, they represent a more innovative stage of phonological development as compared to Old Indian. At the same time, their phonology contains some archaic traits (e.g. the retention of intervocalic consonants) lost in later Middle and New Indo-Aryan. However, it cannot be ruled out that the cited lexemes are very early, phonologically adapted borrowings from Sanskrit.

times.[3] In the cited article Morgenstierne also adduces two examples of Aryan borrowings with very archaic phonology:

1) *-fʌltʌs* 'to break' (Morgenstierne 1945: 93), cf. OIA *sphaṭati* 'bursts', *sphāṭayati* 'causes to split', Old High German *spaltan*, German *spalten* 'to split' < PIE **(s)p(h)el-t-* (LIV: 577; Pokorny 1959: 985–987);
2) *phʌltočiŋ* 'puttees' (Morgenstierne 1945: 93), cf. OIA *paṭṭa-* 'cloth, bandage', Hindi *paṭṭī* 'strip of cloth, ribbon, puttee' (> English *puttee*), Old Church Slavic *platъno* 'cloth, canvas, fabric', Old High German *faltan*, Old English *fealdan* 'to fold' < PIE **pel-t-o-* (Pokorny 1959: 803, 804).

The most conspicuous historical-phonological feature of the above two loans is the consonant cluster *lt* corresponding to a retroflex stop (*ṭ* or *ṭṭ*) in Old Indian. This kind of correspondence suggests that in the source-language, clusters of the type "*l* + dental" were not affected by Fortunatov's law, and this language could, therefore, hardly have been Indo-Aryan or modern Dardic.[4] An Iranian source also seems to be unlikely. Apart from the fact that the usual Iranian reflex of PIE **l* in clusters is *r*, no cognates of the aforecited words are attested in the Iranian languages that are believed to have influenced Burushaski. All of this indicates that etymological stratification of Aryan loans is still far from clear and needs further research.

In order to throw some light on this issue I have conducted a study of Aryan vocabulary contained in Hermann Berger's Burushaski-German dictionary (Berger 1998b). Its main advantage is being the biggest bilingual dictionary of Burushaski and a European language published so far: it includes lexical material representing all the local varieties. Before proceeding to the discussion of particular etymologies, I find it necessary to offer some general considerations.

Probably the most significant fact revealed in the course of working with Berger's dictionary is that a large part, if not the majority of Aryan loanwords are common for all the dialects of Burushaski. The importance of this point can hardly be overstated if one bears in mind the linguistic geography and ethnic history of the region in question. The Burushaski language consists of two divergent lects spoken in the Gilgit-Baltistan administrative territory of Northern Pakistan.[5] The western lect, called Werchikwar, is the vernacular of the Yasin valley, and the eastern one, sometimes termed "Burushaski proper", predominates in the basin of the Hunza river. The latter dialect is in turn subdivided into two subdialects — the Hunza and the Nager (Nagar, Nagir)[6]. The Yasin and Hunza valleys do not directly adjoin each other, being separated by the long valley of Gilgit, nowadays populated by speakers of the Shina language.

[3] This fact is indicated by rock-paintings and inscriptions found in the region (Dani 2001). A corpus of Buddhist manuscripts in Sanskrit was discovered in the valley of Gilgit located adjacent to the Burushaski-speaking area.

[4] Although first discovered in Old Indo-Aryan, Fortunatov's law seems to work also in modern Dardic and Nuristani languages. Cf., e.g. Dardic and Nuristani lexical items belonging to the two above-mentioned etyma: Kashmiri *phaṭun* 'to burst', Shina (Drasi dialect) *phoṭyōno* 'to split', Indus Kohistani *phaṭáṽ* 'to copulate with' (< **sphāṭyate*, Zoller 2005: 288), Kati *pṭe-*, Kamviri *pṭa-* 'flake off; break off (outer layer); explode in small bursts (as wood in a fire)'; Pashai *paṭā* 'strip of skin', Khowar *peṭek* 'scarf, *dupatta*', Kalasha *pāṭi* 'scarf', Indus Kohistani *paṭʰ* 'the piece of leather of a sling into which the stone is placed; the strap of a gun; a plaster; strip, stripe', *pʌṭĩ* 'a long strip of cloth that is wrapped around the legs as traditional trousers', *pʌṭũ* 'a type of cloth from Chitral and Gilgit' (Zoller 2005: 269), Kashmiri *paṭh* 'long strip of cloth from loom', Kamviri *pâṭü* 'turban', Prasun *puṭi, puṭī* '*Rand (eines Gewandes)*' (Buddruss, Degener 2015: 754).

[5] A certain number of Burushaski speakers, reportedly a few hundred, live in the Jammu and Kashmir union territory of India. They are descendants of migrants from the North who settled in the Vale of Kashmir a hundred-plus years ago (Munshi 2019).

[6] Hunza and Nager are two districts (formerly, princely states) situated on the right and left bank of the Hunza river respectively.

Since the easiest, if not the only practicable route connecting Hunza and Yasin runs through the Gilgit valley, there are strong reasons to consider Shina a wedge that had once split the more or less unified Burushaski-speaking area. In other words, before the arrival of Shina speakers the language of Gilgit was most probably Burushaski[7], and Shina migration must have been a crucial factor in the divergence of the Yasin and Hunza-Nager dialects. Based on glottochronological calculations, the beginning of the latter process should be dated to the 10th century A.D., and it seems fair to say that the influx of Shina loans into the eastern (Hunza-Nager) dialect of Burushaski must have started on or shortly after this date. This influx appears to result not only from close economic and political ties between the Gilgit and Hunza valleys, but, probably to a greater extent, from mass immigration of Shina speakers to Hunza and Nager[8].

The Yasin valley most likely remained untouched by this ethnic change. However, it fell under the domination of Khowar-speaking political elite originating from neighboring Chitral. This caused massive spread of Burushaski-Khowar bilingualism[9] and consequently extensive borrowing of Khowar vocabulary into the Yasin dialect. The presence of a very significant Khowar lexical stratum in Werchikwar sharply distinguishes it from Burushaski proper, influenced by Shina. Thus, the sources of modern Dardic loanwords are different for the two main varieties of Burushaski, and Aryan lexical items shared by both of them must, for the most part, have been borrowed from some source-language(s) other than Shina and Khowar. The adoption of such items must have begun before the separation of the Yasin and Hunza-Nager dialects, i.e. more than a millennium ago. Perhaps the two abovementioned phonologically archaic Aryan loans discovered by Morgenstierne belong to this early layer.

It goes without saying that the "pan-Burushaski" Aryan loan vocabulary needs thorough analysis. Such an analysis may not only throw significant light on the history of language contact in the Upper Indus region, but also yield some results useful for the comparative study of Aryan languages. In Berger's dictionary Indo-Iranian borrowings common for all Burushaski varieties are relatively easily recognizable. Though Berger's main goal was to comprehensively document lexical data on the Hunza-Nager dialect[10], Yasin cognates are quoted for all Hunza-Nager words possessing them. This principle is applied to loans too. For many of them, the hypothetical source of borrowing is also specified, referring to R.L. Turner's "Comparative dictionary of the Indo-Aryan languages" (Turner 1966)[11]. In our etymological list that follows, each of Berger's comparisons is reproduced and commented upon[12]. To select mate-

[7] This view seems to be generally accepted by scholars. See e.g. Lorimer 1935; 1937; Jettmar 1975. It is supported, inter alia, by the fact that the Gilgit dialect shows the deepest Burushaski influence in comparison with other Shina varieties. This influence is not confined to the lexicon, but manifests itself on other levels of language structure. For details see Lorimer 1937.

[8] Unlike in Gilgit, this migration did not bring about a total language shift. However, in Hunza and Nager it seems to have given rise to an ethnic group called Shin, whose members are or were the speakers of the Shina language. In Nager the share of this group in the whole population is higher than in Hunza, and it is no wonder that the Nager dialect of Burushaski shows the strongest Shina influence in the vocabulary.

[9] In the 1st half of the 20th century about one third of Yasin's population was bilingual in Burushaski and Khowar (Lorimer 1935).

[10] The dictionary comprises the last volume of Berger's three-volume description of the Hunza-Nager dialect "*Die Burushaski-Sprache von Hunza und Nager*".

[11] Berger's etymological references for tentative Aryan loanwords usually have the form "T+ entry number in Turner's dictionary", T being the abbreviation of "Turner".

[12] If no etymological hypothesis is ascribed to Berger, this means that the etymology in question belongs to the present author.

rial for the list, I used three main criteria: (i) presence of a particular Burushaski word both in the Yasin and Hunza-Nager dialect; (ii) presence of a plausible Aryan etymology; (iii) absence of inherited exact cognates in Shina and/or Khowar. The latter criterion seems to be important because the presence of cognates in both Dardic languages implies the possibility of parallel borrowing from Khowar into Werchikwar and from Shina into Burushaski proper. In a few cases, it seems reasonable to classify a loanword as belonging to the early Aryan vocabulary even though it does not meet some of the above criteria. All such cases are analyzed and discussed separately. Etymologies on the list are numbered and arranged in alphabetical order. Burushaski words are quoted in Berger's transcription[13]; for each of them the German translation from the dictionary is specified.

A list of probable early Aryan loans in Burushaski.

1. Hunza-Nager ***a-/oó-***, Yasin ***a-, ai-*** '*Negation beim Verbum*' (Berger 1998b: 9). According to Berger, borrowed from Indo-Aryan (= OIA, Av. *a-* 'not, un-, non-'). Since the Burushaski verbal negative prefix turned out to have a Sino-Caucasian etymology with cognates in North Caucasian and Sino-Tibetan (Starostin 2005), Berger's view should be regarded highly questionable. However, it seems likely that in certain words outside the verbal system, chiefly in nouns and adjectives with privative and negative meaning (cf., e.g. *abáś* '*Schwierigkeit, Unglück*' (Berger 1998b: 11), *aboóm* '*unglücklich, unselig, Unglück bringend*' (Berger 1998b: 11), *agúuẏo* '*roh, unreif, geschmacklos; wenig oder gar nicht gedüngt; wenig fruchtbar*' (Berger 1998b: 13), *akáṣ* '*Unglück, Pech, was man notgedrungen, wohl oder übel, machen muß*' (Berger 1998b: 14)), the initial *a-* may indeed reflect the Aryan negative and privative prefix. Some of such adjectives and nouns have antonyms with the initial element *su-* and the meaning "possessing something" or "possessing something of good quality". This element is probably also of Aryan origin (see **88**).

2. Hunza-Nager ***abáaśo***, feminine ***abáaśi*** '*von getrübtem Blick, schwachsichtig; in der Nacht schlecht sehend*' (Berger 1998b: 11). Berger compares this word with Shina *awáaśo* '*gelähmt, verkrüppelt*'. It seems, however, more probable that the Burushaski adjective reflects some derivative of the Indo-Iranian root **(s)paś-* 'to see, look, watch' (cf. OIA *paśyati* 'sees, looks', *spaśa-* 'a spy, a guard', Av. *spasiieiti* 'glimpses, espies') with the negative prefix **a-*, the prototype in the donor language being perhaps **apaśya-* 'not seeing, unseeing'. The cited Shina word is, due to semantic difference, unlikely to belong here.

3. Hunza-Nager, Yasin ***astám*** '*gerichtliche Entscheidung, Urteil; Rechtsfall, Prozeß*' (Berger 1998b: 22). Berger's etymology: T 1517 (< **āsthāman-* 'meeting place'). Cf. also Shina *astóom* 'judgement'. This etymology, first advanced by Morgenstierne (1945: 92), seems somewhat farfetched. The most frequent meaning of the word *astám* in Burushaski is evidently 'decision (e.g. judicial), judgement' rather than 'place of trial, courthouse'. More plausible appears to be the connection with PII **stambh-/stabh-* 'to be or make firm, support' (> OIA *stambhatē* 'supports', *stabhita-* 'fixed', Av. *staβra-*, Khotanese *staura-* 'firm', Classical Persian *sitabr* 'thick'). Cf. the semantic development of a derivative from this root in some New Indo-Aryan lan-

[13] Voiceless dental unaspirated and aspirated affricates are transcribed as *ċ* and *ċh* respectively, their palatal and cerebral counterparts as *ć, ćh* and *ç, çh* respectively; *ś* stands for the palatal sibilant, *ṣ* for the cerebral one, *ġ* for the voiced velar fricative, *ẏ* for the voiced retroflex sibilant with simultaneous dorso-palatal narrowing ("ein stimmhafter retroflexer Sibilant mit gleichzeitiger palatal-dorsaler Engebildung" (Berger 1998a: 22)).

guages: Sindhi *ṭhahraṇu* 'to stop, stand, be decided', Punjabi *ṭhahirnā* 'to stop, be settled, be proved', Hindi *ṭhaharnā* 'to stop, decide' < **stabhira-* 'fixed, firm' (Turner 1966: 789). The immediate source of the Burushaski word probably reflects a formation with the preverb *ā* (**āstambha-*).

4. Hunza-Nager, Yasin ***aśtán*** '*Pferdeknecht*' (Berger 1998b: 23). Cf. Shina *aśtóon* id. Following Morgenstierne, Berger compares the Burushaski word with OIA *aśvasthāna-* 'horse-stable' and proposes for it the prototype **aśvasthānin-*[14] (Morgenstierne 1945: 92; Berger 1998b: 23). Morgenstierne's etymology implies an unexplained loss of the suffix *-in-* in the donor language. Phonologically less problematic seems to be the derivation from **aśvatăna-* (cf. OIA *tanati*, *tānayati* 'helps, assists, aids'). The meaning of this hypothetical protoform may be reconstructed as 'horse-helper, the one who takes care of (lit. helps) horses'.

5. Hunza-Nager, Yasin ***ayáś*** '*Himmel*' (Berger 1998b: 24). Berger's etymology: T 1009 (= OIA *ākāśya-* 'of the sky'). Equally plausible seems to be Turner's comparison with OIA *ākāśa-* 'sky' (T 1008). The Old Indian word has a probable cognate in Shina (*hagaái*) where the development of the earlier intervocalic *-k-* is remarkably different from that in the source-language of the Burushaski lexeme.

6. Hunza ***báago*** '*Anteil; Ersatz*', Nager ***báago*** '*anstelle von; von seiten*' (Berger 1998b: 25), Yasin ***bágo*** 'share' (Lorimer 1938: 398). Berger's etymology: T 9434 (= OIA *bhāgya-* 'fortune, lot'). The Aryan prototype **bhāga-* (> OIA *bhāga-*, Av. *bāga-* 'share, portion, fraction', Old Persian **bāga-* 'allotment, plot', Classical Persian *bāγ* 'garden') also cannot be ruled out. Reflexes of PII **bhāgia-* are represented in some Dardic languages (cf. Shina *báago, bagóo*, Kashmiri *bōg* 'share') but not in Khowar, and Yasin *bágo* can, therefore, hardly be a Khowar loan.

7. Hunza-Nager ***badá***, Yasin ***bayá*** '*Schritt, Sohle (von Stiefel, Fuß)*' (Berger 1998b: 29). Morgenstierne compared the Burushaski word with OIA *pada-* 'foot, footprint, step' (Morgenstierne 1935). Initial *b* must be the result of sonorization of an initial tenuis[15].

8. Hunza-Nager ***balá*** '*Vorteil, nützlich*' (Berger 1998b: 32). Berger's etymology: T 9408 (= OIA *bhalla-* 'auspicious', Sindhi *bhalo*, Punjabi, Hindi *bhalā* 'good'). The etymology seems plausible. Similar semantic development took place in Iranian: Classical Persian *bahr* 'profit, benefit' < Proto-Iranian **badra-* 'lucky, happy'[16] (Rastorgueva & Edelman 2003: 44). The word has no attested cognates in Shina and Khowar, which suggests that it may belong to the early Aryan loan vocabulary.

9. Hunza-Nager ***balándo*** '*stark, mächtig*' (Berger 1998b: 33). Berger's etymology: T 9166 (= OIA *balavant-* 'strong'). The source of this word in the donor language probably reflects the secondary thematized base **balaṷanta-*. Note the voicing of the historical voiceless stop after the nasal. Since no cognates of the Burushaski lexeme are attested in Shina and Khowar, it may be tentatively classified as an early Aryan loanword.

10. Hunza, Yasin ***balán*** '*Waagebalken; Querbalken oder Leine zum Wascheaufhängen*', Nager ***balón*** '*Leiste, die den Trichter zur Mühle stützt; Leiste um den Webstuhl; Geländer*' (Berger 1998b: 33).

[14] The supposed meaning is perhaps 'the one in the horse-stable'.

[15] On voicing of initial voiceless consonants in Burushaski see Lorimer 1935 and Morgenstierne 1945.

[16] Etymologically related to OIA *bhadra-* 'lucky, happy', *bhalla-* 'auspicious' (< **bhadla-*).

Berger's etymology: T 11879 (= OIA *vilagna-* 'hanging to'). Cf. also Shina *balóṅ*, *bʌlʊṅ* 'stick to hang clothes on, perch (for a bird)'.

11. Hunza-Nager, Yasin ***balk*** '*Brett; Bord (auf dem man den Brotteig bereitlegt); (bei der Sitar) Fläche über dem Bauch*' (Berger 1998b: 34–35). The word may be compared with OIA *phalaka-* 'board, lath, shield' (< **sphalaka-*), Proto-Iranian **spara-* 'shield' (> Middle Persian *spar*, Classical Persian *sipar* id.), Kashmiri *phal* 'strip of wood'. Burushaski *b* < **ph* due to initial sonorization.

12. Hunza-Nager ***bandeéś -t-*** '*fesseln, binden, festhalten*', Yasin ***bandéś et-*** '*befehlen*' (Berger 1998b: 36). Contamination of two different roots seems to have taken place. With the meaning 'to bind, to tie' the Burushaski word can hardly be separated from OIA *bandhati*, Av. *baṇdaiieiti* 'binds, ties'. With the meaning 'to order, command', however, it may reflect or be semantically influenced by some cognate of Kalasha *bandek* 'to teach; to announce so as to inaugurate, to order or command' (Trail & Cooper 1999: 27). The Kalasha verb is, most probably, etymologically connected with OIA *vandatē* 'praises, worships', Av. *vaṇd-*, Khotanese *van-*, Parthian *wynd-* 'to praise, honor, worship' (Kogan 2019). Shina *bandeéś* '*Befehl*' (Berger 1998b: 36) may be a Burushaski loanword.

13. Hunza-Nager, Yasin ***bas-*** '*(Schnee, Tau) fallen, sich setzen, sich niederlassen*' (Berger 1998b: 41). Cf. OIA *vasati* 'dwells, stays', Av. *vaŋhaiti* 'dwells', Hindi *basnā* 'to dwell', Nepali *basnu* 'to remain, inhabit, sit', Gujarati *vasvũ*, Marathi *vasṇẽ* 'to settle' Kashmiri *wasun* 'to descend, to land from a boat'. A Sino-Caucasian etymology has been proposed for this root. Along with *ōs-* 'to put down, place' it is derived from Proto-Sino-Caucasian **=ăsÁ* 'to be, stay' (Starostin 2005). However, the exact relationship between the two Burushaski reflexes of this prototype is unclear. It seems quite probable, or at least not excluded, that the real continuant of the cited Sino-Caucasian protoform is *ōs-*, while *bas-* is an Aryan loan[17].

14. Hunza-Nager, Yasin ***baṣá*** '*Turban*' (Berger 1998b: 43). Berger compares the word with OIA *praśna-* '*Korb, Turban*', Shina *paṣóo* 'turban'. The meaning 'turban' seems to be extremely rare for the Old Indian lexeme. In addition, the phonological development in Shina is irregular, which makes Berger's comparison very doubtful. Cf., however, OIA *pispṛśati* 'will cause to be adorned', *sparśayatē* 'adorns?' (Turner 1966: 798), the loss of initial *s* and the change **rś* > *ṣ* being regular for Shina. Since the word is "pan-Burushaski", borrowing from Shina is rather unlikely.

15. Hunza-Nager ***baṣçár***, Yasin ***baçár*** '*Guckloch, Schießscharte*' (Berger 1998b: 43). Berger's etymology: T 11647 (= OIA *vikṣāra-* 'a lucky hit on the target'). Cf. also Khowar *baçhár* 'loophole'. The "pan-Burushaski" nature of the word makes borrowing from Khowar improbable.

16. Hunza-Nager, Yasin ***bat*** '*flacher Stein, Steinplatte, Schieferplatte*' (Berger 1998b: 43). Berger's etymology: T 11348 (< **varta-* 'round stone'). Cf. Ashkun, Waigali *wāṭ*, Kati *woṭ*, Tirahi *baṭ* 'stone', Gawar-Bati *wāṭ* 'stone, millstone'; Kalasha *bat*, Kho. *bort*, Bshk. *baṭ*, Tor. *bāṭ*, Maiyã *bhāt*, Palula *băṭ*; Shina *băṭ* 'stone', Kashmiri *waṭh* 'round stone', Romani *barr*, Lahnda, Punjabi *vaṭṭā* 'stone', Khotanese *ūḍāra-* 'crystal', Wakhi *wərt* 'marble, millstone, stone', Kurdish *bar(d)* 'stone').

[17] It should be noted that the initial *b-* of *bas-* remains unexplained if we accept the above Sino-Caucasian etymology of this root.

17. Hunza ***béric̣***, Nager ***béḍic̣***, Yasin ***béḍec̣*** '*Angehöriger der Schmiede- und Musikerkaste in der Gilgit Agency, Zigeuner, "Dom", Schmied, Musiker*' (Berger 1998b: 48). The Burushaski word evidently has cognates in Northwestern Tibetan dialects: Balti *bekar* 'court singer and dancer who improvises poems and songs' (Sprigg 2002: 28), Ladakhi *beda* 'member of the caste that used to be itinerant musicians', *bemo* f. (Norman 2010: 640–641). All the cited forms seem to be derivatives of the common root *be-*. This root is likely to be borrowed from Indo-Iranian and may reflect PII **wād(i)ya-* 'music, musical instrument' (cf. OIA *vādya-* 'musical instrument, music', *vādyakara-* 'musician', *vādayati* 'plays a musical instrument', Ossetic *wadynʒ* 'flute, panpipe'). For details see Kogan 2019.

18. Hunza-Nager ***biík***, Yasin ***behék*** '*Weide (Baum)*' (Berger 1998b: 51). Berger's etymology: T 12097 (= OIA *vēta-* 'cane, reed'). Cf. also Pashai *wēu*, Shina *bĕu*, Classical Persian *bēd* 'willow' < PII **u̯aita-*, **u̯aiti-*. The Burushaski word seems to reflect the prototype **u̯aitaka-* or **u̯aitikā-*.

19. Hunza-Nager ***bijaáṭo, bijiíṭo***, Yasin ***bijíju*** '*fürchterlich, schrecklich aussehend*' (Berger 1998b: 52). Berger's etymology: T 9515 (= OIA *bhīyatē* 'fears'). Cf. also Shina *bižoĭki̯* 'to fear', *bižātŭ* (Bailey 1924: 133), *bijaáṭo* (Berger 1998b: 52) 'timid'. In Werchikwar progressive assimilation seems to have taken place (**bijíṭu* > *bijíju*). The word is probably "pan-Burushaski", so Shina could hardly have been its immediate source.

20. Hunza-Nager, Yasin ***ċápi*** '*Pinzette, Zange*' (Berger 1998b: 69). Cf. Khowar *ċápi* id. The latter word may well be a borrowing because initial dental affricates are rare in the inherited vocabulary of Khowar. Probably, the Burushaski lexeme reflects some derivative of PII **kap-/čap-* 'to catch, snatch, pick, pinch' (> Shina *čap-*, Indus Kohistani *cap-*, Gawar-Bati *cep-* 'to bite', Pashai *čip-* 'to bite off', Kashmiri *cop* 'a bite', Munji *cəb-* 'to pinch', Wakhi *čəp-* 'to pick, pluck')[18]. For semantic development cf. Russian *kusachki* (*кусачки*) 'nippers, tongs, cutting pliers' < *kusat'* (*кусать*) 'to bite'.

21. Hunza-Nager, Yasin ***ċar*** '*Wachtposten, Wächter, pl. Wachleute, Wache*' (Berger 1998b: 69). Berger's etymology: T 4683 (= OIA *cara-* 'spy'). Cf. also Shina *car* 'watchman, guard', Kashmiri *car* 'spy'. Since earlier palatal affricates are not dentalized in Shina, the Shina word is no doubt borrowed.

22. Hunza-Nager ***ċar -t-*** '*zerreißen, zerspalten, aufschneiden, herunterreißen*', ***ċir man-***[19] '*(Kleider) in Stücke gehen, zerreißen*' (Berger 1998b: 69). Berger's etymology: T 4844 (< **cīrayati* 'splits'). Cf. also OIA *cīra-* 'strip (of bark or cloth)', Hindi, Punjabi *cīrnā* 'to tear, split', Shina *ciroiki* 'to split (wood) lengthwise'. The Shina word must be borrowed because of the irregular initial dental affricate (see **21**).

23. Hunza-Nager ***ċhárịṣ***, Yasin ***ċáreṣ*** '*Teil des Türrahmens, d.i. Türpfosten oder Schwelle; pl. Türrahmen, Rahmen (des Webstuhls, der Drechselmaschine)*' (Berger 1998b: 75). Berger's etymology: T 1880 (< **utsāra-* 'going out'). Berger also compares Khowar *carēni* 'doorframe'. Because of

[18] For more Iranian reflexes of this root see (Rastorgueva & Edelman 2003: 221–226). If this etymology is correct, the Khowar word must definitely be considered a loan because in Khowar the old palatal affricate **č* is preserved unchanged, and no dentalization takes place.

[19] Hunza-Nager *-t-*, Yasin *et-* 'to do' and Hunza-Nager, Yasin *man-* 'to become' very often function as light verbs in Burushaski complex predicates.

phonological differences the Khowar word can hardly be the immediate source of the Burushaski forms.

24. Hunza-Nager ***ćar -t-*** '*(Flüssigkeit, Pulver) gießen*' (Berger 1998b: 85). Cf. OIA *kṣarati*, *kṣalati* 'flows, trickles', *kṣārayati* 'causes to flow', *kṣālayati* 'washes', Proto-Iranian **xšar-* 'to flow' (> Persian *šārīdan* 'to trickle', *(āb)šār* 'waterfall', Ossetic *äxsärdzän* 'waterfall' (< **xšar-čana-*), Middle Persian *Xšart* 'the river Jaxartes' (Livshits 2003), Kashmiri *čhalun* 'to wash', *čhar* 'a sprinkle of water etc. from the fingers'), Gawar-Bati *ṭhār* 'rapids in a stream', Palula, Indus Kohistani *ṭhār*, Bashkarik, Kalasha *uṭhār*, Shina *ṭhar* 'waterfall', Khowar *čəreik* (*ṭh*?) 'to pour, shed, drop, sprinkle, scatter (seed), make fall'. The Khowar word can hardly be the source of the Burushaski one because the latter is attested in the Hunza-Nager dialect.

25. Hunza-Nager ***ćarí, ćirí*** '*Heimchen, Grille*' (Berger 1998b: 86, 89). Berger's etymology: T 4845 (= OIA, Pali *cīrī* 'cricket'). Cf. also Khowar *čari* 'bedbug'. The word was also borrowed into Northwestern Tibetan dialects: Balti *ceri*, Ladakhi *cari* 'bedbug' (Sprigg 2002: 41; Norman 2010: 266). Since it is attested in the Hunza-Nager dialect, Khowar can hardly be its immediate source.

26. Hunza-Nager, Yasin ***ćódo*** '*Verhöhnung, Stichelei, Beschimpfung*' (Berger 1998b: 90). Berger's etymology: T 4926-29 (= OIA *cōda-* 'goad', *cōdayati* 'impels; importunes, asks'). Cf. also Pali *cōdaka-* 'one who rebukes', Classical Persian *čust* 'agile'. Shina *ćódo* 'sarcasm, scoffing' should be considered a loanword due to the irregular retention of intervocalic *d*.

27. Hunza-Nager ***ćhaġéni*** '*schwarze Ziege*' (Berger 1998b: 95). Berger's etymology: T 4961 (< **chagana-* 'goat'). Cf. OIA *chaga-*, *chagala-* 'goat', Khowar *čhani*, Pashai *čhanīk* id.

28. Hunza-Nager ***ćhe***, Yasin ***ćéi*** '*Einschnitt (mit der Axt)*' (Berger 1998b: 99). Berger's etymology: T 5064 (= OIA *chēda-* 'section, piece; incision'). Cf. also Kalasha *čhēla* 'piece', Shina *čhɛ* 'cross cut with an axe', Sindhi *chehu* 'incision, end', Hindi *cheu* 'cut, stroke, mark, chip'.

29. Hunza-Nager ***ćhiṣ***, Yasin ***ćiṣç, ćirç*** '*Leiter, schraggelegter Stamm mit eingehauenen Fußkerben; Bahre*' (Berger 1998b: 100). Berger's etymology: T 12704 (< **śritrā-* 'ladder'). Cf. also OIA *śriti-* 'entrance', Ashkun *istrî* 'ladder made of a single log', Waigali *ṣī*, *ṣiŕ*, Kati *ṣīk*, Prasun *čīk*, Kalasha *šīṭ*, Shina *ṣhiç*, Punjabi, Hindi *sīṛhī*, Marathi *śiḍ(h)ī*, Pashto *ṣəl*, Yazghulami *x̌ad* 'ladder'. For the Burushaski forms the prototypes **śriṣṭra-* and **śriṣṭa-* cannot be ruled out.

30. Hunza-Nager ***ṭháci***, ***ćháṭi*** '*Rang, Würde*' (Berger 1998b: 106). Berger's etymology: T 3649 (= OIA *kṣatriya-* 'ruling; one of the ruling order'). If the etymology is correct, the variant *ćháṭi* must have developed due to affricate dissimilation. No Shina or Khowar cognates are attested which gives reason to believe that the Burushaski word may be an early Aryan loan.

31. Hunza ***ṭ(h)ar***, Yasin ***ṭar, ṭir*** '*(Kindersprache und vulgar) Stuhlgang, Pissen, Durchfall*', Nager ***ṭhar*** '*Wasserfall (größerer)*' (Berger 1998b: 106). Berger's etymology: T 3662. Etymologically identical to **24**.

32. Hunza-Nager ***ṭhur*** '*Messer (größeres, z.B. zum Schachten)*', Yasin ***ṭur*** '*Keil, der an der Steinschleuder die zwei Sehnen auseinanderspannt*' (Berger 1998b: 107). Berger's etymology: T 3727 (= OIA *kṣura-* 'razor', *kṣurī-* 'knife, dagger'). Cf. also Kalasha *ṭhūŕi* 'knife, Bashkarik *ṭhur*, Torwali *ṭhū*, Palula *ṭhūr* 'dagger, knife', Shina *ṭūr* 'small knife', Lahnda, Punjabi, Hindi *churī*, Ne-

pali *churi*, Romani *čhuri* 'knife', Awadhi *chūrā*, Nepali *churo* 'razor' < PII **kšura-* < PII **ksuro-* (Mayrhofer 1992: 435–436).

33. Hunza-Nager, Yasin ***damán*** '*Herr, Besitzer, Gott; adj. wirklich, eigen (bei Verwandten, im Gegensatz zu Stief-)*' (Berger 1998b: 112). Berger's etymology: T 6180 (= OIA *damūnas-* 'householder, master'). Cf. also Shina *dabóon*, *dabūn* 'owner, master, landlord, God'. The change *ū* > *a* in the stressed syllable is not noted anywhere else in the Indo-Iranian loan vocabulary of Burushaski. Therefore, Berger's comparison seems questionable. Much more plausible is the etymology advanced by Morgenstierne who compared the Burushaski lexeme with OIA *damana-* 'taming, subduing' (Morgenstierne 1945: 92). It is also possible that the source of the word in question reflected the prototype **damāna-*, i.e. the old medial participle of the root **dam-* 'to tame, subdue'[20]. Such a hypothesis can satisfactorily explain the long vowel in the Shina cognate.

34. Hunza-Nager, Yasin ***dar*** '*Angst, Zittern*' (Berger 1998b: 114). Berger proposes two alternative etymologies: T 6186 (= OIA *dara-* 'fear') or < Balti *darba*, Tib. *ʼdar ba* 'to tremble, shiver'. Borrowing from Tibetan appears much less probable for linguogeographical reasons: Werchikwar where the word in question is attested, being geographically distant from the Tibetan-speaking area, is not and seems to have never been in close contact with Tibetan dialects.

35. Hunza, Nager, Yasin ***darí*** '*Fenster*' (Berger 1998b: 115). Berger's etymology: T 6651 (< **dvara-* 'door'). Cf. OIA *dvār-* 'door, gate', Av. *duuar-* 'gate', Pashai *darī*, Gawar-Bati *derī*, Kalasha *durík*, Palula *darúṛi*, Torwali *deræn*, Indus Kohistani *dʌrī*, Shina *darií*, Kashmiri *dār*, Punjabi, Gujarati *bārī* 'window'.

36. Hunza, Yasin ***daróġo***, Nager ***daráġuẏ*** '*Stock, Stecken, Stab*' (Berger 1998b: 115). Berger's etymology: T 6298 (= OIA *dāru-* 'piece of wood'). Cf. also Old Persian *dāruv*, Av. *dāuru-* 'piece of wood', Pashai *dār* 'firewood', Shumashti, Gawar-Bati *dār* 'wood, tree', Khowar *dar* 'timber, firewood, gallows', Kashmiri *dār* 'log', Kumauni, Nepali, Hindi *dār* 'timber', Persian *dār* 'gallows', Shughni *δōrg* 'wood, log'. The intervocalic *ġ* of the Burushaski words may reflect the old **-k-*. If so, the Aryan prototype should be reconstructed as **dāruka-*.

37. Hunza-Nager, Yasin ***daṣmán*** '*islamischer Priester*' (Berger 1998b: 116). According to Berger, either T 6117 (= OIA *dakṣamāna-* 'being able') or < Urdu *dānišmand*. Cf. also Khowar *daṣmán* 'priest'. Since the Urdu word, as well as its immediate source in Classical Persian (*dānišmand*), does not mean 'priest', the usual meaning being 'a learned, wise, sagacious person', the former of Berger's comparisons seems more plausible. The presence of the lexeme under study in the Hunza-Nager dialect suggests that its Khowar cognate may have been borrowed from Burushaski but hardly vice versa.

38. Hunza ***datáġar***, Yasin ***datáġar, tatáġar*** '*Fieber*', Nager ***datáġar*** '*Malaria*' (Berger 1998b: 116). Berger's etymology: T 5679, 5680 (= OIA *tapta-* 'heated, hot', **taptakārī* 'act of making hot'). Cf. also Av. *hąm.tapta-* 'heated, hot', Tirahi *tata*, Bashkarik *tʌtt*, Palula *tāto*, Shina *tātụ*, Kashmiri *tot*, Sindhi *tato*, Lahnda, Punjabi *tattā*, Kumauni, Nepali *tāto*, Romani *tato*, Khotanese *tauda-*, Persian *taft*, Pashto *tōd* 'hot, warm', Ossetic *tævd* 'hot; heat'. The initial *d* in Burushaski must have developed due to the sonorization of earlier *t*.

[20] With the same root cf. also OIA *damayati* 'subdues', Ossetic *domɨn* 'to tame' < PIE **demh₂-* (LIV: 116).

39. Hunza-Nager ***diṣçík***, Yasin ***giṣṭék, giṣṭík, kištíki, diṣṭík*** '*sonnengetrockneter Ziegel*' (Berger 1998b: 121). Berger hypothetically compares the Burushaski words with OIA *iṣṭakā* 'brick'. Although this comparison suggests itself, it leaves unexplained the irregular initial consonants of both the Hunza-Nager and Yasin forms. As a provisional hypothesis, it may be assumed that the forms starting with dorsals developed due to the influence of Persian *χišt* 'brick', *k* conveying the Persian uvular fricative, and *g* resulting from the sonorization of earlier *k*, whereas the initial *d* of the other two forms arose by analogy with some derivative of the PII **daiǰh-* 'to plaster, anoint; shape, build' (> OIA *dēhī-* 'dam, surrounding wall', Old Persian *didā-* 'rampart, fortress').

40. Hunza-Nager ***dumáṣ*** '*Wolke (Staub, Rauch, Wasser)*' (Berger 1998b: 125). As a tentative etymology, Berger suggests T 6849 (= OIA *dhūma-* 'smoke') and also quotes Shina *dumáṣ* '*Rauch*'. The word, although it evidently contains an Aryan root, seems to have a Burushaski derivational suffix and, therefore, to have been borrowed from Burushaski into Shina, not vice versa. It is remarkable in this connection that the common word for 'smoke' in Shina is *dŭm* and not *dumáṣ*.

41. Hunza-Nager ***gaċhír***, Yasin ***gaċér*** '*Geier*' (Berger 1998b: 141). Berger's etymology: T 4231, 4233 (= OIA *gr̥dhya-* 'greedy', *gr̥dhra-* 'desiring greedily; vulture'). Cf. also Pali *gijjha-* 'greedy; vulture', Waigali, Dameli *grīč*, Gawar-Bati *līnči*, *līnċi*, Palula *grhīnǰo*, Shina *gireč*, Sindhi *gijha*, Lahnda *girijjh* 'kite', Kashmiri *grad*, Punjabi *giddh*, Nepali *giddha*, Hindi *gīdh*, Gujarati *gīdh* 'vulture', Marathi *gīd*(*h*), *gidhāḍ* 'vulture, kite', Yidgha *-γereno* 'eagle' (< **gr̥dnu-*). The etymology looks plausible, though some historical-phonological irregularities need explanation. The source form of the Burushaski word in the donor language must have undergone devoicing of the final voiced affricate with subsequent metathesis (PII **gr̥dhya-* > (**gariǰ*, **gareǰ* >) **gariʒ*, **gareʒ* > **gariċ*, **gareċ* > *gaċhir*, *gaċer*).

42. Hunza-Nager, Yasin ***gamún*** '*Unterteil; (am Baum) Wurzel, Stumpf*' (Berger 1998b: 145). Probably < PII **gambhūna-*, cf. OIA *gambhan-* 'depth', Shina *gabŭn* 'foot of mountain, bottom, tree-trunk' (Morgenstierne 1945: 92).

43. Hunza-Nager, Yasin ***gaṭ*** '*Knoten (auch im Stengel von Pflanzen), Knorren (im Holz); Knöchel (des Fingers); Zyklus, Ablauf*' (Berger 1998b: 150). Berger's etymology: T 4353 (= OIA *granthayati* 'ties'). Cf. also OIA *granthi-*, Hindi, Bengali, Marathi, Gujarati *gāṭh*, Punjabi, Lahnda *gaṇḍh*, Sindhi *g'aṇḍhi* 'knot', Proto-Iranian **graϑa-* > Middle Persian *grih*, Persian *girih*, Khotanese *grratha*, Sogdian *γr'nš*, Ishkashimi *γureẍ*, Ossetic *ælxync'* 'knot'. The immediate source of the Burushaski word most probably reflects the prototype **garthi-* or **gartha-*. The continuants of the PII root **garth-*[21] are found in some Dardic languages and in Pashto: Indus Kohistani *gār*, Brokskat *gaṭhi* 'knot', Pashai *gaṭanā* 'joint', *gaṭh-* 'to tie', Pashto *γaṛəl* 'to twist, spin, plait'. As a loanword, the lexeme in question is also attested in Northwestern Tibetan dialects: Balti, Purik *gaṭ* 'knot, joint of body', Ladakhi (Leh dialect) *chang-gat* 'knee-joint', *gat* 'obstacle' (Kogan 2019).

44. Hunza-Nager ***gukúrus, gukúro***, Yasin ***kukúres*** '*Hündchen, junger Hund; Junges des Schneehuhns*' (Berger 1998b: 159). Berger's etymology: T 3329 (= OIA *kurkura-, kukkura-* 'dog'). Cf. Tirahi *kukri*, Shina *khŭkŭr* (Gilgiti dialect), *kŭkúr* (Gurezi dialect) 'puppy', Kumauni, Nepali, Bengali *kukur* 'dog', Sinhalese *kukurā* 'dog, puppy'. Nuristani, Dardic and Iranian forms with intervocalic *-č-* or its regular reflexes may also belong here: Waigali *kuċúru* (*ċ* < **č*), Bashkarik

[21] This root is, no doubt, a variant of PII **grath-*/**granth-* 'to tie, bind'.

kúčur, Woṭapuri *kuċúr*, Torwali *kuǰű* (*ǰ* < **č*), Indus Kohistani *kuċír* (*ċ* < **č*), Palula *kučuro* 'dog', Kashmiri *kuċur* (*ċ* < **č*) 'puppy', Kurdish *kūčik*, Balochi *kučak(k)*, Parachi, Ormuri *kučuk* 'dog'. The word is sometimes considered onomatopoeic (Mayrhofer 1992: 372).

45. Hunza-Nager, Yasin ***ġaténç*** '*Schwert, zweischneidiger Säbel*' (Berger 1998b: 174). Berger's etymology: T 2858, 2860, 2862 (= OIA *kartari-* 'scissors, knife', *kaṭṭāraka-* 'knife' (< **karttāraka-*), *karttr̥kā-* 'knife, dagger'). A cognate of this word seems to have been borrowed into Balti (cf. Balti *rkat* 'to cut down with a sword')[22]. The details of the sonorization process in Burushaski are not quite clear: one would expect the initial voiced stop *g* rather than fricative *ġ*.

46. Hunza-Nager ***ġuṭúl***, Yasin ***xuṭúl*** '*großer, runder, oben sich verengender Korbbehalter zum Aufbewahren von Getreide*' (Berger 1998b: 183). Berger's etymology: T 3546 (= OIA *kōṣṭha-* 'granary, storeroom'). The Burushaski forms most probably reflect **khuṭul* < **kuṭhul* (cf. Shina *kotul, kuṭúul* 'wattle and mud erection for storing grain'). The cluster *ṣṭh* in OIA *kōṣṭha-* has developed either from PIE **-sth-* (Pokorny 1959: 953) or PIE **-k'th-* (Mayrhofer 1992: 404–405).

47. Hunza-Nager, Yasin ***jaṭ*** '*alt, greis*' (Berger 1998b: 225). Cf. OIA *jarati* 'makes old, grows old', *jīrṇa-* 'old', Av. *zar-* 'to age, grow old', *zarəta-* 'old', Shina *ǰărŭ* 'old', *ǰarižóĭkį* 'to become old', Pashto *zōṛ* 'old' (< **źarta-*). The prototype of the Burushaski word probably reflects the old perfect participle in *-ta-* (**ǰarta-* < **ǰarita-*?), also preserved in Iranian.

48. Hunza-Nager ***jií***, Yasin ***ji*** '*Leben, Seele, Selbst; Geliebter, Liebling; lieb (wie das eigene Leben)*' (Berger 1998b: 226). Berger's etymology: T 5239 (= OIA *jīva-* 'living, living being, vital breath, life'). Cf. also Old Persian *ǰīva-*, Av. *ǰuua-* 'living, alive', Kashmiri *zuw* 'soul, living creature', Punjabi *jīu*, Hindi *jī*, Kumauni *jyū* 'life, soul', Nepali *jiu* 'body, life', Bengali *jīu*, *ji* 'life, spirit, courage'.

49. Hunza ***juú, juúna***, Nager ***juúa, juú-náa***, Yasin ***jū, jûna*** '*etwa: Herr (höfliche Anrede gegenüber jedermann)*' (Berger 1998b: 231). Berger's etymology: T 5240 (= OIA *jīva* 'long live!') or < West Tibetan *žu, ǰu*. Since the word is attested in Werchikwar, borrowing from Tibetan seems improbable. Cf. also Shina *ǰū* 'salutation, Sir', Lahnda *jīu*, Hindi *jī*, Nepali *jiu* 'particle of assent or respect', Kumauni *jyū* 'term of respectful address'. Etymologically related to **48**.

50. Hunza-Nager ***ǰóo***, Yasin ***ǰo*** '*plötzlicher Angriff, Verletzung*' (Berger 1998b: 235). Berger's etymology: T 6640 (= OIA *drōgha-* 'injurious', *drōha-* 'injury, treachery'). Cf. also Old Persian *drauga-*, Av. *draoga-* 'lie, deceit', Shina *ǰō* 'injury; deception, treachery', Sindhi *ḍrohu* 'deceit', Classical Persian *durōγ*, Kurdish *dirō* 'lie'. For the change **dr* > *ǰ* see also **104**.

51. Hunza-Nager ***káṭar***, ***kóṭar***, Yasin ***káṭar*** '*Schnitte (Brot), Stück, Hälfte*' (Berger 1998b: 243). Berger's etymology: T 3241 (= OIA *kuṭṭayati* 'crushes, grinds, pounds'). This comparison is doubtful due to historical-phonological irregularities (the final *r* in the Burushaski forms remains unexplained), as well as semantic differences. More probable seems to be the connection with Sindhi *katiraṇu*, Punjabi, Hindi *katarnā*, Nepali *katranu* 'to clip, shear', Gujarati *kātarvũ*, Marathi *kātarṇẽ* 'to clip', Kashmiri *katarun* 'to cut into slices' (< Indo-Aryan), Hindi *katrā* 'piece', Gujarati *kātrī* 'thin slice'. The cited Indo-Aryan words are believed to reflect some derivatives of OIA *kartari-* 'scissors, knife' (Turner 1966: 145). The source of the Burushaski lexemes may also belong to the same etymon.

[22] For more Aryan etymological parallels to the Balti word see Kogan 2019.

52. Hunza ***kiṣ*** '*ererbtes Land, Landbesitz*' (Berger 1998b: 245). According to Berger, Yasin ***qíçi*** '*der die Hälfte des Erbes allein bekommt*' may be connected with the Hunza word (Berger 1998b: 244). Berger's etymology: T 3448 (= OIA *kr̥ṣi-* 'ploughing'). Cf. also OIA *kr̥ṣati* 'ploughs', *kr̥ṣāṇa-* 'ploughing', Av. *a-karšta-* 'unploughed, uncultivated', Dameli, Khowar *kiṣ-*, Pashai *kiš-* 'to plough', Kati, Dameli, Kalasha *kiṣ*, Waigali *kīṣ* 'ploughing', Pashai *kiṣ* 'cultivated field, crops', Khowar *kiš* 'plough, cultivation', Shina *kiṣī* 'line', Punjabi *kisāṇ*, Hindi, Nepali *kisān* 'farmer, peasant', Sogdian *kštyč* (Buddhist), *kyštyč* (Manichaean) 'cultivated field', Classical Persian *kištan* 'to till, sow'.

53. Hunza-Nager, Yasin ***khan*** '*Fort, Festung, befestigtes Dorf*' (Berger 1998b: 251). Berger's etymology: T 2720 (= OIA *kanthā-* 'wall, town' < Iranian). Phonologically closer, however, seem to be derivatives of PII **k(h)an-* 'to dig' (> OIA *khan-* 'to dig', Proto-Iranian **kan-/*χan-* 'to dig, throw, pile'[23]) with an initial aspirate. Cf., e.g. OIA *khanati* 'digs', *khānya-* 'anything being dug out'. The original meaning in the donor language was perhaps 'surrounded with a moat or earth rampart'. Interestingly, a similar semantic development is attested in Iranian where a number of words for 'town, (fortified) settlement' are formed from the same root (Edelman 2011: 347)[24].

54. Hunza-Nager, Yasin ***khaní*** '*geröstete Weizen- oder Gerstenkörner beliebtes Schnellgericht bei Jägern usw.*' (Berger 1998b: 251). Berger's etymology: T 2661, 2665 (= OIA *kaṇa-* 'a grain of corn', *kaṇika-* 'grain, a meal of parched wheat'). Cf. also Shina *kaná* 'grain of rice with the husk on', Kashmiri *kanɨ* 'granule'. The aspiration of the initial consonant in Burushaski remains unexplained.

55. Hunza-Nager ***khaṅór*** '*Körper von Tieren und Menschen, aus denen die Innereien herausgenommen sind*', Yasin ***khóṅgor*** '*Gerippe*' (Berger 1998b: 251). Berger's etymology: T 2603 (= OIA *kaṅkāla-* 'skeleton'). The initial aspirate, as well as the difference of accent between the Hunza-Nager and Werchikwar forms needs explanation.

56. Hunza-Nager ***laṅ man-*** '*sich bewegen, schütteln, zur Seite springen; aufhören; sich verspäten*', Yasin ***laṅ -mán-, lâm -man-*** '*(Regen) aufhören*' (Berger 1998b: 262). Berger's etymology: T 10951 (= OIA *lamba-* 'pendant'). The comparison is doubtful both phonologically and semantically. Cf., however, OIA *laṅghayati* 'leaps over; transgresses; ascends; prevents', *laṅghya-* 'passable, to be traversed or passed over; to be neglected, negligible', Pashai *laṅg-* 'to descend', Khowar *laṅgik* 'to move, shake, sway', Torwali *laṅg-* 'to pass, cross', Indus Kohistani *lʌ́ŋg karʌ́ṽ* 'to move or push someone or something to the side', Palula *laṅg-* 'to pass, cross, swim, bathe', Shina *lʌ́ŋoiki* 'to pass over, step over; to dismiss, see off', Sindhi *laṅghaṇu* 'to pass over', Lahnda *lagghaṇ*, Punjabi *laṅghṇā* 'to pass'. The Werchikwar form *lâm -man-* may have evolved from *laṅ -mán-* due to assimilation of nasals.

57. Hunza-Nager ***leéṣ***, Yasin ***leṣ*** '*blaugrüner Fasan (Monal pheasant), kommt in Chilas, Astor vor*' (Berger 1998b: 265). Cf. Kalasha *lohiṣṭ* 'male of Himalayan pheasant', Palula *lōwīṣṭ*, Bashkarik *lōîṭ* 'golden oriole', Torwali *lawēṭ* 'male golden oriole', Indus Kohistani *lhéṭʰ* 'a gold pheasant with shining green feathers and a 'crown" (Zoller 2005: 372), Shina *lēṣ* 'peahen'. Dardic words

[23] For numerous Iranian derivatives of this root see Edelman 2011: 199–214, 340–348.

[24] Cf. Proto-Iranian **kanta-*, **kanθa-* > Khotanese *kanthā-*, Sogdian *knδ*, Yaghnobi *kant* 'city, town'. The Iranian source of the above-cited Sanskrit word reflected the same prototype.

are usually compared with the hypothetical Old Indian form **lōhiṣṭha-* 'very red' (Turner 1966: 650). If this etymology is true, the common Dardic prototype should be reconstructed as **laudiṣṭa-*.

58. Hunza-Nager, Yasin ***mágun*** '*Perle, Halskette; Spule, Haspel*' (Berger 1998b: 274). Berger's etymology: T 9731 + 4190 (< **maṇiguṇa-*, cf. OIA *maṇi-* 'jewel, ornament', *guṇa-* 'thread, rope'). The first stem component of the above reconstructed compound may, however, with equal right be compared with OIA *manyā* 'nape of neck', Av. *manaoϑrī-* 'neck, nape', Pashai *mandă*, Gawar-Bati *manda*, Bashkarik *maṇ* 'neck', Shina *māṇi* 'Adam's apple', Sindhi *maṇi* 'vertebrae of neck', Romani *men* 'nape of neck, neck, throat'.

59. Hunza-Nager, Yasin ***man*** '*Plattform aus Erde in Häusern (zum Schlafen und Sitzen), steinern in Gärten zum Sitzen, vor dem Poloplatz für die Honoratioren*' (Berger 1998b: 278). Berger hypothetically compares the word to OIA *mañca-* 'stage, platform, bed' (T 9715). No other examples of the change *ñč* > *n* have been attested. See Starostin 2005 for a possible Sino-Caucasian etymology.

60. Hunza-Nager, Yasin ***mart*** '*steile, abgeböschte Erdwand am Flußufer*' (Berger 1998b: 282). Cf. OIA *mṛttikā-* 'earth, clay', *mārttika-* 'earthen'. Shina *mʌç, moç* 'earth-cliff' is either unrelated to the Burushaski word or has evolved through metathesis (**mārtta-* > **mart* > **matr* > *mʌç*), the retroflex affricate *ç* being the regular reflex of the earlier cluster *tr* in Shina. Morgenstierne (1945: 92) hypothesizes a metathesis in Burushaski (*mərt* < **mətr*). This hypothesis, however, makes the lexeme under study unetymologizable.

61. Nager ***mázur***, Yasin ***mázar, mazúr*** '*Linsen, Hülsenfrucht*' (Berger 1998b: 285). Berger's etymology: T 9924 (= OIA *masūra-* 'lentil'). The change *s* > *z* is also attested in Shina *mázur* 'lentils'.

62. Hunza-Nager ***meéṣ*** '*Ledertasche*' (Berger 1998b: 286). Berger's etymology: T 10343 (<**maiṣya-* 'ovine'). Cf. also OIA *mēṣa-*, Av. *maēša-* 'sheep, ram'. This comparison was first proposed by Morgenstierne (1935: XXII–XXIII). Since no cognates of the Burushaski word are attested in neighboring Indo-Iranian languages, it may well belong to the early Aryan loan vocabulary.

63. Hunza-Nager ***miná***, Yasin ***miná, mená*** '*Ölkuchen, Aprikosenkerne, aus denen das Öl herausgepreßt ist*' (Berger 1998b: 288). Berger's etymology: T 8174 (= OIA *piṇyāka-* 'oil-cake'). Cf. also Shina *pino* 'kernels from which oil has been expressed', Kumauni *pinā* 'oil-seed refuse', Hindi *pīnā* 'oil-cake'.

64. Hunza-Nager ***mindą́q***, Yasin ***manḍą́ġ*** '*Kranich*' (Berger 1998b: 288). Berger's etymology: T 9744 (< **maṇḍādaka-* 'frog-eater', cf. OIA *maṇḍa-, maṇḍūka-* 'frog'). Cf. also Khowar *maḍáġ, manḍáġ* 'heron'. Borrowing from Khowar, theoretically possible for the Yasin dialect, is for linguogeographical reasons unlikely for Burushaski proper.

65. Nager ***mirdík*** '*die seitliche Wand, die beim Ausheben von Erde entsteht; kleine Lehmwand*' (Berger 1998b: 289). Berger's etymology: T 10286 (= OIA *mṛttikā-* 'earth, clay'). A more probable prototype seems to be **mṛdikā-*, cf. OIA *mṛd-, mṛdā-* 'earth, clay'. The sequence *ir* in the Burushaski word may convey either the retained syllabic *ṛ* or its regular reflex in the donor language.

66. Hunza ***muç***, Yasin ***-muṣṭ*** '*Faust, Handvoll*', Hunza ***múçi***, Nager, Yasin ***muṣṭí*** '*Pflugsterz*' (Berger 1998b: 292). Berger's etymology: T 10221 (= OIA *muṣṭi-* 'clenched hand, fist; handful'). Cf. also Av. *mušti-*, Khotanese *muṣṭu*, Sindhi *muṭhi*, Lahnda, Punjabi *muṭṭh*, Hindi, Gujarati,

Marathi *mūṭh*, Nepali *muṭhi*, Classical Persian *mušt*, Shughni *mut*, Wakhi *məst*, Kashmiri *mŏṭh*, Palula, Kalasha, Khowar *muṣṭi*, Pashai *muṣṭī* 'fist'; Gawar-Bati *muṣṭāk* 'fist', *muṣṭ* 'handle of plough'; Shina *muṭ(h)* 'fist', *muṣṭí* 'handle of plough'. The Hunza word for 'fist' can hardly be borrowed from Shina. If this were the case one should expect Hunza **muṭ*.

67. Hunza-Nager, Yasin ***nas*** '*Geruch, Duft; Parfum; Moschus*' (Berger 1998b: 301). Berger's etymology: T 7093 (= OIA *nāsya-* in *nāsya-grāsam* 'nose cord'). Cf. also OIA *nasya-* 'belonging to or being in the nose', Sindhi *nāsa*, Lahnda, Punjabi, Hindi, Marathi, Kashmiri *nās* 'snuff', Gujarati *nās* 'inhalation of smoke through nose'.

68. Hunza-Nager ***páayo***, Yasin ***pâlu*** '*Keil (zum Spalten von Klötzen, zum Befestigen des Kopfes am Polostock)*' (Berger 1998b: 309). Berger's etymology: T 8032 (= OIA *pāṭayati* 'splits, opens'). Hunza-Nager *ẏ*, Yasin *l* < **ṛ* (< **-ṭ-*, **-ḍ-*). The Old Indian verb, as well as the source form of the Burushaski words, reflects PIE **(s)p(h)el-t-* (Pokorny 1959: 985–987). Note the effect of Fortunatov's law in the donor language.

69. Nager ***pítkiṣ*** '*Gallenblase*', Hunza ***pit*** '*Hitze und Jucken*', Yasin ***-pét, -pít*** '*Galle*' (Berger 1998b: 316). Berger's etymology: T 8181 (= OIA *pitta-* 'bile'). Cf. also Shina *pit*, Kashmiri *pĕth*, Sindhi *pitu*, Punjabi *pitt*, Hindi, Gujarati *pit*, Kumauni *piti* 'bile', Assamese *pit* 'bile, gall-bladder, courage'.

70. Hunza, Yasin ***punár***, Nager ***phunár*** '*Blume*' (Berger 1998b: 318). Berger's etymology T 13846 (= OIA *sphuṇṭati* 'expands', *sphuṇḍatē* 'opens') or T 8257 (= OIA *puṇḍari-* 'lotus'). In some Dardic languages there are words for 'flower', etymologically related to OIA *sphuṇṭati*: Shina *phŭnắr*, Chilisso *phundo*, Gowro *phono*, Indus Kohistani *phuṇ(ḍ)*, Bashkarik *phuṇḍ*. This fact argues for the former of Berger's comparisons.

71. Hunza-Nager ***pháan-***, Yasin ***phân-*** '*(Menschen, Tiere) wachsen, schwellen, (Baum) viele Zweige treiben; durchsickern (Wasser aus einem nicht ganz versperrten Kanal)*' (Berger 1998b: 320). Berger hypothesizes an etymological connection with OIA *sphāna-* 'swelling, increasing' (T 13829). The etymology looks quite plausible.

72. Hunza-Nager, Yasin ***phal*** '*(Getreide-)Korn; ein kleines Stückchen (Fleisch)*' (Berger 1998b: 320). Berger's etymology: T 9051 (= OIA *phala-* 'fruit, seed of a fruit, spot on dice'). Cf. also Ashkun *pol*, Kati *pul*, *pyuŕ* 'grain', Waigali *pol* 'a single grain', Tirahi *phallī* 'grain', Pashai *phal* 'fruit, seed', Shumashti *phäl* 'a single grain', Kalasha *phắlak*, Khowar *pholok*, Palula *phalŭṛī*, Shina *phal*, Kashmiri *phol* 'grain', Sindhi *pharu*, Lahnda, Punjabi, Hindi, Kumauni, Nepali *phal*, Assamese, Bengali *phɔl*, Oriya *phaḷa*, Gujarati, Marathi *phaḷ*, Sinhalese *pala* 'fruit'. Note the greater semantic similarity of the Burushaski word to Dardic than to Indo-Aryan.

73. Hunza-Nager ***-phált-*** '*brechen, aufbrechen, ein Loch hineinbrechen, in die Luft sprengen*', Hunza-Nager ***du-phált-***, Yasin ***du-phâlt-*** '*platzen, (Gewehr) losgehen, (Herz) brechen*' (Berger 1998b: 322). For etymology see "Introduction".

74. Hunza ***phaltóoćiṅ*** '*Wickelgamaschen*' (Berger 1998b: 322). For etymology see "Introduction".

75. Hunza-Nager ***pháqo*** '*Art Brot…; Brot überhaupt*', Yasin ***páqu*** '*Brot*' (Berger 1998b: 323). Berger's etymology T 7621 (= OIA *pakva-* 'cooked, ripe'). The source of the Burushaski words may

as well have been the reflex of PII **pāka-*, cf. OIA *pāka-* 'baking, cooking, burning (e.g. of bricks)', Av. *nasu-pāka-* 'cremating (lit. cooking) corpses'.

76. Hunza-Nager, Yasin ***phar*** '*sich drehend, kreisend, zurück*' (Berger 1998b: 323). Berger's etymology: T 9050 (< **pharati* 'moves, turns'). Cf. also Av. *spar-* 'to tread, hasten', Classical Persian *sipardan* 'to travel', Shina *phar* 'turn', *phĭróĭk̟i* 'to turn, return', Kashmiri *phirun* 'to turn over', Punjabi, Hindi *phirnā*, Kumauni *phirṇo* Nepali *phirnu* 'to wander' (< **phirati*).

77. Hunza-Nager ***pharáṭ -t-***, Yasin ***pharáṭ et-*** '*drehen, winden, schrauben*' (Berger 1998b: 324). Berger's etymology: T 7872 (= OIA *parivartatē* 'revolves, turns round'). Cf. also Shina *pharāṭ thoĭk̟i* 'to twist', Pashai *paraṭi-*, Savi *parat-* 'to turn round', Gawar-Bati *parati-* 'to fall', Lahnda *partaṇ* 'to return', Marathi *paratṇẽ* 'to turn, revolve, return'.

78. Hunza-Nager ***pháta***, Yasin ***pháte*** '*hölzerne Schüssel*' (Berger 1998b: 326). Berger's etymology: T 8055 (= OIA *pātra-* 'drinking vessel, dish'). The word must have been adopted from some Middle or New Indo-Aryan language with *(t)t* < OIA *tr* (cf., e.g. Pali *patta-*, *pātĩ* 'bowl', Prakrit *patta-* 'vessel'). Shina *pătị̆* 'large dish' should also be considered a loanword because in Shina the regular reflex of the old cluster **tr* is *ç*.

79. Hunza ***phíimiṅ***, ***phímićiṅ***, Nager ***phíinin***, Yasin ***phémiṅ*** '*kleine Welle, Wellengekräusel, Schaum*' (Berger 1998b: 328). Berger's etymology: T 9108 (= OIA *phēna-* 'foam, froth'). Cf. also Ashkun *piẽṛe*, *pẽṛe* 'scum, foam', Waigali *pŕẽ*, Kati *pŕẽ*, Dameli *phīna*, Pashai *phēn*, Gawar-Bati *phēṇ*, Kalasha *pheṇu*, Khowar *phenu*, Bashkarik *phə̃n*, Palula *phīn*, Shina *fiṇĕ*, Kashmiri *phīnɨ*, Sindhi *phīṇu*, Kumauni, Gujarati, Marathi *pheṇ*, Ossetic *fɨnk* 'foam', Persian *fīnak* 'froth of the sea'. The intervocalic *m* of the Burushaski forms has probably developed from earlier **n* due to the assimilating influence of the initial labial (*phíimiṅ* < **phíiniṅ*, *phémiṅ* < **phéniṅ*).

80. Hunza-Nager ***-phóiṅ*** '*Schulter*', Yasin ***-phúiṅ*** '*Genick*' (Berger 1998b: 333). Berger's etymology: T 13839 (= OIA *sphya-* 'scapula'). Cf. also Sogdian *βyk*' 'shoulder', Shughni *fiyak*, Munji *fiyo* 'shoulderblade', Wakhi *pəy* 'wooden spade', Waigali *pīwə*, Pashai *phīī*, Khowar *phiu* 'scapula', Gawar-Bati *phīo*, Bashkarik *phī* 'shoulderblade', Indus Kohistani *phíā*, Shina *phižu*, Kashmiri *phyok* 'shoulder'. Direct borrowing from Shina or Khowar into Burushaski seems unlikely because of phonological differences.

81. Hunza-Nager, Yasin ***phut*** '*Art Naturgeist, etwa Gnom*' (Berger 1998b: 338). Berger's etymology: T 9552.2 (= OIA *bhūta-* 'spirit, demon'). Initial *ph* in Burushaski no doubt conveys the original aspirate *bh*. This fact suggests that the source of borrowing was an Indo-Aryan rather than Dardic or Iranian lect. Most probably, we are dealing here with a loanword from Sanskrit or some very archaic Middle Indo-Aryan language retaining the old intervocalic *-t-* (cf. e.g. Pali *bhūta-* 'demon').

82. Hunza-Nager ***qarqaámuċ***, Yasin ***qarqámuś*** '*Huhn, Hahn, Henne*' (Berger 1998b: 343). The word may reflect some derivative of PII **kr̥ka-* 'hen, cock', cf. OIA *kr̥kavāku-*, Av. *kahrka-* 'cock', Middle Persian *kark* 'chicken, hen', Kurdish *kerk*, Ossetic *kark*, Wakhi *kərk* 'hen', Kalasha *kakawáŋk* 'chicken, hen' (Trail & Cooper 1999: 146). Shina *kărkāmŭš* 'hen' is probably a Burushaski loanword.

83. Hunza-Nager ***ráaçi***, Yasin ***ráçi*** '*Wächter, Wachtposten; Schutzgeist*' (Berger 1998b: 361). Berger's etymology: T 10557 (= OIA *rakṣin-* 'guard, watcher'). Cf. also Shina *raçhi* 'guard, watch', Kashmiri *rəčh* 'watchman', Hindi *rākhī* 'field-watcher'.

84. Hunza-Nager ***rajaáki***, Yasin ***rajáki*** '*unbezahlte Arbeit für den Mir bzw. den Staat; Zwangsarbeit*' (Berger 1998b: 362). Berger's etymology: T 10694 (= OIA *rājya-* 'kingship, kingdom', Shina *rʌǰi* 'rule, government'). The elements *-aáki* in the Hunza-Nager and *-áki* in the Yasin dialect seem to be etymologically identical to the Burushaski derivational formant *-aáki*[25]. Shina *rʌjáki* 'compulsory communal labor' must have been borrowed from Burushaski.

85. Hunza-Nager, Yasin ***sar*** '*gedrehter Wollfaden, Garn, Kammgarn*' (Berger 1998b: 375). Berger's etymology: T 13260 (= OIA *sara-* 'string', *sarikā-* 'string of pearls'). The word was also borrowed into Northwestern Tibetan dialects: Balti (Skardu dialect) *syar*, Purik, Ladakhi *sar* 'wick' (Kogan 2019).

86. Hunza-Nager ***sarúu-***, Yasin ***sarú-*** '*(Lasten) hin- und her tragen, in Stationen tragen*' (Berger 1998b: 376). Berger's etymology: T 13358 (= OIA *sārayati* 'makes move, removes'). Cf. also Shina *sərоiki* 'to bring in, fetch (water, grass, crops)', Kashmiri *sārun* 'to transport gradually from one place to another (e.g. grain from threshing floor to house), collect', Kumauni *sārṇo* 'to convey, remove, transplant, copy, separate grain from dirt', Garhwali *sārnu* 'to carry a load', Nepali *sārnu* 'to move'.

87. Hunza-Nager ***sínda***, Yasin ***sénde*** '*Fluß*' (Berger 1998b: 379). Berger's etymology: T 13415 (= OIA *sindhu-* 'river, the Indus, flow (in general), sea'). Cf. also Tirahi *sīnd*, Khowar *sin*, Indus Kohistani *sīn*; Shina *sĭn* 'river'. The Old Indian word is traditionally derived from PII **sindhu-* 'natural border' (Mayrhofer 1996: 729–730). It is, however, argued by some scholars (see, e.g. Pinnow 1953; Harmatta 1992) that borrowing from Burushaski into Aryan may have taken place.

88. Hunza-Nager, Yasin ***su-*** '*erstarrtes Präfix "gut", "sehr" in zahlreichen Adjektiven*' (Berger 1998b: 383). Berger compares the OIA prefix *su-* indicating goodness. Sometimes, however, the Burushaski prefix seems to denote possession: *sutúulo* '*schwer*' (Berger 1998b: 386), cf. *atúulo* '*leicht, schwach*' (Berger 1998b: 386), OIA *tulā-* 'balance, weight', Shina *tul* 'weight, strength'. In such cases it may reflect PII **sa-, *sam-* 'together, with, possessing something' (> OIA *sa-*, *sam-*, Av. *ha-, hąm-*), phonologically influenced by **su-*.

89. Hunza-Nager ***śameéy***, Yasin ***śamé*** '*vier Holzpflöke, von denen je einer am Jochbalken des Pflugs an den Seiten der beiden Ochsen befestigt ist, um sie zusammenzuhalten; einer dieser Pflöke*' (Berger 1998b: 390). Berger's etymology: T 12322 (< **śamyēṣā-*, cf. OIA *śamyā-* 'stick, wedge, peg, yoke pin', *īṣā-* 'pole of carriage or plough'). Cf. also Shina *šamēṣ* 'yoke pin'.

90. Hunza-Nager, Yasin ***śaní*** '*(Gemüse-)Garten*' (Berger 1998b: 390). Berger's etymology: T 12323 (= OIA *śayana-* 'bed'). Cf. Woṭapuri *šēn* 'garden', Shina *šʌni, šɛni* 'garden plot, vegetable plot' (Lorimer 1938: 326) with similar semantic development.

91. Hunza-Nager, Yasin ***śaṅ*** '*wach, aufmerksam; Aufmerksamkeit*' (Berger 1998b: 390). Berger's etymology: T 12258 (= OIA *śaṅkā-* 'fear, distrust'). Cf. also Shina *šo̯ṅ* 'care, anxiety; awake, alert', Khowar *šaṅg* 'fear, suspicion'. The word is attested in Northwestern Tibetan dialects and in Zhangzhung: Balti *shang* 'wisdom, sense', Purik *šaŋ* 'consciousness', Ladakhi *shang*

[25] On this formant see Berger 1998a: 211. Notably, Berger quotes examples of it being affixed to borrowed roots: *musulman-aáki* 'Islam', *kaafir-aáki* 'paganism'.

'alertness, awareness, caution, prudence', Zhangzhung *shan* 'consciousness' < Indo-Iranian (Kogan 2019, 2021).

92. Hunza-Nager ***śaráay,*** Yasin ***śará*** '*(Essen) kalt, abgestanden*' (Berger 1998b: 391). Probably a reflex of PII **śăra-*, **śarta-* 'cold', cf. Av. *sarəta-*, Persian *sard*, Pashto *sōṛ*, Wakhi *sɨr* (< **sāra-*), Khowar *šaru* id., Pashai *šōro* 'ice', Shina *šărāvŭ* 'cold (of food)'.

93. Hunza-Nager, Yasin ***śarán*** '*Ruinen, Mauerreste*' (Berger 1998b: 391). Berger's etymology: T 12326 (= OIA *śaraṇa-* 'protecting; shelter, home'). Due to semantic difference, this comparison seems to be far-fetched. Much more probable is the connection with OIA *śaraṇa-* 'falling asunder, bursting', *śr̥ṇāti* 'crushes, breaks', Av. *a-sarə-ta-* 'not discouraged (lit. 'not broken')', Shina *šəroniǰoiki* 'to be destroyed', Gowro, Chilisso *šar-* 'to break'.

94. Hunza-Nager, Yasin ***śat*** '*Stärke, Kraft, Gewalt; sehr viele, eine große Menge*' (Berger 1998b: 392). Berger's etymology: T 12250 (= OIA *śakti-* 'power'). Cf. also Shina *šăt* 'strength, violence', Indus Kohistani *šɅtʰ* 'strength, power'.

95. Hunza-Nager ***śipít***, Yasin ***iśpít*** '*Medicago falcata Linn., Luzerne (als Pferdefutter angebaut)*' (Berger 1998b: 395). Cf. Proto-Iranian **aśu̯a-asti-* 'horse-food' > Middle Persian *aspast* 'clover, alfalfa', Bactrian *aśpista*, Classical Persian *aspist*, *ispist*, Pashto *špēšta* 'alfalfa' (Rastorgueva & Edelman 2000: 245). Borrowing from certain Iranian lect(s) seems probable for the Burushaski words. However, absence of a sibilant before the final *t* in both of them may point to a source form reflecting the prototype **aśu̯a-atti-*. This source form may well have belonged to some Dardic language (cf. Shina *iśpít* 'alfalfa').

96. Hunza ***śiriḍáko***, Nager ***śurdúṅ ḍáko***, Yasin ***śiríduṅ*** '*Hauptpfosten des Hauses, spielt bei Riten eine Rolle*' (Berger 1998b: 395). Cf. Hunza-Nager *ḍáko* '*hölzerne Säule, Dachpfosten, Zeltstange*' (Berger 1998b: 129), Yasin *duṅ* '*Pfosten*' (Berger 1998b: 395). Berger compares the first component with OIA *śrī-* 'light, beauty; welfare, riches' (T 12708). Semantically, however, the connection with OIA *śiras-* 'head' seems to be more probable, the original meaning of the whole compound thus being 'the main (lit. 'head-') pillar'.

97. Hunza-Nager, Yasin ***śuá*** '*gut, fein, nett, ausgezeichnet; bei guter Gesundheit; in guten Verhältnissen lebend; adv. guṭ*' (Berger 1998b: 397). Berger's etymology: T 12532 (= OIA *śubha-* 'bright, auspicious, good'). Cf. also Shina *šo*, Palula *šuwo* 'good', Indus Kohistani *šṓ* 'good, nice; pretty', Kashmiri *šūbun* 'to be beautiful, to be proper'.

98. Hunza-Nager, Yasin ***-śul*** '*Geburtswehen*', Hunza-Nager ***śul*** '*Schmerz (vor Kälte)*' (Berger 1998b: 398–399). Berger's etymology: T 12575 (= OIA *śūla-* 'spike; sharp pain, colic'). Cf. also Bashkarik *šūl* 'wound', Shina *šūl*, Kashmiri *hɨl* 'labour pains', Sindhi *sūru* 'a pain', Punjabi *sūl* 'spike, thorn, colic', Nepali *sul* 'colic', Hindi *sūl* 'sharp-pointed weapon, sharp pain, colic'.

99. Hunza-Nager, Yasin ***ṣaṅ*** '*Bresche, Zugang in der Mauer, der mit Dornen verschlossen wird, Lücke, Durchbruch; Grenze, Grenzlinie, Richtung*' (Berger 1998b: 404). Probably from earlier **ṣaṅk*[26] < PII **śr̥na-ka-*, cf. OIA *śīrṇa-* 'broken (the perfect participle of *śr̥ṇāti* 'breaks')'.

[26] For *ṅ* < **ṅk* see **91**.

100. Hunza-Nager, Yasin ***ṣiṅ*** '*Röhre, Blutgefäß; Garnrolle (aus einer Rohre bestehend); Federkiel; Schwungfeder*' (Berger 1998b: 408). Berger's etymology: T 12583 (= OIA *śr̥ṅga-* 'horn'). Cf. also Dameli, Kalasha, Gawar-Bati, Palula, Shina *ṣiṅ*, Waigali, Shumashti, Bashkarik *ṣīṅ*, Kashmiri *hĕng*, Sindhi *siṅu*, Punjabi *siṅg*, Hindi *sīṅg* 'horn'.

101. Hunza-Nager, Yasin ***taríṅ*** '*Schlauch aus Tierhaut (zum Aufbewahren von Flüssigkeit und für Floße)*' (Berger 1998b: 422). Berger's etymology: T 5700 (= OIA *taraṇa-* 'crossing; raft, boat'). The etymology is quite plausible because a raft made of inflated skinbags is the usual means for crossing mountain rivers in the region.

102. Hunza-Nager ***tíṣṣi***, Yasin ***téṣṭik*** '*Spanne, Abstand von der Spitze des Daumens bis zu der des kleinen Fingers bei gestreckter Hand*' (Berger 1998b: 427). Berger's etymology: T 6343 (= OIA *diṣṭi-* 'measure of length'). Cf. also Dameli, Khowar, Palula *diṣṭ*, Shina *diṭ* 'span'. The devoicing of original initial d in Burushaski remains unexplained.

103. Hunza ***tiṣk***, Yasin ***teṣk*** '*Dolch*', Nager ***tiṣk*** '*Bajonett*' (Berger 1998b: 427). Berger's etymology: T 5839 (= OIA *tīkṣṇa-* 'sharp'). The source form of the Burushaski words in the donor language does indeed seem to be related to the cited Old Indian adjective reflecting a derivative of the PII root **taig-* 'to be sharp, pointed' (> OIA *tējatē* 'is sharp', Av. *taēža-*, Middle Persian *tyz*, Classical Persian *tēz* 'sharp', Classical Persian *tēγ* 'sword, cutting edge'). However, absence of any traces of the earlier nasal in Burushaski suggests **tīkṣaka-* as a more probable prototype.

104. Hunza-Nager ***újo***, Yasin ***újų*** '*Fischotter, Otter*' (Berger 1998b: 453). Berger's etymology: T 2056 (= OIA *udra-* 'an aquatic animal'). Cf. also Av. *udra-*, Middle Persian *udrak*, Ossetic *wɨrd*, Gawar-Bati *uλ*, Bashkarik *ūl*, Palula *ūdr*, Shina *ŭẓŭ*, *ujo*, Kashmiri *wŏdur* 'otter'. For the change **dr* > *j* see also **50**. This change is shared by the donor lect with some Dardic languages, e.g. Shina and Torwali.

105. Hunza-Nager ***wáar-*** '*bedecken mit, als Deckel benützen, drauflegen (um es zu bedecken)*', ***wáariṣ*** '*Deckel*' (Berger 1998b: 462). Yasin ***yáwarkiṣ*** '*Deckel*' (Ibid.) may also belong here. Cf. OIA *vr̥ṇoti* 'covers', Av. *aiβi-vərənuuaiti* 'conceals'.

106. Hunza-Nager ***yaári***, Yasin ***yári*** '*das erste Morgenlicht; Sonnen- oder Mondstrahlen*' (Berger 1998b: 468). Cf. Av. *ayar-* 'day (as opposed to night)', Khowar *yor* 'sun' (*o* < **a*). Phonologically less probable is the connection with Wakhi *(y)ir* 'sun' (< Proto-Iranian **hūr-*).

107. Hunza-Nager ***yáiṅ***, Yasin ***yánji*** '*Mühle, Wassermühle*' (Berger 1998b: 470). Berger's etymology: T 10412 (= OIA *yantra-* 'controlling device; any implement or contrivance', *yantraka-* 'handmill'). Cf. also Dameli *žandra*, *žan* 'mill', Pashai *žantr* 'watermill', Shumashti *žõ*, Woṭapuri *yaṇ*, Gawar-Bati *žãλ*, *ǰõλ*, Kalasha *žõtr*, Bashkarik *yãl*, Palula *yāndr* 'mill', Shina *yō̲r* 'large watermill', Kashmiri *yĕndɨr* 'spinning wheel', Sindhi *jaṇḍru* 'handmill', Lahnda, Punjabi *jandar* 'watermill'. Hunza *yoór* '*Mühle (mit Wasser betrieben)*' (Berger 1998b: 470) is a direct borrowing from Shina.

108. Hunza-Nager ***zan -t-*** '*zerstoßen, zerstampfen, (Kern) aufschlagen, (Menschen) verletzen*' (Berger 1998b: 482). Cf. OIA *hanti*, Av. *ǰainti* 'beats, strikes; kills' < PII **ǰhanti*. The source of borrowing must have been a language characterized by dentalization of old palatal affricates. Therefore, it could not have been Shina where such a process did not take place.

109. Nager ***zaq*** '*(plotzliches, rasch vergehendes) Kopfweh; adj. verwundet, verletzt, zerstoßen*' (Berger 1998b: 483). Probably with initial sonorization < **čhaq*. The source form of the Burushaski lexeme may have reflected the older prototype **čhataka-*. Cf. OIA *kṣataka-* 'wound', *kṣaṇoti* 'injures, hurts', Khowar *čhek* 'illness, pain' < PIE **tk'en-* (LIV: 645). The same word without sonorization has been borrowed into Northwestern Tibetan dialects: Balti *tshak* 'rheumatism', Purik *tshaq* 'pain with difficulty of breathing', *tshak yong* 'to ache', Ladakhi *tshak* 'sprain, pulled muscle, sudden cramp, sudden sharp pain' (Kogan 2019). Shina *zak* 'headache', Khowar *zaq* 'severe pain in the head' (Lorimer 1938: 389) with initial *z*, quite rare for the inherited vocabulary of these languages, must have been adopted from Burushaski.

Data interpretation and discussion

I have listed those Burushaski words whose origin as early ("pan-Burushaski") Indo-Iranian loans seems certain or highly probable[27]. Probably the most conspicuous feature of the presented list is its historical-phonological heterogeneity. One can find quite a number of instances where a particular PII phoneme or phoneme sequence is reflected differently in the same position. Thus, old initial palatal affricates may either remain unchanged (see **25, 26, 48, 49**) or be dentalized, sometimes with subsequent assibilation (see **20, 21, 22, 108**). Word-initially, the old syllabic sonorant **w* (**u̯*) is in one case retained (see **105**) while in other cases it changes to *b* (see **10, 13, 15, 16, 17, 18**). The old intervocalic *d* is preserved in **7** and **26**,[28] and dropped in **17, 28, 64**. The PII cluster **lt* has developed to a cerebral consonant (Proto-Burushaski **ṛ* > Hunza-Nager *ẏ*, Yasin *l*) in **68** and remained unaffected by Fortunatov's law in **73** and **74**. Clusters of the type "*r* + dental stop" either retain their first component (see **60**) or lose it with or without cerebralization of the historical dental (see **16, 43, 45, 47, 51**). All these facts unambiguously suggest that early Indo-Iranian loan vocabulary in Burushaski can hardly be considered a single etymological stratum stricto sensu, but rather should be subdivided into at least two separate strata — an archaic and an innovative one.

Another important fact, revealed while scrutinizing the above lexical material, is the presence of several cases of apparently Dardic historical-phonological development. Such cases include, e.g. the drop of a cluster-initial sibilant before a historical (Indo-Iranian) voiceless aspirate (PII **šth* > **ṣṭh* > *ṭ(h)*, see **46**) and the change of PII **kš* to *ṣh* (see **30, 31, 32, 83**)[29]. If etymology **109** is correct, the reflexes of PIE **tk'* (> *čh* > *ch*) also represent a Dardic historical-phonological isogloss, **čh* being the Proto-Dardic reflex of this cluster. Notably, this feature indicates that, although belonging to the Dardic group, the language(s) under study can hardly be classified with the East Dardic branch[30] where the Proto-Dardic affricate **čh* was cerebralized in the initial position. The deaspiration of old voiced aspirates, noted in quite a

[27] In the dictionary there are also a number of lexemes for which Berger proposes highly hypothetical Aryan etymologies. Some of these etymologies he himself marks with a question. Such cases should be dealt with separately, and I abstain from analyzing them in the present work.

[28] For **7** borrowing from Sanskrit theoretically cannot be ruled out, whereas for **26** Sanskrit origin seems unlikely due to semantic difference. The old *d* could also be retained between the syllabic sonorant **ṛ̯* and a vowel (see **65**).

[29] One instance is attested of the change **kš* > *ć* (see **24**). In this case, the palatal affricate may well have evolved from the cerebral one (*ć* < **ṣ(h)*). If so, Hunza-Nager *ćar -t-* must belong to the innovative etymological stratum, and the palatalization of the earlier cerebral affricate is most probably connected with the dentalization of historical palatal affricates, both transformations representing the same historical-phonological process.

[30] The East Dardic branch of the Dardic group includes Kashmiri, Shina, Palula and Kohistani languages.

number of words on the above list (see **6, 8, 19, 40, 87, 108**), is shared by Dardic, Iranian and Nuristani. However, no other Iranian or Nuristani traits have been attested in our material. Thus far, it seems impossible to derive any of the two aforementioned etymological strata from any particular Dardic language known to us. It is hoped that future research will throw some light on this problem.

There is one more issue that merits further study. As was demonstrated in my previous publications (Kogan 2019; 2020; 2021), there was in the past a lect adjacent to the Burushaski-speaking area and possessing distinctively Dardic features in terms of historical phonology. The only data on this lect available to us at present is a number of loanwords in Northwestern Tibetan varieties of Ladakh and Baltistan, as well as in the now extinct Zhangzhung language. The question arises whether this Indo-Iranian dialect can be closely related or identical to the source-language of at least some Indo-Iranian borrowings in Burushaski. This question can hardly be answered with certainty at the current state of our knowledge. On one hand, several words on our list do have etymological parallels in Balti, Ladakhi, Purik and Zhangzhung (see **17, 25, 43, 85, 91, 109**). On the other hand, some historical-phonological peculiarities of Indo-Iranian loans in Zhangzhung and Northwestern Tibetan (e.g. **a > o, u* before a prevocalic nasal; **ś > y, 0* intevocalically) are not shared by Burushaski words of Indo-Iranian origin analyzed above[31]. Whether this fact should be attributed to a different source or a different date of borrowing remains unclear.

What is, however, clear from all the above is that the traditional view of Indo-Iranian loan vocabulary in Burushaski is somewhat simplistic and for this reason should not be accepted in full. The issue of Burushaski-Aryan contact is far from being satisfactorily explored and hence needs more scholarly attention than it has received so far.

Abbreviations for language names

Av. — Avestan; OIA — Old Indo-Aryan; PIE — Proto-Indo-European; PII — Proto-Indo-Iranian; Tib. — Tibetan.

References

Bailey, T. Grahame. 1924. *Grammar of the Shina language ("Royal Asiatic Society, Prize Publication Fund", vol. VIII).* London: Royal Asiatic Society.

Berger, Hermann. 1998a. *Die Burushaski-Sprache von Hunza und Nager. Teil I. Grammatik.* Wiesbaden: Harrassowitz Verlag.

Berger, Hermann. 1998b. *Die Burushaski-Sprache von Hunza und Nager. Teil III. Wörterbuch Burushaski-Deutsch, Deutsch-Burushaski. Unter Mitarbeit von Nasiruddin Hunzai.* Wiesbaden: Harrassowitz Verlag.

Buddruss, Georg, Almuth Degener. 2015. *Materialien zur Prasun-Sprache des Afghanischen Hindukusch, Teil I: Texte und Glossar.* Harvard University Press.

Dani, Ahmad Hasan. 2001. *History of Northern Areas of Pakistan (Upto 2000 AD).* Lahore: Sang-e-Meel Publications.

Edelman, Dzhoy I. 2011. *Etimologicheskij slovar' iranskix jazykov. T. 4 (i—k).* Moskva: «Vostochnaja literatura» RAN.

Harmatta, J. 1992. The emergence of the Indo-Iranians: The Indo-Iranian languages. In: A. H. Dani, V. M. Masson (eds.). *History of civilizations of Central Asia. Volume I. The dawn of civilization: earliest times to 700 B.C.*: 357–378. Paris: Unesco Publishing.

Jettmar, Karl. 1975. *Die Religionen des Hindukusch (= Die Religionen der Menschheit. Bd. 4, 1). Mit Beiträgen von Schuyler Jones und Max Klimburg.* Stuttgart: Kohlhammer Verlag.

[31] The innovative Aryan etymological stratum in Burushaski shares one single historical-phonological feature with Aryan loan vocabulary in Northwetern Tibetan dialects and Zhangzhung, namely, the dentalization of old palatal affricates.

Kogan, Anton. 2019. On possible Dardic and Burushaski influence on some Northwestern Tibetan dialects. *Journal of Language Relationship* 17(3–4): 263–284.

Kogan, Anton. 2020. Notes on the historical phonology of Indo-Iranian loanwords in Northwestern Tibetan dialects. *Journal of Language Relationship* 18(3–4): 261–275.

Kogan, Anton. 2021. Towards the reconstruction of language contact in the pre-Tibetan Upper Indus region. *Journal of Language Relationship* 19(3–4): 153–165.

LIV = Kümmel, Martin, Helmut Rix (eds.). 2001. *Lexikon der indogermanischen Verben: Die Wurzeln und ihre Primärstammbildungen. 2nd ed.* Wiesbaden: Reichert.

Livshits, Vladimir A. 2003. Drevneje nazvanije Syrdar'ji. *Vestnik drevnej istorii* 1(244): 3–10.

Lorimer, David L. R. 1935. *The Burushaski language. Vol. I. Introduction and grammar.* Oslo: H. Aschehoug & Co. (W. Nygaard).

Lorimer, David L. R. 1937. Burushaski and its alien neighbours: problems in linguistic contagion. *Transactions of the Philological Society* 1937: 63–98.

Lorimer, David L. R. 1938. *The Burushaski language. Vol. III. Vocabularies and index.* Oslo: H. Aschehoug & Co. (W. Nygaard).

Mayrhofer, Manfred. 1992. *Etymologisches Wörterbuch des Altindoarischen. I Band.* Heidelberg: Carl Winter Universitätsverlag.

Mayrhofer, Manfred. 1996. *Etymologisches Wörterbuch des Altindoarischen. II Band.* Heidelberg: Universitätsverlag C. Winter.

Morgenstierne, Georg. 1935. Preface. In: Lorimer, David L. R. (ed.). *The Burushaski language. Vol. I. Introduction and grammar*: VII–XXX. Oslo: H. Aschehoug & Co. (W. Nygaard).

Morgenstierne, Georg. 1945. Notes on Burushaski Phonology. *Norsk Tidsskrift for Sprogvidenskap* 13: 61–95.

Munshi, Sadaf. 2019. *Srinagar Burushaski. A Descriptive and Comparative Account with Analyzed Texts.* Leiden / Boston: Brill.

Norman, Rebecca. 2010. *A Dictionary of the Language Spoken by Ladakhis.* Ms. (unpublished draft).

Pinnow, Heinz-Jürgen. 1953. Zu den altindischen Gewässernamen. *Beiträge zur Namensforschung* 4: 217–234.

Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch.* Bern / München: Francke Verlag.

Rastorgueva, Vera S., Dzhoy I. Edelman. 2000. *Etimologicheskij slovar' iranskix jazykov. T. 1 (a—ā).* Moskva: «Vostochnaja literatura» RAN.

Rastorgueva, Vera S., Dzhoy I. Edelman. 2003. *Etimologicheskij slovar' iranskix jazykov. T. 2 (b—d).* Moskva: «Vostochnaja literatura» RAN.

Sprigg, Richard K. 2002. *Balti-English, English-Balti dictionary.* New York / London: Routledge.

Starostin, Sergei. 2005. *Burushaski etymology.* Available online at: starlingdb.org/cgi- bin/query.cgi?root=config&morpho=0&basename=\data\sinocauc\buruet (accessed 30.05.2023).

Strand, Richard F. 2006. *Khowar lexicon.* Available online at: https://nuristan.info/lngFrameL.html (accessed 30.05.2023).

Trail, Ronald L., Gregory R. Cooper. 1999. *Kalasha Dictionary with English and Urdu compiled by Ronald L. Trail, Gregory R. Cooper.* Islamabad: National Institute of Pakistan Studies, Quaidi-Azam University, Summer Institute of Linguistics.

Turner, Ralph L. 1966. *A Comparative Dictionary of the Indo-Aryan Languages.* Oxford University Press.

Zoller, Claus P. 2005. *A Grammar and Dictionary of Indus Kohistani. Volume 1: Dictionary.* Berlin / New York: Mouton de Gruyter.

А. И. Коган. Об этимологической стратификации индоиранских заимствований в языке бурушаски.

В статье делается попытка уточнить общепринятые представления о происхождении индоиранской заимствованной лексики в языке бурушаски. Проанализировав обширный лексический материал, автор приходит к выводу о том, что весьма значительная часть подобной лексики была усвоена до распада протобурушаскского состояния и не может быть выведена ни из одного традиционно выделяемого языка-донора (шина, кховар, урду, персидского, восточноиранских языков Памира). Показано также, что

общебурушаскский сегмент арийского лексического пласта этимологически неоднороден, причем одним из его источников, вероятно, является некий дардский идиом.

Ключевые слова: язык бурушаски; арийские языки; дардские языки; языковые контакты; лексические заимствования; этимология.

Е. Е. Арманд
НИУ ВШЭ, Москва; armandlena@yandex.ru

Рефлексы древнеиранских превербов в шугнанском языке[1]

Статья посвящена описанию и диахроническому анализу глаголов шугнанского языка, в состав которых входят превербы общеиранского происхождения (*a-*, *an-*/*im-*/*am-*, *bar-*, *bi-*/*be-*/*b-*, *naǩ-*/*naǰ-*/*niǩ-*, *ni-*/*na-*). Основная задача исследования — определить количество, грамматический статус и семантические признаки этих превербов в современном шугнанском глаголе, и понять, каким образом эти признаки коррелируют с семантикой соответствующих превербов на древне- и/или праиранском уровне. Анализ показывает, что некоторые шугнанские глаголы, образованные с помощью превербов, действительно восходят к общеиранскому состоянию; в то же время значительная часть глаголов имеет параллели только в языках шугнано-рушанской группы и не представлена в других иранских языках, что может косвенно свидетельствовать об определенном периоде продуктивности превербного словообразования уже после распада иранского языкового единства. Разбор глаголов сопровождается примерами синхронного употребления и этимологической информацией. Полученные выводы способствуют прояснению некоторых спорных моментов, связанных с глагольным словообразованием в истории иранских языков.

Ключевые слова: иранские языки; праиранский язык; шугнано-рушанские языки; шугнанский язык; превербы; глагольное словообразование.

Введение

Превербом в русскоязычной иранистике называют словообразовательный глагольный префикс, который может отделяться от глагольной основы формообразующими элементами. Этот термин используется как для древних, так и для современных иранских языков, хотя морфологический статус превербов различается по языкам в разные периоды их истории. Так, например, в древнеперсидском языке превербы уже очевидно срослись с глагольной основой, поскольку в графике писались слитно, хотя и могли отделяться от глагольной основы аугментом (префикс *a-* в прошедших временах). В современном персидском языке превербы отделяются от глагольной основы только грамматическим показателем длительности *mi-* и префиксом отрицания *na-*/*ne-*. По отношению к глагольной основе термины префикс и преверб можно было бы считать синонимами, однако некоторые исследователи, например, Д. И. Эдельман, их различают: для глагольных префиксов используется термин «преверб», а для именных — «префикс».

Превербы, как правило, происходят из наречий и могут функционировать в языке одновременно как наречия, предлоги и послелоги. В языках с базовым порядком слов S–O–V служебные слова, указывающие на связь в синтагме между элементами –V– и –O– и относящиеся к обоим членам синтагмы, помещаются между этими элементами и соотносятся синтаксически как с одним членом синтагмы, так и с другим. Синтаксические элементы, располагающиеся между объектом и глаголом, выступают в грамматической

[1] Исследование осуществлено в рамках Программы фундаментальных исследований НИУ ВШЭ. Автор выражает глубокую благодарность А. В. Дыбо и другим рецензентам за многочисленные замечания, позволившие улучшить общую композицию и устранить отдельные недостатки в статье.

функции послелога по отношению к объекту и в функции преверба по отношению к глагольной форме (подробнее о происхождении превербов, предлогов и послелогов см. Гамкрелидзе & Иванов 1984: 320–326). Так, в древних иранских языках (древнеперсидском и авестийском) существовала развитая система превербов, некоторые из которых одновременно могли быть наречиями, предлогами или послелогами. Эти языковые единицы имели в основном направительную семантику (др.-перс. *ā-* ‘к’, *ava-* ‘прочь, вниз’, *upā-* ‘по направлению к’, *ni-* ‘вниз, по’, *patiy-* ‘снова, к тому же, против’ и др.). В авестийском и древнеперсидском языках было порядка 16 превербов, которые сочетались с глагольными корнями самой разной семантики (движения, состояния, действия и т. п.), однако даже в древних языках проследить значение преверба не всегда представляется возможным, поскольку встречаются случаи, когда значения глаголов с превербом и без него не отличаются очевидным образом.

В среднеперсидском языке развилась новая система превербов, т. е. из древнеперсидских наречий появились новые превербы, которые, как правило, со временем поменяли свой морфологический статус, перестав функционировать как самостоятельные слова, превратившись в приглагольные клитики или в послелоги. Например, ср.-перс. *andar*, встречающийся в древнеперсидских надписях только в функции предлога *atar* (Kent 1953: 166), мог быть предлогом места и времени ‘в, на, к’ или превербом при глаголах, например: *rēxtan* ‘лить’ — *andar rēxtan* ‘вливать’, *rasīdan* ‘достигать, прибывать’ — *andar rasīdan* ‘настигать’, сохраняя, как правило, свое исходное наречное значение ‘внутри’.

Во многих иранских языках древние превербы срослись с глагольными основами, и об их существовании знают только специалисты по истории языка, однако в современных языках мы видим следы древних превербов внутри глагольной основы. Например, в классическом персидском языке начальное *ā-* в глагола *āmadan* ‘приходить’, *āvardan* ‘приносить’, *āvēxtan* ‘вешать’ восходит к древнеиранскому превербу *ā-*, а в глаголе *farmudan* ‘приказывать’ реконструируется древний преверб *fra-*.

Мы будем использовать термин «преверб» по отношению к шугнанской глагольной основе, следуя традиции, хотя эти морфологически выделяемые части словоформы (морфемы) правильнее было бы называть префиксами, поскольку они не обладают ни переставимостью, ни отделимостью, т. е. не отделяются от глагольной основы ни самостоятельными словами, ни показателями отрицания.

Основной задачей нашего исследования было определить количество и состав древних превербов, отраженных в шугнанском глаголе, и понять, сохраняется ли значение древних превербов в семантике современного глагола. В отличие от хорошо изученного персидского (а также таджикского и дари), имеющего богатую письменную традицию и зафиксированного на разных этапах своего развития, шугнанский язык до сих пор практически не имеет письменной фиксации, поэтому мы вынуждены сравнивать древнеиранскую систему превербов непосредственно с современной, не имея промежуточных данных.

Шугнанский язык относится к шугнано-рушанской группе, принадлежащей северно-памирской подгруппе восточноиранской группы иранских языков (Edelman & Dodykhudoeva 2009: 773). Внутри восточноиранских языков принято выделять группу памирских языков (термин «памирские языки» основан на географической, а не на генетической близости, см. Wendtland 2009: 173), к которым, кроме шугнанского, относятся также рушанский, язгулямский, ваханский, ишкашимский, бартангский и др. Эти бесписьменные языки распространены на территории современной Горно-Бадахшанской автономной области республики Таджикистан, а также на территории Афганистана,

Пакистана и КНР. Шугнанский язык распространен в Рашкалинском и Шугнанском районах на Западном Памире и в Афганском Бадахшане. Шугнан находится в долине реки Пяндж и в долине ее притоков — Гунт, Шахдара (Рашкалинский район) и Бадж. Число говорящих носителей примерно 100 тысяч человек (Эдельман & Юсуфбеков 1999). Шугнанский язык представлен двумя разновидностями, собственно шугнанской и баджувской; внутри собственно шугнанского выделяется особый шахдаринский говор с барвазским подговором (Карамшоев 1988: 5).

В качестве исходной информации мы использовали две грамматики шугнанского языка: «Лахҷаи баджуви забони шуғнанӣ» (Баджувский диалект шугнанского языка) Д. Карамшоева (1963) и «Грамматикаи забони шуғнонӣ» (Грамматика шугнанского языка) Т. Бахтибекова (1979). Надо сказать, что обычно превербам уделяется не очень много места в грамматических описаниях; как правило, все сводится к их перечислению и примерам префиксальных глаголов. В грамматике Д. Карамшоева (Карамшоев 1963: 177) приводится список из 5 приставок (*пар-, наx̌-/нар-, ин-, бар-, ни-*), а в грамматике Бахтибекова — 10 (*а-, ан-* (бадж. *ин-*), *бар-, би-, пар-, наx̌-/ниx̌-* (бадж. *нар-*), *пи-, ри-, ни-, ѡи-*) (Бахтибеков 1979: 49–50), однако значения превербов в данных грамматиках не определяются и исторические комментарии не даются. Наше исследование показало, что и второй список нельзя назвать полным.

В статье везде, где возможно, для глаголов с превербами приводятся этимологии, которые нам удалось найти в Этимологическом словаре иранских языков В. С. Расторгуевой и Д. И. Эдельман. В том случае, если этимология не была найдена в словаре, мы приводим разделенную на морфемы основу шугнанского глагола, пользуясь данными Л. Р. Додыхудоевой, изложенными в монографии «Шугнанский глагол в историческом освещении».

Для языков с живой префиксальной словообразовательной моделью (например, русского) в традиционной грамматике приняты две стратегии описания префиксов/приставок. Первая («классическая») предполагает сравнение префиксального глагола с соответствующим глаголом без префикса. Такая стратегия используется в традиционно-ориентированной «Грамматике современного русского языка». При использовании этой стратегии исследователь пытается установить, какое значение добавляет префикс в семантику новообразованного глагола. Эта стратегия хорошо применима для живой словообразовательной модели; в случае же, когда новые глаголы таким образом давно не образуются и мы имеем дело с ограниченным списком простых глаголов, прибегать к ней довольно затруднительно. Данная модель описания по отношению к шугнанскому языку осложняется еще и тем, что не для всех глаголов с превербами мы можем найти соответствующий глагол без преверба или глагол с тем же корнем, но с другим превербом.

Синхронный анализ позволяет некоторым образом выделить правдоподобные синхронные значения для превербов и корня, однако, если обратиться к этимологии, оказывается, что историческое семантическое развитие шло не так, как нам кажется по синхронному анализу, а в некоторых случаях оказывается, что глаголы на самом деле не однокоренные. Если мы не будем использовать этимологические данные и анализировать материал строго синхронно, мы можем получить не вполне адекватные результаты, которые лишь усложнят общую картину, так что мы будем вынуждены приписывать глагольным корням и превербам не свойственные им значения.

Рассмотрим один из немногочисленных парадигматических случаев, когда в языке одновременно сохранился глагол без преверба и три глагола с разными превербами. В шугнанском языке есть четыре глагола с корнем *tīd-*:

— беспрефиксальный *tīdow* (презентная основа *tiy*/*ti*/*tê*) — частоупотребительный, многозначный глагол 1) ‘идти, ходить, ездить, уезжать, улетать’; 2) ‘идти, течь, струиться’ (о потоке и т. п.); 3) ‘исходить, распространяться’; 4) ‘дуть (о ветре)’[2];

— глагол *pitidow*, бадж. *pikidow* (презентная основа шугн. *pitin*, бадж. *pikin*) ‘выдергивать; выкапывать; отделять, удалять’;

— глагол *naxtīdow* (презентная основа *naxti*) 1) ‘выходить’ (в разных значениях — извне, откуда-л., из чего-л.); 2) ‘возникать, появляться, объявляться’ и т. п.;

— многозначный глагол *andīdow*, бадж. *indīdow* (презентная основа шугн. *andiʒ*/*andiz*, бадж. *yindiz*/*indiz*) 1) ‘вставать, подниматься; просыпаться, пробуждаться (ото сна)’; 2) ‘уходить, покидать, отходить, отлипать’; 3) ‘увеличиваться (об объеме, размере, количестве и т. п.)’; 4) ‘вырастать, подрастать, расти, прорастать, появляться (о всходах растений)’; 5) ‘начинаться, возникать’.

Если мы не знаем этимологий этих глаголов и исходим из вероятной тождественности их корней (на основании простого фонетического сходства), то значение превербов можно понять следующим образом: *pi-* ‘вверх’, *nax-*’наружу’, *an-*/*in-* ‘вверх’. Однако при такой интерпретации нас должно смутить несовпадение глагольных парадигм глаголов *tīdow* (NPST.1SG *tíyum*, NPST.1SG.SH *tīm*, PST *tūyd*, PST.F/PL *toyd*) и *pitidow* (NPST шугн. *pitin*/бадж. *pikin*, NPST.3SG шугн. *pitīnt* / бадж. *pikīnt*, PST шугн. *pitid* / бадж. *pikid*) и наличие переходности/каузативности в глаголе *pitidow*, которая не отразилась на фонетическом облике (древнеиранский каузативный суффикс *-aia-* обычно дает в шугнанском *-ê-*, есть и поздний суффикс каузативности *-en-*). В этой паре глаголов значение преверба определяется довольно точно, но значение глагольного корня оказывается другим. Глагол *pitidow* восходит к сочетанию древнеиранского преверба **pati-* и глагольного корня *kan-*, где *kan-* ‘1) копать, рыть; вырезать, выцарапывать; 2) вырывать (с корнем), выдергивать, уничтожать; разрушать, 3) сыпать, насыпать; бросать; класть, помещать, наполнять’ (Эдельман 2011: 199, 205), соответственно, значение ‘выдергивать, выкапывать’ логично вытекает из глагольного корня ‘копать, рыть’ и преверба со значением ‘вверх’. Здесь следует отметить, что в шугнанском языке преверб *pi-* системно изменил свое историческое значение, поскольку значение **pati* обычно реконструируется как ‘против, перед, к, у’; также в шугнанском сохранился прямой потомок **pati* — предлог *pi*, указывающий на направление действия вверх от говорящего, нахождение вверху (Эдельман 2020: 243–245).

Теперь обратимся к глаголу *andīdow*, для которого первым значением в Шугнанско-русском словаре отмечается ‘вставать, подниматься; просыпаться, пробуждаться (ото сна)’. В шугнанском языке сохранилось несколько глаголов с превербом *an-* (баджувское *in-*), и обычно этот преверб сохраняет свое историческое значение совместности (*an-*/*in-* восходит к **ham-*/**han-* — наречие ‘вместе’, префикс, преверб со значениями совместности, соединения, объединения в единое целое, сближения, т. е. ‘вместе, совместно; со-’ (Расторгуева & Эдельман 2007: 344)), исключение составляет только глагол *andīdow*. Подробно проанализировав все примеры, приведенные в словарной статье, мы можем предположить две версии семантического перехода.

[2] В статье все значения шугнанских глаголов и языковые примеры приводятся по Шугнанско-русскому словарю Д. Карамшоева (Карамшоев 1988, 1991, 1999), при этом мы пользуемся электронной версией словаря, размещенной на сайте pamiri.online/dict (Makarov, Melenchenko & Novokshanov 2022), поэтому здесь и далее мы не будем размещать конкретные ссылки на словарь. Если мы используем пример, полученный в результате опроса, то после него будет стоять помета «пример из анкеты 2024», поскольку эти примеры были получены в ходе анкетирования информантов-носителей шугнанского языка во время экспедиции ВШЭ на Памир в 2024 г.

Во-первых, если этот глагольный корень изначально имел значение ‘двигаться, передвигаться, перемещаться’, то семантический переход к ‘уходить, покидать, отходить, отлипать’ кажется логичным: в этом случае сохраняется исходное значение глагольного корня, а преверб привносит значение направления движения ‘из исходной точки к …’ и сохраняется значение совместности, собирательности, совместности (*собрался и пошел*). Следующий семантический переход к ‘встать’ (*встал и пошел*), далее — семантический переход к ‘вставать, подниматься; просыпаться, пробуждаться (ото сна)’ и к ‘вырастать, подрастать, расти, прорастать, появляться (о всходах растений)’.

tu memūn-en=en andoyd=o ar čīd=en γal
2SG гость-PL=3PL уходить.PST.PL=Q DOWN дом=3PL пока
‘Твои гости ушли или еще в доме?’

Вторая гипотеза предполагает, что важным компонентом в семантике глагольного корня было значение ‘течь’: допустимо, что изначально глагол *andīdow* использовался для описания, например, весеннего подъема воды в реке, т. е. о ситуации, когда много воды собралось (сохраняется значение преверба), и, как следствие, река поднялась:

šarvidoǰ andī-c xu yed=i yoδ-ǰ
горная река подняться-PF.F и мост=3SG нести-PF
‘Вода в реке прибыла (букв. горная река поднялась) и снесла мост’.

Если такое понимание глагола *andīdow* верно, то элемент значения ‘вверх’ стал главным признаком этого действия, и все остальные значения имеют этот элемент значения (‘вставать, подниматься’, т. е. *переместить тело вверх*; аналогично ‘вырастать (о людях и растениях)’ значит ‘стать выше, подняться над поверхностью’; от ‘вставать, подниматься’ развилось значение ‘просыпаться, пробуждаться (ото сна)’ и т. п.). Обе гипотезы предполагают, что значение древнего преверба *ham-/han-* ушло на второй план или вовсе исчезло.

Вторая стратегия описания превербов предполагает выделение общей части всех глаголов, в состав которых входит тот или иной преверб; в таком случае исследователь пытается определить тот общий, как правило, направительный элемент значения, который объединяет эти глаголы. Мы будем прибегать к этой стратегии для определения статуса древнеиранского преверба в современном состоянии, а именно для понимания общей ситуации — возможно ли на современном уровне говорить о превербе в составе глагола, или же древнеиранский преверб вошел в состав глагольной формы и его выделение уже невозможно. Будем считать, что выделение преверба на синхронном уровне оправдано, если в нашем списке будет два или более глагола, имеющих соответствующие пары или тройки однокоренных глаголов с другим превербом или без него.

Вторая стратегия осложняется тем, что, во-первых, один шугнанский преверб может восходить к двум или даже трем древнеиранским превербам, а во-вторых, за длительное время существования и употребления лексемы, как правило, изменяют свое исходное значение, и мы фиксируем семантические переходы от прямого значения к переносному, которые, к сожалению, не всегда удается объяснить. Поскольку шугнанский язык является бесписьменным и у нас нет свидетельств употребления глаголов в промежуточный период (к тому же на данный момент нет достаточного диалектного материала для промежуточной реконструкции), мы не можем понять, в какой момент развития языка произошло то или иное изменение в значении глагола[3].

[3] Подробно различные стратегии описания префиксов и префиксальной модели словообразования (применительно к русскому языку) разобраны в монографии Кронгауз 1998.

Вероятно, при работе с бесписьменным языком, но при наличии хорошо описанных древних языков (при изучении большинства иранских языков мы сталкиваемся именно с такой ситуацией) следует прибегать к более сложной стратегии, используя все имеющиеся данные — этимологические и синхронные.

Совершенно очевидно, что сейчас в шугнанском языке, как и в большинстве других иранских языков, глагольная лексика пополняется за счет сложно-именных глаголов, тогда как простые глаголы активно вытесняются из языка и новые глаголы, по всей видимости, довольно давно префиксальным способом не образуются; таким образом, мы имеем дело с замкнутым множеством простых глаголов, иногда малоупотребительных, или неизвестных, например, современным городским жителям.

В ходе работы нам удалось установить и пересмотреть имеющийся список превербов, отраженных в шугнанских глаголах, определить, к каким древнеиранским превербам они восходят, а также свести фонетические варианты современных превербов к одному. Так, у нас имеются случаи, когда один древний преверб дает два–три фонетических варианта в шугнанском, и наоборот, — есть случаи, когда один шугнанский преверб восходит к двум или даже трем разным древнеиранским превербам.

Нам удалось выделить 12 превербов, без учета фонетических вариантов (варианты приводятся через косую черту): *a-, an-/im-/an-, bar-, bi-/be-/b-, nax̌-/naɣ̌-/nix̌-, ni-, far-/ri-, par-, pi-/pid-, war-, wi-/we-, si-/zi-*. В настоящей статье мы рассмотрим шугнанские глаголы с шестью превербами: *a-, an-/im-/an-, bar-, bi-/be-/b-, nax̌-/naɣ̌-/nix̌-, ni-* (анализу глаголов с остальными шестью превербами предполагается в дальнейшем посвятить отдельную публикацию).

1. Глаголы с превербом *a-*

В шугнанском языке мы находим семь глаголов с превербом *a-*: *abêx̌tow, aɣ̌īdow, agāntow, aɣ̌êzdow, ažirtow, ažêrtow, arrix̌tow*. Основанием для выделения преверба *a-* на синхронном уровне являются пары глаголов: *abêx̌tow* (презентная основа *abōz*, 3Sg *abêzd*) — *bêx̌tow* (презентная основа *boz*, 3Sg *bêzd, bozd*), шугн. *ažirtow* — бадж. *nižêrtow* (с тем же значением), *arrix̌tow/arix̌tow/arrax̌tow/arax̌tow* (презентная основа *arāz/arrāz*) — соответствующий глагол без преверба *rix̌tow* (презентная основа *raz*) — глагол с другим превербом *birêx̌tow* (презентная основа *birêz*, PST *birōx̌t*); глаголы *aɣ̌īdow, agāntow, aɣ̌êzdow* однокоренные.

Шугнанский преверб *a-* в большинстве случаев восходит к праиранскому **ā-*. В авестийском *ā-* могло быть наречием, предлогом или превербом с различными значениями: направления, местонахождения, нахождения близ чего-либо, около, рядом, совместно, за чем-то и т. п. (Расторгуева & Эдельман 2000: 299). В древнеперсидском преверб *ā-* чаще указывал на направление действия к говорящему ‘сюда, в эту сторону’ (Арманд 2005: 138–141). В современном персидском языке преверб *ā-* отражен, например, в глаголах *āmadan* ‘приходить’ (из *ā-* + *gam-* ‘идти’ основа прошедшего времени и *ā-* + *ay-* основа настоящего времени), *āvardan* ‘приносить’ (из *ā-* + *bar-* ‘нести’), *āxtan* ‘извлекать, обнажать (напр. саблю)’ из **ā-Θn̥xta-* (Hassandoust 2016: 23) и др., т. е. в персидских глаголах во многих случаях ориентация действия к говорящему/субъекту сохраняется. Однако имеются примеры, где проследить значение ‘к себе / к говорящему / к субъекту действия’ проблематично, например, персидский глагол *ārastan* ‘украшать, наряжать; приводить в порядок, устраивать’ из **ā-rāsta-*/**a-rāda-*, где *rād-* = ‘готовить, приводить в порядок’ (Hassandoust 2016: 32). В данном случае можно говорить об ориентации действия по направлению к объекту.

1.1. Глагол *abêx̌tow* (презентная основа *abōz*, 3Sg *abêzd*) ‘глотать, проглатывать; затягиваться (при курении)’, согласно Расторгуева & Эдельман 2000: 288, состоит из двух древних превербов: *a-* и *upa-* и глагольного корня *aź-* ‘гнать, погонять, вести’. Глагол *abêx̌tow* употребляется в современном языке. В шугнанском языке также употребляется глагол без преверба *a-*: *bêx̌tow* (презентная основа *boz*, 3Sg *bêzd, bozd*) ‘посылать, отправлять, отсылать’ из **apa-aźa-aia-*.

(1) шугн.
pero aboz xu bād gap δa
перед глотать.IMP и потом слово давать.IMP
‘Сначала проглоти, а потом разговаривай! (пример из анкеты 2024)’

По всей видимости, в данном случае преверб *a-* имеет направительное значение, ориентируя действие к субъекту, и мы можем приписать ему значение ‘к себе, в себя’, т. е. *abêx̌tow* букв. значит ‘вести к себе, вводить в себя’.

1.2. Шугнанский непереходный глагол *ažirtow* (презентная основа *ažir*) ‘мокнуть, намокнуть, размокать, намокать’ состоит из преверба **ā-* и глагольного корня **gar-* 1) ’увлажнять(-ся); наполнять(-ся) жидкостью’; 2) ‘капать’; переходный/каузативный коррелят — *ažêrtow* (презентная основа *ažêr*) ‘мочить, смачивать, размачивать’ из **ā-gār-aia* (Расторгуева & Эдельман 2007: 152–153). В собственно шугнанском мы не находим ни соответствующего глагола без преверба, ни однокоренного глагола с другим превербом, однако в баджувском диалекте, а также в рушанском, рошорвском, хуфском имеется глагол с тем же значением, образованный от того же корня с превербом *ni-*: бадж. *nižêrtow* (*nižêr*).

(2) шугн.
ku čis yid půst ažū-d=o?
ну-ка смотреть.IMP D2.SG кожа мокнуть-PST=Q
‘Взгляни-ка, не намокла ли шкура?’

(3) шугн.
wuz sā-m půst ažêrt-ow
1SG идти-1SG кожа мочить-INF
‘Я пойду намочить кожу’.

В современном языке этот глагол также употребляется в значении ‘замочить белье, одежду (перед стиркой)’:

(4) шугн.
uz-ta nur vega xu luq ažêr-um
1SG-EMPH сегодня вечер REFL тряпка замочить-1SG
‘Сегодня вечером я замочу белье’ (пример из анкеты 2024).

В глаголах *ažirtow* и *ažêrtow*, составляющих пару по признаку некаузативный — каузативный и означающих соответственно ‘поглощать жидкость’, т. е. ‘вбирать жидкость в себя’, и ‘намочить’, т. е. ‘сделать так, чтобы некоторый объект вобрал в себя жидкость’, преверб *a-* придает действию направление ‘внутрь’, ‘к субъекту’ в некаузативном глаголе и ‘к объекту’ при переходном/каузативном глаголе.

1.3. *arrix̌tow*/*arix̌tow*/*arrax̌tow*/*arax̌tow* (презентная основа *arāz*/*arrāz*) 1) ‘прыгать, подскакивать; тянуться (вверх); подниматься на дыбы’; 2) ‘бить ключом, струей; брызгать’;

3) перен. ‘чваниться, пыжиться, задирать нос, зазнаваться’. В данном случае мы можем выделить преверб на синхронном уровне, поскольку имеется глагол *rix̌tow* (презентная основа *raz*) ‘литься, сыпаться, осыпаться, опадать, высыпаться, выпадать’; также есть глагол с тем же корнем, но с другим превербом *birêx̌tow* (презентная основа *birêz*, PST *birōx̌t*) ‘пить’.

Относительно этимологии этого глагола мнения исследователей расходятся; так, Моргенстьерне возводит этот глагол к *ā-* + *raź-*, а Д. И. Эдельман к *fra-* + *raź-* (*raź-* ‘направлять(ся), выпрямлять(ся); приводить в порядок’) (Эдельман 2020: 434); поскольку у глагола есть вариативность с удвоенным *r*, версия Д. Эдельман кажется более вероятной. При этом преверб *fra-* чаще имеет значение ‘вперед’, в данном же случае мы вынуждены считать значение ‘вверх’ более логичным, на что также указывает предлог *pi*[4] и местоименное наречие *pad-am* ‘там (наверху), туда (наверх)’ в примере:

(5) шугн.
gūx̌t awezůn pi wiδům-at piš-i tu aroz-d pad-am
мясо висячий UP потолок-ADD кошка-EZ 2SG подскакивать-PST UP-D3
‘Мясо висит под потолком, а твоя кошка подпрыгивает туда’.

Также у этого глагола развилось переносное значение, не отмеченное в словаре Д. Карамшоева — ‘держаться надменно, важничать, задирать нос, зазнаваться’. Видимо, оно развилось от значения ‘тянуться (вверх)’:

(6) шугн.
yā ta lap patar-āǰ arroz-d
D3.F.SG там очень вверх-DIR прыгать-3SG
‘Она о себе высокого мнения (пример из анкеты 2024)’.

По всей видимости, то, что в современном словаре подается как второе значение глагола ‘бить ключом, струей; брызгать’, в действительности представляет собой префиксальное образование на *a-* от глагола *rix̌tow* (презентная основа *rāz*) ‘литься, сыпаться, осыпаться, опадать’, который восходит к **raik-* : *rik-* ‘лить(-ся), сыпать(-ся), выливать(-ся); осыпать(-ся)’ (Эдельман 2020: 347). В этом случае мы можем приписать превербу направительное значение ‘вверх’. По утверждению консультанта, глагол *arrix̌tow*/*arix̌tow*/*arrax̌tow*/*arax̌tow* в значении ‘бить ключом, струей; брызгать’ в современном шугнанском языке не употребляется; возможно, это относится лишь к диалекту города Хорога, а в сельской местности или в баджувском диалекте языковая ситуация может быть другой.

(7) бадж.
yā x̌ac pās-pati arax̌-t
D3.F.SG вода подъем прыгать-PST.F
‘Вода бьет струей?’

1.4. Глагол *aǰ̌īdow*/бадж. *awīdow* (презентная основа *aǰ̌as* / бадж. *awas*) 1) ‘ложиться, укладываться, разваливаться’ (чаще о животных); 2) ‘никнуть к земле, стелиться’ (о рас-

[4] В шугнанском языке направительные предлоги различают ориентацию по вертикали и горизонтали, так, предлог *ar* ‘в, на; до (вниз)’, используется для обозначения перемещения к ориентиру ниже говорящего; предлог *pi* ‘в, на, к, до (вверх)’, используется для обозначения перемещения к ориентиру выше говорящего; предлог *tar* ‘в, на, до, к (обычно по горизонтали)’, используется для обозначения перемещения к ориентиру на одном уровне с говорящим.

тениях). В языках шугнано-рушанской группы налицо результат контаминации рефлексов корней **hak-* ‘лежать, покоиться, иметь опору’ и **hāi-* в значениях ‘ложиться, покоиться; лежать’. Глагол *aɣ̌īdow* этимологически состоит из преверба **ā-* и глагольного корня **hak-* : *hač-*/**-šak-* : *-šač-* ‘лежать, покоиться, иметь опору’ (Расторгуева & Эдельман 2007: 341).

(8) шугн.
yid čūšč aɣ̌ūy-d xu pu-d yid
D2.SG ячмень ложиться-PST и портиться-PST D2.SG
‘Этот ячмень полег и сгнил’.

Согласно опросу, проведенному в ходе экспедиции 2024 года, этот глагол употребляется в современном шугнанском языке, однако городские жители Хорога его мало используют, а в кишлаках, по мнению консультанта, он употребляется, в том числе по отношению к людям, как отмечено в словаре Карамшоева, *неодобрительно, шутливо*:

(9) шугн.
tamay-et čīz maδori=n-de maɣ̌-en dastūr aɣ̌ī-c
2PL-2PL что полдень=3PL-COMP овца-PL похожий ложиться-PF.F
‘Чего это вы в обед развалились как овцы?! (пример из анкеты 2024)’

1.5. Также в шугнанском языке сохранился каузативный глагол с тем же корнем *agāntow* / шахдар. *aɣ̌āntow* (презентная основа шугнано-бадж. *agān* / шахдар. *aɣ̌ān*) ‘покрывать, укрывать; укутывать (например, одеялом)’ из **ā-(hi)-šā-nā* или **ā-(hi)-ši-nā*, где *hi-* = ‘стлать, расстилать; укрывать; класть, накладывать; бросать’ (Расторгуева & Эдельман 2007: 321). Соответствующего глагола без преверба нет.

(10) бадж.
bašānd-aθ=en wi agān-č
хороший-INT=3PL D3.M.SG.O укрывать-PF
‘Его хорошенько накрыли (одеялами)’.

1.6. шугн. *aɣ̌êzdow* (презентная основа *aɣ̌êz*) 1) ‘укладывать спать’; 2) ‘придавливать, приминать’; 3) ‘давать отдыхать, давать покой (животным)’. Это каузативный глагол от корня **hak-* : *hač-* / **-šak-* : *-šač-* ‘лежать, покоиться, иметь опору’: **ā-šač-aia-* (Расторгуева & Эдельман 2007: 341–342). Соответствующего глагола без преверба нет.

(11) шугн.
ů ɣ̌in di xu puc aɣ̌êz
эй жена D2.M.SG.O REFL сын укладывать спать.IMP
‘Жена, уложи спать своего сына’.

Поскольку очевидное направительное значение ‘вниз’ вшито в семантику глагольных корней ‘класть’ и ‘лежать’, на долю преверба *a-* в глаголах *aɣ̌īdow*, *agāntow* и *aɣ̌êzdow* остается либо пустое значение, либо же значение ‘нахождение вблизи от субъекта действия’.

Таким образом, преверб *a-*, имеющий довольно размытую семантику и в древних иранских языках, придает действию направление ‘к’ или даже ‘внутрь’ (*abêx̌tow*, *ažirtow*, *ažêrtow*), или значение ‘местонахождение вблизи от субъекта действия’ (*aɣ̌īdow*, *agāntow*, *aɣ̌êzdow*), или значение, не отмеченное для древнеиранских языков — ‘вверх’ (*arrix̌tow*).

2. Глаголы с превербом *an-*

В шугнанско-русском словаре *Д.* Карамшоева зафиксировано шесть глаголов с префиксом шугн. *an-* / бадж. *im-* / (редко) *am-*, и еще один глагол отмечен только в словаре Зарубина: *anǰīvdow*, *anǰafctow*, *ancīvdow*, *angixtow*, *andīdow*, *anduzdow*, *ambīθ*. Мы можем говорить о возможности выделить этот преверб на синхронном уровне, поскольку у трех глаголов из шести в языке сохранились либо однокоренные глаголы без преверба, либо глаголы с другим превербом: шугн. *anǰīvdow*, бадж. *inǰīvdow* (презентная основа шугн. *anǰāv*, бадж. *inǰāv*) — *parǰīvdow* (презентная основа ш.-б. *parǰīv* бадж. *parǰāv*); шугн. *angixtow* / шугн. *angaxtow,* бадж. *ingixtow,* бадж. *yingixtow* (презентная основа шугн. *angaxc*, бадж. *yingaxc*, бадж. *ingaxc*) — *bidgixtow* (презентная основа *bidgaxc*); шугн. *andīdow*, бадж. *indīdow* (презентная основа шугн. *andiʒ*/*andiz*, бадж. *yindiz*/*indiz*) — *tīdow* (презентная основа *ti*, 3SG *tīz-d*) — глагол с тем же корнем, но другим превербом *naxǰtīdow* (презентная основа *naxǰti*, 3SG *naxǰtīz-d*).

Преверб шугн. *an-* / бадж. *im-* / (редко) *am-*, по всей видимости, восходит к древнеиранскому **ham-*/*han-* (перед язычными согласными) : *ha-* — префикс, преверб со значениями совместности, соединения, объединения в единое целое; сближения, т. е. ‘вместе, совместно; со-’ (Расторгуева & Эдельман 2007: 344), ср. также наречие **hama-* ‘вместе’.

2.1. Шугн. *anǰīvdow*, бадж. *inǰīvdow* (презентная основа шугн. *anǰāv*, бадж. *inǰāv*) ‘держать; держать(ся), брать(ся), хватать(ся); поймать, схватить’. Оба эти глагола восходят к **han-* + **kap-*/**xap-*/**čap-* ‘хватать, схватывать; кусать, хватать ртом’ (Эдельман 2011: 238); в шугнанском языке имеется глагол с тем же корнем, но с другим превербом *parǰīvdow* ‘брать; отнимать, отбирать, забирать’. Также в шугнанском встречается специализированный глагол с тем же корнем и префиксом *anǰafctow* ‘браться, приниматься, начинаться (об осадках)’ с суффиксом *-c* из инхоативного суффикса *-sa* и с поздним сокращением корневого **ā*.

(12) шугн.
wuz=um wi duzd daraw-aθ anǰūv-d
1SG=1SG D3.M.SG.O вор INC-ADV держать-PST
‘Я сразу поймал вора’.

У этого глагола развилось переносное значение, не отмеченное в словаре *Д.* Карамшоева — ‘болеть (о животе), прихватить живот после употребления жирной пищи’:

(13) шугн.
biyor=um xub lap marub xū-d xu yu=yi mu anǰūv-d
вчера=1SG хороший очень сметана есть-PST и D3.M.SG=3SG 1SG.O держать-PST
‘Я вчера съел много домашней сметаны, и у меня прихватило живот (пример из анкеты 2024)’.

(14) шугн.
dinyo-ta nur anǰafc-t
белый свет-FUT.DEF сегодня начинаться, идти (об осадках)-3SG
‘Сегодня пойдет дождь (или снег)’.

В глаголе *anǰīvdow* ‘схватить и держать при себе’ преверб *ham-* сохраняет значение совместности, сближения, объединения. Глагол *anǰafctow* в примере (14) прежде всего сохраняет значение инхоативности, но также имеется элемент объединения, поскольку он употребляется только по отношению к осадкам (дождю и снегу); возможно, такое упо-

требление основано на том, что сначала собираются тучи, а потом идет снег или дождь, т. е. значение совместности присутствует (аналогично русскому *собирается дождь*).

2.2. Шугнанский глагол *angixtow* / шугн. *angaxtow,* бадж. *ingixtow,* бадж. *yingixtow* (презентная основа шугн. *angaxc*, бадж. *yingaxc*, бадж. *ingaxc*) 1) ‘попадать; застревать; вонзаться’; 2) ‘задевать, касаться’; 3) ‘заикаться, запинаться’; 4) ‘попадать впросак; оказываться в трудном положении, испытывать трудности, затрудняться’; из **han-kuxsā-* : *han-kuxta-* (**kau-k* : ** kau-č-* и др. ‘скрючиться, свернуться; наклоняться, застрять; уклониться, избегать’) (Эдельман 2011: 361–363). В шугнанском языке имеется глагол с другим превербом, но с тем же корнем *bidgixtow* ‘спотыкаться, запинаться; запинаться, говорить с заминками’.

(15) бадж.
růpcak ya-t xu ar qapqen ingix-t
лиса приходить-PST и DOWN капкан попадать-PST
‘Лиса пришла и попала в капкан’.

Этот глагол употребляется в современном языке во всех перечисленных в словаре значениях, в том числе и в переносном ‘попасть в затруднительное положение’:

(16) шугн.
wam čor tūy-d tar x̌ār xu yā xu ziryot-en=qati angix-t
D3.F.SG.O муж идти.M-PST EQ город и D3.F.SG REFL ребенок-PL=COM попадать-PST
‘Ее муж уехал в город, и она со своими детьми попала в тяжелое положение (пример из анкеты 2024)’.

Ситуации ‘попасть в капкан, застрять’, как и ситуация ‘попасть в затруднительное положение’ подразумевают потерю подвижности, возможности продолжать существовать в прежнем состоянии; то же можно сказать и про ‘заикаться’ — физически не иметь возможность сказать, потерять способность к произнесению. Ни в одном из этих случаев нельзя выделить элемент значения ‘совместность, соединение, объединение’. Хотя ситуация ‘попасть в капкан, застрять’ предполагает соединение всего или части тела животного с капканом, при дальнейшем развитии значений (метафоризации) происходит опущение валентности того, с чем объединяется субъект.

2.3. Шугн. *andīdow*, бадж. *indīdow* (презентная основа шугн. *andiʒ*/*andiz*, бадж. *yindiz*/*indiz*) многозначный глагол 1) ‘вставать, подниматься; просыпаться, пробуждаться (ото сна)’; 2) ‘уходить, покидать, отходить, отлипать’; 3) ‘увеличиваться (об объеме, размере, количестве и т. п.)’; 4) ‘вырастать, подрастать, расти, прорастать, появляться (о всходах растений)’; 5) ‘начинаться, возникать’ из **han-* и *tak-* ‘течь, идти’ в виде основ *-tača-* и **-tač-aia-*; а также каузативный глагол с тем же корнем шугн. *anduzdow*, бадж. *induzdow* (презентная основа *anduz*: *anduzd* бадж. *induz : induzd*) ‘будить, поднимать (с места, с постели)’ (Расторгуева & Эдельман 2007: 345). Поскольку в шугнанском языке имеются глаголы с тем же корнем (многозначный глагол без преверба *tīdow* ‘идти; ходить, ездить; отправляться, уезжать, улетать и др.’; глагол с другим превербом *naxtīdow* ‘выходить (в разн. знач. извне, откуда-л., из чего-л.), возникать, появляться и др.’, выделение преверба в данном случае оправдано.

(17) шугн.
yā wam yax andoy-d xu diga na-nost
D3.F.SG D3.F.SG.O сестра вставать-PST и другой NEG-сидеть-PST
‘Ее сестра поднялась и больше не садилась’.

(18) шугн.
tu memūn-en=en andoyd=o ar čīd=en ɣal
2SG гость-PL=3PL уходить=Q DOWN дом=3PL пока
‘Твои гости ушли или еще в доме?’

(19) шугн.
šarvidoǰ andī-c xu yed=i yoδ-ǰ
горная река подняться-PF.F и мост=3SG нести-PF
‘Вода в реке прибыла (букв. горная река поднялась) и снесла мост’.

(20) шугн.
yu tar xūδm ca ve-d awal
D3.M.SG EQ сон SUBD быть-3SG сначала
wi andůz xu tām as wi peǩc
D3.M.SG.O будить.IMP и тогда EL D3.M.SG.O спрашивать.IMP
‘Если он спит, то сначала разбуди его и потом спроси у него’.

Этот глагол употребляется во всех ситуациях, описывающих природные явления, когда что-то поднимается с земли (ветер, пыль, снег):

(21) шугн.
vega dis xaǩ birušā andūy-d
вчера таким образом твердый вьюга вставать-PST
‘Вчера поднялась сильная метель (пример из анкеты 2024)’.

(22) шугн.
mošin naɣǰad-at sitidax andūy-d
машина проходить-PST-ADD клубы пыли вставать-PST
‘Машина проехала и подняла пыль столбом (пример из анкеты 2024)’.

Глагол *andīdow* уже был подробно рассмотрен во вступительной части статьи, где мы предположили, что, образовавшись от корня глагола *tīdow* ‘идти, ходить, течь и др.’, он изначально мог использоваться для описания, например, весеннего подъема воды в реке, т. е. в ситуации, когда много воды собралось, и, как следствие, река поднялась. Если такое понимание глагола *andīdow* верно, то элемент значения ‘вверх’ стал главным признаком этого действия, и все остальные значения глагола *andīdow* его имеют (‘вставать, подниматься’, т. е. переместить тело вверх; аналогично ‘вырастать (о людях и растениях)’ значит ‘стать выше, подняться над поверхностью’; от ‘вставать, подниматься’ развилось значение ‘просыпаться, пробуждаться (ото сна)’ и т. п.).

Если же сравнивать семантику глагола *andīdow* с соответствующим глаголом без преверба *tīdow*, который употребляется как глагол движения ‘идти’, не зная его этимологии, то нам следовало бы выделить в глаголе *andīdow* значение ‘вверх’, и даже в примере (18) этот элемент значения присутствует, поскольку для того, чтобы уйти, нужно встать (гости в памирском доме сидят на полу), а затем пойти. В примере (19), который иллюстрирует значение ‘увеличиваться (об объеме, размере, количестве, уровне и т. д.)’ глагол *andīdow* также имеет элемент значения ‘вверх’, поскольку, когда вода в реке прибывает, она поднимается.

2.4. Шугнанский глагол *ancīvdow*, бадж. *incīvdow* (презентная основа шугн. *ancav*, бадж. *incāv*) ‘шить, зашивать’ происходит из **ham-* + **darb-*, где *darb-* = ‘соединять, связы-

вать, скручивать’ (Расторгуева & Эдельман 2003: 346). Мы не находим в современном языке соответствующего глагола без преверба или с другим превербом, но с тем же корнем.

(23) шугн.
va wuz tu-rd lef ancā-m
давай 1SG 2SG-LAT одеяло шить-1SG
‘Давай я сошью тебе одеяло’ (пример из анкеты 2024).

Поскольку в самом глаголе ‘шить, зашивать’ заложена семантика сближения, совместности, соединения (процесс сшивания подразумевает соединение двух и более частей ткани), можно говорить, что и в глаголе *ancīvdow* преверб либо дублирует совместность, заложенную в значении глагольного корня, либо не имеет собственного элемента значения.

Этот глагол, как и в русском языке, используется в качестве медицинского термина для описания хирургической операции:

(24) шугн.
operaciya tayor su-t-at waδ γal anc-en we
операция готовый идти-PST-ADD D3.PL еще шить-3PL D3.PL.O
‘Операция прошла хорошо, они сейчас его зашивают (пример из анкеты 2024)’.

Также этот глагол используется в устойчивом выражении *cem ancivdow* ‘не смотреть, отворачиваться’, *букв*. ‘зашить глаза’:

(25) шугн.
uz bof-t xu cem ancā-m
1SG годиться-3SG REFL глаз шить-1SG
‘Мне, что, не смотреть?’ (букв. ‘мне, что, зашить глаза?’) (пример из анкеты 2024)

2.5. Баджувский глагол *ambīθ* (презентная основа *ambīθ*/*ambis*) ‘обваливаться, рушиться’ (в словаре Карамшоева форма инфинитива не приводится) из **ham-pad-s-i̯a-*, где **pad-* = ‘двигаться, идти, ступать; спускаться; класть, ложиться, падать’ (Эдельман 2020: 68). Этот глагол зафиксирован только в баджувском диалекте; жители города Хорога, носители собственно шугнанского, его не знают. Соответствующего глагола без преверба или однокоренного глагола с другим превербом не найдено.

(26) бадж.
wim ziδůn burǰ ambūs-t xu δo-d ar wêδ
D3.F.SG.O кладовка стена обваливаться-PST и падать-PST DOWN арык
‘Стена кладовой рухнула и повалилась в арык’.

Здесь можно говорить, что значение преверба ‘вместе, совместно с’ сохраняется, поскольку ‘обвалиться’ можно сказать только о крупном объекте, в данном случае о стене, т. е. ‘упасть единым целым’. Это единственный глагол с вариантом преверба *ham-*, поскольку глагольный корень начинается с губного согласного.

Таким образом, преверб *an-*/*in-* сохраняет свое исходное значение совместности, сближения, объединения не во всех глаголах; элемент значения ‘совместность’ не обнаружен как минимум в глаголе *angixtow* / шугн. *angaxtow* ‘попадать; застревать; вонзаться и др.’, а также в развившихся переносных значениях у глагола *andīdow*, бадж. *indīdow* и образованного от него каузативного шугн. *anduzdow*, бадж. *induzdow*.

3. Глаголы с превербом *bar-*

В *Грамматикаи забони шуғнонӣ* (Грамматика шугнанского языка) Т. Бахтибеков отмечает наличие преверба *bar-*, который, по всей видимости, является недавним заимствованием из таджикского; этимологически он восходит к **upari*, др.-перс. *upariy* ‘сверху’ (в древних языках было наречием, в ср.-перс. стало предлогом и превербом, в классическом персидском — редко, послелогом), в сочетании с глагольными основами означает направление действия вверх, а также ‘на, к поверхности’. В словаре Карамшоева отмечается, что преверб *bar-* употребляется а) в именной части ряда сложных глаголов, например, *bar-ǰoy s.* ‘устраиваться’ (букв. ‘на место становиться’); *bar-boδ č.* ‘ронять, упускать’ (букв. ‘делать на ветер’); б) в нескольких существительных: *bar-dirůɣ* ‘ложный’; *bardam* ‘здоровый, бодрый’. Во всех приведенных примерах *bar* является предлогом в заимствованных сочетаниях. Он также выделяется в глаголе *bar-nêǰdow* ‘крутить’, где *nêǰdow* = ‘вращать, крутить, вертеть; кружить’, т. е. в этом единственном отмеченном в словаре глаголе преверб никак не меняет значения (примеры, к сожалению, не приведены). Согласно опросу, проведенному в ходе экспедиции 2024 года, в шугнанском языке нет своих глаголов, образованных с помощью преверба *bar*, есть только заимствованные из таджикского языка глаголы с этим префиксом. Поскольку в последние десятилетия наблюдается тенденция к вытеснению таджикских заимствований из шугнанского языка, носители предпочитают не употреблять очевидные таджикские слова при наличии собственного эквивалента.

4. Глаголы с превербом *bi-/be-/b-*

Из двенадцати глаголов с превербом *bi-/be-/b-* пять имеют пары соответствующих глаголов без преверба или с другим превербом: *biridow* (презентная основа *birāw- : birud,* ж.р. *birad*) — *siridow* (презентная основа ш.-б. *sirāw*, бадж. *sirīws*); ш.-б. *biδīvdow*, бадж. *biδāvdow* (презентная основа ш.-б. *biδafc*) — шугн. *niδīvdow*, бадж. *niδāvdow* (презентная основа *niδafc*); *bizīdow* (презентная основа *bizin*) — *zintow* (презентная основа *zin)* — *zīdow* (презентная основа *zīn*); *birêx̌tow* (презентная основа *birêz,* PST *birōx̌t*) — *arrix̌tow/ arix̌tow/arrax̌tow/arax̌tow* (презентная основа *arāz/arrāz*); у глагола *bīftow* (презентная основа *bāf,* 3SG *boft*) есть пара — *firīptow* (презентная основа *firāp*), однако вряд ли носители языка осознают их как однокоренные.

Глагольный префикс *bi-/be-/b-*, по всей видимости, восходит к двум разным праиранским наречиям-превербам: 1. ир. **apa-* наречие или преверб, префикс ‘от, прочь; отдельно; далеко, позади’ (Расторгуева & Эдельман 2000: 182) и 2. **upa-* ‘по направлению к’. В современном шугнанском мы обнаруживаем двенадцать глаголов с превербом *bi-/be-/b-*: *bedow, binêstow, biǰentow, birêwtow, biridow, bêx̌tow, bīftow, biδīvdow, biδemtow, bizīdow, birêx̌tow, bix̌čīdow.* К сожалению, из-за фонетического совпадения и схожей семантики не всегда представляется возможным точно сказать, какой именно древний преверб реконструируется в том или ином глаголе.

4.1. Непереходный глагол *biridow* (презентная основа *birāw- :* 3SG *birud*) ‘перестать питаться материнским молоком, бросить сосать (грудь)’ и каузативный глагол с тем же корнем *birêwtow/birêwdow* (презентная основа *birêw*) ‘отнимать от груди, переставать кормить молоком; отнимать от матки (молодняк)’ восходят к *apa-rāu̯-* и соответственно к **apa-rāu̯-aia-*, где *rau- : ru-* = ‘отделять, удалять, отрезать’ (Эдельман 2020: 392–393). Этот глагол употребляется по отношению как к животным, так и людям. В шугнанском языке

имеется однокоренной глагол с другим превербом *sirêwdow* (презентная основа *sirêw, sirêwt*), *siridow* (ш.-б. *sirāw*, бадж. *sirīws*) ‘отделять, отбирать, выбирать; разнимать’; только в баджувском ‘отнимать от груди’.

(27) шугн.
Aslimo az biš bira-d-at tām mu bob mū-d
Аслимо EL грудь перестать сосать.F-PST-ADD тогда 1SG.O дедушка умирать-PST
‘Когда Аслимо бросила грудь, умер мой дед’.

(28) шугн.
wuz=um parwos xu puc as biš biru-d
1SG=1SG прошлый год REFL сын EL грудь отнимать от груди-PST
‘Я в прошлом году отняла ребенка от груди’.

Оба глагола *biridow* и *birêwtow* содержат элемент значения ‘в направлении от субъекта, прочь, от себя’, т. е. отражают значение древнего преверба *apa-*.

4.2. Непереходный глагол ш.-б. *biδīvdow*, бадж. *biδāvdow* (презентная основа ш.-б. *biδafc*) ‘закрываться (напр., о глазах, рте)’ и каузативный глагол от того же корня *biδemtow* (презентная основа *biδemb*) 1) ‘закрывать (глаза, рот)’; 2) ‘спать, засыпать (букв. закрывать свои глаза)’: из **upa-*, **abi-* и/или **pati-* и глагольного корня **daub-*, **dub-*, **dab-*, **da-m-b-*, **dap-* 1) ‘толкать, прибивать’; 2) ‘прикреплять; пришлепывать; примыкать, приставать, прилеплять(-ся)’; 3) ‘действовать, воздействовать’ (Расторгуева & Эдельман 2003: 398–402). Этимологически его связывают с глаголом без преверба *δīvdow/δůvdow* ‘собирать; рвать; клевать, склевывать’ и с глаголом с другим превербом *niδīvdow*, бадж. *niδāvdow* (презентная основа *niδafc*) ‘прилипать, приклеиваться, приставать’, ср. также каузативный глагол с тем же корнем *niδembtow/niδemtow* (презентная основа *niδemb*) ‘лепить, прилеплять, приклеивать’ (Расторгуева & Эдельман 2003: 399–400).

(Ранее не упоминавшийся преверб **abi-* мог быть наречием, предлогом, превербом или именным префиксом со значением (а) направления к чему-л., на что-л.; (б) предназначенности, усиления; также ‘вокруг’. Ни в каких других шугнанских глаголах его рефлексы нам не встретились, так что его присутствие в этом глаголе сомнительно; с другой стороны, преверб **pati-* в других шугнанских глаголах отражается как *pi-*.)

(29) шугн.
wi cem-en=en biδov-d
D3.M.SG.O глаз-PL=3PL закрываться.NPST.F.3SG
‘Глаза его закрылись’.

(30) шугн.
xu ɣêv biδemb xu tām aboz
REFL рот закрывать.IMP и тогда глотать.IMP
‘Закрой свой рот и затем глотай’.

(31) бадж.
nog xāb=ta vo xu cem biδem-t
хоть бы ночь=IRR быть REFL глаз смыкать-PST
‘Возможно, ночью он будет спать (букв. ‘сомкнет свои глаза’)’.

У этого глагола развилось значение ‘много говорить, болтать без умолку’, а также ‘много есть, есть без остановки’:

(32) шугн.
wam-and-ta wam γev ačaθ na-beδafc-t
D3.F.SG.O-LOC-but D3.F.SG.O рот вовсе NEG-закрывать-3SG
‘Она много болтает / много ест (*букв*. ‘у нее рот никогда не закрывается’) (пример из анкеты 2024)’.

Все примеры в словаре Карамшоева демонстрируют употребление этих глаголов только по отношению к парным элементам — губам или векам глаз, которые смыкаются друг с другом, т. е. движутся по направлению друг к другу; таким образом, с точки зрения семантики здесь более логично видеть следы преверба **upa-*. Для каузативного глагола *biδemtow* переход от значения ‘закрыть глаза’ к значению ‘засыпать’ кажется логичным и не меняет общей картины ‘смыкания/соединения’ парных частей (губ и век).

4.3. Шугнанский глагол *bizīdow* (презентная основа *bizin*) 1) ‘помещать (в загон, хлев и т. д.), загонять (скот, домашнюю птицу)’; 2) ‘арестовывать, заключать в тюрьму’; 3) только в шахдаринском ‘помещать; загонять; вносить, вводить (по отношению к одушевленным и неодушевленным предметам)’ из **upa-gan-/jan-* ‘бить, поражать, колотить, убивать’ (Расторгуева & Эдельман 2007: 140). В некоторых иранских языках этот глагол имеет также значение ‘класть, помещать’. Сохранился глагол без преверба *zintow* (презентная основа *zin*) ‘отнимать, отбирать, вырывать (из рук)’ и *zīdow* (*zīn*) ‘убивать, умерщвлять; губить, уничтожать; бить, лупить, ударять и др.’, а также глагол с тем же корнем, но другим превербом *piǰentow* (презентная основа *piǰen*) ‘продевать, нанизывать’.

(33) шх.
xu vorǰ=at γal na-bizīδ-ǰ=o?
REFL жеребец=2SG пока NEG-загонять-PF=Q
‘Ты еще не загнал свою лошадь [в конюшню]?’

По всей видимости, глагол *bizīdow* развил значение ‘загонять скот (или других домашних животных) внутрь стойла/дома’ из значения ‘бить, колотить’, употребляясь в ситуациях, когда конечной целью действия было ‘поместить животное или человека в огороженное пространство’, т. е. *гнать животное от себя,* побивая его, например, палкой/хворостиной или любым другим предметом. Тогда преверб в составе этого глагола, исходя из семантических соображений, было бы логично восстановить в виде **apa-*.

Значение ‘сажать в тюрьму’ является переносным от ‘загонять в хлев’ (человек уподобляется животному, которое надо ‘поместить в закрытое пространство или помещение, оградив от всех остальных людей’). По данным опроса, употребляясь по отношению к человеку, этот глагол приобретает уничижительную окраску; обычно так говорят о человеке с презрением, с выражением злорадства, приравнивая человека к скотине:

(34) шугн.
wi yi=n mis wev=qati bizī-d
D3.M.SG.O один=3PL тоже D3.PL.O=COM помещать-PST
‘Его тоже с ними посадили’ (пример из анкеты 2024).

4.4. Шугнанский глагол *birêx̌tow* (презентная основа *birêz,* PST *birōx̌t*) ‘пить’. Существует две версии образования этого глагола. Моргенстьерне возводит его к сочетанию преверба *apa-* или *upa-* с глаголом *rix̌tow* ‘лить(-ся), сыпать(-ся)’ (т. е. буквально ‘вливать в себя’, поскольку глагол *rix̌tow* двузалоговый); Д. И. Эдельман возводит его к сочетанию преверба **pari-* с глагольным корнем **aź-* ‘гнать, погонять, вести’ (т. е. буквально ‘вгонять, вносить’)

(Эдельман 2020: 435). Против версии Моргенстьерне говорит тот факт, что парадигмы глаголов *birêx̌tow* и *rix̌tow* не совпадают (презентная основа *rix̌tow* — *raz*, PST *rix̌t*); как отмечается в Арманд & Некушоева 2022, глагол *rix̌tow* используется только по отношению к сыпучим, но не жидким объектам. Однако версия Д. И. Эдельман тоже кажется не вполне надежной, поскольку ни в каких других случаях начальное праиранское *p-* не дает в шугнанском *b-* (в других глаголах преверб **pari-* дает преверб *par-*). Выше уже был рассмотрен глагол с другим превербом *arrix̌tow*/*arix̌tow*/*arrax̌tow*/*arax̌tow* (презентная основа *arāz*/*arrāz*) 1) ‘прыгать, подскакивать; тянуться (вверх); подниматься на дыбы’; 2) ‘бить ключом, струей; брызгать’; 3) перен. ‘чваниться, пыжиться, задирать нос, зазнаваться’.

(35) шугн.
dů́nd čoy=yām birox̌-t idi ačaθ na-bof-t
много чай=1PL пить-PST что совсем NEG-годиться-3SG
‘Мы столько пили чая, что просто невозможно’.

От этого глагола образовано существительное, имя деятеля *birezīǰ* ‘пьющий, пьяница’:

(36) шугн.
ero tu de lāk idi dis birezīǰ
ADR 2SG D2.M.SG.O оставление ведь так пьющий
‘Оставь его, он пьяница!’ (пример из анкеты 2024)

4.5. Шугнанский глагол *bīftow* (презентная основа *bāf*, 3SG *boft*) 1) ‘годиться; быть возможным; 2) быть в состоянии, мочь’ из **upa-ap-* с переходом **p* > *f* при регулярной предконсонантной позиции в 3-м л. ед.ч., где *ap-* = ‘доставать, достигать’ (Расторгуева & Эдельман 2000: 178–181). Д. И. Эдельман относит к этому же корню глагол с другим превербом *firīptow* ‘добираться, достигать, доходить и др.’, однако вряд ли он воспринимается носителями языка как однокоренной.

(37) шугн.
wāδ-ta mis nīstow na-bāf-en wi mala-ti
D3.PL-IN тоже пребывать NEG-мочь-3PL D3.M.SG.O дом-SUP
‘И они не могут жить в том доме’.

Модальный глагол *bīftow* этимологически содержит элемент значения ‘достигать чего-либо’, т. е. исторически преверб **upa-* можно понимать как предельность действия (‘двигаться по направлению к цели’) и как следствие — ‘достичь цели, иметь возможность что-то делать’.

4.6. Шугнанский глагол *bedow* (презентная основа *bes*, NPST.3SG *best*) 1) ‘исчезать, пропадать; теряться’; 2) ‘уходить; отправляться’; 3) *перен.* ‘умирать’. Этот глагол получился из сочетания преверба *apa-* и глагольного корня **ai-* : *i-* ‘идти, двигаться, перемещаться, приходить’ (Расторгуева & Эдельман 2000: 109, 111). По всей видимости, его исходным значением является ‘уходить прочь’; из него развилось значение ‘исчезать, пропадать; теряться’ и далее — переносное значение ‘умирать’.

Соответствующего глагола без преверба или с другим превербом нам не удалось найти. Возможно, в случае глагола *bedow* нам следовало бы говорить о сращении преверба и глагольного корня, поскольку преверб можно выделить только на основании этимологических данных и провести морфемную границу между превербом и глагольным корнем весьма проблематично.

(38) шугн.
soz-aθ=en *na-lův-d* *xu* *be-d-en*
песня-INT-3PL NEG-говорить-PST и уходить-PST-3PL
‘Они даже песен не пели и ушли’.

(39) шугн.
čib *mis* *yůri* *va-d* *yā* *mis* *be-d*
ложка тоже здесь быть-PST D3.F.SG ADD2 исчезать-PST
‘Ложка здесь была и исчезла’.

(40) бадж.
yu *cavor* *meθ=i* *tůў-d* *xu* *be-d*
D3.M.SG четыре день=3SG болеть-PST и умирать-PST
‘Он четыре месяца болел и умер’.

В глаголе *bedow*, по всей видимости, сохраняется значение древнего преверба ‘прочь’, т. е. ‘уйти прочь’ → ‘исчезнуть’; отсюда часто встречающийся в других языках семантический переход в ‘умереть’.

4.7. Шугнанский каузативный глагол *binêstow* (презентная основа *binês*) 1) ‘терять, девать’; 2) ‘тратить, расходовать’; 3) ‘губить, уничтожать, изводить’; 4) ‘гнать, выгонять, прогонять; удалять’; из **apa-nāś-aia-*, где *naś-* ‘терять(-ся), погибать; губить; пропадать, исчезать; скрывать(-ся)’ (Эдельман 2015: 487–490). Согласно проведенному анкетированию, этот глагол употребляется также в значении ‘убирать, прятать, отложить’.

(41) шугн.
wāδ=en *tar* *x̌ābi* *xu* *půnd* *binês-t*
D3.PL=3PL EQ ночная мгла REFL дорога терять-PST
‘Они заблудились в темноте (букв. потеряли свою дорогу)’.

(42) шугн.
māš *wi* *γiδā* *ca-rāng* *binês-ām*
1PL D3.M.SG.O мальчик как терять-1PL
‘Как нам извести этого юношу?’

(43) шугн.
uz-ta *tu* *binês-um* *azůd-and*
1SG-EMPH 2SG терять-1SG D1.EL-LOC
‘Я тебя выгоню отсюда!’ (пример из анкеты 2024)

(44) шугн.
xu *telefun* *benês*
REFL телефон терять.IMP
‘Убери свой телефон!’ (пример из анкеты 2024)

По всей видимости, в шугнанском сохранился только каузативный глагол, и судя по примерам из словаря Карамшоева, в непереходном значении ‘теряться, погибать, исчезать’ он не употребляется. В значении ‘терять, потратить, уничтожить’ присутствует элемент ‘прочь, насовсем’ (‘уничтожить’ или ‘потратить’ значит ‘окончательно потерять связь с некоторым объектом’), как и в примерах из анкеты с не отмеченными в словаре значениями ‘выгнать (прочь)’ и ‘убрать (прочь)’.

4.8. Шугнанский глагол *biγ̌entow*, бадж. *biwentow* (презентная основа шугн. *biγ̌en*, бадж. *biwen*) ‘трясти плодовые деревья (рукой, ногой, но не шестом)’ из **apa-* или *upa-(hi)-šān-aia-* или **apa-(hi)-šan-* с вторичной переогласовкой по каузативному типу, где *han-/-šan-* ‘трясти, стряхивать, сбрасывать; сбивать (плоды с дерева)’ (Расторгуева & Эдельман 2007: 357–358; Додыхудоева 1988: 122).

(45) шх.
dam *mūn* *biγ̌en*
D2.F.SG.O яблоко трясти плодовые деревья IMP.
‘Потряси ту яблоню’.

Глагол *biγ̌entow* имеет довольно ограниченную сферу употребления и не отличается многозначностью; в данном случае трудно определить, имеется ли в нем дополнительный элемент значения ‘прочь’ (*apa-*) или ‘по направлению к’ (*upa-*). Если считать, что действие ‘трясти’ подразумевает движение рукой или с помощью какого-то инструмента (например, палки) от себя и к себе (в разные стороны), то можно говорить, что здесь сохраняется преверб *upa-*, т. е. действие направлено от субъекта; если же здесь скорее семантика ‘трясти до тех, пока плод не упадет’ (этот глагол используется только по отношению к плодовым деревьям), т. е. ‘совсем сбить’, тогда преверб скорее восходит к *apa-*.

4.9. Каузативный глагол *bêx̌tow* (презентная основа *boz*, 3Sg *bêzd*, *bozd*) ‘посылать, отправлять, отсылать’ из **apa-aźa-aia-*, где *aź-* = ‘гнать, погонять, вести’ (Расторгуева & Эдельман 2000: 288), т. е. ‘отводить прочь, отгонять от себя’. При отсутствии в языке глагола без преверба или глагола с другим превербом провести границу между морфемами весьма затруднительно, тем более что в шугнанском языке имеется также, по всей видимости, заимствованный из таджикского (без преверба: кл. перс. *bāxtan*, презентная основа *bāz*, тадж. *бохтан*, презентная основа *боз* ← ср.-перс. *wāzīdan*, презентная основа *wāz*, см. Hassandoust 2016: 350) глагол *bêx̌tow* (презентная основа *boz*, 3Sg *bêzd*, *bozd*) ‘играть’, с совпадающими формами настоящего и прошедшего времени; эти два глагола отличаются только формами перфекта.

(46) шугн.
tar *mu-xez* *wam* *boz*
EQ 1SG.O-APUD D3.F.SG.O посылать.IMP
‘Пошли ее ко мне’.

Интерпретация преверба осложняется тем, что предлог *tar* используется в случае перемещения *к ориентиру* (на одном уровне с говорящим), а послелог *-xez* означает либо локализацию *рядом с* или *перед ориентиром* (‘перед, пред; при, у, около, возле; к’), либо личный локатив (*прийти к врачу*). Исторически здесь скорее преверб *upa-*, а не *apa-*, хотя глагол *bêx̌tow* также содержит направительный элемент значения ‘от, прочь от субъекта’.

4.10. Глагол *bix̌čīdow* (презентная основа *bix̌čār*, PST *bix̌čūd*) 1) ‘черпать, зачерпывать, разливать, насыпать (жидкости и сыпучие вещества)’; 2) ‘хватать, захватывать’ восходит к **upa-skar-a-* (Додыхудоева 1988: 122).

(47) шугн.
dam *sêr* *ar* *būǰīn* *bix̌ča*
D2.F.SG.OBL зерно[5] DOWN мешок сыпать.IMP
‘Насыпай то зерно в мешок’.

[5] Имеется в виду ‘куча зерна на току’.

Согласно проведенному анкетированию, у этого глагола развились переносные значения, не отмеченные в словаре Карамшоева, в первую очередь переносное значение ‘много зарабатывать’, аналогично разговорному русскому выражению ‘грести деньги лопатой’:

(48) шугн.
yu=yi taram kor ca čū-d pul=i bixčū-d
D3.M.SG=3SG там работа когда делать-PST деньги=3SG черпать-PST
‘Когда он там работал, он очень много зарабатывал (пример из анкеты 2024)’.

Также у этого глагола появилось значение *перен*. ‘накинуться, наброситься с руганью (обычно о женщине):

(49) шугн.
uz=um naw xu gāp sar čū-d-at ya=yi mu bixčū-d
1SG-1SG новый REFL слово начало делать-PST-ADD D3.F.SG=3SG 1SG.O черпать-PST
‘Я только начала говорить, как она на меня накинулась (пример из анкеты 2024)’.

Несмотря на неясную этимологию этого глагола, все же можно предположить, что в нем сохраняется направительное значение древнеиранского преверба **upa-* ‘по направлению к’ или **apa-*; тогда действие направлено от субъекта (от себя).

Таким образом, преверб *be-/bi-* восходит:
— к праиранскому превербу **apa-* в глаголах *bedow*, *binêstow*, *biridow*, *birêwtow ~ birêwdow*, *bizīdow* с сохранением направительной семантики ‘от, прочь’;
— предположительно к праиранскому превербу **upa-* в глаголах *bêxtow*, *bīftow*, *biδīvdow* (также с сохранением направительного значения ‘по направлению к’);
— про три глагола *birêxtow, bixčīdow, biγentow* нельзя с уверенностью утверждать, какой из двух превербов в них сохранился.

5. Глаголы с превербом *naǎ-, naў-, niǎ-*

Шугнанские превербы *naǎ-, naў-, niǎ-* из праиранского **niš-*, **niž-*, **nir-* наречие, предлог, преверб, именной префикс ‘прочь, наружу’. В позиции перед звонким согласным корня выступают рефлексы: шугн. *ў*, бадж. *w* (Эдельман 2015: 562, 563). Мы нашли пять глаголов с этими превербами: *naўǰīdow*, *naўʒimtow*, *naǎfidow*, *niǎfentow*, *naǎtīdow*, из них две пары некаузативного и каузативного глагола от одного корня. Для глаголов *naǎfidow* и *naǎfentow* имеется глагол с тем же корнем, но с другим превербом (*sifidow*); у глагола *naǎtīdow* имеется соответствующий глагол без преверба (*tīdow*) и глагол с тем же корнем и другим превербом (*andīdow*).

5.1. Шугнанско-баджувский вариант *naǎfidow*, шугн. *niǎfidow*, бадж. *narfidow* (презентная основа шугн.-бадж. *naǎfiθ*, шугн. *niǎfiθ*, бадж. *narfiθ*) ‘выдернуться; вывихнуться’ из **nis-*, **nir-* (или **ni-* + анлаут корня на **sf-* < **sph-*) + **fan-* : *fa-* / **(s)pan-* : *(s)pa-* ‘двигать(ся)’, часто о движении по вертикали; ‘подниматься; спускаться’, а также каузативный глагол ш.-б. *naǎfentow*, шугн. *niǎfentow*, бадж. *narfentow* (презентная основа ш.-б. вариант *naǎfen*, шугн. *niǎfen*, бадж. *narfen*) ‘выдергивать, вырывать, вытаскивать, отдирать’ без переогласовок, из кауз. **nišfānaḍa-* (Расторгуева & Эдельман 2007: 44–46). В других языках шугнано-рушанской группы (рушанский, хуфский, бартангский, сарыкольский) имеется глагол

с тем же превербом и тем же глагольным корнем. У глагола *naxfīdow* имеется пара — глагол с другим превербом, но с тем же глагольным корнем — *sifīdow.*

(50) шугн.
yu mêx as sitan-and naxf-īd
D3.M.SG гвоздь EL столб-LOC выдернуться-PST.M
'Тот гвоздь выскочил из столба'.

(51) шугн.
markāb xu mêx naxfen-t
осел REFL гвоздь выдергивать-PST
'Осел выдернул свою привязь (колышек)'.

В глаголе *naxfīdow/nixfīdow/narfīdow* преверб *nax-/nix-/nar-* сохраняет значение 'наружу, прочь'.

5.2. Шугнанский многозначный глагол *naxtīdow* (презентная основа *naxti*) 1) 'выходить (в разн. знач. извне, откуда-л., из чего-л.)'; 2) 'возникать, появляться, объявляться'; 3) 'оканчиваться, проходить, быть на исходе (о времени)' и др. (из **niš-tač-a(ya)-*, см. Додыхудоева 1988: 127) состоит из преверба *nax-* и сохранившегося в шугнанском языке глагола без преверба *tīdow* (презентная основа *ti*) 1) 'идти; ходить, ездить; отправляться, уезжать, улетать'; 2) 'идти, течь, струиться (о потоке и т. п.)' и др.; к этому глаголу также есть парный глагол с другим превербом — *andīdow.*

(52) шугн.
kū bīrīn xac tīz-d
гора SUB вода идти-3SG
'Под горой течет река'.

(53) шугн.
buor-and daraxt šīrā naxtīz-d
весна-LOC дерево сок выходить-3SG
'Весной сок сочится из дерева (смола)'.

Если глагол *tīdow* означает движение и перемещение вообще (о людях, времени, жидкостях и т. п.), а траектория движения задается предлогами и послелогами, то глагол *naxtīdow* имеет траекторию движения наружу, т. е. значение преверба 'наружу, прочь' сохраняется.

5.3. Шугнанский глагол *nayǰīdow* (презентная основа шугн. *nayǰīs*, бадж. *nayǰis*) 1) 'проходить; переходить (в разн. знач.); миновать'; 2) 'проходить, миновать (о времени)'; 3) 'обходить, превосходить' (в чем-л.) и каузативный глагол с тем же превербом и с тем же корнем *nayʒimtow* (презентная основа *nayʒimb/nayʒinb*) 1) 'проводить (о времени, мероприятии)'; 2) 'проводить, вести; уводить' восходят к **nir-gəsa-* : *nir-gata-* от корня **gam-* 'идти' (Расторгуева & Эдельман 2007: 117, 125).

(54) шугн.
yu mūysafed daδ as wi čūpůn nayǰīs-t xu tīz-d xu půnd-ti
D3.M.SG старик потом EL D3.M.SG.O чабан проходить-3SG и идти-3SG REFL дорога-SUP
'Тогда старик прошел мимо чабана и пошел своей дорогой'.

(55) бадж.
waz=um wi sitůr δo-d-δo-d xu
1SG=1SG D3.M.SG.O крупный рогатый скот давать-PST-давать-PST и
as Riɣ̌ůvjī=m wi naɣ̌ʒim-t
EL Ригувджи=1SG D3.M.SG.O проводить-PST
‘Я гнал-гнал скот и привел его за Ригувджи’.

(56) бадж.
yā šol-i δo-d qetarmā xu paded naɣ̌ǰā-d
D3.F.SG шаль-3SG давать-PST повязывание и D2.UP проходить-PST
‘Она обвязалась шалью и прошла вверх’.

Оба глагола имеют в модели управления предлог *as*, который придает глаголу направление движения, действия из какого-л. места, откуда-л. или указывает на исходный пункт, точку отсчета во времени, т. е. ‘из, с, от’, так что элемент значения ‘мимо, наружу’ в глаголе сохраняется.

Таким образом, в шугнанском сохраняется не так много глаголов с превербом *nax̌-*, *naɣ̌-*, *nix̌-*, однако наличие для трех глаголов из пяти парных лексем (либо глагола без преверба, либо глаголов с другими превербами) дает основание его выделять, тем более, что во всех сохранившихся глаголах присутствует общий элемент значения древнеиранского преверба **niš-*, **nіž-*, *nir-* ‘прочь, наружу’.

6. Глаголы с превербом *ni-/na-*

В шугнанском языке представлено порядка четырнадцати глаголов с превербом *ni-/na-*: *nīstow*, *nimêδdow/nimêstow*, *naraysentow*, *nixix̌tow*, *nixêx̌tow*, *niδembtow/niδemtow*, *niδīvdow/niδāvdow*, *niɣix̌tow*, *nimêwtow*. Некоторые из них являются заимствованиями из таджикского, например, *nivištow* ‘писать’ и *niwêzdow* ‘играть (на музыкальных инструментах)’. Для глагола *naraysentow* ‘расходовать, кончать (с чем-л.)’ (презентная основа *naraysen*) сохранился глагол с тем же корнем, но без преверба — *redow* ‘оставаться, сохраняться’ (в разных значениях; презентная основа *ris*), однако связь между ними не очевидна. У глагола *nixix̌tow* (презентная основа *nixarθ*) ‘разрушаться, обваливаться; спускаться; опускаться’ имеется соответствующий баджувский глагол от того же корня, с тем же значением, но с другим превербом *rixix̌tow* (презентная основа *rixarθ*); глагол *niδīvdow/niδāvdow* (презентная основа *niδafc*) этимологически связывают с глаголом без преверба *δīvdow/δůvdow* ‘собирать; рвать; клевать, склевывать’ и с упоминавшимся ранее глаголом с другим превербом *biδīvdow*.

Преверб *ni-/na-* из **ni-* — наречие, предлог, преверб и именной префикс ‘вниз, внутрь’. В более поздних языках в ряде случаев превербы и префиксы, продолжающие *ni-*, смыкаются с рефлексами **niš-*, [**nіž-*] / **nir-*, благодаря частичному семантическому сходству и фонетическим совпадениям (Эдельман 2015: 556), однако в шугнанском языке это два разных преверба, поскольку не происходит ни фонетического, ни семантического совпадения.

6.1. Шугнанский глагол *nīstow* (презентная основа *niθ*) 1) ‘сидеть; садиться, усаживаться’; 2) ‘ждать (сидя, стоя); поджидать’ (часто с предлогом *pi*); 3) ‘оставаться; пребывать, проживать, останавливаться (на постой)’; 4) ‘садиться, опускаться (с высоты), приземляться’; 5) ‘приминаться, уминаться’ из **ni-had-i̯a-* или **ni-hid-i̯a-* (**had-*, **-šad-* : *-zd-*, *-žd-*

‘сидеть, садиться; проживать, обитать, пребывать’). Соответствующий каузативный глагол: *nêδdow/nêstow* (презентная основа *nêδ*) 1) ‘сажать, усаживать (кого-л.)’; 2) ‘сажать (растения)’; 3) ‘ставить; устанавливать’; 4) ‘ставить (отметку)’; 5) ‘устраивать, пристраивать кого-л.’ из **ni-hād-a̯ia-* (Расторгуева & Эдельман 2007: 305, 309). В шугнанском языке нет соответствующего глагола без преверба или с другим превербом.

(57) шугн.
yā *beɣam-aθ* *nīθ-t*
D3.F.SG беззаботный-ADV сидеть-3SG
‘Она спокойно сидит’.

(58) шугн.
čorik-en=en *pi* *māš* *nos-č*
муж-PL=3PL UP 1PL сидеть-PF
‘Мужчины поджидают нас (букв. сидящие перед нами)’.

Этот пример интересен использованием предлога *pi*, который чаще имеет значение ‘вверх’, но также может маркировать адресата; в сочетании с глаголом *nīstow* он образует значение ‘ожидать, поджидать’.

(59) шугн.
piroki *mewā* *na-suδ-ǰ* *nêδ-ǰ*
раньше плод NEG-идти-PF сажать-PF
‘Раньше (здесь) не сажали плодовых (деревьев)’.

Основное лексическое значение этого глагола не изменилось с древнеиранского уровня; он зафиксирован в авестийском и в древнеперсидском, где преверб *ni-* придает действию направление движения вниз. Этот глагол сохранился в большинстве иранских языков; по всей видимости, это образование как минимум праиранского уровня.

6.2. Шугнанский каузативный глагол *nimêδdow/nimêstow* (презентная основа *nimêδ*) ‘квасить, готовить кислое молоко’ из **ni-mād-a̯ia-*, где **mad-* = ‘скисать, закисать; сбраживать(-ся), бродить, заквашивать(-ся); загустевать, коагулировать (о молоке)’; ‘вздуваться, подниматься, всходить (о тесте)’. Некаузативный глагол *nimoystow/nimostow/nimêstow* (презентная основа *nimoys/nimos*) ‘заквашиваться (о молоке)’ имеет основу настоящего времени из инхоативной основы **ni-mād-sa-* (Эдельман 2015: 100, 102). В шугнанском не засвидетельствован соответствующий глагол без преверба или с другим превербом, только именное образование *x̌ūmêδ* ‘молочная закваска’.

(60) шугн.
δu-ara *meθ* *kuzāy-en-and* *pāy* *nimêδ-en*
два-три день глиняный кувшин-PL-LOC кислое молоко квасить-3PL
‘Два-три дня в кувшинах квасят молоко’.

Определить значение преверба в этом глаголе довольно затруднительно. Вероятно, дело в том, что объем сквашенного (если слить сыворотку) уменьшается по сравнению с исходным, т. е. уровень в посудине становится ниже.

6.3. Шугнанский глагол *naraysentow* (презентная основа *naraysen*) ‘расходовать, кончать (с чем-л.)’ из **ni-rik-sa-* + поздний суффикс каузатива *-ēn*, где **raik-* : *rik-* = ‘оставаться, оставлять, покидать, отпускать, выпускать, освобождать’ (Эдельман 2020: 342–344). Шуг-

нанский глагол с тем же корнем, но без преверба и каузативного суффикса *redow* (презентная основа *ris*) имеет значения 1) ‘оставаться, сохраняться (в разных знач.)’; 2) ‘отставать, оставаться позади’; 3) ‘утомляться, уставать’; 4) (с предлогом *as*) ‘лишаться чего-л., оставаться без чего-л.; быть не в состоянии что-л. сделать’.

(61) шугн.
di *māš* *yoy̌ǰ=at* *naraysen-t*
D2.M.SG.O 1PL мука́=2SG расходовать-PST
‘Ты покончил с нашей мукой’.

Глагол с превербом развил новое значение ‘тратить, расходовать’; определить значения входящих в него частей по отдельности уже не представляется возможным.

6.4. Шугнанский глагол *nixix̌tow* (презентная основа *nixarθ*) 1) ‘разрушаться, обваливаться’; 2) ‘спускаться; опускаться’ и каузативный глагол *nixêx̌tow* (презентная основа *nixêx̌*/*nixêrθ*) ‘разрушать, обваливать’. Есть соответствующий баджувский глагол от того же корня, с тем же значением, но с другим превербом: *rixix̌tow* (презентная основа *rixarθ*). Этимология глагола неизвестна.

(62) шугн.
mu *zidůn* *nixax̌-t*
1SG.O кладовая разрушаться-PST
‘Моя кладовая обвалилась’.

(63) шугн.
wam *wêδ=en* *nixêx̌-č*
D3.F.SG.O арык=3PL разрушать-PF
‘Тот арык разрушили’.

Действие, обозначенное глаголом ‘разрушаться’, предполагает перемещение субъекта (при некаузативном глаголе) или объекта (при каузативном глаголе) вниз; поскольку разрушение цельного предмета предполагает его падение при потере целостности, можно говорить, что в глаголах *nixix̌tow* и *nixêx̌tow* присутствует элемент значения ‘вниз’.

6.5. Шугнанский глагол *niδīvdow*, бадж. *niδāvdow* (презентная основа *niδafc*) 1) ‘прилипать, приклеиваться, приставать’; 2) ‘приставать, задираться; сцепиться в схватке, драке’; 3) ‘начинать (что-л. делать), приступать к чему-л., приняться за что-л.’ и каузативный глагол с тем же корнем *niδembtow*/*niδemtow* (презентная основа *niδemb*) ‘лепить, прилеплять, приклеивать; стравливать, подбивать на драку’ восходит к сочетанию преверба **ni-* и глагольного корня **daub-* : *dub-*, **dab-*, **da-m-b-*, **dap-* 1) ‘толкать, прибивать’; 2) ‘прикреплять; пришлепывать; примыкать, приставать, прилеплять(-ся)’; 3) ‘действовать, воздействовать’ (Расторгуева & Эдельман 2003: 398). Этимологически этот глагол связывают с глаголом без преверба *δīvdow*/*δůvdow* ‘собирать; рвать; клевать, склевывать’ и с упоминавшимся ранее глаголом с другим превербом *biδīvdow* ‘закрываться (о глазах, рте и т. п.)’ (Расторгуева & Эдельман 2003: 399–400).

(64) шугн.
qoq *sit* *pi* *dewol* *na-niδafc-t*
сухой земля UP стена NEG-прилипать-3SG
‘Сухой ком земли к стене не прилипнет’ (погов. о недолговечности клеветы).

(65) шугн.
di garδā γal mā-niδemb
D2.M.SG.O лепешка пока PROH-лепить.IMP
'Не лепи пока лепешку [к стенам очага]'.

Глагол развил, по всей видимости, новое лексическое значение от 'толкать' к 'прилеплять к чему-либо' (в данном случае к стенке печи), развивая признак действия 'рукой от себя', однако однозначно определить направление действия в данном случае довольно затруднительно, поскольку в примере (64) и в других примерах в словаре Карамшоева при глаголах *niδīvdow* и *niδembtow* используется предлог *pi*, который в большинстве случаев имеет значение 'вверх', а не 'вниз'. Хотя в примере на значение 'стравливать, подбивать на драку' действие 'стравливать (куропаток друг с другом)' подразумевает общее направление вниз от субъекта действия, поскольку птицы находятся на земле, т. е. ниже деятеля:

(66) шугн.
pi mu-nd xu zarīʒ mā-niδemb
UP 1SG.O-LOC REFL куропатка PROH-лепить.IMP
'Не стравливай свою куропатку с моей[6]'.

При сравнении трех однокоренных глаголов *δīvdow — niδīvdow — biδīvdow* можно получить направительные значения: для *δīvdow* ('собирать, рвать, клевать') — 'к себе', для *niδīvdow* ('прилипать, приставать') — 'от, вверх или вниз'; для *biδīvdow* 'закрываться (о глазах, рте)' — 'друг к другу'.

6.6. Шугнанский глагол *niγixtow* (презентная основа шугн. *niγůў*, бадж. *niγaw*/*niγůw*) 'слушать, слышать' из **ni-gauša-* : *ni-gušta-*, где **gauš-* : *guš-* = 1) 'слышать, слушать'; 2) 'звучать, слышаться' (Расторгуева & Эдельман 2007: 247, 249). Соответствующий глагол без преверба или с другим превербом отсутствует. Подобное образование есть во многих иранских языках, как в западных, так и в восточных.

(67) бадж.
tu mu gāp xīč niγůў
2SG 1SG.O слово детальный слушать.IMP
'Ты внимательно слушай мои слова'.

Несмотря на то, что глагол сохранил свое значение, по всей видимости, с праиранского уровня, определить значение преверба *ni-* довольно затруднительно (возможно: *сиди и слушай* — тогда действие направлено вниз от говорящего к слушающему, однако возможно и другое расположение участников ситуации).

6.7. Шугнанский глагол *nimêydow*/*nimêytow* (презентная основа *nimoy*) 'виднеться; казаться' из **ni-mā*[1]*-*, где **mā-* : *mi-* = 'мерить, измерять, отмерять; соразмерять; приводить в порядок, в соразмерность, сооружать'. Рефлексы глаголов с превербом **ni-* в ряде иранских языков частично или полностью контаминированы с рефлексами аналогичных образований от корня **mā*[2]*-* 'делать знаки жестами; манить жестами; показывать(-ся), махать, качать(-ся)' (Эдельман 2015: 316, 321). Каузативный шугнанский глагол *nimêwtow* / ш.-б. *nimêwdow* / брв. *nimêydow* (презентная основа ш.-б. *nimêw*, брв. *nimêy*) 'подавать знаки; манить жестами' из **ni-mā*[2]*-* (Эдельман 2015: 331, 332).

[6] На Памире распространена традиция проведения куропаточьих боев. Для этого в некоторых хозяйствах специально выращивают куропаток.

(68) брв.
yu *čīd* *as* *dar-aθ* *nimêy-d*
D3.M.SG дом EL далекий-ADV виднеться-PST
‘Тот дом виднелся издалека’.

(69) шугн.
wuz=um *tar* *wi* *nimêw-t-at* *yu* *fām-t-i*
1SG=1SG EQ D3.M.SG.O подавать знаки-PST-ADD D3.M.SG понимать-PST-3SG
‘Я подал ему знак, и он понял’.

6.8. Шугнанский глагол *niwêzdow* (презентная основа *niwoz*) ‘играть (на музыкальных инструментах)’ из **ni=wāz=a=* < **ni=wač=* (Додыхудоева 1988: 129).

(70) шугн.
x̌āb=en *dišīd-ti* *δuyůn* *čorik* *rabob-tê* *niwêz-d-at* *soz-en* *lůvd*
ночь=3PL крыша-SUP два мужчина рубаб-SUP играть-PST-ADD песня-PL говорить-PST
‘Ночью на крыше двое мужчин играли на рубабах и пели песни’.

Употребление этого глагола в шугнанском у́же, чем в персидском и таджикском, где перс. *navāxtan* (презентная основа *navāz*) имеет значения 1) ‘играть (на музыкальном инструменте)’; 2) ‘бить, ударять; наносить удар’; 3) ‘ласкать, гладить, поглаживать’ (ср. также тадж. *навохтан* (презентная основа *навоз*) 1) ‘ласкать, нежить, лелеять’; 2) ‘играть на музыкальных инструментах’); однако во всех трех языках этот глагол используют по отношению к игре на струнных музыкальных инструментах (хотя в персидском сфера употребления этого глагола шире — он используется также для игры на барабанах, скрипке и т. п.), где предполагается движение рукой вниз (тадж. *дутор навохтан* ‘играть на дутаре’). Возможно, это старое заимствование в шугнанский, поскольку, во-первых, совпадают основы настоящего времени, а во-вторых, сфера употребления этого глагола в шугнанском по сравнению с персидским и таджикским сужена и специализирована; к тому же сами названия музыкальных инструментов в шугнанский язык заимствованы из персидского/таджикского и относятся к сфере культурной лексики.

Несмотря на то, что преверб *ni-*/*na-* выделяется на синхронном уровне, поскольку он сохраняется в целом ряде глаголов, образующих пары и даже тройки с однокоренными основами, вряд ли можно говорить о сохранности направительного значения ‘вниз’ во всех этих образованиях. По всей видимости, многие из этих глаголов были образованы еще в праиранском: так, глагол *nīstow* ‘сидеть’ имеет параллели во многих иранских языках, в том числе в древнеперсидском и авестийском (ав., др.-перс. *ni-šad-* ‘сидеть’, см. Расторгуева & Эдельман 2007: 306). В древних языках отмечаются и другие глаголы с превербом *ni-*, например, др.-перс. *ni-paiϴ-* ‘гравировать, надписывать’ (Эдельман 2020: 88); ав. *nī-kan-* ‘закапывать, зарывать’; др.-перс. *ni-kan-* ‘сносить, разрушать, уничтожать’ и др. (Шугнанский глагол *nivištow* ‘писать’ заимствован из таджикского и поэтому не был рассмотрен выше). Шугнанский материал показывает, что лишь немногие глаголы с превербом *ni-* сохраняют в своем значении направительный элемент ‘вниз’ — в первую очередь это *nixix̌tow* (презентная основа *nixarθ*) ‘разрушаться, обваливаться; спускаться; опускаться’ и каузативный глагол *nixêx̌tow* (презентная основа *nixêx̌*/*nixêrθ*) ‘разрушать, обваливать’. Остальные глаголы, как правило, развивают новые значения, и определить семантику преверба в них на синхронном уровне затруднительно.

Заключение

Проанализированный языковой материал показывает, что в шугнанском сохранился довольно значительный пласт глаголов с исконно иранскими превербами. Выделение этих превербов внутри глагольной словоформы на синхронном уровне оправдано тем, что внутри каждого из подмножеств глаголов, объединенных общим превербом, мы обнаруживаем хотя бы несколько основ, образующих пару либо с аналогичным глаголом без преверба, либо с другим превербом. При этом, однако, оснований утверждать о сохранности семантики древнеиранских превербов в шугнанских глаголах у нас нет, поскольку семантическая трансформация исходных сложных значений в большинстве случаев привела к их десемантизации.

Ситуация осложняется тем, что мы явно имеем дело с глаголами, образованными в разное время. Так, с высокой вероятностью можно утверждать, что шугнанский глагол *bīftow* ‘годиться; быть возможным; быть в состоянии, мочь’ был образован еще на индоиранском уровне ввиду др.-инд. параллели *irā́r-* ‘достигать, получать’; как в среднеиранских (согдийском и хорезмийском), так и в современных (таджикском, ягнобском, хуфском, бартангском, рушанском) иранских языках сохраняются глаголы, восходящие к праиранскому **ira-ar-* (Расторгуева & Эдельман 2000: 180).

Ниже в Таблице 1 показано, в каких иранских языках сохраняются глаголы, образованные таким же образом, что и рассмотренные нами шугнанские. В первом столбце дается реконструированная для праиранского состояния форма соответствующего глагола; во втором — (при наличии) глагольные формы авестийского и древнеперсидского, (с древнеиндийскими параллелями, если таковые известны); далее следуют перечни среднеиранских, (отдельно) современных западно- и восточноиранских, (отдельно) языков шугнано-рушанской группы и других памирских языков, в которых сохраняются рефлексы данных глаголов; наконец, в последнем столбце помещаются шугнанские глаголы в форме инфинитива. Информация в основном собрана по «Этимологическому словарю иранских языков»; для языков шугнано-рушанской группы также использованы словари из монографий В. С. Соколовой (1959, 1960) и Т. Н. Пахалиной (1971; электронные версии текстов словарей размещены на сайте pamiri.online/dict, см. Makarov, Melenchenko & Novokshanov 2022).

Таблица 1. Шугнанские глаголы с превербами и их корреляты в других иранских языках

№	праиранский	др.-ир.	ср.-ир.	зап.	вост.	шугн.-руш. + др. памирск.	шугн.
1	**ā-apa-aźa-aia-*	–	–	–	–	сар.	*abêxtow*
2	**ā-gār-aia-*	–	–	кл.перс.	–	–	*ažirtow*
3	**ā-raź-* или **fra-raź-*; также **ā-raik-*	–	ср.-перс., парф., согд., хор., х.-с.	–	осет.	–	*arrixtow*
4	**ā-hak-*	–	–	–	–	руш., хуф., барт., рош.	*aỹīdow* *agāntow* *aỹêzdow*
5	**han-kap-*	–	–	–	–	руш., барт., рош.	*anjīvdow*
6	**han-kuxsā-*/**han-kuxta-*	–	х.-с.	–	–	руш., хуф., барт., сар.	*angixtow*

№	праиранский	др.-ир.	ср.-ир.	зап.	вост.	шугн.-руш. + др. памирск.	шугн.
7	**han-tak-*	–	–	–	–	руш., барт., хуф., вах., сар.	*andīdow*
8	**ham-darb-*	др.-инд. *sáṁ-dr̥bdha-*	парф.	–	–	руш., хуф., барт., рош., язг., ишк.	*ancīvdow*
9	**ham-pad-s-i̯a-*	–	ср.-перс.	–	–	руш., язг., сар., ишк., вах.	бадж. *ambīθ/ambis*
10	**apa-rāu̯-* и **apa-rāu̯-a i̯a-*	–	–	–	–	–	*biridow birêwtow*
11	**upa-daub-* или **abi-daub-* или **pati-daub-*	–	согд.	–	ягн., пушту	барт., руш., хуф., сар.	*biδīvdow biδemtow*
12	**upa-gan-*	–	–	–	–	хуф., руш., барт., рош., сар.	*bizīdow*
13	**apa-raź-* или **api-raź-* или **upa-raź-*	–	–	–	–	руш., хуф., барт., сар., язг.	*birêx̌tow*
14	**upa-ap-*	др.-инд. *upā́p-*	согд., хор.	тадж.	ягн.	хуф., барт., руш.	*bīftow*
15	**apa-ai-*	ав. *apa-ay-*	парф., хор.	кл. перс.	пушту	сар., язг., ишк.	*bedow*
16	**apa-nāś-a i̯a-*	–	парф., согд., хор.		ягн.	руш., хуф., барт., рош., сар.	*binêstow*
17	**upa-(hi)-šān-a i̯a-* или **apa-(hi)-šan-a i̯a-*	–	–	–	–	хуф., руш., барт., рош., язг.	*biɣ̌entow*
18	**apa-aźa-a i̯a-* или **upa-aźa-a i̯a-*	ав. *avāzōit* ав. *upāzaiti*	–	–	пушту	мундж., хуф., руш., барт., рош.	*bêx̌tow*
19	**upa-skar-a-*	–	–	–	–	–	*bix̌čīdow*
20	**nis-fan-*	–	х.-с.	–	–	руш., хуф., барт., сар.	*nax̌fīdow*
21	**niš-tač-a(ya)-*	–	–	–	–	барт., хуф., руш.	*nax̌tīdow*
22	**nir-gəsa-/ *nir-gata-*	ав. *niǰasa-*, *ni-ǰāmaya-*	парф., хор., х.-с.		мдж.	руш., барт., сар.	*naɣ̌ǰīdow naɣ̌ʒimtow*
23	**ni-had-i̯a-* или **ni-hid-i̯a-*; **ni-hād-a i̯a-*	ав. *ni-šad-*; др.-перс. *ni-šad-* (кауз. *ni-šādaya-*)	ср.-перс., парф., согд., хор.	совр. перс., тадж., бел., заза, гил. и др.	ягн., пушту, мдж.	руш., хуф., барт., рош., сар., язг., ишк.	*nīstow nêδdow/ nêstow*
24	**ni-mād-a i̯a-*; **ni-mād-sa-*	–	хор.	–	–	сар.	*nimêδdow/ nimêstow; nimoystow*
25	**ni-rik-sa-*	–	–	–	–	–	*naraysentow*
26	**ni-daub-*	–	ср.-перс.	кл. перс.	–	руш., хуф., барт., рош., сгл., вах.	*niδīvdow niδembtow*
27	**ni-gauša-/ *ni-guš-ta-*	–	ср.-перс., парф., согд., хор.	кл. перс., совр. перс., тадж.	осет., пушту, мдж.	руш., барт., язг.	*niɣix̌tow*
28	**ni-mā-*	–	ср.-перс., парф., согд.	кл. перс., совр. перс., тадж. и др.	осет.	руш., хуф., барт., рош., сар., язг.	*nimêydow*
29	**ni-wač-*	ав. *navāzan*	–	кл.-перс., совр. перс., тадж. и др.	–	барт., руш., хуф.	*niwêzdow*

Из таблицы видно, что как минимум несколько шугнанских глаголов безусловно следует относить к праиранскому уровню. Таков, например, глагол *nīstow* (презентная основа *niθ*) ‘сидеть; садиться, усаживаться’, известный уже в авестийском и древнеперсидском, а также сохраняющийся во многих современных иранских языках, как западных, так и восточных — персидском, таджикском, дари, пушту, ягнобском, мунджанском, язгулямском и др. (Расторгуева & Эдельман 2007: 306–309). То же можно сказать и про шугнанский глагол *bedow* ‘исчезать; теряться; уходить, отправляться’, имеющий параллели в авестийском (*apa-ay-* ‘уходить, переходить’), в среднеиранских языках (парфянский, хорезмийский) и в современных языках (кл. персидский, пушту, язгулямский, ишкашимский, сарыкольский, см. Расторгуева & Эдельман 2000: 111). Всего таких глаголов насчитывается 6.

Десять глаголов из нашего списка зафиксированы в среднеиранских и в живых восточно- и западноиранских языках. Еще одну небольшую группу образуют глаголы, имеющие параллели только в восточноиранских языках. Так, глаголы *biδīvdow* и *biδemtow* имеют параллели в согдийском, ягнобском, пушту и в языках шугнано-рушанской группы (бартангском, рушанском, хуфском, сарыкольском); глаголы *nimêδdow*/*nimêstow* и *nimoystow* имеют параллели в хорезмийском и сарыкольском, глагол *naxˇfīdow* — в хотаносакском и в рушанском, хуфском, бартангском и сарыкольском.

В то же время мы находим глаголы с превербами, которые не имеют параллелей ни в древне-, ни в средне-, ни в новоиранских языках, кроме языков шугнано-рушанской группы; таков, например, глагол *naxˇfīdow*, шугн. *nixˇfīdow*, бадж. *narfīdow* (презентная основа шугн.-бадж. *naxˇfīθ*, шугн. *nixˇfīθ*, бадж. *narfīθ*) ‘выдернуться; вывихнуться’ — он имеет параллели только в рушанском (*nawfēdōw* ‘отрываться, выдергиваться’), в хуфском (каузативный глагол *nawfēntōw* ‘выдергивать, выламывать, отдирать’) и в бартангском (*nawfēθ-* : *nawföd* ‘выдергиваться, вываливаться, вылезать’, см. Расторгуева & Эдельман 2007: 46); или ср. шугнанский глагол *aγˇīdow* / бадж. *awīdow* (презентная основа *aγˇas* / бадж. *awas*) ‘ложиться, укладываться; никнуть к земле’, с параллелями в рушанском (основа прошедшего времени *awuyd*), хуфском (основа прошедшего времени *awūyd*), бартангском (*awayst*, основа прошедшего времени *awūyd* ‘лежать (о животных)’). Из разобранных нами порядка 40 глаголов 11 имеют параллели только в языках шугнано-рушанской группы; это дает основания считать, что на прашугнано-рушанском уровне префиксальная словообразовательная модель еще могла быть относительно продуктивной и, учитывая значительную взаимную близость этих языков, продержалась в этом статусе дольше, чем в других ветвях иранской группы.

При этом, однако, необходимо отметить, что большинство глаголов, как простых, так и с превербами, в шугнанском развивают новые лексические значения, и в некоторых случаях полностью утрачивают исходную семантику древнеиранского преверба. В единичных случаях мы даже наблюдаем радикальные изменения направительного значения преверба, как в случае с глаголом *andīdow* ‘вставать, подниматься; просыпаться, пробуждаться (ото сна); уходить, покидать, отходить, отлипать; увеличиваться, вырастать, подрастать, расти, прорастать, появляться (о всходах растений); начинаться, возникать’, для которого на синхронном уровне мы вынуждены выделять значение ‘вверх’, а не ‘совместность, соединение’, присущее отраженному в этом глаголе древнеиранскому превербу *ham-*.

В дальнейшем мы планируем подвергнуть аналогичному анализу данные других памирских языков, чтобы понять, специфичны ли подобные изменения только для шугнанского или же они демонстрируют аналогичные развития в близкородственных языках.

Список сокращений и глоссы

ав. — авестийский; бадж. — баджувский; барт. — бартангский; бел. — белуджский; брв. — барвазский; вах. — ваханский; гил. — гилянский; др.-инд. — древнеиндийский; др.-перс. — древнеперсидский; кл. — классический; мдж. — мунджанский; осет. — осетинский; парф. — парфянский; перс. — персидский; рош. — рошорвский; руш. — рушанский; сар. — сарыкольский; сгл. — сангличский; совр. — современный; согд. — согдийский; ср.-перс. — среднеперсидский; хор. — хорезмийский; х.-с. — хотаносакский; хуф. — хуфский; шугн. — шугнанский; ш.-б. — шугнано-баджувский; шх. — шахдаринский; ягн. — ягнобский; язг. — язгулямский.

ADD — аддитивная частица и союз; ADV — адвербиализатор; APUD — локализация рядом с или перед ориентиром (также личный локатив); COORD2 — аддитивная частица; D1 — указательное местоимение 1-й серии (Ich-Deixis); D2 — указательное местоимение 2-й серии (Du-Deixis, Der-Deixis); D3 — указательное местоимение 3-й серии (Jener- Deixis); DIM — диминутив; DOWN — перемещение к ориентиру ниже говорящего; EL — элатив; EQ — эквалатив, перемещение к ориентиру на одном уровне с говорящим; EZ — изафет (тадж.); F — женский род; IMP — императив; INF — инфинитив; INT — интенсификатор; IRR — будущее, хабитуалис; LAT — латив; LOC — локатив; M — мужской род; NEG — общий показатель отрицания; NPST — непрошедшее время; O — косвенный падеж; PF — перфект; PL — множественное число; PROH — отрицание императива; PST — прошедшее время; Q — вопросительная частица; REFL — рефлексив; SG — единственное число; SUBD — подчинительный союз; SUP — суперэссив; UP — перемещение к ориентиру выше говорящего; 1 — 1-е лицо; 2 — 2-е лицо; 3 — 3-е лицо.

Литература

Арманд, Е. Е. 2005. Семантика древнеперсидских превербов. В кн.: И. С. Смирнов (ред.). *Аспекты компаративистики 1. Orientalia et Classica. Труды Института восточных культур и античности VI*: 135–161. Москва: Издательство РГГУ.

Арманд, Е. Е., Ш. С. Некушоева. 2022. Глаголы ‘лить(ся)’ и ‘сыпать(ся)’ в шугнанском языке. *Родной язык* 1: 73–109.

Бахтибеков, Тӯпчӣ. 1979. *Грамматикаи забони шугнонӣ*. Душанбе: Дониш.

Гамкрелидзе, Т. В., Вяч. Вс. Иванов. 1984. *Индоевропейский язык и индоевропейцы. Том 1*. Тбилиси: Издательство Тбилисского университета.

Додыхудоева, Л. Р. 1988. *Шугнанский глагол в историческом освещении*. Душанбе: Дониш.

Карамшоев, Д. К. 1963. *Лаҳҷаи баджуви забони шугнанӣ*. Душанбе: Академия наук Таджикской ССР.

Карамшоев, Д. К. 1988. *Шугнанско-русский словарь. Том 1*. Москва: Восточная литература.

Карамшоев, Д. К. 1991. *Шугнанско-русский словарь. Том 2*. Москва: Восточная литература.

Карамшоев, Д. К. 1999. *Шугнанско-русский словарь. Том 3*. Москва: Восточная литература.

Кронгауз, М. А. 1998. *Приставки и глаголы в русском языке: семантическая грамматика*. Москва: Языки русской культуры.

Пахалина, Т. Н. 1971. *Сарыкольско-русский словарь*. Москва: Наука.

Расторгуева, В. С., Д. И. Эдельман. 2000. *Этимологический словарь иранских языков. Том 1.* Москва: Восточная литература.

Расторгуева, В. С., Д. И. Эдельман. 2003. *Этимологический словарь иранских языков. Том 2.* Москва: Восточная литература.

Расторгуева, В. С., Д. И. Эдельман. 2007. *Этимологический словарь иранских языков. Том 3.* Москва: Восточная литература.

Соколова, В. С. 1959. *Рушанские и хуфские тексты и словарь*. Москва / Ленинград: Изд-во Акад. наук СССР.

Соколова, В. С. 1960. *Бартангские тексты и словарь*, Москва / Ленинград: Изд-во Акад. наук СССР.

Эдельман, Д. И. 2011. *Этимологический словарь иранских языков. Том 4*. Москва: Восточная литература.

Эдельман, Д. И. 2015. *Этимологический словарь иранских языков. Том 5*. Москва: Наука / Восточная литература.

Эдельман, Д. И. 2020. *Этимологический словарь иранских языков. Том 6*. Москва: Наука / Восточная литература.

Эдельман, Д. И., Ш. П. Юсуфбеков. 1999. Шугнанский язык. В кн.: В. С. Расторгуева (ред.). *Языки мира. Иранские языки III. Восточноиранские языки*: 225–242. Москва: Индрик.

References

Armand, Elena E. 2005. Semantika drevnepersidskix preverbov. In.: I. S. Smirnov (ed.). *Aspekty komparativistiki 1. Orientalia et Classica. Trudy Instituta vostochnyh kul'tur i antichnosti VI*: 135–161. Moskva: Izdatel'stvo RGGU.

Armand, Elena E., Sh. S. Nekushoeva. 2022. Glagoly 'lit'(sja)' i 'sypat'(sja)' v shugnanskom jazyke. *Rodnoj jazyk* 1: 73–109.

Bahtibekov, Tūpchī. 1979. *Grammatikai zaboni shugnonī*. Dushanbe: Donish.

Dodykhudoeva, Leila R. 1988. *Shugnanskij glagol v istoricheskom osveshchenii*. Dushanbe: Donish.

Edel'man, Dzhoy I. 2011. *Etimologicheskij slovar' iranskix jazykov. Tom 4*. Moskva: Vostochnaja literatura.

Edel'man, Dzhoy I. 2015. *Etimologicheskij slovar' iranskix jazykov. Tom 5*. Moskva: Nauka / Vostochnaja literatura.

Edel'man, Dzhoy I. 2020. *Etimologicheskij slovar' iranskix jazykov. Tom 6*. Moskva: Nauka / Vostochnaja literatura.

Edel'man, Dzhoy I., Sh. P. Jusufbekov. 1999. Shugnanskij jazyk. In: V. S. Rastorgueva (ed.). *Jazyki mira. Iranskije jazyki III. Vostochnoiranskije jazyki*: 225–242. Moskva: Indrik.

Edelman, D. (Joy) I., Leila R. Dodykhudoeva. 2009. The Pamir languages. In: Gernot Windfuhr (ed.). *The Iranian languages*: 773–786. London / New York: Routledge.

Gamkrelidze, Tamaz V., Vjach. Vs. Ivanov. 1984. *Indoevropejskij jazyk i indoevropejcy. Tom 1*. Tbilisi: Izdatel'stvo Tbilisskogo universiteta.

Hassandoust, Mohammad. 2016. *Farhang-e rišešenāxti-ye zabān-e fārsi. Jeld-e 1–5* [*An Etymological Dictionary of the Persian Language*]. Tehrān: The Academy of Persian Language and Literature.

Karamshoev, Dodkhudo K. 1963. *Laxchai badzhuvi zaboni shugnanī*. Dushanbe: Akademija nauk Tadzhikskoj SSR.

Karamshoev, Dodkhudo K. 1988. *Shugnansko-russkij slovar'. Tom 1*. Moskva: Vostochnaja literatura.

Karamshoev, Dodkhudo K. 1991. *Shugnansko-russkij slovar'. Tom 2*. Moskva: Vostochnaja literatura.

Karamshoev, Dodkhudo K. 1999. *Shugnansko-russkij slovar'. Tom 3*. Moskva: Vostochnaja literatura.

Kent, Roland G. 1953. *Old Persian: Grammar, Texts, Lexicon*. New Haven, Connecticut: American Oriental Society.

Krongauz, Maxim A. 1998. *Pristavki i glagoly v russkom jazyke: semanticheskaja grammatika*. Moskva: Jazyki russkoj kul'tury.

Makarov, Yury, Maksim Melenchenko, Dmitry Novokshanov. 2022. Digital Resources for the Shughni Language. In: Atul Kr. Ojha, Sina Ahmadi, Chao-Hong Liu, John P. McCrae (eds.). *Proceedings of the Workshop on Resources and Technologies for Indigenous, Endangered and Lesser-Resourced Languages in Eurasia within the 13th Language Resources and Evaluation Conference*: 61–64. Marseille: European Language Resources Association.

Pakhalina, Tatiana N. 1971. *Sarykol'sko-russkij slovar'*. Moskva: Nauka.

Rastorgueva, Vera S., Dzhoy I. Edel'man. 2000. *Etimologicheskij slovar' iranskix jazykov. Tom 1*. Moskva: Vostochnaja literatura.

Rastorgueva, Vera S., Dzhoy I. Edel'man. 2003. *Etimologicheskij slovar' iranskix jazykov. Tom 2*. Moskva: Vostochnaja literatura.

Rastorgueva, Vera S., Dzhoy I. Edel'man. 2007. *Etimologicheskij slovar' iranskix jazykov. Tom 3*. Moskva: Vostochnaja literatura.

Sokolova, Valentina S. 1959. *Rushanskije i hufskije teksty i slovar'*. Moskva / Leningrad: Izd-vo Akademii nauk SSSR.

Sokolova, Valentina S. 1960. *Bartangskije teksty i slovar'*, Moskva / Leningrad: Izd-vo Akademii nauk SSSR.

Wendtland, Antje. 2009. The position of the Pamir languages within East Iranian. *Orientalia Suecana* LVIII: 172–188.

Elena Armand. Reflexes of Old Iranian preverbs in th Shughni language

In this paper, I present a descriptive and diachronic analysis of a set of Shughni verbs which include preverbs of Common Iranian origin (*a-*, *an-*/*im-*/*am-*, *bar-*, *bi-*/*be-*/*b-*, *naǩ-*/*naɣ̌-*/*niǩ-*, *ni-*/*na-*). The main objective of the study is to determine the quantity, grammatical status, and semantic features of these preverbs in the modern Shughni verb system, and to understand how these features correlate with the semantics of the corresponding preverbs at the Old and/or Proto-Iranian level. My analysis shows that some Shughni verbs formed by means of preverbs actually go back to the Common Iranian state; at the same time, a significant portion of such verbs have parallels only in the languages of the Shughni-Rushani group and are not represented in other Iranian languages, which may indirectly indicate a certain

period of productivity of preverbal word formation after the collapse of the Iranian linguistic unity. The analysis is accompanied by examples of synchronic usage and etymological information. The findings help clarify some controversial issues related to verbal word formation in the history of Iranian languages.

Keywords: Iranian languages; Proto-Iranian; Shughni-Rushani languages; Shughni language; preverbs; verb formation.

А. О. Бадеев
НИУ ВШЭ / Институт языкознания РАН, Москва; badeev@iling-ran.ru

Указательные местоимения в древнеиранских языках: основные типы употребления и реконструкция праиранской системы[1]

В статье рассматривается употребление указательных местоимений в древнеиранских языках. В соответствии с дейктической семантикой указательных местоимений определяются границы супплетивных парадигм и серий. Впервые на древнеиранском материале устанавливается соответствие каждой парадигмы местоимений с определенным типом дейксиса. На основе изученных типов употребления в древнеиранских языках, данных других иранских языков, а также в сопоставлении с древнеиндийским реконструируется праиранская система демонстративов.

Ключевые слова: иранские языки; древнеиранские языки; авестийский язык; древнеперсидский язык; дейксис; демонстративы; указательные местоимения; праиранский язык.

Настоящая статья представляет собой исследование дейксиса и указательных местоимений (демонстративов) в древнеиранских языках. В основе своей оно имеет анализ различных контекстов употребления указательных местоимений в авестийском и древнеперсидском языках, а также сопоставление двух текстов авестийской традиции с опорой на представления о дейксисе и демонстративах в лингвистической типологии. Для достижения этого необходимо: 1) определение супплетивных парадигм местоимений (также доказательство того, что элементы внутри них кореферентны и принимают значения, характерные для местоимений данной парадигмы); 2) определение значений, свойственных для каждой парадигмы; 3) распределение местоимений каждой из парадигм по трем базовым дейктическим значениям — то есть ближнему, среднему и дальнему дейксису.

Раздел 1 обозревает две письменные традиции на древнеиранских языках: авестийскую и древнеперсидскую, фокусируясь на генетико-диалектологическом статусе авестийского. В разделе 2 приводятся сведения о древнеиранских указательных местоимениях, почерпнутые из существующих описаний, а также об используемых в работе подходах к дейксису и демонстративах. В разделе 3 приводится реконструкция основ указательных местоимений в древнеиранских языках, обсуждаются особенности супплетивизма, приводится пример парадигмы. В подразделах 3.1–3.4 даны примеры употребления рефлексов этих основ в авестийском и древнеперсидском языках, приводится соответствие каждой группы местоимений с определенным типом дейксиса, уточняются сведения об актуальных парадигмах указательных местоимений. Раздел 4 представляет собой анализ употребления указательных местоимений в двух текстах авестийской традиции, Ардвисур-Яште и Хадохт-Наске. В разделе 5 представлена реконструкция праиранской системы указательных местоимений на основе изученных типов употребления в древнеиранских языках, данных других иранских языков, а также в сопоставлении с древнеиндийским. Выводы обозначены в разделе 6. В приложении даны парадигмы для всех серий указательных местоимений в древнеиранских языках.

[1] Исследование осуществлено в рамках Программы фундаментальных исследований НИУ ВШЭ.

1. Древнеиранские языки. Хронология. Генетико-диалектологический статус

Начало древнеиранской языковой эпохи принято отсчитывать от разделения индоиранской общности на индоарийскую и собственно иранскую ветви (не позднее 2000 г. до н.э.), конец же соотносится с IV — III вв. до н. э., временем составления последних текстов на древнеперсидском языке (Оранский 1979: 63; Skjaervo 2009: 43). К древнеиранским языкам хронологически можно отнести авестийский, древнеперсидский, мидийский и скифский. До нас дошли древнеперсидские и авестийские тексты, в то время как сведения о мидийском и скифском языках крайне ограничены и сводятся в первую очередь к ономастике, а также ряду мидийских заимствований в древнеперсидском. В дальнейшем речь пойдет об особенностях древнеперсидского и авестийского, так как мидийский и скифский языковой материал не позволяет сделать выводы о системе демонстративов (указательных местоимений) в этих языках.

Древнеперсидский язык относится к юго-западной группе иранских языков, в то время как принадлежность авестийского языка к той или иной иранской подгруппе установить затруднительно. Это связано с наличием в нем признаков, характерных как для западноиранских, так и для восточноиранских языков. Сложность заключается также в том, что авестийский язык делится на два ключевых языковых слоя: староавестийский и младоавестийский, генетико-диалектологическое соответствие между которыми также является предметом споров.

В частности, А. Панаино в своей работе высказывает сомнения насчёт того, что староавестийский и младоавестийский — это две стадии одного того же языка, как некогда считалось (Panaino 2007). Несомненно, староавестийский диалект в целом более архаичен, а младоавестийский демонстрирует больше инноваций. Однако отдельные элементы морфологии (парадигм склонений имен, глагольной флексии) в младоавестийском более архаичны и сходны с теми, что можно наблюдать в древнеиндийском; в староавестийском же, наоборот, в этих местах наблюдаются инновации. Эти изоглоссы позволяют говорить о том, что младоавестийский язык может иметь своим предком некую языковую форму, близкородственную староавестийскому, но имевшую ряд различий (Panaino 2007: 24). Вместе с тем по данным в первую очередь фонетики, а также морфологии и синтаксиса подтверждается типологическая близость младоавестийского и древнеперсидского языков (в отличие от староавестийского) (Skjaervo 2009: 44).

1.1. Авестийская традиция. Тексты. Диалектология

На староавестийском или диалекте Гат записаны собственно Гаты (Y.28-34, Y.43-51, Y.53), Ясна Хаптанхаити или «Ясна Семиглавая» (Y.35-41), а также разнообразные цитирования, вступительные и завершающие пассажи в Младшей Авесте. Гаты — «Песнопения» имеют четко выраженную метрику и по мнению многих исследователей являются проповедями в поэтической форме, созданными самим Заратуштрой. Между текстами Гат «вложена» Ясна Хаптанхаити, составленная в древнейшей поэтической форме, отличной от поэтической формы Гат и имеющей ритмические и риторические параллели в италийских языках, см. Watkins 1995: 232–240. Тексты Младшей Авесты содержатся в следующих книгах: Ясна, Висперед, Видевдад, Яшты, Хорде Авеста, и в отдельных фрагментах (Хадохт-наск, Нирангистан и др.). Между младоавестийскими текстами Ясны также «вложены» Гаты.

Считается, что на староавестийском говорили на востоке Иранского нагорья в середине — начале II тысячелетия до н.э. (Malandra 1983: 16; Skjaervo 2009: 46; Martínez & de

Vaan 2014: 1). Это время приблизительно соответствует времени создания Ригведы, язык которой сближают по уровню архаичности с диалектом Гат. Тогда же были составлены в первоначальном виде Гаты и Ясна Хаптанхаити. При этом младоавестийский был разговорным языком в начале — середине I тысячелетия до н.э. В то же время фиксируются древнеперсидский, мидийский и скифский языки. Предположительно, тогда и создается канон Авесты в устной форме, а Гаты Заратуштры становятся частью большой Авесты, составленной на младоавестийском языке (Skjaervo 2009: 46).

Таким образом, Авеста передавалась исключительно устно, пока в VI веке н.э. не был составлен письменный корпус авестийских текстов. Несомненно, существовало несколько школ, передававших Авесту, в том числе и на письме, ср. согдийскую запись молитвы Ашем Воху в работе Cantera 2022: 222. С записью Авесты в VI веке традиция устной передачи не была прервана, и разные школы продолжили существовать. И хотя после исламизации Ирана значительная часть манускриптов была утрачена, синтез устной и письменной практик сохранил корпус авестийских текстов в том виде, в котором мы имеем их сегодня, см. Cantera 2022: 223–227. При этом древнейшие сохранившиеся рукописи Авесты датируются XIII в. Влияния передатчиков и переписчиков, говоривших на других языках, а также обычные ошибки, накопившиеся за это время, несомненно присутствуют в священном тексте.

А. Панаино подвергает сомнению единое происхождение авестийской традиции, приводя определенные историко-культурные аргументы. Во-первых, религиозные идеи, содержащиеся в текстах Гат — более инновационные, чем отдельные тексты, написанные на младоавестийском языке. Гимны Хаоме и Тиштрии, сказание о Йиме и т. п. являются частью более древней устной традиции, восходящей к праиндоиранскому уровню (Panaino 2007: 25). Стоит привести в пример хотя бы Йиму, который, обладая ролью культурного героя и царя в Младшей Авесте, присутствует также и в индуистском пантеоне (Яма), и в религиозных представлениях нуристанцев-кафиров в виде верховного бога и культурного героя (Имра — Мара — Ямараджа). Несомненно, образ Йимы сложился ещё до распада индоиранской общности, а по мнению отдельных исследователей пантеон во главе с Йимой — общеарийское наследие в религиозной системе кафиров (Йеттмар 1986: 121).

Наконец, если между жизнью Заратуштры и возникновением устных текстов на младоавестийском разница в четыре века, неясно, как зороастрийские жрецы могли так долго довольствоваться скудными и трудно сопоставимыми между собой староавестийскими текстами Гат и Ясны Хаптанхаити (Panaino 2007: 26). Сохранившиеся староавестийские фрагменты представляют, вероятно, лишь самую малую часть той древнейшей традиции. Действительно, книги Авесты состоят из разнообразных фрагментов, при этом значительный вклад в это разнообразие вносит та устная традиция, которая легла в основу Младшей Авесты.

1.2. Древнеперсидская письменная традиция

Древнеперсидский представлен одним диалектом — это язык клинописных надписей ахеменидских царей, создание которых можно датировать относительно четко: 522 — около 350 гг. до н. э. Преимущественно это надписи Дария I (522 — 486 гг. до н. э.) и Ксеркса I (486 — 465 гг. до н. э.) (Kent 1950: 6; Skjaervo 2009: 43). Доподлинно неизвестно, насколько язык клинописных надписей соответствовал разговорному языку как ахеменидской элиты, так и их персоязычных подданных. Та же форма языка, которая впервые была зафиксирована при Дарии, уже могла быть непонятной для его ближайших пре-

емников. В поздних надписях фиксируются многочисленные ошибки, расхождения с нормой текстов VI в. до н. э. (Skjaervo 2009: 47). Язык персов поздней империи Ахеменидов, зафиксированный в надписях, оставленных потомками Ксеркса I, принято называть поздним древнеперсидским. В них можно усмотреть также перестройку системы именного склонения, выход из употребления форм имперфекта, отпадение конечных безударных гласных основы и прочие изменения, характеризующие его переходное состояние по отношению к среднеперсидскому (de Vaan & Lubotsky 2012: 207).

2. Указательные местоимения (демонстративы) в древнеиранских языках. История изучения. Представления о дейксисе и демонстративах

Изучение древнеиранских указательных местоимений и их дейктических значений началось с изысканий сравнительно-исторических лингвистов, которые выделяли древнеиранские лексемы вместе с их когнатами в других индоевропейских языках, сравнивали их значение и употребление. Реконструируемые праиндоевропейские корни, отраженные в указательных местоимениях языков индоевропейской семьи, К. Бругманн предложил соотносить с четырьмя способами выражения дейктических значений: ich-deixis, du-deixis, der-deixis и jener-deixis[2] (Brugmann 1904). Эти способы отражают различные виды указания на референта. Так, к ich-deixis относится указание на область в пространстве, ассоциируемую с говорящим, к du-deixis — на область, которая понимается говорящим как домен адресата. Тип der-deixis понимался как слабо-дейктическое или анафорическое указание на референта, индифферентное по отношению к положению референта в пространстве. Впоследствии в лингвистической науке значения типа der-deixis были отнесены к широкому классу эндофорических значений, не являющихся собственно дейктическими, см. Diessel 1999: 113. Последний тип, jener-deixis включал указание на удаленный объект, находящийся «по ту сторону», то есть исключаемый из области как говорящего, так и адресата.

Здесь нельзя не отметить, что изначально дейктические типы применялись для описания индоевропейских местоименных морфем и прочно ассоциировались с ними, однако уже в более поздних работах им стало придаваться значение языковых универсалий: «дейктические слова были бы неспособны передавать данные логические функции, если бы у них не было с самого начала соответствующих внутренних предпосылок» (Bühler 1934).

Представлениями о дейксисе как об универсальной категории руководствуется и настоящая работа. В частности, здесь приняты обозначения дейктических значений, основанные на выводах, сделанных Э. Ландером и Л. Хагеман (Lander & Haegeman 2018). Выделяется три серии указательных местоимений со значениями: 1) близкий к говорящему, 2) близкий к адресату, 3) удаленный от говорящего и адресата. Для упрощения эти значения именуются также близким, средним и дальним, как принято в большинстве типологических описаний. Вместе с тем постулируется синкретизм значений (Lander & Haegeman 2018: 123). Это означает, что в одном местоимении может совмещаться два и более из них. Например, синкретическое местоимение X, ко-

[2] Ср. hic- и istic-deixis вместо ich- и du-deixis в работе Я. Ваккернагеля (Wackernagel 2010). Это уточнение требовалось для того, чтобы подчеркнуть: указание производится в отношении некой области, соотносимой с одним из локуторов, а не напрямую на говорящего/адресата, как можно было бы предположить, исходя из обозначения, предложенного Бругманом.

торое сочетает значения 1 и 2, указывает на область говорящего и/или адресата в зависимости от контекста; а местоимение, которое сочетает все три значения, является нейтральным, его употребление не зависит от степени удаленности референта от дейктического центра. Иначе говоря, как бы ни было много (или мало) морфологических средств для выражения удаленности в языке, в строго дейктическом смысле они будут сводиться к трем базовым значениям или их комбинациям.

В настоящей работе также рассматривается употребление указательных местоимений в слабо-дейктических или эндофорических контекстах. Под этим объединяющим названием понимается анафорическое, дискурс-дейктическое и рекогнитивное употребление. Все три употребления рассматриваются как производные от собственно дейктического или экзофорического употребления и являются связующим звеном в процессе грамматикализации указательных местоимений (Diessel 1999: 110). Референт анафорического дейктического местоимения — именная группа, упомянутая в предыдущем фрагменте текста. Так анафорические указательные местоимения связывают адресата с предыдущим дискурсом. Дискурс-дейксис также отсылает адресата к предыдущему дискурсу, но с той разницей, что референт дискурс-дейктического указательного местоимения не упоминается в тексте напрямую. При этом референтом может быть фрагмент текста, или некая ситуация, которая описывалась ранее, но не называлась. Наконец, рекогнитивное указательное местоимение, наоборот, не отсылает к предыдущему дискурсу, его референт — новый для дискурса, однако предполагается, что знание о нем уже есть у адресата. Рекогнитивное употребление маркирует ситуацию, в которой нарочито активируется это знание адресата (Diessel 1999: 109). К последнему близко примыкает также отслеживающий тип, используемый для того, чтобы сконцентрировать внимание собеседника на референте (Visser 2020: 105).

Большое внимание уделено древнеиранским, в первую очередь авестийским указательным местоимениям в «Древнеиранском словаре» Х. Бартоломе и «Учебнике авестийского» Г. Райхельта, ученика Х. Бартоломе (Bartholomae 1904; Reichelt 1909). Для каждого местоимения приводятся примеры его употребления в различных контекстах, выделяются его дейктические функции соответственно представлениям о дейксисе в ту эпоху. Г. Райхельт пользуется терминологией К. Бругмана и соотносит указательные местоимения авестийского языка с каждым из четырех типов дейксиса (Reichelt 1909: 279-280). В более поздних описаниях не предпринимались попытки уточнить значения множественных указательных местоимений древнеиранских языков. Сложилась традиция разделять их в грамматиках на три серии: две «ближних» (без строгого разграничения между ними основ и функций) и одну «дальнюю» (Соколов 1961; Соколов 1979; Виноградова 2000) или не разделять на серии вовсе (Johnson 1917; Kent 1950; Beekes 1988; Kellens 1991; Martínez & de Vaan 2014). Таким образом, авестийский и древнеперсидский дейксис и указательные местоимения не получили должного внимания и нуждаются в подробном изучении.

Методология настоящей работы включает в себя анализ широкого спектра контекстов, в которых употребляются демонстративы древнеиранских языков. Однако в первую очередь мы считаем необходимым остановиться на их дейктическом употреблении. Как уже упоминалось ранее, экзофорические значения являются первичными по отношению эндофорическим — дискурсивным. Орфография авестийских примеров в работе приведена в соответствии с таковой в базе текстов индоевропейских языков TITUS (Thesaurus Indogermanischer Text- und Sprachmaterialien, https://titus.uni-frankfurt.de/texte/etcs/iran/airan/avesta/avest.htm). Древнеперсидский языковой материал в статье соответствует представленному в монографии Schmitt 2009, в которой древнеперсидские клино-

писные надписи даны на языке оригинала в транскрипции с переводом на немецкий язык. Перевод примеров основывается преимущественно на значениях, приведенных в той же монографии, а также в словаре Schmitt 2014.

3. Реконструкция основ указательных местоимений. Типы супплетивизма. Словоизменение

Указательные местоимения в древнеиранских языках возводятся к следующим праиранским основам[3] (Эдельман 1990; Dunkel 2014):

1. **a-* < и.-е. **e-* или **h_1e-*.

2. **aḭ-* : **i-* (< и.-е. **eḭ-* : **i-*) связаны чередованием, восходящим к праиндоевропейскому, ср. реконструированную парадигму и.-е. **h_1e-* : **i-* в работе Beekes 2011: 229; от этого корня образованы инновационные основы **ima-* (< и.-е. **i-mó-* с добавлением форманта **-m-*), также основа женского рода **iḭ-* < и.-е. **i-h_2-*.

3. **anā* < и.-е. **e nó-h_1*, образовано при помощи сложения основы указательного местоимения **e-* и дейктической частицы **-no-*. Таким же образом образовано др.-инд. указательное местоимение ближней серии в инстр.п. *enā́*/*anéna* (Kupfer 2002: 49). Связано со ст.-слав. *onъ* ‘тот’, лит. *añs*, *anás* ‘тот’, тем не менее, степень родства неясна. Реконструкция ст.-слав. и лит. местоимений из и.-е. указательного местоимения **áno-* (родственно упомянутой выше частице), приводимая Г. Дункелем, оспаривается за неимением точных когнатов в других индоевропейских языках (Dunkel 2014: 55; Peyrot 2017). В древнеиранских языках встречается только в инстр.п. ед.ч. мужского рода. Это ограничение на употребление не имеет до сих пор обстоятельного объяснения. Ср., возможно, окончание тематических основ, а также указательных местоимений др.-инд. ед.ч. инстр.п. *-ena*, которое вытесняет старый инструментальный падеж на *-ā* (Debrunner & Wackernagel 1930: 34, 498). Аналогично в хотаносакском ед.ч. инстр.п. *-äna*, *-ina*. Р. Э. Эммерик полагает, что такое развитие в индоиранских языках связано с распространением окончаний *n*-склонения на тематическое, ср. х.-с. им.п. *tcei’mä* : инстр.п. *tcei’mäna* (ав. *čašman-*); им.п. *balysä* : инстр.п. *balysäna* (х.-с. *balysa-*) (Emmerick 1968: 258). Основы **a-*, **aḭ-*, **iḭ-*, **ima-*, **anā* образуют супплетивную парадигму в авестийском и древнеперсидском.

4. **ha-* : **ta-* (< и.-е. **so-* : **to-*) также связаны чередованием, в результате которого образуют супплетивную парадигму в древнеиранских языках. Традиционно супплетивизм этих основ сводится к праиндоевропейскому разграничению одушевленности/неодушевленности, активности/неактивности, утраченному в древнеиранских языках (Семереньи 1980: 219; Эдельман 1990: 57). Сходные паттерны супплетивизма свойственны для основ всех серий указательных местоимений в древнеиранских языках. Они будут рассмотрены подробнее в этом же разделе.

5. **aiša-* : **aita-* происходят из сложения двух и.-е. местоименных основ **eḭ-* + **só-* (с внутренним сандхи), **eḭ-* + **tó-*. Аналогично основам **ha-*, **ta-* они образуют супплетивную парадигму.

6. **hāu-*, **hu̯a-*[4] происходят из сложения и.-е. местоименной основы **só-* и дейктической частицы **u*. Рефлексы и.-е. **u* встречаются в ряде указательных местоимений славянских и арийских языков (ср. др.-инд. *asaú*), где они реализуются как некий дейктиче-

[3] Для упрощения парадигмы ниже перечислены основы мужского и среднего родов, основы женского рода синхронно образуются преимущественно тривиально при помощи удлинения конечного гласного основы.

[4] Только в староавестийском, иначе предлагается читать как **hau̯*, см. (Martínez & de Vaan 2014: 20).

ский элемент (Мейе 1938: 333). В древнеиранских языках рефлексы **hāu-*, **hu̯a-* встречаются только в форме именительного падежа единственного числа.

7. **au̯a-* также является результатом сложения и.-е. местоименной основы **e-* и дейктической частицы **éu̯* (полная ступень от **u*), то есть содержит тот же дейктический элемент **u* и объединяется в супплетивную парадигму с **hāu-*, **hu̯a-* в древнеиранских языках (ср. также ст.-слав. *ovъ*).

Древнеиранские указательные местоимения демонстрируют сразу несколько типов супплетивизма. Первый тип является наиболее распространенным и наследуется непосредственно из праиндоевропейского. В древнеиранских языках он служит морфологическим средством для обозначения родовых различий. Для древнеиранских указательных местоимений, как и для имен, выделяются три рода: мужской, средний и женский. Категория рода здесь традиционно выражена в основе. Так, основы женского рода отличаются от основ мужского и среднего родов преимущественно удлинением конечного гласного основы[5]: ср. ав. м., ср.р. *aēta-*, ж.р. *aētā-*. В то же время основы среднего рода отличаются от основ мужского и женского родов отсутствием супплетивной основы в номинативе, см. таблицы 1, 2.

Таблица 1. Супплетивизм основ указательных местоимений в младоавестийском языке, ср. Ефимов 1975: 107.

Падеж	Мужской род	Средний род	Женский род
Им. п.	***aēm***	*imaṯ*	***īm***
Вин. п.	*imәm*	*imaṯ*	*imąm*
Им. п.	***hō***, ***hā***	*taṯ*	***hā***
Вин. п.	*tәm*	*taṯ*	*tąm*
Им. п.	***aēša***, ***aēšō***	*aētaṯ*	***aēšā***
Вин. п.	*aētәm*	*aētaṯ*	*aētąm*
Им. п.	***hāu***, *aom*	*auuaṯ*	***hāu***
Вин. п.	*aom*	*auuaṯ*	*auuąm*

Для древнеперсидского сходным образом:

Таблица 2. Супплетивизм основ указательных местоимений в древнеперсидском языке, ср. Brandenstein & Mayrhoffer 1964: 67–68.

Падеж	Мужской род	Средний род	Женский род
Им. п.	***iyam***	*ima*	***iyam***
Вин. п.	*imam*	*ima*	*imām*
Им. п.	***hau̯***	*ava*, *avaš-ci*	***hau̯***
Вин. п.	*avam*	*ava*, *avaš-ci*	*avām*

Основы, выделенные выше полужирным, встречаются только в им.п. ед.ч. мужского и женского родов, в то время как все остальные позиции в парадигме занимают другие основы. Поскольку супплетивизм такого типа своими корнями уходит в праиндоевро-

[5] В системе гласных авестийского языка *ə* < **a*, *ą* < **ā* перед носовыми согласными, окончание им.п. *-ō* < **-ah* < **-as*.

пейский, объяснение его следует искать в грамматике этого языка. Вероятно, что праиндоевропейский был языком типа subject pro-drop, то есть языком, в котором местоименный субъект в простом предложении опускался (Lehmann 1974: 40). Чтобы заполнить пустующую позицию местоименного субъекта, привлекалась супплетивная основа. Существуют анатолийские параллели, которые позволяют предположить, что имена среднего рода в праиндоевропейском не могли занимать позицию субъекта переходного глагола, и потому заменялись лексическими синонимами: ср. **wodor* ‘вода’ (ср.р.) и **xap-* (*h₂ep-*) ‘(текущая) вода’ (ж.р.); **nebhos* ‘(облачное) ‘небо’ (ср.р.) и **dyēu-* ‘небо’ (м.р.) (Yakubovich 2021: 139). Исходя из данных древнеиранских, а также и других индоевропейских языков, можно заключить, что праиндоевропейские местоимения среднего рода образовывали одну падежную форму на месте форм именительного и винительного падежей в мужском-женском родах, ср. Plank 2016. По-видимому, в среднем роде существовала альтернативная стратегия кодирования актантов, в которой субъект непереходного глагола и объект переходного маркировались одним падежом. Ср. местоименные энклитики хеттского языка, которые могут занимать позиции непереходного субъекта и переходного объекта, но не допускаются в позиции переходного субъекта (Yakubovich 2021: 137).

В им.п. мн.ч. супплетивные основы, по-видимому, отсутствовали, см. Beekes 2011: 229; Kapović 2017: 86. В древнеиранских языках формы указательных местоимений всех трех родов во мн.ч. им.-вин.п. образуются от одной основы. Кроме того, во всех падежах кроме им.-вин. формы среднего рода совпадают с мужским. Таким образом, собственно формы среднего рода представлены только в им.-вин.п.

В основе второго типа супплетивизма лежит разделение на прямые (им.-вин.) и косвенные падежи. В прямых падежах (кроме мужского и женского рода в им.п. ед.ч.) употребляются *ima-* основы, в то время как в косвенных — *a-* основы. Второй тип полностью доминирует во мн.ч. Наконец, третий тип заключается в том, что отдельная форма встраивается в парадигму в инстр.п. ед.ч. мужского рода. Это *ana* в парадигме местоимения *aē-/ima-/ī-/a-/ana*; также *tā* в парадигме *ha-/ta-/a-* (см. таблицы 6, 7 в приложении). Отсутствие формы *ana* в парадигме местоимения *ha-/ta-/a-* при наличии в ней форм от основы *a-* служит аргументом в пользу того, что *ana* является независимой основой и не может быть воспринята как *a-* + *na* окончание инстр.п. Два последних паттерна супплетивизма являются, по-видимому, общей индоиранской инновацией, ср. парадигму др.-инд. указательного местоимения *a-/i-/ena-* (Kupfer 2002: 49).

В древнеиранских языках противопоставляются формы единственного, двойственного и множественного числа. Число выражено флективным способом, при этом для местоимений характерна флексия, отличная от именной. Ср., например, форманты **-hm-*, **-hi̯-* < и.-е. **-sm-*, **-si̯-* в склонении основ дательного, отложительного, местного падежей (Эдельман 1990: 59–60; Beekes 2011: 227). Двойственное число имеет наиболее бедную падежную парадигму, что является одной из предпосылок исчезновения данной категории в среднеиранский период (Молчанова 1975: 205). Так, в авестийском двойственное число засвидетельствовано для основ указательных местоимений *a-*, *ima-*, *ta-* в трех падежных формах (в ед.ч. у указательных местоимений выделяют семь падежей). Указательные местоимения в древнеперсидском не имеют форм двойственного числа.

3.1. Рефлексы **a-*, **ai̯-*, **ii̯-*, **ima-*, **anā*

Рефлексы основ **a-*, **ai̯-*, **ii̯-*, **ima-*, **anā* засвидетельствованы в качестве основ указательных местоимений в древнеперсидском и авестийском языках. В древнеперсидском

рефлексы **a-*, **anā* отражены в формах *ahyāyā* (местный, также родительный-дательный падежи) и *anā* (инструментальный-отложительный падежи), рефлексы **i̯-* — в формах именительного падежа мужского и женского рода. Формы, образованные от **ima-*, засвидетельствованы для всех остальных падежей. В свою очередь, в авестийском возводимые к основам **ai̯-*, **i̯i̯-*, **ima-* местоимения закреплены за именительным и винительным падежами. Рефлексы основы **a-* встречаются во всех падежах, кроме именительного и винительного. Рефлексы **anā* встречаются только в инструментальном падеже. Иначе в диалекте Гат, где рефлексы **ima-* не представлены, в отличие от местоимений, образованных от основ **a-*, **ai̯-*, **i̯i̯-*, **anā* (Beekes 1988: 139).

Многочисленные примеры, где указательные местоимения **a-*, **ai̯-*, **i̯i̯-*, **ima-*, **anā* кореферентны, то есть разные местоимения отсылают к одному объекту, подтверждают необходимость объединения этих форм в единую супплетивную парадигму в древнеперсидском и авестийском языках (ср. также объединение рефлексов **ha-*, **ta-*, **a-* в парадигму в подразделе 3.2). Ср. пример (1) из авестийского, где употребление рефлексов **a-*, **ima-*, **anā* свидетельствует о таковой кореферентности. В качестве референтов видим *zaoϑra-* ‘возлияние’ и *barəsma-* ‘Барсман’. Интересующие нас рефлексы указательных местоимений и их перевод здесь и в дальнейшем выделены в тексте полужирным.

(1) ***ahmiia*** *zaoϑre* ***imaṯ*** *barəsma* *āiiese* *yešti*
D.N.SG.LOC libation.SG.LOC D.N.SG.ACC Baresman.SG.ACC get.PRS.1SG.MID worship.INF
ana *barəsmana* ***imam*** *zaoϑrąm* *āiiese* *yešti*
D.N.SG.INS Baresman.SG.INS D.N.SG.ACC libation.SG.ACC get.PRS.1SG.MID worship.INF
‘В **этом** возлиянии **этот** Барсман заполучаю поклонением, **этим** Барсманом **это** возлияние заполучаю поклонением’ (Y.2.1).

Аналогично в древнеперсидском (2), где в качестве референта указательных местоимений *ahyāyā* и *imam* видим *dipi-* ‘надпись’.

(2) *ava* ***ahyāyā*** *dipiyā* *nai̯* *nipištam* *avahya* *=rādī* *nai̯*
D.N.SG.NOM D.F.SG.LOC inscription.SG.LOC NEG write.PPRT.N.SG.NOM D.N.SG.DAT =RSN NEG
nipištam *mā* *taya* *haya* *aparam* ***imām*** *dipim*
write.PPRT.N.SG.NOM NEG REL.N.SG.NOM REL.M.SG.NOM afterward D.F.SG.ACC inscription.SG.ACC
patipr̥sāti *avahyā* *paru* *ϑadayāti* *taya* *=manā* *kr̥tam*
read.SUBJ.3SG.ACT D.SG.DAT much seem.SUBJ.3SG.ACT REL.N.SG.NOM =1SG.GEN do.PPRT.N.SG.NOM
nai̯ *=šim* *ima* *vr̥navātai̯* *duruxtam* *maniyātai̯*
NEG =3SG.M.ACC D.N.SG.ACC convince.SUBJ.3SG.MID lie.PPRT.N.SG.NOM believe.SUBJ.3SG.MID
‘То не было написано в **этой** надписи; для того (затем) не было написано, чтобы, кто потом прочтет **эту** надпись, тому не показалось, что мною сделано слишком много, чтобы (не получилось так, что) его это не убедит (и он) сочтет, что я солгал’ (DB § 58).

За этой группой указательных местоимений прочно закреплено значение ближнего дейксиса. В контекстах близости к говорящему, а также в контекстах противопоставлений близкое ≠ далекое употребляются местоимения этой группы. Ср. примеры (3, 4) из авестийского языка и пример (5) из древнеперсидского. Во всех этих примерах при указании на удаленные объекты всегда употребляется местоимение ав. *auua-*, др.-перс. *ava-*. Отметим факт противопоставления *aē-/ima-/ī-/a-/ana* ≠ *auua-*.

(3) *niuuaēδaiiemi* *haṇkāraiiemi* <…> ***aiŋhåsca*** *zəmō* ***auuaiŋheca*** *aṣ̌nō*
announce.PRS.1SG.ACT carry out.PRS.1SG.ACT D.F.SG.GEN earth.SG.GEN D.M.SG.GEN heaven.SG.GEN
‘Я возвещаю (и) провожу (эту Ясну) ... для **этой** земли, для **тех** небес’ (Y.1.16).

(4) *upa kuua =ciṯ* ***aiŋhå̊*** *zəmō* <…> *upa* ***auue***
upon where =EMPH D.F.SG.GEN earth.SG.GEN upon D.M.PL.NOM
starō yōi haptōiriŋga
star.PL.NOM REL.M.PL.NOM Haptoiringa.PL.NOM
‘Во всяком месте **этой** земли … где **те** звезды Хапторинг …’. (Yt.12.22, 28).

(5) *baga vazr̥ka auramazdā haya avam asmānam*
god.SG.NOM great.M.SG.NOM Ahuramazda.SG.NOM REL.M.SG.NOM D.M.SG.ACC sky.SG.ACC
adā haya ***imām*** *būmīm adā*
create.IMF.3SG.ACT REL.M.SG.NOM D.F.SG.ACC earth.SG.ACC create.IMF.3SG.ACT
‘Ахурамазда, великий бог, который создал **то** небо, который создал **эту** землю’ (DZc § 1).

Отдельно необходимо отметить употребление указательных местоимений этой серии в перформативных контекстах. Перформативное высказывание предполагает равноценность высказывания действию. Перформативное высказывание не является строго дейктическим, его можно отнести к эндофорическому дейксису. Пример такого эндофорического употребления указательного местоимения приводит в своей работе Э. Кёниг:

(6) *I* ***hereby*** *promise you never to touch hard drinks again*
‘Настоящим я обещаю тебе никогда больше не прикасаться к крепким напиткам’ (König 2020: 38).

В примере (6) указательное местоимение английского языка *hereby* маркирует перформативный контекст. В таких контекстах в авестийском языке употребляется местоимение *ana*; возможно, и другие местоимения ближней серии следует интерпретировать таким образом (ср. примеры 7, 8). Ср. также перформативные высказывания с указательными местоимениями в русском языке: *Сим я провозглашаю мир!*; *Это — священные слова*.

(7) ***ana*** *buiiå̊ zauuanō.sāsta*
D.N.SG.INS become.OPT.2SG.ACT invoke.NOM.SG.F
ana *buiiå̊ huiiaštatara*
D.N.SG.INS become.OPT.2SG.ACT good praise.PPRT.CMP.NOM.SG.F
‘Да будешь ты **сим** призванной, да будешь ты **сим** самой почтенной’ (Yt.5.9).

(8) ***imå̊sə*** *tē haoma gāϑå̊* <…>
D.F.PL.NOM 2SG.GEN Haoma.SG.VOC Gatha.PL.NOM
ime *həŋti aršuxδa vācō*
D.M.PL.NOM COP.PRS.3PL.ACT straight spoken.M.PL.NOM word.PL.NOM
‘**Се —** твои Гаты, о Хаома ... **се —** правдивые слова’ (Y.10.18).

Следует упомянуть и об употреблении местоимений, образованных от основы *a-* в анафорическом значении, а также производном от него значении личного местоимения 3-го лица в авестийском. В существующих описаниях упоминается притяжательное значение форм ед.ч., род.п. м.р. *ahe*, *aiŋhe* (Martínez & de Vaan 2014: 72). Помимо этого, образованное от основы *a-* местоимение 3-го лица употребляется и в других формах, как в примере (9). Ср. также примеры (11, 12).

(9) *spəŋtaṯ haca maniiaoṯ zaraϑuštra* ***aēšąm***
holy.M.SG.ABL from spirit.SG.ABL Zarathustra.SG.VOC D.M.PL.GEN
ciϑrəm vahištāaṯ =ca manaŋhaṯ
appearance.PL.NOM best.N.SG.ABL =CONJ thought.SG.ABL
О творец! Где пребывают умерших души, Фраваши святые? ‘У святого духа, о Заратуштра, **их** проявления, у лучшей мысли’ (H.2.40).

3.2. Рефлексы **ha-*, **ta-*

В древнеперсидском языке непосредственно от основ **ha-*, **ta-* не образуется указательных местоимений. Однако от этих же основ образуются путем сложения относительные местоимения *haya-/taya-* < и.-е. **só h_2i̯o-/*tó h_2i̯o-* (Dunkel 2014: 738). При этом сохраняется супплетивная парадигма, см. таблицу 3.

Таблица 3. Относительные местоимения древнеперсидского языка им.-вин.п. (Brandenstein & Mayrhoffer 1964: 71).

Падеж	Мужской род	Средний род	Женский род
Им. п.	***haya***	*taya*	***hayā***
Вин. п.	*tayam*	*taya*	*tayām*

В авестийском рефлексы **ha-*, **ta-* засвидетельствованы в качестве указательных местоимений. Как в младоавестийском, так и в диалекте Гат они встречаются в формах именительного, винительного и инструментального падежей (Beekes 1988: 140). Указательные местоимения *ha-*, *ta-* употребляются в авестийском языке в анафорических контекстах, ср. примеры (10, 11). Они также засвидетельствованы в контекстах катафоры, как в примере (12). Вероятно, следует видеть в примере (12) также дискурсивную (временную) анафору.

(10) *frōiṯ =mē* ***taṯ*** *nąma* *framrūiδi*
well =1SG.GEN D.N.SG.ACC name.SG.ACC speak.IMP.2SG.ACT
aṣ̌āum *ahura* *mazda*
righteous.M.SG.VOC Ahura.SG.VOC Mazda.SG.VOC
«Наше имя, о Спитама Заратуштра … оно самое действенное»: ‘Промолви мне **это** имя, о праведный Ахура Мазда’ (Yt.1.5).

(11) *tūm* ***tā*** ***aēibiiō*** *xṣ̌aiiamna*
2SG.NOM D.N.PL.ACC D.M.PL.DAT possess.MPRT.NOM.SG.F
nisirinauuāhi *arəduuī* *sūre* *anāhite*
give over.SUBJ.2SG.ACT Ardvi.SG.VOC Sura.SG.VOC Anahita.SG.VOC
«Девы с чревом бесплодным, жаждущие господина, будут просить у тебя сильного мужа; женщины, находящиеся на пороге рождения, будут просить у тебя хороших родов»: ‘Ты все **это** им, имея, воздашь, о Ардви Сура Анахита!’ (Yt.5.87).

(12) ***hā*** ***ahmāi*** *aṣ̌iš* *ərənāuui* ***taṯ*** ***ahmāi***
D.F.SG.NOM D.M.SG.DAT reward.SG.NOM get.INJ.3SG.PASS D.N.SG.NOM D.M.SG.DAT
jasaṯ *āiiaptəm* *yaṯ* *=hē* *puϑrō* *us.zaiiata*
come.INJ.3SG.ACT boon.SG.NOM REL.N.SG.NOM =3SG.GEN son.SG.NOM born.INJ.3SG.MID
‘**Такая ему** награда досталась, **такой к нему** пришел дар: у него родился сын’ (Y.9.7).

Анафорические указательные местоимения *ha-*, *ta-* грамматикализовались также в значении личного местоимения 3-го лица. Личные местоимения 3-го лица в древнеиранских языках представлены только в энклитической форме в косвенных падежах, см. парадигму для авестийского языка (Виноградова 2000: 25). Полная форма личного местоимения 3-го лица соответствует формам указательных местоимений *ha-*, *ta-* как в примере (13) или форме указательного местоимения *a-* (в тех падежах, где у *ha-*, *ta-* отсутствует косвенная форма), как в примерах (11, 12). Таким образом, можно говорить об

ещё одной супплетивной парадигме анафорического местоимения (= личное местоимение 3-го лица) *ha-*/*ta-*/*a-*, где формы именительного, винительного и инструментального падежей местоимений *ha-*, *ta-* дополняются формами, образованными от *a-* в остальных косвенных падежах, см. таблицу 7 в приложении. Ср. также примеры употребления *a-* в подразделе 3.1. О грамматикализации анафорического местоимения в личное местоимение 3-го лица см. (Diessel 1999: 119–120).

(13) *sraošō* *aṣ̌iiō* <…> ***hō***
Sraosha.SG.NOM holy.M.SG.NOM D.M.SG.NOM
vərəθraja *drujəm* *jaγništō*
victorious.M.SG.NOM Druj.SG.ACC smashing.SUP.M.SG.NOM
‘Святой Сраоша … **он** побеждающий, разбивающий Друдж’ (Yt.11.3).

Г. Райхельт указывает на употребление *ha-*, *ta-* в контексте ближнего дейксиса, см. примеры (14, 15) (Reichelt 1909: 283). Действительно, в примере (14) встречается противопоставление *ha-* ≠ *hāu*. В то же время мы не склонны видеть в этом единичном случае экзофорическое употребление *hō*, так как в тексте присутствует очевидная анафора. В V.22.1 (пример 15) *taṯ nmānəm* на первый взгляд употребляется не анафорически, ведь с этого фрагмента начинается текст. Однако тот же фрагмент V.22.1 повторяется еще в двух местах в тексте (V.22.8, 22.14), и там уже анафора присутствует — в предыдущих перед этими двумя стихами видим *nmānəm* в (V.22.7, 22.13). Значит, V.22.1 (15) — это вырванный из контекста фрагмент, фактически ставший формулой, которая по недоразумению оказалась на первом месте в тексте.

(14) ***hāu*** ***=ca*** *aiiå* *narå* *vohu* *manō*
D.SG.NOM =CONJ D.M.DU.GEN man.DU.GEN good.N.SG.ACC thought.SG.ACC
jāgərəbuštarō *aŋhaṯ* *yō* *gəuš* *uruθβarə*
grab.APRT.CMP.SG.NOM COP.SUBJ.3SG.ACT REL.M.SG.NOM cow.SG.GEN abdomen.SG.ACC
hąm.pāfrāiti *yaθa* ***hāu*** *yō* *nōiṯ* *iθa* ***hō***
fill.SUBJ.3SG.ACT REL.ADV D.SG.NOM REL.M.SG.NOM NEG thus D.M.SG.NOM
aspərənō.mazō *upamərətō* ***hāu*** *aspərənō.mazō*
worth of Asperena.SG.NOM almost dead.M.SG.NOM D.SG.NOM worth of Asperena.SG.NOM
‘**Тот** из **этих** двух человек лучше обретает Воху Мано, кто наполняет живот (мясом) быка, чем **тот**, кто не делает этого; **он** — практически мертв; **тот** же (выше его на) величину Асперены’ (V.4.48).

(15) *azəm* *yō* *ahurō* *mazdå*
1SG.NOM REL.M.SG.NOM Ahura.SG.NOM Mazda.SG.NOM
<…> *yasə* ***taṯ*** *nmānəm* *ākərənəm*
REL.M.SG.NOM D.N.SG.ACC house.SG.ACC create.IMF.1SG.ACT
‘Я Ахура Мазда, … сотворил **этот** дом’ (V.22.1)

Ср. также единичный пример, где *taṯ* появляется в контексте указания на объект близкий к адресату (16). Ср. с примерами (17, 18, 19), где в контексте среднего дейксиса употребляется указательное местоимение *aēša-*/*aēta-*. Нехарактерное употребление такого рода можно объяснить совпадением нескольких контекстов, ведь налицо также пример анафоры. Вероятно, и здесь *taṯ* употребляется в анафорическом значении.

(16) *xṣ̌aθrəm* *ahurāi* *cinasti*
kingdom.SG.ACC Ahura.SG.DAT acknowledge.PRS.3SG.ACT

taṯ *mazda* *=tauua* *xšaϑrəm*
D.N.SG.NOM Mazda.SG.VOC =2SG.GEN kingdom.N.SG.NOM
‘Он царство Ахуре причисляет: “**Это**, о Мазда, твое царство!” ’ (Y.19.14).

3.3. Рефлексы **aiša-*, **aita-*

Рефлексы этих основ засвидетельствованы в полном объеме в авестийском (только в младоавестийском диалекте), а в древнеперсидском только в форме именительно-винительного падежа (Schmitt 2014). В диалекте Гат они не встречаются, что, впрочем, может быть чистой случайностью (Beekes 1988: 139), тогда как в младоавестийском они встречаются во всех падежах.

В работе (Martínez & de Vaan 2014: 74) высказано предположение о том, что в младоавестийском диалекте формы *aēša-*, *aēta-* постепенно вытеснили формы *ha-*, *ta-*, от которых были образованы — ведь последние встречаются только в именительном, винительном и инструментальном падежах, а формы *aēša-*, *aēta* представлены полной парадигмой. Однако в диалекте Гат основы *ha-*, *ta-* зафиксированы в тех же трех падежах (Beekes 1988: 139). Неизвестно, были ли представлены в нем другие формы, однако такое совпадение наводит на мысль о том, что других падежных форм *ha-*, *ta-* могло не быть вовсе. Когнаты *aēša-*, *aēta-* представлены в том числе и в древнеиндийском как *esá-*, *etá-*, см. Mayrhofer 1992: 272. Таким образом, отсутствие их в диалекте Гат говорит скорее об утрате в нем этих основ, а не об инновации в младоавестийском.

В авестийском *aēša-*/*aēta* в составе супплетивной парадигмы употребляется в контекстах среднего дейксиса, т. е. указания на область, близкую адресату (в том числе психологически). Ср. примеры (17, 18, 19) из трех различных текстов традиции. Указанные контексты не являются анафорическими, референты *aēša-*/*aēta* можно назвать близкими адресату, значит, указанные контексты являются с большой долей вероятности контекстами среднего дейксиса.

(17) *fratarōtaire* *gātuuō* *niṣādaiiōiš* ***aēta*** *humata* ***aēta***
eminent.CMP.M.SG.LOC place.SG.LOC seat.OPT.2SG.ACT D.N.SG.INS good think.PPRT.SG.INS
hūxta ***aēta*** *huuaršta*
D.N.SG.INS good say.PPRT.SG.INS D.N.SG.INS good do.PPRT.SG.INS
«Сознание праведника говорит душе праведника»: ‘Ты усадил (меня) на более выдающееся место (твоими) **этими** добрыми мыслями, **этими** добрыми словами, **этими** добрыми делами’ (H.2.14).

(18) *xvarata* *narō* ***aētəm*** *miiazdəm* *yōi* *=dim*
eat.IMP.2PL.ACT man.PL.VOC D.M.SG.ACC Myazda.SG.ACC REL.M.PL.NOM =3SG.M.SG.ACC
haŋhāna *aṣā* *=ca* *frārəti* *=ca*
deserve.MPRT.PL.NOM righteousness.SG.INS =CONJ fervor.SG.INS =CONJ
‘Вкушайте, мужчины, **эту** Мьязду, кто из вас есть заслужившие её праведностью и рвением’ (Y.8.2).

(19) *sraoṣō* *aṣiiō* *drujim* *apərəsaṯ* <…>
Sraosha.SG.NOM holy.M.SG.NOM Druj.SG.ACC ask.IMF.3SG.ACT
kō *=tē* ***aētaēṣąm*** *arṣnąm* *paoiriiō*
Q.M.SG.NOM =2SG.GEN D.M.PL.GEN man.PL.GEN first.M.SG.NOM
‘Святой Сраоша у Друдж спросил: «Кто из **этих** твоих мужчин первый?» ’ (V.18.33).

В эндофорических контекстах *aēša-/aēta-* встречается как дейктическое местоимение предшествующего дискурса — это дискурс-дейктическое употребление (20).

(20) *kaṯ* ***aētahe*** *paiti varšta* *šiiaoϑna*
Q.N.SG.NOM D.M.SG.GEN for perform.PPRT.N.PL.ACC deed.PL.ACC
ciciϑβå *azaēta*
know.PF.APRT.M.SG.NOM counteract.OPT.3SG.MID
Сначала следует описание злодеяния. Затем Заратуштра спрашивает Ахурамазду: ‘Как совершенные деяния **этого (рода)** он, знающий (о совершенном), искупит?’ (V.18.68).

Многочисленные повторы *aēša-/aēta-* в отдельных фрагментах позволяют зафиксировать также отслеживающее употребление этого местоимения. Если в случае рекогнитивного употребления указательное местоимение встречается в тексте однократно и вводит референта в дискурс, то отслеживающий тип обозначает употребление такого местоимения при каждом следующем появлении референта в дискурсе (Visser 2020: 105).

Проиллюстрируем такое употребление примером из текста Зам-Яшта (Yt.19). Со стиха 19.45 начинается фрагмент текста, посвященный борьбе за Хварно огня и змея. Здесь первый раз во фрагменте упоминается Хварно (21):

(21) *uγrəm* *axvarətəm* *xvarənō*
mighty.N.SG.ACC unseizable.N.SG.ACC Khvarenah.SG.ACC
mazdaδātəm *yazamaide* <…>
by Mazda create.PPRT.N.SG.ACC worship.PRS.1PL.MID
‘Мощное неуловимое Хварно, сотворенное Маздой мы почитаем …’ (Yt.19.45)

Затем следует отступление, после которого нам напоминают, что было такое Хварно. Здесь уже с указательным местоимением *aēta-* (22):

(22) ***aētaṯ*** *xvarənō* *haŋgərəfšāne* *yaṯ* *axvarətəm*
D.N.SG.ACC Khvarenah.SG.ACC capture.SUBJ.1SG.MID REL.N.SG.ACC unseizable.N.SG.ACC
Так думал огонь Ахура Мазды: ‘**Это** Хварно неуловимое я захвачу’ (Y.19.47).

Все следующие вхождения *xvarənō* в этом фрагменте также с указательным местоимением *aēta-* (23, 24, 25):

(23) *yezi* ***aētaṯ*** *niiāsåŋhe* *yaṯ* *axvarətəm* *frā* *ϑβąm*
if D.N.SG.ACC capture.SUBJ.2SG.MID REL.N.SG.ACC unseizable.N.SG.ACC PVB 2SG.ACC
=paiti *apāϑa* <…>
=against put an end.PRS.1SG.ACT
Огонь обращается к змею: ‘Если ты захватишь **это** (Хварно) неуловимое, я уничтожу тебя …’ (Yt.19.48)

(24) ***aētaṯ*** *xvarənō* *haŋgərəfšāne* *yaṯ* *axvarətəm*
D.N.SG.ACC Khvarenah.SG.ACC capture.SUBJ.1SG.MID REL.N.SG.ACC unseizable.N.SG.ACC
Тогда уже змей думает: ‘**Это** Хварно неуловимое я захвачу’ (Yt.19.49)

(25) ***aētaṯ*** *xvarənō* *frapinuuata* *auui* *zraiiō* *vouru.kašəm*
D.N.SG.NOM Khvarenah.SG.NOM fall down.INJ.3SG.MID into sea.SG.ACC Vouru Kasha.SG.ACC
И в итоге Хварно не достается ни тому, ни другому: ‘Упало **это** Хварно в море Ворукаша’ (Yt.19.51).

Таким образом, указательное местомение *aēta-* атрибутивно или субстантивно употребляется в значении отслеживающего дейксиса в Yt.19.47–51. Различные контексты употребления указывают на то, что это — не экзофорическое употребление. Отслеживающий дейксис маркирует референта, за которым должно быть закреплено внимание адресата. В этом фрагменте текста как раз Хварно, переходящее из рук в руки, является объектом нашего интереса, на что указывает в том числе употребление указательного местоимения *aēta-*.

Употребление др.-перс. *aita-*, когната авестийского *aēta-*, также связано с отдельным типом эндофорического дейксиса. В немногочисленных примерах (26, 27) употребление *aita-* обусловлено указанием на элемент из предыдущего дискурса, в тексте прямо не названный — это дискурс-дейктическое употребление указательного местоимения. Контекстов, где *aita-* употреблялось бы в ином значении, нами засвидетельствовано не было. Так, вероятно, в столь узком употреблении сохранился в древнеперсидском языке рефлекс древнеиранского **aita-*.

(26) *antar* ***aitā*** *dahyāva* *tayaį* *upari* *nipištā* *ayaų̊da*
among D.F.PL.ACC country.PL.ACC REL.N.PL.NOM above write.PPRT.F.PL.NOM riot.IMF.3SG.ACT
‘Среди **этих** стран, описанных выше, (одна) пребывала в волнении’ (XPh § 4).

(27) *mām* *auramazdā* *pātu* *hacā* *gastā* *utā* *=maį*
1SG.ACC Ahuramazda.SG.NOM protect.IMP.3SG.ACT from evil.SG.ABL CONJ =1SG.GEN
viϑam *utā* *imām* *dahyāų̊m* ***aita*** *adam* *auramazdām*
house.SG.ACC CONJ D.F.SG.ACC country.SG.ACC D.N.SG.ACC 1SG.NOM Ahuramazda.SG.ACC
jadiyāmi ***aita*** *=maį* *auramazdā* *dadātu*
ask.PRS.1SG.ACT D.N.SG.ACC =1SG.DAT Ahuramazda.SG.NOM give.IMP.3SG.ACT
‘Да защитит Ахурамазда от зла меня и дом мой, и страну эту; **это** я прошу у Ахурамазды; **это** пусть даст мне Ахурамазда’ (DNa § 5).

3.4. Рефлексы *aų̊a-, *hāu-, *hų̊a-

Местоимения **aų̊a-*, **hāu-*, **hų̊a-* содержат в себе дейктический элемент **u*. Как в авестийском, так и в древнеперсидском языке рефлексы **hāu-*, **hų̊a* встречаются только в форме именительного падежа мужского — женского родов. Они образуют супплетивную парадигму с рефлексами **aų̊a-*. Ср. сходное распределение супплетивных основ в парадигме *aēša-*/*aēta-* в подразделе 3.3.

Типичным экзофорическим употреблением рефлексов этих основ является указание на объект, удаленный как от говорящего, так и от адресата. В примерах выше (3, 4, 5, 14) уже освещалось противопоставление рефлексов **aų̊a-*, **hāu-* местоимениям ближнего дейксиса. Ср. также пример (28) из авестийского без такого противопоставления, где референты *auua-* и *hāu* — также удаленные объекты.

(28) *anusō* *zī* *spitama* *zaraϑuštra* *aēšą* *yā*
unwillingly for Spitama.SG.VOC Zarathustra.SG.VOC D.F.SG.NOM REL.F.SG.NOM
paiti.irista ***auuaṯ*** *huuarə* *ā.tāpaite* *anusō*
pollute.INJ.3SG.MID D.N.SG.NOM sun.SG.NOM shine.PRS.3SG.MID unwillingly
hāu *må̊* *anusō* ***auue*** *stārō*
D.M.SG.NOM moon.SG.NOM unwillingly D.M.PL.NOM star.PL.NOM
‘Невольно, ведь, о Спитама Заратуштра, эта [Друдж Насу], которая осквернила, [на неё] **то** солнце светит, невольно **та** луна, невольно **те** звезды’ (V.9.41).

Однако дейктическое употребление рефлексов **aṷa-* этим не исчерпывается. В ряде контекстов как древнеперсидского, так и авестийского языка у **aṷa-* встречается значение среднего дейксиса, описанное выше на примерах (18, 19, 20) для *aēša-/aēta-*.

Ср. примеры (30, 31) из авестийского языка. В примере (30) *varō* ‘Вар’ согласуется с указательным местоимением *auuaṯ*, когда речь идет об объекте, близком к адресату (тому, кто этот Вар приготовил). В том случае, когда говорящим становится готовивший Вар, употребляется местоимение *imaṯ* (31).

(30) *aētaṯ =tē jasāni auuaŋhe azəm yō ahurō*
D.ADV =2SG.DAT come.SUBJ.1SG.ACT help.SG.DAT 1SG.NOM REL.M.SG.NOM Ahura.SG.NOM
mazdå auui ***auuaṯ*** *varō uzdātəm*
Mazda.SG.NOM to D.N.SG.ACC Var.SG.ACC prepare.PPRT.N.SG.ACC
‘Тогда я, Ахура Мазда, приду на помощь тебе, к **этому** приготовленному (тобой) Вару’ (Yt.12.4).

(31) *uruuaϑąm =ciṯ zbaiiemi auui* ***imaṯ*** *varō uzdātəm*
friend.SG.ACC =EMPH invoke.PRS.1SG.ACT to D.N.SG.ACC Var.SG.ACC prepare.PPRT.N.SG.ACC
‘Призываю (его, Рашну) другом к **этому** приготовленному (мной) Вару’ (Yt.12.5).

В отличие от ряда сходных контекстов для *aēša-/aēta-*, нам удалось засвидетельствовать только один бесспорный контекст среднего дейксиса для *auua-* в авестийском языке, в тексте Рашн-Яшта (30). Возможно, тому виной общая скудность контекстов среднего дейксиса в авестийских текстах. Мы также можем предположить, что за выбором демонстратива *aēša-/aēta-* или *auua-* в контекстах среднего дейксиса стоят диалектные различия, причем диалект, выбирающий стратегию *auua-*, хуже отражен в Авесте. Нельзя исключать и искажение текста Яшта. Рашн-Яшт почти целиком состоит из повторов ритуальных формул и плохо сохранился, так что проследить употребление демонстративов в нем в целом затруднительно.

Сходным образом и в древнеперсидском языке засвидетельствовано употребление *ava-* в отношении референта, близкого собеседнику, ср. примеры (32, 33). Более того, не лишено основания утверждение, что *ava-* в древнеперсидском — единственное указательное местоимение для обозначения среднего дейксиса. По крайней мере, так обстоит дело для тех немногих контекстов обращения к собеседнику, которые отражены в древнеперсидском корпусе. Если это так, то руководствуясь теорией синкретизма дейктических значений (Lander & Haegeman 2018), возможно постулировать синкретизм значений среднего и дальнего дейксиса для древнеперсидского языка. Это означает, что, хотя указательные местоимения в языке и представлены двумя сериями, противопоставленными по удаленности, употребление местоимения *ava-* шире, используясь также в контекстах среднего дейксиса.

(32) *utā taya kunavāhi* ***ava*** *=taiy*
CONJ REL.N.SG.ACC make.PRS.2SG.ACT D.N.SG.ACC =2SG.DAT
auramazdā ucāram kunaṷtu
Ahuramazda.SG.NOM successful.N.SG.ACC make.IMP.3SG.ACT
‘И что ты делаешь, **то** тебе Ахурамазда пусть сотворит успешным’ (DB § 66).

(33) ***avaš*** *=ci āxšnudi taya paratar ϑanhyāti*
D.N.SG.ACC =EMPH hear.IMP.2SG.ACT REL.N.SG.ACC beyond-that is-stated.SUBJ.3SG.ACT
‘**То** услышь, что помимо этого сказано’ (DNb § 11).

Эндофорическое употребление рефлексов **aua-*, **hāu-* также представлено в текстах Авесты и надписях ахеменидских царей. Ср. пример (34), где *auua-* употребляется катафорически, при этом катафора выступает здесь также во временном значении. То есть *auuaṯ* относится к тому референту, который означает некоторое возможное будущее событие. Ср. примеры из русского языка: ***Я прошу об этом, чтобы завтра я мог восславить Ахурамазду*[6] и *Я прошу о том, чтобы я завтра мог восславить Ахурамазду*.

(34) ***auuaṯ*** *āiiaptəm* *dazdi* *=mē* *aṣ̌iš* *vaŋuhi*[7]
D.N.SG.NOM boon.SG.NOM give.IMP.2SG.ACT =1SG.GEN Asha.SG.NOM good.F.SG.NOM
yā *bərəzaiti* *yaṯ* *bauuāni* *aiβi.vaniiå̊*
REL.F.SG.NOM grow-high.PRS.3SG.ACT REL.N.SG.NOM become.SUBJ.1SG.ACT victorious.M.SG.NOM
vīspe *daēuua* *māzaniia* <…> *pairi.tacaṯ* *pairi.jasaṯ*
all.M.PL.ACC Daeva.PL.ACC giant.PL.ACC run-near.INJ.3SG.ACT approach.INJ.3SG.ACT
aṣ̌iš *vaŋuhi* *yā* *bərəzaiti* *vīṇdāṯ*
Asha.SG.NOM good.F.SG.NOM REL.F.SG.NOM grow-high.PRS.3SG.ACT obtain.INJ.3SG.ACT
təm *yānəm* *Haošiianŋhō* *Paraδātō*
D.M.SG.ACC gift.SG.ACC Haoshyangha.SG.NOM Paradhata.SG.NOM
‘**Такой** дар, прошу, дай мне, о Аша высокая, чтобы я стал побеждающим всех дэвов-великанов … Примчалась Аша прекрасная, высокая, и получил тот дар Хаошьянха Парадата’ (Yt.17.25–26).

Ср. также анафорическое употребление со значением временной анафоры (35), противоположное катафоре в примере (34):

(35) *ciṯ* ***auuaṯ*** *vacō* *ās* *ahura*
what.N.SG.NOM D.N.SG.NOM word.SG.NOM COP.INJ.3SG.ACT Ahura.SG.VOC
mazda *yaṯ* *=mē* *frāuuaocō*
Mazda.SG.VOC REL.N.ACC.SG =1SG.DAT proclaim.INJ.2SG.ACT
‘Что **то** было за слово, о Ахура Мазда, которое ты мне изрёк? ’ (Y.19.1).

Что касается употребления *hau̯* в древнеперсидском языке, оно за исключением нескольких примеров субстантивно; *hau̯* практически завершило переход из разряда указательных местоимений в личное местоимение 3-го лица единственного числа, ср. (36).

(36) *utā* *1* *martiya* *bābiruviya* *nadintabai̯ra* *nāma*
CONJ one man.SG.NOM Babilonian.SG.NOM Nadintabaira.SG.NOM name.SG.ACC
ai̯nai̯rahyā *puça* ***hau̯*** *udapatatā* *bābirau̯*
Aniri.SG.GEN son.SG.NOM D.M.SG.NOM rebel.IMF.3SG.MID Babilon.SG.LOC
‘И один муж, вавилонянин по имени Надинтабайра, сын Анири; **он** восстал в Вавилоне’ (DB § 16).

При этом во всех атрибутивных примерах данное указательное местоимение демонстрирует рекогнитивное употребление. В примере (37) *hau̯* употребляется атрибутивно и не имеет референта в предшествующем дискурсе фрагмента (Фраорт в тексте вообще упоминается впервые). Оба этих условия — основные для определения рекогнитивного дейксиса (Diessel 1999: 105–107).

[6] * здесь означает, что, если *об этом* является вершиной относительной клаузы *чтобы завтра я мог восславить Ахурамазду*, то предложение под * неграмматично.

[7] По-видимому, следует читать *vaŋ^{v}hī-* (Martínez, de Vaan 2014: 58), о чем говорит также нулевое окончание в номинативе в примере (34).

(37) ***hauꞑ*** *fravartiš* *haya* *mādaiꞑ* *xšāyaϑiya*
D.M.SG.NOM Fraortes.SG.NOM REL.M.SG.NOM Media.SG.LOC king.SG.NOM
‘**Тот** Фраорт, который царь в Мидии’ (DB § 31).

Итак, нам удалось выявить основные типы употребления демонстративов, которые встречаются в текстах Авесты (кроме текстов Гат и Ясны Хаптанхаити) и древнеперсидских царских надписей. Несомненно, встречаются исключения и расхождения, и их не всегда легко объяснить. Тем не менее, к этим типам сводится большинство употреблений указательных местоимений в обсуждаемом корпусе, что будет рассмотрено более подробно на примере двух избранных текстов авестийской традиции в разделе 4.

4. Употребление указательных местоимений в текстах Ардвисур-Яшта (Yt.5) и Хадохт-Наска (H)

В тексте Абан-Яшта или Ардвисур-Яшта (Yt.5) насчитывается 132 стиха. В нем также содержится множество повторов, обеспеченных главным рефреном — жертвоприношением и просьбой к Ардвисуре Анахите одного из героев (или злодеев), затем её согласием или отказом. Единственная крупная вставка, прерывающая череду почти непрерывных обращений — это фрагмент Yt.85-96. Такая структура текста обеспечивает множество идентичных или почти идентичных употреблений указательных местоимений. В частности, *tąm*, *taꞑ*, *ana*, *auuaꞑ*, *ahmāi* повторяются многократно.

Всего указательных местоимений в тексте Яшта: 32 от основы *a-*; 6 от основы *ima-*; 62 от основы *ana*; 119 от основы *ha-*/*ta-*; 6 от основы *aēša-*/*aēta-*; 50 от основы *auua-*. В их число не вошли при подсчете наречные формы на *-aꞑ* и *-ϑa*: *āaꞑ* ‘так, тогда, затем’ (образовано от *a-*), также *auuauuaṇt-* ‘такой, столько’ (образовано от *auua-*).

Местоимения основы *a-* употребляются в контекстах ближнего дейксиса и ограниченно (Yt.5.4 = 5.102, Yt.5.5) в анафорических; также в значении личного местоимения 3-го лица (*ahmāi*, *aēibiiō*) см. выше примеры (11, 12). Анафорическое употребление *aēṣąm*, *aiŋhåsca* в примерах (38, 39) продиктовано отсутствием формы родительного падежа множественного числа у местоимения *ha-*, *ta-* (сходным образом для форм личного местоимения 3-го лица).

(38) <…> *yeŋhe* *hazaŋrəm* *vairiianąm* *hazaŋrəm* *apaγžāranąm*
REL.F.SG.LOC thousand.SG.NOM lake.PL.GEN thousand.SG.NOM inflow.PL.GEN
kasciꞑ *=ca* ***aēšąm*** *vairiianąm* *kasciꞑ* *=ca* ***aēšąm*** *apaγžāranąm*
INDEF.M.SG.NOM =CONJ D.M.PL.GEN lake.PL.GEN INDEF.M.SG.NOM =CONJ D.M.PL.GEN inflow.PL.GEN
caϑβarə.satəm *aiiarə.baranąm* *huuaspāi* *naire* *barəmnāi*
forty.SG.NOM day ride.PL.GEN good horse.M.SG.DAT man.SG.DAT carry.MPRT.M.SG.DAT
‘…у которого тысяча заливов, тысяча притоков, и каждый из этих заливов, и каждый из этих притоков — в сорока сутках езды всаднику на хорошей лошади несущемуся’ (Yt.5.4).

(39) ***aiŋhåsca*** *=mē*[8] *aēuuaŋhå* *āpō*
D.F.SG.GEN =1SG.DAT one.F.SG.GEN river.SG.GEN
‘**этой** одной реки’ (Yt.5.5).

[8] В. И. Абаев утверждает, что *mē* здесь употребляется в значении *dativus ethicus* в речи рассказчика (Абаев 2016: 92). Альтернативная трактовка заключается в том, что в стихе все же представлена речь самой

Употребление *ana* обусловлено контекстами перформативных высказываний, преимущественно с одним и тем же контекстом, повторяющимся на протяжении всего текста (Yt.5.9, 5.11, 5.13, 5.15 и т. д.), см. пример (7). Местоимения от основы *ha-*/*ta-* широко представлены в тексте в разных формах (*hō*, *hā*, *tąm*, *taṯ*, *tā*, *tē*) в контекстах анафоры (в том числе дискурсивной), а также в качестве местоимения 3-го лица, не всегда по смыслу отличного от анафорического местоимения, см. примеры (11, 40, 41, 42). В примере (42), как и в (16) анафорическое местоимение *ha-*, *ta-* употребляется при передаче прямой речи.

(40) ***hō*** *auuaϑa* *vazata* *ϑri.aiiarəm* *ϑri.xšaparəm*
D.M.SG.NOM D.ADV fly.INJ.3SG.MID three days.SG.ACC three nights.SG.ACC
‘**Он** так носился три дня и три ночи’ (Yt.5.62).

(41) ***hā*** *=hē* *bāzauua* *gəuruuaiiaṯ* *moṣu*
D.F.SG.NOM =3SG.GEN arm.DU.ACC take up.INJ.3SG.ACT quickly
taṯ *ās* *nōiṯ* *darəγəm*
D.N.SG.NOM COP.INJ.3SG.ACT NEG long
‘**Она** его на руки взяла; быстро **то** было, не долго’ (Yt.5.65)

(42) <…> *uiti* *vacəbiš* *aojanō:* ***tā*** *bā*
thus word.PL.INS speak.MPRT.M.SG.NOM D.N.PL.NOM indeed
aṣa ***tā*** *aršuxδa*
truthful.N.PL.NOM D.N.PL.NOM straight-spoken.N.PL.NOM
‘…так словами говоря: **это** правдивые, **это** правильные (слова)’ (Yt.5.76–77).

Местоимение *aēša-*/*aēta-* употребляется в тексте в нескольких контекстах дискурс-дейксиса, см. пример (43). В этом же примере см. указательное местоимение дальней степени *auuaṯbiiō*.

(44) ***aēta*** *yasna* ***aēta*** *vahma* <…> *paiti.auua.jasa*
D.M.SG.INS prayer.SG.INS D.M.SG.INS invocation.SG.INS come-down-back.IMP.2SG.ACT
arəduuī *sūre* *anāhite* *haca* ***auuaṯbiiō*** *stərəbiiō*
Ardvi.SG.VOC Sura.SG.VOC Anahita.SG.VOC from D.M.PL.ABL star.PL.ABL
Пассаж, *завершающий гимн*: ‘Силой этой молитвы, этого призыва… о Ардви Сура Анахита, спустись назад с тех звезд! ’ (Yt.5.132).

Местоимения, образованные от основы *auua-*, употребляются в тексте как в значении дальнего дейксиса (44), так и в контекстах временной катафоры, см. пример (45). Особо следует отметить употребление *auuaṯ* в примере (46). Строго говоря, это анафорическое употребление, скорее нехарактерное для *auua-*, однако лексикализовавшееся в этой формуле (она повторяется по всему тексту: в стихах 5.19, 5.23, 5.27, 5.31, 5.35, 5.39, 5.43, 5.47, 5.51, 5.55, 5.59, 5.66, 5.70, 5.74, 5.79, 5.83, 5.99, 5.106, 5.110, 5.114, 5.118. При этом в 5.66 (и в некоторых других отрывках) — формула не предваряется эксплицитным выражением просьбы (*auuaṯ āiiaptəm dazdi mē*), но формула повторяется все равно. В формуле вместе с тем присутствует анафорическое местоимение *taṯ* (не следует ли здесь предполагать некую ошибку, возникшую при передаче текста?).

Анахиты и *mē* употреблено в родительном падеже ‘моей’. Фрагмент Yt.5.1–5.9 испещрен различными вставками, меняются рассказчики и объект описания (антропоморфная богиня или река), так что можем здесь предположить, что это говорит Анахита, и Yt.5.5 просто вырван из контекста. Тогда и употребление указательного местоимения *aiŋhåsca* следует отнести к ближнему дейксису.

(45) ***auuaṯ*** *āiiaptəm* *dazdi* *=mē* <…> *yaϑa* *azəm*
D.N.SG.ACC boon.SG.ACC give.IMP.2SG.ACT =1SG.GEN REL.ADV 1SG.NOM
urpətəm *xšaϑrəm* *bauuāni* *vīspaṇąm* *viuṇąm*
highest.N.SG.NOM power.SG.NOM become.1SG.SUBJ.ACT all.F.PL.GEN land.PL.GEN
‘**Такую** награду/удачу дай мне… чтобы я высшей властью стал над всеми странами’ (Yt.5.22).

(46) *daϑaṯ* *ahmāi* ***taṯ*** ***auuaṯ*** *āiiaptəm*
give.INJ.3SG.ACT D.M.SG.DAT D.N.SG.ACC D.N.SG.ACC boon.SG.ACC
‘Дала ему **это/эту**, **такую** награду’ (Yt.5.22).

Наконец, в тексте местоимение *ima-* употребляется в контексте ближнего дейксиса:

(47) <…> *yaϑa* *vīspå* ***imå*** *āpō*
REL.ADV all.F.PL.NOM D.F.PL.NOM water.PL.NOM
yå *zəmā* *paiti* *frataciṇti*
REL.F.PL.NOM earth.SG.INS upon flow forth.PRS.3PL.ACT
‘…словно все **эти** воды, которые по земле текут’ (Yt.5.15 = 5.96, 5.102, 5.121).

Также *ima-* встречается в одном пассаже, где достоверно определить контекст не удается (Yt.5.95). Пассаж определенно испорчен и труден для понимания, *imå paiti.vīsəṇte* сложно трактовать потому, что непонятно значение собственно глагола *paiti.vīsəṇte*, а также референт местоимения *imå*, ср. (Lubotsky 2022: 221–222). По этой причине он не глоссируется в настоящей работе.

Хадохт-Наск насчитывает 59 стихов, разбитых на две части, 17 и 42 стиха каждая. Вторая часть также разделяется на два смысловых фрагмента. Первый фрагмент описывает судьбу, постигшую после смерти душу праведника, второй — душу грешника. Между двумя фрагментами много сходств, второй многократно цитирует первый. В том числе повторяются и конструкции с указательными местоимениями.

В тексте Хадохт-Наска нами выявлено 57 употреблений указательных местоимений. Из этих 48 употреблений 20 приходится на *aēša-*/*aēta-*, 12 на *ha-*/*ta-*, 6 на *hāu*, 5 на *auua-*, 9 на *ima-*, 5 на *a-*.

Местоимение *hō* повторяется 6 раз в третьем стихе текста, см. пример (48). Употребление *hō* здесь анафорическое. Для такого употребления характерны повторы, то есть анафора в литературоведческом смысле.

(48) *yō* *aṣəm* *staoiti* <…> ***hō*** *mąm* *staoiti*
REL.M.SG.NOM Asha.SG.ACC praise.PRS.3SG.ACT D.M.SG.NOM 1SG.ACC praise.PRS.3SG.ACT
yim *ahurəm* *mazdąm* ***hō*** *āpəm* *staoiti*
REL.M.SG.ACC Ahura.SG.ACC Mazda.SG.ACC D.M.SG.NOM water.SG.ACC praise.PRS.3SG.ACT
hō *ząm* *staoiti* ***hō*** *gąm* *staoiti*
D.M.SG.NOM earth.SG.ACC praise.PRS.3SG.ACT D.M.SG.NOM cow.SG.ACC praise.PRS.3SG.ACT
hō *uruuarå* *staoiti* ***hō*** *vīspa* *vohū*
D.M.SG.NOM plant.PL.ACC praise.PRS.3SG.ACT D.M.SG.NOM all.N.PL.ACC good.N.PL.ACC
mazdaδāta *aša.ciϑra* *staoiti*
by-Mazda-give.PPRT.N.PL.ACC with-seed of-Asha.N.PL.ACC praise.PRS.3SG.ACT
‘Кто Ашу восхваляет, … **тот** меня, Ахура Мазду восхваляет, **тот** воду восхваляет, **тот** землю восхваляет, **тот** корову восхваляет, **тот** растения восхваляет, **тот** все благое, сотворенное Маздой, содержащее семя Аши восхваляет’ (H.1.3).

Местоимение *taṯ* встречается в тексте во фрагментах H.2.8, H.2.26 (повтор), дважды в H.2.18 и H.2.36 (повтор). Во всех случаях прослеживается анафора. Ср. употребление *taṯ* в примере (49), где описывается пища, подносимая праведным после смерти. Аналогично употребление *təm* в контексте анафоры см. в H.2.8, H.2.26.

(49) *xvarəϑanąm =he bərətanąm <...>* ***taṯ***
food.PL.GEN =3SG.DAT carry.PPRT.N.PL.GEN D.N.SG.NOM
asti yūnō <...> ***taṯ*** *nāirikaiiāi*
COP.PRS.3SG.ACT young man.SG.GEN D.N.SG.NOM woman.SG.DAT
'Из пищи, ему (праведнику) принесенной... **эта** — юноши... **эта** — женщине' (H.2.18)

В нескольких отрывках (H.1.4, H.2.41) встречаем рекогнитивное употребление *aēša-/aēta-* (всего 3), как в примере (50). Указательное местоимение *aēṣ̌ō* здесь вводит нового референта в дискурс.

(50) *āaṯ uṣ̌aŋhąm para frārətōiṯ* ***aēṣ̌ō*** *mərəγō*
then dawn.SG.ACC towards come towards.OPT.3SG.ACT D.M.SG.NOM bird.SG.NOM
yō parō.darš ***aēṣ̌ō*** *mərəγō*
REL.M.SG.NOM Parodarsh.SG.NOM D.M.SG.NOM bird.SG.NOM
yō karətō.dąsuš āϑrō vācəm surunaoiti
REL.M.SG.NOM Kareto-dasu.SG.NOM fire.SG.GEN voice.SG.ACC listen.PRS.3SG.ACT
'Тогда ближе к рассвету приходит **эта** птица Пародарш, **эта** птица Карето-дасу, слушает голос огня' (H.2.41).

Во фрагменте (H.2.1-6) можно обозначить употребление *aēša-/aēta-* в значении отслеживающего дейксиса. Существительное *xšapan-* 'ночь' последовательно (6 раз) употребляется с местоимением *aēta-*. Заратуштра спрашивает у Ахура Мазды, что должно случиться на такую-то ночь, и Ахура Мазда отвечает, что на такую-то ночь должно случиться следующее. Аналогичная серия вопросов и ответов представлена и в другом фрагменте текста (H.2.19–24). Подробнее об употреблении отслеживающего дейксиса см. примеры (21–25).

Дейктическое употребление *aēta-* в тексте Хадохт-Наска уже рассматривалось выше в примере (18), где представлен диалог души и сознания праведника, это H.2.14. В этом фрагменте *aēta-* употребляется трижды. В H.2.36 повторяется та же конструкция, что и в H.2.14, *aēta-* употребляется в диалоге души и сознания грешника.

Подобный контекст диалога, в котором также можно было бы рассмотреть контекст среднего дейксиса — это стих H.2.11, где употребляется указательное местоимение *auua-*. Однако появление *auua-* в следующем стихе H.2.12, где мы не ожидаем видеть средний дейксис, указывает на то, что *auua-*, вероятно, употребляется в обоих примерах (51, 52) временной удаленности. Эта временная удаленность в контексте соотносится с поступками и качествами из прошлой, земной жизни праведника. Аналогичным образом конструкции повторяются в H.2.29, H.2.30.

(51) *cišca ϑβąm cakana* ***auua*** *masana =ca*
INDEF.M.SG.NOM 2SG.ACC love.PF.3SG.ACT D.N.SG.INS greatness.SG.INS =CONJ
vaŋhana =ca sraiiana =ca hubaoiδita =ca
virtue.SG.INS =CONJ fairness.SG.INS =CONJ sweet-scentedness.SG.INS =CONJ
vərəϑrająsta =ca paiti.duuaēšaiiaṇta =ca yaϑa
victorious-strength.SG.INS =CONJ counteracting-enemies.SG.INS =CONJ REL.ADV

yaţ *=me* *saδaiiehi*
REL.N.ACC.SG =1SG.DAT reveal.PRS.2SG.ACT
‘Всякий любил тебя за **то** величие и добродетель, и справедливость, и благовонность, и победоносную силу, и противостояние врагам, как (те,) которые ты мне являешь’ (H.2.11).

(52) *tum* *mąm* *cakana* <…> ***aiua*** *masana* *=ca*
2SG.NOM 2SG.ACC love.PF.3SG.ACT D.N.SG.INS greatness.SG.INS =CONJ
vaŋhana *=ca* *sraiiana* *=ca* *yaϑa* *yaţ* *=te* *saδaiiemi*
virtue.SG.INS =CONJ fairness.SG.INS =CONJ REL.ADV REL.N.ACC.SG =2SG.DAT reveal.PRS.1SG.ACT
‘Ты меня любил… за **то** величие и добродетель, и справедливость, как (те), которые я тебе являю’ (H.2.12).

В тексте также встречается одно дейктическое противопоставление, которым обусловлено употребление указательных местоимений *ima-* и *aiua-* в пассаже H.1.16, ср. пример (53). К дейктическому употреблению *ima-* см. также пример (54). В примерах (44, 47, 54) отсутствует явный контраст между близким и далеким, однако использование местоимений *a-*, *ima-* для (психологически) близких объектов (*gaēϑā-* ‘мир, бытие, творение’ (Yt.5.73, 5.109, 5.113), *ap-* ‘вода’, в тексте Яшта «воды, которые по земле текут», *karšuuar-* ‘страна, континент’) и *aiua-* для далеких (*star-* ‘звезда’) может быть семантически обусловлено противопоставлением земного мира неземному.

(53) <…> *yaţ* *antar* *=əca* *ząm* *antar* *=əca* *asmanəm*
REL.N.SG.NOM into =CONJ earth.SG.ACC into =CONJ heaven.SG.ACC
imąm *=ca* *ząm* ***aiua*** *=ca* *raocå̊*
D.F.SG.ACC =CONJ earth.SG.ACC D.N.PL.ACC =CONJ *luminous.PL.ACC*
‘… которое на земле и в небесах, **этой** земле, **тех** светлых [небесах]’ (H.1.16).

(54) *kā* *aēiua* *ašō.stūitiš* *yā*
Q.F.SG.NOM one.F.SG.NOM praise-of-Asha.SG.NOM REL.F.SG.NOM
vīspəm ***imaţ*** *karšuuarə*
all.N.SG.ACC D.N.SG.ACC country.SG.ACC
‘Какое восхваление Аши одно (стоит) всей **этой** страны?’ (H.1.14).

Наконец, в примере (55) демонстрируется употребление *hāu* в катафорической конструкции. Оно же повторяется в тексте шестикратно.

(55) *kā* *aēiua* *ašō.stūitiš* <…> ***hāu*** *bā*
Q.F.SG.NOM one.F.SG.NOM prayer of Asha.SG.NOM D.F.SG.NOM indeed
ašāum *zaraϑuštra* *yąm* *bā* *nā* *pasca* *fraŋuharəiti*
righteous.M.SG.VOC Zarathustra.SG.VOC REL.F.SG.ACC indeed man.SG.NOM after eat.PRS.3SG.ACT
haomahe *hutahe* *ašəm* *staoiti* *frastauuanō*
haoma.SG.GEN press.PPRT.M.SG.GEN Asha.SG.ACC praise.PRS.3SG.ACT glorify.MPRT.M.SG.NOM
‘Какое восхваление Аши … **такое**, о праведный Заратуштра, когда муж после вкушения хаомы выжатой, Ашу славит, прославляющий’ (H.1.8-9).

Местоимение *a-* употребляется анафорически / в значении 3-го лица, так же, как и в тексте Ардвисур-Яшта, ср. пример (9) выше.

Таким образом, употребление указательных местоимений в текстах Ардвисур-Яшта и Хадохт-Наска в целом соответствует основным типам употреблений, выделенных нами

для каждой из основ на основании общего анализа авестийского корпуса. Диалектные различия в употреблении указательных местоимений при анализе двух данных текстов нами не были выявлены. Описанные выше особенности употребления указательных местоимений позволяют сделать вывод о четырех сериях указательных местоимений для авестийского языка, а именно ближней, средней, дальней и анафорической. Анафорическое употребление (а также производное из него употребление в значении местоимения 3-го лица) закреплено за *ha-*/*ta-*/*a-*; катафорическое — за *auua-*/*hāu*. Контексты дискурс-дейксиса (в том числе его отслеживающего типа) закреплены за местоимением *aēša-*/*aēta-*.

Теперь можно уверенно сказать, что отнесение *ha-*/*ta-*, а также *aēša-*/*aēta-* к ближней серии без «разграничения основ и функций» (Соколов 1979: 185; Виноградова 2000: 24) несостоятельно. Анафорическое значение, ключевое для *ha-*/*ta-*, а также *a-* в отдельных косвенных падежах, должно послужить весомым доводом для того, чтобы выделять отдельную анафорическую серию. Это местоимение в своем значении соответствует др.-инд. *sá-*, *tá-* — которое также не выражает значений дейктической удаленности, однако представлено, в отличие от авестийского, полной парадигмой (Елизаренкова 1982: 244–245).

Иначе можно представить местоимение *ha-*/*ta-* как принадлежащее области der-deixis, не выражающей значения удаленности от дейктического центра. Эта модель принята в сравнительно-исторических описаниях начала XX века, однако и в них местоимение *ha-*/*ta-* не воспринимается как однозначно анафорическое, ср. примеры из (Reichelt 1909: 283), уже рассматривавшиеся в нашей работе (14, 15).

5. Реконструкция праиранской системы указательных местоимений

В работе Kupfer 2002 приводится реконструкция системы указательных местоимений древнеиндийского языка, см. таблицу 4. Все представленные в таблице местоимения обладают древеиранскими когнатами. К. Купфер анализирует примеры на употребление и приходит к выводу, что древнеиндийская система была двухчастной.

Таблица 4. Употребление местоимений др.-инд. языка (Kupfer 2002: 370).

Близкое/видимое для говорящего	Далекое/невидимое для говорящего	Анафорическое	Эмфатическое
idám	**avád*	*tád*	*etád*

Мы можем с уверенностью сказать, что данные авестийского языка позволяют реконструировать праиранскую систему указательных местоимений как трехчастную, ср. таблицу 5, где представлены четыре серии указательных местоимений. Совпадение с древнеиндийским нарушается только в позиции рефлексов **aiša-*, **aita-*.

Таблица 5. Четыре серии указательных местоимений праиранского языка.

I серия (ближний дейксис)	II серия (средний дейксис)	III серия (дальний дейксис)	Анафорическая серия
**a-*, **aḭ-*, **iḭ-*, **ima-*, **anā*	**aiša-*, **aita-*	**aṵa-*, **hāu-*	**ha-*, **ta-*; возможно, **a-*

На позицию **aiša-*, **aita-* в системе на месте среднего дейксиса указывают данные не только авестийского, но и других иранских языков. В генетически неоднородных памирских языках в результате лингвистических контактов возобладала трехчастная система,

в которой в качестве указательного местоимения среднего дейксиса выступают рефлексы **aiša-*, **aita-*. Рефлексы **aiša-*, **aita-* в значении среднего дейксиса прослеживаются и в таких среднеиранских языках, как согдийский, хотаносакский, парфянский (Sims-Williams 1994; Wendtland 2009: 181; Durkin-Meisterernst 2014: 282). Отсутствие в том или ином иранском языке рефлексов **aiša-*, **aita-* должно быть связано с утратой в нем соответствующих дейктических категорий и качественным изменением системы. Так, в юго-западных иранских языках этот процесс начался, по-видимому, достаточно рано, что прослеживается уже на примере двухчастной системы древнеперсидского языка.

6. Заключение

«Форм именительного падежа от различных местоименных основ в языке Авесты вообще несколько больше, чем хотелось бы для установления стройной системы» — писал С. Н. Соколов (Соколов 1979: 185). Тем не менее, система местоимений авестийского языка, действительно, более сложная и развернутая, чем древнеперсидская, все же поддается описанию. Мы можем выделить четыре серии указательных местоимений. Первая серия включает в себя местоимения, образующие супплетивную парадигму *aē-/ima-/ī-/a-/ana*. Вторая — местоимение *aēša-/aēta-* (а также, возможно, местоимение *auua-* в диалекте, представленном в Yt.12). Третья — местоимения *auua-/hāu*, которые также составляют супплетивную парадигму. Четвертая — анафорическое местоимение *ha-/ta-/a-*. Из данного описания следует, что *a-* присутствует сразу в двух сериях, образуя специфический синкретический тип. Кроме того, некоторые типы эндофорического употребления также закреплены за отдельными сериями или местоимениями. Так, *ana* употребляется в контекстах перформативного высказывания, за *aēša-/aēta-* закреплен тип дискурс-дейксиса (в том числе его отслеживающего типа), а за *auua-/hāu* — катафора.

В то же время диалектные различия в употреблении указательных местоимений на материале двух различных текстов авестийской традиции не выявляются. Едва ли они могут служить объяснением столь распространенного среди основ указательных местоимений супплетивизма. При этом супплетивизм основ не мешает закреплению за каждой из парадигм одного или нескольких значений, а различия в выражении этих значений между основами, образованными от разных корней, в одной и той же парадигме не прослеживаются. Наконец, *ha-/ta-/a-* в авестийском засвидетельствовано также на пути грамматикализации в личное местоимение 3-го лица (то есть местоимение может выражать как анафорические значения, так и значения местоимения 3-го лица в различных падежных формах).

Для древнеперсидского мы показали, что указательные местоимения распределяются по двум сериям. Основы *iya-*, *ima-*, *a-*, *anā* относятся к одной супплетивной парадигме и покрывают значения ближнего дейксиса. Указательное местоимение *ava-*, традиционно характеризуемое как дальний дейксис, также употребляется в контекстах, где говорящий обращается к адресату, что позволяет говорить о покрытии им области среднего дейксиса и синкретизме средний + дальний дейксис (в терминах Lander & Haegeman 2018). Хотя древнеперсидские основы *ava-*, *haṷ* и следует отнести к супплетивной парадигме, исходя, по крайней мере, из сведений морфологии, употребление *haṷ* в экзофорических контекстах на корпусе древнеперсидских текстов не прослеживается. Местоимение *aita-* засвидетельствовано лишь в эндофорическом дискурс-дейктическом употреблении. В целом система указательных местоимений древнеперсидского языка схожа с авестийской, однако сильно упрощена, а грамматическое выражение дейктиче-

ских значений в ней ослаблено. Полные парадигмы указательных местоимений авестийского и древнеперсидского языков см. в Приложении 1 (таблицы 6–11).

Новые представления о дейксисе и демонстративах в древнеиранских языках должны оказаться полезными для будущих переводов и толкований древних текстов. Полученные данные также являются релевантными для описания систем указательных местоимений в других иранских языках.

В то же время, возможны расхождения в трактовках отдельных форм, а мотивацию употребления того или иного указательного местоимения не всегда удается определить точно. Пересекаются различные дейктические значения: например, анафоры, катафоры/временной анафоры и временной катафоры у местоимений *ha-/ta-/a-* и *auua-/hāu* — так что они встречаются в довольно схожих контекстах. Экзофорические контексты пересекаются с эндофорическими (в особенности при передаче прямой речи) и на материале древних текстов не всегда можно принять одно верное решение. Лучший статистический анализ употребления демонстративов могли бы дать данные корпуса с глоссированными текстами и тщательно сделанным подстрочным переводом. Такого ресурса для древнеиранских языков, к сожалению, не существует.

Наконец, пока без должного внимания остались указательные местоимения в текстах Гат и Ясны Хаптанхаити. Это несомненно более сложный материал, написанный на отличном от всей остальной Авесты языке, небольшой на фоне остальной Авесты, но имеющий множество различных трактовок и переводов. Изучение указательных местоимений в древнеавестийском корпусе требует отдельной работы.

Приложение 1. Полные парадигмы указательных местоимений авестийского и древнеперсидского языков

Таблица 6. Парадигма указательного местоимения I серии *aē-/ima-/ī-/a-/ana* (ближний дейксис, авестийский язык)

Число	Падеж	Мужской род	Средний род	Женский род
Ед. ч.	Им. п.	*aēm*	*imaṯ*	*īm*
	Вин. п.	*imәm*	*imaṯ*	*imąm*
	Род. п.	*ahe, aiŋ́he*		*aiŋ́hå, aiŋ́håsca*
	Дат. п.	*ahmāi*		*aiŋ́hāi*
	Отлож. п.	*ahmāṯ*		*aiŋ́hāt*
	Инстр. п.	*ana*		*aiia*
	Местн. п.	*ahmi, ahmiia*		*aiŋ́he*
Дв. ч.	Им. п.	*ima*	—	—
	Род. п.	*aiiå̊*	—	—
	Дат. п.	*ābiia*		*ābiiā*
Мн. ч.	Им. п.	*imą, ime*	*ima, imå*	*imå*
	Вин. п.	*imą, ime*	*ima, imå*	*imå*
	Род. п.	*aēšąm*		*å̊ŋhąm*
	Дат. п.	*aēibiiō*		*ābiiō*
	Инстр. п.	*aēibiš*		—
	Местн. п.	*aēšu, aēšuua*		*āhuua*

Таблица 7. Парадигма анафорического указательного местоимения *ha-*/*ta-*/*a-* (авестийский язык)

Число	Падеж	Мужской род	Средний род	Женский род
Ед. ч.	Им. п.	*hō, hā*	*taṯ*	*hā*
	Вин. п.	*təm*	*taṯ*	*tąm*
	Род. п.	*ahe, aiŋ́he*		*aiŋ́hå, aiŋ́håsca*
	Дат. п.	*ahmāi*		*aiŋ́hāi*
	Отлож. п.	*ahmāṯ*		*aiŋ́hāt*
	Инстр. п.	*tā*		*aiia*
	Местн. п.	*ahmi, ahmiia*		*aiŋ́he*
Дв. ч.	Им. п.	*tā*	*tē*	—
	Вин. п.	*tā*	*tē*	—
	Дат. п.	*ābiia*		*ābiiā*
Мн. ч.	Им. п.	*tē*	*tā*	*tå*
	Вин. п.	*tą, tə*	*tā*	*tå*
	Род. п.	*aēšąm*		*åŋhąm*
	Дат. п.	*aēibiiō*		*ābiiō*
	Инстр. п.	*aēibiš*		—
	Местн. п.	*aēšu, aēšuua*		*āhuua*

Таблица 8. Парадигма указательного местоимения II серии *aēša-*/*aēta-* (средний дейксис, авестийский язык)

Число	Падеж	Мужской род	Средний род	Женский род
Ед. ч.	Им. п.	*aēša, aēšō*	*aētaṯ*	*aēšā*
	Вин. п.	*aētəm*	*aētaṯ*	*aētąm*
	Род. п.	*aētahe*		*aētaŋ́hå, aētaiiå*
	Дат. п.	*aētahmāi*		—
	Отлож. п.	*aētahmāṯ*		—
	Инстр. п.	*aēta*		*aētaiia*
	Местн. п.	*aētahmi, aētahmiia*		—
Дв. ч.	Им. п.	*aēta*	*aēte*	—
	Вин. п.	*aēta*	*aēte*	—
	Род. п.	*aētaiiå*		—
Мн. ч.	Им. п.	*aēte, aētaēca*	*aēta, aētāciṯ*	*aētå*
	Вин. п.	*aētą, aētə*	*aēta, aētāciṯ*	*aētå*
	Род. п.	*aētaēšąm*		*aētaŋhąm*
	Дат. п.	*aētaēibiiō*		*aētāibiiō*
	Инстр. п.	—		—
	Местн. п.	*aētaēšu, aētaēšuua*		—

Таблица 9. Парадигма указательного местоимения III серии *аuиа-/hāu* (дальний дейксис, авестийский язык)

Число	Падеж	Мужской род	Средний род	Женский род
Ед. ч.	Им. п.	*hāu, аот*	*аuиаṯ*	*hāu*
	Вин. п.	*аот*	*аuиаṯ*	*аuиąm*
	Род. п.	*аuиаŋ́he, аuиаheса*		*аuиаiŋ́hå*
	Дат. п.	—		*аuиаiŋ́hāi*
	Отлож. п.	—		*аuиаiŋ́hāṯ*
	Инстр. п.	*аuиа*		—
	Местн. п.	—		—
Мн. ч.	Им. п.	*аuие*	*аuиа, аuиå*	*аuиå*
	Вин. п.	*аuие, +аuиə̄*	*аuиа, аuиå*	*аuиå, аuиаŋhå*
	Род. п.	*аuиаēšąm*		—
	Дат. п.	—		—
	Инстр. п.	*аuиāiš*		—
	Местн. п.	—		—

Таблица 10. Парадигма указательного местоимения I серии *iуа-/ima-/a-/anā* (ближний дейксис, древнеперсидский язык)

Число	Падеж	Мужской род	Средний род	Женский род
Ед. ч.	Им. п.	*iyam*	*ima*	*iyam*
	Вин. п.	*imam*	*ima*	*imām*
	Род. п.	—		*ahyāyā*
	Дат. п.	—		*ahyāyā*
	Отлож. п.	*anā*		
	Инстр. п.	*anā*		—
	Местн. п.	—	—	*ahyāyā*
Мн. ч.	Им. п.	*imaiy*	*imā*	*imā*
	Вин. п.	*imaiy*	*imā*	*imā*
	Род. п.	*imaišām*		—
	Дат. п.	*imaišām*		—
	Отлож. п.	*imaibiš*		—
	Инстр. п.	*imaibiš*		
	Местн. п.	—		—

Таблица 11. Парадигма указательного местоимения II серии *ava-*/*haų* (дальний дейксис, древнеперсидский язык)

Число	Падеж	Мужской род	Средний род	Женский род
Ед. ч.	Им. п.	*haų*	*ava*, *avaš-ci*	*haų*
	Вин. п.	*avam*	*ava*, *avaš-ci*	*avām*
	Род. п.	*avahyā*		—
	Дат. п.	*avahyā*		—
	Отлож. п.	*avanā*		*avanā*
	Инстр. п.	*avanā*		*avanā*
	Местн. п.	—		—
Мн. ч.	Им. п.	*avaį*	*avā*	*avā*
	Вин. п.	*avaį*	*avā*	*avā*
	Род. п.	*avaįšām*		—
	Дат. п.	*avaįšām*		—
	Отлож. п.	—		—
	Инстр. п.	—		—
	Местн. п.	—		—

Список сокращений и глоссы

1SG — первое лицо единственное число; 2SG — второе лицо единственное число; 3SG — третье лицо единственное число; 1PL — первое лицо множественное число; 2PL — второе лицо множественное число; 3PL — третье лицо множественное число; ABL — аблатив, отложительный падеж; ACC — аккузатив, винительный падеж; ACT — активный залог глагола; ADV — адвербиализатор; APRT — активное причастие; CMP — сравнительная степень; CONJ — союз; COP — копула; D — указательное местоимение / демонстратив; DAT — датив, дательный падеж; DU — двойственное число; F — женский род; GEN — генитив, родительный падеж; IMF — имперфект; IMP — императив, наклонение глагола; INDEF — неопределенное местоимение; INJ — инъюнктив (наклонение глагола), в младоавестийском выражает простое прошедшее время; INS — инструментальный (творительный) падеж; LOC — локатив, местный падеж; M — мужской род; MID — медиальный залог глагола; MPRT –медиальное причастие; N — средний род; NEG — отрицание; NOM — номинатив, именительный падеж; OPT — оптатив, желательное наклонение глагола; PF — перфект; PL — множественное число; PRS — презенс, настоящее время; PASS — пассивный залог глагола; PPRT –пассивное причастие; PVB — преверб; REL — относительное местоимение; RSN — послелог со значением причины; SG — единственное число; SUBJ — субъюнктив (конъюнктив), наклонение глагола; SUP — суперлатив, превосходная степень; VOC — вокатив, звательный падеж; ав. — авестийский язык; вин.п. — винительный падеж; дат.п. — дательный падеж; др.-инд. — древнеиндийский язык; др.-перс. — древнеперсидский язык; ед.ч. — единственное число; ж.р. — женский род; и.-е. — праиндоевропейский язык; им.п. — именительный падеж; инстр.п. — инструментальный падеж; лит. — литовский язык; местн.п. — местный падеж; мн.ч. — множественное число; м.р. — мужской род; отлож.п. — отложительный падеж; род.п. — родительный падеж; ср.р. — средний род; ст.-слав. — старославянский язык; х.-с. — хотаносакский язык.

Литература

Абаев, В. И. 2016. *Авеста. Лекции В. И. Абаева (1982–1983 гг.) с анализом и комментариями*. Цхинвал: Республика.

Виноградова, С. П. 2000. Авесты язык. В: В. С. Расторгуева и др. (ред.). *Языки мира. Иранские языки. III. Восточноиранские языки*: 10–37. Москва: Индрик.

Елизаренкова, Т. Я. 1982. *Грамматика ведийского языка*. Москва: Наука.

Ефимов, В. А. 1975. Категория рода. В: В. С. Расторгуева (ред.). *Опыт историко-типологического исследования иранских языков. Том II. Эволюция грамматических категорий*: 7–116. Москва: Наука.

Йеттмар, Карл. 1986. *Религии Гиндукуша (пер. с нем.)*. Москва: Наука.

Мейе, Антуан. 1938. *Введение в сравнительное языкознание индоевропейских языков (пер. с фр.)*. Москва–Ленинград: Государственное социально-экономическое издательство.

Молчанова, Е. К. 1975. Категория числа. В: В. С. Расторгуева (ред.). *Опыт историко-типологического исследования иранских языков. Том II. Эволюция грамматических категорий*: 200–249. Москва: Наука.

Оранский, И. М. 1979. Введение. В: В. С. Расторгуева (ред.). *Основы иранского языкознания. Древнеиранские языки*: 10–128. Москва: Наука.

Расторгуева, В. С., Д. И. Эдельман. 2000. *Этимологический словарь иранских языков. Том 1*. Москва: Восточная литература.

Семереньи, Освальд. 2002. *Введение в сравнительное языкознание (пер. с нем.)*. Москва: УРСС.

Соколов, С. Н. 1961. *Авестийский язык*. Москва: Издательство восточной литературы.

Соколов, С. Н. 1979. Язык Авесты. В: В. С. Расторгуева (ред.). *Основы иранского языкознания. Древнеиранские языки*: 129–233. Москва: Наука.

Эдельман, Д. И. 1990. *Сравнительная грамматика восточноиранских языков: Морфология. Элементы синтаксиса*. Москва: Наука.

References

Abaev, V. I. 2016. *Avesta. Lekcii V. I. Abaeva (1982–1983 gg.) s analizom i kommentarijami*. Chinval: Respublika.

Bartholomae, Christian. 1904. *Altiranisches Wörterbuch*. Strassburg: Verlag von Karl J. Trübner.

Beekes, Robert S. P. 1988. *A Grammar of Gatha-Avestan*. Leiden / New York / København / Köln: Brill.

Beekes, Robert S. P. 2011. *Comparative Indo-European Linguistics*. Amsterdam / Philadelphia: John Benjamins Publishing Company.

Brandenstein, Wilhelm, Manfred Mayrhofer. 1964. *Handbuch des Altpersichen*. Wiesbaden: Harassowitz.

Brugmann, Karl. 1904. *Die Demonstrativpronomina der indogermanischen Sprachen (eine bedeutungsgeschichtliche Untersuchung)*. Leipzig: Abhandl. Der philol.-hist. der königlich-sächsischen Gesellschaft der Wissenschaften.

Bühler, Karl. 1934. *Sprachtheorie*. *Vol. 2*. Jena: Fischer.

Cantera, Alberto. 2022. On the Edge between Literacy and Orality: Manuscripts and Performance of the Zoroastrian Long Liturgy. *Oral Tradition* 35(2): 211–250.

Debrunner, Albert, Jacob Wackernagel. 1930. *Altindische Grammatik III. Band: Nominalflexion — Zahlwort — Pronomen*. Göttingen: Vandenhoeck und Ruprecht.

Diessel, Holger. 1999. *Demonstratives. Form, function, and grammaticalization*. Amsterdam / Philadelphia: John Benjamins.

Dunkel, George E. 2014. *Lexikon der indogermanischen Partikeln*. *Band 2*. *Lexicon*. Heidelberg: Carl Winter Universitätsverlag.

Durkin-Meisterernst, Desmond. 2014. *Grammatik des Westmitteliranischen (Parthisch und Mittelpersisch)*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.

Edelman, D. I. 1990. *Sravnitel'naja grammatika vostochnoiranskix jazykov: Morfologija. Elementy sintaksisa*. Moskva: Nauka.

Efimov, V. A. 1975. Kategorija roda. In: V. S. Rastorgueva (ed.). *Opyt istoriko-tipologicheskogo issledovanija iranskix jazykov*. *Tom II. Evol'ucija grammaticheskix kategorij*: 7–116. Moskva: Nauka.

Elizarenkova, T. Ja. 1982. *Grammatika vedijskogo jazyka*. Moskva: Nauka.

Emmerick, Ronald Eric. 1968. *Saka Grammatical Studies*. London / New York / Toronto: Oxford University Press.

Gippert, Jost (ed.). 1996–2008. *TITUS Text Collection: Avest. Avesta Corpus*. Available online at: https://titus.uni-frankfurt.de/texte/etcs/iran/airan/avesta/avest.htm [Accessed 18.01.24].

Jettmar, Karl. 1986. *Religii Gindukusha* (transl. from the German). Moskva: Nauka.

Johnson, Edwin Lee. 1917. *Historical Grammar of the Ancient Persian Language*. New York: Cincinnati American Book Co.

Kapović, Mate. 2017. Proto-Indo-European morphology. In: Mate Kapović (ed.). *The Indo-European Languages*. *Second Edition*. Abingdon: Routledge.

Kellens, Jean. 1991. L'avestique de 1972 à 1990. *Kratylos* 36: 1–31.

Kent, Ronald G. 1950. *Old Persian: Grammar, Texts, Lexicon*. New Haven: American Oriental Society.

König, Ekkehard. 2020. Beyond exophoric and endophoric uses: Additional discourse functions of demonstratives. In: Åshild Næss et al. (eds.). *Demonstratives in discourse* (Topics at the Grammar-Discourse interface 6): 21–42. Berlin: Language Science Press.

Kupfer, Katharina. 2002. *Die Demonstrativpronomina im Rigveda*. Frankfurt am Main / Berlin / Bern / Bruxelles / New York / Oxford / Wien: Lang.

Lander, Eric, Liliane Haegeman. 2018. Syncretism and Containment in Spatial Deixis. In: Lena Baunaz et al. (eds.). *Exploring Nanosyntax*: 116–148. New York: Oxford University Press.

Lehmann, Winfred Philipp. 1974. *Proto-Indo-European Syntax*. Austin: University of Texas Press.

Lubotsky, Alexander. 2022. Young Avestan *pašne* and its etymology. *Acta Linguistica Petropolitana* 18–1: 216–225.

Malandra, William W. (ed.). 1983. *An introduction to ancient Iranian religion: Readings from the Avesta and Achaemenid inscriptions*. Vol. 2. Minneapolis: University of Minnesota Press.

Martínez, Javier, Michiel de Vaan. 2014. *Introduction to Avestan*. Leiden / Boston: Brill.

Mayrhofer, Manfred. 1992. *Etymologisches Wörterbuch des Altindoarischen*. I Band. Heidelberg: Carl Winter Universitätsverlag.

Meillet, Antoine. 1938. *Vvedenije v sravnitel'noje jazykoznanije indojevropejskix jazykov* (transl. from the French). Moskva / Leningrad: Gosudarstvennoje social'no-ekonomicheskoje izdatel'stvo.

Molchanova, E. K. 1975. Kategorija chisla. In: V. S. Rastorgueva (ed.). *Opyt istoriko-tipologicheskogo issledovanija iranskix jazykov. Tom II. Evol'ucija grammaticheskix kategorij*: 200–249. Moskva: Nauka.

Oranskij, I. M. 1979. Vvedenie. In: V. S. Rastorgueva (ed.). *Osnovy iranskogo jazykoznanija. Drevneiranskije jazyki*: 10–128. Moskva: Nauka.

Reichelt, Hans. 1909. *Awestisches Elementarbuch*. Heidelberg: Carl Winter Universitätsverlag.

Panaino, Antonio. 2007. Chronologia Avestica: tra cronologia linguistica e storia religiosa (Fililogia e storia del testo avestico, 1.) In: Antonio Panaino et al. (eds.). *Disputationes iranologicae Vindobonenses* I: 7–34. Wien: Verlag der Österreichischen Akademie der Wissenschaften.

Peirot, Michaël. 2017. Slavic *онъ*, Lithuanian *anàs* and Tocharian A *anac*, *anäṣ*. In: Bjarne Simmelkjær Sandgaard Hansen et al. (eds.). *Usque ad radices. Indo-European studies in honour of Birgit Anette Olsen*: 633–642. Copenhagen: Museum Tusculanum Press.

Plank, Frans. 2016. *The uniquely suppletive inflection of the Indo-European demonstrative pronoun *to-/*so- 'that', and how pertinacious deficits of neuters help to explain it*. Talk given at the Berkeley Germanic Linguistics Roundtable. Berkeley: University of California.

Rastorgujeva, V. S., D. I. Edel'man. 2000. *Etimologicheskij slovar' iranskix jazykov. Tom 1*. Moskva: Vostochnaja literatura.

Schmitt, Rüdiger. 2009. *Die altpersischen Inschriften der Achaimeniden: Editio minor mit deutscher Übersetzung*. Wiesbaden: Reichert.

Schmitt, Rüdiger. 2014. *Wörterbuch der altpersischen Königsinschriften*. Wiesbaden: Reichert.

Szemerényi, Oszvald. 2002. *Vvedenije v sravnitel'noje jazykoznanije* (transl. from the German). Moskva: URSS.

Sims-Williams, Nicholas. 1994. The triple system of deixis in Sogdian. *Transactions of the Philological Society*. Vol. 92(l): 41–53.

Skjærvø, Prods Oktor. 2009. Old Iranian. In: Gernot Windfuhr (ed.). *Iranian languages*: 43–195. Abingdon / New York: Routlege.

Sokolov, S. N. 1961. *Avestijskij jazyk*. Moskva: Izdatel'stvo vostochnoj literatury.

Sokolov, S. N. 1979. Jazyk Avesty. In: V. S. Rastorgueva (ed.). *Osnovy iranskogo jazykoznanija. Drevneiranskie jazyki*: 129–233. Moskva: Nauka.

de Vaan, Michiel, Alexander Lubotsky. 2012. Old Persian. In: Holger Gzella (ed.). *Languages from the World of the Bible*: 194–208. Boston / Berlin: Walter de Gruyter.

Vinogradova, S. P. 2000. Avesty jazyk. In: V. S. Rastorgueva et al. (eds.). *Jazyki mira. Iranskie jazyki. III. Vostochnoiranskie jazyki*: 10–37. Moskva: Indrik.

Visser, Eline. 2020. Tracking and recognitional use of Kalamang opa: Demonstrative of cognitive accessibility. In: Åshild Næss et al. (eds.). *Demonstratives in discourse* (*Topics at the Grammar-Discourse interface 6*): 103–122. Berlin: Language Science Press.

Wackernagel, Jacob. 2010. *Vorlesungen über Syntax mit besonderer Berücksichtigung von Griechisch, Lateinisch und Deutsch 1920–1924. Vol. 2*. Cambridge: Cambridge University Press.

Watkins, Calvert. 1995. *How to kill a dragon: aspects of Indo-European poetics*. Oxford / New York: Oxford University Press.

Wendtland, Antje. 2009. The Position of the Pamir Languages within East Iranian. *Orientalia Suecana* LVIII: 172–188.

Yakubovich, Ilya. 2021. Hittite Agent Form. In: A. I. Kogan et al. (eds.). *Problemy obshchej i vostokovednoj lingvistiki. Sochetaemost' jazykovyx edinic i jazykovyje modeli. Pam'ati Z. M. Shal'apinoj (1946–2020)*: 124–141. Moscow: Institute of Oriental Studies.

Artyom Badeev. Demonstrative pronouns in Old Iranian languages: main types of usage and the reconstruction of the Proto-Iranian system

In this paper I discuss the use of demonstratives in Old Iranian languages. On the basis of the deictic semantics of demonstratives, the boundaries of suppletive paradigms are defined. For the first time, the correspondence of each paradigm to a particular type of deixis is established on Old Iranian data. Using this data, as well as data from other Iranian languages and its comparison with Old Indian, I present a systematic reconstruction of the Proto-Iranian system of demonstratives.

Keywords: Iranian languages; Old Iranian languages; Avestan language; Old Persian language; deixis; demonstratives; Proto-Iranian language.

Issue 2 / Часть 2

Articles / Статьи

В. Н. Харьков[*], *А. Ю. Урманчиева*[†], *Н. А. Колесников*[‡], *А. А. Зарубин*[§],
Л. В. Валихова[**], *И. Ю. Хитринская*[††], *И. А. Грунтов*[‡‡], *И. М. Егоров*[§§],
В. М. Лемская[***], *А. В. Дыбо*[†††], *В. А. Степанов*[‡‡‡]

[*] Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; vladimir-kharkov@medgenetics.ru

[†] Институт лингвистических исследований РАН; urmanna@yandex.ru

[‡] Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; nikita.kolesnikov@medgenetics.ru

[§] Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; aleksei.zarubin@medgenetics.ru

[**] Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; larisa.valikhova@medgenetics.ru

[††] Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; khitrinskaya@tnimc.ru)

[‡‡] Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; Институт языкознания РАН, Москва; grountov@mail.ru

[§§] Институт финно-угроведения/уралистики Университета Людвига Максимилиана, Мюнхен;, i.m.jegorow@gmail.com

[***] ФГБОУ ВО «Томский государственный педагогический университет», Томск; lemskaya@tspu.edu.ru

[†††] Институт языкознания РАН, Москва; adybo@mail.ru

[‡‡‡] Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; vadim.stepanov@medgenetics.ru

Популяционная генетика и лингвистическая компаративистика: филогения и филогеография гаплогруппы N1a2b и лингвистическая предыстория самодийцев[1]

В работе проведено сопоставление данных этногенетики по гаплогруппе N1a2b Y-хромосомы и лингвистики самодийской языковой группы. Гаплогруппа N1a2b унаследована современными народами от более раннего населения уралоидного антропологического типа, заселявшего территорию Южной Сибири. При детальной реконструкции филогении этой гаплогруппы показана этноспецифичная структура сублиний N1a2b. Лингвистическая предыстория самодийцев реконструируется путем анализа и стратификации лексических заимствований. Релевантные для данной статьи результаты по лингвистической филогенетике получены методом ручного анализа общих инноваций (для самодийских языков) и байесовской лексикостатистики (для тюркских языков). Между результатами генетического и лингвистического анализа обнаружены следующие соответствия. Регион начального распространения гаплогруппы N1a2b в целом соотносится с прародиной самодийцев, установленной по лингвистическим данным. Специфическая близость лексики маторского и ненецких языков находит подтверждение в истории сублинии N1a2b1b1. Эта сублиния представлена у тундровых ненцев (но не у энцев) и у хакасов качинцев, на язык которых как раз и могли перейти маторы. Наличие сублинии N1a2b1b1 у казымских хантов объясняется их вторичными контактами с ненцами, о чем также свидетельствуют ненецкие культурные заимствования. Сублиния N1a2b1b2a1 обнаруживается у тувинцев и монголов, что коррелирует с тувино-монгольскими языковыми контактами. Полученные современные совместные результаты данных популяционной генетики и лингвистики хорошо дополняют друг

[1] Исследование выполнено за счет гранта Российского научного фонда № 22-64-00060, https://rscf.ru/project/22-64-00060/.

друга и дают достоверное подтверждение языковому родству и генетико-демографическим процессам в популяциях Сибири.

Ключевые слова: этногенетика; Y-хромосома; лингвистическая предыстория; самодийские народы.

Введение

Анализ структуры генофондов популяций человека с использованием диаллельных и микросателлитных маркеров Y-хромосомы является одним из наиболее современных и результативных методов изучения генетического разнообразия популяций человека, позволяет проводить комплексную оценку популяционной истории, выполнять филогенетические реконструкции отдельных линий, филогеографические исследования гаплогрупп, анализ миграционных процессов. Несмотря на бурное развитие методов, позволяющих изучать генофонды популяций на уровне полных геномов, использование маркеров Y-хромосомы по-прежнему является актуальным для решения целого ряда вопросов популяционной генетики (Underhill et al. 2000; Jobling & Tyler-Smith 2003; Karmin et al. 2015; Watanabe et al. 2019; Esteller-Cucala et al. 2023).

Территория Южной и Западной Сибири — уникальный регион с точки зрения разнообразия естественно-географических условий и этнокультурного многообразия проживающего здесь населения. Одним из достаточно древних и относительно широко распространенных на этой территории по данным антропологии, этнологии и лингвистики является самодийский компонент, который связан с одной из основных миграционных волн в Южной, Западной Сибири и Арктике.

Современные народы, говорящие на самодийских языках, занимают значительное пространство: от Белого моря до Таймыра на севере, а также среднее и нижнее Приобье. В настоящее время это лесные и тундровые ненцы, лесные и тундровые энцы, нганасаны и селькупы, разделенные на несколько территориальных групп. Еще относительно недавно некоторые самодийские этносы или субэтносы (камасинцы, маторы, койбалы, карагасы, тайгийцы, кашинцы) проживали в Саянах. Сейчас они полностью растворились в основном тюркском населении (Дульзон 1960). Поскольку эти самодийцы сыграли существенную роль в формировании тюркских этносов саянского региона, генетические и лингвистические данные, относящиеся к популяциям Южной Сибири, необходимо рассматривать в комплексе для самодийских и тюркских народов.

Проблема корректного соотнесения данных генетики и лингвистики, их объединения в комплексный междисциплинарный подход является чрезвычайно актуальной, поскольку именно попытки сопоставления генетических и лингвистических данных зачастую вызывают пристальное внимание специалистов широкого круга направлений, связанных с реконструкцией истории населения различных регионов.

Процессы ассимиляции и популяционной дифференциации также сыграли свою роль в формировании современных народов Сибири. Самодийцы относятся к уральскому антропологическому типу, который по своим характеристикам занимает промежуточное положение между монголоидной и европеоидной расами (Рогинский & Левин 1978). Внутри самодийской группы различия по антропологическому типу сильно варьируются. К настоящему времени исследования, предполагающие выяснение соотношения лингвистической и генетической истории самодийцев, проводились на ограниченном материале и с не очень высокой точностью (Волков 2013; Karafet et al. 2018). Все самодийские языки можно разделить на две группы: северосамодийские (тундровый

и лесной ненецкий, тундровый и лесной энецкий, нганасанский, а также маторский) и южносамодийские (селькупский язык, представленный пятью диалектами, а также камасинский).

Многие сибирские этносы антропологически и генетически довольно гетерогенны (например, алтайцы, тувинцы, хакасы, ненцы и эвенки), а исследованные до настоящего времени выборки не отражают всей многокомпонентности генофонда, характеризуя лишь часть имеющегося разнообразия. Исследование направлено на анализ состава генофондов современного коренного населения Южной и Западной Сибири.

Исследование состава генофондов современных популяций человека по различным генетическим маркерным системам, реконструкция их происхождения, скоррелированные с современными представлениями об истории языков данных этносов, имеют важное значение для комплексного понимания формирования различных этно-территориальных групп. Проблематика, касающаяся анализа структуры и соотношения различных генетических компонент у народов Сибири, генеалогического ветвления их языков, а также межъязыковых заимствований, как материальных, так и структурных, несмотря на относительно высокий уровень изученности, имеет ряд невыясненных вопросов. Генетика в этой связи дает важные результаты для междисциплинарного исследования, поскольку развитие современных методов анализа популяционного генофонда позволило вывести этногенетические исследования на совершенно новый уровень.

История этнических групп, населяющих Сибирь, была насыщена различными событиями и претерпела изменения популяционного, языкового и культурного состава за счет миграций различных популяций, относящихся к разным языковым семьям и группам (енисейской, самодийской, угорской, тюркской, монгольской, тунгусо-маньчжурской, юкагирской и др.). Генетико-демографические процессы в популяциях, колебания численности, метисационные события, миграции и естественный отбор оказывают влияние на структурированность генетического разнообразия в геномах отдельных индивидов и популяциях в целом.

Для маркеров Y-хромосомы показан наибольший уровень генетических отличий между популяциями в сравнении с любыми другими генетическими системами. Это приводит к значительной степени территориальной и межэтнической генетической дифференциации и высокому уровню специфичности гаплотипов. Высокие показатели межпопуляционного и межгруппового разнообразия Y-хромосомы делают ее маркеры наиболее дифференцирующими из всех, применяемых в популяционной генетике и геномике человека, и подчеркивают эффективность анализа гаплотипов Y-хромосомы для выявления различий между популяциями и при изучении миграций (Dupanloup et al. 2003; Barbujani et al. 2003). Кроме характеристики генетического разнообразия, дифференциации и состава популяционного генофонда, исследование маркеров Y-хромосомы очень важно для филогенетических реконструкций различных гаплогрупп.

Гаплогруппа N1a2b (ранее обозначавшаяся как N2 и N1b), неравномерно распределенная в популяциях Сибири и Восточной Европы, представляет собой маркер доисторической связи между сибирскими и европейскими популяциями. Она связана с древней общностью населения уралоидного антропологического типа, заселявшего территории к востоку и западу от Уральских гор. Наибольшая частота гаплогруппы N1a2b приходится на территорию Западной Сибири и Алтае-Саянского региона (Zerjal et al. 1997; Karmin et al. 2015; Ilumäe et al. 2016; Karafet et al. 2018). Именно эта гаплогруппа доминирует у всех этносов, принадлежащих к группе самодийских языков уральской языковой семьи, в Западной Сибири — 92% у нганасан, 78% у энцев и 74% у тундровых ненцев (Karafet et al. 2002; Kharkov et al. 2021). Ее частота в Южной Сибири, в частности,

у хакасов и тувинцев, также отражает вклад в их современный генофонд самодийских этносов, заселявших эту территорию ранее (см. Таблицу 1).

Распространение N1a2b на территории Европы связано с расселением предков современных финно-угорских народов, поскольку данная гаплогруппа характерна для этих народов Урала и Поволжья, где она составляет более 20% в некоторых популяциях (удмурты, коми) (Tambets et al. 2004; Karmin et al. 2015). Присутствие гаплогруппы N1a2b у татар, башкир и русских связано с включением в состав этих народов финно-угорского генетического субстрата. Западной границей ареала этой гаплогруппы является Финляндия, где она обнаруживается у единичных образцов (Lappalainen et al. 2006; Rootsi et al. 2007). Предположительным местом возникновения гаплогруппы N1a2b является Южная Сибирь.

В диахронической лингвистике существует две группы методов, дающих результаты, релевантные для истории этнических групп и подходящие для непосредственного сравнения с результатами других наук, в первую очередь популяционной генетики, археологии и антропологии. Первая группа методов связана с изучением дивергентных процессов, установлением языкового родства и построением филогенетических деревьев. Вторая группа методов сосредотачивается вокруг изучения конвергентных процессов, т. е. заимствований и эффектов ареальной диффузии.

Филогенетические деревья в сравнительно-историческом языкознании бывают двух типов: качественные и количественные. Качественное дерево строится на основании выявленных в родственных языках/диалектах процессов фонетических, морфологических и лексических изменений. Последовательность прохождения этих процессов может быть упорядочена по времени на основании результатов, наблюдаемых в каждом из языков. Сопоставление последовательностей дает возможность определить, какие из процессов проходили в общем состоянии этих языков, а какие могли идти уже только после разделения общего состояния на отдельные группы и далее на отдельные языки. Такое дерево отражает относительную хронологию языковых изменений, но допускает абсолютную датировку только по внешним по отношению к процедуре данным (письменные памятники, заимствования между языками с письменной традицией, лексика, связанная с датированными технологиями и т. п.). Количественные деревья строятся лексикостатистическим методом, основанным на количественных сходствах фрагментов базового словаря сравниваемых языков. В пределах базисной лексики предполагается относительно равномерная скорость внутренних замен, что допускает абсолютную датировку (Старостин 2007). Лексикостатистические матрицы обрабатываются алгоритмами, практически идентичными с теми, что используются для филогенетического анализа на основе ДНК. Лексикостатистические методы могут быть применены к любым достаточно документированным языкам/диалектам, поскольку основаны на анализе универсальных наборов концептов. Их основное ограничение состоит в том, что они игнорируют различие между синапоморфией и симплезиоморфией.

Взаимодействие этнических групп в прошлом может быть реконструировано на основании заимствований. Заимствоваться могут практически любые элементы языковой системы, но наиболее надежно идентифицируются лексические заимствования. Лексические заимствования могут стратифицироваться с привлечением сведений о топологии филогенетического древа и относительной хронологии фонетических процессов. Количество заимствований и распределение по семантическим группам дает сведения об интенсивности и характере взаимодействий этнических групп. Различаются адстратные и субстратные заимствования. Адстратные заимствования указывают на культурный обмен и торговые связи между этническими группами. Они охватывают небольшое коли

Таблица 1. Частоты встречаемости сублиний гаплогруппы N1a2b Y-хромосомы у коренных народов Сибири

Популяции	N1a2b1 (B478, Z35080 xB169)	N1a2b1b (B169 xB175)	N1a2b1b~ (B478, Z35080 xB169)	N1a2b1b1 (Y68212)	N1a2b1b1 (VL65, Z35102)	N1a2b1b1a ~ (B170 xZ35104)	N1a2b1b1 b (B172)	N1a2b1b2b1 (B178, PF3415)	N1a2b2a2 a~ (B478 xB169)	N1a2b2b1 ~ (Z35076)	N1a2b1b2a1 ~ (B228, Z35125)
Ненцы	8,3 %	-	-	-	-	29,4 %	12,6 %	-	-	-	-
Ханты	-	-	-	30,6 %	-	-	3,3 %	-	-	2,6 %	-
Хакасы качинцы	-	-	-	92,0 %	-	-	-	-	-	-	-
Чулымцы	-	-	-	-	52,6 %	-	-	-	-	-	-
Хакасы сагайцы	-	-	-	-	-	-	-	-	37,9 %	-	-
Шорцы	-	-	-	-	-	-	-	-	11,1 %	-	-
Челканцы	-	-	-	-	-	-	-	-	20,0 %	-	-
Тувинцы	-	2,4%	-	-	-	-	-	23,9%	-	-	0,9%
Южные алтайцы	-	-	-	-	-	-	-	1,5%	-	-	-
Калмыки	-	0,6 %	-	-	-	-	-	2,5 %	-	-	-
Якуты	-	-	0,7 %	-	-	-	-	-	-	-	-
Буряты	-	-	0,4 %	-	-	-	-	-	-	-	-

чество семантических полей и редко проникают в базисную лексику. Субстратные заимствования указывают, что в истории данной популяции имел место языковой сдвиг, т. е. переход с одного языка на другой. Они охватывают многие семантические поля, в том числе проникая и в базисную лексику; могут иметь характерные фонотактические паттерны, контрастирующие с унаследованной от праязыка лексикой. Данные топонимики также часто могут указывать на наличие субстрата (Дульзон 1950: 175–183; Дульзон 1953; Дульзон 1954: 59–94; Дульзон 1973).

С точки зрения генетики при распаде популяции и расширении кластера гаплотипов в одной из дочерних популяций этот кластер будет ограничен только этой дочерней популяцией (при условии, что взаимосвязь между популяциями незначительна). Следовательно, оценка его датировки будет определяться не самим разделением (т. е. расхождением популяции), а событием расширения кластера, которое могло произойти в любой момент эволюции дочерней популяции, во время или после разделения.

С позиции лингвистики инновации, которые обнаруживаются в одном языке/диалекте, но отсутствуют в другом, возникают после физического разделения протопопуляции носителей языка на частично изолированные субпопуляции. Лингвистическая датировка разделения, таким образом, указывает на момент времени, когда языковые различия были закреплены, а не на возникновение этого процесса. Следовательно, начало дивергенции у населения могло иметь место за несколько поколений до времени, определяемого по лингвистическим данным.

Целью настоящего исследования является взаимная верификация результатов, которая становится возможна благодаря тому, что и генетические, и лингвистические исследования позволяют строить филогенетические деревья и датировать точки отделения ветвей (Balanovsky et al. 2011; Mallory et al. 2019). Вычисление временны́х промежутков по данным лингвистики основано на подсчете расхождений, образовавшихся после разделения языков выборки. Аналогичным образом, временны́е отрезки по данным генетики Y-хромосомы рассчитываются исходя из количества мутаций, возникших после начала расширения гаплогруппы.

В работе выявлен вклад самодийского генетического субстрата на основе генотипирования различных SNP и STR маркеров Y-хромосомы, филогенетического и филогеографического анализа линии N1a2b Y-хромосомы, маркирующей самодийский генетический субстрат. На основании данных о популяционно-специфичной структуре гаплогрупп Y-хромосомы выполнена детальная характеристика и реконструкция формирования современного генетического разнообразия исследуемых популяций, что позволяет точнее реконструировать историю формирования генофондов современных сибирских народов.

Экспериментальная часть

М а т е р и а л ы. В работе использовались образцы ДНК мужчин из различных популяций. Забор первичного биологического материала (венозной крови) у доноров производили с соблюдением процедуры письменного информированного согласия на проведение исследования. На каждого донора составлялась анкета с краткой родословной, указанием этнической принадлежности и мест рождения предков. Индивида относили к данной этнической группе на основании его собственной этнической идентификации, его родителей и места рождения. Для анализа YSTR-гаплотипов гаплогруппы N1a2 были использованы 592 образца ДНК мужчин принадлежащих к ее основной азиатской

ветви: ненцев (N=256), хантов (N=48), хакасов сагайцев (N=88), хакасов качинцев (N=43), чулымских тюрок (N=21), шорцев (N=11), тувинцев (N=110), якутов (N=2), бурятов (N=1), челканцев (N=3) и калмыков (N=9).

Методы исследования гаплотипов Y-хромосомы. Для изучения состава и структуры гаплогрупп Y-хромосомы в исследование были включены две системы генетических маркеров: диаллельных локусов, представленных SNP, и полиаллельных высоковариабельных микросателлитов (YSTR). С помощью 445 SNP-маркеров определяли принадлежность образцов к различным сублиниям N1a2b. Классификация гаплогрупп дана в соответствии с данными Международного общества генетической генеалогии (ISOGG). Анализ STR-гаплотипов внутри гаплогрупп проводили с использованием 46 микросателлитных маркеров нерекомбинирующей части Y-хромосомы (YSTR) (DYS19, 385a, 385b, 388, 389I, 389II, 390, 391, 392, 393, 426, 434, 435, 436, 437, 438, 439, 442, 444, 445, 448, 449, 456, 458, 460, 461, 481, 504, 505, 518, 525, 531, 533, 537, 549, 552, 570, 576, 635, 643, YCAIIa, YCAIIb, GATAH4.1, Y-GATA-A10, GGAAT1B07, DYF387; Kharkov et al. 2020). STR-маркеры генотипировали с помощью капиллярного электрофореза на генетическом анализаторе ABI Prism 3730 и Нанофор-05. Генотипирование SNP-маркеров проводили с помощью ПЦР и последующего анализа фрагментов ДНК с помощью ПДРФ-анализа. Экспериментальные исследования проведены на базе Центра коллективного пользования научно-исследовательским оборудованием «Медицинская геномика» (НИИ медицинской генетики Томского НИМЦ). Построение медианных сетей гаплотипов Y-хромосомы проводили с использованием программы Network v. 10.2.0.0 (www.fluxus-engineering.com) по методу медианных сетей Бандельта (Bandelt et al. 1999). Оценку возраста генерации наблюдаемого разнообразия гаплотипов в гаплогруппах проводили методом ASD (Zhivotovsky et al. 2004) на основании средних квадратичных отличий в числе повторов между всеми маркерами. Возраст поколения принимался равным 30 годам, темп мутирования 0,0021 на локус на поколение.

Полногеномное секвенирование мужчин с гаплогруппой N1a2b выполнено методом массового параллельного секвенирования NGS (Next Generation Sequencing) на 84 образцах ДНК, принадлежащих к азиатской линии N1a2b. Секвенирование с покрытием x30 проводили на приборе Illumina NovaSeq. Выполнен биоинформатический анализ первичных данных секвенирования геномов, включающий оценку качества полученных данных секвенирования для проанализированных образцов. Анализ данных полногеномного секвенирования образцов показал, что качество сгенерированных прочтений было достаточно высоким. Эти геномные данные были включены в базу данных для ее расширения и дальнейшей биоинформатической обработки общего массива образцов. Анализ и извлечение генотипов из данных секвенирования проводились с использованием Genome Analysis Toolkit 4 (https://software.broadinstitute.org/gatk).

Для привязки генотипов образцов к уже описанным SNP и обнаружения новых мутантных позиций было применено два подхода:

1) фильтрация данных с пропущенными в различных образцах позициями, уровнем покрытия и ошибочных аллелей в одной позиции. К сожалению, в результате этого подхода было отсеяно очень большое количество SNP, включенных в список, особенно это касалось многих терминальных SNP. Это удаляло также и много не описанных ранее SNP, которые очень важны для более точной детализации филогении различных линий;

2) после этого были удалены все позиции без мутаций в файлах этих гаплогрупп. Фильтрация данных в автоматическом режиме не проводилась. Далее эти данные в табличном виде были перенесены в Excel и начат отбор мутантных позиций и отсев оши-

бочных вручную на всем наборе образцов мужчин, принадлежащих к этим гаплогруппам. Это позволило оставить практически все привязанные к базе SNP, обозначаемые различными шифрами (A, AM, B, BY, C, CTS, F, FGC, FT, L, M, MF, VL, S, SUR, TY, Y, YP, Z, ZS), а также существенно увеличить количество новых YSNP.

Полученные после отбора данные по позициям YSNP для всех секвенированных образцов были использованы для создания отдельных vcf-файлов с достоверными мутантными позициями. С помощью программ Y-LineageTracker и StarBeast3 были построены филогенетические деревья для N1a2 на уровне отдельных образцов (см. рисунки 1, 2 в онлайн-приложении к настоящей работе). Построение филогенетических деревьев с использованием пакета программ BEAST основано на использовании байесовской статистики (метода Монте-Карло для марковских цепей), вовлекающей сложный математический аппарат. Эти деревья более точно отображают разделение различных ветвей внутри гаплогрупп. Все построенные деревья полностью согласуются с количеством YSNP в отдельных линиях и демонстрируют разделение образцов на этноспецифичные группы.

Подробная филогения гаплогруппы N1a2b построена по результатам NGS-секвенирования полных геномов различных образцов мужчин. Данные по Y-хромосомным SNP получены после биоинформатической обработки данных, привязки описанных ранее мутаций к указанным позициям и отбора новых терминальных SNP для более точного разделения сублиний внутри этой гаплогруппы. Детальное разделение на отдельные ветви показывает специфичность разных линий для всех исследованных этносов, субэтносов и популяций.

Для гаплогруппы N1a2b по всем секвенированным образцам была выполнена точная привязка к детальным этноспецифичным вариантам отдельных сублиний Y-хромосомы, за счет принадлежности к уже описанным мутантным SNP. У многих образцов мужчин из этих гаплогрупп, относящихся к сибирским популяциям, было обнаружено большое количество новых, не описанных ранее SNP. В общем массиве маркеров они составляют около 20%, но для различных образцов их доля в терминальных SNP порой доходит до 100%. Эти результаты позволяют гораздо точнее разделить филогенетическую структуру этой линии, детализировать их происхождение.

Лингвистические методы. Лингвистический подход включал в себя исследование заимствования лексики между различными уральскими идиомами (тундровый ненецкий и казымский хантыйский) методами сравнительно-исторического языкознания и контактной лингвистики (Бурлак & Старостин 2005; Campbell 2013), анализ положения маторского языка внутри самодийской подгруппы путем изучения общих инноваций (Campbell 2013), объединяющих маторский язык с северносамодийскими и южносамодийскими языками, анализ языковых контактов тюрков и самодийцев, тюрков и енисейцев, тюрков (тувинцев и чулымцев) и монголов, что позволяет датировать контакты, а также получать данные о древних миграциях носителей этих языков (Epps 2015).

Результаты и обсуждение

При определении прародины гаплогрупп необходимо использовать несколько критериев. Во-первых, одним из критерев является высокая частота гаплогруппы, указывающая на возможное место ее происхождения. Но максимальная частота не может служить однозначным показателем центра зарождения гаплогруппы, поскольку и на

периферии ареала гаплогруппы популяции могут испытывать сильное влияние дрейфа, что может привести к эффекту основателя по данной линии у конкретного этноса или группы популяций. Во-вторых, наличие предковой по отношению к исследуемой линии и близкородственных сестринских гаплогрупп также указывает на прародину. В большинстве случаев местом возникновения гаплогруппы может быть регион с высокой частотой базальных ветвей этой клады. Третьим критерием является разнообразие YSTR-гаплотипов в пределах изучаемой гаплогруппы. Территорию с наибольшим разнообразием какой-то родственной группы гаплотипов принято рассматривать как вероятную зону их общего происхождения. Там, где гаплогруппа впервые достигла значимой частоты и откуда она затем расширяла свой ареал, должны были накопиться многочисленные мутации и сформироваться множество дочерних гаплотипов. В месте происхождения показатели разнообразия должны быть максимальны. В отдаленных популяциях гаплотипическое разнообразие должно быть мало, поскольку с миграционными потоками они получают не репрезентативную выборку гаплотипов предковой популяции, а лишь случайную часть спектра.

Построенное филогенетическое древо азиатской гаплогруппы N1a2b (рисунок 3 в онлайн-приложении к настоящей работе) показывает разделение образцов на различные ветви в соответствии с теми специфичными сублиниями, которые определяются терминальными SNP. Близость секвенированных образцов мужчин, принадлежащих к этой гаплогруппе, полностью совпадает с их этнической привязкой и данными их анкетирования.

Сублинии гаплогруппы N1a2b в популяциях Алтае-Саян и лингвистические данные

Самая древняя ветвь этой линии присутствует в популяциях Алтая и Саян N1a2b2a2a~ (B478, Z35080, Z35081, Z35082, Z35083, Z35084, YP3079 (xZ35088,B168,B169)). К ней принадлежат образцы челканцев, хакасов-сагайцев и шорцев. Большинство из них формируют попарные ветви, с различными терминальными SNP, объединяющие эти образцы в соответствии с их этническим происхождением. Относительно небольшая частота этой сублинии в популяциях северных алтайцев обусловлена тем, что этот базовый субстратный самодийский компонент был достаточно сильно замещён в течение последних полутора тысяч лет тюркскими и монгольскими группами, но его присутствие в генофонде современных популяций южно-сибирских этносов демонстрируют наличие вклада древнего самодийского населения. Присутствие этой линии с этноспецифичными терминальными SNP и отличающимися гаплотипами говорит о их наследовании от достаточно древних общих предков, входивших в состав раннего населения этих территорий.

Близость образцов, принадлежащих к этнически родственным популяциям, может быть сопоставлена с лингвистической классификацией. Генетические образцы, для которых можно предполагать самодийский субстрат, были взяты у носителей тюркских языков.

В Южной Сибири в исторически засвидетельствованный период представлены следующие группы тюркских языков (связанные с затронутыми в исследовании популяциями): саянская (тувинский с диалектами и тофаларский); з-группа (среднечулымский, мрасский диалект шорского, хакасский с диалектами, а именно кызыльским, качинским, сагайским, шорским); алтайская, делящаяся на южно-алтайскую подгруппу (алтай-кижи, теленгитский, телеутский), и северно-алтайскую (кумандинский, челканский, ту-

баларский; возможно, также кондомский диалект шорского). Алтайская группа, по последним данным (Egorov et al. 2022), представляет собой раннее ответвление кыпчакской группы. С антропологической точки зрения традиционная культура алтайцев больше связана со степью, чем культура саянской и з-групп, так что алтайцы могут быть более новыми пришельцами в Южную Сибирь, чем носители первых двух групп тюркских языков. Датировка появления тюрков в Южной Сибири связывается обычно с алтайским походом Маодуня (шаньюй сюнну, 201 г. до н.э., см. напр., Хелимский 2000). Неплохо датируются, например, контакты тюрков с енисейцами. Если ориентироваться на реконструкцию праенисейской фонетики в Старостин 1982 и праенисейский лексикон, реконструированный в Старостин 1995, а также расширенный и учитывающий новейшую литературу вариант его, опубликованный в интернете по адресу starlingdb.org (ToB), то распад праенисейского языка датируется глоттохронологически II в. до н.э., когда выделяются две ветви — «кетская» и «коттская». От «кетской» отделяется в V в. н.э. пумпокольский язык, затем в XII в. н.э. распадаются кетский и югский языки. Коттская группа в VI в. н.э. распадается на коттский и аринский.

Соответственно выделяется три пласта старых тюркских заимствований в енисейские: (1) пратюркские заимствования в праенисейский (не позднее рубежа эры); (2) заимствования, происшедшие не позднее VI в. н.э. (только в «коттской» ветви); (3) заимствования, происшедшие не позднее XII в. н.э. (только в кетско-югской подветви), см. Дыбо 2007. Это предполагает по крайней мере три волны контактов — и, возможно, три волны переселений.

Сублинии гаплогруппы N1a2b у ненцев

Вторая специфичная сублиния N1a2b1 (B478, Z35080, Z35081, Z35082, Z35083, Z35084 (xZ35088,B169)), являющаяся чуть менее древней, обнаружена у тундровых ненцев восточной части ЯНАО. К ней принадлежат исключительно образцы ненцев, которые по мужской линии имеют принадлежность к энецким родам, вошедшим в состав тундровых ненцев. Эта линия является наследием более раннего населения тундровой территории, которое появилось там до прихода предков ненцев, и эта более ранняя миграционная волна самодийцев обладала другим составом сублиний и гаплотипов. Ещё одна очень специфическая сублиния N1a2b1b~ обнаружена у двух якутов и одного бурята.

Сублинии гаплогруппы N1a2b у тувинцев

Следующая ветвь присутствует у тувинцев, якутского эвенка и калмыков. У тувинцев она разделена на три сублинии, отличающиеся по терминальным SNP и кластерам гаплотипов. С максимальной частотой у тувинцев представлен ее вариант N1a2b1b2b1 (B178, PF3415, Z35147, Z35149, Z35152). К нему относятся два образца южных алтайцев. Её возраст составляет 1442 года (SD=368 лет). Четыре образца тувинцев принадлежат к очень редкой параллельной линии N1a2b1b2a1~ (B228, Z35125, Z35127, Z35128). Ранее она была обнаружена у монголов (Ilumäe et al. 2016). К третьей тувинской сублинии (xB175, Z35117, Z35118) относятся 10 образцов. Монголы и тувинцы находились в интенсивных контактах, как минимум, с XIII века. В XIV–XVII веках наблюдалось сильное влияние ойратов (ближайших родственников современных калмыков), позже преобладало влияние халхамонгольского языка. Заимствования шли, главным образом, в одном направлении — из монгольского языка в тувинский. Тувинцы жили, фактически, в ситуации билингвизма, что отразилось в значительном объеме лексики, заимствованной из монгольских языков

(более 1500 лексем по оценке в Khabtagaeva 2006). Интенсивность контактов и старый билингвизм, сохраняющийся и до настоящего времени, подтверждаются проникновением монгольских заимствований в базисную лексику. В различных диалектах тобаской группы монголизмы отмечены в следующих слотах 110-словной версии списка Сводеша: ‘все’, ‘земля’, ‘зеленый’, ‘дождь’, ‘песок’, ‘семя’, ‘маленький’, ‘хвост’, ‘женщина’. Что касается самодийских контактов языков саянской группы (тувинского и тофаларского), то известно существенное количество тюркизмов, определяемых как саянские, в матор-тайги-карагасской группе, и небольшое количество заимствований из матор-тайги-карагасского в тувинском, больше в тофаларском (см. ниже).

Сублинии гаплогруппы N1a2b у чулымских тюрков

Следующая специфичная линия выявлена у чулымских тюрков (VL65, Z35095, Z35099, Z35102) и является наиболее близкой к сестринской ей линии N1a2b1b1 у хакасов качинцев. Медианная сеть гаплотипов (см. рисунок 4 в онлайн-приложении) демонстрирует у чулымцев звездообразную филогению с недавним эффектом основателя и доминированием по частоте предкового гаплотипа. Кластер чулымских гаплотипов равноудален от всех качинских сеоков. Возраст этого кластера у чулымцев составил 667 лет (SD=194 года), что достаточно хорошо совпадает с предположительным временем расселения предков чулымцев на север.

При исследовании истории тюркского населения среднего течения р. Чулым А. П. Дульзон обращался прежде всего к данным топонимики (Дульзон 1950), а также номенклатуре родства и свойства (Дульзон 1954) в сравнении с материалом языков других народов Сибири. Лингвистические данные, обработанные А. П. Дульзоном, указывают на наличие субстратов в языке чулымских тюрков — прежде всего енисейского (кетского) и самодийского (селькупского). Распространение тюркского языка в среднем течении р. Чулым, с точки зрения А. П. Дульзона (Дульзон 1973), происходило в три этапа: начиная с VII–VIII вв. с верховьев реки по направлению вниз; в XII–XIII вв. с юга вниз по р. Кие; не позднее XV в. — с юго-запада вверх по р. Чулыму (либо вниз по р. Яе). По данным интерпретации гидронимов среднего Причулымья, тюркизация этой территории произошла путем наслоения на три кетоязычные народности (ассаны, пумпоколы, арины; Дульзон 1973). Помимо этого, на основе комплекса этнолингвистических и археологических данных отмечается «преемственная связь причулымского населения» с селькупами (Дульзон 1953; Дульзон 1954).

Следует отметить наблюдения о чулымско-тюркских монголизмах. В. И. Рассадин (1980) на материале, собранном В. В. Радловым (нижнечулымские диалекты) и А. П. Дульзоном (диалектная принадлежность не указывается), устанавливает пласт монгольской лексики в чулымско-тюркском языке. Вопрос о монгольском субстрате в данном случае не ставился; пласт монголизмов представляет собой единичные лексические единицы (не более 50), проникшие в чулымский язык в XIII-XIV вв., либо, возможно, через более поздние контакты с джунгарскими ойратами (Рассадин 1980).

Сублинии гаплогруппы N1a2b у хантов и ненцев

Все остальные исследованные образцы разделяются на две крупные субклады. К первой из них (N1a2b1b1-Y68212, Y70717, Y70315, Y70327) относятся ханты и хакасы-качинцы, которые разделились между собой достаточно давно. Эта сублиния хантов близка к хакасам-качинцам (Z35093, Z35097, Z35103; Valikhova et al. 2022). К гаплогруппе N1a2b2b1~

(Z35076) относятся три образца казымских хантов. Медианная сеть гаплотипов (см. рисунок 5 в онлайн-приложении) также демонстрирует у хантов звездообразную филогению с недавним эффектом основателя и преобладанием по частоте предкового гаплотипа. Возраст этого кластера у хантов составил 858 лет (SD=338 лет).

Вторая сублиния также разделяется на две ветви. Первая N1a2b1b1a~ (B170, B171, Z35091, Z35092, Z35093, Z35095, Z35096, Z35097, Z35099, Z35102, Z35103, Z35105 и др. (xZ35104, Z35106, Z35107)). К ней принадлежат практически подавляющее большинство ненцев фратрии Харючи и многие образцы, относящиеся к фратрии Вануйто. Вторая сублиния N1a2b1b1b (B172, Z35108, Z35109, Z35112 и др.) также присутствует у ненцев и хантов, которая унаследована ими от ненецких предков по мужской линии. К ней принадлежат все ненцы из фратрии Вануйто, относящиеся к родам Вануйто, Пуйко и Яунгат, и хантыйского по происхождению рода Пурунгуй.

Структура этого филогенетического древа и соотношение возраста отдельных сублиний хорошо согласуется с данными антропологии и лингвистики, которые описывают территорию Саян как исходную территорию расселения самодийских групп.

Филогения гаплогруппы N1a2b и лингвистическая история самодийцев

На данном филогенетическом дереве с точки зрения лингвистической истории самодийцев наиболее примечательной представляется линия N1a2b1b1. Как можно видеть по схеме, представленной на Рисунке 3 «Разделение сублиний гаплогруппы N1a2b Азиатской ветви по этноспецифичным терминальным SNP», эта линия представлена у ненцев, хантов и хакасов-качинцев. Эта линия обнаружена у казымских хантов, что, вероятно, можно объяснить их контактами с тундровыми ненцами. Лингвистически эти контакты отражаются в достаточно значительном пласте культурной и ландшафтной лексики, заимствованной из тундрового ненецкого в северные диалекты мансийского и хантыйского языков (заимствований в хантыйский при этом значительно больше). Основной корпус этих заимствований был выделен еще В. Штейницем (Steinitz 1959). В статье А. Ю. Урманчиевой (Урманчиева 2021) показано, что эти заимствования сохраняют целый ряд архаичных черт фонетики тундрового ненецкого. Историческая фонетика обско-угорских языков указывает на то, что эти заимствования произошли не позднее XV века.

Приведем заимствования из ненецкого в казымский хантыйский: т. нен. *xabt°* ‘кастрированный олень’ > каз. хант. *χɔpti*; т. нен. *xabt°rka* ‘бесплодная важенка, которую дрессируют для езды’ > каз. хант. *xɔptərka*; т. нен. *maly°cya* ‘малица’ > каз. хант. *mɔlʲ‘śi* [иногда: *mɔlʲəpśi*]; т. нен. *nyada* ‘ягель’ > каз. хант. *ńɔta*; т. нен. *pan°* ‘нижний край (например, нюка); подол одежды’ > каз. хант. *pɔn̩*; т. нен. *parka* ‘парка’ > каз. хант. *pərχa*; т. нен. *tati* ‘младшая жена (при многожёнстве)’ > каз. хант. *tɔtti*; т. нен. *syaqmey°* ‘менструация’ > каз. хант. *śɔm‘ŋa*; т. нен. *ŋaw°ka* ‘авка’ > каз. хант. *awkka*; т. нен. *mador* ‘группа, толпа, стадо, стая’ > каз. хант. *matər*; т. нен. *syan°* ‘две плоские костяшки, скреплённые ремнём на узде оленьей упряжки’ > каз. хант. *śaṇəw, -ŭw*; т. нен. *syar* ‘кожа (под волосами) > каз. хант. *śar*, обд. хант. *śár* ‘верхний слой шкуры’; т.нен. *nyamsyer* ‘невозмутимый, хладнокровный (о человеке)’ > каз. хант. *ńŏmža*; т. нен. *nyəb°lyuy°* ‘неблюй (осенний телёнок); телячья шкура, снятая примерно в сентябре, пригодная для малицы’ > каз. хант. *ńŏpl̩əw, -l̩ŭw*; т. нен. *yədalabcy°h* ‘скребок для выделки шкур’ > каз. хант. *jŏtʲl̩əpši*; т. нен. *nyəw°di°* ‘телёнок годовалой важенки; внебрачный ребёнок’ > каз. хант. *ńŏwti*; т. нен. *nəmna* ‘лоншак, олень-самец по второму году’ > каз. хант. *nŏmna*; т. нен. *məlk°* ‘комолый’ > каз. хант. *mŏlχa*; т. нен. *nyulək°* ‘мягкий’ > каз. хант. *ńŭl̩ak* пух; т. нен. *xuŋkər°* ‘выемка между ягодицами (у человека)’ > каз. хант. *šăš-χǫŋχart* ‘на крупе идущее от спины вниз углубление, выемка в крестце

(и спине) лошади'; т. нен. *tíncyah* (*h* < *ŋ*) 'аркан для поимки оленей' > каз. хант. *tǐńśaŋ*; т. нен. *túcyah* (*h* < *ŋ*) 'женская сумочка для швейных принадлежностей (из оленьей шкуры или цветного сукна)' > каз. хант. *tǔťśaŋ*; т. нен. *syobya* 'капюшон (малицы или совика)' > каз. хант. *śǫma*; т. нен. *myuy°ko* 'поднючье (нижняя зимняя покрышка чума мехом внутрь)' > каз. хант. *mŏjəkkǫ*.

Таким образом, представленность линии N1a2b1b1 у казымских хантов можно объяснить вторичными контактами с тундровыми ненцами. Тогда эта линия остается представленной у тундровых ненцев и хакасов-качинцев, причем парадоксальным образом тундровые ненцы на рассматриваемом филогенетическом дереве оказываются теснее связаны с хакасами-качинцами, чем даже с энцами. Но, возможно, лингвистические данные помогут объяснить этот парадокс. Традиционно ближайшими лингвистическими родственниками тундрового ненецкого считались северносамодийские языки: ближайший к тундровому лесной ненецкий, достаточно близкие к обоим ненецким лесной и тундровый энецкий и занимающий достаточно обособленную позицию в северносамодийской подгруппе нганасанский. Но, как показано в работе Helimski 1997, к северносамодийским, несмотря на свое крайне южное положение, относился также маторско-тайгийско-карагасский язык. Е. А. Хелимский проверил всю известную маторскую лексику самодийского происхождения методом сепаратных изоглосс, т. е. он выделял такие лингвистические параллели, которые связывают маторский либо с южносамодийскими (селькупским и/или камасинским), либо с северносамодийскими (ненецким и/или энецким и/или нганасанским). Маторско-южносамодийских лексических изоглосс, по его подсчетам — 38, маторско-северносамодийских — 97 (Helimski 1997). На основании этого Е. А. Хелимский приходит к выводу о том, что маторский относится к северносамодийским языкам, к ненецко-энецкой подгруппе.

В работе Урманчиева 2023 проведено дополнительное исследование маторско-северносамодийских связей. Был несколько изменен метод: не только определялись лексические параллели, связывающие маторский с южно- либо с северносамодийскими, но также оценивалось, должны ли выделенные параллели рассматриваться как сохранение общего архаизма (что может быть случайным явлением) либо как общая инновация (поскольку инновации являются ключевыми для определения подгруппы в лингвистической классификации). В таблице 2 ниже учитываются только такие параллели, которые могут претендовать на статус общих инноваций (из числа которых удалены однозначно вторичные параллели контактного происхождения). Параллели, которые сложно охарактеризовать как архаизмы/инновации и общие архаизмы из подсчетов исключаются.

Таблица 2. Лексические инновации, связывающие маторский с южно- и северносамодийскими языками

	общие инновации
– с камасинским и селькупским	4
– с селькупским	4
– с камасинским	3
– всего с южносамодийскими	*11*
– с ненецким, энецким и нганасанским	14
– с ненецким и энецким	12
– с энецким	0
– с ненецким	**8**
– всего с северносамодийскими	*34*

Как видно из этой таблицы, число параллелей, которые могут трактоваться как общие генетические инновации, у маторского существенно (в три раза) выше с северносамодийскими, чем с южносамодийскими. Е.А. Хелимский отмечает бóльшую близость маторского к энецко-ненецкой подгруппе (Helimski 2022). Но интересно, что внутри этой подгруппы маторский демонстрирует отчетливую близость с ненецким (строка таблицы, выделенная полужирным): иными словами, ненецкий и маторский разошлись позже, чем ненецкий и энецкий. Если теперь мы предположим, что среди предков современных хакасов-качинцев были носители маторского, настолько же в генетическом отношении близкие к ненцам, насколько маторский близок к ненецкому в лингвистическом отношении, это позволит нам объяснить, почему на полученном методами генетики филогенетическом дереве хакасы-качинцы близки к тундровым ненцам. Естественно, это достаточно прямолинейная и упрощенная картина: говоря о «носителях маторского», мы должны иметь в виду скорее какой-то из диалектов маторско-тайгийско-карагасского континуума — возможно, даже тот, о котором не сохранилось вообще никаких сведений. Но принципиально картину можно очертить именно таким образом, как это сделано выше.

Медианные сети YSTR-гаплотипов мужчин, относящихся к этой кладе, демонстрируют значительные этнические различия и специфичность спектра гаплотипов для отдельных этносов и субэтнических групп. Практически все основные этно-специфичные сублинии показывают звездообразную филогению этих гаплотипов, которые демонстрируют эффект основателя и во многих случаях достаточно активный демографический рост предковых групп. Наиболее сильный эффект основателя выявлен у ненцев, что связано с переходом их на тундровое оленеводство и значительным увеличением их численности из поколения в поколение. У хакасов присутствуют две разные сублинии, которые связаны с базовым самодийским населением, разделённым территориально и вошедшим в состав сагайцев и качинцев. Южные сагайцы из Аскизского и Таштыпского районов демонстрирует генетическую близость к шорцам и челканцам. Качинцы и чулымцы также достаточно близки друг другу, но более удалены от хакасов-сагайцев.

Что касается качинцев, следует отметить, что генетический материал забирался только у северных качинцев, носителей так называемого белоиюсского говора качинского языка. Именно эта группа демонстрирует в языковом отношении (как, впрочем, и по фамилиям) сильное влияние кызыльского адстрата (а, может быть, и субстрата): наличие заимствований из кызыльского диалекта с переходом **ni-* > *i-*, презенса с *odur-* и под. Так что, возможно, эти генетические данные касаются кызыльцев и чулымцев, а не качинцев и чулымцев. Южные сагайцы (прежде всего бельтыры), напротив, и в языковом, и в этнологическом отношении связаны с шорцами, поскольку с XVII в. зафиксированы повторяющиеся миграции шорцев на территорию нынешней южной Хакасии. Что касается языковых контактов с различными самодийскими группами, то лингвистам известны следующие факты (Дыбо 2007).

1. Установлено как минимум 19 достаточно надежных заимствований из пратюркского в прасамодийский, т. е. отражающихся в северосамодийских языках: ПСам. **jemńə-* ‘латать’ < ОТ **jama-* ‘латать, штопать’, ПСам. **jęsəj-* ‘ставить чум’ < ОТ **jasa-* ‘строить, устраивать’, ПСам. **kępu* ‘оса’ < ОТ **Kapuŋ* (?) ‘шмель’, ПСам. **päjmå* ‘сапог’ < ПТ **baʎmak* ‘сапог, обувь’, ПСам. **kåŋ* ‘правитель’ < ПТ **kān* ‘правитель’, ПСам. **ta(ə)j* ‘ветвь, сук’ < ОТ **dal* ‘ветка, ива’, ПСам. **pât-* ‘погружаться, заходить (о солнце)’ < ПТ **bat-* ‘погружаться, тонуть, заходить (о солнце)’, ПСам. **jekâ* ‘близнец’ < ПТ **(h)ęjkiř* ‘близнецы’, ПСам. **wekänä*/**wekɜrɜ* ‘осетр’ < ПТ **bEkre* ‘вид

осетровой рыбы', ПСам. **jokə̂-* (? **jok-*) 'sich verirren' < ПТ **jōk* 'нет', **jōk-ad-* (отыменной глагол), **jōk-a-l-* (рефлексив отыменного глагола) 'потеряться, исчезнуть', ПСам. **tet²tə̂* (? **tettə̂*) 'четыре' < ПТ **dȫrt* 'четыре', ПСам. **juntз* 'лошадь' < ОТ **junt* 'лошадь, кобыла', ПСам. **jür* 'сто' < ПТ **jü̈ř* 'сто', ПСам. **jür* 'жир' < ПТ **ü̈ř* 'жир', ПСам. **ken* 'ножны' < ПТ **kījn* 'ножны', ПСам. **kį̄r* 'седой, светлый, белый' < ОТ **Kī̆r* 'серый, седой', ПСам. **ki*, **kilз* 'соболь' < ПТ **kīλ* 'соболь', ПСам. **kə̂psə̂* 'шаманская ложка для окроплений' < ПТ **Kamič* 'черпак'. Заимствования из прасамодийского в пратюркский более сомнительны и малочисленны, их насчитывается всего четыре: ОТ **K(i)aja* 'скала, гора' < ПСам. **koəjə* 'гора, водораздел', ОТ **kās* 'кора' < ПСам. **käsa* 'кора'; ОТ **tī̆t* 'лиственница/кедр' < ПСам. **tį̄tę̄ŋ* 'кедр'; ОТ **bün* (**büjn*?) < ПСам. **wę̄ńз* 'мясной или рыбный суп, бульон'. Это должны быть очень ранние контакты, до распада северных и южных самодийцев, датируемого приблизительно рубежом эры.

2. Относительно ранние, т. е. не из современных тюркских языков Южной Сибири, заимствования в матор-тайги-карагасский (Хелимский 2000); обсуждение см. в Дыбо 2007, с возможной передатировкой в "современные" — 3 vs 2 слова из гипотетического пра-южносамодийского во все тюркские языки Южной Сибири: тув. *kada* 'тетя', 'бабушка, свекровь', *kadaj* 'жена, старуха, баба', тоф. *qadaj*, качин. *qadaj* < ЮСам. < ПСам **kåtå* 'бабушка'; тув. *dazɨl*; шор. *tazɨl*; галт. *t.azɨl*; туба *tazɨl*; куу *tazɨl* 'корень' < ЮСам **tasV* < ПСам **tę̄nsə ~ *cę̄nsə* (> ПССам **tę̄nsə*); МТК **üd8* 'корова, крупный рогатый скот' < тув. *ud* 'вол, бык'; МТК **kaj* 'гусь' < ПТ **Kāř* 'гусь'; Камас.-койб. *boi* 'лед' < ОТ **būz* 'лед'.
3. Относительно современные заимствования в матор-тайги-карагасский и в камасинский (там, где это можно определить — в матор-тайги-карагасский из языков саянской группы, в основном тофаларского, и языков z-группы, в основном хакасских диалектов, в том числе сагайского; в камасинский — из языков з-группы, хакасских и шорских диалектов, в том числе качинского и кызыльского) — всего больше 500 (Joki 1952; Helimski 1997). Относительно современные заимствования из МТК в тувинский и тофаларский — 3: тув. *ire* 'дед', *irek* 'дед, старик', тоф. *ire* 'дедушка, отец мужа, прадед, предок по муж. линии, медведь' < матор. *irei* 'медведь' < ПСам. **iräj* 'дед, старик'; тоф. *tarhiš* 'годовалый олененок (неодобрительное название)' < ПСам **cərkəj-* 'олененок'; тоф. *čoɣdi̯* этноним < матор. *čagədi* 'старый'. Из камасинского в з-группу — 4: хакас. *сидік*, саг. *седік* 'тяжелый, трудный' < камас. *sed-em* тж. < ПСам. **seč-* тж. (Helimski 1995); хак. кач. (XIX в.) *мугнэ* 'росомаха' < койб. камас. *мугна* < ПСам. **wiŋkəncä*; хак. койб. (уст.) *ка* 'лось' < койб. камас. *ка* < ПСам. **kå*; хак. кач. (XIX в.) *мятъ* 'выдра' < койб. камас. *мятъ* < ПСам. **töt*. Ср. еще кальку: кызыл. *харға айы* март *(букв.* вороны месяц) — камас. *bārik'i* 'первый месяц весны' (*bāri* ворона).

Данных по языковым контактам с самодийцами, выделяющих разные зоны контакта для кызыльцев (или белоиюсских качинцев; + чулымцы) и шорцев, пока не обнаруживается. Исторически (по картам) видно, что маторы могли контактировать тесно с койбалами, которые, по материалам Кастрена, несомненно, хакасы и по некоторым параметрам ближе к качинцам. Северные качинцы и кызыльцы в засвидетельствованное время контактируют с камасинцами. Структурные изоглоссы с маторским также не показывают разделения «север — юг» для диалектов з-группы; например, специфически маторская депалатализация начального **ń-* происходит в кызыльском, качинском, сагайском и шорском, не затрагивает чулымский; специфически маторский процесс *č->s-/_Vs* нахо-

дит свою параллель в кызыльском, но там это совсем новый процесс, идущий на протяжении XX в., и с несколько иными условиями распределения. С другой стороны, например, данные по истории распространения показателей творительного падежа в хакасских диалектах (Дыбо & Мальцева 2021) позволяют предположить миграцию койбалов-качинцев в зону контакта с кызыльцами в течение первой половины XX века. Учитывая отсутствие генетических данных по усть-абаканским, староиюсским качинцам и кызыльцам, следует признать, что как лингвистические, так и генетические данные нуждаются в дальнейшем исследовании.

Генотипирование новых высокоинформативных SNP-маркеров показывает, что в пределах гаплогруппы N1a2b1 Y-хромосомы наблюдается разделение не только на этноспецифические линии, но и на более узкие сублинии и кластеры гаплотипов, вплоть до индивидуального уровня. Большинство этнических генофондов характеризуются преобладанием или даже полным доминированием специфичных SNP по этой гаплогруппе при использовании нескольких сотен YSNP. Это позволит подробно и очень точно описать территориальную и межэтническую генетическую дифференциацию по специфичности различий YSNP и YSTR-гаплотипам.

Присутствие у всех исследованных популяций разных этноспецифичных вариантов гаплогруппы N1a2b подтверждает отсутствие миграций носителей этой гаплогруппы и обмена генами на протяжении последних нескольких сотен лет. Основным фактором разделения этой гаплогруппы на отдельные сублинии явилась генетическая изоляция разных предковых самодийских групп коренного населения и их дальнейший демографический рост.

Значительная частота этой гаплогруппы в Южной Сибири, в частности, у хакасов, тувинцев и северных алтайцев, вероятнее всего отражает вклад в их генофонд самодийских этносов, заселявших эту территорию ранее. Очевидно, в Алтае-Саянском регионе они не были вытеснены или истреблены более поздними мигрантами, а ассимилированы и включены в состав современных тюркоязычных групп. При расселении на север предков нганасан, энцев и ненцев происходило их смешение с древним аборигенным населением и увеличение численности людей в популяциях при смене поколений.

С точки зрения лингвистики наблюдается изменение состава лексики в записях словарных материалов, сделанных на Саянах с XVIII по середину XIX в. В начале этого периода мы видим самодийские словари с вкраплением тюркских заимствований, в конце — тюркские словари с единичными самодийскими заимствованиями (Дыбо & Мальцева 2021).

Местом начала распространения на территории Сибири гаплогруппы N1a2b является регион Алтая и Саян. Территория, соответствующая современному ареалу тувинцев, алтайцев и хакасов, была первичным очагом генерации разнообразия и экспансии численности предковой группы носителей гаплогруппы N1a2b. По мере расселения носителей гаплогруппы сформировались вторичные очаги генерации разнообразия и экспансии численности — на территории Западной Сибири и ЯНАО. По лингвистическим данным прародина самодийцев находилась примерно в этом же регионе. Часть самодийцев мигрировала оттуда на север около рубежа нашей эры. В позднейшее время оставшееся в Саяно-Алтайском регионе самодийское население было ассимилировано сибирскими тюркскими народами.

Расширенный состав терминальных SNP для выявленных гаплогрупп позволил детально охарактеризовать и проанализировать различия в гаплотипической структуре отдельных этноспецифичных сублиний, определить их родство, следы экспансии численности в генофонде исследованных популяций.

Заключение

В рамках исследования на основе данных о полных геномах и структуре сублиний N1a2b Y-хромосомы выполнено детальное описание самодийского компонента в генофонде современных популяций Южной и Западной Сибири, оценка его вклада и реконструкция формирования современного генетического разнообразия исследуемых популяций, что позволяет точнее охарактеризовать историю генофондов современных сибирских народов. Получены новые результаты сопоставления данных этногенетики и лингвистики.

Гаплогруппа N1a2b маркирует древнюю популяционную общность населения уралоидного антропологического типа, заселявшего территории к востоку от Уральских гор. Распространение этой гаплогруппы на территории Западной и Южной Сибири связано с расселением самодийцев. Наибольшая частота в популяциях северных самодийцев — нганасан, ненцев и энцев служит этому прямым подтверждением. Полученные результаты позволяют по-новому взглянуть на историю формирования ненцев, энцев, хантов, хакасов, шорцев, чулымских тюрков, тувинцев, северных алтайцев, телеутов, сибирских татар и других народов. Детально охарактеризована филогения гаплогруппы N1a2b Y-хромосомы, являющейся основой генофонда многих этносов Сибири.

Филогенетический анализ Y-хромосомных сублиний и гаплотипов доказывает, что центром происхождения и расселения носителей самодийского компонента в Южной и Западной Сибири является территория современного Алтая и Саян. Полученные результаты хорошо согласуются с данными лингвистики, этнологии и антропологии о вкладе самодийского компонента в формирование различных народов Алтая-Саян и историческими ареалами самодийских и других языков уральской языковой семьи.

Обнаружение большого количества SNP-маркеров, которые определяют филогенетически недавние ветви для различных гаплогрупп Y-хромосомы и данные об их распространении в современных популяциях, позволили провести отбор специфичных маркеров для популяций. Использование таких высокоинформативных этноспецифичных SNP-маркеров Y-хромосомы является одним из наиболее перспективных инструментов для анализа региональных и этнических популяционных генофондов на уровне субэтносов, родов и локальных территориальных групп.

Все исследованные выборки характеризуются преобладанием и даже полным доминированием специфичных SNP по гаплогруппе N1a2b. Сопоставление проводится в первую очередь по терминальным SNP для различных сублиний и консервативным, более медленно мутирующим STR-маркерам, которые обладают наибольшей этнодифференцирующей информативностью.

Корректность полученного методами генетики филогенетического дерева, отражающего взаимную близость популяций, была верифицирована путем сопоставления с выводами об этнической истории данных этносов, полученными лингвистическими методами. Это сопоставление позволило подтвердить в нескольких важных аспектах правильность соотнесения гаплогруппы N1a2b с самодийским компонентом. Как минимум некоторые из разделений данной гаплогруппы на сублинии коррелируют с современными представлениями о сложных процессах смены языка, адстратных и субстратных влияний, оставляющих свой след в языке: на основе таких специфических лингвистических маркеров реконструируются различные контактные ситуации, в которые в разные периоды своей истории были вовлечены носители самодийских языков. При этом необходимо отметить, что данное междисциплинарное исследование показало *взаимную* верификацию выводов: для лингвистов столь же важно, что реконструируемые ими про-

цессы находят подтверждение в результатах, полученных методами других наук. Перечислим вкратце наиболее существенные корреляции между данными генетики и лингвистики:

- регион, в котором в Сибири началось распространение гаплогруппы N1a2b, в целом близок к той зоне, которую лингвисты очерчивают как прародину самодийцев, то есть как ту территорию, на которой жили носители реконструируемого самодийского праязыка и на которой начались процессы его диалектной дивергенции;
- наиболее загадочно с лингвистической точки зрения положение маторского языка в группе самодийских. Этот язык, угасший в XIX веке, географически был самым южным из самодийских языков (северные предгорья Саян), однако сравнительно-исторические исследования маторской лексики, проводившиеся в последние десятилетия, показывают, что он принадлежит к группе северносамодийских языков. Е. А. Хелимский (Helimski 1997) делает вывод о том, что маторский относится к ненецко-энецкой подгруппе; в диссертации А. Ю. Урманчиевой (Урманчиева 2023) делается попытка обосновать особую близость внутри этой подгруппы маторского к ненецкому. Однако лингвистический материал по маторскому достаточно скуден: мы располагаем только списками слов (в донаучной записи, которая требует тщательной и не всегда возможной лингвистической интерпретации), материалов по грамматике этого языка до нас практически не дошло, так как его носители были ассимилированы тюркоязычными народами данного региона. В этой ситуации (слишком смелые выводы, сделанные на ограниченном материале) очень важно получить независимое подтверждение лингвистической гипотезы. Прямое подтверждение мы находим в истории сублинии N1a2b1b1: она представлена у тундровых ненцев и у хакасов качинцев (на язык которых как раз и могли перейти маторы), но не у энцев;
- наличие сублинии N1a2b1b1 у казымских хантов объясняется их вторичными контактами с ненцами, что отражается в многочисленных культурных заимствованиях из ненецкого в северные диалекты хантыйского;
- на примере детального разбора фрагмента филогенетического дерева — поиска лингвистических корреляций истории сублинии N1a2b1b1 — мы показали перспективность таких междисциплинарных исследований. Валидность этой методики подтверждена также тем, что сублиния N1a2b1b2a1, обнаруженная у тувинцев и монголов, находит соответствие в истории тувинско-монгольских языковых контактов. Несомненно, столь же многообещающие результаты ожидаются и в отношении истории других этносов, говорящих или говоривших на самодийских языках. Однако это требует существенно более полного охвата популяций Сибири генетическими обследованиями. Построенное на современном этапе генетическими методами филогенетическое дерево заведомо неполно: на нем отсутствует информация по нганасанам, лесным ненцам, селькупам, информация по носителям енисейских языков, поэтому крайне желательным было бы расширение исследовательской базы.

Полученные результаты не противоречат общепринятым версиям этногенеза самодийских народов, но позволяют более детально описать этот процесс. Информация о структуре генофонда популяций Южной и Западной Сибири по Y-хромосомным сублиниям является важным дополнением к существующим лингвистическим, а также антропологическим и археологическим данным о их происхождении, формировании, миграциях, демографических процессах и родственных связях с другими народами.

Онлайн-приложение к статье (доступно по адресу jolr.ru/jlr22/kharkov.zip) содержит следующие иллюстративные материалы:

Рисунок 1. Филогенетическое древо гаплогруппы N1a2 (StarBeast3).

Рисунок 2. Филогенетическое древо гаплогруппы N1a2 (Y-LineageTracker).

Рисунок 3. Разделение сублиний гаплогруппы N1a2b азиатской ветви по этноспецифичным терминальным SNP.

Рисунок 4. Медианная сеть YSTR-гаплотипов гаплогруппы N1a2b1 у чулымцев.

Рисунок 5. Медианная сеть YSTR-гаплотипов гаплогруппы N1a2b1b1 у хантов.

Список сокращений

арин. — аринский; галт. — горноалтайский (алтай-кижи); каз. — казымский (диалект хантыйского); камас. — камасинский; кач. — качинский (диалект хакасского); койб. — койбальский; кот. — коттский; матор. — маторский; МК — язык(и) словаря Махмуда Кашгарского; МТК — маторско-тайгийско-карагасский; огуз. — огузский; обд. — обдорский (диалект хантыйского); ОТ — общетюркский; ПЕн. — праенисейский; ПСам. — прасамодийский; ПТ — пратюркский; саг. — сагайский (диалект хакасского); т.нен. — тундровый ненецкий; тоф. — тофаларский; тув. — тувинский; хак. — хакасский; хант. — хантыйский; шор. — шорский.

Литература

Бурлак, С. А., С. А. Старостин. 2005. *Сравнительно-историческое языкознание*. Москва: Издательский центр «Академия».

Волков, В. Г. 2013. Древние миграции самодийцев и енисейцев в свете генетических данных. *Томский журнал лингвистических и антропологических исследований* 1(1): 79–96.

Дульзон, А. П. 1950. Древние смены народов на территории Томской области по данным топонимики. *Ученые записки Томского государственного педагогического института* VI: 175–187.

Дульзон, А. П. 1953. Поздние археологические памятники и проблема происхождения чулымских татар. *Ученые записки Томского государственного педагогического института* X: 127–133.

Дульзон, А. П. 1954. Термины родства и свойства в языках Нарымского края и Причулымья. *Ученые записки Томского государственного педагогического института* XI: 59–94.

Дульзон, А. П. 1960. Этнический состав древнего населения Западной Сибири по данным топонимики. В кн.: *Доклады делегации СССР: XXV Международный конгресс востоковедов:* 1–10. Москва: Изд-во Вост. лит.

Дульзон, А. П. 1973. Диалекты и говоры тюрков Чулыма. *Советская тюркология* 2: 16–29.

Дыбо, А. В. 2007. *Лингвистические контакты ранних тюрков. Лексический фонд: пратюркский период.* Москва: Восточная литература.

Дыбо, А. В., В. С. Мальцева. 2021. Хакасские диалекты: конец XIX в. и начало XXI в. *Oriental Studies* 14(6): 1353–1367.

Рассадин, В. И. 1980. Монголизмы в чулымско-тюркском языке В кн.: И. Д. Бураев (ред.). *Монголо-бурятские заимствования в сибирских тюркских языках:* 48–51. Москва: Наука.

Рогинский, Я. Я., М. Г. Левин. 1978. *Антропология: Учебник для студентов университетов. 3-е изд*. Москва: Высшая школа.

Старостин, С. А. 2007. *Труды по языкознанию.* Москва: Языки славянских культур.

Старостин, С. А. 1982. Праенисейская реконструкция и внешние связи енисейских языков. В кн.: Е. А. Алексеенко (ред.). *Кетский сборник. Антропология, этнография, мифология, лингвистика:* 144–237. Ленинград: Наука.

Старостин, С. А. 1995. Сравнительный словарь енисейских языков. В кн.: С. А. Старостин (ред.). *Кетский сборник. Лингвистика*: 176-315. Москва: Языки русской культуры.

Урманчиева, А. Ю. 2021. Фонетические особенности ненецких заимствований в обско-угорских языках. *Уралоалтайские исследования* 40(1): 101–123.

Урманчиева, А. Ю. 2023. *Реконструкция лингвистического ландшафта Западной Сибири (на материале самодийских языков): диссертация на соискание ученой степени доктора филологических наук*. Санкт-Петербург.

Хелимский, Е. А. 2000. *Компаративистика, уралистика. Лекции и статьи*. Москва: Языки русской культуры.

References

Balanovsky, O., K. Dibirova, A. Dybo, O. Mudrak, S. Frolova, E. Pocheshkhova, M. Haber, D. Platt, T. Schurr, W. Haak et al. 2011. Parallel Evolution of Genes and Languages in the Caucasus Region. *Molecular Biology and Evolution* 28(10): 2905–2920.

Bandelt, H.-J., P. Forster, A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16(1): 37–48.

Barbujani, G. et al. 2003. A recent shift from polygyny to monogamy in humans is suggested by the analysis of worldwide Y-chromosome diversity. *Journal of Molecular Evolution* 57: 85–97.

Burlak, S. A., S. A. Starostin. 2005. *Sravnitel'no-istoricheskoje jazykoznanije*. Moskva: Izdatel'skij centr «Akademija».

Campbell, Lyle. 2013. *Historical linguistics: an introduction. 3rd edition*. Edinburgh University Press.

Dul'zon, A. P. 1950. Drevnije smeny narodov na territorii Tomskoj oblasti po dannym toponimiki. *Uchenyje zapiski Tomskogo gosudarstvennogo pedagogicheskogo instituta* VI: 175–187.

Dul'zon, A. P. 1953. Pozdnije arxeologicheskije pam'atniki i problema proisxozhdenija chulymskix tatar. *Uchenyje zapiski Tomskogo gosudarstvennogo pedagogicheskogo instituta* X: 127–133.

Dul'zon, A. P. 1954. Terminy rodstva i svojstva v jazykax Narymskogo kraja i Prichulym'ja. *Uchenyje zapiski Tomskogo gosudarstvennogo pedagogicheskogo instituta* XI: 59–94.

Dul'zon, A. P. 1960. Etnicheskij sostav drevnego naselenija Zapadnoj Sibiri po dannym toponimiki. In: *Doklady delegacii SSSR: XXV Mezhdunarodnyj kongress vostokovedov*: 1–10. Moskva: Izdatel'stvo vostochnoj literatury.

Dul'zon, A. P. 1973. Dialekty i govory t'urkov Chulyma. *Sovetskaja t'urkologija* 2: 16–29.

Dupanloup, I., L. Pereira, G. Bertorelle, F. Calafell, M. J. Prata, A. Amorim, G. A. Barbujani. 2003. A Recent Shift from Polygyny to Monogamy in Humans Is Suggested by the Analysis of Worldwide Y-Chromosome Diversity. *Journal of Molecular Evolution* 57(1): 85–97.

Dybo, Anna V. 2007. *Lingvisticheskije kontakty rannix t'urkov. Leksicheskij fond: prat'urkskij period*. Moskva: Vostochnaja literatura RAN.

Dybo, Anna V., Vera S. Mal'ceva 2021. Hakasskije dialekty: konec XIX v. i nachalo XXI v.. *Oriental Studies* 14(6): 1353–1367.

Egorov, I. M., A. V. Dybo, A. S. Kassian. 2022. Phylogeny of the Turkic Languages Inferred from Basic Vocabulary: Limitations of the Lexicostatistical Methods in an Intensive Contact Situation. *Journal of Language Evolution* 7(1): 16–39.

Epps, Patience. 2015. Historical linguistics and socio-cultural reconstruction. In: Claire Bowern, Bethwyn Evans (eds.). *The Routledge handbook of historical linguistics*: 579–597. London / New York: Routledge.

Esteller-Cucala, P., M. Palmada-Flores, L. F. K. Kuderna, C. Fontsere, A. Serres-Armero, M. Dabad, M. Torralvo, A. Faella, L. Ferrández-Peral, L. Llovera et al. 2023. Y chromosome sequence and epigenomic reconstruction across human populations. *Communications Biology* 6(1): 623.

Helimski, Eugen. 1995. Laut- und Wortgeschichte der Türksprachen. *Turkologica* 26: 75–95.

Helimski, Eugen. 1997. *Die matorische Sprache. Studia uralo-altaica 41.* University of Szeged.

Helimski, Eugen. 2022. Zur Stellung des Matorischen innerhalb der samojedischen Sprachen. In: Valentin Gusev, Anna Urmanchieva, Aleksandr Anikin (eds.). *Siberica et Uralica: In memoriam Eugen Helimski* (*Studia uralo-altaica 56*): 479–495. University of Szeged.

Helimskij, E. A. 2000. *Komparativistika, uralistika. Lekcii i stat'ji.* Moskva: Jazyki russkoj kul'tury.

Ilumäe, A. M., M. Reidla, M. Chukhryaeva, M. Järve, H. Post, M. Karmin, L. Saag, A. Agdzhoyan, A. Kushniarevich, S. Litvinov et al. 2016. Human Y Chromosome Haplogroup N: A Non-trivial Time-Resolved Phylogeography that Cuts across Language Families. *The American Journal of Human Genetics* 99(1): 163–173.

ISOGG = *International Society of Genetic Genealogy*. On-line resource: www.isogg.org. Accessed on 07.12.2023.

Jobling, M. A., C. Tyler-Smith. 2003. The human Y chromosome: an evolutionary marker comes of age. *Nature Reviews Genetics* 4: 598–612.

Joki, Aulis J. 1952. *Die Lehnwörter des Sajansamojedischen. Mémoires de la Société Finno-Ougrienne 103.* Helsinki: Suomalais-Ugrilainen Seura.

Karafet, T. M., L. P. Osipova, M. A. Gubina, O. L. Posukh, S. L. Zegura, M. F. Hammer. 2002. High levels of Y-chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life. *Human Biology* 74(6): 761–789.

Karafet, T. M., L. P. Osipova, O. V. Savina, B. Hallmark, M. F. Hammer. 2018. Siberian genetic diversity reveals complex origins of the Samoyedic-speaking populations. *American Journal of Human Biology* 30(6): e23194.

Karmin, M., L. Saag, M. Vicente, M. A. Wilson Sayres, M. Järve, U. G. Talas, S. Rootsi, A. M. Ilumäe, R. Mägi, M. Mitt et al. 2015. A recent bottleneck of Y chromosome diversity coincides with a global change in culture *Genome Research* 25(4): 459–466.

Khabtagaeva, Baiarma. 2006. The Intensity of Mongolian Influence in the Tuva Language. *Acta Orientalia* 59(1): 85–94.

Kharkov, V. N., L. M. Novikova, I. Y. Khitrinskaya, V. G. Volkov, V. A. Stepanov, O. V. Shtygasheva, F. A. Luzina. 2020. Gene Pool of Khakass and Shors for Y Chromosome Markers: Common Components and Tribal Genetic Structure. *Russian Journal of Genetics* 56(7): 849–855.

Kharkov, V. N., L. V. Valikhova, E. L. Yakovleva, V. N. Serebrova, N. A. Kolesnikov, T. I. Petelina, I. Yu. Khitrinskaya, V. A. Stepanov. 2021. Reconstruction of the Origin of the Gydan Nenets Based on Genetic Analysis of Their Tribal Structure Using a New Set of YSTR Markers. *Russian Journal of Genetics* 57(12): 1414–1423.

Lappalainen, T., S. Koivumäki, E. Salmela, K. Huoponen, P. Sistonen, M. L. Savontaus, P. Lahermo. 2006. Regional differences among the Finns: a Y-chromosomal perspective. *Gene* 376(2): 207–215.

Mallory, James, Anna Dybo, Oleg Balanovsky. 2019. The Impact of Genetics Research on Archaeology and Linguistics in Eurasia. *Russian Journal of Genetics* 55(12): 1472–1487.

Rassadin, Valentin I. 1980. Mongolizmy v chulymsko-t'urkskom jazyke. In: I. D. Bur'ajev (ed.). *Mongolo-bur'atskije zaimstvovanija v sibirskix t'urkskix jazykax*: 48–51. Moskva: Nauka.

Roginskij, Ja. Ja., M. G. Levin. 1978. *Antropologija: Uchebnik dl'a studentov universitetov*. *3-e izd*. Moskva: Vysshaja shkola.

Rootsi, S., L. A. Zhivotovsky, M. Baldovic, M. Kayser, I. A. Kutuev, R. Khusainova, M. A. Bermisheva, M. Gubina, S. A. Fedorova, A. M. Ilumäe et. al. 2007. A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *European Journal of Human Genetics* 15(2): 204–211.

Starostin, S. A. 1982. Praenisejskaja rekonstrukcija i vneshnije sv'azi jenisejskix jazykov. In: E. A. Alekseenko (ed.). *Ketskij sbornik. Antropologija, etnografija, mifologija, lingvistika*: 144–237. Leningrad: Nauka.

Starostin, S. A. 1995. Sravnitel'nyj slovar' jenisejskix jazykov. In: S. A. Starostin (ed.). *Ketskij sbornik. Lingvistika*: 176–315. Moskva: Jazyki russkoj kul'tury.

Starostin, S. A. 2007. *Trudy po jazykoznaniju*. Moskva: Jazyki slav'anskix kul'tur.

Steinitz, Wolfgang. 1959. Zu den samojedischen Lehnwörtern im Ob-Ugrischen. *Ural-Altaische Jahrbücher* XXXI: 426–453.

Tambets, K., S. Rootsi, T. Kivisild, H. Help, P. Serk, E. L. Loogväli, H. V. Tolk, M. Reidla, E. Metspalu, L. Pliss et al. 2004. The western and eastern roots of the Saami — the story of genetic "outliers" told by mitochondrial DNA and Y chromosomes. *The American Journal of Human Genetics* 74(4): 661–682.

ToB = Starostin, S. A. et al. 1998–2005. *The Tower of Babel. An etymological database project.* Online resource available at: starlingdb.org.

Underhill, P. A., P. Shen, A. A. Lin, L. Jin, G. Passarino, W. H. Yang, E. Kauffman, B. Bonné-Tamir, J. Bertranpetit, P. Francalacci, M. Ibrahim, T. Jenkins, J. R. Kidd, S. Q. Mehdi, M. T. Seielstad, R. S. Wells, A. Piazza, R. W. Davis, M. W. Feldman, L. L. Cavalli-Sforza, P. J. Oefner. 2000. Y chromosome sequence variation and the history of human populations. *Nature Genetics* 26(3): 358–361.

Urmanchieva, Anna Ju. 2021. Foneticheskije osobennosti nenetskix zaimstvovanij v obsko-ugorskix jazykax. *Uralo-altajskije issledovanija* 40(1): 101–123.

Urmanchieva, Anna Ju. 2023. *Rekonstrukcija lingvisticheskogo landshafta Zapadnoj Sibiri (na materiale samodijskix jazykov)*. *Doctoral thesis*. Saint-Petersburg.

Valikhova, L. V., V. N. Kharkov, A. A. Zarubin, N. A. Kolesnikov, M. G. Svarovskaya, I. Yu. Khitrinskaya, O. V. Shtygasheva, V. G. Volkov, V. A. Stepanov. 2022. Genetic Interrelation of the Chulym Turks with Khakass and Kets according to Autosomal SNP Data and Y-Chromosome Haplogroups. *Russian Journal of Genetics* 58(10): 1228–1234.

Volkov, V. G. 2013. Drevnie migracii samodijcev i enisejcev v svete geneticheskix dannyx. *Tomskij zhurnal lingvisticheskix i antropologicheskix issledovanij* 1(1): 79–96.

Watanabe, Y., I. Naka, S. S. Khor, H. Sawai, Y. Hitomi, K. Tokunaga, J. Ohashi. 2019. Analysis of whole Y-chromosome sequences reveals the Japanese population history in the Jomon period. *Scientific Reports* 9(1): 8556.

Zerjal, T., B. Dashnyam, A. Pandya, M. Kayser, L. Roewer, F. R. Santos, W. Schiefenhövel, N. Fretwell, M. A. Jobling, S. Harihara et al. 1997. Genetic relationships of Asians and Northern Europeans, revealed by Y-chromosomal DNA analysis. *The American Journal of Human Genetics* 60(5): 1174–1183.

Zhivotovsky, L. A., P. A. Underhill, C. Cinnioğlu, M. Kayser, B. Morar, T. Kivisild, R. Scozzari, F. Cruciani, G. Destro-Bisol, G. Spedini et al. 2004. The Effective Mutation Rate at Y Chromosome Short Tandem Repeats, with Application to Human Population-Divergence Time. *The American Journal of Human Genetics* 74(1): 50–61.

Vladimir N. Kharkov, Anna Y. Urmanchieva, Nikita A. Kolesnikov, Aleksey A. Zarubin, Larisa V. Valikhova, Irina Yu. Khitrinskaya, Ilya A. Gruntov, Ilya M. Egorov, Valeriya M. Lemskaya, Anna V. Dybo, Vadim A. Stepanov. Population genetics and comparative linguistics: phylogeny and phylogeography of haplogroup N1a2b and the linguistic prehistory of the Samoyed peoples

In this paper, we carry out a comparison of ethnogenetic data on the Y chromosome haplogroup N1a2b and linguistics of the Samoyed language group. Haplogroup N1a2b was inherited by modern peoples from an earlier population of the Uraloid anthropological type that inhabited the territory of Southern Siberia. A detailed reconstruction of the phylogeny of this haplogroup shows the ethnospecific structure of the N1a2b sublineages. The linguistic prehistory of the Samoyed peoples is reconstructed based on the analysis and stratification of loanwords. The results on linguistic phylogeny that are relevant for the present paper are obtained using manual analysis of common innovations (for the Samoyedic languages) and Bayesian lexicostatistics (for the Turkic languages). The following correspondences between the results of the genetic and linguistic analyses were uncovered. The region of the initial spread of haplogroup N1a2b roughly corresponds to the area identified as the Samoyedic homeland based on linguistic data. Specific proximity of Mator to the Nenets languages finds support in the history of N1a2b1b1 lineage. This lineage is attested among the Tundra Nenets and Khakas Kachins, whose language could be acquired by the former Mator-speaking population. The presence of N1a2b1b1 lineage among Kazym Khanty is explained by their secondary contacts with Nenets, entailed by the Nenets loanwords in the cultural vocabulary of the Northern Khanty dialects. N1a2b1b2a1 sublineage has been found among Mongols and Tuvans, correlating with the history of the Mongol-Tuvinian language contacts. The undertaken detailed analyses of the phylogeny of N1a2b1b1 lineage confirms the credibility of these conclusions. The obtained modern joint results from population genetics and linguistics complement each other fairly well and provide reliable confirmation of linguistic kinship and genetic-demographic processes in the populations of Siberia.

Keywords: ethnogenetics; Y-chromosome; linguistic prehistory; Samoyed people.

Н. Р. Сумбатова
Институт языкознания РАН, Москва; nina.sumbatova@gmail.com

Соответствия постувулярных согласных в даргинских языках[1]

Задачей работы является исследование соответствий постувулярных (фарингальных и ларингальных) согласных в языках даргинской группы нахско-дагестанской языковой семьи. Описаны соответствия между этими согласными в девяти даргинских идиомах (акушинском, муиринском, кадарском, мегебском, тантынском, ицаринском, чирагском, кайтагском, кубачинском). На основании этих соответствий предлагается реконструкция фрагмента системы согласных и общее описание переходов от реконструируемых прафонем к их рефлексам в современных языках. Особое внимание уделяется признаку фарингализации, которая в части даргинских языков выступает как просодическая характеристика слога, а в части — как сегментный признак гласных.

Ключевые слова: нахско-дагестанские языки; даргинские языки; постувулярные согласные; ларингальные согласные; фарингальные согласные; фарингализация; реконструкция консонантных систем.

1. Введение

1.1. Предварительные замечания

Даргинские языки образуют отдельную группу внутри нахско-дагестанской языковой семьи. По данным лексикостатистики (Коряков & Сумбатова 2007), даргинская группа языков разделилась более 2000 лет тому назад. В нее входит не менее 20–30 идиомов разного уровня. Число и точный состав образующих даргинскую группу языков и диалектов, а также внутренняя классификация даргинских языков до сих пор составляет предмет дискуссии (см., например, Муталов 2021).

На наш взгляд, решение проблемы классификации даргинских языков должно основываться на реконструкции прадаргинской фонологии и моделировании звуковых изменений, которые происходили после разделения прадаргинского языка. В настоящее время существует две реконструкции прадаргинской фонологии — фрагмент общесеверокавказской реконструкции С. Л. Николаева и С. А. Старостина (Nikolaev & Starostin 1994), а также словарь даргинских корней О. А. Мудрака (Мудрак 2016). Реконструкция С. Л. Николаева и С. А. Старостина основана на данных четырех диалектов (акушинского, урахинского, чирагского и кубачинского), из которых два — акушинский и урахинский — очень близки друг к другу. В работе О. А. Мудрака использованы данные семи даргинских идиомов — литературного, урахинского, мегебского, ицаринского, кайтагского, кубачинского, чирагского; в ней предложена реконструкция прадаргинской фо-

[1] Работа выполнена при поддержке гранта РНФ 22-28-01648. Автор выражает искреннюю признательность О. А. Мудраку за предоставление базы данных, использованных в его работе; М. А. Даниэлю за полевые данные по урцакинскому диалекту; А. С. Абдуллаеву, Н. А. Вагизиевой, М. Х. Мамаеву, Р. О. Муталову за помощь в подготовке данных для этой работы; В. Ф. Выдрину за помощь в работе над текстом; анонимному рецензенту за ценные и конструктивные замечания.

нологии и кратко описаны переходы от прадаргинского состояния к современным языкам его выборки. Однако это описание, на наш взгляд, недостаточно подробно, недостаточно ясно и поэтому не всегда убедительно. Кроме того, мы привлекаем материалы сравнительно недавних исследований даргинских языков, которых не было в распоряжении О. А. Мудрака и тем более С. Л. Николаева и С. А. Старостина.

Помимо реконструкций, в ряде работ можно найти полезные описания соответствий между даргинскими идиомами (например, в Магометов 1976, Кадибагамаев 1998, Темирбулатова 2004, Шахбанова 2011), однако в большинстве случаев речь идёт о парных соответствиях — чаще всего между языком, исследуемым в соответствующей работе, и даргинским литературным.

Мы поставили перед собой общую задачу исследовать регулярные фонетические соответствия между даргинскими языками, с тем чтобы попытаться в дальнейшем построить свой вариант реконструкции прадаргинской фонологической системы и описать фонетические переходы, которые произошли в отдельных языках даргинской группы. В этой работе мы рассмотрим один из наиболее сложных фрагментов даргинской системы согласных, а именно согласные фарингального и ларингального места образования. Мы выясним, какие соответствия между этими согласными представлены в даргинских языках, постараемся предложить реконструкцию фрагмента системы согласных и показать, какие рефлексы реконструируемые согласные дали в современных языках даргинской группы.

Работа имеет следующую структуру: в разделе 1.2 речь идет о классификации даргинских языков; в 1.3 приводятся сведения о даргинских языках, необходимые для понимания работы. В разделе 2 содержатся наиболее общие сведения о соответствиях постувулярных согласных в даргинских языках. Конкретные соответствия подробно разбираются в разделах 3, 4 и 5. Там же предлагается реконструкция подсистемы постувулярных согласных. Наконец, в заключении повторяются основные выводы работы и формулируются некоторые вопросы, на которые еще предстоит ответить.

1.2. Классификации даргинских языков и материал исследования

В разное время предлагалось большое число классификаций даргинских диалектов или языков, в основном основанных на отдельных фонологических или грамматических признаках (см. обзор классификаций в Муталов 2021).

Долгое время общепринятой была классификация даргинских языков на две ветви (северную и южную, иногда их называли также языками «акушинского типа» и языками «цудахарского типа» соответственно). В том или ином виде эта классификация принята, например, в работах Быховская 1940; Абдуллаев 1954; Хайдаков 1985; Мусаев 2002; Муталов 2002. В основе этой классификации лежат фонетические признаки, основной из которых — наличие геминированных согласных: в «южных» языках они есть, в «северных» нет. Учитывалась также форма местоимений 1-го лица и личных показателей (в «северных» языках ‘я’ — *nu*, в «южных» *du*; показатель 1-го лица и 2-го лица множественного числа в «северных» языках — *=ra*, показатель 2-го лица единственного числа — *=ri* (*=re*), в «южных» — соответственно *=da* и *=di* (*=de*).

К числу «северных» языков обычно относят акушинский, губденский, кадарский, урахинский, мегебский; к числу «южных» — цудахарский, кубачинский, аштынский, ицаринский, чирагский, кайтагский, мугинский, амухский. «Южных» языков гораздо больше, чем северных, среди них больше мелких идиомов, на которых говорят в одном или двух-трех селениях.

В классификации Р. О. Муталова (2021) сохраняется членение даргинских языков на две ветви, однако деление на идиомы внутри каждой ветви отличается от более ранних классификаций: в северной ветви выделяется два языка (акушинский с большим числом диалектов и мегебский), а внутри южной ветви — четыре (сирхя-цудахарский, кубачинский, кайтагский и чирагский).

Наиболее последовательной следует признать классификацию, сделанную Ю. Б. Коряковым (2021) на основе лексикостатистики (1). В этой классификации выделяется 15 языков, часть которых имеет довольно много диалектов. Классификация Ю. Б. Корякова опровергает бинарное деление даргинских идиомов на две ветви: «южным» языкам традиционных классификаций соответствует три разных группы даргинских языков — южнодаргинская группа, включающая пять отдельных языков и большое число более мелких идиомов, чирагская и кайтагская.

(1) Классификация из работы Коряков 2021:
Северно-центрально-даргинская группа
Севернодаргинский язык (**акушинское наречие**; урахинское наречие; верхнемулебкинское наречие; мюрего-губденское наречие; мугинское наречие; **кадарское наречие**)
Цудахарский язык
Усиша-бутринский язык
Гапшиминский язык
Муиринский язык
Мегебский язык
Южнодаргинская группа
Юго-западно-даргинский язык (**тантынское наречие**; сирхинское наречие; верхневуркунское наречие)
Санжи-**ицаринский** язык
Санакари-чахрижский язык
Амузги-ширинский язык
Кубачинский язык
Аштынский язык
Чирагская группа
Чирагский язык
Кайтагская группа
Кайтагский (хайдакский) язык
Шаринский язык

В нашем исследовании представлены девять даргинских идиомов, а именно: акушинский; кадарский; мегебский; муиринский; тантынский; ицаринский; чирагский; кайтагский; кубачинский. Эти девять идиомов представляют все ветви даргинской группы по классификации Ю. Б. Корякова (в (1) они выделены жирным шрифтом). Акушинский и кадарский являются диалектами («наречиями») севернодаргинского языка, который, в свою очередь, входит в северно-центрально-даргинскую группу; к этой же группе относятся мегебский и муиринский идиомы (они входят в нее как отдельные языки). Тантынский, ицаринский и кубачинский идиомы представляют южнодаргинскую группу: тантынский — один из идиомов юго-западно-даргинского языка, ицаринский представляет собой один из двух вариантов санжи-ицаринского языка; наконец, кубачинский язык входит в южнодаргинскую группу как отдельный язык. Кубачинский хорошо известен как язык, сильно отличающийся от других даргинских, что породило разнооб-

разные версии о его особом происхождении. Наконец, чирагский и кайтагский языки образуют отдельные ветви даргинской группы.

В этой работе мы будем опираться на классификацию Ю. Б. Корякова. В то же время, хотя в целом мы не согласны с традиционным членением языков на южную и северную ветвь, мы будем иногда говорить о даргинских языках как о «южных» или «северных» (эти термины мы будем брать в кавычки). В ряде случаев это оказывается удобным потому, что некоторые фонетические соответствия противопоставляют языки именно таким образом. По традиционной классификации, из анализируемых в данной статье идиомов, к северным языкам относятся акушинский, кадарский, мегебский идиомы, к южным — кубачинский, ицаринский, тантынский, муиринский, кайтагский и чирагский. Даргинский литературный язык, созданный на основе акушинского диалекта, можно условно отнести к северной ветви.

Часть использованных нами данных получена в ходе полевых исследований, часть взята из ранее опубликованных работ[2]. Основные источники следующие:

- акушинский диалект — словник в книге van den Berg 2001, около 800 лексических единиц;
- кубачинский язык —словарь Магомедов & Саидов-Аккутта 2017, около 7000 слов;
- тантынский идиом — словник из работы Сумбатова & Ландер 2014 с некоторыми дополнениями, всего около 1400 лексем;
- ицаринский идиом — словник из работы Sumbatova & Mutalov 2003 и словник, собранный Р. О. Муталовым (Mutalov 2020), всего около 1100 лексем;
- кадарский идиом — словарь, собранный Н. А. Вагизиевой (не опубликован), объем около 1800 лексем; частично пересекается с данными в работе Вагизиева 2018;
- чирагский язык — данные из базы данных О.А. Мудрака, любезно предоставленные автором; в базу данных вошли преимущественно слова, содержащиеся в книгах Кибрик & Кодзасов 1988, 1990, всего около 800 значений (число чирагских лексем больше из-за возможности нескольких способов передачи одного значения);
- мегебский язык — словарь, собранный экспедициями НИУ ВШЭ (Musaev, Morozova & Daniel 2020), около 1100 лексем; словник в грамматике Магометов 1982.
- муиринский язык — полевые данные, собранные в 2018—2021 году в экспедициях РГГУ[3]; всего около 1400 лексем;
- кайтагский (хайдакский) язык — словарь Гасанова 2011, около 4300 лексем; частично также слова из базы данных О. А. Мудрака (2016) и из грамматики Темирбулатова 2004.

На основании перечисленных источников мы составили базу данных, слова в которой сведены в компаративные серии. В настоящее время в базе представлено около 2500 корней, некоторые из которых, однако, представлены только в одном-двух идиомах и пока не могут быть использованы для наших целей.

[2] Наш набор языков пересекается с материалом О. А. Мудрака (Мудрак 2016), где использованы данные семи даргинских идиомов: литературного, урахинского, мегебского, ицаринского, кайтагского, кубачинского, чирагского. Однако конкретные источники данных у О. А. Мудрака другие (см. Мудрак 2016: 6–7), за исключением чирагского и частично ицаринского. В частности, О. А. Мудрак не имел возможности пользоваться большими словарями литературного и кубачинского языка (Юсупов 2017; Абдуллаев 2017; Магомедов & Саидов-Аккутта 2017), которые ко времени создания его работы еще не вышли.

[3] Руководитель экспедиций — Н. Р. Сумбатова; в сборе данных участвовали (в алфавитном порядке): Олег Валиулин, Софья Гайдученко, Маргарита Иванова, Иван Калякин, Виктория Позднякова, Александр Ряпосов, Елена Сокур, Софья Ярошевич. Особенно большой вклад внесли Виктория Позднякова и Диана Пчелинцева.

Все слова нашей базы мы проверили по недавно вышедшему диалектологическому словарю С. М. Темирбулатовой (2022) и дополнили информацией из этого словаря; в случае расхождения полевых данных или других источников со словарем Темирбулатовой мы пытались проверить данные у носителей языка (такая возможность у нас была для муиринских, тантынских, ицаринских, реже для мегебских данных) или по дополнительным источникам.

Акушинский словник в книге van den Berg 2001 невелик, поэтому в тех компаративных сериях, где нужного акушинского слова нет ни в словнике, ни в словаре С. М. Темирбулатовой, мы использовали слова даргинского литературного языка из словаря Юсупов 2017. Эти случаи отмечены буквой (*Л*).

Слова идиомов, собранные в ходе полевых исследований (мегебского, тантынского, муиринского, ицаринского, кадарского) в базе данных представлены в фонологической записи. Кубачинская, кайтагская, акушинская лексика, а также вся лексика, позаимствованная из словаря Темирбулатова 2022, даны в виде транслитерации орфографической (кириллической) записи в соответствующих словарях. К сожалению, в кубачинском словаре не всегда передается фарингализованность гласных (кроме *a*), а кайтагский словарь (Гасанова 2011) часто расходится с данными из грамматики Темирбулатова 2004 (последняя, судя по всему, гораздо более надежна). Акушинская лексика из книги van den Berg 2001 и чирагская из базы данных О. А. Мудрака приведены к тому варианту транскрипции, который используется для других идиомов (знаки транскрипции приводятся ниже, в разделе 1.3.1).

Все единицы в нашей базе даны либо в виде начальной формы лексемы (для существительных это форма абсолютива, не имеющая флексий), либо в виде корня (в последнем случае в конце ставится дефис). В частности, в виде корня даны глаголы[4].

Для выявления фонетических соответствий мы старались не привлекать заимствованную лексику и идеофоны. В тех немногих случаях, где идеофоны все же использованы, это специально оговаривается.

1.3. Сведения о фонологии даргинских языков

1.3.1. Фонологические системы даргинских языков

Число гласных в даргинских языках сравнительно невелико: как правило, имеется только четыре основных гласных (*a, u, e, i*). Во всех идиомах, насколько это известно, есть также фарингализованные гласные, их число может варьировать от одного (*aˤ*) до четырех (*aˤ, uˤ, eˤ, iˤ*)[5]. Кроме того, в некоторых идиомах (например, в кубачинском и ицаринском) гласные противопоставлены по долготе.

Напротив, системы согласных довольно богаты. В таблицах 1 и 2 показаны типичные системы «северных» и «южных» даргинских языков (в соответствии с традиционной классификацией). Как видно из таблиц, основное различие между этими группами идиомов заключается в наличии/отсутствии геминированных согласных: в «южных» языках смычные согласные одного места образования образуют четверки вида «звонкий — глухой простой — глухой геминированный — абруптивный», а в северных им соответствуют тройки «звонкий — глухой — абруптивный», например *b — p — p: — p’* в южных

[4] Глаголы даны в двух видах (СВ/НСВ); о выражении вида в составе глагольного корня см. раздел 1.3.3.

[5] Фонологический статус фарингализованных гласных не всегда ясен. Поскольку проблема фарингализации тесно связана с темой этой работы, мы подробнее обсуждаем ее в разделе 1.3.2.

языках и *b — p — p'* в северных. Фрикативные согласные в даргинских языках не могут быть абруптивными, а аффрикаты чаще всего не бывают звонкими, соответственно группы одного места образования в южных языках образуют тройки (например, *z — s — s:* или *c — c: — c'*), а в северных — пары (*z — s* или *c — c'*).

В большинстве даргинских языков[6] имеются также лабиализованные согласные, которые не показаны в Таблицах 1 и 2. Лабиализованными могут быть все согласные, кроме сонантов, губных и, видимо, дентальных фрикативных, однако фонологический статус лабиализованных не всегда ясен. В этой работе лабиализованные согласные и их соответствия систематически не обсуждаются[7].

Таблица 1. Согласные фонемы в «южных» даргинских языках

	Смычные	Фрикативные	Аффрикаты	Сонанты
Губные	b p p: p'			m, w
Дентальные	d t t: t'	z s s:	c c: c'	n, l, r
Альвеолярные		ž š š:	č č: č'	j
Велярные	g k k: k'	(ɣ[8]) x x:		
Увулярные	q q: q'	ʁ χ χ:		
Фарингальные	ʢ	[ʕ] ħ		
Ларингальные	ʔ	h		

Таблица 2. Согласные фонемы в «северных» даргинских языках

	Смычные	Фрикативные	Аффрикаты	Сонанты
Губные	b p p'			m, w
Дентальные	d t t'	z s	c c'	n, l, r
Альвеолярные		ž š	č č'	j
Велярные	g k k'	(ɣ) x		
Увулярные	q q'	ʁ χ		
Фарингальные	ʢ	ħ		
Ларингальные	ʔ	h		

Система согласных акушинского и кадарского идиомов, представленных в нашей выборке, соответствует Таблице 2. В тантынском идиоме система такая, как показано в Таблице 1, включая фонему /ʕ/. В прочих идиомах выборки (муиринском, ицаринском, кубачинском, кайтагском, чирагском) система согласных соответствует Таблице 2, однако ларингала /ʕ/ нет.

[6] В том числе во всех идиомах нашей выборки, кроме акушинского.

[7] Анонимный рецензент отметил отсутствие обсуждения лабиализованных согласных как недостаток нашей работы. Мы согласны с тем, что без такого обсуждения реконструкция остается неполной, но все же решили не включать его в эту работу, так как соответствия лабиализованных довольно сложны и требуют самостоятельного исследования, которое мы надеемся со временем представить (см., например, 'запах' (начальная согласная): кадар. *wa*ˁʔ, акуш. *ga*ˁʔ, муир. *k:a*ˁʔ, ицар., кайтаг., кубач. *k:*ʷ*a*ˁʔ; 'слово, звук' (конечная согласная): кадар. , акуш. , муир., тант. *dew*, мегеб. *deh*ʷ, ицар. *deh*).

[8] В ряде идиомов звонкая велярная фонема /g/ произносится как фрикативный звук [ɣ] (иногда в свободном варьировании с [g]), однако по своим фонологическим и морфонологическим свойствам соотносится со смычными глухими /k/ и /k:/.

В мегебском языке ситуация не совсем ясна. Это язык «северного» типа. В описании Магометов 1982: 8–9 отмечается, что в мегебском имеются интенсивные (геминированные) согласные, звонкие аффрикаты *ʒ* и *ǯ*, два звонких велярных — фрикативный и смычный, а также «интенсивный фарингальный глухой». Фонологичность этих звуков не обсуждается. В работе Moroz 2019 наличие геминированных отрицается, а звонкие аффрикаты и звонкий фрикативный *ɣ* рассматриваются как аллофоны фонем /z/, /ž/ и /g/ соответственно. Описание постувулярных согласных в этой статье отличается от привычной картины. Во-первых, постулируется три, а не четыре постувулярных фонемы (без учета лабиализованных), а именно фарингальный /ħ/, ларингальные /ʔ/ и /h/. Во-вторых, рассматривается несколько аллофонов этих фонем, прежде всего [н] и [ʡ] (аллофоны /ħ/ и /ʔ/ соответственно).

Предметом нашего исследования являются соответствия постувулярных, то есть фарингальных и ларингальных согласных. Как видно из Таблиц 1 и 2, устройство фарингальных и ларингальных подсистем согласных в целом сходно в различных идиомах. Большинство современных даргинских языков имеет в этой зоне четыре согласных: два смычных — /ʔ/ и /ʡ/ — и два фрикативных — /h/ и /ħ/. Кроме того, в одном из идиомов нашей выборки — тантынском — присутствует звонкий фрикативный фарингальный согласный /ʕ/.

Вполне вероятно, что фонетическая реализация постувулярных согласных различается в разных идиомах даргинской группы, о чем косвенно свидетельствуют расхождения в терминах, используемых для их характеристики в разных работах. Если для согласных *ʔ* и *h* практически все авторы используют термины, указывающие на ларингальную зону артикуляции (*гортанные* или *ларингальные*, в англоязычных работах — *glottal*), то с *ʡ* и *ħ* ситуация сложнее: их определяют либо как фарингальные (*pharyngeal*), как в работах Темирбулатова 2004, Магометов 1963, van den Berg 2001, Sumbatova & Mutalov 2003, либо как *надгортанниковые* (*epiglottal*) — как в Абдуллаев 1954, Ganenkov 2019a, 2019b[9].

Тем не менее можно констатировать, что в большинстве даргинских идиомов, в том числе во всех идиомах нашей выборки (кроме, возможно, мегебского) подсистема постувулярных включает два глухих фрикативных согласных (h ~ ħ), которые противопоставляются как ларингальный и фарингальный или эпиглоттальный, а также два смычных (ʔ ~ ʡ), также ларингальный и фарингальный/эпиглоттальный.

В этой работе мы будем использовать одинаковые термины (ларингальный и фарингальный) и одинаковые символы (те, что фигурируют в таблицах 1 и 2) для постувулярных согласных всех языков нашей выборки, даже если фонетически эти обозначения окажутся не совсем точными.

Существенные проблемы при сопоставлении лексических единиц создает гортанная смычка /ʔ/. Ее присутствие во многих случаях трудно определить, так как звук [ʔ] во многих идиомах не встречается в начале слова и между гласными. В позиции между гласными, а иногда и в начале слова, однако, могут быть существенные аргументы за то, чтобы постулировать /ʔ/ из морфологических или морфонологических соображений. Например, в большинстве даргинских языков на стыке гласных происходит стяжение (правила которого могут различаться по языкам), и в случаях, когда такого стяжения нет, логично постулировать присутствие гортанной смычки. Так, в (2b) две гласных *а* на стыке морфем стягиваются в одну, а в (2a) стяжения не происходит, что можно объяснить за счет присутствия в суффиксе начального ларингала /ʔ/.

[9] В работе Moroz 2019 /ħ/ характеризуется как *pharyngeal*, а [ʡ] (аллофон /ʔ/) — как *epiglottal*. В книге Forker 2020: 20 использован двойной термин *pharyngeal/epiglottal*.

(2) Даргинский литературный

a. q:aˤmaala /q:aˤma-ʔala/
изогнутый-N.ABSTR
‘извилина’

b. kages /ka-ag-es/
DOWN-попасть.PFV-INF
‘попасть’

Другой пример — формы порядковых числительных. В большинстве даргинских языков они состоят из основы числительного и причастия глагола ‘сказать’, для которого постулируется корень, состоящий из одной согласной /ʔ/. Если при этом на стыке основы числительного и причастия встречаются две гласных, то стяжения также не происходит, ср. (3a), где постулируется ларингал /ʔ/, препятствующий стяжению *a* и *i*, и (3b), где те же гласные стягиваются в сочетание [aj]:

(3) Даргинский литературный

a. caibil /ca-ʔ-ib-il/
один-сказать.PFV-PRET-ATTR
‘первый’

b. kajses /ka-is-es/
DOWN-кроить.PFV-INF
‘кроить’

Мотивированно постулировать /ʔ/ в случае нулевой реализации можно только при хорошем знании правил морфонологии, что в нашем случае не всегда доступно. Мы будем пользоваться данными в том виде, в каком они представлены в имеющихся у нас словарях и базах данных, в случае необходимости оговаривая наличие фонемы /ʔ/, не имеющей фонетической реализации.

1.3.2. Проблема фарингализации в даргинских языках

Практически во всех описаниях даргинских языков встречаются термины *фарингализация* и/или *фарингализованный*. Как правило, речь идет о фарингализованных гласных, реже — о фарингализованных согласных (см., например, Кадибагамаев 1998: 38–40 для сирхинского; Магометов 1963: 47–50 для кубачинского). Многие авторы говорят о фарингализации как о просодической характеристике даргинских слов.

Фонетически фарингализация представляет собой сужение гортани при произнесении тех или иных речевых сегментов, однако нас больше интересует ее фонологический статус. В большинстве описаний даргинских языков фарингализация и свойства фарингализованных звуков описаны недостаточно ясно и/или недостаточно подробно. В частности, не всегда ясно, является ли фарингализация (в данном идиоме) просодическим признаком, и если да, то какова сфера ее действия (слог, несколько слогов, фонетическое слово), являются ли «простые» и соответствующие фарингализованные гласные (и согласные, если они есть) разными фонемами или нет, каково распределение простых и фарингализованных звуков. Хотя в этой статье у нас нет возможности дать подробный обзор проблемы фарингализации в даргинских языках, попробуем зафиксировать основные моменты, которые существенны для нашей работы[10].

Во-первых, даргинские языки различаются по набору гласных, которые противопоставляются по признаку фарингализованности. Если взять языки нашей выборки, то в четырех из них — акушинском, кадарском, муиринском и кайтагском — фарингализован-

[10] Мы учли описания следующих языков: литературного даргинского (Абдуллаев 1954), акушинского (van den Berg 2001), кубачинского (Магометов 1963), мегебского (Магометов 1982, Moroz 2019), кайтагского (Темирбулатова 2004), ицаринского (Sumbatova & Mutalov 2003), санжинского (Forker 2020), тантынского (Сумбатова & Ландер 2014), худуцкого (Ganenkov 2019b), чирагского (Ganenkov 2019a), сирхинского (Кадибагамаев 1998).

ным может быть только гласный *aˤ*, в ицаринском и чирагском — *aˤ* и *uˤ*, в тантынском — *aˤ*, *uˤ*, *eˤ*, в кубачинском, аштынском и ширинском (Беляев 2019; 2021) — все четыре гласных (*aˤ*, *uˤ*, *eˤ*, *iˤ*)[11]. Это, в частности, означает, что в кубачинском, аштынском и ширинском фарингализация может быть фонетически реализована на любом слоге.

Во-вторых, рассмотрим вопрос о том, является ли фарингализация просодическим признаком или сегментным признаком гласных. Ответ необязательно будет одинаковым для разных даргинских идиомов. В большинстве даргинских языков есть корреляция между фарингализованностью гласной и характеристиками соседних согласных. Чаще всего корреляция наблюдается в последовательности CV, то есть с гласным коррелирует предшествующий ему согласный, однако в некоторых языках корреляция наблюдается и в последовательности VC. Корреляция выглядит следующим образом: (а) рядом с ларингальными согласными возможны только нефарингализованные гласные, (б) рядом с фарингальными согласными (иногда также рядом с некоторыми увулярными) возможны только фарингализованные гласные, (в) рядом с прочими согласными воможны любые гласные; (г) гласные, не противопоставленные по признаку фарингализации, сочетаются с любыми согласными. В (4) такая корреляция иллюстрируется на примере последовательности CV в даргинском литературном языке: (4a), (4b) — простая гласная после ларингальной согласной; (4d), (4e) — фарингализованная после фарингальной согласной; примеры (4c) и (4f) демонстрируют допустимость простой и фарингализованной гласной после согласной *qː*; прочие примеры демонстрируют допустимость любых согласных перед гласной *i*, не имеющей фарингализованного коррелята (4g–4k), и в позиции после сонанта перед другой согласной (4l–4p).

(4) Даргинский литературный

a.	**ha**la	b.	q’ur**ʔa**n	c.	**qː a**m					
	‘перед’		‘Коран’		‘племя’					
d.	**ħa**ˤb	e.	**ʔa**ˤnd	f.	**qː a**ˤm					
	‘три’		‘узел’		‘миска’					
g.	**ħi**	h.	**ʔi**niz	i.	**hi**mi	j.	azir**ʔi**bil	k.	**si**ris	
	‘кровь’		‘источник’		‘желчь’		‘тысячный’		‘иней’	
l.	dur**ħ**-ni	m.	ʔaˤrʔ-ni	n.	tar**h**-ni	o.	war**ʔ**-ni	p.	har**g**-ni	
	мальчик-PL		курица-PL		ласка-PL		мед-PL		навес-PL	
	‘мальчики’		‘куры’		‘ласки’ (жив.)		‘мед’ (собир.)		‘навесы’	

Корреляция между фарингализованностью гласной и качеством соседней согласной может присутствовать в идиоме независимо от числа гласных, противопоставленных по фарингализованности. В частности, она есть в акушинском, где имеется только одна фарингализованная гласная, и в кубачинском, где их четыре.

Иначе говоря, в даргинском литературном (а также в акушинском, ицаринском, санжинском и других языках) фарингальные и ларингальные согласные (вопрос об увулярных мы здесь не рассматриваем), а также простые и фарингализованные гласные дополнительно распределены в одних контекстах и противопоставлены в других, как показано в Таблице 3[12].

[11] С мегебским языком не все ясно: в грамматике Магометов 1982 представлен только один фарингализованный гласный (*aˤ*); в работе Г. Мороза (Moroz 2019) указано, что фарингализованными могут быть все гласные, в том числе *oˤ*, для которого отсутствует нефарингализованный вариант. В словнике мегебского языка (Musaev, Morozova & Daniel 2020), которым мы пользовались, встречаются *aˤ* и *oˤ*, но другие фарингализованные гласные практически отсутствуют.

[12] Таблица 3 имеет иллюстративный характер, поэтому в ней показаны не все возможные позиции согласных и проигнорированы увулярные согласные, которые иногда по крайней мере частично ведут себя как фарингальные.

Таблица 3. Сочетаемость простых и фарингализованных гласных с согласными

Позиция	Пример контекста	Согласные			Примеры
		Фарингальные	Ларингальные	Другие	
А. Перед гласными, противопоставленными по фарингализованности	_a	—	ha, ʔa	Ca	4abc
	_aˤ	ħaˤ, ʡaˤ	—	Caˤ	4def
Б. Перед гласными, не имеющими фарингализованного корелята	_i	ħi, ʡi	hi, ʔi	Ci	4fhijk
В. После сонанта (перед согласным или на конце слова)	r_#, r_C	Rħ(C), Rʡ(C)	Rh(C), Rʔ(C)	RC(C)	4lmnop

В квадрате, выделенном жирной линией, фарингальные и ларингальные согласные (или простые и фарингализованные гласные) находятся в дополнительном распределении; в прочих клетках нашей таблицы они противопоставлены[13]. Зона дополнительного распределения тем больше, чем большее число гласных имеет фарингализованный коррелят и чем в большем числе позиций наблюдается корреляция между фарингализованностью гласной и качеством согласной.

Примером идиома, где зона дополнительного распределения максимальна, может служить кубачинский, в котором, как было сказано выше, все четыре гласные могут быть как простыми, так и фарингализованными (контекстов типа Б в кубачинском нет).

Примером языка, где зона дополнительного распределения практически отсутствует, может служить муиринский, в котором, как и в акушинском, фарингализованной может быть только гласная *aˤ*, но, в отличие от акушинского, после фарингальных согласных *ħ* и *ʡ* допустимы как фарингализованные, так и простые гласные, как показано в (5) и (6)[14].

(5) Муиринский

a. **ʡa**ri	b. **ʡa**t	c. ba**ʡa**gan
‘заяц’	‘узел’	‘широкий’
d. **ʡa**ˤt'a	e. **ʡa**ˤt	f. waˤr**ʡa**ˤ
‘лягушка’	‘кадык’	‘курица’

(6) Муиринский

a. **ħa**na	b. a**ħa**l	c. dar**ħa**
‘шрам’	‘гость’	‘сын’
d. **ħa**ˤka	e. **ħa**ˤrc'	f. im**ħa**ˤ
‘шуба’	‘куница’	‘корень’

Если в языке есть контексты типа В (Таблица 3), как в литературном и акушинском, то фарингальные и ларингальные согласные в этом языке фонологически противопоставлены. Что касается гласных, то чем больше в языке зона дополнительного распределения, тем более естественным будет решение считать фарингализацию просодической

[13] В тех позициях, которые не попали в нашу таблицу, — например, после гласной, — ситуация может быть разной в зависимости от морфонологии конкретного идиома.

[14] В муиринском выбор фарингализованной/простой гласной контекстно обусловлен только после ларингальных согласных и в абсолютном начале слова. В обеих этих позициях допустима только простая гласная *a*, но не фарингализованная *aˤ*: *hark* ‘открытый’; *χurhar* ‘легкое’; *arcan* ‘птица’; *at:a* ‘отец’.

характеристикой слога или более протяженной единицы[15]. Такой подход выглядит разумным для кубачинского, но гораздо менее очевидным для акушинского и тем более для муиринского. Действительно, фарингализация рассматривается как просодический признак в описаниях языков, имеющих четыре фарингализованных гласных (кубачинского — Магометов 1963, сирхинского — Кадибагамаев 1998, худуцкого — Ganenkov 2019b, чирагского — Ganenkov 2019a, аштынского — Беляев 2021) и как сегментный признак гласного — в описании идиомов, имеющих только фаригализованный *aˤ* (акушинского — van den Berg 2001, кайтагского — Темирбулатова 2004).

К трактовке фарингализации имеет отношение еще одно явление, отмеченное во многих даргинских языках. Рассмотрим его на примере ицаринского идиома. Ицаринский имеет два фарингализованных гласных, правила сочетаемости простых/фарингализованных гласных с фарингальными/ларигальными согласными в целом такие же, как в акушинском. Как и в других даргинских языках, в ицаринском некоторые грамматические морфемы имеют два варианта, с простым и с фарингализованным гласным, и выбор варианта обусловлен предшествующей корневой морфемой[16]: если в корне фарингализованный гласный, то фарингализация как бы распространяется на следующий слог, содержащий грамматический показатель, как в (7):

(7) Ицаринский

a.	b-at-aq-ur	b.	b-aˤrq:-aˤq-ur
	N-оставлять.PFV-CAUS-PRET		N-учить.PFV-CAUS-PRET
	‘заставил оставить’		‘заставил учить’

Такая зависимость напоминает сингармонизм, однако можно показать, что она имеет скорее морфологический, чем фонологический характер. Во-первых, «фарингализующими» могут быть морфемы, не содержащие фарингализованных гласных, как в (8b). Во-вторых, способность вызывать фарингализацию в соседнем слоге является свойством именно морфемы, а не слога: ср., например, сегментно омонимичные корни в (8a) и (8b), один из которых вызывает фарингализацию последующего суффикса, а другой нет. Способность взывать фарингализацию в соседнем слоге является лексическим признаком морфемы.

(8) Ицаринский

a.	b-irq:-u-j	b.	b-irq:-uˤ-j	c.	d-uˤbχ:-uˤ-j
	N-копать.IPFV-TH-SUBJ		N-пачкать.IPFV-TH-SUBJ		NPL-молотить.IPFV-TH-SUBJ
	‘копать’		‘пачкать’		‘молотить’

Правила «расстановки» фарингализации в ицаринском (и близкородственном ему санжинском, см. Forker 2020: 26–29) вряд ли могут быть описаны как полностью просодические. Тем не менее складывается ощущение, что такие правила возникли в результате преобразования системы, в которой фарингализация на более раннем этапе была просодическим признаком (к этой мысли мы вернемся в разделе 3.3).

В этой работе перед нами не стоит задача описания фонологических систем всех языков, задействованных в нашей работе. Для нас важно иметь представление о соотно-

[15] Нелогично, однако, одновременно выделять фарингализованные и простые гласные фонемы и вводить просодический признак фарингализации, как это делается в Forker 2020, где на стр. 26 фарингализация вводится как признак гласных, а на стр. 28 — как супрасегментный признак.

[16] Возможно и обратное влияние: некоторые аффиксы подвергают фарингализации предшествующий корень, см. Sumbatova & Mutalov 2003: 4–6.

шении простых и фарингализованных гласных, а также знать, что во всех исследуемых языках (даже в кубачинском) фарингальные и ларингальные согласные фонологически противопоставлены.

1.3.3. Структура слога и сочетания согласных

Структуру слога в даргинских языках можно выразить формулой (C)V(R)(C), где V — любой гласный, C — любой согласный, R — согласный некоторого узкого класса, содержащего прежде всего сонанты (*w, j, m, n, r, l*), иногда также *b, s, ž, š* (набор согласных, допустимых в позиции R, слегка расходится по даргинским языкам). Иначе говоря, в начале слога стечения согласных недопустимы, а в конце ограничены. На стыках слогов, однако, возможны сочетания практически любых согласных. В (9a) приводятся одноморфемные слова, иллюстрирующие все возможные структуры слога, в (9b) — сочетания, возникающие на стыках слогов (все примеры в (9) из литературного даргинского языка).

(9) Даргинский литературный

a.	V	*u* ‘имя’
	VC	*it* ‘тот’
	VR	*aw* ‘четыре’
	VRC	*arc* ‘серебро’
	CV	*ga* ‘уголь’
	CVC	*maq* ‘береза’
	CVR	*t’em* ‘запах’
	CVRC	*q’anq’* ‘журавль’
b.	VCV	*musa* ‘место’
	в том числе VRV	*mura* ‘сено’
	VCCV	*žahsi* ‘молодой’
	в том числе VCRV	*bucrub* ‘сироп’
	VRCV	*t’erxa* ‘палка’
	VRRV	*χurmi* ‘липы’ (множ. ч.)
	VRCCV	*barχsi* ‘смежный, совместный’
	в том числе VRCRV	*t’unt’ri* ‘мухи’

Что касается постувулярных согласных, то их позиционное распределение в целом такое же, как у других шумных. Слова в (10) иллюстрируют все допустимые позиции таких согласных (H — постувулярный согласный, | — граница слога).

(10) Даргинский литературный

#HV	в начале слова (перед гласной)	*ħi* ‘кровь’, *hih* ‘дыхание’, *ʔaˤħ* ‘добро’, *ʔaˤnd* ‘узел’
V|HV	между гласными	*baˤʔusi* ‘широкий’, *meħe* ‘мозг’
VH#	после гласной в конце слова	*daˤħ* ‘лицо’, *beʔ* ‘хозяин’
VH|RV	после гласной перед сонантом	*bahlasi* ‘спокойный’
VH|CV	после гласной перед шумным согласным	*duhbi* ‘слова’, *baˤħsi* ‘противоположный’
VRH#	после сонанта в конце слова	*urħ* ‘чужак’, *derħ* ‘этаж’
VR|HV	после сонанта перед гласным	*durħaˤ* ‘мальчик, сын’, *ʔaˤrʔaˤ* ‘курица’
VRH|CV	после сонанта перед шумным согласным	*zaˤrħsi* ‘свежий’

VRH|RV	между сонантами	*durħni* ‘мальчики, сыновья’, *χurʁri* ‘легкие’
VC|HV	после шумного согласного перед гласным	*masħu* ‘омовение ног’

При исследовании соответствий согласных, включая постувулярные, некоторую сложность создают согласные в корнях глаголов. В даргинских языках существуют довольно жесткие ограничения на структуру глагольных корней. Абсолютное большинство корней имеет одну из следующих структур: VC, *CL*-VC, VRC, *CL*-VRC, RVC, где *CL*- обозначает позицию показателя именного класса. При этом внутри корня выражается категория глагольного вида (совершенный/несовершенный вид), так что большинство глагольных корней существует в двух видовых вариантах. Эти варианты могут различаться наличием слота для классного согласования, наличием и местом сонанта R, выбором гласной V, как показано в (11).

(11) Муиринский

Значение	Корень СВ	Корень НСВ
‘находить’	*CL-ark:*	*CL-urk:*
‘писáть’	*CL-elk’*	*luk’*
‘бросать’	*ih*^{w}	*irh*^{w}
‘ломать’	*CL-urʔ*	*uʔ*
‘шить’	*CL-ep:*	*ip:*
‘сеять’	*CL-aʔ*	*CL-alʔ*
‘измерять’	*umc*	*umc*

Когда мы обсуждаем соответствия согласных в глагольных корнях, позицию согласного С нельзя охарактеризовать однозначно: во-первых, позиция С может различаться в разных видовых вариантах одного корня (в одном виде согласный стоит после гласного, в другом — после сонанта); во-вторых, в абсолютном большинстве глагольных форм непосредственно за корнем следует грамматический показатель, который в большинстве случаев начинается с гласной (как в 12abc), но есть и показатели, начинающиеся с согласной (12d). Кроме того, в формах так называемого краткого оптатива глагол выступает в виде чистой основы, то есть справа от корня никаких показателей нет (согласный корня оказывается на конце слова, 12e).

(12) Муиринский

a. b-ar**q’**-ib
N-делать.PFV-PRET
‘сделал’

b. b-ir**q’**-u
N-делать.IPFV-FUT.3.TR
‘сделает’

c. b-ar**q’**-ara / b-ir**q’**-ara
N-делать.PFV-INF / N-делать.IPFV-INF
‘сделать’ / ‘делать’

d. b-ar**q’**-ni / b-ir**q’**-ni
N-делать.PFV-MSD / N-делать.IPFV-MSD
‘тот факт, что сделал’ / ‘тот факт, что делал’

e. b-ar**q’** / b-ir**q’**
N-делать.PFV(OPT) / N-делать.IPFV(OPT)
‘пусть сделает!’ / ‘пусть делает!’

Ниже при описании соответствий ларингальных и фарингальных согласных мы будем оговаривать те случаи, когда они встречаются в составе глагольного корня.

2. Соответствия постувулярных согласных: общие сведения

Соответствия постувулярных согласных — один из наиболее сложных фрагментов системы соответствий даргинских идиомов. Дополнительную трудность вызывает то об-

стоятельство, что фонема /ʔ/ часто имеет нулевую реализацию, поэтому не всегда есть возможность решить, присутствует ли она в некотором слове синхронно (1.3.1). Объективные сложности усугубляются заметным числом ошибок и расхождений в словарях, причем это, к сожалению, касается не только полевых данных, но даже и некоторых опубликованных словарей. Поэтому те наблюдения, которые будут сделаны в разделах 3–5, следует считать предварительными.

Постувулярные согласные участвуют в большом числе соответствий, причем иногда они соответствуют согласным другого места образования, чаше всего увулярным, но не только, например:

Таблица 4. Примеры соответствий постувулярных согласным других мест образования

	кадар.	мегеб.	акуш.	муир.	тант.	ицар.	кайт.	чираг.	кубач.
лягушка, жаба	ʔaˤtˈa	ʔaˤtˈa		ʔaˤtˈa	ʔaˤtˈa	ʔaˤtˈa	ʔaˤtˈa	**ʁ**ˤàtˈa	ħitˈa
угол внутренний				**q**amz	ʕaˤmza	ʔuˤmza	**q**:aˤmzi		
посох	ʔaˤsa	ʔansa	ʔaˤjsa		ʔiˤsa	**q**'aˤs:a		as:a	ħaˤs:a
слово, звук	de**w**	deh^{w}	dew	dew	dew	deh			
уздечка (второй слог)		hurhur	urhur	urhur		ur**x**:ur		ur**x**:ùr	u**x**:uj

Часть соответствий такого рода вполне регулярна, однако в этой статье мы рассмотрим только те регулярные соответствия, в которых участвуют исключительно фарингальные и ларингальные согласные.

Соответствия постувулярных согласных рассмотрены в работе Мудрака (Мудрак 2016: 34–35) на материале семи даргинских идиомов, список которых частично совпадает с нашим. Для четырех рядов соответствий согласных современных идиомов О. А. Мудрак реконструирует четыре прафонемы (Таблица 5, строки 1–4) — в целом те же, что представлены в большинстве современных языков. Последняя строка таблицы отражает ряд соответствий с билабиальным *w*, который в нашей работе не рассматривается (Мудрак реконструирует для этого ряда соответствий прафонему *ɣ̊). О существовании звонкого фарингального согласного /ʕ/ у Мудрака, видимо, не было информации.

Таблица 5. Реконструкция ларингалов согласно Мудрак 2016: 35.

ПДарг	акуш.	урах.	мегеб.	ицар.	хайд.	куб.	чир.
*ʔ	ʔ / ʕ[17]	ʔ / ʕ	ʔ / ʕ	ʔ / ʕ (ʡ)	ʔ / ʕ	ʔ / ʕ	ʔ / ʕ
*ʡ	ħ / h	ħ / h	ħ / h	ʔ / ʕ (ʡ)	ʔ / ʕ (ʡ), (h/ħ)	ʔ / ʕ	ʕ, (h, ħ)
*ḫ	ʕ / ʔ	ʕ / ʔ	ʕ (ħ)	h / ħ / ʕ	h / ħ / ʕ	h / ħ	h / ħ / ʕ
*ħ	h / ħ	h / ħ	h / ħ	h / ħ	h / ħ	h / ħ	h / ħ
*ɣ̊	w	w	w	ɣ	w	w	w

Работа Мудрака, очевидно, неполна. В ней не только не учитывается наличие звонкого фрикативного фарингального /ʕ/, но и не описаны условия выбора того или иного реф-

[17] Обозначения в работе Мудрака не совпадают с нашими; в частности, символом ʕ у него обозначен «айн», «эпиглоттальный смычный ларингал» (Мудрак 2016: 35), а у нас — звонкий фрикативный фарингальный согласный.

лекса для тех ларингалов, которые присутствуют в его модели — хотя, как видно из его таблицы, практически во всех ее клетках представлено более одного рефлекса. Кроме того, в тексте работы нет подробного обоснования этой реконструкции[18]. Все это приводит нас к необходимости рассмотреть соответствия постувулярных еще раз.

Основные типы соответствий представлены в Таблице 6. Как видно из таблицы, имеются тривиальные или почти тривиальные соответствия (для согласных *ʔ*, *h*, *ħ*), близкое к тривиальному соответствие *ʔ* и нуля, а также два ряда соответствий, где, упрощая ситуацию, можно говорить о противопоставлении «северных» и «южных» языков: в строке (5) в «северных» языках смычный ларингал, а в «южных» чаще всего фрикативный, в строке (6) — наоборот. При этом муиринский язык, вопреки традиционным классификациям и в соответствии с классификацией Ю. Б. Корякова (2021), примыкает к «северным» идиомам.

Легко видеть, что наша Таблица 6 довольно хорошо соотносится с таблицей О. А. Мудрака: наши строки 1–2 соответствуют первой строке в Таблице 5, строки 3—4 соотносятся со строкой 4, строки 5 и 6 — со строками 3 и 2 соответственно. Кроме того, есть довольно много слов, которые не укладываются в приведенные здесь соотношения; мы постараемся упомянуть и прокомментировать по крайней мере некоторые из них.

Таблица 6. Основные классы соответствий ларингалов в даргинских языках

	кадар.	мегеб.	акуш.	муир.	тант.	ицар.	кайт.	чираг.	кубач.
(1)	ʔ	ʔ	ʔ	ʔ	ʔ	ʔ	ʔ	ʔ	ʔ
(2)	ʔ/Ø	ʔ/Ø	ʔ/Ø	ʔ/Ø	ʔ/Ø	ʔ/Ø	ʔ/Ø	ʔ/Ø	ʔ/Ø
(3)	h	h	h	h	h	h	h	h	h
(4)	ħ	ħ	ħ	ħ	ħ	ħ	ħ	(χ)	ħ
(5)	ʔ	Ø / ʔ	ʔ	ʔ	ħ / ʔ / h	h	h	h / ʔ / ħ	h
(6)	ħ	ħ	ħ	ħ	ʕ / ħ / ʔ / Ø	ʔ, ʔ/Ø	ʔ, ʔ/Ø	ʔ, ʔ/Ø	ʔ, ʔ/Ø

В следующих трех разделах мы рассмотрим ряды соответствий, представленные в Таблице 6: ряды (1—4) будут проанализированы в разделе 3, ряд 5 — в разделе 4 и, наконец, ряд 6 — в разделе 5.

3. Тривиальные и близкие к ним соответствия ʔ, ʔ, ħ и h

3.1. Соответствия ʔ и ʔ

Слов с тривиальным соответствием *ʔ* довольно много. Заметную часть таких слов составляют заимствования, в основном арабские[19], но довольно много и исконной лексики,

[18] В работе Nikolaev & Starostin 1994: 117–120 реконструируются те же четыре фонемы плюс, под вопросом, пятая, *hI*, — фарингализованный фрикативный ларингал.

[19] Примеры арабских заимствований, в которых представлен ларингал ʔ: ‘жизнь’ — акуш., тант., кайт., кадар. *ʔaˤmru*, мегеб. *ʔumru*, кубач. *ʔuˤmru*; ‘вина’ — акуш., тант. *ʔaˤjib*, кадар. *ʔaˤip*, кубач. *ʔaˤjab*; ‘час’ — акуш., мегеб., кайт. *saʔaˤt*, кадар. *saʔaˤt*, ицар. *s:aˤʔaˤt*, кубач. *saˤħaˤtʼ*; ‘пятница’ — кадар., муир., тант., ицар., кайт., кубач. *žumaˤʔ*, акуш. *ʤumaˤʔ*; ‘змея’ — акуш., чираг., кубач. *maˤlʔun*, мегеб. *malʔun*.

примеры которой представлены в Таблице 7 (более точные количественные данные будут приведены ниже). Как видно из таблицы, согласный *ʔ* в таких словах возможен в разных позициях (конец слова; между гласными; после сонанта перед гласной), но в абсолютном большинстве примеров он встречается в начале слова, чаще всего перед *aˤ*[20].

Таблица 7. Тривиальные соответствия согласного ʔ

	кадар.	мегеб.	акуш.	муир.	тант.	ицар.	кайт.	чираг.	кубач.
земля, глина	ʔaˤnži	ʔaˤnža	ʔaˤnži	ʔaˤnčːa	ʔaˤnčːi	ʔaˤnčːa	ʔaˤnčːa	ʔančːija	ʔaˤnčːaˤ
жирный	ʔaˤwa-	ʔaˤwa-	ʔaˤwa-	ʔaˤwa-	ʔaˤwá-	ʔaˤʁaˤ-	ʔaˤwa-		
хороший	ʔaˤħ-	ʔaˤχ-	ʔaˤħ-	ʔaˤħ-	ʕaˤχ-	ʔaˤχː-	ʔaˤħ-	aχˤ-[21]	ʔaˤχː-
родник	ʔaˤniz		ʔiniʒ	ʔinic	ʔinic		ʔaˤnic	ʔincː	
запах	waˤʔ		gaˤʔ (*Л*)	kːaˤʔ		kːʷaˤʔ	kːʷaˤʔ		kːʷaˤʔ
широкий	CL-aˤʔ-uj	CL-aˤ(ʔ)-ul		CL-aʔa	CL-uʔaˤ-				
мерзнуть	CL-aˤrʔ-/ CL-irʔ-	CL-aˤrʔ-/ CL-irʔ-	CL-aˤrʔ-/ CL-irʔ-	CL-aˤrʔ-/ CL-irʔ-	CL-aˤrʔ-/ CL-irʔ-				CL-aˤʔ-/ CL-uˤʔ-
красть	CL-iʔ-/ CL-ilʔ-	CL-iʔ-/ CL-ilʔ-	CL-iʔ-/ CL-ilʔ-		CL-iʔ-/ CL-ilʔ-			CL-iʔʷ-/ CL-ilʔʷ-	CL-iʔ-/ CL-ilʔ-

Что касается согласного *ʔ*, то поскольку его присутствие на синхронном уровне определяется не во всех позициях и в целом он встречается не слишком часто, мы не знаем ни одной серии, где *ʔ* синхронно присутствовал бы во всех языках в ненулевой реализации.

Рассмотрим ряды соответствий, в которых *ʔ* в одних языках соответствует нулю в других. Почти во всех таких сериях *ʔ* обнаруживается в позиции после сонанта, причем непоследовательно: чаще он фиксируется в языках северной ветви, однако, как правило, не во всех и без очевидных закономерностей. Кроме того, в большинстве языков имеется глагольный корень ‘сказать’, состоящий в совершенном виде из одного согласного *ʔ*. Чаще всего этот корень синхронно реализуется как фонетический ноль, однако присутствие фонемы *ʔ* восстанавливается (1.3.1).

Таблица 8. Соответствия ʔ : Ø

	кадар.	мегеб.	акуш.	муир.	тант.	ицар.	кайт.	чираг.	кубач.
сказать	ʔ-	ʔ-	ʔ-		ʔ-	ʔ-		ʔ- (ʔ)	ʔ-
вдова	gamʔa	gamʔa	gamʔa	kːama	kːamʔa	kːama	kːama	kːama	kːama
истинный	mar-	marʔ-	mar-	mar-	mar-	mar-	mar-	mar-	
боб	qarʔa	qara	qara	qara	qara		qara	qàra	qa

Для этих и аналогичных слов можно реконструировать ларингал, который в некоторых языках дает рефлекс *ʔ*, а в других ноль. Вполне вероятно, что слов, имевших в своем со-

[20] Перед *и* (*иˤ*) тривиальное соответствие *ʔ* зафиксировано в идеофоне ‘кукарекать’: акуш., кадар. *ʔиʔи-*, мегеб. *ʔоˤʔоˤ-*, тант., ицар., *ʔиˤʔиˤ-*.

[21] Это слово, а также слово ‘широкий’ в мегебском имеют на месте *ʔ* ноль или *ʔ*.

ставе *ʔ*, гораздо больше (особенно в позиции начала слова), но в силу синхронных морфонологических и фонетических правил ларингал *ʔ* в этих словах не представлен ни в одном современном языке.

Помимо серий, показанных в таблицах 7 и 8, есть небольшое число слов, где в разных языках представлены *ʔ* или ноль, причем очевидных правил распределения по языкам или по позициям не видно.

Таблица 9. Нерегулярные соответствия ʔ : Ø

	кадар.	мегеб.	акуш.	муир.	тант.	ицар.	кайт.	чираг.	кубач.
девять	urč'eb-	urč'eb-	urč'em	urč'im-	urč'em	irč'em	irč'am-	arčàm-	ʔuˤč'uˤm
молотильная доска		irgi	irgi (*Л*)	ʔirč:i	irk:i		irk:i		i:č:e
камин	ʔaˤnʁi		anq:i	anq:i	aq:i	ʔaˤq:i		àq:a	aq:e

3.2. Соответствия h и ħ

В отличие от компаративных серий с *ʔ*, серий с тривиальными или почти тривиальными соответствиями согласных *h* и *ħ* немного (Таблицы 10 и 11 ниже).

В большинстве слов с тривиальными соответствиями *ħ* (Таблица 10) этот согласный стоит перед гласной, чаще всего *a*ˤ (реже *u*/*u*ˤ) — либо в начале слова, либо после *r*; он представлен также в конце слога и в глагольном корне (после гласной/после сонанта).

Отдельная проблема в сериях с *ħ* связана с чирагским языком. Списки слов для этого идиома не очень большие, а лексика сильно отличается от прочих языков, поэтому в большинстве серий, где во всех прочих языках представлен ларингал *ħ*, в чирагском соответствующей единицы нет. В двух словах согласному *ħ* всех прочих языков в чирагском соответствует *χ*, однако мы не можем быть уверенными в том, что нет других вариантов[22].

Таблица 10. Тривиальные и близкие к тривиальным соответствия ħ

	кадар.	мегеб.	акуш.	муир.	тант.	ицар.	кайт.	чираг.	кубач.
нутряной жир	ħaˤli	ħoˤli[23]	ħaˤli	ħali	ħaˤli	ħuˤI	ħaˤli	χˤùle	ħuˤle
ребенок	durħaˤ	doˤrħoˤ	durħaˤ	darħa	durħaˤ	duˤrħuˤ	daˤrħaˤ	darχˤà	du:ħuˤ
внутрь		CL-uħna-	CL-uħna	CL-uħna	CL-uˤħnaˤ				
бороться			CL-aˤħ-/ CL-irħ-	CL-iħ-/ CL-irħ-	CL-aˤħ-/ CL-irħ-	CL-aˤħ-/ CL-iħ-			

В сериях с тривиальными соответствиями *h* во всех словах, кроме одного, *h* стоит в начале слова перед *a*; в одном слове (‘медленный’) — после гласной *a* перед сонантом *l*.

[22] Например, полностью тривиальное соответствие представлено в идеофоне ‘ржать’ (акушинский, кадарский, мегебский, муиринский, ицаринский *ħiħi*, тантынский *hiˤħiˤ*, кубачинский *ħiħe*), но мы не используем идеофоны для установления соответствий. См. также чирагское слово ‘лопух’ в Таблице 12.

[23] Авторы мегебского словаря используют символ *o*ˤ для обозначения фонемы, которая соответствует фарингализованному *u*ˤ в других идиомах.

Таблица 11. Тривиальные соответствия h

	кадар.	мегеб.	акуш.	муир.	тант.	ицар.	кайт.	чираг.	кубач.
быстрый	halak-					halak-	halak-		halak-
легкий, нетрудный	hamad-		hamad-				hamat-		
вспоминать (коверб[24])	han		han	han	han	han	han		han-
медленный		bahla-	bahla-		bahla-	bahla-		bahlà-	bahla-

Как и для пары *ʔ*/*ʔ*, есть небольшое число компаративных серий (практически все они приводятся в Таблице 12), в которых *h* одних идиомов нерегулярным образом соответствует *ħ* других. В некоторых сериях (‘осел’ — в начале слова, ‘день’) *h* или *ħ* в одних идиомах соответствуют нулю в других. Пока мы не видим четких закономерностей в распределении *h*, *ħ* и нуля по языкам, можно лишь сделать наблюдение, что в северной ветви *ħ* встречается чаще. Согласные *ħ*/*h* в этих словах встречаются в разных позициях (в начале слова, в конце слова после гласной, между гласными)[25]. Интересно, что во втором слоге слова ‘осел’ в чирагском снова встречается *χ*, которое соответствует *h* и *ħ* других языков.

Таблица 12. Нерегулярные соответствия ħ : h : Ø

	кадар.	мегеб.	акуш.	муир.	тант.	ицар.	кайт.	чираг.	кубач.
лопух	ħaq'uli	ħaq'ur	ħaˤkul	[qːart]-ħaˤkul	haq'uli	haq'ul	haq'lej	haq'ùle	ħaˤqul
солнце, день	bari	barħi	barħi	barħi	beri	bar	bari	bàre	be
осел (начало слова; начало второго слога)	ħamħa	hamħa	emħe	himħa	ħaˤmħaˤ	amha	emha	àmχa	ħaˤmħaˤ
сыворотка		neħ	neħ (*Л*)	niħ	neħ	maˤħ	naˤħ	maːˤ	nihe

3.3. Проблема фарингализации и начало реконструкции

В разделах 3.1—3.2 мы рассмотрели часть соответствий постувулярных согласных (строки 1-4 из Таблицы 6), а именно — тривиальные и близкие к ним соответствия, а также те соответствия, в которых представлены только согласные, однородные по способу образования (или только смычные, или только фрикативные). В Таблице 13 рассмотренные соответствия приводятся еще раз, с указанием количества лексических единиц и позиций, в которых они засвидетельствованы.

Единственная хорошо представленная позиция — начало слова (перед гласной). В этой позиции появляются согласные *ʔ*, *ħ*, *h* (строки 1, 3, 4). *Л*аринга*л* *ʔ* (строка 2), как

[24] Коверб — первая часть сложного глагола, не употребляется вне его состава. В отличие от собственно глагольных корней, не выражает вид и, как правило, не изменяется.

[25] Ср. также идеофон ‘дышать’ (последний согласный): акушинский, кадарский, мегебский, муиринский, кайтагский *hih*, тантынский *seħ*, ицаринский *sːiħ*, кубачинский *sih*.

Таблица 13. Количественные данные для соответствий ʔ : ʔ и ħ : h[26]

	Соответствие	Всего	Перед гласной в начале слова)	Между гласными	После сонанта (перед гласной)	После гласной (в конце слова)	После гласной (перед сонантом)	После сонанта в глаг. корне	После гласной в глаг. корне
1	ʔ (везде)	22	14	2		2		4 {2}	{2}
2	ʔ : Ø	4	1		3				
3	ħ (: чираг. χ?)	7	4		1		1	1	{1}
4	h (везде)	7	6				1		
5	ʔ : Ø	3	3						
6	ħ : h : Ø	6	1	2	2	1			

правило, не идентифицируется в начале слова. Во всех позициях, кроме начала слова, количество имеющихся слов недостаточно для каких-либо обобщений.

Мы считаем, что для реконструкции прадаргинских постувулярных согласных ключевым является признак фарингализации. В разделе 1.3.2 мы привели сведения о фарингализации в различных даргинских языках, где попытались показать, что этот признак устроен по-разному в разных идиомах и что по крайней мере в некоторых из них (например, в кубачинском) его естественно считать просодической характеристикой слога или более крупного сегмента. Мы знаем также, что признак фарингализации тесно связан с согласными определенного места артикуляции: ларингальные согласные в большинстве даргинских языков несовместимы с фарингализованными гласными, а фарингальные, наоборот, несовместимы с простыми (при наличии соответствующего фарингализованного), см. примеры в (6) и Таблицу 3.

Связь фарингальных согласных с признаком фарингализации хорошо видна на данных нашей выборки: почти во всех компаративных сериях, входящих в описанные выше соответствия и содержащих *ħ* или *ʔ*, можно обнаружить след фарингализации — хотя бы в некоторых языках находится слово, содержащее фарингализованную гласную в одном из слогов, соседних с фарингальным согласным. Посмотрим, например, на корень ‘осел’ в Таблице 12. В кадарском, мегебском, акушинском и муиринском в нем содержится по меньшей мере один фарингальный согласный, но нет фарингализованных гласных. Однако такие гласные есть в тантынском и кубачинском.

Кроме того, в разделе 1.3.2 речь шла о том, что во многих даргинских языках существует явление, сходное с сингармонизмом по признаку фарингализации, но на синхронном уровне имеющее скорее морфологический характер. Это тоже заставляет предположить, что в этих языках фарингализация имела своей сферой действия не отдельные фонемы, а более крупные единицы.

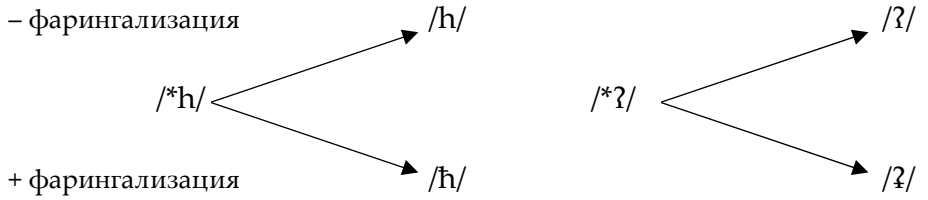
Предположим, что на прадаргинском уровне фарингализация была просодическим признаком, который распространялся по крайней мере на слог (фарингализованный слог, в свою очередь, мог вызывать фарингализацию в соседних слогах). Фонетически фарингализация скорее всего реализовывалась и на гласных, и на согласных. В этом слу-

[26] Здесь и ниже, в Таблицах 17 и 19, учитываются только позиции в пределах одной морфемы. Глагольные корни показаны в двух последних колонках, каждый корень учтен два раза, если он имеет два видовых варианта. Один из вариантов дан обычным образом, второй в фигурных скобках. При подсчете общего числа лексических единиц два видовых варианта глагольного корня учтены как одна единица.

чае для уже рассмотренных рядов соответствий (Таблица 13) достаточно реконструировать две исходных фонемы — *ʔ и *ħ. Можно допустить, что на прадаргинском уровне ларингальные и фарингальные согласные (ʔ / ʡ и h / ħ) не были противопоставлены как самостоятельные фонемы, но функционировали как позиционные варианты фонем /*ʔ/ и /*h/ соответственно, причем выбор зависел от наличия фарингализации: в условиях фарингализации (в соседстве с фарингализованным гласным) эти две фонемы реализовывалась как [ʡ] и [ħ], а в ее отсутствие — как [ʔ] и [h].

Со временем в части даргинских языков признак фарингализации перестал реализовываться на некоторых гласных (прежде всего на передних), но различия между фарингальными и ларингальными согласными (ʔ / ʡ и h / ħ) сохранились даже в отсутствие фонетического различия гласных по фарингализованности. Более того, когда эти гласные перестали различаться по фарингализованности, противопоставление фарингальных и ларингальных согласных стало фонологическим (ср. *himi* ‘желчь’ и *ħi* ‘кровь’ — даргинский литературный).

(13) Развитие фарингальных и ларингальных согласных



В тех слогах, где фарингализация продолжала реализовываться фонетически, произошло изменение статуса фарингализации: из просодического признака она превратилась в дифференциальный признак гласных (тех, на которых фарингализация сохранилась, чаще всего *aˤ*). Различия в современных парах типа *q:am* ‘племя’ и *q:aˤm* ‘миска’ (даргинский литературный) мы приписываем разным гласным фонемам, а не просодическому признаку фарингализации. При этом дополнительное распределение фарингальных и соответствующих им ларингальных в таких слогах сохранилось.

В Таблице 14 показана эволюция от прадаргинского к современным идиомам типа акушинского, в которых допустим только один фарингализованный гласный (*aˤ*). В качестве примера мы взяли ларингал *ħ в начале слога.

Таблица 14. Реконструкция перехода от прадаргинского к акушинскому: слоги структуры HV

Прадаргинский уровень			Современный уровень		Примеры
[*ha]	[*ħaˤ]		/ha/	/ħaˤ/	hala ‘перед’; ħaˤli ‘нутряной жир’
[*hu]	[*ħuˤ]	>	/hu/, /ħu/		hum ‘зоб’; ħu ‘ты’
[*he]	[*ħeˤ]		/he/, /ħe/		hes ‘старший’; emħe ‘осел’
[*hi]	[*ħiˤ]		/hi/, /ħi/		himi ‘желчь’; ħi ’кровь’

Предположение (13) объясняет тенденции к распределению постувулярных согласных в современных языках — например, явное преобладание фарингальных согласных перед *aˤ*: именно этот гласный может быть фарингализованным во всех языках даргинской группы[27].

[27] В то же время непонятно, почему в рядах (1) и (3) Таблицы 13 совсем нет слов, начинающихся с сочетания фарингального согласного с узким гласным. Ответ на этот вопрос будет частично представлен в разделе 4.

В то же время наша гипотеза предсказывает, что в каждом конкретном слове все языки выборки должны последовательно иметь либо *h*, либо *ħ* (а также либо *ʔ*, либо *ʔ*), хотя, как мы знаем, есть слова, для которых это неверно: это формы с нерегулярными соответствиями *ʔ : ʔ : Ø* и *ħ : h : Ø* (Таблицы 9 и 12). По-видимому, в части диалектов фарингализация в каких-то условиях исчезла еще до фонологизации различия между фарингальными и ларингальными согласными. Условия такого развития пока неясны.

Еще один вопрос, который возникает в связи с нашей гипотезой, — это противопоставление *a* и *aˤ* после фарингальных согласных в муиринском и некоторых других идиомах. Наша гипотеза не объясняет, как оно стало возможным. Рассмотрим примеры муиринских слов на *ħa, ħaˤ, ʔa* и *ʔaˤ* (Таблица 15):

Таблица 15. Слова, начинающиеся на *ħa, ħaˤ, ʔa* и *ʔaˤ* в муиринском, и их соответствия в других идиомах

	кадар.	мегеб.	акуш.	**муир.**	тант.	ицар.	кайт.	чираг.	кубач.
				ħa, ʔa					
подушка	ʔaˤnija	ʔaˤnela	ʔaˤnala	**ʔanila**	ʔuˤnela			ʔänuna	
узел	ʔaˤnd	ʔaˤndari	ʔaˤnd	**ʔat**					
широкий	CL-aˤʔuj	CL-aˤ-ul		**CL-aʔa-**	CL-uʔaˤ-				
сын, мальчик	durħaˤ	doˤrħoˤ	durħaˤ	**darħa**	durħaˤ	duˤrħuˤ		darχˤà	duːħuˤ
осел (второй слог)	ħamħa	hamħa	emħe	**himħa**	ħaˤmħaˤ	amha	emha	àmχa	ħaˤmħaˤ
жир нутряной	ħaˤli	ħoˤli	ħaˤli	**ħali**	ħaˤli		ħaˤl-la	χˤùle	ħuˤle
				ħaˤ, ʔaˤ					
лягушка	ʔaˤtʼa	ʔaˤtʼa		**ʔaˤtʼa**	ʔaˤtʼa	ʔaˤtʼa	ʔaˤtʼa	ʁˤàtʼa	ħitʼa
жирный	ʔaˤwa-	ʔaˤwa-		**ʔaˤwa-**	ʔaˤwá-	ʔaˤʁaˤ-	ʔaˤwa-		
тайный	diʔaˤm		diʔaˤn	**diʔaˤn-**	diʔaˤn-		diʔaˤn-		
лопух	ħaqʼuli	ħaqʼur		**[qːart]-ħaˤkul**	haqʼuli		haqʼlej	haqʼùle	ħaˤqul
шуба				**ħaˤka**	ħaˤka	ħaˤka			
черемша				**ħaˤnku**			ħaˤnkur		ħaˤnku

Можно предположить, что противопоставление *ħaˤ ~ ħa* и *ʔaˤ ~ ʔa* в муиринском исторически носило позиционный характер: похоже, что гласная *aˤ* в определенных контекстах утратила фарингализацию. У нас пока нет реконструкции даргинских гласных и мы не можем указать эти контексты точно. Можно, однако, заметить, что после *ħa* или *ʔa* в муиринском, как правило, идет сонант и хотя бы в одном из идиомов на месте *a* обнаруживается *u*, *uˤ* или *oˤ*; напротив, в тех слогах, где в муиринском *ħaˤ* или *ʔaˤ*, во всех когнатах также представлена гласная *aˤ*[28].

Таким образом, мы предполагаем, что на прадаргинском уровне в словах типа ‘жир нутряной’ корень был фарингализованным и, возможно, содержал гласный, отличный от *a*; обозначим его *U*. Впоследствии согласная фонологизировалась как /*ħ*/, а гласная в муиринском перешла в простое *a*; в других идиомах она сохранила фарингализацию и перешла в *uˤ (oˤ)* или *aˤ*, как показано в (14):

[28] Косвенным подтверждением вторичности муиринского случая может служить тот факт, что в кадарском, где и *a* и *aˤ* возможны после *ħ (ʔ)*, распределение *ħa* / *ħaˤ* в целом не совпадает с муиринским, ср. в таблице 15 слова ‘мальчик, сын’, ‘лопух’, ‘жир нутряной’.

(14) */hUli/$^{+\text{фарингализация}}$ > */ħU^{ʕ}li/ > /ħali/ (муиринский)
/ħa^{ʕ}li/ (кадарский)
/ħu^{ʕ}le/ (кубачинский)

Если распределение *a* и *a*^{ʕ} после фарингальных согласных в муиринском действительно возникло как позиционное, то возможность употребления *a* и *a*^{ʕ} после фарингальных в этом идиоме не противоречит нашей гипотезе.

4. Соответствия смычных и фрикативных ларингалов (ʔ/ʔ vs. h/ħ)

Небольшая группа слов последовательно противопоставляет постувулярный смычный ларингал в «северных» языках фрикативному в «южных». При этом в «северных» языках это почти всегда *ʔ*, а в «южных» встречаются *h* и *ħ* (*h* гораздо чаще). В двух случаях такое соответствие представлено в начале слова, на остальные позиции также приходится по одному-два примера.

Таблица 16. Соответствия ʔ ~ ħ

	кадар.	мегеб.	акуш.	муир.	тант.	ицар.	кайт.	чираг.	кубач.
назад			ʔela	ʔila	ħela	hela		hila-	hila-
яблоко	ʔinc	inc	ʔinc	ʔinci	ħinc	hinci	hinci	ħinc	hinze
дешевый			durʔe-	durʔa-	durʔa^{ʕ}-[29]	durha-	durha-	durha-	duha-
камыш		č'a^{ʕ}ʔa^{ʕ}	č'iʔa^{ʕ}						č'iħa^{ʕ}
сноп				buʔmi	buhemi		biham	bahàm	buhum
хозяин	CL-eʔ	weʔ	CL-eʔ	CL-aʔ	CL-e^{ʕ}ʔ	CL-ah	wah	wah	CL-ah

В Таблице 17 ниже приводятся количественные данные для всех соответствий, рассмотренных в разделах 3 и 4. Ряды, восходящие, в соответствии с нашими предположениями, к одной прафонеме, объединены в одной строке (строке 1 в Таблице 17 соответствуют строки 1–2 из Таблицы 13, аналогично строке 2 соответствуют строки 3–4). Нерегулярные соответствия (*ʔ : ʔ : Ø*, *ħ : h : Ø*) в таблицу не включены.

Таблица 17. Количественные данные для соответствий *ʔ : ʔ*, *ħ : h* и *ʔ : ħ*

	Прафонема	Рефлексы	Всего	Перед гласной (в начале слова)	Между гласными	После сонанта (перед гласной)	После гласной (в конце слова)	После гласной (перед сонантом)	После сонанта в глаг. корне	После гласной в глаг. корне
1	*ʔ	ʔ, ʔ, 0	26	15	2	3	2		4 {2}	{2}
2	*h	ħ, h, Ø	14	10		1		2	1	{1}
3	*hʔ	ʔ ~ h	6	2	2	1	1			

[29] Смычный ларингал в словах ‘дешевый’ и ‘хозяин’ в тантынском представляет собой очевидное исключение — возможно, это заимствование из одного из северных языков, которые в целом характерны для тантынского (см. подробнее раздел 5), но не исключено и позиционное распределение (*ʔ* после сонанта и на конце слова, *ħ* перед гласной). Для более определенных выводов пока недостаточно данных.

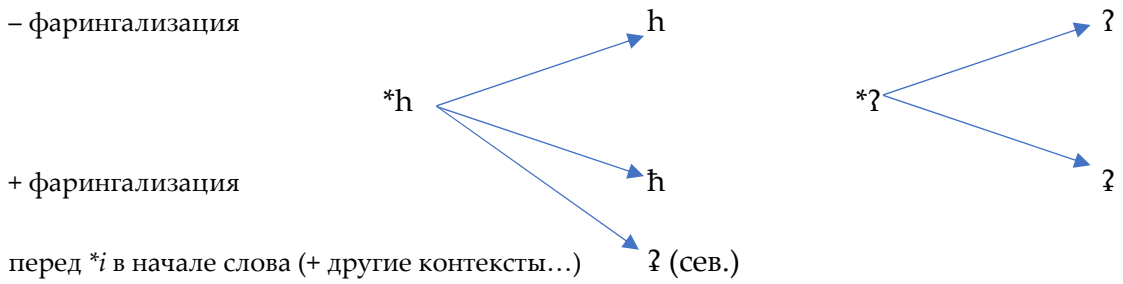
Все три группы соответствий возможны в позиции начала слова перед гласной, однако в словах с соответствием *ʔ : ħ* (Таблица 16) во всех идиомах мы видим после ларингала узкий гласный. Проблема в том, что у нас всего две таких серии (‘назад’, ‘яблоко’). В рядах соответствий, рассмотренных в предыдущих разделах, ни одного слова с такими гласными после постувулярного в начальной позиции не было. Иначе говоря, в позиции начала слова ряд 3 (*ʔ ~ ħ*) находится в дополнительной дистрибуции как с группой слов в первой строке (*ʔ, ʔ, ∅*), так и со словами во второй (*ħ, h, ∅*). То есть, если опираться исключительно на позицию начала слова, то можно допустить, что серии в ряду 3 (Таблица 16) могут восходить к любой из ранее постулированных согласных (**ʔ, *h*).

В остальных позициях (кроме начала слова) ряд 3 ближе к дополнительной дистрибуции с рядом 2, однако слов в этих позициях так мало, что серьезные выводы вряд ли возможны. Мы тем не менее считаем, что ряд 3 действительно представляет ту же прафонему, что и ряд 2. В таком случае получается, что в севернодаргинских языках согласный **h* переходил в *ʔ* перед каким-то узким гласным (скорее всего **i*), а также, возможно, в других позициях (между каким-то гласными? после каких-то гласных в конце слова?). Противоположное предположение (переход **ʔ* в *h/ħ*) означало бы, что мы допускаем общую инновацию для всех «южных» языков, включая чирагский и кайтагский, в то время как данные лексикостатистики показывают, что так называемые «южные» языки, в отличие от «северных», не образуют генетического единства.

Мы не обсудили вопрос о фарингализации в словах с соответствием *ʔ : ħ/h*. Поскольку во всех «северных» языках почти без исключений мы находим в этой группе фарингальный *ʔ*, а в «южных» встречается и *ħ*, и *h* (причем *h* гораздо чаще), можно осторожно предположить, что переход **h* > *ʔ* происходил только в условиях фарингализации, причем в части южных языков (ицаринском, кайтагском, возможно, чирагском) на слогах, начинающихся с *hi*, фарингализация исчезала, а в тантынском сохранялась, возможно, давая разный результат в зависимости от позиции. Но для сколько-нибудь надежных выводов пока недостаточно данных.

Таким образом, пока мы выделили два прадаргинских ларингала, которым соответствуют три возможных рефлекса:

(15) Развитие согласных **ʔ* и **h*



5. Соответствия фрикативных vs смычных ларингалов (h/ħ : ʡ/ʔ)

Мы не рассмотрели еще одну большую группу соответствий (строка 6 в Таблице 6), устроенных практически зеркально по отношению к соответствиям, рассмотренным в разделе 4: в этой группе фрикативный согласный в «северных» языках в основном соответствует смычному в «южных». По сравнению с соответствиями, рассмотренными в разделе 4, компаративных серий такого типа довольно много, более 25. Основных рядов со-

ответствий в этой группе четыре: их различает рефлекс, представленный в тантынском идиоме (*ʕ, ħ, ʔ, Ø*). В Таблице 18 мы рассматриваем эти ряды отдельно (расхождения между *ʔ, ʔ* и нулем в других идиомах мы обсудим ниже).

Таблица 18. Компаративные серии ħ : ʕ/ħ/ʔ/Ø : ʔ/ʔ/Ø

	кадар.	мегеб.	акуш.	муир.	тант.	ицар.	кайт.	чираг.	кубач.
Группа 1. ħ : ʕ : ʔ/ʔ									
ты	ħu	ħu	hu	ħu	ʕuˤ	u	i	u	u
метка		ħarzi	ħerzi	ħirc:i	ʕeˤrc:i	arc:i	erc:i	ʔàrc:e	
глаз	ħuli	ħuli	ħuli	ħuli	ʕuˤli	ʔuˤl	uli	ʔùle	ule
рот, клюв	muħli	muħli	muħli	muħli	muˤʕuˤli			muʔùl	mule
восемь	gaħ-	gaħ-	geħ	k:aħ-	k:eˤʕ	k:a:	gaʔ-	k:aˤʔ-	k:a
Группа 2. ħ : h : ʔ/ʔ									
работа	ħaˤnki	ħaˤnči	ħaˤnči	ħaˤnči	ħaˤnči	ʔaˤči	ʔaˤču	ʔaˤč	ʔaˤče
красный	ħint'aga	ħunt'-	ħunt'en-	ħint'in	ħint'in-	it'in	it'in-	ʔat'ìn-	it'in-
река	ħark'	ħarkʷ'	ħerk’	ħirk'	ħeˤrkʷ'	arkʷ'	erk'	ʔarkʷ'	akʷ'
Группа 3. ħ : ʔ : ʔ/ʔ									
овод, слепень	ħimiz	ħumerz	ħunz		ʔimiz				umic
гнить	CL-erħ-/ uħ-	CL-erħ-/ uħ-	CL-erħ-/ uħ- (*Л*)	CL-erħ-/ uħ-	CL-erʔ-/ CL-uʔ-	CL-erʔ-/ CL-uʔ	CL-erʔ/ʔ	CL-VrʔCL-iʔ	
Группа 4. ħ : Ø : ʔ/ʔ									
лицо	baˤħ	baˤħ	baˤħ	baˤħ	CL-aˤ	daˤʔ	baˤʔ	daːˤ	CL-aˤj
мозг	meħa	meħʷa	meħe	maħħa	meˤ	mua	maʔ	maʔ	ma:

В нескольких сериях нет тантынского рефлекса, поэтому неясно, к какому ряду соответствий их следует относить (например, ‘обрыв, край’: кадарский, мегебский *beħ*, акушинский *baˤħ*, кайтагский *baʔ*, чирагский *baʔ*, кубачинский *baj*).

Рассмотрим все эти группы более подробно (количественные данные приводятся в Таблице 19).

Таблица 19. Количественные данные для соответствий «фрикативный в северных языках vs смычный в южных»

Группа	Соответствие (северн. ~ тант. ~ южн.)	Всего	Перед гласной (в начале слова)	Между гласными	После сонанта (перед гласной)	После гласной (перед сонантом)	После гласной (в конце слова)	После сонанта в глаг. корне	После гласной в глаг. корне
1	ħ : ʕ : ʔ/ʔ/Ø	12	10	2			1		
2	ħ : ħ : ʔ/ʔ/Ø	6	5		1				
3	ħ : ʔ : ʔ/ʔ/Ø	2	1					1	{1}
4	ħ : Ø : ʔ/ʔ/Ø	3					3		
5	ħ : ʔ : ʔ/ʔ/Ø	5			1 {1}		3		

Постувулярные согласные в группе 4 близки к дополнительному распределению с согласными в группах 1 и 2. Можно предположить, что Ø на конце в тантынском является рефлексом той же фонемы, что представлена в группе 1 и/или 2. Большая часть согласных в группах 1 и 2 находится в позиции начала слова, причем в обоих случаях возможны разные гласные, так что дополнительного распределения между группами 1 и 2 нет. В строке 3 всего два слова, поэтому тут наиболее вероятны какие-то случайные отклонения.

В группе 1 в тантынском идиоме представлен редкий для даргинских языков ларингал ʕ. В тантынском не очень много морфем с этим согласным, однако ларингал ʕ зафиксирован в таких частотных единицах, как личные местоимения второго лица *ʕuˤ* ‘ты’ и производное от него *ʕuˤxːa* ‘вы’, а также отрицательный префикс *ʕaˤ-*. Само наличие этого согласного, его дистрибуция и соответствия ставят несколько вопросов.

Первый из них — вопрос о том, является ли согласный ʕ инновацией или архаизмом. Насколько нам известно, ларингал ʕ обнаружен пока только в трех даргинских идиомах: помимо тантынского, это не вошедшие в нашу выборку диалекты селений Худуц и Урцаки[30]. Хотя наши данные о худуцком и особенно урцакинском идиоме неполны[31], мы уверены в том, что речь не идет о трех независимых инновациях: ларингал ʕ в этих идиомах зафиксирован в когнатных единицах (см. ниже).

Тантынский, худуцкий и урцакинский идиомы относятся к юго-западно-даргинскому языку (Коряков 2021), но, насколько это известно, не образуют внутри этого языка какого-либо единства[32]. Все три селения расположены в сравнительно высокогорной части даргинского ареала, для которой в целом характерно присутствие архаичных черт. При этом Танты, Худуц и Урцаки не расположены рядом, между ними и поблизости есть селения, в которых ларингал ʕ отсутствует (в точности это известно, в частности, про идиомы селений Кунки, Ицари, Санжи, Цугни и т. д.). Следовательно, общая инновация вряд ли возможна, и перед нами скорее архаичный признак. Можно допустить, что в прадаргинском существовал аналогичный ларингал *ʕ, который переходил в смычный постувулярный (ʡ, ʔ) в «южных» языках и в *ħ* в «северных». Это допущение, однако, не отвечает на вопрос о том, сколько прафонем представляют четыре группы соответствий, показанных в Таблице 18.

Различия между группами 1, 2 и 3 (ʕ, *ħ*, ʡ в тантынском) в принципе можно объяснить за счет междиалектных заимствований, особенно для группы 3, так как в ней совсем мало слов. Что касается группы 2, где в тантынском представлен *ħ*, то можно предположить, что это влияние северных диалектов, что вполне логично, так как селение Танты расположено на стыке зоны юго-западно-даргинского и зоны севернодаргинских языков. С северной стороны ближайшим к нему (5–6 км) является селение Гапшима с языком северно-центрально-даргинской группы, а сразу за Гапшимой — районный центр Аку-

[30] Как отмечается в работе Беляев 2021, этот звук также сохранился в произношении некоторых носителей ширинского диалекта, например в слове (ʕ)*äqˀ* ‘стадо’. В позиции после предполагаемого *ʔ гласные в ширинском фарингализованы, но этот тип фарингализации акустически и артикуляционно отличается от «первичной» фарингализации, встречающейся в других позициях; подробнее см. в Беляев 2021. Сходную ситуацию описывает Д. С. Ганенков (Ganenkov 2019a) для чирагского языка. В этом идиоме согласный ʕ пал, но его след виден в *эпиглоттализованности* следовавших за ним гласных, например, в словах *uˀ* ‘ты’ and *uˀle* ‘глаз’ (признак эпиглотализованности отличается от признака фарингализованности).

[31] Для идиома селения Урцаки мы располагаем только стословным списком (полевые данные, любезно представленные М. А. Даниэлем), в котором согласный ʕ представлен только в семи словах.

[32] Ширинский и чирагский идиомы, о которых шла речь в сноске 30, не относятся к юго-западно-даргинскому языку: ширинский входит в южнодаргинскую группу и образует в ней отдельную диалектную подгруппу вместе с амузгинским (Беляев 2019); чирагский образует отдельную подгруппу в составе даргинской группы.

ша, где говорят на акушинском диалекте. Традиционно тантынцы активно общаются с носителями северных даргинских диалектов — в частности, при посещении Акуши и Ая-базара, расположенного в Левашинском районе Дагестана. С южной стороны от с. Танты, через перевал, расположены селения, где говорят на диалектах юго-западно-даргинского языка (Ургани, Цугни, Урари, Гуладты и др.), однако эти селения меньше и, судя по всему, тантынцы посещают их гораздо реже, чем Акушу. В то же время количество слов в группе 2 довольно велико, что все же заставляет сомневаться в том, что фонетический облик этих слов — результат междиалектного влияния.

Чтобы попробовать разобраться с этим, обратимся к данным диалектов с. Урцаки и Худуц. В стословнике идиома с. Урцаки представлено семь единиц, содержащих ларингал ʕ. Для худуцкого мы располагаем словником примерно в 650 единиц и грамматическим очерком диалекта (Ganenkov 2019b). Обнаруживается, что в худуцком, а иногда и в урцакинском идиоме ʕ обнаруживается в некоторых из тех слов, где в тантыском мы видим *ħ*[33] (примеры в Таблице 20). Более того, согласный ʕ представлен здесь в слове ‘гнить’, где в тантынском мы имеем *ʔ*. Видимо, наше предположение о вторичности согласных *ħ* и *ʔ* в тантынских словах групп 2 и 3 было верным — для всех слов групп 1–3 мы можем предположить исходный ларингал *ʕ.

Таблица 20. Звонкий фрикативный ларингал в идиомах селений Танты, Худуц, Урцаки

	с. Танты	с. Худуц	с. Урцаки
кровь	ʕeˤ		ʕe
глаз	ʕuˤli	ʕul	ʕuli
три	ʕaˤb-	ʕab-	
восемь	kːeˤʕ	kːaʕ-	
ты	ʕuˤ	ʕu	ʕu
префикс отрицания	ʕaˤ-	ʕa-	ʕa-
работа	ħaˤnči*	ʕači	
река, ручей	ħeˤrkʼʷ	ʕerkʼʷ	
гнить	CL-erʔ-	CL-erʕ-	

Таким образом, соответствия, рассмотренные в этом разделе, позволяют реконструировать еще один прадаргинский ларингал (*ʕ), к которому восходит ʕ в некоторых современных идиомах (тантынском, худуцком, урцакинском), *ʔ*/ʔ в прочих языках «южного» типа и *ħ* в «северных» даргинских языках. В конце слова *ʕ в тантынском выпадает (группа 4 в Таблице 18).

Вопрос о фарингализации в слогах с *ʕ не вполне ясен, но поскольку примеров на соответствия с этой фонемой довольно много, то какие-то предположения строить можно. Во-первых, в «северных» языках *ʕ дает *ħ практически без исключений, то есть независимо от фарингализации. В «южных» языках рефлексы *ʕ различаются (Таблица 21), причем в части языков результат, судя по всему, зависит от наличия исходной фарингализации в слоге (третья колонка):

[33] Обратная ситуация зафиксирована в слове ‘рот’: тант. *muˤʕuˤli*, худуц. *muˤħli*, урцак. *muʕli* (худуцкое слово найдено в словнике, который нуждается в дополнительной проверке).

— в тантынском, как мы знаем, есть примеры с разными рефлексами *ʕ (ʕ, *ħ*, *ʔ*, ноль), но практически во всех словах с исходным *ʕ присутствует фарингализация, то есть, очевидно, согласный *ʕ вызывал фарингализацию в соседнем слоге, даже если исходно ее там не было;

— в ицаринском и кайтагском, судя по всему, *ʕ давал разные результаты в зависимости от фарингализованности слога: *ʔ* в условиях фарингализации и ʔ/ноль в ее отсутствие. Корреляция между ицаринским и кайтагским почти полная, ср. примеры в (21) (исключение — ‘глаз’);

— в кубачинском *ʕ всегда дает ноль перед *u*, *i*; на конце слова возможны *j* или ноль (правила распределения пока неясны); перед кубачинским *a* возможны и *ʔ*, и ноль (ср. ‘три’ и ‘лето’), причем наблюдается совпадение по фарингализованности с ицаринским и кайтагским.

Ситуация с чирагским сложнее, так как есть неясность с представлением данных: *Д.* С. Ганенков, которому известно о существовании согласного *ʕ в ряде даргинских идиомов, пишет в Ganenkov 2019a, что на месте исчезнувшего *ʕ в чирагском обнаруживается ноль с эпиглоттализацией последующего гласного (во всех приведенных в его очерке примерах *ʕ находился в начале слова), но не сказано, верно ли это для всех вхождений *ʕ — судя по всему, нет, ср. чирагское слово ‘три’ (подробный разбор этой проблемы, разумеется, не входил в задачи грамматического очерка). Мы добавили примеры из очерка в Таблицу 21 (вторая колонка справа), но для каких-либо выводов этих примеров недостаточно. Если же опираться на записи из более ранних источников (преимущественно Кибрик & Кодзасов 1988; 1990), то почти во всех словах с исходным *ʕ обнаруживается *ʔ* (за исключением слова ‘лицо’).

Таблица 21. Рефлексы исходного *ʕ в «южных» даргинских языках

	Позиция	Фаринг.	тант.	ицар.	кайт.	чираг.	чираг. (Ganenkov 2019a)	кубач.
колесо	#_u	+	ʕuˤla	ʔuˤla	ʔula	ʔùla		ule
глаз	#_u	+	ʕuˤli	ʔuˤl	uli	ʔùle	uˀle[34]	ule
три	#_a	+	ʕaˤb	ʔaˤ:b	ʔaˤb-	ʔàˤb-	ʔàˤb-	ʔaˤb
работа	#_a	+	ħaˤnči	ʔaˤči	ʔaˤču	ʔaˤč		ʔaˤče
лицо	a_#	+	CL-aˤ	daˤʔ	baˤʔ	da:ˤ	daˀ[35]	CL-aˤj
ты	#_u	–	ʕuˤ, ʕuˤx:a	u	i	ʔu	uˀ	u
префикс отрицания	#_a	–	ʕaˤ-	a-	a-		aˀ-	(a-)
река	#_E	–	ħeˤrkʼʷ	arkʼʷ	erkʼ	ʔarkʼʷ	aˀrkʼ-	akʼʷ
рот, клюв	V_V		muˤʕuˤli			muʔùl		mule
гнить	корень глаг.	–	CL-erʔ-/ CL-uʔ-	CL-erʔ-/ CL-uʔ	CL-erʔ/ CL-uʔ	CL-VrʔV		CL-iʔ
мозг	V_#	–	meˤ	mua	maʔ	maʔ		ma:

[34] В работе *Д.* С. Ганенкова диакритический знак ˤ обозначает эпиглоттализацию, но поскольку у нас этот знак используется для обозначения фарингализации, мы обозначили эпиглоттализацию «зеркальным» значком ˀ.

[35] *Д.* С. Ганенков отмечает, что, в отличие от Кибрик & Кодзасов 1990, он не слышит долготы в этом слове.

6. Заключение

В заключение представим картину развития ларингалов в даргинских языках так, как она вырисовывается в соответствии с нашими предположениями.

(1) Мы постулировали для прадаргинского языка три постувулярных согласных: **ʔ*, **ħ*, **ʕ* и предположили, что в прадаргинском языке действовал просодический признак фарингализации, который распространялся по меньшей мере на слог и реализовывался фонетически как фарингализованное произношение гласных, а скорее всего — и согласных. Под воздействием фарингализации фонемы /*ʔ/, /*h/ реализовывались как [*ʢ] и [*ħ] соответственно.

(2) Со временем в большинстве даргинских языков реализация признака фарингализации на части гласных стала невозможной. По этому параметру современные даргинские языки различаются: в некоторых из них фарингализация по-прежнему возможна на всех гласных (в нашей выборке это кубачинский; аналогичная ситуация в близкородственных кубачинскому аштынском и ширинском, а также худуцком и чирагском). В других языках фарингализованными могут быть от одной до трех согласных (раздел 1.3.2).

Хотя фарингализация некоторых гласных стала невозможной, различие согласных *ħ vs. h* и *ʔ vs. ʢ* сохранилось и фонологизировалось: в результате в современных даргинских языках возникли фонемы /ħ/, /ʢ/.

(3) В «северных» даргинских языках (в нашей выборке это акушинский, кадарский, мегебский, муиринский), помимо этого, произошел переход **h* > *ʔ* в позиции перед некоторым узким гласным (предположительно **i*) и, возможно, в каких-то других контекстах.

(4) Согласный **ʕ* в «северных» языках утратил звонкость и перешел в **ħ* (видимо, во всех контекстах), а в прочих языках перешел в один из смычных постувулярных согласных (*ʔ/ʔ*), сохранившись лишь в некоторых малых (одноаульных) идиомах, которые, по классификации Ю. Б. Корякова, относятся к юго-западно-даргинскому языку.

Разумеется, наши предположения оставляют нерешенными довольно много проблем и объясняют не все слова (в особенности это относится к соотношению *ħ/h* и *ʔ/ʔ*) — вероятно, часть из них можно будет объяснить, когда мы будем располагать мотивированной реконструкцией гласных и всех остальных согласных.

Кроме того, необходимо гораздо подробнее рассмотреть отдельные даргинские идиомы. Для многих из них диахронические исследования затруднены из-за отсутствия качественных описаний фонологии и морфонологии, а также из-за проблем с фиксацией фонологических признаков в имеющихся словарях. Будем надеяться, что в разрешении этих проблем нам поможет документация даргинских языков, которая в последние годы идет очень активно.

Сокращения

В тексте работы: C — любой согласный; H — постувулярный согласный; R — сонант или другой согласный, способный входить в кластер (см. раздел 1.3.3); V — гласный; *Л* — слово из словаря даргинского литературного языка (Юсупов 2017).

В глоссах: ATTR — атрибутивный показатель, CAUS — каузатив, DOWN — преверб ‘вниз’, FUT — будущее время, INF — инфинитив, IPFV — имперфектив, M — мужской класс, MSD — масдар, N — неличный класс единственного числа, NPL — неличный класс множественного числа, N.ABSTR — абстрактное существительное, OPT — оптатив, PFV — перфектив, PL — множественное число, PRET — претерит, SUBJ — конъюнктив, TH — тематический элемент (показатель одной из глагольных основ), 3 — 3е лицо, TR — переходный глагол.

Литература

Абдуллаев, З. Г. 2017. *Даргинско-русский словарь: 12 000 слов, терминов и выражений [Дарган-урус словарь]*. Махачкала: Институт ЯЛИ ДНЦ РАН.

Абдуллаев, С. Н. 1954. *Грамматика даргинского языка (фонетика и морфология)*. Махачкала: Дагестанский филиал АН СССР.

Беляев, О. И. 2019. Место ширинского и амузгинского среди диалектов даргинского языка. *Томский журнал лингвистических и антропологических исследований* (*Tomsk Journal of Linguistics and Anthropology*) 2(24): 20–38.

Беляев, О. И. 2021. Два типа фарингализации в ширинском даргинском: предварительные наблюдения. В кн.: Т. А. Майсак, Н. Р. Сумбатова, Я. Г. Тестелец (ред.). *Дурхъаси хазна. Сборник статей к 60-летию Р. О. Муталова*: 18—29. Москва: Буки Веди.

Быховская, С. Л. 1940. Пережитки inclusiv-а и exclusiv-а в даргинских диалектах. В кн.: И. И. Мещанинов (ред.). *Язык и мышление IX*: 85–100. Москва / Ленинград: Академия наук СССР.

Вагизиева, Н. А. 2018. *Лексико-фразеологические особенности кадарского диалекта даргинского языка*. Диссертация на соискание ученой степени канд. филол. наук. Махачкала: Дагестанский государственный педагогический университет.

Гасанова, У. У. 2011. *Хайдакско-русский словарь*. Махачкала.

Кадибагамаев, А. А. 1998. *Сирхинский диалект даргинского языка (фонетическая система).* Махачкала: ДНЦ РАН.

Кибрик, А. Е., С. В. Кодзасов. 1988. *Сопоставительное изучение дагестанских языков. Глагол*. Москва: Издательство МГУ.

Кибрик, А. Е., С. В. Кодзасов. 1990. *Сопоставительное изучение дагестанских языков. Имя. Фонетика*. Москва: Издательство МГУ.

Коряков, Ю. Б. 2021. Даргинские языки и их классификация. В кн.: Т. А. Майсак, Н. Р. Сумбатова, Я. Г. Тестелец (ред.). *Дурхъаси хазна. Сборник статей к 60-летию Р. О. Муталова*: 139–154. Москва: Буки Веди.

Коряков, Ю. Б., Н. Р. Сумбатова. 2007. Даргинские языки. В кн.: *Большая российская энциклопедия. Т. 8*: 328–329. М.: Большая российская энциклопедия.

Магомедов, А. Дж., Н. И. Саидов-Аккутта. 2017. *Кубачинско-русский словарь*. Москва: Наука.

Магометов, А. А. 1963. *Кубачинский язык (исследование и тексты).* Тбилиси: Изд-во АН Грузинской ССР.

Магометов, А. А. 1976. Корреляция системы консонантов в диалектах даргинского языка. *Ежегодник иберийско-кавказского языкознания III*: 219–225.

Магометов, А. А. 1982. *Мегебский диалект даргинского языка.* Тбилиси: Мецниереба.

Мудрак, О. А. 2016. *Даргинские основы. Т. 1*. Москва: Языки народов мира.

Мусаев, М.-С. М. 2002. *Даргинский язык*. Москва: Academia.

Муталов, Р. О. 2002. *Глагол даргинского языка*. Махачкала: ДГУ.

Муталов, Р. О. 2021. Классификация даргинских языков и диалектов. *Социолингвистика 3 (7)*: 8–25.

Сумбатова, Н. Р., Ю. А. Ландер. 2014. *Даргинский говор селения Танты. Грамматический очерк. Вопросы синтаксиса*. Москва: Языки русской культуры.

Темирбулатова, С. М. 2004. *Хайдакский диалект даргинского языка*. Махачкала: ДНЦ РАН.

Темирбулатова, С. М. 2022. *Диалектологический словарь даргинского языка.* Махачкала: Алеф.

Хайдаков, С. М. 1985. *Даргинский и мегебский языки*. Москва: Наука.

Шахбанова, П. Г. 2011. *Фонетические и морфологические особенности карбачимахинского говора даргинского языка*. Диссертация на соискание ученой степени канд. филол. наук. Махачкала: Дагестанский государственный педагогический университет.

Юсупов, Х. А. 2017. *Дарган-урус словарь Даргинско-русский словарь: около 40 000 слов и фразеологических выражений*. Махачкала: Институт ЯЛИ ДНЦ РАН.

References

Abdullaev, S. N. 1954. *Grammatika darginskogo jazyka (fonetika i morfologija)*. Makhachkala: Dagestanskij filial AN SSSR.

Abdullaev, Z. G. 2017. *Darginsko-russkij slovar': 12 000 slov, terminov i vyrazhenij*. Makhachkala: Institut JaLI DNC RAN.

Beljaev, O. I. 2019. Mesto shirinskogo i amuzginskogo sredi dialektov darginskogo jazyka. *Tomskij zhurnal lingvistichekix i antropologicheskix issledovanij (Tomsk Journal of Linguistics and Anthropology) 2(24)*: 20–38.

Beljaev, O. I. 2021. Dva tipa faringalizacii v shirinskom darginskom: predvaritel'nyje nabl'udenija. In: T. A. Majsak, N. R. Sumbatova, Ya. G. Testelets (eds.). *Durqasi khazna. Sbornik statej k 60-letiju R. O. Mutalova*: 18–29. Moscow: Buki Vedi.

van den Berg, Helma. 2001. *Dargi folktales. Oral stories from the Caucasus. With an introduction to Dargi grammar*. Leiden: Research School CNWS.

Bykhovskaja, S. L. 1940. Perezhitki inclusiva i exclusiva v darginskix dialektax. In: I. I. Meshchaninov (ed.). *Jazyk i myshlenije IX*: 85–100. Moscow / Leningrad: Akademija nauk SSSR.

Forker, Diana. 2020. *A grammar of Sanzhi Dargwa* (*Languages of the Caucasus 2*). Berlin: Language Science Press.

Ganenkov, Dmitry. 2019a. *Chirag Dargwa*. Ms.

Ganenkov, Dmitry. 2019b. *Khuduts Dargwa*. Ms.

Gasanova, U. U. 2011. *Khajdaksko-russkij slovar'*. Makhachkala.

Jusupov, Kh. A. 2017. *Dargan-urus slovar'. Darginsko-russkij slovar': okolo 40 000 slov i frazeologicheskix vyrazhenij*. Makhachkala: Institut JaLI DNC RAN.

Kadibagamaev, A. A. 1998. *Sirxinskij dialect darginskogo jazyka (foneticheskaja sistema).* Makhachkala: DNTs RAN.

Khajdakov, S. M. 1985. *Darginskij i megebskij jazyki.* Moscow: Nauka.

Kibrik, A. E., S. V. Kodzasov. 1988. *Sopostavitel'noe izuchenije dagestanskix jazykov. Glagol*. Moscow: Moscow State University.

Kibrik, A. E., S. V. Kodzasov. 1990. *Sopostavitel'noe izuchenije dagestanskix jazykov. Im'a. Fonetika.* Moscow: Moscow State University.

Korjakov, Yuri. 2021. Darginskije jazyki i ix klassifikacija. In: T. A. Majsak, N. R. Sumbatova, Ya. G. Testelets (eds.). *Durqasi khazna. Sbornik statej k 60-letiju R. O. Mutalova:* 139–154. Moscow: Buki Vedi.

Korjakov, Yuri, Nina Sumbatova. 2007. Darginskije jazyki. In: *Bol'shaja rossijskaja enciklopedija. Vol. 8:* 328–329. Moscow: Bol'shaja rossijskaja enciklopedija.

Magomedov, A. Dzh., N. I. Saidov-Akkutta. 2017. *Kubachinsko-russkij slovar'.* Moscow: Nauka.

Magometov, A. A. 1963. *Kubachinskij jazyk (issledovanije i teksty).* Tbilisi: Izdatel'stvo AN Gruzinskoj SSR.

Magometov, A. A. 1976. Korrel'acija sistemy konsonantov v dialektax darginskogo jazyka. *Ezhegodnik iberijsko-kavkazskogo jazykoznanija III:* 219–225.

Magometov, A. A. 1982. *Megebskij dialekt darginskogo jazyka.* Tbilisi: Metsniereba.

Moroz, George. 2019. Phonology of Mehweb. In: Michael Daniel, Nina Dobrushina, Dmitry Ganenkov (eds.). *The Mehweb language: Essays on phonology, morphology and syntax*: 17–37. Berlin: Language Science Press.

Mudrak, O. A. 2016. *Darginskije osnovy*. *Vol. 1*. Moscow: Jazyki narodov mira.

Musaev A., V. Morozova, M. Daniel. 2020. *Mehweb Jena wordlist*. Ms. (Linguistic Convergence Laboratory, HSE University, Moscow). Available online at: https://linghub.ru/mehwebdict/.

Musaev, M.-S. M. 2002. *Darginskij jazyk*. Moscow: Academia.

Mutalov R. O. 2020. *Itsari Jena wordlist*. Ms.

Mutalov, R. O. 2002. *Glagol darginskogo jazyka*. Makhachkala: Dagestan State University.

Mutalov, R. O. 2021. Klassifikacija darginskix jazykov i dialektov. *Sociolingvistika 3(7)*: 8–25.

Nikolaev, Sergei L., Sergei A. Starostin. 1994. *A North-Caucasian Etymological Dictionary*. Moscow: Asterisk.

Shakhbanova, P. G. 2011. *Foneticheskije i morfologicheskije osobennosti karbachimaxinskogo govora darginskogo jazyka.* Dissertation for the degree of Candidate of Sciences. Makhachkala: Dagestan State Pedagogical University.

Sumbatova, N. R., Yu. A. Lander. 2014. *Darginskij govor selenija Tanty. Grammaticheskij ocherk. Voprosy sintaksisa.* Moscow: Jazyki russkoj kul'tury.

Sumbatova, Nina, Rasul Mutalov. 2003. *A Grammar of Icari Dargwa*. München: LINCOM.

Temirbulatova, S. M. 2004. *Xajdakskij dialekt darginskogo jazyka*. Makhachkala: DNTs RAN.

Temirbulatova, S. M. 2022. *Dialektologicheskij slovar' darginskogo jazyka.* Makhachkala: Alef.

Vagizieva, N. A. 2018. *Leksiko-frazeologicheskije osobennosti kadarskogo dialekta darginskogo jazyka*. Dissertation for the degree of Candidate of Sciences. Makhachkala: Dagestan State Pedagogical University.

Nina Sumbatova. Correspondences of post-uvular consonants in Dargwa languages

In this paper, I study the correspondences of the post-uvular (that is, pharyngeal and laryngeal) consonants in the Dargwa languages of the Nakh-Dagestanian language family. The

study is based on the data of nine lects of the Dargwa language group (Aqusha, Muira, Kadar, Mehweb, Tanti, Itsari, Chirag, Kaytag, Kubachi). I also suggest a reconstruction of a fragment of the Proto-Dargwa consonant system and describe the transitions from the reconstructed phonemes to their reflexes in modern languages. The reconstruction shows that the modern systems of postuvular consonants in Dargwa have been strongly influenced by the development of pharyngealization, an important phonological feature of the modern varieties of Dargwa.

Keywords: Nakh-Dagestanian languages; Dargwa languages; post-uvular consonants; laryngeal consonants; pharyngeal consonants; pharyngealization; reconstruction of consonant systems.

Mikhail Zhivlov
Russian State University for the Humanities / HSE University, Moscow; zhivlov@yandex.ru

Hokan I: A review of comparative studies[1]

The article reviews the history of Hokan studies, from the formulation of the Hokan hypothesis by Dixon & Kroeber in 1912–1913 to the present day. Despite more than a hundred years of research, there is yet no consensus on the validity of the Hokan hypothesis. The article argues that one of the reasons for this is the fact that many attempts to compare Hokan languages used a non-standard methodology, in which the study stops at listing the observed sound correspondences, instead of the classical comparative method, which requires a search for complementary distributions and a comprehensive reconstruction of protophonology.

Keywords: comparative method; Hokan languages; Native American languages; genealogical classification of languages.

0. Introduction

The Hokan hypothesis, combining a number of language isolates and small families in California into a single Hokan stock, was proposed more than a hundred years ago. Despite the long history of the proposal, there is still no consensus on its validity. Although putative Hokan languages remain severely underdocumented, now we have much more data on them than was available to researchers in the 20th century. Time has come for a reassessment of the hypothesis in the light of new data. Such a reassessment must necessarily begin with a survey of what has already been done in the field of comparative Hokan studies. In the words of Lyle Campbell (1997: 290), "[a] thorough understanding of the Hokan hypothesis requires a knowledge of its history".

The present paper is organized as follows: Section 1 is dedicated to the first period of Hokan studies, from the original proposal by Dixon & Kroeber to the work of Sapir. Section 2 surveys an intermediate period from the 1940s to the 1950s, when little has been done on the Hokan hypothesis. Section 3 offers a survey of the period from the 1950s to the 1980s, focusing on the work of Mary R. Haas and her students. Section 4 is dedicated to attempts of general synthesis by such scholars as K.-H. Gursky, D. Leshchiner, T. Kaufman and others in the 1980s. Section 5 surveys the work on Hokan done from the 1990s to the present day. Section 6 discusses general methodological problems relevant to Hokan comparison.

1. The birth of the Hokan hypothesis

Although links between, e.g., Yuman, Seri and Tequistlatecan (Brinton 1891: 109–112), Shasta and Palahinihan (Dixon 1905), Chimariko, Shasta and Palahinihan (Dixon 1910: 335–339) had been proposed earlier, the Hokan hypothesis *sensu stricto* dates from 1912–13, when Dixon &

[1] The study was carried out with the assistance of the "Laboratory Of Unnecessary Things" of the Independent University of Moscow. I thank Fernando O. de Carvalho, Albert Davletshin, Yoram Meroz, George Starostin and an anonymous reviewer for their valuable advice. All remaining errors are mine alone.

Kroeber (1912, 1913b) announced their findings on the classification of Californian languages in a brief note, published first in "American Anthropologist" and then in "Science". Their proposed new families were Penutian, consisting of Maidu, Wintun, Miwok, Costanoan and Yokuts; Hokan, consisting of Shasta (including Shastan proper, Achumawi and Atsugewi), Chimariko, Pomo, probably Karuk and possibly Yana; and Ritwan, consisting of Yurok and Wiyot. No evidence was presented in favor of these hypotheses. Then followed a somewhat more substantial paper (Dixon & Kroeber 1913a), where Esselen and Yuman were added to Hokan and a new family was proposed — Iskoman, consisting of Chumash and Salinan. The evidence for Hokan presented in this paper consisted of structural similarities and five sets of lexical cognates — words for 'tongue', 'eye', 'water', 'stone' and 'sleep', each represented in no less than five out of seven postulated branches of Hokan. The authors also tentatively suggested that Seri may belong to Hokan and that Hokan and Iskoman may be related.

Kroeber (1915) connects Seri and Oaxaca Chontal with Yuman and Hokan in general. He lists 35 comparative sets. Kroeber also discusses sound correspondences, but from a decidedly pre-Neogrammarian point of view, e.g.: "This Chontal-Seri correspondence f : x is corroborated by Seri-Mohave f : h in number 15—the fricative character is retained, but the point of articulation changed" (Kroeber 1915: 282).

The next researcher to take up work on Hokan was Edward Sapir. Sapir (1917) presents 192 sets of comparisons between Yana and other Hokan languages; some additional Hokan cognate sets without Yana are also included. Sapir's comparisons are semantically quite straightforward; in most cases the meanings of compared forms are either identical or very close. Some of Sapir's comparisons must be discarded in the light of new data or reconstruction within accepted small families. Consider his etymology of the word for 'man':

60. Yana *ʻisi* "man, male, husband"
Shas. *ic* "man"; New River Shasta *gè-ʼic*
Chim. *itci*, *itri* "man"
S. Pomo *atcai* "man"; N. Pomo *tca* "person"; C. Pomo *tcatc*; S. W. Pomo *atca*
Chum. *-isūyix* "husband"
Chon. *acans* "person" (Sapir 1917: 9).

Checking the Yana, Shasta and Chimariko forms across more modern sources, we can see that these look just as comparable as the forms given by Sapir: Central and Northern Yana *hisi* 'man', Yahi *hihsi* 'man; husband, husband's brother' (Sapir & Swadesh 1960: 98), Shasta *ʔís* 'person; Indian' (Bright & Olmsted 1959: 37), Chimariko *ʔiṭi* 'husband, man' (Conathan 2002: 31). Moreover, we can add Achumawi *ís* 'person; Indian' (Nevin 2020) to the comparison. The Pomoan forms, however, go back to a Proto-Pomo form hardly compatible with the Yana, Shasta, Achumawi and Chimariko forms above: Kashaya *ʔacaˇc* 'person, Indian, man', Southern Pomo *ʔač·ay*, Northern Pomo *čáʔ* 'man, person', Central Pomo *cá·ċ* 'person', Northeastern Pomo *tá·t-ka·*, Southeastern Pomo *ca-wi*, Eastern Pomo *ká·k^{h}* 'man, Indian' < Proto-Pomo *ʔaká·kʔ 'man' (McLendon 1973: 81). The Chumash (more precisely, Ineseño) word *-isūyix* 'husband' was taken from Kroeber (1904: 42). "Samala-English Dictionary", based on data collected by J. P. Harrington, gives this word as *is^{h}ɨ'y* 'husband', from *is-* 'one's own' + *hɨ'y* 'male', a shortened form of *ɨhɨ'y* 'man, male' (The Santa Ynez Band of Chumash Indians 2007: 148). Thus, the comparison with Chumash must be rejected. Finally, Lowland Chontal *sans* 'person, human being, living creature' has an irregular plural *-san^{y}uʔ* (O'Connor 2013: 226), pointing to a root *san-*. The comparison with the forms above does not look promising.

Nevertheless, a surprisingly large number of Sapir's comparisons in (Sapir 1917) stand the test of time. It is interesting to look at some of the recurrent sound correspondences noted by

Sapir. Yana *y* corresponds to Chimariko *s* or *c* [š] in Yana *ʼīya* 'trail', Chimariko *hissa* 'trail'; Yana *wêyu* 'horn', Chimariko *h-owec* 'antlers, horn'; Yana *-ya* 'female', Chimariko *-sa* 'female' (Sapir 1917: 9, 16, 23). Yana *pʼu* corresponds to Chimariko *xu* in Yana *pʼô-* 'to blow', *pʼu-sā-* 'to smoke', Chimariko *-xu-*, *-xuc-* 'to blow'; Yana *pʼū-* 'to swim', Chimariko *-xū-* 'to swim'; Yana *pʼuiʼ-* 'to be fat', Chimariko *-xu-* 'fat' (adj.); Yana *pʼun-* 'to paint', Chimariko *-xol-* in *-po-xolxol* 'to paint' (Sapir 1917: 13). Despite the obviously inadequate quality of data available at that time, Sapir (1917) remains one of the most persuasive collections of comparanda supporting the Hokan hypothesis.

The final statement of Dixon & Kroeber's views on Hokan was their short monograph "Linguistic families of California" (1919). Here, Hokan includes Karuk, Chimariko, Shastan (including Achumawi and Atsugewi), Pomoan, Yana, Washo, Esselen, Salinan, Chumash, Yuman, Seri and Tequistlatecan (Oaxaca Chontal). No new evidence for Hokan as such is adduced in the 1919 publication. The authors explain this in the following way: "The evidence submitted by the writers as to the unity of these languages is admitted by them to be but slender. Yet they feel themselves absolved from the obligation of presenting further formal proofs through the publication by Dr. Sapir of a recent essay devoted nominally to the determination of the position of Yana in the Hokan stock, but in effect rendering as full a proof of the actuality of the family as could be demanded" (Dixon & Kroeber 1919: 103). However, since the inclusion of Washo in Hokan was a new idea, the authors give two lists of Washo-Hokan parallels, one compiled by themselves and another by Sapir. What is more interesting, the authors provide a "Historical Introduction", where they explain how they arrived at their classification. In the wake of Dixon & Kroeber's (1903) typological survey of Californian languages, the authors became interested in lexical similarities between these languages. However, from the start they interpreted these similarities as due to contact between unrelated languages[2]: "As evidences of similarities between this and that language accumulated, they were indeed noted, but were consistently interpreted as instances of one unrelated language borrowing either material or machinery from another" (Dixon & Kroeber 1919: 49). A systematic comparison was attempted: "About two hundred and twenty-five English words were selected on which material was most likely to be accessible in reasonably accurate and comparable form, and the known native equivalents in sixty-seven dialects of the twenty-one stocks were entered in columns. Comparisons were then instituted to determine all inter-stock similarities that seemed too close or too numerous to be ascribed to coincidence. The purpose of the study was threefold: first, to ascertain the nature and degree of borrowing between unrelated languages; second, to trace through these borrowings any former contacts or movements of language groups not now in contact; third, in the event of any relationship existing between languages then considered unrelated, to determine this fact." (Dixon & Kroeber 1919: 49). The resulting "lexicostatistical matrix" made no sense in terms of language contact: "Families some distance apart on the map often had more stems in common than those in juxtaposition; if the remote group was regarded as once in contact with the one with which it shared most words, it must have been in contact also with others with which it shared but few words. ... Finally, in a mood rather of baffled impotence, an interpretation of the cases of most abundant resemblance as due to genetic relationship was applied. At once difficulties yielded, and arrangement emerged from the chaos" (Dixon & Kroeber 1919: 50). The number of resemblances in the "lexicostatistical matrix" between branches of what would be called Hokan ranged from 4 (Karok-Yana) to 21 (Shasta-Pomo). The authors themselves admit that their comparisons

[2] Fernando O. de Carvalho (p.c.) points out that these interpretations by Dixon & Kroeber make sense within the general Boasian outlook of North American linguistics at the time (see Campbell 1997: 62–66, 72).

"might or might not be considered as including radical words due to a common origin: they certainly included words not due to such common origin but derived by loan. ... Further, the comparisons used being avowedly superficial, that is, not based on analysis, a certain number of false coincidences were bound to have crept in" (Dixon & Kroeber 1919: 52). Still, it is clear that there was more to Dixon & Kroeber's Hokan data than just five lexical sets published in Dixon & Kroeber (1913a). This is important in view of the following misleading statement by Lyle Campbell: "Hokan had the shakiest of origins. ... This hypothesis was based on inspectional resemblances involving only five words in these languages: 'tongue', 'eye', 'stone', 'water', and 'sleep'" (Campbell 1997: 290). Of course, the unfortunate fact that Dixon and Kroeber did not publish their evidence, thus setting a precedent for later researchers in the field, does not mean that their evidence was confined to five words.

Sapir (1920a) demonstrates an interesting paradigmatic morphological similarity between Chimariko and Salinan: verbal subjective pronominal prefixes have the following form in both languages (Table 1; see also Zhivlov 2018).

Table 1. Chimariko and Salinan pronominal prefixes, after Sapir (1920a)

		Chimariko	Salinan
Sing.	1	y-, i-	e-
	2	m-	m-
	3	h-	-
Plur.	1	a-; ya-	a-
	2	q-	k- (subject of 2nd per. plur. imperative)
	3	h-	-

Sapir's (1920c) review of J. Alden Mason's monograph "The Language of the Salinan Indians" contains, apart from many valuable remarks on Salinan synchrony, some diachronic observations. The most interesting is the comparison of plural infixes in Salinan and Yana, worth quoting in full: "One of the most interesting and irregular features of Salinan is the formation of the plural of nouns and of the plural and iterative of verbs. No less than a dozen distinct types and a large number of irregular formations are discussed and illustrated by Mason, the great majority of them involving a suffixed or infixed *-t-*, *-n-*, or *-l-*. Significantly analogous plurals, often of great irregularity though of less frequency, are found in Yana ; e. g. such Salinan plurals as *ṭ-eṭeyitinai* ARROWS (sing. *ṭ-eṭeyini*') and *anetem* SEVERAL REMAIN (sing. *anem*) offer more than a cursory parallel to such Yana forms as *mut'djauṭ'i-wi* CHIEFS (sing. *mu·djaup'a·*), *k'uru·-wi* SHAMANS (*-r-* < *-d-*; sing. *k'u·wi*), *sa·dimsi-* SEVERAL SLEEP (sing. *samsi-*, *sams-*). The Salinan type with infixed *-h-*, *-x-* (e. g. *mehen-* HANDS, sing. *men-*; *kaxau* SEVERAL SLEEP, sing. *kau*) may be analogous to such Yana forms as *dja·li-* SEVERAL LAUGH (from **djahali-* ?), sing. *djal-*" (Sapir 1920c: 306). Similar infixed plurals are attested in Chontal, Seri and Pomoan (Langdon 1990; Zhivlov 2018).

Sapir's second large paper on Hokan (Sapir 1920b), actually written before "The position of Yana...", tries to demonstrate the connection between Hokan and Coahuiltecan — a grouping of languages in Texas and northeastern Mexico proposed by Swanton (1915). Swanton's Coahuiltecan includes Tonkawa, Comecrudo, Cotoname, Coahuilteco, Karankawa and Atakapa. Sapir does not question the coherence of Coahuiltecan as a group, but tries to connect it as a whole to Hokan. Almost half of the comparisons in this article are between just one of the

Hokan branches and Coahuiltecan and thus are irrelevant for the Hokan hypothesis in the narrow sense. If Coahuiltecan (or some parts of it) was related to Hokan, we would expect Hokan-Coahuiltecan parallels to have a better distribution within Hokan proper.

Sapir (1921) gives Salinan additions to comparisons from (Sapir 1920b), as well as 28 new Salinan-Hokan comparisons. The paper also briefly discusses the position of Washo within Hokan.

Sapir (1925) presents the evidence for relating the small Mesoamerican family of Tlapanec-Subtiaba to Hokan-Coahuiltecan. We know now that Tlapanec-Subtiaba forms part of the Otomanguean family (Campbell 2017a, 2017b). Even if Hokan is related to Otomanguean as a whole (Kaufman 2016), it is unlikely that many of Sapir's comparisons will hold. The importance of this paper for Hokan studies lies not in comparisons with Subtiaba, but in the fact that it is here that Sapir laid out his views on Hokan morphology, specifically, nominal, adjectival and verbal prefixes. At least one of Sapir's findings in this area was confirmed by later research. Sapir (1925: 506–512) reconstructed Proto-Hokan stative prefix *m- on the basis of the rather inadequate data he had at the time. Traces of such a prefix were later found in Eastern Pomo (McLendon 1975: 48), Yuman (Miller 2001: 60) and Seri (Marlett 1981: 70–71).

While Dixon & Kroeber did not offer much evidence in support of Hokan, Sapir's work represented a huge step forward in collecting and organizing such evidence. In his works on Hokan, Sapir used the method now known as "mass comparison". He understood very well the limitations of this type of work: "Some of these comparisons are doubtful at best and a number of them will, on maturer knowledge, have to be discarded. A certain amount of groping in the dark cannot well be avoided in the pioneer stage of such an attempt as this" (Sapir 1920b: 289). However, a careful application of the comparative method, as in Sapir's own work on Uto-Aztecan (Sapir 1913, 1919), was impossible here because the only phonologically adequate documented Hokan language at the time was Yana, described by Sapir himself. Still, Sapir does not limit himself to comparisons of lexical lookalikes; he notes recurrent sound correspondences (1917: 6, 7, 8, 9, 10, 13, 16, 21, 23, 27, 1920b: 284, 286, 287), provides preliminary Hokan reconstructions[3] (1917: 33); discusses Hokan root structure (1925: 493-495), derivation (1925), and bits of paradigmatic morphology (1920a). Surprisingly, there were no published attempts to systematically reconsider Sapir's evidence for Hokan and Hokan-Coahuiltecan in the light of new data on the languages involved.

2. A temporary standstill

The following decades saw no new work on Hokan in the proper sense. The only papers published in that period were attempts to add new languages to Hokan.

Rivet's (1942) article is an unsuccessful attempt to show Hokan affiliation of a fragmentarily attested Colombian isolate, Yurumanguí. "The result of the dearth of Yurumangui data combined with the fecundity of Rivet's etymological imagination is that most of the morphological analyses he proposed rest on hypothetical affiliations with Hokan, and so remain undemonstrated" (Poser 1992: 216). If anything, Rivet's work shows that with the loose application of the "mass comparison" approach it is possible to "demonstrate" Hokan affiliation of just about any language.

[3] "Sapir was aware that the available data, except for Yana, was phonologically underspecified if not totally inadequate. I reckon that he intended his pHok starred forms as an approximation to what a full reconstruction would be when phonologically adequate data was available" (Kaufman 1989: 80).

Harrington's (1943) article is broadly similar: it attempts to show a Hokan affiliation for a South American language, this time Quechua. The results are similarly unconvincing. Unlike Rivet, Harrington tries to set up sound correspondences, but his approach can only be characterized as pre-Neogrammarian. Cf. the following sound correspondence with examples:

> *', q, y—'*
> Q iwa, plant, tree; Subt i·ci, tree; Q qalyu, for *'alyu, tongue; Chim -pen, tongue; Q yawar, blood; Sal a·ˊkat, blood (Harrington 1943: 337).

Greenberg & Swadesh (1953) propose that Jicaque (Tol), spoken in Honduras, is a Hokan language. The actual scope of the comparison is wider, since it includes not just "Hokan-Coahuiltecan", but also "Gulf" languages and occasionally other languages from Sapir's Hokan-Siouan superstock. Some comparisons look plausible, e.g,

> *house*: Jicaque wa; Yuma ʔava; Esselen iwa-no; Chimariko ʔawa; Yana wa-wi; Comecrudo wamak; Subtiaba g^{w}a (Greenberg & Swadesh 1953: 219).

More typical, however, are comparisons like the following:

> *mouth*: Jicaque lam, lala; Tonkawa kala; Comecrudo xal; Atakapa kat; Yuma axaʔa; Pomo xacita; Chumash ïk; Chitimacha ša; Tunica šohu; Karankawa emi-akoy. Earliest form perhaps xa-la, with development in two directions as lala and axa, ixa (Greenberg & Swadesh 1953: 219).

The "reconstruction" *xa-la, which is not based on regular sound correspondences, allows the authors to compare such dissimilar forms as Jicaque *lala* and Chumash *ïk*.

3. Mary Haas and the Survey of California Indian Languages

The next period of Hokan studies, which lasted from the 1950s to the 1980s, was dominated by the work of Mary R. Haas and her associates at the Survey of California Indian Languages, established in 1953 in the University of California at Berkeley under the leadership of Haas and Murray B. Emeneau. The main purpose of the Survey was to organize fieldwork on endangered languages of California. The results of this fieldwork gave new impetus to comparative studies on Hokan and Penutian.

Haas (1954) attempts to reconstruct the Proto-Hokan-Coahuiltecan word for 'water'. While Haas discusses sound correspondences for the word in question, she does not provide any evidence that these correspondences recur in at least one other word. Thus, the methodology used in this paper only superficially resembles the traditional comparative method. Haas' reconstructions are actually "pre-reconstructions" in the sense of Peiros (1997).

Bright (1954) compares data from Karuk, Chimariko, Shasta, Achumawi and Atsugewi with the aim of establishing regular sound correspondences. He correctly notes that irregularities in sonorant correspondences may be caused by morphophonemic alternations of nasal and non-nasal sonorants, of a type attested in Karuk.

Bright (1956) attempts to carry out experimental glottochronological counts for eight southern Hokan-Coahuiltecan ("Hokaltecan") languages: Seri, Oaxaca Chontal, Subtiaba-Tlapanec, Jicaque, Comecrudo, Yuma, Salinan, and Tonkawa. The author uses strict objective criteria of what counts as "similar" forms. In addition, dissimilar forms may be counted as matches based on recurrent sound correspondences. Still, it is intuitively clear that the criteria used by Bright result in many spurious matches, such as Seri *ítak* ~ Chontal *ekał* 'bone', counted as a match on the basis of the "correspondence" Seri *k* ~ Chontal *k*, or Seri *ʔánoł* ~ Salinan *i̱népuk*^{h} 'arm', with the "correspondence" Seri *n* ~ Salinan *n*. The fact that such matches are not distinguished from obvious similarities like Seri *ʔaX* ~ Chontal *axáʔ* 'water' or Seri *ʔapł* ~

Chontal *ípaL* ‘tongue’ deprives the resulting counts of any value. See also Marlett’s (2008) criticism of Bright’s Seri-Salinan comparisons.

In a series of three short papers, Olmsted (1956, 1957, 1959) tries to show that Achumawi and Atsugewi do not form a subgroup with Shasta within what he calls Northern Hokan. Olmsted uses his own field materials on Achumawi and Atsugewi. Unfortunately, the phonological accuracy of his field records leaves much to be desired (Nevin 1998: 15–18), which makes the work in question less useful than it might have been.

Jacobsen (1958) compares Washo — a language he himself worked on — with Karuk, for which a good description (Bright 1957) was available. Some of his comparisons are striking: Washo *í·biʔ* ~ Karuk *ʔípih* ‘bone’, Washo *-á·gal* ‘into or in the mouth’ ~ Karuk *-kara* ‘into one’s mouth’, Washo *-íwl* ‘to roll’ ~ K *-ívruh* in *ikrívruh* (sg.) ‘to roll (as a ball)’, *ʔinívruh-* (du.-pl.) ‘to roll (as balls)’, Washo *-í·šib* ‘to be straight, correct, right’ ~ Karuk *ʔíšip* ‘to extend, be in a line’, Washo *-ámad* ‘through a tubular space’ ~ Karuk *-vara* ‘in through a tubular space’, Washo *-íliʔ* ‘up, upwards’ ~ Karuk *-rih* ‘up’. Since these comparisons find no counterparts in other Hokan languages, and even Jacobsen himself does not postulate a specific relationship of Washo and Karuk, a contact explanation seems preferable. It is interesting that many of these parallels are locative-directional suffixes/postpounds — a category whose spread in the Northern California language area must have involved language contact (Mithun 2007). Jacobsen lists sound correspondences between Washo and Karuk, but his approach is rather mechanistic. The comparison of Washo *émle* ‘heart’ and Karuk *iθva·y* ‘breast, chest, heart’, according to Jacobsen, illustrates the following correspondences: W *e* ~ K *i*, W *e* ~ K *a·*, W *m* ~ K *v*, W *l* ~ K zero, W zero ~ K *y*, W zero ~ K *θ*. Each of these correspondences, save the last, is supported by at least one additional example. While no segment is left unaccounted for, one may doubt the realism of the whole picture.

One of the parts of Morris Swadesh’s introduction to the Yana Dictionary (Sapir & Swadesh 1960) is dedicated to “Yana-Hokan Notes”. Swadesh compares the basic vocabulary of Yana and Chimariko and finds 22 “likely cognates” among 113 semantically equivalent word pairs. Swadesh correctly emphasizes the necessity of recurrent sound correspondences: “The principal fact in this connection is that the compared elements show consistent phonetic relationships, as strikingly illustrated by the correspondence of Chimariko /x/ to Yana /p/ before rounded vowel in four cognate sets, including "swim" and "blow," listed above, and two others ("fat," "smear") given by Sapir” (Sapir & Swadesh 1960: 17). He also notes the importance of lexical sets represented in many daughter languages: “Even more striking is the multilanguage agreement; such words as "water," "two," "drink," "eat," "tongue," "thou" show cognates in six to ten different languages or language families” (Sapir & Swadesh 1960: 18).

Haas (1963) starts from an observation that certain cases of long vowels in Shasta go back to earlier sequences *VmV with intervocalic *m. This *m completely disappears in Shasta, but is reflected as /w/ in Okwanuchu. This finding allows Haas to propose a number of Proto-Hokan reconstructions with such sequences: PH *išamaruƙa ~ *išamaƙaru ‘ear’, PH *č-imapasi ~ *imačipasi ‘liver’ (cf. Sapir’s *ipasi), *imaraƙʷi ~ *imaƙʷari ‘navel’, *imapaki ~ *amipaki ‘neck (nape)’, *yamari ~ *imari ~ *irama ‘nose’, *amiraẋʷa ~ *amiẋʷara ~ *s/kiraẋʷa ~ *s/kiẋʷara ‘nails (claws)’, *iš/čarima ~ *iš/čamira ‘arm’, *ipari ~ *ipawari ~ *ipariwa ‘tongue’, *as/cima ~ *is/cama ‘sleep’. Proto-Hokan, as it emerges from these reconstructions, is a language with rather long words, from three to five syllables, with no consonant clusters and with frequent metathesis[4].

[4] Fernando O. de Carvalho (p.c.) points out that Haas’ willingness to go on with such “protoform stuffing” and to postulate rampant metathesis likely stems from her prior work on Muskogean (see, e.g., Haas 1969: 40–42). In the Muskogean domain this kind of strategy, yielding three syllable proto-forms for disyllabic, widely

Now, one can hardly object to the comparison of such forms as, e.g., Antoniano Salinan *-épal* 'tongue', Southeastern Pomo *bal* 'tongue', Yuma *i·pálʸ* 'tongue', for which Haas reconstructs PH *ipari — whether or not supported by recurrent correspondences, we at least have here a clear case of phonetic similarity. But the same cannot be said about Barbareño Chumash *elʔew* 'tongue' and Washo *álŋ* 'to lick', which are given under PH variant reconstruction *ipariwa. The problem is that the correspondence of Chumash and Washo zero to a labial stop in other languages is not confirmed by additional examples, and thus, is not recurrent. Comparing dissimilar forms not connected by recurrent correspondences may allow one to call cognates any two forms at will. Another problem is that the length of the reconstructed forms allows comparisons between words which supposedly continue different parts of a longer form, so that "cognates" may share no cognate material at all. Thus, both Karuk *ti·v* 'ear' and Barbareño Chumash *tuʔ* 'ear' are said to go back to Proto-Hokan *išamaruƙa 'ear', but the Karuk word continues the *išam- part, whereas the Chumash word reflects *-ruƙ- (Haas 1963: 46). In the same way, Yana *ima* 'liver' and Chimariko *-ši* 'liver' are given under Proto-Hokan *č-imapasi 'liver', but the Yana form reflects *-ima-, while the Chimariko form reflects *-si (Haas 1963: 47). The device of metathesis allows Haas to compare words that are unrelated within one uncontroversial family: Cocopa *iya·kal* 'navel' (supposedly from PY *imakʷáli < PH *imaƙʷari) is compared to Yuma *malʸpú·* 'id.' (< PY *imalikʷí < PH *imaraƙʷi), with the following comment: "PYu. *kʷ > p in Maricopa, Yuma, and Diegueno; > k in Kahwan and Cocopa." (Haas 1963: 48). Actually, Proto-Yuman *kʷ is retained as such in all Yuman languages (Langdon & Munro 1980: 126). Lyle Campbell (1997: 294) was certainly right to conclude that "[i]n the absence of a more fully developed proposal for the historical phonological developments, one might suspect, for example in the case of 'ear', that Karuk *t<i·v* (< means that the sound is assumed to have undergone a change of assimilation), Chimariko *-sam*, and Chumash *tuʔ* may not really be cognate forms in genetically related languages and that they may not derive from Haas's proposed proto form **išamaruk'a/*išamak'aru* (1963b:46); similarly, Achomawi *owè·*> 'liver' is a stretch from the assumed Proto-Hokan **č-imapasi/*imačipasi* (1963b:47); and Chumash *top'o*, Achomawi *alu*, and Washo *í>ʔb* 'navel' are a far leap from each other and from the proposed Proto-Hokan **imarak'ʷi/*imak'ʷari*".

Haas (1964) compares Yana and Karuk — two Hokan languages for which published dictionaries were available at the time. Comparisons with other Hokan languages are also adduced, but sound correspondences are listed only for Yana and Karuk. In what would become a model for further binary comparisons, correspondences, including non-recurrent ones, are simply listed, and no attempt is made to show that they form a coherent system or to assign protophonemes to each correspondence. Haas also points out the problem of "intersection" between Hokan and other Californian families: Penutian, Yukian and Ritwan. She compares her Proto-Hokan reconstructions to words with the same meaning in non-Hokan languages of California. Thus, Proto-Hokan *imaraƙʷi ~ *imaƙʷari 'navel' is compared to such "Penutian" forms as Chukchansi Yokuts *čˇut-*, Central Sierra Miwok *pó·ṭi-*, Rumsen *lop-*, Wintu *naq* and Maidu *betéke*. What these comparisons demonstrate is not the extent of lexical diffusion in California, but rather the ability of Haas' method to integrate almost any word from any language under polysyllabic reconstructions unsupported by recurrent sound correspondences.

McLendon (1964) presents 149 lexical comparisons between Eastern Pomo and Yana. Data from other Hokan languages are included wherever possible. While comparisons between

diverging reflexes (e.g. Choctaw *fani* 'squirrel' : Creek *íłu* 'id.' < Proto-Muskogean *ixʷaNi/u), with different languages retaining different "parts" of the etymon, was apparently successful, and some Muskogean languages like Creek do in fact have very frequent cases of metathesis.

Eastern Pomo and Yana mostly look reasonable, it is not always clear whether the compared forms from other languages are intended as cognates, and if yes, what exactly is being compared. Consider the set for 'belly':

> Pe *wi·ní* 'pregnant' : Yn, Yy *wil(la)*. ... Cf. Ps *wi·ní* 'pregnant'; K *višva·n* 'belly'; S *ʔíčnit* 'belly'; Yuy *i·tó* 'belly' (McLendon 1964: 129).

Southern Pomo *wi·ní* 'pregnant' is straightforwardly related to the Eastern Pomo form. Karuk *višva·n* 'belly' may originally be a compound whose second part may be related to Pomoan and Yana forms. Shasta *ʔíčnit* 'belly' is perhaps adduced because the *-ni-* part can be compared with *-ní* in Pomoan words, but nothing in the Shasta form corresponds to *wi-* of Pomoan and Yana words. It is a complete mystery to me why Yuma *i·tó* 'belly' is given in this set. This case shows how important it is to have a working model of sound correspondences for the family as a whole — without it, there is no clear boundary between plausible and implausible comparisons.

Silver (1964) compares Shasta and Karuk, although forms from other Hokan languages are occasionally adduced. Some of the comparisons may represent genuine Hokan cognates: 'drink' S *-ic·-*, K *ʔíš*; 'rock' S *ʔíc·aʔ*, K *ʔasa*; 'sleep' S *-icmas-*, K *ʔasiṽ*; 'water' S *ʔác·a*, K *ʔá·s*. Some other comparisons, without parallels in other Hokan languages, are likely to be areal loans: 'gopher' S *xáras·aʔ*, K *axra·s*; 'sack' S *pú·ʔas*, K *pu·viš* (cf. also Yurok *puwi·š* 'sack'); 'dog salmon' S *ʔičmú·n·a*, K *ačvu·n*.

Turner (1967) tries to show that Oaxaca Chontal languages are not related to Seri. He lists eight "cognates" between Highland Chontal and Seri on the 100-word Swadesh list plus twelve additional "cognates" from a bigger 500-word list and comes to the following conclusion: "Evidently all of the above cognates are either accidental or the result of borrowing. Because of the lack of actual cognates, it seems unnecessary to even attempt to describe systematic phonemic correspondences". Turner's lack of familiarity with historical linguistics is evident from the following passage: "It would be difficult to explain how two related languages could have different grammatical categories, and such is the case with Seri and Chontal". A list of supposedly insurmountable differences between the two languages includes the following: "Seri identifies blue and green by different words whereas Chontal does not have separate words for these two colors. ... Seri has separate words for the urine of men and women as well as separate words for the act of urination by men or women; Chontal does not have this feature". Bright (1970) in his rebuttal remarks "(1) that it is impossible to prove the LACK of relationship between ANY two languages, and (2) that the positive evidence for the relationship of Seri and Tequistlatec is in fact greater than admitted by Turner".

Langdon (1974) is a detailed survey of the history of Hokan studies up to 1970. Her final evaluation is quite optimistic: "We may conclude that Hokan-Coahuiltecan as conceived by SAPIR has withstood the test of time fairly well. Seriously disputed as to membership are only Tonkawa and Karankawa, and a new member, Jicaque, is likely. Less clear is the validity of SAPIR's subgroupings, which, however, have no serious competitors to replace them, except the very cautious view that no two of the families grouped under the label Hokan-Coahuiltecan can be shown as yet to be closer to each other than to the others. ... While a full demonstration of the validity of the Hokan-Coahuiltecan hypothesis is not yet a reality, there is a growing sense of excitement as convergent results are reported" (Langdon 1974: 86).

Silver (1974) discusses suffixes/postpounds in names of plants, trees and bushes. She draws attention to parallels with non-Hokan Californian languages and suggests, e.g., that the Yurok suffix of plant-tree-bush terms *-epˀ* may be borrowed from the Karuk morpheme *(ʔ)i·p* with the same function.

Turner (1976) shows that noun plural formation follows different rules in Chontal and Seri; he uses this fact as a further support for his argument that the languages are unrelated.

Silver (1976) raises the problem of "morphemization", i.e. fossilization of morpheme sequences, "which often has the result of obscuring diachronically significant underlying phonological representations of morpheme strings". According to Silver, aspirated and glottalized consonants in Hokan languages can go back to sequences like *CVx, *CVh and *CVʔC.

James M. Crawford (1976) compares Chimariko and Yuman, presenting 134 lexical sets. While many comparisons are quite good, e.g., Chimariko *'awa*, Cocopa *wá* 'house', Chimariko *'awu*, Cocopa *wí* 'mountain' or Chimariko *hoxu*, Cocopa *i·xú* 'nose', in some other cases long dissimilar forms are compared on the basis of one matching segment. E.g., the set for 'smooth' compares Chimariko *luyu'it*, *luyu'-* with Cocopa *xlqáy*, Yawapai *halowári* (these two Yuman forms evidently have different roots), with the following comment: "[o]nly laterals are compared". There are other cases where the Yuman forms within one set are not related to each other, as in the set 'bird', where Cocopa *šá* and Mohave *ʔaciyér* are unrelated. The author lists sound correspondences between Chimariko and Proto-Yuman. Out of 45 consonant correspondences, 14 are encountered only in one example (i.e., they are non-recurrent), and nine more are encountered in two examples. This is one more case of a binary comparison where analysis stops at the stage of listing the correspondences. No attempt is made to distinguish between chance coincidences and actual correspondences, or to integrate the correspondences into a coherent system.

Judith G. Crawford (1976) compares Seri and Yuman languages. The paper contains 227 comparative sets; forms from other Hokan branches are also adduced. Sound correspondences are listed, but not analyzed; no reconstruction is attempted. In some cases, long forms are compared based on one "corresponding" segment only; e.g., Seri *kkap* 'to fly' is compared to Mohave *himán* 'fly, get up', Havasupai *mánika* 'fall' with the comment "Only S *p* and Yuman *m* are being compared". Some other cases involve loose semantics, as in the comparison of Seri *ʔáapXa* 'cottontail rabbit' with Cocopa *prxá·w* 'fox'. See also criticism of this work by Marlett (2007).

Waterhouse (1976) gives a detailed reconstruction of Proto-Chontal (Tequistlatecan), based on a systematic comparison of Highland and Lowland Chontal. She also provides several lists of comparisons with other languages: "Comparison with Yuman languages", "Comparison with Seri", "Comparisons with other Hokan languages", "Comparisons with Karankawa and Cariban" and "Comparison with Tlappanec (and Otomanguean?)". Many of her comparisons with Yuman languages look quite promising, e.g.:

Paipai *yipúk* ~ Lowland Chontal *-epúh* 'nape'; Cocopa *ipáɬ* ~ Lowland Chontal, Highland Chontal *-paL* 'tongue'; Tipai *tú* 'stomach' ~ Lowland Chontal, Highland Chontal *-etú* 'navel'; Cocopa *ix^{w}áṭ* ~ Lowland Chontal, Highland Chontal *-Wac'* 'blood'; Yuma *ʔamáṭ* ~ Lowland Chontal, Highland Chontal *amác'* 'earth'; Diegueño *ʔak^{w}í* 'cloud' ~ Lowland Chontal, Highland Chontal *akwí* 'rain'; Kiliwa *amʔál*, Diegueño *ʔamal* ~ Lowland Chontal, Highland Chontal *ʔmaL* 'century plant'; Mohave *naʔáy* 'father (w.s.)' ~ Highland Chontal *aʔáyiʔ* 'father'; Yuma *ʔaxán* 'good' ~ Lowland Chontal *hanáʔ* 'all right'; Yavapai *ʔaháʔ* ~ Lowland Chontal, Highland Chontal *aháʔ* 'water'; Diegueño *ʔíɬ* 'worm' ~ Lowland Chontal, Highland Chontal *ʔiL* 'flea'; Cocopa *išáɬ*, Paipai *ṣál* 'hand, arm' ~ Lowland Chontal *asáL*, Highland Chontal *išáL* 'wing'; Paipai *ṣmá* ~ Lowland Chontal *sma-*, Highland Chontal *šmay-* 'to sleep'; Paipai *x^{w}íx*, Mohave *hah^{w}é·k* ~ Lowland Chontal *Weloo-*, Highland Chontal *We-* 'to smell'; Paipai *oʔo x^{w}áy* ~ Lowland Chontal *We-*, Highland Chontal *Way-* 'to smoke'; Yuma *ʔu·n^{y}é* ~ Lowland Chontal, Highland Chontal *áne* 'road'; Yavapai *k^{y}úli* 'long' ~ Lowland Chontal *kul^{y}íʔ*, Highland Chontal *akulíʔ* 'far'.

Lyle Campbell (1976) supports the Jicaque-Hokan hypothesis with a list of 30 resemblances between Jicaque, Chontal, Comecrudo, Seri, Proto-Pomo and Proto-Palaihnihan. According to

him, “[t]he Jicaque-Hokan hypothesis looks even better when the full spectrum of Hokan languages is compared”.

Oswalt (1977) discusses in detail the Proto-Pomoan word *hibal ‘tongue’ and its development in Pomoan languages. He also briefly lists cognates from other Hokan branches. This type of an in-depth treatment of an individual etymology is unfortunately quite rare in Hokan studies. Especially interesting is Oswalt’s demonstration that Pomoan languages have a fossilized noun prefix *Hi-, frequent in terms for body parts.

Jacobsen (1979) is a detailed survey of Hokan inter-branch comparisons. Jacobsen gives some figures to show to what extent errors in badly controlled Washo data recorded by Dixon and Kroeber can vitiate the comparisons: “In my opinion at least 10 (16.7%) in the Dixon and Kroeber list, and 14 (13%) in the Sapir list fall into this category. Most of the mistakes involve misanalysis of the forms, usually failure to recognize the presence of more than one morpheme”. A special critical section deals with Gursky (1974)[5]. Analyzing two-member sets from Gursky’s compilation, Jacobsen comes to the conclusion that 89 sets represent probable loanwords: “These 89 dubious sets represent 20.5% of the two-member sets, and 12.6% of the total number of sets”. He also finds “the strong onomatopoetic component in at least 23 of the sets, mostly bird names”.

Langdon (1979) compares Pomoan and Yuman — the two branches of Hokan for which solid reconstructions were available. Langdon’s work is unlike other “binary comparisons” produced at the time. Instead of merely listing sound correspondences, she tries to show how comparison with Yuman can shed light on the problems of Pomoan reconstruction. According to Langdon (1979: 594-595), “Hokanists recognize that the concept of "Hokan" or for that matter, of any of the other groupings subsuming such ancient connections, differs somewhat from that of the prototypical "language family" in that the relationships it encompasses, while probably partly genetic in the traditional sense, must have been also affected by early as well as ongoing non-genetic areal pressures”. One might argue that prototypical “language families” in this sense, i.e. cases where genetic relationship is not accompanied by subsequent contact, simply do not exist. Langdon’s advice is quite reasonable: “the rational way to approach the "Hokan problem" is to work within the strong hypothesis that Hokan is in fact a language family in the classical sense, thus keeping alive the eventual aim of fully established sound correspondences and solidly reconstructed meaningful elements in the context of their grammatical structure, while simultaneously welcoming and actively seeking evidence for diffused and universal traits”. Langdon shows that some superficial similarities between Pomoan and Yuman languages are due to parallel development. Thus, voiceless resonants, found in Eastern Pomo and Diegueño, result from independent innovations. Langdon suggests that contrast between /t/ and /ṭ/, known to be secondary in Yuman languages (in Proto-Yuman, [t] occurred pre-stress, [ṭ] — post-stress), may also be secondary in Pomoan. She notes that “among the affixes reconstructed by Oswalt, only *t and *t^{h} appear” (Langdon 1979: 603). Langdon also demonstrates the parallelism between the effect of stress in Yuman, where consonants are never lenited in pre-stress (root-initial) position, and Pomoan, where pre-stress (root-initial) consonants are accompanied by a “laryngeal increment” (/h/ or /ʔ/). A similar process is attested in Yana (Langdon: 1979: 614-615). The article also includes a discussion of morphosyntactic parallels between Pomoan and Yuman, and a list of 49 lexical comparisons between the two families. Most of Langdon’s comparisons look much more probable than average comparisons by other authors of that period.

Webb (1979) compares terms for ‘tree’ and particular tree species across Hokan family. Each of her comparisons is represented in many daughter branches, thus minimizing the risk

[5] For this work see Section 4 below.

of areal loans. Despite the statement that "[t]he cognate lexeme sets are recognized on the basis of recurrent consonant correspondences", no such correspondences are given in the paper. Some of the compared forms are unreliable. E.g., Webb gives Yana *ʔaxa* 'tree', a form I was unable to find in the primary sources. Webb's inclusion of some Proto-Uralic forms as external comparanda raises the question of just how easy it is to find accidental similarities in the absence of recurrent sound correspondences.

The main part of Campbell & Oltrogge's (1980) article is a detailed reconstruction of Proto-Tol. The paper also contains a brief discussion of external affiliation of Tol. The authors suggest a specific relationship between Tol and Oaxaca Chontal, listing 57 probable cognates. While some of their comparisons do indeed look probable, e.g., Proto-Tol *lo 'leaf' ~ Highland Chontal *-la* 'leaf', some others can be shown to be wrong. Take, for example, the following comparison: "*to be sick:* WJ pte-vé *sick,* EJ peʔ, paȼ *die;* -tafba *sick*; -fwana *sick*" (Campbell & Oltrogge 1980: 221). Highland Chontal *-tafba* 'sick' is actually *ɬ-aštaf-ba* 'sick person', derived from the verb *d-eštaf-ʔma* 'to make sick; to get sick'. This form is unrelated to Highland Chontal *ɬ-af-gwana* 'sick one', derived from *ʔal-gwana* 'sickness'. The authors' conclusion that "regardless of the ultimate outcome of distant genetic research involving Hokan generally ... the Jicaque-Tequistlatec relationship will hold up" (Campbell & Oltrogge 1980: 222–223) is certainly too optimistic.

Webb (1980) compares Esselen with other Hokan languages. Again, some of her data are unreliable: I was unable to find the source of Washo *amatu-* 'land/earth'.

Webb (1981) lists Hokan comparisons for some kinship terms as well as terms such as 'man' and 'woman'. Here, too the compared material is highly unreliable: the word *xmár* 'boy/child', listed as Chimariko, is actually Yuman.

While Hokan comparative studies of that period used new accurate field data, they failed to create a working model of comparative Hokan. This is at least partly due to a nonstandard methodology which is discussed in detail in Section 6 below.

4. Attempts at synthesis

The next period in Hokan studies, partly overlapping with the previous period, was characterized by attempts at synthesis.

The work of Karl-Heinz Gursky (1964, 1965, 1966a, 1966b, 1968, 1974, 1988, 1989, 1990, 1995) logically continues that of Sapir. Like Sapir, Gursky was initially interested in large-scale classification. In his early work he compares Hokan with two poorly attested isolated languages, Quinigua (1964) and Waikuri (1966b), as well as with "Gulf" and the wider "Algonquian-Gulf" (1965, 1966a, 1968) — hypothetical macrofamilies suggested by Mary Haas. Later, Gursky turned to collecting lexical and grammatical comparative sets for Hokan itself. Especially important is Gursky's magnum opus "Der Hoka-Sprachstamm" (1974) with 707 comparative sets — the largest published compilation of Hokan lexical comparisons. Subsequently, Gursky published a three-part addenda et corrigenda (1988, 1989, 1990). While Gursky occasionally gives preliminary Proto-Hokan reconstructions (marked with # rather than with the usual asterisk in order to emphasize their tentative nature), he thought that a reconstruction of Proto-Hokan would be premature (1974: 173–175). Although methodologically his work adopts the "mass comparison" approach, it is strikingly different from the most widely known examples of this approach, such as Greenberg (1987). Gursky pays close attention to phonological accuracy of the forms compared, trying to verify all the forms with the help of the most reliable sources. He painstakingly distinguishes between forms which can be transcribed phonologically, and forms known only in pre-phonological notation. His attention to

details led him to an unexpected discovery that Olmsted's Achumawi dictionary (Olmsted 1966), based on Jaime de Angulo's fieldnotes, contains a number of Eastern Pomo words. These words were taken from de Angulo's manuscript comparison of Achumawi and Eastern Pomo, compiled with the aim to show that these two languages are unrelated (Gursky 1987). In his paper "Some grammatical evidence for the Hokan stock" Gursky (1995) gives 138 comparative sets for grammatical morphemes. This is undoubtedly the most extensive compilation of grammatical evidence for Hokan. One can agree with Terrence Kaufman's assessment of Gursky's work: "The contribution of Gursky to comparative Hokan studies is in my view quite valuable and probably not given as much attention as it merits. I was amazed at the number of undoubtedly valid grammatical comparisons found in Gursky 1966b. Although his scope of comparison was broader that just Hokan, the Hokan data can be examined by itself. ... The work of K-H Gursky in assembling likely Hokan lexical cognate sets (especially Gursky 1974) deserves the highest praise also. Even though he used all the previously published work of other scholars, he has found much that is new, he has judiciously split apart sets that were unlikely in the first place, and has brought all this together in the compass of two or three major articles" (Kaufman 1989: 111).

In his well-known monograph "Language in the Americas", Joseph Greenberg (1987) gives a list of 168 exclusively Hokan comparisons plus 100 Hokan comparisons with parallels in other "Amerind" languages. Greenberg's Hokan is maximally inclusive: it consists of "Achomawi (including Atsugewi), Chimariko, Chumash, Coahuilteco, Comecrudo, Cotoname, Esselen, Jicaque, Karankawa, Karok, Maratino, Pomo, Quinigua, Salinan, Seri, Shasta (including Konomihu, probably a distinct language), Subtiaba (including Tlappanec), Tequistlatec (sometimes called Chontal of Oaxaca), Tonkawa, Waicuri, Washo, Yana, Yuman, and Yurumangui" (1987: 132). Thus, Greenberg calls Hokan what others called "Hokan-Coahuiltecan". Greenberg was an advocate of the method of "mass comparison" and his Hokan comparisons are assembled based on superficial similarity rather than sound correspondences. Greenberg's treatment of Salinan and Yurumangui data was criticized by Poser (1992).

The years 1988–1990 saw two attempts at reconstructing Proto-Hokan based on regular sound correspondences. One of these was made by Dmitry Leshchiner with the assistance of Sergei Nikolaev (Leščiner 1989, 1990; Leshchiner & Nikolaev 1992). Leshchiner reconstructs Proto-Hokan consonants, vowels and tones, giving their reflexes in Achumawi, Karuk, Yana, Proto-Pomoan, Proto-Yuman, Seri and Tequistlatecan. The reconstructions are accompanied by 295 cognate sets, which makes Leshchiner's work the only one with both cognate sets and explicit Proto-Hokan reconstructions. Unfortunately, Leshchiner's reconstruction has a number of serious drawbacks. First, his Proto-Hokan consonant and vowel systems (Tables 2 and 3) are much larger than any system in any of the attested daughter languages. He reconstructs 58 consonants and 11 vowels (not counting vowel length). Typologically, this system looks rather unexpected when compared to attested phonological systems of Hokan and other Californian languages.

Table 2. Proto-Hokan consonants, after Leshchiner & Nikolaev (1992).

^{m}b	b	d	ḍ								
^{m}p’	p’	t’	ṭ’	ƛ’	č’	č^{w}’	k’	(k^{w}’)	q’	q^{w}’	
^{m}p^{h}	p^{h}	(t^{h})	ṭ^{h}	ƛ^{h}	č^{h}	č^{wh}	k^{h}	k^{wh}	q^{h}	q^{wh}	
^{m}p	p	t	ṭ	ƛ	č	č^{w}	k	k^{w}	q	q^{w}	
^{m}f	f	s		ɬ	š	š^{w}	x	x^{w}	ẋ	ẋ^{w}	
^{m}w	w	r	ṛ	l	y						ʔ
	m	n	ṇ				ŋ				

Table 3. Proto-Hokan vowels, after Leshchiner & Nikolaev (1992).

i	ü	ɨ	u
e	ö	ə	o
ɛ		a	ɔ

Second, probable borrowings were not systematically excluded: "When we weren't sure that the words compared are not true cognate, then this item remains in the dictionary. We note the cases which strongly seem to be borrowings" (Leshchiner & Nikolaev 1992: 366). This strategy can only lead to proliferation of spurious sound correspondences, which in turn leads to overbloated protophonology. Third, many comparisons exhibit loose semantics, leading one to suspect that the forms compared are neither true cognates nor loans. Thus, the list of Hokan body parts (Leshchiner & Nikolaev 1992: 373–380) contains the following comparisons: 'cheek ~ forehead'; 'nose ~ forehead'; 'brain, Adam's apple, head ~ eye'; 'spleen ~ milt ~ body, meat'; 'back ~ shoulder blade ~ elbow'; 'calf of leg ~ elbow'; 'shoulder blade ~ arm ~ finger, hand ~ thigh, inside of leg'; 'palm of hand ~ arm above elbow ~ thigh'; 'ankle ~ nail'. Finally, the regularity of correspondences is also in doubt. As an example, we can look at the reflexes of the alleged Proto-Hokan phoneme *^{m}b (see Table 4). This phoneme should give /v/ in Karuk, at least word-initially, but out of four examples with Karuk reflexes, two have /v/, one /p/ and one /ʔ/. Yana should have /m/, but out of five cases with Yana reflexes, three have /m/ and two /p/ without any complementary distribution. Moreover, none of Leshchiner's papers give any comments on complementary distribution of reflexes. Overall, Leshchiner's work cannot be used as a working model in further attempts to reconstruct Proto-Hokan.

Table 4. Reflexes of Proto-Hokan *^{m}b in Leshchiner & Nikolaev's (1992) reconstruction.

PHo*	Ač	K	Ya	PPo*	PYu*	Se	Ts
*^{m}b	m/p-	v-	m-	b	m-	?-/-p-	ø-
*^{m}bāk'ő 'back'	pōk	vásih	mak'i 'back, backbone'	*bak'ő	*mák		(?) la-picúlâ? 'buttocks'
*si^{m}b[ü] 'moss'	simi?túl (PR sa**p̄'**a)			(Pe) [xālé] sībú 'tree moss'			
*xä^{m}bu 'carrot'	hamuč (PR h'ām'ač) 'wild carrot'			(Pe) šībú			
*^{m}bā̄$_{\circ}$xʌ 'bear'	**w**ah' (PR woh')		mūxa (N) 'porcupine'		*max$_{\circ}$á 'badger', *max$_{\circ}$áta 'bear'		
*^{m}b[ō]wʌ́ (~-ā̄- ?) 'elk'	páw (PR puwwāwi 'cow')	**p**ŭfič 'deer, venison'	**p**a(na) 'deer, deer meat'	*bo?ó 'to hunt; elk'	*ma?úl 'antelope'	?ap 'venado buro'	
*^{m}bHéšö 'meat'	**m**ísuč (PR mīsuč')	**ʔ**íš	**p**asi (N), **p**ahsi (Y)	*bixe		ipp̣ási	l-išík'
*^{m}baka 'wood-pecker sp.'	**m**akmáka			(Pe) bākáka			
*^{m}būk'ā́ 'worm'	amuq (PR āmoq' 'caterpillar, grub, maggot')	vákaj	cāmuk'u (N) 'acorn worm'				ɬ-âmúšâk^{h} 'large white worm sp.'

In 1989, Terrence Kaufman, a leading authority on Mesoamerican languages, published his work "A Research Program for Reconstructing Proto-Hokan: First Gropings". The 119-page paper contains a survey of Hokan phonological systems, tables of sound correspondences between Hokan branches and a list of 1118 Proto-Hokan reconstructions. Kaufman's Hokan includes the following languages: Pomoan, Chimariko, Yana, Karuk, Shastan, Achumawi-Atsugewi[6], Washo, Esselen, Salinan, Yuman, Cochimí, Seri, Coahuilteco, Comecrudoan, Chontal, Tol. Kaufman reconstructs the following phonological system for Proto-Hokan (see Tables 5, 6).

Table 5. Proto-Hokan consonants, after Kaufman (1989)

p	ṭ	t	¢	č	k^y	q	k^w	7
p'	ṭ'	t'	¢'	č'	$k^{y'}$	q'	$k^{w'}$	
p^h	$ṭ^h$	t^h	$¢^h$	$č^h$	k^{yh}	q^h	k^{wh}	
f		θ	s	š	x^y	ẋ	x^w	h
	r							
	l	l^y						
m	n	n^y						
					y		w	

Table 6. Proto-Hokan vowels, after Kaufman (1989)

i		u
e		o
	a	
length /:/		
stress /ˊ/		

This work might have put an end to the debate on the validity of Hokan, were it not for one thing: Kaufman lists his reconstructions but omits the data underlying those reconstructions. The absence of reflexes in Kaufman's work makes its evaluation a hard, although not impossible, task. Kaufman's Proto-Hokan phonological system looks typologically and areally plausible for a Californian language, and his reconstructions look more realistic compared to those by other authors. Compare Kaufman's #(a-)l^yafu 'navel' (Kaufman 1989: 133) with Haas' *imarak̆^{w}i ~ *imak̆^{w}ari 'navel' or Leshchiner's *ʔăHlúp'ʌ 'navel'. Kaufman's protoform looks more similar to such forms as Chimariko *-onapu* 'navel', Karuk *ʔárup* 'navel, navel cord', Shasta *ʔé·raw* 'navel' and Proto-Yuman *ml^{y}pu(:) 'navel' (Miller 2018: A61).

Nevertheless, some of Kaufman's reconstructions must be rejected in the light of new evidence. For instance:

(a) #saHmaH [s ~ š] 'heart' (Kaufman 1989: 133) is based on the following comparison by Gursky (1974: 193):

Eastern Pomo *sa·ma·i* 'heart', Highland Chontal *-unšahmaʔ* 'heart'.

[6] It is widely believed that Achumawi and Atsugewi constitute a language family (or a subfamily of Hokan), sometimes called Palaihnihan. There is a published reconstruction of Proto-Palaihnihan phonology (Olmsted 1964). However, Nevin (2019) has persuasively shown that Olmsted's reconstruction is based on faulty methodology and inaccurate data. According to Nevin, "a proto-language ancestor common to Achumawi and Atsugewi has not been reconstructed, and they cannot be claimed to be more closely related to each other than either is to the Shastan languages or perhaps even to Yana" (Nevin 2019: 48).

In fact, the Highland Chontal form is derived from the verbal stem *-unšá-* 'to breathe' with the instrumental suffix *-hmaʔ*, therefore, the comparison should be rejected.

(b) #i-7ip 'navel' (Kaufman 1989: 133) is based on the following comparison by Judith G. Crawford (1976: 316):

Seri *ʔatóošx iʔíip* 'navel', Washo *íʔib* 'navel'

The Seri word literally means 'where the umbilical cord is', from *ʔatoošx* 'umbilical cord' and *iʔiip* 'where it is', derived from the verb *kaap* 'to be' (Moser & Marlett 2010: 76, 379, 435). While the similarity of Washo and Seri forms is striking, it is undoubtedly due to chance resemblance.

(c) #panA [a ~ o] 'river' (Kaufman 1989: 160) is based on a comparison by Campbell & Oltrogge (1980: 222):

Tol *sɨ-pones* 'river', Highland Chontal *-banaʔ* 'river'

The Tol word is most likely a compound of *ʔɨsɨ́* 'water' (Dennis & Royce de Dennis 1983: 55) and *pɨné* 'big' (Dennis & Royce de Dennis 1983: 30).

In some other cases, Kaufman's reconstructions can be amended based on new data.

(d) #x̣owK 'coal' (Kaufman 1989: 142) is probably based on the following entry in Gursky (1974: 183): "COAL Ach //həwk// *coal*; Sh //xúk^{w}// *coal (ember), charcoal*; Chi kōwa *coals*". Judging by Kaufman's reconstruction and correspondences, the Chimariko word does not belong here. The diphthong in #x̣owK is apparently based on the Achumawi word, for which the online "Achumawi Dictionary" gives variants *ɬúk* and *ɬuq̓* (Nevin 2020). Thus, we can safely assume that *ow should be replaced with *u in the reconstruction of this word.

The cases above are, however, exceptional, and on the whole Kaufman's comparisons hold water remarkably well. For example, Kaufman reconstructs the following pair of words for Proto-Hokan: #a-px̣a 'shit' and #i-px̣a 'guts' (Kaufman 1989: 133). While some daughter languages preserve only one of the two words, the opposition between the two is preserved in Southern Pomo *ʔahp^{h}a* 'excrement', *ʔihp^{h}a* 'intestines, guts' (Oswalt 1981: 17, 22) and Chimariko *ápx̣a* 'excrement', *ípx̣a* 'intestines' (Berman 2001: 1053). Since there is no evidence of contact between Chimariko and Pomoan, this remarkable parallel is best viewed as inherited from the common ancestor. The two words share the same root but have different fossilized prefixes. These prefixes are not synchronically attested in Pomoan and Chimariko, but they can be identified with Kaufman's Proto-Hokan #7a:- 1. 'absolutive of intimately possessed noun'; 2. 'substance or mass noun prefix' and #Hi:- 1. 'body-part prefix'; 2. 'possessed state of intimately possessed noun' (Kaufman 1989: 117), thus confirming Kaufman's morphological reconstruction. Another pair of reconstructions having the same prefixes is #a:-x^{y}á7 'water' and #i-x^{y}a7 'juice' (Kaufman 1989: 131). Here, the opposition of the two prefixed forms is preserved in Seri *ʔaχ* 'water' vs. *iχ* 'liquid, sap, juice' and Highland Chontal *l-ahaʔ* 'water' vs. *l-iha* in *l-iha-ʔmał* 'mescal (lit. its-water the maguey)' (Turner & Turner 1971: 179, 194). Unlike the case with Chimariko and Pomoan, Seri words contain productive prefixes *ʔa-* 'absolute (without explicit possessor)' (Moser & Marlett 2010: 298–299) and *i-* 'third person possessor' (Moser & Marlett 2010: 386). In this case we deal not with isolated lexical parallels, but with reconstructed items that can be morphologically segmented on the Proto-Hokan level.

Kaufman (2015)[7] provides a model of Proto-Hokan grammar and a list of approximately 150 reconstructions of grammatical morphemes. Once again, the reflexes in daughter lan-

[7] This paper is a slightly updated version of a manuscript written in 1989.

guages are omitted, but, unlike Kaufman 1989, this study lists the languages on which a particular reconstruction is based.

Langdon (1990) offers a new approach to Hokan morphological comparison: "The alternative then suggested itself to focus on elements of the stem without prejudging whether they function as prefixes, roots, or suffixes. Crucial here is Jacobsen's (1980) notion of bipartite verb stems in Washo, where two basic parts of a stem are postulated but where each part can consist of a root. So instead of thinking of a Hokan verb stem as consisting of one (or more) prefixes plus a root, it seemed more appropriate to focus simply on how many distinct parts a verb stem consists of, and the semantic patterns in which they combine. The consequence is a view of Hokan verb forms as consisting of a sequence of elements (possibly all basically lexical) whose various combinations and recombinations in the daughter languages reshape the identifying criteria of the distinction between root and affix so that the issue of what is a root vs. what is an affix, while important for the synchronic structure of each language, is perhaps less crucial historically" (p. 66). Langdon reconstructs the following Hokan kernel stem structure (Table 7):

Table 7. Hokan stem structure, after Langdon (1990)

I	II	III	IV	V
Bipartite I	Plural infixed	Bipartite II	Direction/Motion	Plural

"The basic insight — if that is what it is — is that in a particular language and from the perspective of comparative Hokan, one cannot assume that prefixes must be compared to prefixes (and only prefixes) and roots only to roots" (1990: 67). For example, "Bipartite I" yields so-called instrumental prefixes in, e.g., Pomoan, but so-called primary stems in Yana.

5. Hokan studies on the wane

The period from the 1990s up to the present is characterized by a decline of interest in the Hokan hypothesis. Symbolically, the last of the Hokan-Penutian conferences, which were regularly held since the 1970s, took place in the year 2000 (Golla 2011: 304–305).

Poser (1995) argues against Greenberg's (1987: 132) assertion that binary (rather than multilateral) comparisons have done much harm to Hokan studies. Poser points out that even so-called "binary comparisons" often included forms from all branches of Hokan, being thus essentially multilateral in Greenberg's sense. "Ironically, the real problems with Hokan classification are due to the very causes that make Greenberg's methodology questionable. What is controversial about Greenberg's methodology is not his comparison of many languages. It is his failure to establish the systematic correspondences between languages that remove the possibility of chance similarity, his lack of concern for loans, and his willingness to postulate relationship on the basis of minute amounts of evidence, precisely the weaknesses that affect parts of Hokan. ... If we are to learn more about which languages really belong to Hokan and how the family is to be subgrouped, it will be via the program laid out by Haas (1963), namely, the acquisition of more and better data, the establishment of phonological correspondences, and the reconstruction of Proto-Hokan and intermediate proto-languages. It is not from the assemblage of lists of vaguely resemblant words that we can expect to improve our knowledge, but from the fieldwork that continues to be done, the study of the vast store of Harrington material, and such initiatives as Kaufman's (1988) reconstruction of Proto-Hokan" (Poser 1995: 142).

Lyle Campbell's (1997) monograph "American Indian Languages: The Historical Linguistics of Native America" contains a section with a detailed critique of the Hokan hypothesis. His justified criticism of Haas' methodology was already cited above. Nevertheless, some of Campbell's criteria seem overly strict. Thus, by his count, twenty-six comparisons in Haas (1964) are to be rejected as "short forms" (1997: 294). The problem is that many Hokan languages have short roots. Thus, Jamul Tiipay roots have the canonical structure (C)V(C) (Miller 2001: 11). Excluding short forms from comparison amounts to declaring that the comparative method is inapplicable to languages with short canonical roots (e.g., many languages of Southeast Asia). Campbell's verdict on Hokan is rather pessimistic: "Given the reservations expressed here concerning the many Hokan studies, but also taking into account Langdon's more promising comparisons (and Kaufman's optimism—see the next subsection), I conclude that it is by no means clear or even likely that there was a proto language from which some or most of the putative Hokan languages diverged long ago, but that this hypothesis is fully worthy of continued research". Regarding Kaufman's (1989) work, Campbell says: "While Kaufman's proposals have stimulated some other linguists to accept more positive attitudes toward Hokan, they can be evaluated appropriately only after he presents the lexical evidence upon which they are based. Therefore, for the present, we are left with essentially the same uncertainty that has always attended the Hokan hypothesis — there certainly is enough there to make one sympathetic to the possibility of genetic relationship, and yet the evidence presented to date is not sufficient to confirm the hypothesis, regardless of which languages are included".

Mixco (1997) reviews the work of Mary R. Haas on Hokan. According to him, "Haas revealed herself to be an exemplar of a truly "neogrammarian" stamina in the face of the most persistently intractable problems that tax the student of the linguistic prehistory of the Americas (or of the world, for that matter). Time and time again, she treats Hokan as a hypothesis worthy of continuing intellectual devotion. At no point does the knotty tangle of obstacles in the path of a definitive reconstruction discourage her from further efforts". Mixco defends Haas' use of metathesis in Hokan comparisons, pointing out that metathesis frequently occurs within the uncontroversial Yuman family.

In his publication of Sapir's Chimariko linguistic material, Berman (2001) lists several comparisons between Chimariko, Proto-Pomo and Proto-Yuman (see Table 8). Berman also notes that Chimariko suffixes *-kˁ* 'motion towards here' and *-m* 'motion towards there' resemble Yuma suffixes *-k* 'towards here' and *-m* 'towards there'. The pair *-k 'hither' / *-m 'thither' also occurs in Achumawi, Atsugewi and Shasta (Zhivlov 2018: 154–157).

Table 8. Chimariko-Pomo-Yuman comparisons, after Berman (2001)

	Chimariko	Proto-Pomo Oswalt	Proto-Pomo McLendon	Proto-Yuman
earth	*ám·a*	**ʔa(h)ma·*	**ʔahmáṭ/ʔamáṭ*	**ʔ-mat*
eat	*hám·a* to eat	**maʔa-*	**maʔá* food	**ma·* eat mush
excrement	*ápx̣a*	**ʔahpʰa*	**ʔahpʰá*	
hair (body) / fur	*hím·ɪ* hair (of any kind)	**zi(h)me*	**cihmé/ci·me*	**-mi(y)(s)*
intestines	*ípx̣a*	**ʔihpʰa*		**pxa* (Kiliwa *pʰaʔ*)
moon	*ála, allaʳ* moon, sun	**ʔala·ša*	*q'alá·(xa) /ʔal·á·(xa)*	**xlʸʔa·*
water	*äqʻǎ*	**ahqʰa*	**ahqhá*	**ʔ-xa*

Good (2002), using material from Californian Hokan languages, attempts to show that Proto-Hokan had a three-vowel system (*i *a *u) rather than a five-vowel system (*i *e *a *o *u), reconstructed by Kaufman (1989). However, Good's comparisons are purely structural, in the sense that no actual cognate sets are discussed. While this is understandable given that Kaufman (1989) does not list the reflexes of his Hokan protoforms, Good's conclusions can be verified or falsified only by looking at actual correspondences as manifested in cognate sets.

Marlett (2007) reviews the evidence for the relationship between Seri, Yuman and Oaxaca Chontal and comes to the conclusion that the relationship has not been demonstrated. Appendix 1 includes a revision of the resemblances between these languages proposed by earlier researchers, with detailed comments on the morphological structure of Seri forms.

Marlett (2008) gives a very useful critique of Bright's (1956) Seri-Salinan comparisons. He concludes that "[w]hen the least likely candidates are eliminated, the percentage of "correlates" between Seri and Salinan is lower than between Seri and the Otomanguean data presented as "Supanec." This would suggest that the points of comparison may be coincidences—not evidence of a genetic connection. ... Seri and Salinan should be listed as isolates until adequate evidence of a relationship with some language is published and evaluated".

In his book "California Indian Languages", Victor Golla gives a brief survey of "The Hokan Phylum" (Golla 2011: 82–84) as well as a sketch of the linguistic prehistory of Hokan languages (Golla 2011: 242–248). Golla summarizes his position on the Hokan and Penutian hypotheses in the following way: "My position on the historical validity of these hypothetical relationships is somewhat more positive than that of such methodological conservatives as Lyle Campbell (1997), although considerably more cautious than that of such enthusiasts as Paul Radin (1919), Morris Swadesh (1959), and Joseph Greenberg (1987), or even Edward Sapir in his less restrained moments (as, for example, in Sapir 1921c). While the categories "Hokan" and "Penutian" are undoubtedly meaningful—no one would ever classify Shastan as Penutian or Klamath-Modoc as Hokan, and the consensus is now clear that Chumash and Yukian belong to neither (nor to a phylum of their own)—we continue to debate the nature of the historical relationships these categories imply. … I hope, however, that I make clear the highly speculative nature of such hypotheses, and I try to indulge my predilection for them only when, in my view, they appear to cast some useful light on the past" (Golla 2011: 239).

Jany (2013) is a comprehensive bibliography on Hokan languages. The author takes a neutral stance on the question of their genealogical relationship.

Haynie (2014) is a publication of one of the chapters of the author's PhD thesis (Haynie 2012). The work applies a sophisticated statistical methodology to the question of deep relationship between California languages, comparing the initial segments of words in search for recurrent sound correspondences. The results are rather unfavorable for the Hokan hypothesis: "Both in its measure of relationship strength and in its evaluation of statistical significance, this test of linguistic relationships among California languages offers no support for Hokan or Penutian" (Haynie 2014: 439). It must be noted, however, that the limitation of comparison to initial segments does not suit well languages with developed prefixation, where prefixes can easily get fossilized. Consider the following words for 'tongue' from Haynie's wordlists: Achumawi *iplē·*, Cocopa *mpał^{y}* and Southeastern Pomo *bal* (Haynie 2012: 283). It is clear that the historically relevant comparison in this case should include Achumawi *p*, Cocopa *p* and Southeastern Pomo *b* rather than Achumawi *i*, Cocopa *m* and Southeastern Pomo *b*.

Jany (2017) reviews features of polysynthesis in the so-called Northern Hokan languages and comes to the conclusion that "[s]tructural similarities stem from language contact rather than from genetic affiliation".

Zhivlov (2018) gives a systematic survey of some aspects of reconstructible Hokan morphology, namely person and number markers, lexical prefixes, plural infixes, and directional suffixes. The article ends with the following conclusion: "While it has been frequently pointed out that areal diffusion in Northern California makes it hard to distinguish between genetic relationship and borrowing (Jany 2016), some results of the present study suggest a more optimistic outlook. Thus, if the hypotheses offered in this paper are correct, some features of Proto-Hokan (plural infixes and a simple directional system opposing 'hither' and 'thither' forms) are better preserved by languages outside of the Northern California linguistic area." (Zhivlov 2018: 158).

In his monograph on Esselen, Shaul (2019) provides evidence for the Hokan affiliation of this language. In Shaul's own words, "When I began the research for this reworking of Esselen, I was skeptical about the Hokan hypothesis. I am no longer: Kaufman's comparative treatment of Hokan (2015) gave me ways of explaining things in Esselen that I could not otherwise explain". Shaul gives Esselen reflexes for Kaufman's Hokan reconstructions and adds a number of Hokan reconstructions of his own.

Shaul (2020a) provides similar evidence for the Hokan affiliation of Salinan. He defines a "Hokan common core", consisting of 279 reconstructions with wide distribution in Hokan. Out of these, 126 have reflexes in Salinan. These reflexes are listed in Appendix A together with Proto-Hokan reconstructions, but without reflexes in other Hokan branches. This makes it difficult to verify proposed Hokan etymologies of Salinan words. One more problem is that Shaul's works on Salinan (2020a, 2020b) completely ignore one of the two phonologically reliable sources of Salinan data — J.P. Harrington's fieldnotes (the other reliable source — William Jacobsen's fieldnotes — is less extensive). With data that are largely unreliable in respect to their phonology, it is much harder to establish regular sound correspondences.

Lyle Campbell's (2024) new monograph "The Indigenous Languages of the Americas: History and Classification" has a section on the Hokan hypothesis. Compared to the similar section in Campbell (1997), it contains a detailed discussion of Terrence Kaufman's work on Hokan. Among Kaufman's 1118 Hokan comparisons, Campbell finds 242 which occur only in a single linguistic area, 15 with considerable semantic difference in the languages compared, 58 onomatopoetic forms, 17 nursery forms, 23 possible borrowings, as well as a number of short forms and words found also in languages beyond Hokan (Campbell 2024: 348–352). Campbell does not deny, however, that there is a core of very stable words found in most of Hokan languages. These words "must include at least forms with the following meanings: arm/shoulder, ashes, body/meat/deer, bone, dog, earth/ground, eye, grease/fat, hand/give, knee, many, skin, sleep, stone, tail, tongue, and water" (Campbell 2024: 386). He entertains the idea that "the thirty or so words mentioned above that seem like possible cognates that turn up across many of the "Hokan" languages might be inherited not from some Proto-Hokan, but from some Ur-American language of long ago, maybe even one spoken by some of the first inhabitants in the New World". Another possible option is "to suspect that these thirty or so words are not evidence of a Hokan genealogical connection among those particular languages but rather are instead pan-Americanisms", i.e. "those seemingly similar forms that appear to pop up repeatedly in broad comparisons of American Indian languages not known to be related to one another, probably a collection of accidentally similar forms, loanwords (maybe including some Wanderwörter), unrecognized onomatopoeia, sound symbolism, and the like" (Campbell 2024: 353). However, Campbell makes no attempt to demonstrate that these "core" Hokan comparisons actually occur in other Native American language families, and without such a demonstration the idea that they go back to a deeper level than Hokan remains speculative. Campbell's overall assessment of Hokan remains pessimistic: "I, however, remain very

skeptical about Hokan in general. Nevertheless, there are those tantalizing recurrent similarities that need explanation, and it would be difficult to reject Kaufman's work entirely, and so research on possible Hokan connections should continue. However, my strong belief is that if there ever was a Hokan genetic unity, it was so far in the past and so much has changed in the languages since then that it will never be possible to confirm a genetic relationship among those languages on the basis of the information that remains today" (Campbell 2024: 353).

6. Conclusion

There are two possible answers to the question of why there is still no convincing demonstration of Hokan relationship — or, vice versa, no convincing attempt to explain away all similarities as due to contact. One possibility is that the problem lies in the nature of evidence — it is so meager that the comparative method cannot be applied (either because the languages are not related or because of the great time depth of the family). This is essentially the view of Poser (1995: 140), when he says that "[m]any of the links were originally posited on the basis of very slim evidence, consisting of unsystematic similarities in a very small number of words". Cf., however, the opposite point of view voiced by Kaufman (1989: 64): "It is a common observation that there is a very small number of etymologies (usually characterized as being "about thirty") that have reflexes in most of the Hokan families and isolates, and that there do not seem to be any more new and good sets forthcoming. This is not a totally appropriate characterization nor does it show a very constructive attitude. A number of new correspondences and etymologies have been dug up in the course of the binary comparisons referred to above, and one should not expect proto-forms to survive in most of the daughter languages. A random scatter is quite good enough and is in any case what is to be expected".

Another possible answer, which I find more compelling, is that the problem lies with the methodology used by the majority of the proponents of Hokan. From this point of view, two types of studies can be delimited. One type uses the method of mass comparison, i.e. comparisons of words based on superficial resemblances rather than on sound correspondences. Studies of this type include Sapir (1917, 1920c, 1925) and Gursky (1974, 1995). At the time of Sapir, there was no alternative to mass comparison in Hokan studies, since there were no phonologically accurate materials on most Hokan languages, and consequently, no possibility to apply the comparative method. As noted above, Gursky used mass comparison because he thought that applying the comparative method on the level of Proto-Hokan would be premature. While mass comparison can be useful as a first step in collecting the materials for further application of the comparative method, its usefulness is limited, and no conclusive demonstration of distant language relationship can be achieved by this method.

The second type of studies, mostly authored by Mary Haas and her students, applies a methodology which for the lack of a better name I will call "partial comparative method". The difference of this methodology from the standard comparative method was not commented upon previously. The traditional comparative method is a cyclic procedure. It involves finding sound correspondences, looking for complementary distributions between them, and then assigning each correspondence or a group of correspondences to a phoneme of the protolanguage. The next step involves getting rid of comparisons that do not fit in the system and adding new comparisons that do show systematic, even if initially unexpected, correspondences. Then one should revise the correspondences in search of new complementary distributions. The initial set of comparisons is thus constantly revised in order to arrive at more precise protoforms. As a result one gets a coherent system of sound correspondences. The "partial

comparative method" is different. It requires listing sound correspondences found in the data and sometimes grading them by frequency. No attempt is made to revise the correspondences in order to fit them into one system. This allows a reasearcher to retain multiple intersecting correspondences without complementary distributions. Cf. the following subset of Seri-Yuman correspondences in J. G. Crawford (1976: 321–322): t ~ *t, t ~ *ṭ, t ~ *s, t ~ *š, s ~ *s, s ~ *š, š ~ *s, š ~ *š, š ~*t, š ~ *ṭ. The traditional comparative method would require either finding complementary distributions or rejecting some of the correspondences together with lexical comparanda containing them. Another feature of the "partial comparative method" is the looseness when it comes to the formal accountability of compared words. This is obvious in all those cases where entire wordforms are brought into comparison as evidence for the correspondence of individual segments only, with no account being offered for the residue.

The crucial stages missing in the "partial comparative method" are searching for complementary distributions and rejecting those comparisons which do not fit the system. As a result of this, such studies contain many quite improbable or downright wrong comparisons, which is one of the reasons why the Hokan hypothesis lost its appeal to many researchers in the following decades.

Only a small minority of Hokan studies attempt to apply the traditional comparative method. Leshchiner & Nikolaev's (1992) use of the comparative method, while seemingly more sophisticated than the "partial comparative method" of Haas and her students, is marred by inconsistencies in sound correspondences and semantic looseness, which is why their results cannot serve as a proof of Hokan hypothesis.

The only study so far to consistently employ the traditional comparative method is Kaufman (1989), but the absence of actual data makes it quite difficult to verify its results.

Overall, while the genetic nature of relationship between Hokan languages seems plausible in the light of some works reviewed above, especially that by Sapir and Kaufman, none of the studies published so far can be said to constitute a definite proof of the Hokan hypothesis. Still, most of these studies contribute at least something to the growing list of probable comparisons between Hokan branches. The fact that in many of Hokan studies such comparisons are buried beneath a heap of "etymological rubbish" should not discourage us from trying to verify or falsify these comparisons in the light of currently available data.

References

Berman, Howard. 2001. [Sapir's] Chimariko linguistic material. In: Victor Golla, Sean O'Neill (eds.). *The Collected Works of Edward Sapir XIV: Northwest California Linguistics*: 1039–1080. Berlin / Boston: De Gruyter Mouton. doi:10.1515/9783110879803.1039.

Bright, William. 1954. Some Northern Hokan relationships: a preliminary report. In: *Papers from the Symposium on American Indian linguistics held at Berkeley July 7, 1951* (University of California Publications in Linguistics 10): 63–67. Berkeley / Los Angeles: University of California Press.

Bright, William. 1956. Glottochronologic counts of Hokaltecan material. *Language* 32(1): 42–48. doi:10.2307/410651.

Bright, William. 1957. *The Karok language* (University of California Publications in Linguistics 13). Berkeley / Los Angeles: University of California Press.

Bright, William. 1970. On linguistic unrelatedness. *International Journal of American Linguistics* 36(4): 288–290. doi:10.1086/465123.

Bright, William & David L. Olmsted. 1959. A Shasta vocabulary. *Kroeber Anthropological Society Papers* 20: 1–55.

Brinton, Daniel G. 1891. *The American race: a linguistic classification and ethnographic description of the native tribes of North and South America*. New York: N. D. C. Hodges.

Campbell, Eric W. 2017a. Otomanguean historical linguistics: Past, present, and prospects for the future. *Language and Linguistics Compass* 11(4): e12240. doi:10.1111/lnc3.12240.

Campbell, Eric W. 2017b. Otomanguean historical linguistics: Exploring the subgroups. *Language and Linguistics Compass* 11(7): e12244. doi:10.1111/lnc3.12244.

Campbell, Lyle. 1976. The linguistic prehistory of the Southern Mesoamerican periphery. In: *Fronteras de Mesoamérica: 14a mesa redonda*: vol. 1, 157–183. Mexico: Sociedad Mexicana de Antropología.

Campbell, Lyle. 1997. *American Indian languages: the historical linguistics of Native America* (Oxford Studies in Anthropological Linguistics 4). New York: Oxford University Press. doi:10.1093/oso/9780195094275.001.0001.

Campbell, Lyle. 2024. *The indigenous languages of the Americas: history and classification*. New York: Oxford University Press. doi:10.1093/oso/9780197673461.001.0001.

Campbell, Lyle & David Oltrogge. 1980. Proto-Tol (Jicaque). *International Journal of American Linguistics* 46(3): 205–223. doi:10.1086/465655.

Conathan, Lisa. 2002. Split intransitivity and possession in Chimariko. In: Lisa Conathan, Teresa McFarland (eds.). *Proceedings of the 50th Anniversary Conference, June 8-9, 2002, University of California at Berkeley* (Survey of California and Other Indian Languages, Report 12): 18–31. Berkeley: Department of Linguistics, University of California. Available online at: https://escholarship.org/uc/item/01c141b1.

Crawford, James M. 1976. A comparison of Chimariko and Yuman. In: Margaret Langdon, Shirley Silver (eds.). *Hokan Studies: Papers from the First Conference on Hokan Languages held in San Diego, California April 23-25, 1970* (Janua Linguarum: Series Practica 181): 177–191. The Hague / Paris: Mouton. doi:10.1515/9783110819113-012.

Crawford, Judith G. 1976. Seri and Yuman. In: Margaret Langdon, Shirley Silver (eds.). *Hokan Studies: Papers from the First Conference on Hokan Languages held in San Diego, California April 23-25, 1970* (Janua Linguarum: Series Practica 181): 305–324. The Hague / Paris: Mouton. doi:10.1515/9783110819113-020.

Dennis, Ronald K. & Margaret Royce de Dennis. 1983. *Diccionario Tol (Jicaque)–Español y Español–Tol (Jicaque)*. Tegucigalpa: El Instituto Lingüístico de Verano en colaboración con El Instituto Hondureño de Antropología e Historia.

Dixon, Roland B. 1905. The Shasta-Achomawi: a new linguistic stock with four new dialects. *American Anthropologist* 7(2): 213–217. doi:10.1525/aa.1905.7.2.02a00030.

Dixon, Roland B. 1910. The Chimariko Indians and language. *University of California Publications in American Archaeology and Ethnology* 5(5): 293–380.

Dixon, Roland B. & Alfred L. Kroeber. 1903. The native languages of California. *American Anthropologist* 5(1): 1–26. doi:10.1525/aa.1903.5.1.02a00020.

Dixon, Roland B. & Alfred L. Kroeber. 1912. Relationship of the Indian languages of California. *American Anthropologist* 14(4): 691–692. doi:10.1525/aa.1912.14.4.02a00120.

Dixon, Roland B. & Alfred L. Kroeber. 1913a. New linguistic families in California. *American Anthropologist* 15(4): 647–655. doi:10.1525/aa.1913.15.4.02a00050.

Dixon, Roland B. & Alfred L. Kroeber. 1913b. Relationship of the Indian Languages of California. *Science* 37(945): 225–225. doi:10.1126/science.37.945.225.

Dixon, Roland B. & Alfred L. Kroeber. 1919. Linguistic families of California. *University of California publications in American archaeology and ethnology* 16(3): 47–118.

Golla, Victor. 2011. *California Indian languages*. Berkeley: University of California Press.

Good, Jeff. 2002. The vowel systems of California Hokan. In: Lisa Conathan, Teresa McFarland (eds.). *Proceedings of the 50th Anniversary Conference, June 8-9, 2002, University of California at Berkeley* (Survey of California and Other Indian Languages, Report 12): 32–46. Berkeley: Department of Linguistics, University of California. Available online at: https://escholarship.org/uc/item/1jd6h9mn.

Greenberg, Joseph H. 1987. *Language in the Americas*. Stanford, California: Stanford University Press.

Greenberg, Joseph H. & Morris Swadesh. 1953. Jicaque as a Hokan language. *International Journal of American Linguistics* 19(3): 216–222. doi:10.1086/464222.

Gursky, Karl-Heinz. 1964. The linguistic position of the Quinigua Indians. *International Journal of American Linguistics* 30(4): 325–327. doi:10.1086/464792.

Gursky, Karl-Heinz. 1965. Ein lexikalischer Vergleich der Algonkin-Golf- und Hoka-Subtiaba-Sprachen (1). *Orbis* 14: 160–215.

Gursky, Karl-Heinz. 1966a. Ein Vergleich der grammatikalischen Morpheme der Golf-Sprachen und der Hoka-Subtiaba-Sprachen. *Orbis* 15: 511–537.

Gursky, Karl-Heinz. 1966b. On the historical position of Waikuri. *International Journal of American Linguistics* 32(1): 41–45. doi:10.1086/464877.

Gursky, Karl-Heinz. 1968. Gulf and Hokan-Subtiaban: new lexical parallels. *International Journal of American Linguistics* 34(1): 21–41. doi:10.1086/464993.

Gursky, Karl-Heinz. 1974. Der Hoka-Sprachstamm. Eine Bestandsaufnahme des lexikalischen Beweismaterials. *Orbis* 23: 170–215.

Gursky, Karl-Heinz. 1987. *Achumawi und Pomo: Eine besondere Beziehung?* (Abhandlungen der Völkerkundlichen Arbeitsgemeinschaft 57). Nortorf.

Gursky, Karl-Heinz. 1988. *Der Hoka-Sprachstamm: Nachtrag I* (Abhandlungen der Völkerkundlichen Arbeitsgemeinschaft 58). Nortorf.

Gursky, Karl-Heinz. 1989. *Der Hoka-Sprachstamm: Nachtrag II* (Abhandlungen der Völkerkundlichen Arbeitsgemeinschaft 63). Nortorf.

Gursky, Karl-Heinz. 1990. *Der Hoka-Sprachstamm: Nachtrag III* (Abhandlungen der Völkerkundlichen Arbeitsgemeinschaft 65). Nortorf.

Gursky, Karl-Heinz. 1995. Some grammatical evidence for the Hokan stock. In: Michael Dürr, Egon Renner, Wolfgang Oleschinski (eds.). *Language and culture in Native North America: studies in honor of Heinz-Jürgen Pinnow* (LINCOM Studies in Native American Linguistics 2): 138–153. München.

Haas, Mary R. 1954. The Proto-Hokan-Coahuiltecan word for "water." In: *Papers from the Symposium on American Indian linguistics held at Berkeley July 7, 1951* (University of California Publications in Linguistics 10): 57–62. Berkeley / Los Angeles: University of California Press.

Haas, Mary R. 1963. Shasta and Proto-Hokan. *Language* 39(1): 40–59. doi:10.2307/410761.

Haas, Mary R. 1964. California Hokan. In: William Bright (ed.). *Studies in Californian linguistics* (University of California Publications in Linguistics 34): 73–87. Berkeley / Los Angeles: University of California Press.

Haas, Mary R. 1969. *The prehistory of languages* (Janua Linguarum Series Minor 57). The Hague: Mouton. doi:10.1515/9783110881646.

Harrington, John P. 1943. Hokan discovered in South America. *Journal of the Washington Academy of Sciences* 33(11): 334–344.

Haynie, Hannah J. 2012. *Studies in the history and geography of California Languages*. University of California, Berkeley Ph.D. Available online at: https://escholarship.org/uc/item/03c2z9tr.

Haynie, Hannah J. 2014. Deep relationships among California languages. *Diachronica* 31(3): 407–447. doi:10.1075/dia.31.3.04hay.

Jacobsen, Jr., William H. 1958. Washo and Karok: an approach to comparative Hokan. *International Journal of American Linguistics* 24(3): 195–212. doi:10.1086/464456.

Jacobsen, Jr., William H. 1979. Hokan inter-branch comparisons. In: Lyle Campbell, Marianne Mithun (eds.). *The languages of Native America: historical and comparative assessment*: 545–591. Austin / London: University of Texas Press. doi:10.7560/746244-011.

Jany, Carmen. 2013. Hokan languages. In: *Oxford Bibliographies Online: Linguistics*. Oxford University Press. doi:10.1093/obo/9780199772810-0113.

Jany, Carmen. 2017. The Northern Hokan area. In: Michael Fortescue, Marianne Mithun, Nicholas Evans (eds.). *The Oxford Handbook of Polysynthesis*: 260–283. 1st edn. Oxford University Press. doi:10.1093/oxfordhb/9780199683208.013.34.

Kaufman, Terrence. 1989. A research program for reconstructing Proto-Hokan: first gropings. In: *Papers from the 1988 Hokan-Penutian Languages Workshop, Held at the University of Oregon, Eugene, Oregon, June 16–18, 1988*: 50–168. Eugene, OR: Department of Linguistics, University of Oregon.

Kaufman, Terrence. 2015. Some hypotheses regarding Proto-Hokan grammar. Unpublished manuscript. Available online at: https://www.researchgate.net/publication/340925991_Some_hypotheses_regarding_proto-Hokan_grammar_Kaufman_2015.

Kaufman, Terrence. 2016. Tlapaneko-Sutiaba, OtoMangean, and Hokan: where Greenberg went wrong. Unpublished manuscript. Available online at: https://www.researchgate.net/publication/340925787_subtitle_Tlapaneko-Sutiaba_OtoMangean_and_Hokan_where_Greenberg_went_wrong.

Kroeber, Alfred L. 1904. Languages of the coast of California south of San Francisco. *University of California publications: American archaeology and ethnology* 2(2): 29–80.

Kroeber, Alfred L. 1915. Serian, Tequistlatecan, and Hokan. *University of California Publications in American Archaeology and Ethnology* 11(4): 279–290.

Langdon, Margaret. 1974. *Comparative Hokan-Coahuiltecan studies: A survey and appraisal*. The Hague: Mouton. doi:10.1515/9783110887839.

Langdon, Margaret. 1979. Some thoughts on Hokan with particular reference to Pomoan and Yuman. In: Lyle Campbell, Marianne Mithun (eds.). *The languages of Native America: historical and comparative assessment*: 592–649. Austin / London: University of Texas Press. doi:10.7560/746244-012.

Langdon, Margaret. 1990. Morphosyntax and problems of reconstruction in Yuman and Hokan. In: Philip Baldi (ed.). *Linguistic change and reconstruction methodology* (Trends in Linguistics Studies and Monographs 45): 57–72. Berlin / New York: Mouton de Gruyter. doi:10.1515/9783110886092.57.

Langdon, Margaret & Pamela Munro. 1980. Yuman numerals. In: Kathryn Klar, Margaret Langdon, Shirley Silver (eds.). *American Indian and Indoeuropean studies: papers in honor of Madison S. Beeler* (Trends in Linguistics Studies and Monographs 16): 121–135. The Hague / Paris / New York: Mouton Publishers. doi:10.1515/9783110808681.121.

Leščiner, D. R. 1989. Rekonstrukcija nazvanij častej tela v jazykax sem'i xoka. In: *Lingvističeskaja rekonstrukcija i drevnejšaja istorija Vostoka. Materialy k diskussijam na Meždunarodnoj konferencii (Moskva, 29 maja — 2 ijunja 1989 g.). Čast' 1*: 159–166. Moskva: Izdatel'stvo "Nauka."

Leščiner, D. R. 1990. Rekonstrukcija tonov i dolgot v prajazyke sem'i xoka. In: *Sravnitel'no-istoričeskoe jazykoznanie na sovremennom ètape. Pamjati V. M. Illič-Svityča. 6-9 fevralja 1990 g. Tezisy dokladov*: 34–36. Moskva: Institut slavjanovedenija i balkanistiki AN SSSR.

Leshchiner, Dmitry & Sergei Nikolaev. 1992. Proto-Hokan roots. In: Vitaly Shevoroshkin (ed.). *Nostratic, Dene-Caucasian, Austric and Amerind: materials from the First International Interdisciplinary Symposium on Language and Prehistory, Ann Arbor, 8.-12. November 1988* (Bochum Publications in Evolutionary Cultural Semiotics 33): 366–402. Bochum: Universitätsverlag Dr. Norbert Brockmeyer.

Marlett, Stephen A. 1981. *The structure of Seri*. La Jolla CA: University of California at San Diego Ph.D. Available online at: http://lengamer.org/admin/language_folders/seri/user_uploaded_files/links/File/Marlett_1981_Seri_Dissertation.zip.

Marlett, Stephen A. 2007. Las relaciones entre lenguas "hokanas" en México: ¿Cuál es la evidencia? In: Cristina Buenrostro, Samuel Herrera Castro, Yolanda Lastra, Fernando Nava, Juan José Rendón, Otto Schumann, Leopoldo Valiñas, María Aydeé Vargas Monroy (eds.). *Clasificación de las lenguas indígenas de México: Memorias del III Coloquio Internacional de Lingüística Mauricio Swadesh*: 165–192. Mexico City: Universidad Nacional Autónoma de México & Instituto de Investigaciones Antropológicas & Instituto Nacional de Lenguas Indígenas.

Marlett, Stephen A. 2008. The Seri–Salinan connection revisited. *International Journal of American Linguistics* 74(3): 393–399. doi:10.1086/590087.

McLendon, Sally. 1964. Northern Hokan (B) and (C): a comparison of Eastern Pomo and Yana. In: William Bright (ed.). *Studies in Californian linguistics* (University of California Publications in Linguistics 34): 126–144. Berkeley / Los Angeles: University of California Press.

McLendon, Sally. 1973. *Proto Pomo* (University of California Publications in Linguistics 71). Berkeley: University of California Press.

McLendon, Sally. 1975. *A grammar of Eastern Pomo* (University of California Publications in Linguistics 74). Berkeley / Los Angeles / London: University of California Press.

Miller, Amy. 2001. *A grammar of Jamul Tiipay* (Mouton Grammar Library 23). Berlin / New York: Mouton de Gruyter. doi:10.1515/9783110864823.

Miller, Amy. 2018. Phonological developments In Delta-California Yuman. *International Journal of American Linguistics* 84(3): 383–433. doi:10.1086/697588.

Mithun, Marianne. 2007. Grammar, contact and time. *Journal of Language Contact* 1(1): 144–167. doi:10.1163/000000007792548378.

Mixco, Mauricio J. 1997. Haas's Hokan: dead end, or gateway to the future? *Anthropological Linguistics* 39(4): 680–694.

Moser, Mary Beck & Stephen A. Marlett. 2010. *Comcaac quih yaza quih hant ihiip hac: cmiique iitom — cocsar iitom — maricaana iitom = Diccionario seri — español — inglés: con índices español — seri, inglés — seri*. 2nd edn. Hermosillo, Son.: Editorial UniSon : Plaza y Valdés Editores.

Nevin, Bruce. 1998. *Aspects of Pit River phonology*. University of Pennsylvania PhD Thesis.

Nevin, Bruce. 2019. Why Proto-Palaihnihan is neither. Unpublished manuscript. doi:10.13140/RG.2.2.33917.20966.

Nevin, Bruce. 2020. Achumawi Dictionary. Available online at: https://www.webonary.org/odissi/. [Accessed 13 February, 2024].

O'Connor, Loretta. 2013. *Latyaygi — English — Español: A trilingual dictionary of Lowland Chontal of Oaxaca* (Languages of the World/Dictionaries 48). Muenchen: LINCOM GmbH.

Olmsted, D. L. 1956. Palaihnihan and Shasta I: labial stops. *Language* 32(1): 73–77. doi:10.2307/410654.

Olmsted, D. L. 1957. Palaihnihan and Shasta II: apical stops. *Language* 33(2): 136–138. doi:10.2307/410725.

Olmsted, D. L. 1959. Palaihnihan and Shasta III: dorsal stops. *Language* 35(4): 637–644. doi:10.2307/410603.

Olmsted, D. L. 1964. *A history of Palaihnihan phonology* (University of California Publications in Linguistics 35). Berkeley / Los Angeles: University of California Press.

Olmsted, D. L. 1966. *Achumawi dictionary* (University of California Publications in Linguistics 45). Berkeley / Los Angeles: University of California Press.

Oswalt, Robert L. 1977. The Hokan word for "tongue": the Pomo evidence. In: James E. Redden (ed.). *Proceedings of the 1976 Hokan-Yuman languages workshop, held at University of California, San Diego, June 21-23, 1976* (University Museum Studies 11): 69–78. Carbondale, Illinois: University Museum and Art Galleries, Southern Illinois University.

Oswalt, Robert L. 1981. *Southern Pomo word list and map of native place names in the Warm Springs area*. San Francisco: U.S. Army Corps of Engineers.

Peiros, Ilia. 1997. Macro families: can a mistake be detected? In: Irén Hegedűs, Peter A. Michalove, Alexis Manaster Ramer (eds.). *Indo-European, Nostratic, and Beyond: Festschrift for Vitalij V. Shevoroshkin* (JIES Monograph 22): 265–292. Washington DC.

Poser, William J. 1992. The Salinan and Yurumanguí data in Language in the Americas. *International Journal of American Linguistics* 58(2): 202–229. doi:10.1086/ijal.58.2.3519756.

Poser, William J. 1995. Binary comparison and the history of Hokan comparative studies. *International Journal of American Linguistics* 61(1): 135–144. doi:10.1086/466247.

Rivet, Paul. 1942. Un dialecte hoka colombien: le Yurumangí. *Journal de la Société des Américanistes* 34: 1–59. doi:10.3406/jsa.1942.2334.

Sapir, Edward. 1913. Southern Paiute and Nahuatl, a study in Uto-Aztekan. *Journal de la Société des Américanistes* 10(2): 379–425. doi:10.3406/jsa.1913.2866.

Sapir, Edward. 1917. The position of Yana in the Hokan stock. *University of California publications in American archaeology and ethnology* 13(1): 1–34.

Sapir, Edward. 1919. Southern Paiute and Nahuatl, a study in Uto-Aztekan, Part II. *Journal de la Société des Américanistes* 11(1): 443–488. doi:10.3406/jsa.1919.3856.

Sapir, Edward. 1920a. A note on the first person plural in Chimariko. *International Journal of American Linguistics* 1(4): 291–294. doi:10.1086/463727.

Sapir, Edward. 1920b. The Hokan and Coahuiltecan languages. *International Journal of American Linguistics* 1(4): 280–290. doi:10.1086/463726.

Sapir, Edward. 1920c. The language of the Salinan Indians. *International Journal of American Linguistics* 1(4): 305–309. doi:10.1086/463730.

Sapir, Edward. 1921. A supplementary note on Salinan and Washo. *International Journal of American Linguistics* 2(1/2): 68–72. doi:10.1086/463735.

Sapir, Edward. 1925. The Hokan affinity of Subtiaba in Nicaragua. *American Anthropologist* 27(3–4): 402–435, 491–527. doi:10.1525/aa.1925.27.3.02a00040.

Sapir, Edward & Morris Swadesh. 1960. *Yana dictionary* (University of California Publications in Linguistics 22). Berkeley / Los Angeles: University of California Press.

Shaul, David Leedom. 2019. *Esselen studies: language, culture, and prehistory* (LINCOM Studies in Native American Linguistics 83). Muenchen: Lincom GmbH.

Shaul, David Leedom. 2020a. *Salinan and English lexicon* (Languages of the World Dictionaries 74). Muenchen: LINCOM GmbH.

Shaul, David Leedom. 2020b. *Salinan language studies* (LINCOM Studies in Native American Linguistics 85). Muenchen: LINCOM GmbH.

Silver, Shirley. 1964. Shasta and Karok: a binary comparison. In: William Bright (ed.). *Studies in Californian linguistics* (University of California Publications in Linguistics 34): 170–181. Berkeley / Los Angeles: University of California Press.

Silver, Shirley. 1974. Some Northern Hokan plant-tree-bush forms. *The Journal of California Anthropology* 1(1): 102–109.

Silver, Shirley. 1976. Comparative Hokan and the Northern Hokan languages. In: Margaret Langdon, Shirley Silver (eds.). *Hokan Studies: Papers from the First Conference on Hokan Languages held in San Diego, California April 23-25, 1970* (Janua Linguarum: Series Practica 181): 193–202. The Hague / Paris: Mouton. doi:10.1515/9783110819113-013.

Swanton, John R. 1915. Linguistic position of the tribes of Southern Texas and Northeastern Mexico. *American Anthropologist* 17(1): 17–40. doi:10.1525/aa.1915.17.1.02a00030.

The Santa Ynez Band of Chumash Indians. 2007. *Samala-English dictionary: A guide to the Samala language of the Ineseño Chumash people*.

Turner, Paul R. 1967. Seri and Chontal (Tequistlateco). *International Journal of American Linguistics* 33(3): 235–239. doi:10.1086/464966.

Turner, Paul R. 1976. Pluralization of nouns in Seri and Chontal. In: Margaret Langdon, Shirley Silver (eds.). *Hokan Studies: Papers from the First Conference on Hokan Languages held in San Diego, California April 23-25, 1970* (Janua Linguarum: Series Practica 181): 297–303. The Hague / Paris: Mouton. doi:10.1515/9783110819113-019.

Turner, Paul R. & Shirley Turner. 1971. *Dictionary: Chontal to Spanish-English, Spanish to Chontal*. Tucson, Arizona: The University of Arizona Press.

Waterhouse, Viola Grace. 1976. Another look at Chontal and Hokan. In: Margaret Langdon, Shirley Silver (eds.). *Hokan Studies: Papers from the First Conference on Hokan Languages held in San Diego, California April 23-25, 1970* (Janua Linguarum: Series Practica 181): 325–343. The Hague / Paris: Mouton. doi:10.1515/9783110819113-021.

Webb, Nancy M. 1979. The semantic domain "tree": Hokan lexical evidence. In: James E. Redden (ed.). *Proceedings of the 1978 Hokan languages workshop, held at the University of California, San Diego, June 27-29, 1978* (Occasional Papers on Linguistics 5): 1–7. Carbondale, Illinois: Department of Linguistics, Southern Illinois University.

Webb, Nancy M. 1980. Esselen-Hokan relationships. In: James E. Redden (ed.). *Proceedings of the 1979 Hokan languages workshop, held at the University of California, Los Angeles, June 26-28, 1979* (Occasional Papers on Linguistics 7): 72–77. Carbondale, Illinois: Department of Linguistics, Southern Illinois University.

Webb, Nancy M. 1981. Some notes on Hokan person terms. In: James E. Redden (ed.). *Proceedings of the 1980 Hokan languages workshop, held at the University of California, Berkeley, June 30-July 2, 1980* (Occasional Papers on Linguistics 9): 78–85. Carbondale, Illinois: Department of Linguistics, Southern Illinois University.

Zhivlov, Mikhail. 2018. Some morphological parallels between Hokan languages. *Journal of Language Relationship / Вопросы языкового родства* 16(2): 138–161.

М. А. Живлов. Хока I: обзор исследований по сравнению хоканских языков

В статье рассматривается история хоканской гипотезы, от её формулировки Диксоном и Крёбером в 1912-1913 годах до наших дней. Несмотря на более чем сто лет исследований, до сих пор нет единого мнения относительно обоснованности хоканской гипотезы. В статье утверждается, что одной из причин этого является тот факт, что многие попытки сравнения хоканских языков использовали нестандартную методологию, в которой исследование останавливается на перечислении наблюдаемых фонетических соответствий, вместо классического сравнительного метода, требующего поиска дополнительных распределений и достаточно полной реконструкции фонологии праязыка.

Ключевые слова: сравнительный метод; языки хока; индейские языки; генеалогическая классификация языков.