

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

№ 23, issue 3—4



Moscow 2025

Российский государственный гуманитарный университет
Институт языкознания Российской Академии наук

Вопросы языкового родства

Международный научный журнал

№ 23, часть 3—4

ВЯР

Москва 2025

Advisory Board:

William H. BAXTER (Ann Arbor, Michigan)

Václav BLAŽEK (Brno)

Larry M. HYMAN (Berkeley)

Frits KORTLANDT (Leiden)

Alexander LUBOTSKY (Leiden)

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

Редакционный совет:

В. БЛАЖЕК (Брно)
У. БЭКСТЕР (Анн Арбор)
В. Ф. ВЫДРИН (Париж)
Ф. КОРТЛАНДТ (Лейден)
А. ЛУБОЦКИЙ (Лейден)
А. Ю. МИЛИТАРЕВ (Москва)
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

Вопросы языкового родства: Международный научный журнал / Рос. гос. гуманитар. ун-т; Рос. акад. наук. Ин-т языкознания; под ред. Г. С. Старостина. — М., 2025. — № 23/3—4. — x + 217 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, 2025. — № 23/3—4. — x + 217 p.

ISSN 2219-3820

[http://www.jolr.ru/
gstarst1@gmail.com](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, г. Москва, вн. тер. г. муниципальный округ Тверской,
Миусская пл., д. 6, стр. 6

Свидетельство о регистрации СМИ
ПИ № ФС77-73409 от 03 августа 2018 г.
Периодическое печатное издание

Подписано в печать 19.02.2026. Выход в свет 27.02.2026
Формат 60×90/8. Уч.-изд. л. 16,5. Усл. печ. л. 28,5
Тираж 1050 экз. Свободная цена
Заказ № 2320

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

Table of Contents / Содержание

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

Issue 3 / Часть 3

Articles / Статьи

<i>Svetlana Burlak</i> . Sources of lexical replacement and synonymy in Swadesh wordlists for proto-languages	189
[С. А. Бурлак. О возможных источниках лексических замен и синонимии в праязыковых списках базисной лексики]	
<i>Elena Shvedova, Yuri Koryakov, Elizaveta Zabelina</i> . Quantifying lexical distances among Nudiz, Mahmudi, and Verin Dvin Urmi (North-Eastern Neo-Aramaic)	207
[Е. Е. Шведова, Ю. Б. Коряков, Е. А. Забелина. Количественная оценка лексических различий между нудезнайским, мамедайским и верин-двинским урмийским (северо-восточные новоарамейские языки)]	
<i>Vera Tsukanova, Ilya Yakubovich</i> . Hurrian and Ugaritic sibilants: Essay of phonetic interpretation	276
[В. Л. Цуканова, И. С. Якубович. К фонетической интерпретации хурритских и угаритских сибилантов]	

Issue 4 / Часть 4

Articles / Статьи

<i>Andrey Nikulin</i> . A phonological reconstruction of Proto-Dâw–Hup	307
[А. В. Никулин. Фонологическая реконструкция праязыка дыу–хупской группы]	
<i>Mikhail Zhivlov</i> . Hoka III: Velar obstruents and glottal stop	344
[М. А. Живлов. Хока III: велярные шумные и гортанная смычка]	
<i>Anastasia Krylova</i> . Relative chronology of a series of phonetic changes in Old Indo-Aryan clusters based on data from the Kullui language	375
[А. С. Крылова. Относительная хронология ряда фонетических изменений древнеиндийских кластеров на материале языка куллуи]	
<i>М. Е. Васильев</i> . Оценка достоверности исходных данных при лексикостатистических расчётах	389
[Mikhail Vasilyev. Evaluating the reliability of source data in lexicostatistical analysis]	

Contributors

Svetlana Burlak — doctor of sciences (Philology), senior researcher, Institute of Oriental Studies of the Russian Academy of Sciences (Moscow); svetlana.burlak@bk.ru

Yuri Koryakov — candidate of sciences (Philology), senior researcher, Linguistic Convergence Laboratory, HSE University (Moscow); senior researcher, Institute of Linguistics of the Russian Academy of Sciences (Moscow); ybkoryakov@gmail.com

Anastasiya Krylova — candidate of sciences (Philology), researcher, Department of Asian and African languages, Institute of Oriental Studies of the Russian Academy of Sciences (Moscow); krylova_anastasi@bk.ru

Andrey Nikulin — Ph.D., faculty member, Instituto de Formação Superior Indígena Takinahakỹ, Federal University of Goiás; nikulin@ufg.br

Elena Shvedova — research assistant, Linguistic Convergence Laboratory, HSE University (Moscow); senior Lecturer, School of Linguistics, HSE University (Moscow); shvedovalena98@gmail.com

Vera Tsukanova — Ph.D., research fellow, Philipps-Universität Marburg; vera.tsukanova@uni-marburg.de

Mikhail Vasilyev — researcher, Institute of Oriental Studies of the Russian Academy of Sciences (Moscow); mvhumanity@gmail.com

Ilya Yakubovich — doctor of sciences (Philology), research fellow, Philipps-Universität Marburg; ilya.yakubovich@uni-marburg.de

Elizaveta Zabelina — research assistant, Linguistic Convergence Laboratory, HSE University (Moscow); junior research fellow, Institute for Linguistic Studies of the Russian Academy of Sciences (Saint-Petersburg); windmill161@gmail.com

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

Бурлак, Светлана Анатольевна — доктор филол. наук, профессор РАН, ведущий науч. сотрудник Института востоковедения РАН (Москва); svetlana.burlak@bk.ru

Васильев, Михаил Евгеньевич — науч. сотрудник Института востоковедения РАН (Москва); mvhumanity@gmail.com

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

Забелина, Елизавета — стажер-исследователь Международной лаборатории языковой конвергенции НИУ ВШЭ (Москва); младший научный сотрудник Института лингвистических исследований РАН (Санкт-Петербург); windmill161@gmail.com

Коряков, Юрий — канд. филол. наук, старший научный сотрудник Международной лаборатории языковой конвергенции НИУ ВШЭ (Москва); старший научный

сотрудник Института языкознания РАН (Москва); ybkoryakov@gmail.com

Крылова, Анастасия Сергеевна — канд. филол. наук, научный сотрудник Института востоковедения РАН (Москва); krylova_anastasi@bk.ru

Никулин, Андрей Владимирович — Ph.D., научный сотрудник федерального университета штата Гояс; nikulin@ufg.br

Цуканова, Вера Львовна — Ph.D., научный сотрудник Марбургского университета; vera.tsukanova@uni-marburg.de

Шведова, Елена — стажер-исследователь Международной лаборатории языковой конвергенции НИУ ВШЭ (Москва); старший преподаватель Школы лингвистики НИУ ВШЭ (Москва); shvedovalena98@gmail.com

Якубович, Илья Сергеевич — доктор филол. наук, научный сотрудник Марбургского университета; ilya.yakubovich@uni-marburg.de

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 jolr.rsuh@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, bd. 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 (jolr.rsuh@gmail.com). По различным вопросам с редакцией журнала можно также связаться по обычной почте:

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

Issue 1 / Часть 1

Articles / Статьи

Sources of lexical replacement and synonymy in Swadesh wordlists for proto-languages

In the practice of step-by-step reconstruction, situations are encountered where more than one word competes for the same slot in the protolanguage Swadesh-wordlist, and there is no way to dismiss one of the potential synonyms using standard methods. At the same time, in real languages it is almost always possible to choose which lexeme is the most suitable one for a particular slot in the wordlist. The difficulties in making such justifications for proto-languages may be related to both linguistic and sociolinguistic factors. Since we do not know the social landscape of the proto-language, as well as its evolution in descendant populations, there may not be enough material to draw conclusions about whether a word that was restricted to certain social situations in the proto-language (for example, used only by one gender, or only in relation to those higher or lower in the hierarchy, or only in conversations with children, etc.) could have become the primary term in one of the descendant branches after a change in the social structure. The practice of taboo leads to an increase in the number of synonyms in the language: at any given moment, one word is primary, but this can vary over time, so some descendant populations may inherit one word as primary, while others inherit a different one. Additionally, in descendant branches, phonetic or morphological alignments could have diverged in different directions, causing bases that once formed a single (possibly suppletive) paradigm to become the only representatives of the corresponding meaning in each of the descendant branches. The more time has passed since the separation of the proto-language, the harder it is to identify the alternations that caused the initial differences between the observable forms, so that they look apparently unrelated to each other. Perhaps the main difficulty in identifying the primary word lies in the fact that the proto-language may have had a differently structured nominative grid, so that contexts which appear homogeneous to the present observer were only considered so during the existence of the descendant languages, with some of them adopting one word as primary and others selecting a different one.

Keywords: Swadesh wordlist; synonymy; lexical replacement; comparative method; semantic reconstruction; language relationship; lexicostatistics.

In the practice of step-by-step reconstruction, situations are encountered where more than one word competes for the same slot in the Swadesh wordlist for the reconstructed protolanguage, and there is no way to dismiss one of the potential synonyms as derived, borrowed, areally restricted, or having undergone a more likely semantic shift. Nevertheless, in actually observed (currently existing and anciently documented) languages, it is almost always possible to justify the choice of a word for inclusion in the list by dismissing rarer or stylistically marked variants. In accordance with the principle of uniformitarianism, a similar situation should be assumed for proto-languages as well. Therefore, within the research conducted by the Moscow School of comparative linguistics, the task is set, first, to determine which of the potential synonyms was the primary one in the proto-language (Kassian et al. 2010), and second, to develop methods that should allow this to be done in all possible cases (see, for example, Starostin 2010: 100–110).

In fact, the only case of acceptable synonymy in a proto-list is so-called transitional synonymy, when “an older word is gradually being ‘ushered out’ by a more recent replacement” (Kassian et al. 2010: 48). In such a process, there is a period during which they are approximately equal in overall frequency (see Fig. 1), and different idiolects prefer different variants for the same contexts. However, at any given time, such cases are few (at most, 1–2 per 100-wordlist), so they do not significantly affect statistical calculations. Yet this situation highlights the following aspect of the innate human language capacity: humans are adapted to effortlessly acquire several different names for the same element of environmental *realia*. This adaptation was beneficial for small hunter-gatherer groups who, on one hand, communicated more frequently within the group than outside it (which facilitated language divergence through the mechanism of clustering human social networks, see Jackson 2019; Milroy & Gordon 2008), and on the other hand, were forced to communicate with representatives of neighboring groups for genetic exchange (the minimum viable population size for large primates is at least a thousand individuals; see, e.g., Frankham et al. 2014; Harcourt 2002), while the size of a human group sustaining itself by hunting and gathering is limited by environmental capacity and averages 25 people (Khrisanfova & Perevozchikov 2005: 108).

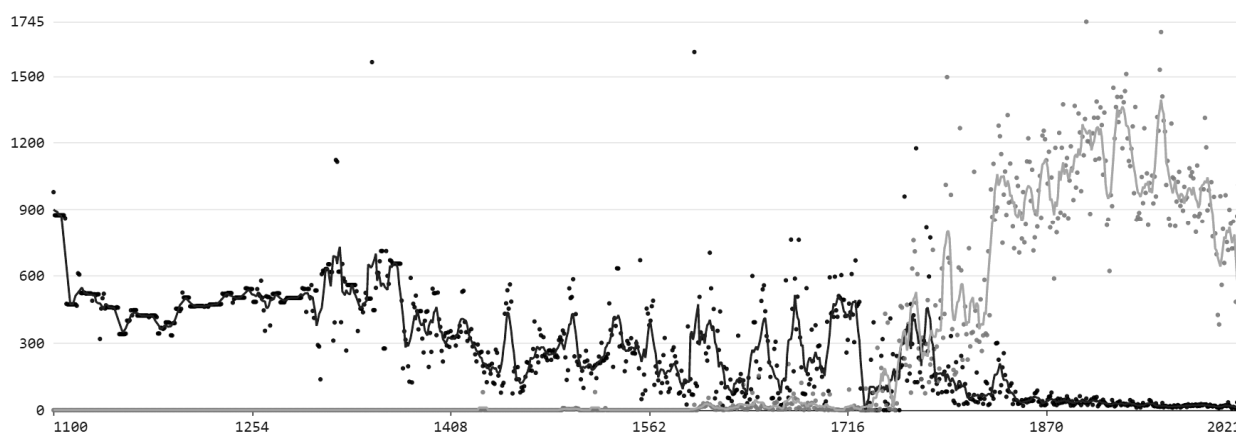


Fig. 1. Frequency of the lexemes *glaz* (grey) and *oko* (black) ‘eye’ in the Russian language of 12th–21st centuries AD (according to Russian National Corpus, see RNC). It can be seen that during the 18th century a lexical replacement takes place: the frequency of the word *glaz* grows rapidly while the frequency of the word *oko* decreases.

Nevertheless, even when collecting data from currently existing languages where direct elicitation from speakers is possible, many words on the list reveal insurmountable or poorly resolvable synonymy. In the case of a proto-language, where such opportunities do not exist, identifying the primary synonym can be challenging.

Difficulties in identifying the primary word filling the corresponding slot in the protolanguage wordlist may be related to both intra-linguistic and sociolinguistic reasons.

From a sociolinguistic perspective, any language can be represented by different social groups of speakers (by gender, age, occupation, etc.), as well as by different communication situations — and in all these cases, different names may be used for the same element of reality. For example, in modern Japanese, the idea of ‘to give’ is expressed by the verb 下さる *ku-dasaru*, if the giver is of higher status than the recipient, and by 差し上げる *sashiageru* (< originally ‘to raise’) in the opposite case. In Classical Ancient Chinese (see Starostin 2019: 159), the meaning ‘to die’ was expressed by the verb 薨 *hōng*, when referring to a high-ranking official, and by the verb 死 *sǐ* in other cases; the meaning ‘to kill’ was expressed by the verb 弑 *shì*, when referring to someone of higher hierarchical status (a prince, a father, etc.), and by the verb 殺 *shā* in other cases. Sometimes texts or recordings from different consultants allow one

to establish the most frequent and stylistically neutral variant (see, for example, the detailed philological analysis in Starostin 2019), but it is quite possible that compilers of dictionaries (especially older ones) may have overlooked such distinctions, and the inability to choose a single variant can affect reconstruction.

The influence of social hierarchy is particularly evident in personal pronouns: for example, in Japanese, different speakers in different situations use as the 1st person singular pronoun the words 私 *watashi* (in female and formal speech) / *watakushi* (in more formal speech), 僕 *boku* ((in male informal speech), 俺 *ore* (also in men's speech, in even more informal situations), etc. In Kassian et al. 2010: 65, 83, 85 it is recommended to exclude any polite forms when collecting pronouns — and this is feasible when collecting field data from a living language or during the philological analysis of an ancient language represented by an extensive text corpus. However, over the course of language development, polite forms can displace those that were originally neutral. For example, in standard Khmer, the word *khnom* (etymologically 'slave') is a neutral 1Sg pronoun, while *ʔaŋ* (the original 1Sg word for 'I') is used as impolite or "as neutral in situations where expressing politeness is not required" (Pogibenko 2013: 30) (and therefore, accordingly, should not be included in the 100-wordlist).

The diversity of pronominal roots is not restricted to politeness forms. Thus, in the Kulan dialect of the Onya Darat Austronesian language (Borneo), spoken by relatively small communities, the choice of pronoun depends on how the generation of the speaker and the generation of the listener relate: for example, when speaking to one's child (or nephew), a person calls themselves *maaq*, and their interlocutor *omo*; if a person is speaking to their brothers, sisters, or parents (as well as their brothers or sisters), they call themselves *oko*, and the interlocutor *omo* (if of the same generation) or *okam* (if of an older generation, and it is the generation, not age, that matters; see Tadmor 2015: 84–85).

Even if some of these variants are accepted as the most neutral in some synchronic state of the language, it is quite possible that with a change in lifestyle, everything can change: for example, when speakers of this dialect move to cities where they have to communicate with persons whose generation they do not know, they develop some system that is different from the original one (Tadmor 2015: 96) — and in different places, such changes can proceed in different ways. Changes in lifestyle are also possible in non-urban cultures living by hunting and gathering or primitive agriculture (see, e.g., Henrich 2016: 174–175).

In the languages of Southeast Asia, pronouns are used less frequently in principle — if a participant of a communication act can be designated by kinship, seniority, title, clan affiliation, etc., the corresponding noun will be used in the sentence (Tadmor 2015: 79). Due to such practices the probability of borrowing personal pronouns increases (which would be undetectable if the source language becomes extinct), and so does the probability that some pronouns may be replaced by nouns.

Another sociolinguistic reason for synonymy is the difference between male and female speech. For example, in Japanese (Holmes & Wilson 2022: 224), women's main word for 'to eat' is 食べる *taberu*, while men's word is 喰う *kuu*; the meaning 'stomach' is expressed as お腹 *onaka* by women and as 腹 *hara* by men (although the word 'stomach' does not belong to the 100-wordlist, frequent patterns of semantic shift imply that it can always become the main designation for 'belly'). In the Kokama-Kokamilla language (Tupi family, South America), men would say *ra* '(s)he', *raepe* 'then', *ramua* 'other', while women would use words *ya* '(s)he', *yaepe* 'then', *yamua* 'other' (Vallejos-Yopán 2015).

Male and female variants may often differ phonetically. For example, in Bengali, in some words, women pronounce *l*, while men pronounce *n* (Holmes & Wilson 2022: 224); in Chukchi, male variants with *r*, *č*, and *rk*, *rg*, would until recently be respectively pronounced by women

as *c* and *cc*, e.g., the word ‘walrus’ in male speech was *rərka*, and in female speech *cəcca*, etc. (Skorik 1961: 33); this state of affairs may be caused by men often marrying representatives of a neighboring dialect or a closely related language (Dunn 1999: 27 with lit.).

In a number of cases, the male and female variants differ morphologically. For example, in the Yana language (N. America), men add suffixes, so the meaning ‘person’ is expressed by the form *yaa* in female speech and the form *yaana* in male speech; the meaning ‘fire’ is expressed by the forms *?au* and *?auna* respectively (Holmes & Wilson 2022: 224). In the Yanyuwa language (Australia; Holmes & Wilson 2022: 224), men use different class prefixes than women, e.g., for ‘meat’ women use the form *ni-warnnyi*, and men *na-warnnyi*. In the Lakota language (White Hat 1999: 18), men and women use different enclitics, e.g., the meaning ‘this is good’ in female speech is rendered as *wašte kšto*, and in male speech as *wašte yelo*.

In male and female speech, the same language elements can occur with different frequencies: for example, in Canadian French, both women and men sometimes drop *l* in the constructions *il y a* and *il fait* (as in *il fait beau* ‘the weather is good’, *il fait froid* ‘it is cold’ etc.), but men do it more often (Holmes & Wilson 2022: 227).

It is clear that in a developed literary language with a large text corpus, careful philological analysis will allow the most neutral variant to be identified; but if such a practice exists in a language used exclusively in everyday communication situations (and there is no reason why such a language could not later become a proto-language for some group), choosing a more neutral variant from male and female or higher/lower one may be difficult (moreover, when working with a limited number of consultants, only one of the existing variants may be included in the dictionary).

Synonyms can originate from so-called “child-directed speech” (or “nursery language”, which is “a language subsystem considered convenient to talk to small children” (Yeliseyeva et al. 2017: 7). Words from this register can displace the original words, surpassing them in frequency, cf., for example, the frequencies of the words *mama* and *mat* ‘mother’ in modern Russian (according to the Russian National Corpus).

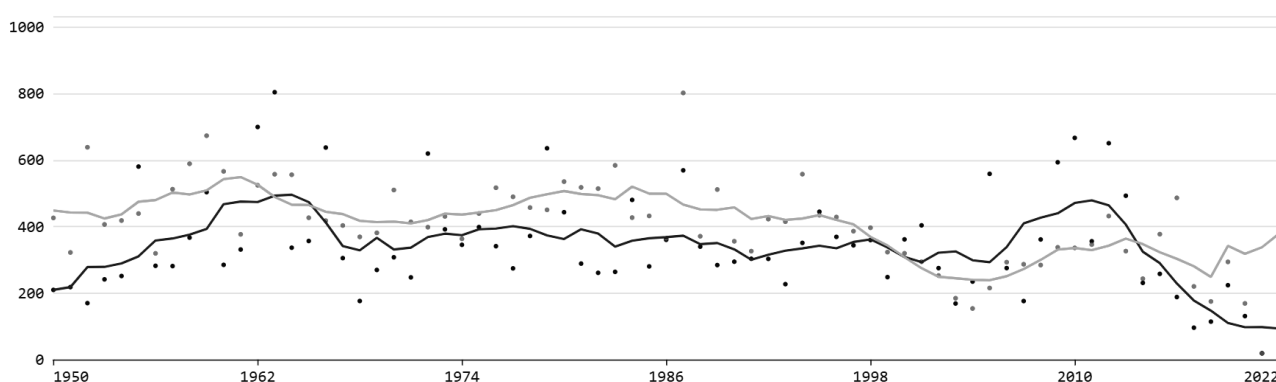


Fig. 2. The frequency of the lexemes *mama* (black) and *mat* (grey) ‘mother’ in the Russian language in 20th–21st centuries (according to Russian National Corpus).

It can be seen that at certain time intervals the word *mama* is used more frequently than *mat*. (Note that, although the word ‘mother’ does not belong to the Swadesh 100-wordlist, it can serve as a source for the word meaning ‘woman’, cf. Lith. *móteris* ‘woman’ < PIE **méh₂tēr* ‘mother’). Likewise, the word for ‘breast’ also can originate from the child-directed speech (Starostin 2010: 93).

Nursery forms are not obliged to be phonetically similar to the corresponding forms in standard adult language (cf., e.g. Rus. *mashina* and *bibika* ‘car’, *jest* and *am-am* ‘to eat’, Engl.

tummy ‘belly’). Since such subsystems exist in many languages, it is probable that they existed also in proto-language time and could supply the sources for lexical replacement.

If a culture has a developed practice of tabooing, for many meanings there can be sets of complete synonyms, allowing one to freely replace one with another if necessary. For example, in the Misima language (Misima Island, Papua New Guinea), where the names of all deceased, as well as words similar to them, are tabooed, there are five words meaning ‘fire’ (Simons 1982: 203). At the same time, different lexemes can act as the most frequent and neutral words at different times (and for different speakers). P. Black notes (Black 1997: 57) that many lexemes of various Australian languages, which in the 1930s had been cited as examples of tabooed words, were recorded in dictionaries of the 1980s as being in active use. On Tahiti island at the end of the 17th century, the word *pō* ‘night’ (< Proto-East-Polynesian **pō*) became tabooed because it was part of the dynastic name of the island’s rulers, Pomare, and the word *ru’i* (originally ‘dark, blind’) began to be used instead, but in the 19th century, some speakers switched back to use *pō*, while others continued to use *ru’i* (Belikov 2006: 24–25).

Theoretically, there is a possibility that the Proto-Indo-European lexemes **péh₂wr* and **h₁ng^wnís* ‘fire’ could have been such synonyms, but then in some descendant languages one of these words became the main designation for fire, while the other one chose a different option (thus, if only the Italic and Indo-Iranian branches of this family had survived, the lexeme **h₁ng^wnís* would be recognized as the main Indo-European word for ‘fire’, since it exists in both branches). Note, however, that such hypotheses should not be accepted without considerable proof (see below).

Furthermore, different speakers may treat different synonyms as stylistically neutral variants, and when a population splits into two, the percentage of those who consider one or another variant stylistically neutral may be unequal in different parts; as a result, one daughter population most often will choose one synonym as stylistically neutral, while the other one will prefer the alternative. During reconstruction, it will then be impossible to determine which variant was more often chosen as stylistically neutral in the original population (and which, accordingly, should be considered the most stylistically neutral for the proto-language).

A word that was the most common at some point may over time move to the linguistic periphery and then return to the status of the main carrier of the corresponding semantics. Perhaps such was the situation with the Indo-European root **g^hel-*, expressing the meaning ‘yellow’ (> Skt. *hári-* ‘yellow, green’, Greek *χλωρός* ‘green, yellow, greenish-yellow’, Rus. *жёлтый* ‘yellow’). In Latin, a replacement occurred: instead of this root, the main designation for ‘yellow’ became the word *fulvus* (GLD 2025), while the derivate of the root **g^hel-* — *galbinus* ‘greenish-yellow’ — moved to the periphery; however, it is precisely this word (and not *fulvus*) which is inherited as the main designation of yellow color in the modern Romance languages (French *jaune*, Romanian *galben*), cf. Normanskaya 2005.

If a community has widespread diglossia (this may also happen in non-literate cultures) and one of the languages is used as the main means for everyday communication, it lacks a high register, and accordingly, words of the lower register become the most frequent. Such processes may trigger for certain words a shift to ‘stylistically neutral’ from (possibly) initially ‘pejorative’, cf. Tok Pisin *klok* ‘heart’ (Belikov 2006: 94) (< Engl. *clock*), Proto-Slavic **rĕkà* ‘hand’ (from a verb meaning ‘to gather’, cf. Lithuanian *riñkti*), Khmer (Surin dialect) *ηuap* ‘to die’ (cf. standard Khmer *ηiap* ‘to wither (of a plant), perish (of an animal), die (of a person, pejorative)’ (Pogibenko 2013: 30). If the idiom occupying the position of a higher register is closely related to the one used in everyday communication, synonymous lexemes can appear in the same contexts for the same speakers. Thus, in examples given in the dictionary of the Deulino village dialect of Russian, the lexemes *ženčina* ‘woman’ and *baba* ‘woman (colloquial/pejorative)’

appear in the same contexts, cf. *dodielistaja zenčina* and *dodielistaja baba* ‘shrewd woman’ from the same consultant within the same text (Ossovetskiy 1969: 145). When compiling the Swadesh wordlist for this dialect, the variant *zenčina* was chosen (GLD 2025), but there is a chance that a different selection of consultants for the dictionary (or different research attitudes of the list compiler) would lead to a different result.

In general, dialects are characterized by a situation where features of one’s own and neighboring idioms are distributed in speech with some non-zero frequency, which can increase and decrease for various reasons. The following situation, documented in one of the Vladimir-Volga dialects, is indicative: “In 2012, immediately after meeting, the first conversation with EM (over two hours) was recorded. The proportion of examples with *yokan’ye* (a phonological feature — S.B.) in this recording is small: 22% <...>, and the number of examples with *yokan’ye* noticeably increased towards the end of the conversation (apparently as EM’s communication with dialectologists became more relaxed). In this conversation, EM approached the first group of consultants in terms of *yokan’ye* — persons with non-peasant professions. In 2014, we visited EM twice; both times she received us very hospitably, as old acquaintances. Most of the first conversation was casual talk, and towards the end, EM was offered a questionnaire on the declension and stress of masculine nouns. In this conversation, the proportion of forms with *yokan’ye* turned out to be significant: more than 55%, i.e., “at the level” of women of her age with incomplete secondary education, farm workers (group 3). The next morning, the questionnaire survey was continued, but EM was already focused on problems of speech correctness and throughout the recording discussed with us the difference between dialect forms and standard ones. For her, this was a conversation on professional topics with fellow philologists. The percentage of forms with *yokan’ye* in this conversation turned out to be lower: 33% (42% in place of *e), i.e., between the data for the first and second groups of informants. Thus, EM partly controls *yokan’ye* in her speech: it becomes less prominent in a «semi-official» communication situation (with unfamiliar visitors from Moscow) and in a «professional» communication situation. The proportion of examples with *yokan’ye* in EM during more relaxed communication corresponds to the archaic variety of dialect usage. Possibly, other consultants with a low proportion of *yokan’ye* are capable, like EM, of controlling its use, however, we do not have data on this” (Dyachenko et al. 2018: 56).

If the language has a developed folklore tradition, it can preserve more archaic lexemes. For example, Tuamotu preserves an archaic style called *fangu* (the language of sacred chants), where many words that have been replaced in ordinary speech due to taboo are preserved: thus, the meanings ‘water’ and ‘to drink’ in ordinary Tuamotu are expressed by the word *komo*, while in *fangu* the original words *vai* ‘water’ and *inu* ‘to drink’ are preserved; for ‘big’, not the standard *tooreu* is used, but rather the original *nui*; for ‘night’, not the standard *ruki*, but the original *poo*, etc. (White 1968: 65). When a population splits, the stylistic attitudes of speakers can change — differently in different parts.

The listed situations are relevant for the case when a proto-language has split into two daughter branches, and one variant is chosen in one of them, and the other in the other one. But, generally speaking, such “semantic criss-crossing” (using the terminology of the Moscow School of comparative linguistics, see Starostin 2010: 103–104) can also be seen in cases when in each of the groups both variants are present. This may be a consequence of the law of homological series in hereditary variability, originally formulated by N. I. Vavilov for plants, but apparently relevant for any objects evolving naturally. This law states that “species and genera that are genetically close are characterized by similar series of hereditary variability with such regularity that, knowing the series of forms within one species, one can predict the finding of parallel forms in other species and genera” (Vavilov 1935).

If the language experienced several periods of dialect fragmentation and subsequent “koineization”, traces of such stylistic differences, as well as cases of transitional synonymy at different stages of its existence, may, after a significant time, look like proto-language synonymy. The matter can also be complicated by the fact that when borrowing from closely related idioms, phoneme-by-phoneme recalculation is often applied (Burlak & Starostin 2005: 63), which makes borrowed words indistinguishable from inherited ones and masks lexical replacements.

Synonyms for concepts frequently encountered in everyday life can sometimes be created literally out of nowhere. As was already noted by Igor Dyakonov (1984: 8), concepts perceived as particularly important for a particular culture, the ones that are often talked about, usually have many synonyms. One can give an example from modern Russian: cat lovers, who often have to talk about their pets (and who usually experience strong positive emotions towards them), call them not only by the common words *koshka* ‘cat’ and *kot* ‘tomcat’, but also by such words as *kosha*, *kotyara*, *kotik*, *kotejka*, *kotofej*, *kote*, *kotenitsa*, *kotyandra*, *koshundra*, *koshandra*, *koshendra*, *koshak*, *koshuk*, *koshakan*, *koshkatan*, *kosharik*, *koshara*, *koshanya*, *koshe*, *koshko*, *kotanych*, *kyso* (Iomdin 2023).

Such an abundance of variants apparently has a neurobiological basis: research shows that, if a word is repeated too often, it loses its meaning — both for the speaker and the listener: as experiments show, after about 30 repetitions of the same word in a row, it suddenly begins to seem meaningless (the “jamais vu” effect, as described in Moulin et al. 2021). This is not surprising: extinction of response to a too frequently repeated stimulus is characteristic of all living organisms (starting with unicellular ones). When producing and perceiving rare words, additional neural circuits are involved (Black 2003; Vlasova et al. 2015; Malyutina et al. 2012).

Moreover, for different speakers, different words will represent the main (most frequent and stylistically neutral) variant, and at times this can vary even for the exact same speaker. Whether such an abundance of names extends only to cultural vocabulary or can also affect some basic meanings is unknown.

In addition to such cases of true synonymy, where words truly have the same meaning, and frequency and stylistic characteristics differ either among different speakers or in different communication situations, there are a number of situations where either the same original root becomes unrecognizable due to phonetic or morphological changes, or two different roots, initially differing in meaning or morphological properties, lose these differences, and different descendant languages choose either one or the other as the main one. Let us consider these cases in more detail.

As a result of phonetic changes, a root can lose some phonemes, and then reanalysis can occur. For example, in modern Spanish, the meaning ‘to eat’ is expressed by the root *com-*, cf. *comer* ‘to eat’, *comida* ‘food’ (this same root is present also in Portuguese and Galician). This happened because the Latin root *ed-* ‘to eat’ (inf. *edere*) underwent the loss of *d* in most forms, and at some point speakers, to ensure better understanding, began to replace *edere* with *com-edere* ‘to eat up’. The root *-e-* in *comer* ceased to be distinguished, and the former prefix *com-* (‘with’) became the carrier of the meaning ‘to eat’. If only Spanish, Portuguese, Galician, and, for example, Slavic or Germanic languages remained from the Indo-European family, the researcher would have to deal with proto-language synonymy: two roots would be reconstructed for the meaning ‘to eat’, **ed-* and **com-*, with no observable semantic difference between them.

Roots are especially affected in languages which, due to a restructuring of phonotactics, lose polysyllabicity. For example, in Mon-Khmer and Kra-Dai languages transitioning from polysyllabism to monosyllabism, initial syllables (“presyllables”) become “weak”: they cannot

appear alone, without a strong syllable nearby, and are characterized by a number of restrictions — they have no tone, no consonant clusters are possible at the beginning, the vowel is reduced, and there is no final consonant. This leads to a large number of phonetic changes in initial syllables.

Furthermore, in such presyllables, variability like “*khr-* ~ *khl-*, *phr-* ~ *phl-*, *khr-* ~ *phr-*, etc.” (Samarina 2007: 46) is often noted, with different speakers possibly preferring certain variants or using both freely (*ibid.*). For example, in the Ruc language (Vietic group of the Mon-Khmer branch of the Austroasiatic family, Vietnam) in words like *rata*³ ~ *lata*³ ‘stone’ or *rəka*¹ ~ *təka*¹ ‘chicken’, free variation of the initial consonant is observed (*ibid.*: 43). If such variation occurred at the moment of the proto-language’s binary split, one variant could gain greater frequency in one daughter branch, and another one in the second one.

Such presyllables could be former prefixes, and in this case, despite the variability, roots could be reconstructed correctly, allowing cognates to be identified. But if root morphemes stood in the initial position in the past (and in the final position there were either also root or suffixal morphemes), then the more time has passed between the split of the proto-language and its descendants available for study, the less likely it is to determine the original identity of the roots, and the higher the probability that reflexes of the same proto-language root will look like etymologically unrelated synonymous morphemes without any semantic distribution.

The situation would look the same if the original phonetic differences were associated with male / female speech, politeness, taboo, or were stylistic variants. Furthermore, the more time has passed since the split of the proto-language, the more phonetic changes have occurred and the more lexemes have been lost, so the initial conditions of phonetic changes may be impossible to establish due to the material being too limited.

False synonymy can be due to morphological reasons. For example, in Atlantic languages (a West African family), the noun class marker can be prefixal, suffixal, or circumfixal. Thus, in the Bapen language (Tenda group of the northern branch), the word for ‘tree’ is *a-tu-ã* (Pl. *ba-tu-bã* indicates that its root is *tu*); in other languages of the Tenda group, class markers are represented by prefixes, cf. Basari *a-tax* (Pl. *ba-tax*), Bedik *ga-tɔ* (Pl. *ba-tɔ*) ‘tree’ (Pozdnyakov 1993: 52). Markers represented as suffixes in Fula (Fula-Serer group of the northern branch) appear as circumfixes in Serer (*ibid.*: 27–31). This state of affairs may be a consequence of grammaticalization processes. For example, in Swahili, the expression ‘these Swahili people’ looks like *wa-suahili ha-wa*, where the class marker *wa* appears both at the beginning and at the end of the noun phrase. With the development of grammaticalization according to the standard model, first the demonstrative pronoun turns into a definite article, then the article begins to accompany every noun and becomes simply a noun marker (see, e.g., Kuteva et al. 2019: 27). In doing so, the corresponding morpheme undergoes phonetic and prosodic erosion. Subsequently, redundancy can be reduced in one way or another: either the initial or the final part of the marker can be omitted. Taking into account the fact that the class marker can be determined by the phonetic characteristics of the lexeme (Sumbatova 2022), a situation of reanalysis is quite realistic. Thus, for the word ‘tongue’ in Proto-Atlantic, the form **d-em-əd* ~ **r-em-ər* is reconstructed (Pozdnyakov 1993: 146–147); subsequently it developed into **de-məd* (root **məd*) in some languages and into *dem-əd* (root **dem*) in others, cf. Kisi *demu-le* (Pl. *demu-ã*), Landuma *da-mer* (Pl. *sə-mer*). The more time has passed since the split of the proto-language, the less material researchers have to trace the grammaticalization process, and accordingly, the more likely it is to obtain two seemingly unrelated synonyms like **lem* and **mel* during reconstruction.

Word formation can also be a source of false synonymy. For example, in Nicobarese languages, the concept ‘palm (of hand)’ is conveyed by the words *el-ti*: (Car language) and *mòh-ti*: (Teressa language). The second part in these lexical complexes goes back to the same proto-

language root, while the initial parts are different: *el* means ‘inner part’, *mòh* means ‘tip, protrusion, nose’ (Pogibenko 2018: 203). Such formations often represent, in essence, so-called “micro-syntactic” constructions where components can be freely replaced by similar ones. The more time has passed since the split of the proto-language, the higher is the probability that in at least one of the daughter branches, the part that represented the proto-language root was lost — for purely phonetic reasons — and then in the reconstruction, seemingly unrelated lexemes like *elt* and *moh* will appear. Languages prone to this type of word formation often repeat the corresponding processes as the semantic motivation of the previous lexical complex becomes obscured. For example, in modern Khmer, the word *cɹvmuh*, which arose from the lexical complex ‘hole’ + ‘nose’ (< Old Khmer *muh* ‘nose’) and originally meant ‘nostril’, came to mean ‘nose’, and for the meaning ‘nostril’, a new lexical complex of words with the same meaning is used: *rɔnthəəʔ cɹvmuh*, where the second component means ‘nose’, while the first goes back to a borrowing from Sanskrit *randhra* ‘hole’ (Pogibenko 2018: 203–204).

Perhaps cases of this type, like the case of Spanish *com-* ‘to eat’, should be treated as lexical replacements. But it is not entirely clear whether replacements of this type have the same speed as replacements occurring due to semantic shifts, or a different one.

Another reason for the appearance of synonymy in a proto-language wordlist can be suppletion. In a proto-language — just as in any extant language — words could have a suppletive paradigm, cf. English *go* and *went*, Tocharian A *y-* (Pres.) / *kālk-* (Pt.) ‘to go’. The paper Kassian et al. 2010 proposes to choose the morphologically simplest stem. Sometimes this approach may conflict with the principle of choosing the most frequent variant, since for different lexemes, various morphological forms can have unequal frequency. So, for example, according to data from the SynTagRus corpus (part of the Russian National Corpus, see, e.g., Timoshenko et al. 2021, for the word *viетка* ‘branch’ the most frequent form is the instrumental plural, for the words *gora* ‘mountain’ and *ruka* ‘hand’, the prepositional plural.

In accordance with the principle of choosing the most frequent of the synonymous designations for the same concept, “for telic verbs, apparently, the perfective aspect form should be taken, and for atelic verbs — the imperfective aspect form: the frequencies of the corresponding dictionary units differ very much” (Burlak 2021: 298).

At the same time, in different sets of texts (and, apparently, for different consultants), frequencies can have even opposite distributions. For example, in modern standard Russian, according to data from the Russian National Corpus (as of February 16, 2025), the forms *chelovek* ‘person (sg.)’ and *lyudi* ‘people (pl.)’ (in all cases together) have approximately equal frequency: *chelovek* occurs 553,208 times, *lyudi* — 571,061 times, but in the Old Russian corpus, the plural *lyudie* (1445 examples) occurs twice as often as the singular *chelověkъ* (713 examples), while in the Dialect corpus the ratio is reversed: for 410 uses of the singular *chelovek*, only 7 examples of the plural *lyudi* were found. When a population splits (which is necessary for the split of a proto-language), speakers who more often use one of the suppletive morphological forms could predominate in one of the daughter populations while people of the other one could more often use the other form; subsequently, this could lead to paradigm unification (in one daughter population based on one stem and in the other, based on its counterpart), and in this case, during step-by-step reconstruction, two seemingly unrelated synonyms will compete for the same slot in the proto-language wordlist. The more time has passed since the split of the proto-language, the less chance there is to find any evidence that in the proto-language both corresponding stems were present and formed one suppletive paradigm with such-and-such distribution of forms by grammatical meaning.

The grammatical meanings by which suppletive stems are distributed can be very different. For example, in many languages spoken on the northwestern coast of North America, stems

used for singular and plural subjects in the verb are suppletive (Campbell 2004: 304–305). The verb ‘to give’ can distinguish stems depending on who is given to.

Initially, suppletion could have developed due to the rarity of phonetic alternation. For example, for Old Church Slavonic, the forms of the infinitive *gъnati* ‘to drive, to chase’ and the present tense *ženetъ* can be meaningfully considered as suppletive (in Polivanova 2013: 72, 430, 588 they are recognized as allomorphs with an anomalous relation), despite the fact that they exhibit regular ablaut: *e*-grade in the present, zero-grade in the infinitive. The more time has passed since the split of the proto-language, the less likely it is to find enough material to discover the rule for the distribution of allomorphs.

The most frequent source of false synonymy is semantics: words can be replaced due to connotations (like the Latin designation for ‘black’, *ater*, was replaced in all Romance languages, see Normanskaya 2005: 180) or due to emphasis: words that had emphatic highlighting could lose it, becoming the most frequent and neutral designations of the corresponding concepts (in this case, new emphatically highlighted words could take their place), or be lost and replaced by new emphatically highlighted words. In different daughter branches, these processes could proceed differently, and the more time has passed since the split of the proto-language, the less chance there would be to trace this, and then two synonyms will compete with equal grounds for a slot in the proto-language wordlist: both the one that was originally primary and the one that was originally stylistically marked.

Furthermore, any language segments reality into fragments denoted by individual words in ways that are distinct from any other language. In the paper Kassian et al. 2010, trying to formalize the principles of compiling 110-item wordlists and develop a meticulous algorithm for identifying the most basic meanings from a set of close ones, contexts are provided that should help the researcher choose the most basic lexeme from those with similar meanings. For example, for the meaning ‘to lie’, it is proposed to take the “verb of continuative/durative action. It is generally not recommended to fill in the slot with the inchoative verb ‘to lie down’, although such practice is rather frequent (and, in truth, both notions are quite commonly expressed with the same root). Should be strictly separated from lying in particular postures (‘lie on one’s side’, ‘lie face down’ etc.)” (Kassian et al. 2010: 66). For modern languages, where consultant questioning is possible, making such distinctions is quite feasible, but in the process of language development, any one such verb (‘to lie down’, ‘to lie on one’s side’, ‘to lie face down’, etc.) could undergo a semantic shift in one of the daughter branches and become the main verb for the meaning ‘to lie’, so that if the corresponding root is lost in other branch(es), the semantic development between the descendant of the proto-language word ‘to lie’ and the proto-language word ‘to lie down’ (or similar) may be impossible to establish.

Words for ‘man’ and ‘woman’ can originate from kinship terms (see above), and kinship terms in many languages are subdivided into referential and appellative. For example, in Tamil, the meaning ‘mother’ is expressed with “the two main roots, *ammā* and *tāy*, of which the root *tāy* is referential, and *ammā* is appellative” (Smirnitskaya 2024: 64). If a long time has passed since the split of the proto-language, semantic shifts could have occurred in the daughter populations, as a result of which the meaning ‘woman’ originated in one language from that word for ‘mother’ which in the proto-language was appellative, and in another — from the one that was referential; in this case, during reconstruction, two absolutely equal synonyms will compete for the slot ‘woman’, and even if it would be possible to establish that each of them came from a word meaning ‘mother’, it would not be possible to choose the “most basic” lexeme.

Contexts formulated on the base of the study of a large number of languages allow many semantic distinctions to be removed, but hardly all of them. For example, for the meaning

‘tail’, the authors of Kassian et al. 2010 propose to take the word that applies to mammals (thereby dismissing separate terms for the tail of a bird, fish, etc.) and is used in contexts: “1. Animals have tails, man does not have a tail. 2. (The beast) is swishing its tail to drive away the flies. 3. He cut off (the beast’s) tail” (Kassian et al. 2010: 79). However, a language may have an opposition between a fluffy tail (like a fox’s) and a non-fluffy one (like a cow’s) or similar. For example, in Icelandic, mammals have tails such as *dindill* ‘short tail (like a sheep’s or seal’s)’, *tagl* ‘horse’s tail’, *hali* ‘tail with a tuft at the end (like a lion’s or cow’s)’, *rófa* — one of the most common tails, such as a cat’s, mouse’s, or fox’s (but its first meaning, according to the dictionary, is not ‘tail’ but ‘coccyx’) and *skott*, the same as *rófa*. In the context “to cut off the tail”, any of these is acceptable. If we choose the most frequent variant, then one must choose *rófa* (if its first meaning would not prevent this) or *skott* (if focusing more on semantics than frequency), and if we rely on the context “wave the tail, driving away insects”, then *hali* should be preferred (as in the database GLD 2025) or *tagl*.

It is not always helpful to use context in order to remove the distinction between a large (usually predatory) bird and a small bird: the contexts “1. Something is moving in the bushes, I cannot tell if it is a bird or an animal. 2. Birds lay eggs, animals and people bear children. 3. There is a bird flying on high, I cannot tell what kind”, recommended in Kassian et al. 2010: 52, may not yield a result, since distinguishing a far-soaring eagle from a closely fluttering sparrow is not difficult for people who know nature well. The context “Birds lay eggs, animals and humans bear children” may appear confusing, just like Russian or English speakers would find confusing a context in which a researcher accustomed to grouping birds with reptiles would ask them to translate how they (diapsids) all lay eggs.

The 400-item wordlist of basic vocabulary, currently undergoing testing as part of the Moscow School’s “Comparative Onomasiological Database for the Ancestral States of Eurasian Linguistic Families” (<https://starlingdb.org/new400/>), includes other meanings that in languages can be represented by more than one word. For example, the meaning ‘to fall’ (also included in the “Leipzig-Jakarta wordlist” as proposed in Haspelmath & Tadmor 2009) can be expressed in four different ways: “movement from a higher surface to a lower one (*the vase fell off the table*), loss of vertical orientation (*the vase fell and the water spilled on the tablecloth*), falling-collapse (*the house fell down*) and detachment (*the dress fell off the hanger*)” (Reznikova et al. 2020: 9). The meaning ‘to love (someone, not something)’ in Ancient Greek is conveyed by 4 verbs (initially, apparently, corresponding to relationships in the nuclear family), cf. *στέργω* (initially describing love of parents for their children), *ἀγαπάω* (initially describing love of children for their parents), *φιλέω* (initially describing love of siblings for each other) and *ἐρᾶω* (initially describing love of the spouses for each other).

In a number of cases, the possibility of the absence of a single basic lexeme with the necessary meaning is recognized in Kassian et al. 2010. For 8 words from the 110-word list, the paper recommends to accept synonyms if there is no general term, namely:

- 1) ‘to fly’, if there are different roots for an iterative verb ‘to fly’ (*Where is the bird flying?*) and a habitual verb ‘to fly’ (*Little chicks do not yet know how to fly*);
- 2) ‘to give’, if a language makes a clear and regular distinction between ‘to give (to 1st/2nd person)’ and ‘to give (to 3rd person)’;
- 3–4) ‘man’ and ‘woman’, if a language has different terms for members of different age groups (in such cases, the respective lexemes for the age of ca. 20–45 are allowed to be treated as synonymous);
- 5) ‘that’, if a language has a more elaborated deixis system;
- 6) ‘not’, if different tense / aspect etc. forms use different negative markers and no preferences can be easily formulated;

7) ‘mouth’, if a language possesses two separate terms, one for ‘oral cavity’ (*His mouth is full of water*) and another for ‘external parts of the mouth’ (*He has a very wide mouth*);

8) ‘tooth’, if a language has separate terms for ‘back tooth, molar’ and ‘front tooth, incisor’.

Other 5 slots are supposed to be split into several: these are pronouns (‘I’ and ‘you’, where differences between direct and oblique stems are common, and ‘we’, where besides this difference, a clusivity distinction is often also present), the words ‘round’ (where characteristics of disc-shaped and spherical objects are often distinguished) and ‘thin’ (where terms characterizing flat and elongated objects may be distinguished, like a thin sheet and a thin rope). For the word ‘warm’, the possibility of splitting into ‘warm’ and ‘hot’ is provided, and for the word ‘year’, the possibility of not filling the slot in the absence of a general term is envisaged.

The grounds for distinguishing some *realia* from others, naming them with separate words, can be very diverse, and if the intuition of the list compilers and native speakers does not coincide (or semantic distinctions are not well researched), words with differing meanings can look like synonyms. For example, Russian has the words *pepel* and *zola*, both meaning ‘ashes’. In the contexts proposed in Kassian et al. 2010: 51 (“1. The campfire has left only ashes. 2. The wind scattered the ashes. 3. He scooped up a handful of ashes”), both are acceptable, with different speakers producing different variants as their first reaction, and neither of them in these contexts appears to be more frequent than the other. Cf. (about a person burning money in a fireplace): “Bumping into something hot, he would pull his hand back — and again thrust it into the fireplace. Having gathered a smoldering pile of ashes (*zola*), he carried it through the bedroom to the bathroom, carelessly spilling it on the floor... In two goes, he transferred almost all the ashes (*pepel*) without residue in the fold of his robe and with it many small embers, from which the robe smoked” (M. A. Osorgin “Svidetel’ istorii” 1932, cit. via RNC. Translated by DeepSeek — S.B.); “A slow rain of ashes (*zola*) fell. It covered the river water with a gray coating. Sometimes birch leaves, turned into ashes (*pepel*), flew down from the sky” (Paustovsky, “Zheltiy svet”, 1936, cit. via RNC, translated by DeepSeek — S.B.).

The greater frequency of the word *pepel* (it is used 4847 times in the main RNC corpus, while *zola* is only encountered 2893 times) is likely provided by collocations like *vulkanicheskiy pepel* ‘volcanic ash’, *prevratit’ v pepel* ‘to turn to ashes’, or *posypat’ golovu pepлом* ‘to pour ashes on one’s head (in mourning)’. In the Old Russian corpus there are 26 examples of *pepel* and its variant *popel*, while *zola* does not occur at all (although the lexeme *zola* is of Proto-Slavic origin); in Old Russian there are 50 examples of *pepel* and 49 examples of *zola*.

From a fire or conflagration, both *pepel* and *zola* remain in roughly equal measure. Both could be used for treatment, cf. “Take a live lizard and burn it to ashes (*pepel*). Then sprinkle those ashes (*pepel*) on the horse’s sore...” and “And however many years old the horse will be, take that many fresh eggs and burn them and make ashes (*zola*). And mix it with that milk and pour it into the horse’s mouth, it will be healthy” — both examples are from a medical book on horses dating to the third quarter of the 17th century, cit. via RNC (translated by DeepSeek — S.B.). With the spread of tobacco smoking, the collocations *tabachnyj pepel* and *tabachnaja zola*, both meaning ‘tobacco ash’, initially competed in roughly equal measure: in the RNC from the early 19th century to the present, there are 28 examples of “tobacco *pepel*” and 20 of “tobacco *zola*”), but in the 20th century, *zola* loses ground: *papirosnaja zola* ‘cigarette ash’ ends in 1938, and *tabachnaja zola* ‘tobacco ash’ — in 1980.

Nevertheless, careful analysis of contexts shows that these lexemes are not fully synonymous. Cf.: “And what’s the difference: *pepel* and *zola*? Which is cleaner? Which is more terminal? *Pepel*, of course — *zola* is even used for fertilizing” (M. I. Tsvetaeva “Istorija odnogo posv’ashchenija”, 1931, cit. via RNC, translated by DeepSeek — S.B.). According to RNC data, only *pepel* can be volcanic, and only *zola* can be taken from a stove. Only *pepel* can fall off

(*osypat'sya*), fall (*padat'*), and fly (*letat'*), but only *zola* can be produced (*poluchat'sya*), serve (*sluzhit'*), and contain smth. (*soderzhat'*); only *pepel* can be shaken off (*str'akhnut'*), blown away (*sdut'*), and scattered (*razveyat'*), but only *zola* can be raked apart (*razgresti*) and scooped out (*vygrebat'*), while both can be raked (*razgrebat'*); one can be reborn (*vozrodit'sya*) only from *pepel*, but bake (*ispech'*) something only in *zola*; only the lexeme *pepel* enters into coordinative constructions with *almaz* 'diamond', *okurok* 'cigarette butt', *lava* 'lava', and *razvalina* 'ruin', but only *zola* — with *navoz* 'manure', *golov'oshka* 'ember', and *sazha* 'soot'. Overall, the lexeme *zola* is used when the corresponding substance is used for something, cf. the following examples (cit. via RNC, translated by DeepSeek — S.B.):

“When worms attack this tree, one should water its stump with vinegar, or sprinkle it with ashes (*zola*)” (V. A. Levshin. *Vseobshcheje i polnoje domovodstvo*, part V, 1795);

“To destroy it — i.e., the earth flea — they sprinkle flax seedlings with ashes (*zola*), I read in one book” (A. N. Engelgardt. *Pis'ma iz derevni. Pis'mo tretje*, 1872);

“They removed the worm, sprinkled the cabbage with ashes (*zola*) against the fleas, watered the garden” (N. A. Leykin. *Derevenskaja prelestnitsa*, 1908);

“I fumigate the field with smoke, sprinkle it with ashes (*zola*) from a basket” (I. A. Aramilev. *V lesakh Urala. Novyj mir*, 1941);

“Burdenko pressed against someone's gates and watched excitedly as she walked along the high wooden sidewalk, sprinkled with stove ashes (*zola*) so it wouldn't be slippery” (Pavel Nilin. *Interesnaja zhizn'*, 1969–1980);

“In the valleys of the Pamirs, local residents have long sprinkled fields with ashes (*zola*) and earth to accelerate spring snowmelt” (L. R. Serebr'annyj, A. V. Orlov. *Ledniki v gorax*, 1985);

“And now it was enough to eat, say, a sprat, drink, say, some Cahors, catch a cold, moving between snowdrifts along paths sprinkled with ashes (*zola*) in the settlement's streets, and the old ailment made itself known, weakly, quietly, but noticeably” (Asar Eppel. *Poka i poskol'ku*, 1991);

“All cuts, especially on roots, are sprinkled with wood ashes (*zola*) — roots dusted with ashes (*zola*) receive an additional stimulus for development” (Valerija Iljina. *Pokor'ajushchije vysotu. Nauka i zhizn'*, 2009);

“To obtain potash, dense wood was burned to ashes (*zola*), the ashes (*zola*) were boiled in water for several days in specially adapted copper pots” (Arkadij Kuramshin. *Elementy. Zamechatel'nyy son professora Mendeleeva*, 2019).

If we are talking about *pepel*, the contexts are completely different: besides the collocations listed above, *pepel* is primarily what is obtained as a result of burning something or someone, cf. the following examples (cit. via RNC, translated by DeepSeek — S.B.):

“For that, considerable booty was obtained, and the whole village was turned into ashes (*pepel*)” (F. I. Soymonov. *Opisanije Kaspijskogo mor'a*, 1763);

“Do not burn me to ashes (*pepel*) with your anger!” (A. A. Bestuzhev-Marlinskij. *Mulla-Nur*, 1836);

“He was cast out of a window, massacred, burnt, and his ashes (*pepel*) were blown abroad from a cannon's mouth, to the four winds of heaven!” (A. S. Pushkin. *Kapitanskaja dochka*, 1836);

“... she only turned pale, hastily burned the letter and blew its light ashes (*pepel*) onto the floor” (M. Yu. Lermontov. *Kn'agin'a Ligovskaya*, 1836–1837);

“His path was strewn with corpses and ashes (*pepel*) of burned villages and cities” (N. N. Alekseev. *Brat na brata*, 1904);

“When the hut burned down, the bravest ventured to rummage in the ashes (*pepel*)...” (Yevgenij Vodolazkin. *Lavr*, 2012);

“There among the burned corals and ashes (*pepel*), Albert Ghiorso discovered atoms of element number 99...” (Arkadij Kuramshin. *Elementy. Zamechatel’nyy son professora Mendeleeva*, 2019).

Even when *pepel* is used for something, it is necessarily indicated what exactly was burned (for *zola* such an indication is not obligatory), cf. (cit. via RNC, translated by DeepSeek. — S.B.):

“After that, this earth is covered with a smaller layer of mixed matter, which contains two parts of ashes (*pepel*), burned from hardwoods, and three parts of quicklime...” (M. V. Lomonosov. *O varenii selitry*, 1741);

“If he was strong enough, he uprooted stumps, burned the whole pile of felled wood and sowed bread on the earth fertilized with ashes (*pepel*), putting it under hay” (N. I. Berezin. *Peshkom po karel’skim vodopadam*, 1903).

Furthermore, *pepel* is used in metaphorical contexts to a greater extent than *zola*, cf., for example (cit. via RNC, translated by DeepSeek. — S. B.): “A dry face. Eyes extinguished, as if sprinkled with ashes (*pepel*)” (G. Ya. Baklanov. *Pyad’ zemli*, 1959).

Also, *pepel* (but not *zola*) denotes such a specific thing as the gray coating that forms on the surface of just extinguished coals, cf. (cit. via RNC, translated by DeepSeek. — S. B.):

“...a coal that had just gone out, covered in places with *pepel*...” (F. P. Wrangel. *Puteshestvije po severnym beregam Sibiri i po Ledovitomu mor’u*, 1841);

“Despite the hot season, embers covered with *pepel* glowed in the fireplace” (F. F. T’utchev. *Na skalax i dolinax Dagestana*, 1903).

Thus, apparently, speakers of Russian choose the lexeme *pepel* when they want to emphasize that these are the remains of something (or someone) burned, and the lexeme *zola* when they want to emphasize that the corresponding substance is used for a specific purpose. Interestingly, it was the Russian word *zola* that was chosen as the main word for the slot ‘ashes’ in Kassian et al. 2010, despite the fact that the lexeme *pepel* has a higher frequency, and this decision appears to be well justified.

Incidentally, it is possible that the situation with PIE **péh₂wr̥* and **h₁ng^wnís* ‘fire’ was also of this sort. Since these two nouns display different genders (**péh₂wr̥* is neuter, while **h₁ng^wnís* is masculine), they may have originally referred to ‘fire’ as ‘substance’ and ‘fire’ as a ‘personal’ entity respectively (see Yakubovich 2021; Schmidt 2010 etc.). An anonymous reviewer suggests that ‘fire as substance’ would be a more natural candidate for lexicostatistical purposes.

Apparently, distinctions of this sort are further evidence of the absence of any presumed innate universal, language-independent mental lexicon suggested in some works — such as “mentalese” (Pinker 1994) or “universal object code” (Zhinkin 1998). Each language at each moment in time clusters the infinite variety of elements of surrounding reality in different ways and on different grounds. Certainly, some similar clustering options can be found quite often, but at any moment the next language available for research can draw boundaries between concepts in its own way.

The differences in ways of clustering *realia* are now actively studied within lexical typology (see, e.g., Reznikova et al. 2020). Perhaps someday a list of possible semantic distinctions, completely exhausting the possibilities of the human mind, will be compiled. But until this is done, it probably makes sense to adapt the methodology so that it can work with lists that take into account the possibility of “synonyms” (such that the real differences between their meanings are not removed by the suggested contexts) appearing at any stage of reconstruction.

That said, these observations should not be interpreted as a call to allow any number of synonyms for any word in the wordlist of any proto-language at any stage of reconstruction. For most of the described processes, a large amount of time must pass: for instance, unrecog-

nizable synonymy which arises due to phonetic or morphological reasons usually requires a chain of phonetic changes, including potentially irregular ones, that obscure the original phonemic composition, and the number of daughter branches must be minimal. (Perhaps this is why language isolates, which have no close relatives, in many cases cannot be convincingly assigned to any macro-family.) When a language does have close relatives, the methodology of step-by-step reconstruction (involving the maximum available amount of data), such as the one adopted by the Moscow School of comparative linguistics, in most cases makes it possible to determine which of the synonyms was the primary one in the proto-language or, at least, could be considered an optimal choice next to any competition (Starostin 2010: 102).

Nevertheless, all of this means that it is always necessary to accompany any reconstructions in onomasiological databases with comments listing their potential quasi-synonyms with references to semantic distinctions that were already tested and excluded (thus, when a new semantic distinction is found, it will be easy to see which reconstructed lexemes need to be re-tested). In those cases where it is impossible to identify one primary candidate, it makes sense to include all of them into reconstructed wordlists, assigning equal (or unequal) probabilities to each one. This would make it easier to reexamine them when a new semantic distinction (or additional sociolinguistic information) comes to light.

References

- Belikov, Vladimir. 2006. *Konvergentnyje protsessy v lingvogeneze (Convergence in linguogenesis)*. Doctoral dissertation: Moscow.
- Black, Paul. 1997. Lexicostatistics and Australian languages: Problems and prospects. In: Darrell T. Tryon, Michael Walsh (eds.). *Boundary rider: essays in honour of Geoffrey O'Grady. Pacific linguistics, C-136*: 51–69. Canberra: The Australian National University.
- Black, Sheila R. 2003. Review of Semantic Satiation. In: S. P. Shohov (ed.). *Advances in psychology research. Vol. 26*: 63–74. Nova Science Publishers.
- Burlak, Svetlana. 2021. Ustoichivost' i chastotnost': jest' li korrel'atsija? (Stability and frequency: is there a correlation?). *Journal of Language Relationship* 19(3–4): 293–307.
- Campbell, Lyle. 2004. *Historical Linguistics: An Introduction. 2nd edition*. Cambridge: The MIT Press.
- Dunn, Michael John. 1999. *A Grammar of Chukchi*. Ph.D. thesis: Australian National University.
- Dyakonov, Igor'. 1984. Sravnitel'noe jazykoznanije, istorija i drugije smezhnyje nauki (Comparative linguistics, history and neighboring sciences). In: *Lingvisticheskaja rekonstrukcija i drevnejshaja istorija Vostoka. Tezisy i doklady konferentsii. Part 2*: 3–20. Moskva: Institut vostokovedeniya RAN.
- Dyachenko, S. V., A. V. Malysheva, A. V. Ter-Avanesova. 2018. Iz nabl'udenij nad fonetikoj vladimirsko-povolzhskix govorov: jokan'je i pozitsionnaja m'agkost' zh, sh (Some observations on the Vladimir-Povolzh'ye dialect phonetics: yokan'ye and positionally determined palatal articulation of zh and sh). *Trudy Instituta russkogo jazyka imeni V. V. Vinogradova* 3(17): 41–68.
- Frankham, Richard, C. J. A. Bradshaw, B. W. Brook. 2014. Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation* 170: 56–63.
- GLD 2025 = Starostin, George (ed.). 2011–2025. *The Global Lexicostatistical Database*. Moscow: Higher School of Economics, & Santa Fe: Santa Fe Institute. Available online at: <https://starlingdb.org/new100/> [accessed September 1, 2025].
- Harcourt, Alexander H. 2002. Empirical estimates of minimum viable population sizes for primates: Tens to tens of thousands? *Animal Conservation Forum* 5(3): 237–244. doi:10.1017/S1367943002002287
- Haspelmath, Martin, Uri Tadmor (eds.). 2009. *Loanwords in the world's languages: a comparative handbook*. Berlin: Walter de Gruyter GmbH & Co.
- Henrich, Joseph. 2016. *The secret of our success. How culture is driving human evolution, Domesticating our species, and making us smarter*. Princeton University Press.

- Holmes, Janet, Nick Wilson. 2022 *An introduction to sociolinguistics. 6th edition*. London: Routledge.
- Iomdin, Boris. 2023. Telegram-kanal «Uznel novoe slovo» (Telegram channel “I’ve got to know a new word”). Available online at: <https://t.me/UznelSlovo/627>.
- Jackson, Matthew O. 2019. *The Human Network: How Your Social Position Determines Your Power, Beliefs, and Behaviors*. New York: Knopf Doubleday Publishing Group.
- Kassian, Alexey, George Starostin, Anna Dybo, Vasily Chernov. 2010. The Swadesh wordlist. An attempt at semantic specification. *Journal of Language Relationship* 4: 46–89.
- Khrisanfova, Elena, Ilya Perevozchikov. 2005. *Antropologija (Anthropology)*. Moskva: Nauka.
- Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog, Seongha Rhee. 2019. *World lexicon of grammaticalization. 2nd ed.* Cambridge University Press.
- Malyutina, S. A., O. V. Dragoy, A. G. Petrushevskiy, O. N. Fedina, M. V. Ivanova, D. A. Sevan, E. F. Gutyrchik. 2012. Nejrofiziologicheskiye korrel’aty chastotnosti slova pri jego nazyvanii (Neurophysiological correlates of the word frequency while citing it). In: *Mezhdunarodnyj simpozium po nejroimidzhingu: fundamental’nyje issledovaniya i klinicheskaja praktika*: 83–85. Moskva: MGPPU.
- Milroy, Lesley, Matthew Gordon. 2008. The concept of social network. In: Lesley Milroy, Matthew Gordon. *Sociolinguistics: Method and interpretation*: 116–133. Oxford: John Wiley & Sons.
- Moulin, Chris J. A., Nicole Bell, Merita Turunen, Arina Baharin, Akira R. O’Connor. 2021. The induction of jamais vu in the laboratory: word alienation and semantic satiation. *Memory* 2021 Aug. 29(7): 933–942. doi: 10.1080/09658211.2020.1727519.
- Normanskaya, Julia V. 2005. *Razvitije i genezis sistem tsvetooboznachenij v drevnix indoevropskix jazykax (Development and genesis of the color name systems in ancient Indo-European languages)*. Moskva: ZAO «S&K».
- Ossovetskiy, I. A. (ed.). 1969. *Slovar’ sovremennogo russkogo narodnogo govora (d. Deulino R’azanskogo rajona R’azanskoj oblasti (Dictionary of a modern Russian dialect: Deulino village, Ryazan’ region))*. Moskva: Nauka.
- Pinker, Steven. 1994. *The Language Instinct: How the Mind Creates Language*. New York: William Morrow and Company.
- Pogibenko, Tamara. 2013. *Avstroaziatskije jazyki: problemy grammaticheskoy rekonstruktsii (Austro-Asiatic languages: the problems of grammar reconstruction)*. Moskva: Institut vostokovedeniya RAN.
- Pogibenko, Tamara. 2018. Leksicheskij kompleks v jazyke ma (Lexical complex in the Ma language). *Vestnik Instituta vostokovedeniya RAN* 5: 200–211.
- Pozdnyakov, Konstantin. 1993. *Sravnitel’naya grammatika atlanticheskix jazykov (Comparative grammar of the Atlantic languages)*. Moskva: Nauka.
- Polivanova, Anna. 2013. *Staroslav’anskij jazyk. Grammatika. Slovare (Old Church Slavonic language. Grammar. Dictionaries)*. Moskva: Russkij fond sodejstvija obrazovaniju i nauke.
- Reznikova, T. I., E. V. Rakhilina, D. A. Ryzhova. 2020. Glagoly padeniya v jazykax mira: frejmy, parametry i tipy sistem (Verbs of falling in the world’s languages: frames, parameters and system types). *Acta Linguistica Petropolitana* 16(1): 9–63. doi: 10.30842/alp2306573716101
- RNC = 2003–2025. *The Russian National Corpus*. Available online at: ruscorpora.ru [accessed on September 1, 2025].
- Samarina, Irina. 2007. Protsessy monosillabizatsii v jazykax materikovoj Jugo-Vostochnoj Azii (Monosyllabization processes in the languages of continental Southeast Asia). *Voprosy filologii* 1(25): 41–49.
- Schmidt, Karl Horst. 2010. Miscellanea Anatolica. In: J. Klinger et al. (eds.). *Investigationes Anatolicae: Gedenkschrift für Erich Neu*: 241–252. Wiesbaden: Harrassowitz.
- Simons, Gary F. 1982. Word taboo and comparative Austronesian linguistics. In: Amran Halim, Lois Carrington, S. A. Wurm (eds.). *Papers from the third international conference on Austronesian linguistics, Vol. 1: Currents in Oceanic. Pacific linguistics, C-74*: 157–226. Canberra: The Australian National University.
- Skorik, P. J. 1961. *Grammatika chukotskogo jazyka (Grammar of the Chukchi language)*. Moskva / Leningrad: Izdatel’stvo Akademii nauk SSSR.
- Smirnitskaya, Anna. 2024. *Dravidijskie terminy rodstva: tipologija semanticheskix perexodov (Dravidian kinship terms: the typology of semantic shifts)*. Moskva: Institut vostokovedeniya RAN.
- Starostin, George. 2010. Preliminary lexicostatistics as a basis for language classification: A new approach. *Journal of Language Relationship* 3: 79–116.
- Starostin, George. 2019. Chinese basic lexicon from a diachronic perspective: implications for lexicostatistics and glottochronology *Journal of Language Relationship* 17(2): 153–176.

- Sumbatova, Nina. 2022. Landuma: a case of radical alliterative agreement. *Journal of African Languages and Linguistics* 43(1): 83–118. doi:10.1515/jall-2022-8893.
- Tadmor, Uri. 2015. When culture grammaticalizes: The pronominal system of Onya Darat. In: R. De Busser, R. J. LaPolla (eds.). *Language Structure and Environment. Social, cultural, and natural factors*: 77–98. Amsterdam / Philadelphia: John Benjamins Publishing Company.
- Timoshenko, S. P., L. L. Iomdin, S. A. Gladilin, E. S. Inshakova. 2021. SinTagRus v sostave NKRJa: novye vozmozhnosti (SynTagRus within the Russian National Corpus: New Possibilities). In: V. P. Zaxarov (ed.). *Korpusnaya lingvistika — 2021. Trudy mezhdunarodnoj konferentsii*: 31–43. Sankt-Peterburg: Skifija-Print.
- Vallejos-Yopán, Rosa. 2015. La indexicalidad de género en kukama-kukamiria desde una perspectiva tipológica. In: Zarina Estrada Fernández, Ana Fernández Garay, Albert Alvarez Gonzalez (eds.). *Estudios de lenguas amerindias 3: Escenarios de diversidad lingüística*: 195–222. Hermosillo, Sonora: UNISON.
- Vavilov, Nikolay. 1935. *Zakon gomologicheskix r'adov v nasledstvennoj izmenchivosti* (The law of homology series in genetical mutability). Moskva/Leningrad: Sel'hozgiz.
- Vlasova, R. M., V. E. Sinitsyn, E. V. Pechenkova. 2015. The Effect of Word Frequency on the Brain Correlates of Object Naming in Russian. *The Russian Journal of Cognitive Science* 2(1): 24–40.
- White Hat, Albert. 1999. *Reading and writing the Lakota language*. Salt Lake City: University of Utah Press.
- White, Ralph Gardner. 1968. Borrowing and Taboo in Eastern Polynesia. *The Journal of the Polynesian Society* 77(1): 57–73.
- Yakubovich, Ilya. 2021. Hittite agent form. In: A. I. Kogan, A. S. Panina (eds.). *Problemy obshchej i vostokovednoj lingvistiki. Sochetamost' jazykovyx jedinit i jazykovye modeli. Pam'ati Z. M. Shal'apinoj (1946–2020)*: 124–141. Moskva: Institut vostokovedenija RAN.
- Yelisejeva, M. B., E. A. Vershinina, V. L. Ryskina. 2017. *Makarturovskij oprosnik: russkaja versija. Otsenka rechevogo i kommunikativnogo razvitiya detej rannego vozrasta. Normy razvitiya. Obraztsy analiza. Kommentarii* (The MacArthur Communicative Development Inventories: Russian version. Assessment of speech development in infants. Development norms. Examples of analysis. Comments). 2nd edition, revised and augmented. Ivanovo: LISTOS.
- Zhinkin, Nikolay. 1998. *Jazyk — rech — tvorcestvo: Issledovanija po semiotike, psixolingvistike, poetike* (Language, speech, creation: Research in semiotics, psycholinguistics, poetics). Moskva: Labirint.

С. А. Бурлак. О возможных источниках лексических замен и синонимии в праязыковых списках базисной лексики

При ступенчатой реконструкции встречаются ситуации, когда на одно и то же место в протосписке претендует более одного слова, и при этом нет возможности отвести один из потенциальных синонимов стандартными способами. Вместе с тем, в реально наблюдаемых языках практически всегда можно обосновать выбор слова для включения в список. Трудности проведения такого обоснования для праязыков могут быть связаны как с социолингвистическими, так и с внутриязыковыми причинами. Поскольку нам неизвестна социальная картина праязыка, а также её эволюция в дочерних популяциях, может не хватить материала для выводов о том, что слово, которое в праязыке было ограничено какими-то социальными ситуациями (например, употреблялось только одним из гендеров, или только по отношению к выше/ниже стоящему в иерархии, или только в разговоре с детьми и т. п.), могло при изменении общественного уклада стать основным в какой-то из дочерних ветвей. К увеличению числа синонимов в языке ведёт практика табу: в каждый конкретный момент основным является одно из слов, но в разные моменты это может различаться, так что одни из дочерних популяций унаследуют в качестве основного одно слово, другие — другое. Кроме того, в дочерних ветвях в разные стороны могли пойти фонетические или морфологические выравнивания, так что основы, прежде составлявшие одну (возможно, супплетивную) парадигму, стали единственными представителями соответствующего значения в каждой из дочерних ветвей. Чем больше времени прошло от разделения праязыка, тем труднее обнаружить те чередования, которые вызвали первоначальное различие между наблюдаемыми формами, выглядящими как несводимые друг к другу. Возможно, главная труд-

ность для выделения основного слова состоит в том, что в праязыке могла быть иначе устроенная номинационная решётка, так что те контексты, которые нынешнему наблюдателю представляются однородными, стали считаться таковыми лишь во время существования языков-потомков, при этом одни из них в качестве основного взяли одно из слов, другие — другое.

Ключевые слова: стословный список Сводеша; синонимия; лексические замены; сравнительно-исторический метод; семантическая реконструкция; языковое родство; лексикостатистика.

*Elena Shvedova**, *Yuri Koryakov***, *Elizaveta Zabelina****

* HSE University, Linguistic Convergence Laboratory (Moscow); shvedovalena98@gmail.com

** HSE University, Linguistic Convergence Laboratory (Moscow); Institute of Linguistics of the Russian Academy of Sciences (Moscow); ybkoryakov@gmail.com

*** HSE University, Linguistic Convergence Laboratory (Moscow); Institute of Linguistic Studies of the Russian Academy of Sciences (St.-Petersburg); windmill161@gmail.com

Quantifying lexical distances among Nudiz, Mahmudi, and Verin Dvin Urmī (North-Eastern Neo-Aramaic)¹

This study documents and analyzes lexical data from four Christian North-Eastern Neo-Aramaic varieties: Mahmudi, Nudiz, Verin Dvin Urmī, and Urmia Urmī, focusing on the previously undescribed Mahmudi and Nudiz. We provide correspondences from these lects for an extended 226-item basic vocabulary list collected for this study with etymologies, cognates from earlier Aramaic, and loanword sources. Cognate share calculations reveal that all four varieties belong to a single language. Notably, Mahmudi and Verin Dvin Urmī — spoken in the same village for 67 years — exhibit stronger convergence than with their genealogical relatives (Nudiz and Urmia Urmī respectively), highlighting contact-driven divergence from inherited patterns.

Keywords: Semitic languages; Aramaic language; North-Eastern Neo-Aramaic; dialectology; lexicostatistics; language contact; lexical semantics.

1. Introduction

One of the approaches to language documentation and comparative linguistic studies is the compiling and analysis of basic vocabulary. As a result of those, one can establish regular sound correspondences, measure lexical distances between lects and eventually build a language classification. With this in mind, we provide 200+ word lists with etymological notes for four North-Eastern Neo-Aramaic varieties: Nudiz, Mahmudi, Verin Dvin Urmī and Urmia Urmī, with the focus on the first two, unstudied and highly endangered varieties. This study contributes to the subgrouping of Neo-Aramaic along with other recent works on Neo-Aramaic lexicology such as Barsky et al. 2018a on Mlahso, Barsky et al. 2018b on Turoyo, and Lyavdansky 2021 on Literary Christian Urmī, as well as more general works by Mutzafi (2005; 2006), Krotkoff (1985), etc.

North-Eastern Neo-Aramaic (NENA) constitutes a complex dialect continuum comprising approximately 150 distinct varieties. Historically these varieties were spoken by Christian and Jewish communities in northern Iraq, north-western Iran, and south-eastern Turkey. Christian NENA speakers commonly identify themselves as “Assyrians”. Throughout the 20th century the majority of NENA-speaking communities were displaced from their ancestral homelands due to persecution and are now dispersed across the globe (Khan 2011). As a result, nowadays

¹ The study is funded by the Russian Science Foundation, Project No. 24-28-01009 «Areal and typological description of the Neo-Aramaic varieties in Armenia». We are grateful to the speakers of Modern Aramaic languages who consulted us. We would also like to thank an anonymous reviewer for the valuable comments, and our colleagues Alexey Lyavdansky and Alina Russkikh for their help.

most NENA varieties are either extinct or critically endangered. Despite significant progress in their documentation in recent decades (see Khan 2011: 710 for overview), the internal classification of NENA remains poorly understood (Kim 2008). This is partly due to the underdocumentation or complete lack of documentation of many varieties, particularly those from south-eastern Turkey. Consequently, there is an urgent need to study these varieties and fill the gaps with reliable linguistic data, while such information is still accessible.

This paper examines two previously unstudied Christian NENA (Assyrian) varieties originating in south-eastern Turkey and belonging to the Hakkari NENA cluster (Khan 2018a: 191) — **Nudz** (abbreviated to Nud) and **Mahmudi** (abbreviated to Mahm). For the past 67 years Mahmudi has been spoken in a single village, Verin Dvin, Armenia, where Christian Urmi variety (further called VD Urmi) is also spoken. Since Mahmudi and **Verin Dvin Urmi** have been in close contact for decades, we include data on VD Urmi in our research. The data of spoken **Christian Urmi of Urmia, Iran**² (further called U Urmi), one of the most thoroughly described NENA varieties (Khan 2016), are also provided for comparison.

The paper is structured as follows. Section 2 includes an overview of the methods and materials used. Conventions of data representation are discussed in Section 3. Section 4 contains the list of words with etymological notes. Section 5 presents Results and Discussion.

2. Methods and materials

2.1. NENA varieties concerned in the paper

As mentioned in the Introduction, the varieties studied in the paper belong to the North-Eastern Neo-Aramaic subgroup of Aramaic languages. Aramaic languages form a distinct group within the Semitic family of the Afro-Asiatic macrofamily (Kogan 2015: 377). The earliest inscriptions in the Aramaic languages date back to the 9th century BCE, so the written history of Aramaic covers a period of almost 3000 years. Aramaic languages are traditionally divided into two main subgroups: Western and Eastern Aramaic, based on shared morphological innovations (Beyer 1986). There are also different classifications of Aramaic languages based on the time of their existence. According to Beyer (1986), attested Aramaic languages can be separated into three groups: Old Aramaic (850 BCE — 200 CE), Middle Aramaic (200 — 700 CE) and Modern Aramaic. However, none of the written Old and Middle Aramaic languages can be considered the direct ancestor of NENA varieties. Khan (2007: 8–12) compares Jewish and Christian spoken NENA with the earlier literary forms of Aramaic — Babylonian Talmudic Aramaic and Classical Syriac. The author discusses phonological, morphological and lexical features, and concludes that NENA varieties are “the descendants of vernacular forms of speech that existed for centuries side by side with, but concealed by, the literary forms of Aramaic” (Khan 2007: 11). The scheme in Figure 1 illustrates the genealogical relationships between Middle and Modern Aramaic languages following Hoberman (1989: 5).

The present paper focuses on North-Eastern Neo-Aramaic varieties, while the material of earlier Eastern Aramaic languages is used only for comparison in order to confirm the Aramaic origin of the words. Basic sociolinguistic information on the four NENA varieties in question — Nudz, Mahmudi, Verin Dvin Urmi and Urmia Urmi — is given below.

² In this paper, we use ‘Urmi’ as a linguonym, ‘Urmia’ to refer to the town and the lake in Iran, and ‘Urmiya’ to refer to the Aramaic-speaking village in Russia.

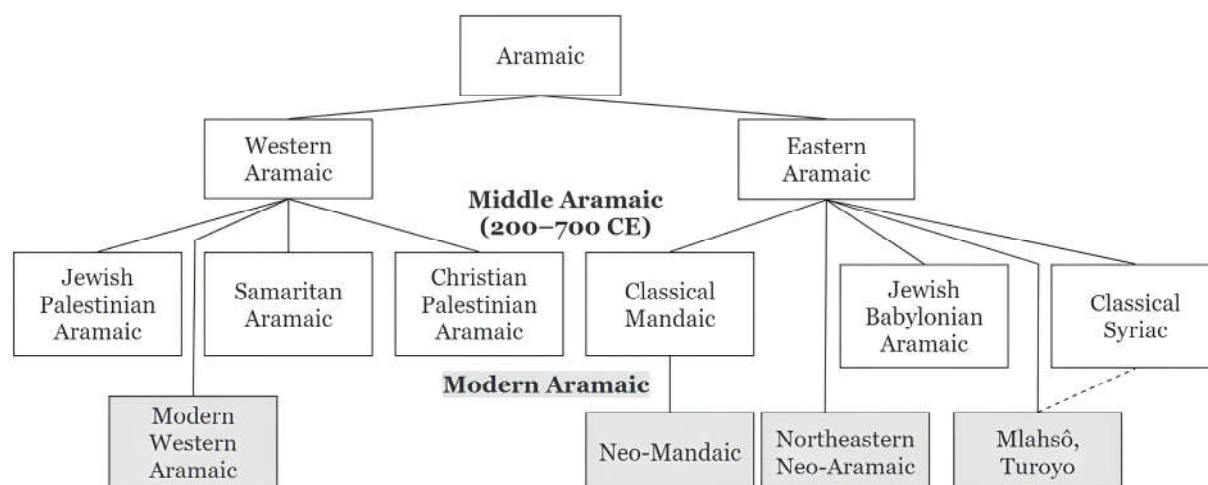


Figure 1. The genealogical relationships between Middle and Modern Aramaic languages following Hoberman (1989: 5)

The **Nudiz** variety (autonym: *lišan nudəzné*) is a Christian NENA variety formerly spoken in the Nudiz (Nördüz) area, in the Van District (Ottoman Empire) in today's south-eastern Turkey (Wilmschurst 2000: 292–293), see map in Figure 2. As a result of the massacres of 1915, all Nudiz speakers left their homeland and fled to other countries, including Syria (Talay 2008: 27) and the Russian Empire. In Russia, they were resettled in the Lutheran German colonies in the north-west of present-day Azerbaijan, along with Van and Bohtan speakers (Teule, Kessel 2012: 46). There they lived mainly in Helenendorf (also known as Yelenino or Khanlar, today's Göygöl), Grünfeld (Kalininkend, today's Vurğun), Annenfeld (Annino or Shamkhor, today's Shamkir), Georgsfeld (Georgiyevsk, Leninfeld or Leninkend, today's Çınarlı), Elizavetpol (Kirovabad or Ganja) (Sado 2017: 45–48). In 1949, the Nudiz Assyrians of Azerbaijan were sent into exile to the Tomsk Oblast, Siberia. They could not return until the late 1950s. While some managed to come back to Azerbaijan, others went to Georgia or Russia. Between 1980 and 1990, many Assyrians of Azerbaijan and Georgia moved to the south of Russia: Krasnodar Krai (in particular to the Assyrian village of Urmiya and the village of Dinskaya) and Stavropol Krai, due to the difficult situation in Transcaucasia. Today, the Nudiz variety is poorly documented, understudied, and highly endangered (Zabelina 2023). We are not aware of any compact settlement of Nudiz speakers, nor have we attested any case of on-going language transmission, the youngest speaker being 65 years old.

The **Mahmudi**³ variety (autonym: *lišan mamədə / mamədəy / bne gol*) was originally spoken by Christians from the Mahmudi Canton in the Van District (Ottoman Empire), located north of the Hakkari Mountains, in eastern Turkey. According to Wilmschurst (2000: 310–311), the ancestors of Mahmudi (and Van dialects) migrated there in the early 18th century from the Jilu and Txuma districts in the Hakkari Mountains. Nowadays, speakers of Mahmudi live in Armenia, and a group of Assyrians who originated in the same area as Mahmudi live in Georgia, where they are known as Saranaye (Sado 2021: 25). According to various unofficial estimates, Mahmudi speakers arrived in Armenia before the tragic events of the early 20th century, around the 1830s (Hovhannisyan 2007) or 1850s (Dum-Tragut 2013), and founded a village called Gyol Aysor. However, in 1958 they were resettled by the authorities to the village

³ We use the name Mahmudi based on the toponym in Badalyan (2015: 20–21). Other forms used elsewhere with reference to the region where the variety was originally spoken are: Mahmedāye, Serai d'Mahmideh (Wilmschurst 2000: 310–311), Saray-Maxmudia (Dum-Tragut 2013: 343) and Mamed (form mentioned by some speakers).

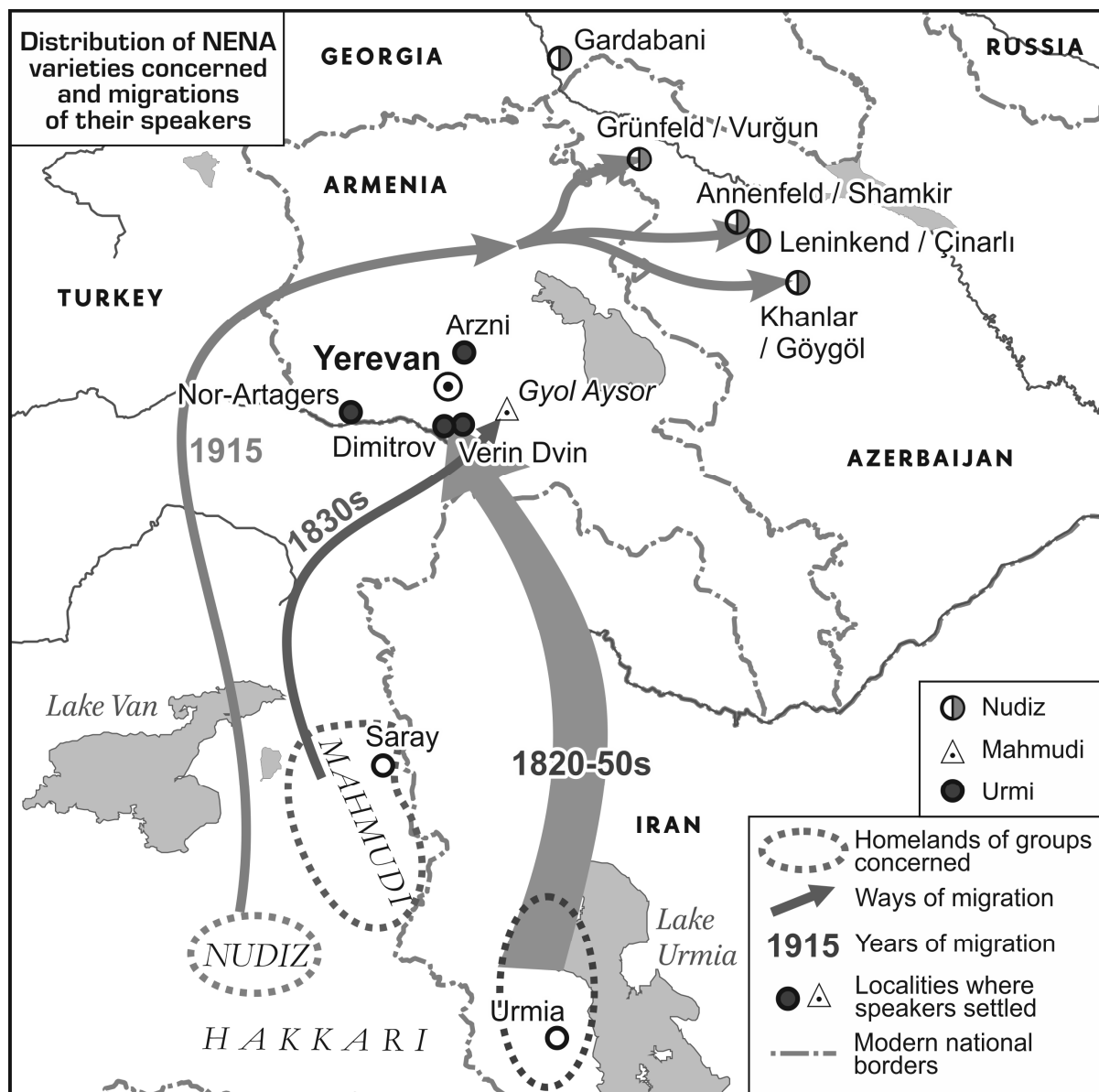


Figure 2. Map of distribution of NENA varieties in question and migrations of their speakers (by Yuri Koryakov)

of Verin Dvin, which was already inhabited by Christian Urmi speakers and Armenians, since the territory of Gyol Aysor became part of the Khosrov Forest State Reserve (Dum-Tragut 2013: 343). The Mahmudi variety has so far remained almost undocumented. Even though the Assyrian identity is common to both Christian Urmi and Mahmudi speakers, the latter have lived as a minority in Verin Dvin for many decades, which has led to a certain leveling: the gradual loss of specific features of their variety. Today, we can observe the phenomenon of dialect shift among Mahmudi speakers, most of whom are bilingual in Christian Urmi (and often also in Russian and Armenian), but not vice versa — Christian Urmi speakers know at most a few words specific to Mahmudi. We suggest that these sociolinguistic conditions have led to changes in Mahmudi and its convergence with Christian Urmi.

The **Verin Dvin Urmi** variety (autolinguonym: *lišanət urməžnáy*) is spoken in the village of Verin Dvin (Ararat Province). The Verin Dvin Assyrians are the largest community of Christian Urmi speakers in Armenia, with a total population of 1,862 in 2023 (AM: 317). Their village was founded no later than 1828 (year of the local church foundation (Dum-Tragut 2013: 348)) by the early Assyrian resettlers from the Urmia region in north-western Iran (Dum-Tragut 2013: 343). There are three other villages in Armenia where Christian Urmi has been

spoken since the 1830s: Arzni, Dimitrov and Nor-Artagers (Shagriyar); some speakers can be found in Yerevan and other big cities. However, the vitality status of Neo-Aramaic in localities other than Verin Dvin is lower.

Urmia Urmi (Christian Urmi of Urmia area) has been spoken west of Lake Urmia in Iran. It is the most extensively documented and studied NENA variety (Khan 2016), partly because it served as a foundation for a literary language (Murre-van den Berg 1999). While some Assyrian Christian communities can still be found in the Urmia region, most speakers migrated during the 19th and 20th centuries, settling in the Caucasus, North America, Europe, and Australia (Khan 2016.1). Today, Christian Urmi is spoken all over the globe, with local dialectal variations and numerous borrowings. Verin Dvin Urmi is one of these local varieties with specific regional features. However, Christian Urmi core vocabulary predictably remains stable. To demonstrate it, we provide lexical entries for Christian Urmi taken from Khan 2016.

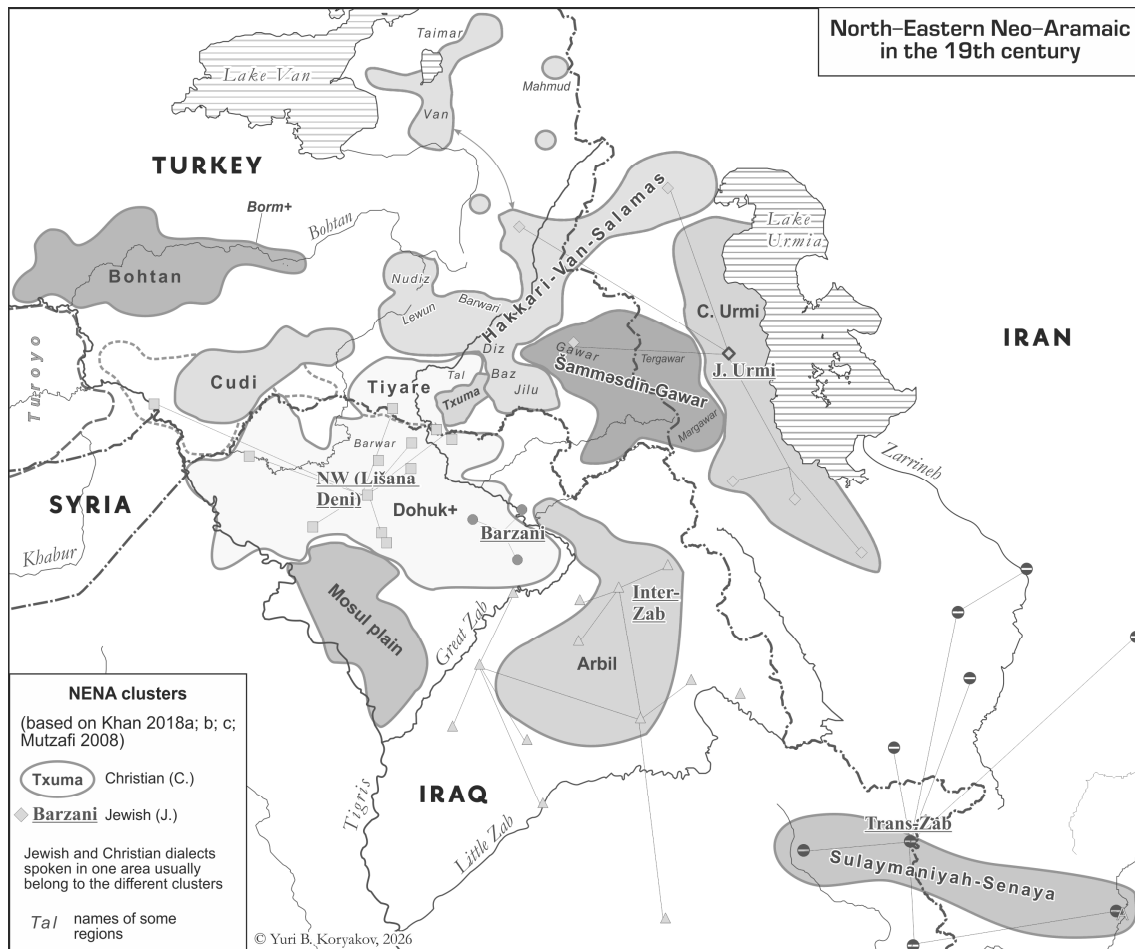


Figure 3. North-Eastern Neo-Aramaic in the 19th century (based on Khan 2018a; b; c; Mutzafi 2008; by Yuri Koryakov)

Significantly, both the Nudiz and Mahmudi varieties originally belong to the Hakkari cluster of NENA (Khan 2018a: 191), see Figure 3. According to our fieldwork observations, they share several diagnostic features in phonology and morphology. For instance, in Nud and Mahm, the reflex of **ū* merged with /u/ or fronted to [y]. Meanwhile, in the Urmi varieties, it was diphthongized as [uj] (Khan 2016.1: 25). Another remarkable phonological feature is the raising of /a/ in Nud and Mahm to mid or high vowels, cf. similar process in many Hakkari NENA (Fox 2015; Ovsjannikova et al., in print). In terms of morphology, one of the remarkable features is the present tense copula paradigm: in Nud and Mahm, the 1st person singular copula form is =*ana*, a characteristic feature in Van (Cereteli 1963) and surrounding areas: Sara

(*ənna*), Diz, Barwar, Timur (*na*), and Jilu (*na*) (Talay 2008: 202). In contrast, the Urmi varieties are characterized by the 1st person singular copula *=ən*. Another feature is that the cardinal numeral ‘one’ has a gender distinction: masculine *xe* vs. feminine *xda*, while the Urmi varieties have lost this distinction and the Urmi form *xa* is gender neutral. The gender distinction is attested in most NENA dialects originating in Turkey (Talay 2008: 423). A more detailed description of the differences and historical correspondences in phonology and morphology will be presented in our other papers.

Therefore, the Nudiz and Mahmudi varieties are expected to be linguistically closer to each other than to the Urmi varieties. However, different sociolinguistic changes of the past two centuries, including migrations and contacts with other languages, have caused linguistic changes. Nudiz speakers, after fleeing their homeland in 1915, lived in Azerbaijan among the speakers of other varieties of the Hakkari cluster, which might have helped preserve original traits. In contrast, Mahmudi has been spoken in Armenia since the middle of the 19th century, in close contact with Urmi, particularly during the last 67 years in Verin Dvin. Thus, we hypothesize that contemporary Mahmudi may exhibit both **retention** of archaic Hakkari features and contact-induced **convergence** with Urmi, even in its basic vocabulary.

2.2. Methods

Lexical distance is one of the measures of similarity between lects based on their lexicon. Lists of basic vocabulary are normally used for this measure, since cultural vocabulary tends to be less stable and prone to borrowing (Tadmor 2009: 64). There are several popular lists of basic vocabulary used in literature. For this paper, we used a set of 226 meanings based on a combination of three sets already widely used in linguistics: the Swadesh 100-meaning set (Swadesh 1955), the Swadesh 200-meaning set (Swadesh 1952), and the Leipzig-Jakarta 100-meaning set (Tadmor 2009). For a more precise choice of words, the semantic specifications and diagnostic contexts, defined in Kassian et al. 2010 for 100-item Swadesh list words and IE-CoR Meanings (created for Heggarty et al. 2023), were used.

We compiled the wordlists on the basis of the elicitation data gathered from the speakers of the NENA varieties, except for Urmia Urmi. Russian was used as an intermediary language. Two speakers of Nudiz, originating from Khanlar (today’s Göygöl) and Leninkend (today’s Shamkir) in Azerbaijan, provided translations during our meetings in Moscow in 2025. We also obtained translations from six speakers of Mahmudi and five speakers of Verin Dvin Urmi during two field trips to Verin Dvin in 2024 and 2025. The Urmia Urmi data were mainly taken from Khan’s dictionary (UD). In addition, in the case of close synonyms, we considered the frequency of words in the examples and/or texts (Khan 2016.4; Ovsjannikova et al. 2022–). We added notes like “(less frequent)” or “(more frequent)” wherever it was relevant. If both words are neutral and sufficiently frequent for a given meaning, they are provided as synonyms.

3. Data representation

3.1. Transcription

For Nud, Mahm and VD Urmi we use the transcription developed within several previous projects on NENA (Zabelina et al. 2022–; Ovsjannikova et al. 2022–). For U Urmi we keep the transcription of Khan 2016.

There is a three-way distinction for stops in all varieties concerned:

- voiced: [b, d, dʒ, ɟ/g];
- aspirated/lax: [p^h, t^h, tʃ^h, c^h/k^h];
- unaspirated/tense: [p, t, tʃ, c/k], marked with a dot in our transcription (ṗ, ṭ, ṭʃ, ṭʃ/k) and with a circumflex by Khan (ṗ̆, ṭ̆, ṭ̆ʃ, ṭ̆ʃ/k).

The voiceless unaspirated velar/postvelar stop [k] is outside this three-way distinction and is represented as *k* in VD Urmi and Mahm, *q* in Nud and *k̰* in U Urmi.

Other consonants are represented as follows.

- š, ž, ʏ, x, ɣ are used for [ʃ, ʒ, j, χ, ʁ].
- č, č̣, ġ (Khan's č, č̣, ġ) represent affricates [tʃ^h, tʃ, dʒ].
- The glottal stop [ʔ] is absent in Nud, Mahm and VD Urmi and is rendered with ' in U Urmi, following Khan 2016.
- The graphemes *w, v, s, z, h, m, n, l, r* are used as in IPA.

Realisation and representation/transcription of the **original velars** depend on the variety. In Nudiz, they retain their **velar** nature with slight palatalization in some contexts. In the northern dialects of the Urmi Plain, these consonants became **palatal** stops /c, c^h, ɟ/, but in the southern dialects they fronted further and became **affricates** /tʃ, tʃ^h, dʒ/ (Khan 2016.1: 109–112). The latter process also occurred in all Caucasian varieties of Urmi, including VD Urmi, so that they became identical to the original affricates č, č̣, ġ⁴. The situation in Mahmudi is heterogeneous: **both realisations** (affricate and velar) are common, so we provide explicit transcriptions for all Mahmudi words with these consonants in square brackets after the word. Realisation and representation/transcription of velars/palatals and affricates are summarized in Table 1.

Nud	Mahm	VD Urmi	U Urmi (Khan 2016)
Velar/palatal/affricates			
k [k ^h]	c [tʃ ^h /k ^h]	c [tʃ ^h]	c [tʃ ^h]/[c ^h]
k̰ [k]	ç [tʃ/k]	ç [tʃ]	ç̣ [tʃ]/[c]
g [g]	g [dʒ/g]	g [dʒ]	ɟ [dʒ]/[ɟ]
q [k]	k [k]	k [k]	k̰ [k]
Original affricates			
č [tʃ ^h]	č	č	č
č̣ [tʃ]	č̣	č̣	č̣
ğ [dʒ]	ğ	ğ	j

Table 1. Velars/palatals and affricates with IPA correspondences

The following vowel phonemes are distinguished: /i, e, ə, a, o, u/. The distribution and phonemic status of /ə/ warrant further research (Khan 2016.1: 49).

The suprasegmental emphasis is marked by the symbol “+” at the beginning of the word, following Khan 2016.1. The stress usually falls on the penultimate syllable, otherwise we indicate it by the acute accent.

⁴ Except in the village of Arzni (Armenia), where the former are realized as [tʃ^h, tʃ, dʒ] respectively.

3.2. Diagnostic forms

Nouns are given in singular except for pluralia tantum nouns. **Adjectives** are given in masculine singular.

For **verbs**, we provide the consonantal root and the inflection class (pattern⁵) according to the notation used by Khan (2016.1: 260–261). In Nudiz and Mahmudi, we also provide three basic forms (in parentheses) for each verb, the so-called Present, Progressive and Past stems, because the verbal systems of these varieties have not been described in the literature. A more detailed description of the functions and distribution of these forms can be found in Zabelina et al. 2022–. The Present and Progressive can appear as bare stems in some forms (3m.sg and under negation correspondingly), so we were able to collect stems without inflectional morphology. The Past stem always requires overt L-suffixes for subject indexing, and we provide the 3rd person plural forms as forms representing the past tense stem, since the 3rd person plural L-suffixes show additional variation between *-lun* and *-le*. For Urmi varieties, we do not provide the diagnostic forms for regular (so-called strong) verbs, because they can be built according to the models described by Khan (2016.1: 260–261). These models are given in Table 2. For irregular (weak) verbs, three forms are given for VD Urmi, and only Present is given for U Urmi, as it is the only form provided in UD.

Pattern	Present (3M.SG)	Progressive	Past (3PL)
I	C ₁ aC ₂ əC ₃ (<i>patax</i>)	bəC ₁ C ₂ aC ₃ a (<i>bəptaxa</i>)	C ₁ C ₂ əC ₃ -lun (<i>ptəx-lun</i>)
II	C ₁ aC ₂ əC ₃ (<i>basəm</i>)	C ₁ aC ₂ uC ₃ ə/a (<i>basumə</i> , marginal: <i>basuma</i>)	C ₁ uC ₂ əC ₃ -lun (<i>busəm-lun</i>)
III	maC ₁ C ₂ əC ₃ (<i>madməx</i>)	maC ₁ C ₂ uC ₃ ə/a (<i>madmuxə</i> , marginal: <i>madmuxa</i>)	muC ₁ C ₂ əC ₃ -lun (<i>mudməx-lun</i>)
QI	C ₁ aC ₂ C ₃ əC ₄ (<i>jarjəš</i>)	C ₁ aC ₂ C ₃ uC ₄ ə/a (<i>jarjušə</i> , marginal: <i>jarjuša</i>)	C ₁ uC ₂ C ₃ əC ₄ -lun (<i>jurjəš-lun</i>)

Table 2. Models for constructing diagnostic forms in Urmi varieties (Khan 2016.1). The following verbal roots were used as examples: *ptx* ‘to open’, *bsm* ‘to heal’, *mdmx* ‘to cause to sleep’, *jrjš* ‘to drag’

3.3. Structure of word entries

In the next section we provide the wordlists, where all the entries are organized as in example 1 below.

(1) Example of the word entry

MOTHER

-1 Nud **da**

-1 Mahm **da**

1 VD Urmi **yəmma**

1 U Urmi **yəmma** ‘mother’ (UD 342); cf. *çoç* ‘mother’ (UD 127); *+mama* ‘mummy’ (child language) (UD 227)

Source:

#-1 < KJ. *dê*, *deya*, *diya*, *diya*, *dya* ‘mother’ (KED 147).

#1 MEA: *emmā* ‘mother’ (SL 52); *immā* ‘mother’ (DJBA 116); *ɛma* ‘mother’ (MD 352).

⁵ Other terms mentioned in the literature are *stems*, *bases*, *forms*, *measures*, *derivation*. Cf. also terms in some other languages: *ʿawzān* in Arabic, *binyanim* in Hebrew, *stirpes* in Latin, *Stämme* in German, and *porody* in Russian.

The words from one variety for each meaning are given in a single line. The line for each language begins with the cognate index(es). Words with the same index are considered cognates for the purposes of further calculations. Negative indices mark borrowings, which are treated differently in the calculations. In the case of synonyms, several indices are given.

Words treated as basic for a certain meaning are given in bold, other words (in light italics after cf.) are provided for comparison. Variants derived from the same root are separated by comma, and words derived from different roots are separated by semicolon. Notes on usage and/or frequency in different dialects are given in parentheses. Glosses are given after words when their meaning does not completely coincide with the entry's meaning. Words under the U Urmi line are taken from Khan 2016; page number after the UD abbreviation and glosses from the dictionary are given for each word.

The etymological data are given in the second part of a word entry. For words of clear Aramaic origin, we provide related lexemes from one or more Middle Eastern Aramaic languages (after the siglum "MEA"): Classical Syriac (SL), Jewish Babylonian Aramaic (DJBA), Classical Mandaic (MD). There is no complete Proto-NENA reconstruction, but we provide PNENA etymons where they are available. Where we suppose that a NENA word is or may be a loanword, we provide possible candidates for the source (language + word + reference). In many cases it is unclear which language is the direct source of a borrowing, since similar words exist in many languages of the region (e.g. Azerbaijani, Kurdish and Persian). In such cases, words from several languages are provided.

4. List of words

This section contains all the wordlists, sorted by the alphabetical order of the entries (in English).

ALL

1 Nud **kullə**

1 Mahm **cullə** [kullə]

1 VD Urmi **cullə**

1 U Urmi **cullə** 'all; the whole' (UD 123)

Source:

MEA: *kul*, *kol* 'all, every' (SL 622); *kullā* 'all, every, entire, each' (DJBA 559); *kul* 'all, each, every, the whole, entire' (MD 206).

AND

1 Nud **u**

0 Mahm see Comments

-1 VD Urmi **i**, but see Comments

1 U Urmi **u** 'and' (UD 85); cf. *ham*, *+ham* 'also' (< Pers. *ham*) (UD 152); *mən*, *m-* 'with' (UD 218–219)

Source:

#1 MEA: *w* 'and' (SL 357); *w* 'and' (DJBA 394); *u* 'and' (MD 155).

#-1 < Russian *i* 'and'.

Comments:

In Mahm and VD Urmi, the most frequent strategy of coordination is juxtaposition; some speakers occasionally use different linkers such as *ham=da* ‘also=ADD’, *xa-xta* ‘also’ or comitative preposition.

ANIMAL

-1 Nud **+hayvan**

-1 Mahm **hayvanat**

-1, -2 VD Urmi **+hayvan (+hayvanat)** ‘animal, especially domestic one’; **+ğanavar** ‘animal, especially wild one’

-1 U Urmi **+hayvan, +heyvan** ‘animal, tame animal’ (UD 158); cf. *+janavar* ‘beast, predatory (animal), monster’ (UD 173)

Source:

#-1 < Kj. *h’eywan* ‘animal’ (KED 276) (or < Azer. *heyvan* ‘animal, brute, beast; cattle’ (AIL 339) < Pers. *heyvān* ‘animal; living being’ (ODP)) < Arab. *ḥayawān* ‘animal, beast’ (DMWA 220).

#-2 < Kj. *canewar* ‘wild animal, beast, creature’ (KED 83), but in Azer. *canavar* ‘wolf’ < Pers. *jān-var* ‘animal, beast’ (ODP).

ANT

1 Nud **šəkvana**

1 Mahm **šəcvana** [šəkvana]

1 VD Urmi **šəcvana**

1 U Urmi **šəcvana** ‘ant’ (UD 293)

Source:

Cf. MEA: *šušmānā* ‘ant’ (SL 1538), *šumšmānā*, *šuššəmānā* ‘ant’ (DJBA 1121), *šušmana* ‘ant’ (MD 458). Mutzafi (2004: 244) compares NENA forms with the Syriac *šəkbā* ‘louse’ (TS 576) since a direct phonetic correspondence between the MEA and NENA lexemes for ‘ant’ is not evident.

ASH(ES)

1 Nud **+qəṭma**

1 Mahm **+kəṭma**

1 VD Urmi **+kuṭma**

1 U Urmi **+kəṭma** ‘ash’ (UD 192)

Source:

MEA: *qəṭmā* ‘ash’ (SL 1353); *qīṭmā* ‘ash’ (DJBA 1011); *giṭma*, *gaṭma* ‘ash’ (MD 89).

AT / NEAR (PREP.)

1 Nud **kəs**

1 Mahm **cəs** [kəs]

1 VD Urmi **cəs**

1 U Urmi **cəs** ‘at, near, in the presence of, at the home of’ (UD 120)

Source:

< **kəsl-i* < **kəsla*/**kislā*? ‘loin flank, side of the body’ (Mutzafi 2006: 93–97; 2008: 357; Fassberg 2010: 283). MEA: *kslyn* ‘loins’ (DJPA 286).

BACK (N)

1 Nud **+xasa**

1 Mahm **+xasa**

1 VD Urmi **+xasa**

1 U Urmi **+xasa** ‘back; support, protection’ (UD 336)

Source:

MEA: *haššā* ‘hip, haunch’ (SL 482); *haršā* ‘loin’ (DJBA 484); *halša*, *haša* ‘hip, back’ (MD 122).

BAD

1 Nud **xərba**

-1 Mahm **pis**

-1 VD Urmi **pis**

1, -1 U Urmi **xərba** ‘bad, nasty, disgusting, mean, evil’ (UD 328); **pīs** ‘bad, worthless, disgusting, nasty’ (UD 256)

Source:

#1 < a noun, which has come to be used as an adjective (Khan 2016.1: 556); MEA: *harbā* adj. ‘deserted; rotten; wanting, needing; empty, without substance’, n. ‘destruction, devastation’ (SL 486); *harbtā* ‘ruin’ (DJBA 481); *harba*, *hirba* ‘destruction, devastation, ruin’ (MD 126).

#-1 < Kj. *p’is* ‘dirty, impure, foul’ (KED 465) or Azer. *pis* ‘bad’ (AIL 641; UD 256).

BARK (N)

1, 2 Nud **qalpət +qariya**, lit. ‘peel/skin of tree’; **gəldət +qariya**, lit. ‘skin/hide of tree’

1, 2 Mahm **kalpət ilana**, lit. ‘peel/skin of tree’; **gəldət kaysa** [gəldət kajsa], lit. ‘skin/hide of wood’

1, 2 VD Urmi **kalpət ilana**, lit. ‘peel/skin of tree’; **gəldət ilana**, lit. ‘skin/hide of tree’

-1 U Urmi **jallaxta** ‘bark of a tree or branch’ (UD 170); cf. *kalpa* ‘shell of nut, fruit or animal, pod, peel; bark’ (UD 177); *ǰalda* ‘skin, fur’ (UD 165)

Source:

Sources for *+qariya* & *ilana* see under TREE.

#1 MEA: *qlāptā* ‘bark, shell’ (SL 1375); *qlāptā* ‘peel, shell’ (DJBA 1022); *qlapta* ‘shell, hard casing’ (MD 413).

#2 see SKIN.

#-1 < Kurd. *çûlik* ‘skin, peel, shell, bark’ (DKF 318).

BECAUSE

-1 Nud **sabab**; cf. also general complementizer *t*

-1, -2, 1 Mahm **sabab**; **čuncət** [tʰuntʰət]; **ka-hada**

-1, -2, 1 VD Urmi **sab**; **čuncət**; **ka-but-hada**

-1, -2, 1 U Urmi **mən-sābab**, **m-sābab**, **sābab**, **sab** ‘because’ < *sābab*, *sabab* ‘cause’ (UD 272); **čuncət**, **čun** ‘because, since, as’ (UD 131); **ka-but-hada**, **kat-but-hada** ‘because’

Source:

#-1 < Pers. *sabab* ‘because’ < Arab. *sabab* ‘reason, cause, motive, occasion’ (DMWA 392).

#-2 < Azer. *çünki* ‘because, as, since, for’ (AIL 180), Sor. *çünke, çunku* ‘because’ (KED 111), Pers. *čon, čun, čun-ke* ‘because, since, as’ (ODP).

#1 < *ka* ‘to’ (prep.) (UD 175) + *but* ‘about’ (UD 107) + *hada, hadax* ‘thus’ (UD 151).

Comments:

In Mahm and VD Urmi, the most frequent strategy of causal clause subordination is juxtaposition; the subordinators listed above are used rarely.

BELLY

1 Nud **kisa**

1 Mahm **cisa** [tʰisa ~ kisa]

1 VD Urmi **cisa**

1 U Urmi **cisa** ‘belly (stomach + intestines)’ (UD 121); cf. *labba* (3) ‘heart; intention; stomach’ (UD 197)

Source:

MEA: *karsā* ‘belly, stomach’ (SL 655); *karsā* ‘stomach, rumen, womb’ (DJBA 603); *karsa* ‘belly, stomach; womb, uterus’ (MD 201).

BERRY

There is no specific word for this meaning in Nud, Mahm and VD Urmi. The cognates of the word *tuyta* mean ‘mulberry’.

1 U Urmi **tuyta** ‘berry, especially mulberry; mulberry tree’ (UD 309)

Source:

MEA: *tūta* ‘mulberry’ (SL 1635); *tuta* ‘mulberry’ (MD 484). Widespread word across the Near East from Akkadian *tuttu* to Sanskrit तूत (tūta), तूद (tūda); Semitic forms are likely borrowed from Iranian, but the ultimate origin is unknown.

BIG

1 Nud **+gora**

1 Mahm **+gora** [gora]

1, -1 VD Urmi **+gura; yacca** ‘big; very big (for some speakers)’

1 U Urmi **+jura** ‘big, great, old, senior, superior’ (UD 169); cf. *yacca* ‘huge, enormous, giant’ (UD 340)

Source:

#1 Probably < Kj. *gir, gur* ‘large, big’ (KED 213; UD 169).

#-1 < Azer. *yekə* ‘big, large, tremendous’ (AIL 841).

BIRD

1, 2 Nud **səpərra; +ṭayra**

1 Mahm **spərra**; cf. *+ṭera* ‘big bird; eagle’

1 VD Urmi **səpra**; cf. *+ṭera* ‘big bird; eagle’

2, 1 U Urmi **+ṭera** ‘bird’ (UD 317); **səpra** ‘sparrow, species of sparrow’ (UD 278); according to examples from the texts, the word *səpra* is more often used in the meaning ‘(little) bird’ (Khan 2016.4)

Source:

#1 MEA: *ṣəprā* ‘bird’ (SL 1298); *ṣiprā* ‘bird, fowl’ (DJBA 962); *ṣupra*, *ṣipra* ‘little bird, sparrow’ (MD 390; 394).

#2 MEA: *ṭayrā* ‘bird, birds’ (SL 528).

TO BITE

1 Nud **+qrṭ** I (+*qarəṭ* | +*bəqraṭa* | +*qrəṭle*)

1 Mahm **+krṭ** I (+*karəṭ* | +*bəkraṭa* | +*krəṭle*)

1 VD Urmi **+krṭ** I

1 U Urmi **+krṭ** I ‘to bite (not necessarily breaking the skin), to bite off, to gnaw, to pick (a bone)’ (UD 188); cf. +*nØs*, +*nys* I ‘to sting, to bite (harmful bite of animal penetrating flesh); to be bitten’ (UD 247)

Source:

MEA: *qrṭ* pe. ‘to gnaw, to wound with the beak’ (SL 1405); *ṭrq* pe. ‘to hit, sting, bite’ (DJBA 519); *qrṭ* pa. ‘to take a bite’ (DJPA 576).

BITTER

1 Nud **+mayra** ‘bitter, hot, spicy’

1 Mahm **+mayra**; cf. *sirupa* ‘spicy’

1 VD Urmi **+mayra**; cf. +*saripa* ‘spicy’

1 U Urmi **+mayra** ‘bitter’ (UD 236); cf. +*saruypa* ‘pungent (spice); biting (cold)’ (UD 283)

Source:

MEA: *marrīrā* ‘bitter’ (SL 834); *mārīr* ‘bitter, cruel’ (DJBA 710); *mrira* ‘bitter, bad’ (MD 278).

BLACK

1 Nud **kuma**

1 Mahm **cuma** [kuma]

1 VD Urmi **cuma**

1 U Urmi **cuma** ‘black’ (UD 124)

Source:

MEA: *ʔukkāmā* ‘black; pupil of the eye’ (SL 15); *ʔukkām* ‘black; pupil, algae’ (DJBA 88); *ʔukma* ‘blackness; darkness’ (MD 343). Cf. *kōmā* ‘black’ (SL 608), the Syriac word is attested only in the “Book of Medicines,” which contains folk medicine recipes notoriously replete with Neo-Aramaic lexis and may be a borrowing from Neo-Aramaic (Ljavidanskij 2014: 53).

BLOOD

1 Nud **dəmma**

1 Mahm **dəmma**

1 VD Urmi **dəmma**

1 U Urmi **dəmma** ‘blood’ (UD 143)

Source:

MEA: *dmā* ‘blood’ (SL 307); *dmā* ‘blood’ (DJBA 340); *dma*, *adma* ‘blood’ (MD 111, 8). See also Khan 2016.1: 199 on the gemination after an epenthetic *ə.

TO BLOW

1 Nud **np̄x** I (*nap̄x* | *bənp̄axa* | *p̄axle*)

1 Mahm **np̄x** I (*nap̄x* | *bənp̄axa* | *np̄axle*)

1 VD Urmi **np̄x** I

1 U Urmi **np̄x** I ‘to blow, to blow away’ (UD 241); cf. *nšm* I ‘to blow (wind)’ (UD 241)

Source:

MEA: *np̄h* pe. ‘to blow, blow into’ (SL 929); *np̄h* pe. ‘to blow, inflate’ (DJBA 760); *np̄a*, *np̄h* pe. ‘to breathe out, fan with breath, blow up, inflate’ (MD 303), *nhp* pe. ‘to blow up (a flame), fan (a flame), breathe (fire)’ (MD 291).

BONE

1 Nud **gərma**

1 Mahm **gərma** [dʒərma]

1 VD Urmi **gərma**

1 U Urmi **jarma** ‘bone’ (UD 163)

Source:

MEA: *garmā* ‘bone’ (SL 261); *garmā* ‘bone’ (DJBA 302); *girma* ‘bone, knuckle’ (MD 92).

BREAST (CHEST)

-1 Nud **+sadra**

-1 Mahm **+sadra**

-1 VD Urmi **+sadra**

-1 U Urmi **+sadra** ‘breast, chest’ (UD 281)

Source:

< Arab. *ṣadr* ‘chest, breast’ (DMWA 507); also Pers. *ṣadr* ‘breast’ (CPED 783). Mutzafi 2008: 382 supposes that it is an old loan.

BREAST (TIT)

1 Nud **bəzza**

1 Mahm **bəzza**

1 VD Urmi **bəzza**

1 U Urmi **bəzza** ‘nipple, teat; breast’ (UD 103)

Source:

MEA: *bezzā* ‘breast, teat’ (SL 132).

TO BREATHE

-1 Nud **napas/nafas** (n) **grš** I, lit. ‘to pull the breath’

-1 Mahm **napas** (n) **grš** [g- ~ dʒ-] I, lit. ‘to pull the breath’

-1 VD Urmi **napas** (n) **grš** I / **škl** I, lit. ‘to pull / to take the breath’

-1 U Urmi **napas** (n) **jrš** I / **škl** I ‘to breathe’ (UD 240)

Source:

Appears to be a calque from Kj. *nefes kišandin* ‘to suck in (air), to inhale, to breathe’, lit. ‘breath, respiration’ + ‘draw, pull, extract’ (DKF 1078) with direct borrowing of nominal part < Kj. *nefes* ‘breath, respiration’ < Arab. *nafas* ‘breath, whiff; puff’ (DMWA 985; UD 240). Also cf. MEA: *npš* pe. ‘to breathe out, exhale’ (SL 937) denominative < *naḫšā* ‘soul, person, self’ (SL 938).⁶

TO BURN (TRANS.)

1 Nud **mØqd** III (*miqəd* | *muqúd* | *muqədle*)

1 Mahm **mØkd** III (*mikəd* | *mukuda* | *mukədle*)

1 VD Urmi **mkØd**, **mØkd** III (*makkəd*, *makəd* | *makuda* | *mukədle*)

1 U Urmi **mkØd**, **mØkd** III (*maḱḱəd*, *maḱəd*) ‘to cause to burn, to cause to be consumed by fire, to set on fire; to heat (oven), to scorch; to fire (clay)’ (UD 205)

Source:

MEA: *yqd* af. ‘to burn s.t.’ (SL 580); *yqd* af. ‘to set on fire’ (DJBA 540); *yqd* pe. ‘to burn (intr.)’ (MD 193).

TO CARRY

1 Nud **lbl** II (*libəl* | *lubúl* | *lubəlle*)

1 Mahm **lbl** II (*libəl* | *lubula* | *lubəlle*)

1 VD Urmi **lbl** II

1 U Urmi **lbl** II ‘to take away, to carry away, to carry, to bear, to conduct, to overcome’ (UD 195)

Source:

MEA: *ybl* pa. ‘to lead, conduct forward; to deliver’; af. ‘to lead, move; to bring before; to bear’ (SL 560); *ybl* af. ‘to carry, bring’ (DJBA 522); *ybl* pe., af. ‘to carry a load, bear, bring’ (MD 188).

CHILD

1, 2 Nud **yal(a)** sora ‘boy; child’

1, 2 Mahm **yala sura** [jala syra] ‘boy; child’

1, 2 VD Urmi **yala sura** ‘boy; child’

1, 2 U Urmi **yala sura** ‘child’ (UD 340) < *yala* ‘boy’ + *sura* ‘small, young’ (UD 280). According to examples from the texts, the meaning ‘child’ can also be conveyed by the word *yala* ‘boy, son, adolescent, young man’ (UD 340).

Source:

#1 old loan < Arab. *ḡayyil* (sg.), *ḡiyāl* (pl.) ‘family, household’ (Mutzafi 2008: 399; DMWA 658), cf. MEA *ialda* ‘boy, child, boy-child’ (MD 186).

#2 see SMALL

⁶ We acknowledge the reviewer’s suggestion that the entire construction could be analyzed as a calque.

CHILD_KIN

1, 2 Nud **kulpat** ‘one’s child (kin), offspring’; **yala sora** ‘boy; child’

2 Mahm **yala sura** [jala syra] ‘boy; child’

2 VD Urmi **yala sura** ‘boy; child’

2 U Urmi **yala sura** ‘child’ (UD 340); cf. +*zarra* ‘descendant, offspring’ (UD 351); *calpat*, *culpat* ‘nuclear family; wife’ (UD 119)

Source:

#1 < Kj. *k’ulfet* ‘wife; woman; family, children’ (KED 335) or < Azer. *külfət* ‘family; household; wife’ (Hc 448); both < Arab. *kulfa* ‘discomfort, trouble, inconvenience; lady’s maid’ (DMWA 837; UD 119).

#2 see CHILD

CLOUD

1 Nud **+ayva**

1 Mahm **+ayva**

1 VD Urmi **+ayva**

1 U Urmi **+ayva** ‘cloud, rain cloud’ (UD 92)

Source:

MEA: *ṣēḇā* ‘dark cloud, cloudiness’ (DJBA 850); *aiba* ‘cloud, fog, mist, darkness’ (MD 14); cf. *aybā* ‘thick clouds’ (TS 411); “the Syriac word is attested only in medieval Syriac lexicons and may be a borrowing from Neo-Aramaic” (Barsky et al. 2018b: 82).

COLD (NOT ONLY WEATHER)

1 Nud **+qayra**

1 Mahm **+kayra**

1 VD Urmi **+kayra**

1 U Urmi **+kayra** ‘cold’ (UD 191)

Source:

MEA: *qarrirā* ‘cold’ (SL 1409); *qarrīr* ‘cold’ (DJBA 1043); *qarir* (MD 403). See also Khan 2016.1: 548 on consonant dissimilation in **qarrira* > *+kayra*.

TO COME

1 Nud **ØyØ, Øšy** I (*ṭe* | *bəššá* | *šile*)

1 Mahm **ØyØ, Øšy** I (*ṭe* | *bəššá* | *šile*)

1 VD Urmi **Øty, yty** I (*atə* | *bitaya* | *tile*)

1 U Urmi **Øty, yty** I (*atə* | *bitaya*, *bitá* | *tilun*) ‘to come, to arrive; to come back, to return’ (UD 79; Khan 2016.1: 355)

Source:

MEA: *ʔty* pe. ‘to come’ (SL 110); *ʔty* pe. ‘to come’ (DJBA 176); *ata* pe. ‘to come, arrive’ (MD 41).

Comments:

On the specific shift **t* > *š* in some Hakkari varieties, including Nud, Mahm and Upper Tiyyari see Talay (2008: 68–69), Zabelina (2023).

CORRECT (RIGHT)

- 1 Nud **duz** [dyz]
 -1 Mahm **duz** [dyz]
 -1 VD Urmi **duz**
 -1 U Urmi **dūz** ‘straight, flat; right, upright; correct, exact’ (UD 146)

Source:

- # < Azer. *düz* (AIL 243; UD 146) or < Kj. *dûz* ‘flat, even, level, straight’ < Turkish *düz* ‘flat, straight’ (KED 170).

TO COUNT

- 1 Nud **bny** I (*banə* | *bəbna* | *bnile*)
 1 Mahm **bny** I (*banini* | *bnə* | *bnile*)
 1 VD Urmi **bny** I (*banini* | *bnaya* | *bnile*)
 1 U Urmi **mny** I (*manə*) ‘to count’ (UD 207); **bny** I (*banə*) ‘to count’ (UD 99); cf. *xžbn* QI ‘to reckon, to count, to estimate, to calculate’ (UD 327)

Source:

- # MEA: *mny* pe. ‘to count, number’ (SL 781); *mny* pe. ‘to count’ (DJBA 686); *mna* pe. ‘to count, number, reckon’ (MD 274).

TO CRUSH

- 1 Nud **+blç** I (*+baləç* | *+bəblaça* | *+bləçle*); cf. **+mrç** I ‘to step on’
 1, 1, 2 Mahm **+blç** I (*+baləç* | *+blaça* | *+bləçle*); **bls** I (*baləs* | *blasa* | *blasla*); **+mrç** I (*+marəç* | *+mraça* | *+mrəçle*)
 1, 2 VD Urmi **+blç** I; **+mrç** I
 1, 1, 2 U Urmi **+blç** I ‘to crush, to mash (with fist); to smash, to crunch’ (UD 109); **bls** I ‘to bruise (tr. and intr.), to crush; to stamp on; to pulp, to mash, to crumple, to wrinkle’ (UD 98); **+mrç** I ‘to squash, to crush by squeezing with fingers; to become crushed; to press (to extract juice)’ (UD 229)

Source:

- #1 MEA: *pls* pa. ‘to demolish’ (DJBA 913); “A ... sound symbolism is identifiable in the following doublets, in which the verb expressing the more forceful action is distinguished from the other member of the pair by being emphatic and exhibiting the affrication of the sibilant /s/ to /ç/: *baləs* I ‘to bruise’ *+baləç* I ‘to crush, to mash (with fist)’” (Khan 2016.1: 182–183).
 #2 MEA: *mrs* pe. ‘to press; to crush’ (SL 836); *mrs* pe. to crush, rub (DJBA 710), *mrs* pe. to compress, press down, bruise, crush down, squash (MD 279). “Affrication of stops and sibilants in emphatic settings have developed in a number of ... verbs expressing forcefulness or hardness, e.g. *+marəç* I ‘to squash’ < **marəs* (cf. Syr. *mrās*)” (Khan 2016.1: 183).

TO CUT

- 1 Nud **prm** I (*parəm* | *bəprama* | *prəmle*); cf. **+kṭkṭ** QI (*+qəṭqəṭ* | *+qəṭqúṭ* | *+qəṭqəṭle*) ‘to chop/cut into many pieces’
 1, 2 Mahm **prm** I (*parəm* | *prama* | *prəmle*); **+qṭØ**, **+qṭy** I (*+qəṭṭini* | *+bəqṭa* | *+qṭile*); cf. **+kṭkṭ** QI (*+kaṭkəṭ* | *+kuṭkuṭa* | *+kuṭkəṭle*) ‘to chop/cut into many pieces’

2 VD Urmi **+kṭØ, +kṭy** I (+kaṭṭə | +bəkṭaya | +kṭile); cf. *prm* I ‘to slaughter’; +kṭkṭ QI ‘to chop/cut into many pieces’

1, 2 U Urmi **prm** I ‘to cut; to cut off, to chop off; to hew’ (UD 250); **+kṭØ, +kṭy** I (+kaṭṭə) ‘to cut, to cut off; to be cut off (intr.); to cut down (tree); to chop off’ (UD 190); cf. +kṭkṭ QI ‘to chop; to cut up, to split (firewood), to hew into pieces’ (UD 190)

Source:

#1 MEA: *prm* pe. ‘to cut up’; pa. ‘to cut apart, tear to pieces’ (SL 1242); *prm* pe. ‘to chop, hack down’ (DJBA 934).

#2 MEA: *qṭṭ* pe. ‘to cut, to cut off’ (SL 1354); *qṭṭ* pe. ‘to bite off’; pa. ‘to cut off; to kill, execute’ (DJBA 1007); *gṭa* pe. ‘to cut (off), sever’ (MD 87).

DAY

1 Nud **yuma**

1 Mahm **yuma**

1 VD Urmi **yuma**

1 U Urmi **yuma** ‘day; sun’ (UD 342)

Source:

MEA: *yaṭwā* ‘day’ (SL 568); *yōmā* ‘day, sun’ (DJBA 529); *ium* ‘day’ (MD 274).

TO DIE

1 Nud **myy, myt** I (*mayə* | *bəmyá* | *mətle*)

1 Mahm **myt** I (*met* | *myata* | *mətle*)

1 VD Urmi **myt** I (*mat* | *myata* | *mətle*)

1 U Urmi **myt** I (*mayət, māt*) ‘to die’ (UD 217)

Source:

MEA: *mwt* pe. ‘to die’ (SL 731); *mwt* pe. ‘to die’ (DJBA 650); *mut* pe. ‘to die’ (MD 263).

TO DIG

1 Nud **xṗr** I (*xapər* | *bəxpara* | *xṗire*)

1 Mahm **xṗr** I (*xapər* | *bəxpara* | *xṗerre*)

1 VD Urmi **xṗr** I

1 U Urmi **xṗr** I ‘to dig, dig up, to loosen (earth)’ (UD 323)

Source:

MEA: *ḥpr* pe. ‘to dig’ (SL 482); *ḥpr* pe. ‘to dig’ (DJBA 478); *hpr* pe. ‘to dig’ (MD 151).

DIRTY

1 Nud **šəxtana**

1 Mahm **šəxtana**

1 VD Urmi **šəxtana**

1 U Urmi **šəxtana** ‘dirty, soiled, easily soiled, e.g. dress; impure, ethically corrupt’ (UD 295)

Source:

<šəxta ‘dirt’; MEA: *šuḥtā* ‘rust, patina’ (SL 1524); *šuta* ‘rust, verdigris’ (MD 458)

+ adjectival derivational suffix *-ana* (Khan 2016.1: 448).

TO DO

1 Nud **Øvd** I (*ʔavəd* | *vada* | *vədle*)

1 Mahm **Øvd** I (*avəd* | *vada* | *vədle*)

1 VD Urmi **Øvd** I (*avəd* | *vada* | *vədle*)

1 U Urmi **Øvd** I (*avəd* | *vada* | *vədlun*) ‘to do; to make; to perform, to create’ (UD 80; Khan 2016.1: 320)

Source:

MEA: *ʔbd* pe. ‘to work; to make, do, produce, complete’ (SL 1054); *ʔbd* pe. ‘to do, make, produce, prepare’ (DJBA 836); *abd*, *ʔbd*, *awd*, *ʔwd* pe. ‘to do, perform, do work, act, make’ (MD 2).

DOG

1 Nud **kəlba**

1 Mahm **cəlba** [tʰəlba]

1 VD Urmi **calba**

1 U Urmi **calba** ‘dog’ (UD 116)

Source:

MEA: *kalbā* ‘dog’ (SL 622); *kalbā* ‘dog’ (DJBA 580); *kalba* ‘dog’ (MD 197).

TO DRINK

1 Nud **šty** I (*šat* | *bəštā* | *štīle*)

1 Mahm **šty** I (*šatini* | *bəštā* | *štīle*)

1 VD Urmi **šty** I (*šat* | *bəštaya* | *štīle*)

1 U Urmi **šty** I (*šatə*) ‘to drink’ (UD 291)

Source:

MEA: *šty* pe. ‘to drink’ (SL 1614); *šty* pe. ‘to drink’ (DJBA 1184); *šta* pe. ‘to drink, imbibe, be satisfied, refreshed’ (MD 476).

DRY

1 Nud **bīruza** [biryza]

1 Mahm **bīruza** [biryza]

1 VD Urmi **baruyza**

1 U Urmi **baruyza**, **briza** ‘dry’ (UD 100)

Source:

< *barəz* I ‘to dry (intr.)’ (UD 99). The further etymology is uncertain. Cf. Arab. *barāz*- ‘plain, or wide expanse of land, without trees’ (Lane 1863: 187; Mutzafi 2008: 340) or Kj. *bêzuret*, *bêzuret* ‘sterile, barren, infertile’ (KED 59).

DULL

There is no specific word for this meaning in Nud, Mahm and VD Urmi, the meaning is expressed by construction “not to be sharp”.

-1 U Urmi **cut** ‘blunt’ (UD 124)

Source:

< Azer. *küt* ‘blunt; dull’ (UD 124; AIL 451)

DUST

-1 Nud **+toz**

-1 Mahm **+toz**

-1 VD Urmi **+toz**

-1 U Urmi **+toz** ‘dust’ (UD 313)

Source:

< Kj. *t’oz* ‘dust; powder; dirt’ (KED 628) or < Az. *toz* ‘dust; powder’ (AIL 771).

EAR

1 Nud **naviyya**

1 Mahm **naviyya**

1 VD Urmi **nata**

1 U Urmi **nata** ‘ear’ (UD 242)

Source:

naviyya < **naØiØa* < **naṭiṭa* < **ʔannāhātā* < **ʔēdnāhātā* (Mutzafi 2005: 84);

MEA: *ʔednā*, pl. *ʔednātā* ‘ear’ (SL 10); *ʔūdānā* ‘ear’ (DJBA 85); *ʔudna* (MD 342); cf. 17 cent. Telkepe *nhāṭa* < Syr. *ʔednāhātā* (Mutzafi 2005: 84; Mutzafi 2008: 366).

EARTH

1 Nud **upra**

1 Mahm **upra**

1 VD Urmi **upra**

1 U Urmi **upra** ‘earth, soil, clay; land’ (UD 85)

Source:

MEA: *ʔapṛā* ‘dust; earth, soil’ (SL 1124); *ʔapṛā* ‘earth, dust, powder’ (DJBA 875); *apra* ‘dust, ashes’ (MD 32).

TO EAT

1 Nud **+Øxl, +yxl** I (+*taxəl* | +*bixala* | +*xəlle*)

1 Mahm **+Øxl, +yxl** I (+*axəl* | +*bixala* | +*xəlle*)

1 VD Urmi **+Øxl, +yxl** I (+*axəl* | +*bixala* | +*xəlle*)

1 U Urmi **+Øxl, +yxl** I (+*axəl* | +*bixala* | +*xəllun*) ‘to eat, to be fed; to consume’ (UD 91; Khan 2016.1: 289)

Source:

MEA: *ʔkl* pe. ‘to eat’ (SL 41), *ʔkl* pe. ‘to eat’ (DJBA 129); *akl, ʔkl* pe. ‘to eat’ (MD 16).

EGG

1 Nud **bita**

1 Mahm **bita**

1 VD Urmi bita**1 U Urmi bita** ‘egg’ (UD 105)

Source:

MEA: *bēṣtā* ‘egg’ (SL 143); *bēṣtā, bēṣtā* ‘egg’ (DJBA 204); *baia, biia, bita* ‘egg’ (MD 47, 60, 64).

EYE

1 Nud +ayna**1 Mahm +ayna****1 VD Urmi +ayna****1 U Urmi +ayna** ‘eye’ (UD 92)

Source:

MEA: *ṣaynā* ‘eye’ (SL 1097); *ṣēnā* ‘eye, sight’ (DJBA 855); *ayna* ‘eye’ (MD 15).

TO FALL

1 Nud lpl, ypl, Øpl I (*lapəl* | *bipala* | *pəlle*)**1 Mahm npl I** (*napəl* | *bənpala* | *pəlle*)**1 VD Urmi npl I** (*napəl* | *bənpala* | *npəlle, pəlle*)**1 U Urmi npl I** (*napəl*) ‘to fall, to fall down’ (UD 240)

Source:

MEA: *npl* pe. ‘to fall’ (SL 931); *npl* pe. ‘to fall’ (DJBA 761); *npl* pe. ‘to fall’ (MD 303).

FAR

1 Nud rəqqa**1 Mahm rəkka****1 VD Urmi rəkka****1 U Urmi rəxka** (northern Urmi variety), **rəkka** (southern Urmi variety) (adj.) ‘distant, remote, far’ (UD 266)

Source:

< noun **rəḥqā* (Khan 2016.1: 170); cf. MEA: *rḥq* ‘to go away’ (SL 1458); *rḥq* ‘to be far away’ (DJBA 1071); *rḥq* ‘to be far’ (MD 427).

FAT

1 Nud məšxa**1 Mahm məšxa**; cf. *tarba* ‘fascia (layer of loose tissue, containing fat, between meat and skin)’**1 VD Urmi məšxa**; cf. *tarba* ‘fascia’**2 U Urmi tarba** ‘fat, tallow, lard’ (UD 305); cf. *məšxa* ‘oil, fat’ (UD 220); *čəzva* ‘rendered fat of tail used for cooking’ (UD 131)

Source:

#1 MEA: *mešhā* ‘oil; fat, tallow’ (SL 844); *mīšhā* ‘oil, fat’ (DJBA 673); *miša* ‘oil, unguent; fat (melted down), butter’ (MD 270).#2 MEA: *tarbā* ‘fat’ (SL 1663); *tarbā* ‘fat’ (DJBA 1230); *tirba* ‘fat of animals’ (MD 486).

FATHER

1 Nud **baba**

1 Mahm **baba**

1 VD Urmi **baba**; cf. *ḡaḡḡuna* ‘father’ (in Arzni, Armenia) < Kj. *kek, kak, kekê* ‘elder brother’ (KED 303; UD 127)

1 U Urmi **baba** ‘father; grandfather’ (UD 96)

Source:

A widespread onomatopoeic word. MEA: *bb?* ‘dad, father’ (onom.) (SL 115); *baba* ‘father’ (MD 45); cf. Kj. *bav, bab, baba* (KED 26); Az. *baba* ‘grandfather’ (AIL 82).

TO FEAR

1 Nud **+zdy** I (+*zadā* | +*bəzdā* | +*zdile*)

1 Mahm **+zdy** I (+*zadini* | +*bəzdā* | +*zdile*)

1 VD Urmi **+zdy** I (+*zadini* | +*bəzdaya, bəzdā* | +*zdile*)

1 U Urmi **+zdy** I (+*zadā*) ‘to fear, to be afraid, to be timid, to be anxious’ (UD 349)

Source:

Cf. *t*-stems of the root *zwʕ* in MEA: *zwʕ* ettaf. ‘to be shaken, terrified’ (SL 374); *zwʕ* ethpe. ‘to be removed’; ethpa. ‘to tremble, be moved’ (DJBA 405). For Jewish Betanure Mutzafi (2008: 400) suggest the following chain of changes: *zd?* < **zdʕ* < *(?*ə*)*dziʕ* < *?*ətziʕ* (*t*-stem of the root *zwʕ*).

FEATHER

-1 Nud **parra**

-1 Mahm **parra**

-1 VD Urmi **parra**

-1 U Urmi **parra** ‘feather’ (UD 251)

Source:

< Pers. *par, parr* ‘a wing, a feather’ (CPED 239); Kj. *p’er* ‘feather, wing’ (KED 439); Azer. *pər* ‘wing’ (AIL 638).

FEW / LITTLE

1 Nud **čumma** ‘few, little’

2 Mahm **xačča**

2 VD Urmi **xačča**

2 U Urmi **xačča** (mod., adv.) ‘a little, a few, some; slightly’ (UD 321); cf. *xa-pərça* ‘a little’ < *pərça* ‘small quantity, small piece’ (UD 262); *zarra* n.m. ‘a little, a jot’ (< Azer. *zərrə* < Arab.) (UD 346); *çam, çam* part. ‘little, less, seldom; wanting’ (UD 127)

Source:

#1 Etymology is unknown; cf. Azer. *kəm*, Pers. *kam*, Kj. *kêm* ‘few, small in number, little, a little bit, a small amount’ (KED 317).

#2 < **xa* (see ONE) & Turkic suffix *-ča* (Khan 2016.1: 181).

TO FIGHT

1 Nud **+płš** I (+*płəš* | +*płəša*, +*bəpłəša* | +*płəšle*)

1 Mahm **+płš** I (+*płəš* | +*płəša* | +*płəšle*)

1 VD Urmi **+płš** I

1 U Urmi **+płš** I ‘to fight, to battle’ (UD 263)

Source:

MEA: *płš* pe. ‘to pierce, break through; to inquire carefully into; to extend, stretch out’ (SL 1203). “In some words an original aspirated stop **p* has shifted to a tense unaspirated stop /*p̰*/. This has occurred in both emphatic and plain words, e.g. +*płəš* ‘he fights’ ” (Khan 2016.1: 163).

FIRE

1 Nud **nura** [nyra]

1 Mahm **nura** [nyra]

1 VD Urmi **nuyra**

1 U Urmi **nuyra** ‘fire, bonfire’ (UD 244)

Source:

MEA: *nūrā* ‘fire’ (SL 904); *nūrā* ‘fire’ (DJBA 738); *nura* ‘fire’ (MD 294).

FISH

1 Nud **nuna** [nyna]

1 Mahm **nuna** [nyna]

1 VD Urmi **nuyna**

1 U Urmi **nuyna** ‘fish’ (UD 244)

Source:

MEA: *nūnā* ‘fish’ (SL 900); *nūnā* ‘fish’ (DJBA 737); *nuna* ‘fish’ (MD 294).

FIVE

1 Nud **xamša**

1 Mahm **xamša**

1 VD Urmi **xamša**

1 U Urmi **xamša** ‘five’ (UD 322)

Source:

MEA: *ḥamšā* ‘five’ (SL 468); *ḥamšā* ‘five’ (DJBA 472); *hamša* ‘five’ (MD 124).

TO FLOAT

There is no specific word for this meaning in Nud, Mahm and VD Urmi.

1 U Urmi **sxy** I (*saxə*) ‘to swim, to bathe; to float; to be washed or purified’ (UD 276)

Source:

See TO SWIM.

TO FLOW

There is no specific word for this meaning in Nud, Mahm and VD Urm, the verb ‘to go’ is used instead.

-1 U Urm **+jry** I (+*jarə*) ‘to flow (of water, river)’ (UD 174)

Source:

< Az. *cərayan etmək* ‘to flow, to pass’ (AIL 155) (cf. also Pers. *jāri šodan* ‘to flow’ (ODP)) < Arab. *jarā* ‘to flow, stream (water)’ (DMWA 121). Alternatively, cf. Syriac *grr* ‘to leak, trickle’ (TS 76), suggested by Mutzafi (2008: 355) as an etymology of JNA Betanure *jāyer* ‘to urinate’.⁷

FLOWER

1 Nud **bəbla**

2 Mahm **varda**

2 VD Urm **varda**

2 U Urm **varda** ‘rose; flower’ (UD 318)

Source:

#1 Etymology is unknown. Cf. Shapytnaya/Levun *bubla* ‘flower’ (Zabelina et al. 2022–); Barwari Bala *bibla* ‘camomile’ (Khan 2008: 1068).

#2 MEA: *wardā* ‘rose, flower’ (SL 359); *wardā* ‘rose’ (DJBA 395); *uarda* ‘flower, rose’ (MD 155).

TO FLY

1 Nud **prx** I (*parəx* | *praxa* | *prəxle*)

1 Mahm **prx** I (*parəx* | *praxa* | *prəxle*)

1 VD Urm **prx** I

1 U Urm **prx** I ‘to fly, to fly away’ (UD 251)

Source:

MEA: *prḥ* pe. ‘to fly’ (SL 1235); *prḥ* pe. ‘to fly’ (DJBA 930); *phr* pe. ‘to fly’, *pra* pe. ‘to fly’ (MD 366, 377).

FLY (N.)

1 Nud **dədva**

1 Mahm **dədva**

1 VD Urm **dədva**

1 U Urm **dədva** ‘fly, an insect’ (UD 142)

Source:

MEA: *dedbā* ‘fly’ (DJBA 328), *didbia* (pl.) ‘fly’ (MD 106f), cf. *dabbābā*, *debbābā* ‘fly’ (SL 268).

FOG

0 Nud There is no special word, occasional borrowings from Russian are used.

-1 Mahm **miž**

⁷ We thank the reviewer for highlighting potential sources of this construction.

-1 VD Urmi **miž**

-1 U Urmi **məž** ‘mist, fog, haze’ (UD 220)

Source:

< Kj. *mij* ‘fog, mist’ (KED 385).

FOOT

1 Nud **aqla**

1 Mahm **+akla**

1 VD Urmi **akla**

1 U Urmi **’akla** ‘leg; foot’ (UD 76)

Source:

MEA: *ʕaqūlā* ‘curvature, curve’ (SL 1129); *ʕāqūlā* ‘cove, inlet’ (DJBA 876).

Cf. Mutzafi comments on *ʕaqūlā* ‘shank, leg’ (TS 425): “but this is surely a vernacular, possibly early Neo-Aramaic word rather than a classical Syriac one. <...> The proto-form **ʕaqūlā* denoted, therefore, a curve of the foot, shank or leg or, perhaps, a crooked foot or leg. If the latter sense is valid, the use of **ʕaqūlā* in the sense of ‘leg, foot’ may have started as a slang or jocular word” (Mutzafi 2014: 186).

FOREST / WOODS

-1 Nud **meša**

-1 Mahm **meša**

-1 VD Urmi **meša**

-1 U Urmi **meša** ‘wood, forest’ (UD 220)

Source:

< Kj. *mêše* ‘forest, woods’ (KED 382); Azer. *meşə* ‘wood(s), forest’ (AIL 554; UD 220).

FOUR

1 Nud **+arpa**

1 Mahm **+arpa**

1 VD Urmi **+arpa**

1 U Urmi **+’arpa** ‘four’ (UD 90)

Source:

MEA: *ʔarbaʕ* ‘four’ (SL 94); *ʔarbāʕ*, *ʔarbaʕ* ‘four’ (DJBA 163); *arba* ‘four’ (MD 36).

TO FREEZE (ABOUT HUMANS)

1 Nud **xbl II** (*xibəl* | *xubúl* | *xubəlle*)

1 Mahm **xbl II** (*xibəl* | *xubula* | *xubəlle*)

1 VD Urmi **xbl II**

1 U Urmi **xbl II** ‘to become cold, to feel cold, to shiver from cold’ (UD 320)

Source:

Khan (2016.1: 425) compares *xabəl II* ‘to become cold’ with Syr. *ħabbəl* ‘to travail’ < *ħēblā* ‘travail’; cf. MEA: *ħbl* pe. ‘to give birth; to conceive; to bring forth, produce; to long for’; pa. ‘to be in travail, labor’ (SL 406); *ħbl* ‘to be in travail’ (DJBA 427).

FRUIT (ANY PRODUCT OF PLANT)

1 Nud **+tunta**1 Mahm **+tunta**1 VD Urmi **+tunta**1 U Urmi **+tunta** ‘(any) fruit, produce’ (UD 317); cf. *yeməš* ‘fruit’ (UD 342); *perə* ‘fruits’ (UD 255); *picə* ‘fruit’ (UD 256)

Source:

MEA: *taḥnā* ‘burden, sack’ (SL 544); *ṭunā*, *tunā* ‘load, burden, fruit’ (DJBA 508), *ṭuna* ‘burden, load; fig. fruit’ (MD 177).

FULL

1 Nud **məlya**1 Mahm **məlya**1 VD Urmi **məlya**1 U Urmi **məlya** ‘full; nourishing; chubby, portly, fat’ (UD 218)

Source:

Resultative participle < *malə* ‘to fill (intr. and tr.)’ (UD 206); MEA: *mly* ‘to fill up’ (SL 768); *mly* ‘to be full’ (DJBA 678); *mła* ‘to fill, be full’ (MD 272).

TO GIVE

1 Nud **yvØ, yv, Øhv** I (*yavvə* | *biyava* | *hule*)1 Mahm **yvl, yvØ, yØv** I, QI (*yavvəl* | *yava* | *yuvvəlle*)1 VD Urmi **yvl, yvØ, yØv** I, QI (*yavvəl* | *yava* | *yuvvəlle*)1 U Urmi **yvl, yvØ, yØv** I, QI (*yavvəl* | *yava* | *yuvvəllun*) ‘to give; to hand over, to deliver; to give out’ (UD 341; Khan 2016.1: 359)

Source:

MEA: *yhb* pe. ‘to give’ (SL 565); *yhb* pe. ‘to give’ (DJBA 526); *yhb, ṛhb, ahb* pe. ‘to give’ (MD 189). On the historical changes in this root see Khan (2016.1: 359–360).

TO GO

1 Nud **Øzl, yzl** I (*ṭazəl* | *bizala* | *zəlle*)1 Mahm **Øzl, yzl** I (*azəl* | *bizala* | *zəlle*)1, 2 VD Urmi **Øzl; rxš, Øxš, Ørš** I (*azəl* | *bərrəšša* | *xəšle*)1, 2 U Urmi **Øzl; rxš, Øxš, Ørš** I (*azəl* | *bərraxša, bərrəšxa, bərrəšša, bəšša* | *xəšlun*) ‘to go’ (UD 81; Khan 2016.1: 351)

Source:

#1 MEA: *ṛzl* pe. ‘to go’ (SL 24); *ṛzl* pe. ‘to go, travel’ (DJBA 100); *azl* pe. ‘to go, move on’ (MD 12).
#2 MEA: *rḥš* pe. ‘to move’ (SL 1458); *rḥš* pe. ‘to creep, move’ (DJBA 1071); *rḥš* pe. ‘to creep, crawl, move, stir, pullulate, bring forth, bring about’ (MD 427).

Comments:

In Urmi varieties, “this verb is expressed by the root Ø-z-l in the present template and the active participle but by a lexically suppletive root in other forms. The suppletive root is

derived historically from **r-x-š*, though the initial /r/ is elided in most forms” (Khan 2016.1: 351).

GOOD

-1 Nud **+spay**

1 Mahm **ǧəns**

1 VD Urmi **ǧəns**, cf. *+spay* as a literary equivalent

-1, 1 U Urmi **+spay** ‘good, pleasant, fine, well’ (UD 286); **jəns**, **jīs**, **ḡīs** ‘good, agreeable, pleasant; (adv.) well’ (UD 172)

Source:

#-1 < Kj. *spehî* ‘pretty, beautiful’ (KED 560; UD 286).

#1 < Azer. *cins* ‘race, breed’, metaphorical ‘thorough-bred’ < Arab. *jins* ‘kind, sort, variety, species, race’ (DMWA 141) < Greek *γενος* (Mutzafi 2018: 22-27; UD 172).

GRASS

1 Nud **ǧəlla**

1 Mahm **ǧəlla** [dʒəlla]

1 VD Urmi **ǧəlla**

1 U Urmi **jəlla** ‘grass’ (UD 165)

Source:

MEA: *gellā* ‘straw, dried grass’ (SL 231); *gillā* ‘straw, chaff’ (DJBA 280); *gila* ‘straw’ (MD 90).

GREEN

1 Nud **mila**

2 Mahm **kina**

2 VD Urmi **kina**

2 U Urmi **kina** ‘green’ (UD 183); cf. *mila* ‘blue, green’ (UD 220)

Source:

#1 ‘green’ < ‘blue’, since in most NENA varieties *milana* or *mila* means ‘blue’, but cf. Aqra (Xərpa), Nəxla (Gerbish) *milana*, Diyana-Zariwaw *mūlana*, Marga *mila* ‘green’ (NENA DB). Further etymology is unclear; cf. Syr. *nīlā* ‘indigo’ (TS 338) < Pers. *nil* ‘indigo, bluing’ (MD); or Pers. *mīnā* ‘enamel; a blue, blueish green, green colour’ (CPED 1364; UD 220).

#2 Etymology is unclear; cf. MEA: *qūnāḡā* ‘blue’ (SL 1335), *qūnāḡā* ‘dark blue’ < Greek *κυανέος* (SL 1337). For semantic shift cf. Jewish Urmi *qənya* ‘green, blue; bruised’ (Khan 2008: 568).

GUTS

1 Nud **+mayə** (pl.)

1 Mahm **+mayə** (pl.)

1 VD Urmi **+may** (pl.)

1 U Urmi **+maya** ‘guts, intestines; womb’ (UD 235); cf. *cisa* ‘belly (stomach + intestines)’ (UD 121)

Source:

MEA: *mḡayyā* ‘intestines’ (SL 800); *mḡayyā* ‘intestines’ (DJBA 694); *mia* ‘intestines, gut’ (MD 265).

HAIR

1 Nud **kavsa**

1 Mahm **cavsa** [kavsa]

1 VD Urmi **cosa**

1 U Urmi **cosa** ‘(head) hair’ (UD 65); cf. +*masta*, +*musta* ‘hair (single strand of hair or all the hair); animal hair’ (UD 236)

Source:

MEA: *qawṣatā* ‘lock of hair, ringlet, curl’ (SL 1341). But there is also another possible etymology < *kawsā* < Syr. *sawkā* ‘branch, twig’ (SL 978) (Mutzafi 2006: 89-90), even though a development of the meaning from ‘branch’ to ‘hair’ is typologically unusual.

HAND

1 Nud **ida**

1 Mahm **ida**; cf. *drana* ~ *drena* ‘arm’

1 VD Urmi **ida**; cf. Archaic *drana* ‘hand’

1 U Urmi **ida** ‘hand; arm’ (UD 84); cf. +*drana* ‘arm’ (UD 149)

Source:

MEA: *īdā* ‘hand’ (SL 31); *yḏā* ‘hand, possession’ (DJBA 523); *īda* ‘hand’ (MD 341).

HARD

1 Nud **qam** ‘hard, strong’

2 Mahm **kavya**

2 VD Urmi **kuya**

2 U Urmi **kuya** ‘hard, hardened, avaricious; strong, strengthened; strongly; loud’ (UD 185); cf. in figurative sense *kəšya* ‘hard; thick, of liquid, dense; strict’ (UD 182)

Source:

#1 Etymology is unknown.

#2 Resultative participle < *kavə* I ‘to become hard; to become strong; to stiffen; to be established’ (Khan 2016.1: 550; UD 181). MEA: *qwy* pa. ‘to remain, wait; to remain unburied; to persevere; to await; to make remain, last’ (SL 1327); *qwa* (Arab. *q-w-y*) ‘to be(come) strong’ (MD 405), “the Mandaic sense ‘fortify’ is surely influenced by the sense of the Arabic cognate *qwy* ‘to be or become strong’” (DMWA 802; CAL).

HE, SHE

1 Nud +**av** (m.); **ay** (f.)

1 Mahm +**av(un)** (m.), **ayən** (f.)

1 VD Urmi +**avən** (m.), **ayən** (f.)

1 U Urmi +**av**, +**avun** ‘he’ (UD 91), **ay**, **ayən** ‘she’ (UD 81)

Source:

PNENA **ahu* (m.), **ahi* (f.) (Hoberman 1990: 85); MEA: *hū* ‘he’ (SL 333), *hī* ‘she’ (SL 339), *haw* ‘that one’ (SL 333); *hu* ‘he; that’ (MD 133); *hī*, *hīia* ‘she’ (MD 151).

HEAD

- 1 Nud **riša**
 1 Mahm **riša**
 1 VD Urmi **riša**
 1 U Urmi **riša** ‘head’ (UD 266)

Source:

MEA: *rēšā* ‘head’ (SL 1462); *rēšā* ‘head, top part’ (DJBA 1078); *riša* ‘head, top’ (MD 434).

TO HEAR

- 1 Nud **+šmy** I (+*šam* | +*bəšmá* | +*šmile*)
 1 Mahm **+šmy** I (+*šamini* | +*bəšmá* | +*šmile*)
 1 VD Urmi **+šmØ, +šmy, +šmmy** I (+*šammə* | +*bəšmaya* | +*šmile*)
 1 U Urmi **+šmØ, +šmy, +šmmy** I (+*šammə* | +*bəšmaya*, +*bəšmá* | +*šmilun*) ‘to hear; to listen, to obey, to be obedient’ (UD 298; Khan 2016.1: 308)

Source:

MEA: *šmē* pe. ‘to hear, listen to’ (SL 1574); *šmē* pe. ‘to hear’ (DJBA 1158); *šma* pe. ‘to hear, listen’ (MD 469).

HEART

- 1 Nud **ləbba**
 1 Mahm **ləbba**
 1 VD Urmi **ləbba**
 1 U Urmi **ləbba** ‘heart; intention; stomach’ (UD 197)

Source:

MEA: *lebbā* ‘heart’ (SL 666); *libbā* ‘heart’ (DJBA 623); *liba* ‘heart’ (MD 234).

HEAVY

- 1 Nud **+yaqura**
 1 Mahm **+yakura**
 1 VD Urmi **+yakuyra**
 1 U Urmi **+yakuyra** ‘heavy, burdensome; slow, inactive, sedate’ (UD 342)

Source:

MEA: *yaqurā* ‘heavy (in weight)’ (SL 581); cf. *yqr* ‘to be heavy’ (SL 582); *yqr* ‘to increase in value’ (DJBA 540); MD *yqr* ‘to honor, respect’.

HERE

- 1 Nud **laxxa**
 1 Mahm **laxxa**
 1 VD Urmi **laxxa**
 1 U Urmi **laxxa**, **’axxa** ‘here; around here, nearby’ (UD 197)

Source:

l-axxa, lit. ‘to here’; MEA: *hārkā* ‘here’ (SL 354), *hākā* ‘here’ (DJBA 381); *haka* adv. of place ‘here’, with prep. *lhaka* ‘hither’ (MD 120).

TO HIDE

- 1 Nud **+ṭšy** II (+ṭiši | +ṭušúv | +ṭušile)
- 1 Mahm **+ṭšy** II (+ṭišini | +ṭušuvva | +ṭušile)
- 1 VD Urmi **+ṭšy** II (+ṭaš | +ṭašuya | +ṭušile)
- 1 U Urmi **+ṭšy** II (+ṭašə) ‘to hide, to conceal, to reserve’ (UD 315)

Source:

MEA: *ṭšy* pe. ‘to hide; to hide o.s.; to be hidden’; pa. ‘to hide, conceal from; to hide o.s.’ (SL 557); *ṭšy* pe. ‘to hide’ (DJBA 520); *ṭša* ethpa. ‘to hide, conceal’ (MD 183).

TO HIT

- 1 Nud **mxy** I (*max* | *bəmxā* | *mxile*)
- 1 Mahm **mxy** I (*maxini* | *mxā* | *mxile*)
- 1 VD Urmi **mxy** I (*max* | *mxaya* | *mxile*)
- 1 U Urmi **mxy** I (*maxə*) ‘to beat, to strike, to hit’ (UD 215)

Source:

MEA: *mhy* pe. ‘to strike’ (SL 738); *mhy* pe. ‘to strike, hit’ (DJBA 655); *mha* pe. ‘to strike, hit’ (MD 258).

TO HOLD

- 1 Nud **dvq** I (*davəq* | *bədvāqa* | *dvəqle*)
- 1 Mahm **dvk** I (*davək* | *bədvāka* | *dvəkle*)
- 1 VD Urmi **dvk** I
- 1 U Urmi **dvk** I ‘to seize, to catch, to hold’ (UD 140)

Source:

MEA: *dbq* pe. ‘to stick to, adhere; to follow; to be numbed; to press to’; pa. ‘to stick, join together; to bind’ (SL 271); *dbq* pe. ‘to adhere, stick; to touch, reach’; pa. ‘to stick, paste to s.t.’ (DJBA 312); *dbq* pe., pa. ‘to cleave to, stick to, snatch, seize’ (MD 101).

HORN

- 1 Nud **qana**
- 1 Mahm **kana**
- 1 VD Urmi **kana**
- 1 U Urmi **kana** ‘horn; trumpet’ (UD 179)

Source:

MEA: *qarnā* ‘horn’ (SL 1412); *qarnā* ‘horn’ (DJBA 1044); *qarna* ‘horn’ (MD 403).

HOUSE

- 1 Nud **biya**
 1 Mahm **biya**
 1 VD Urmi **beta**
 1 U Urmi **beta** ‘house, building’ (UD 104)

Source:

biya < **beya* < **bea* < **beṭa*; MEA: *baytā* ‘house’ (SL 144), *bēṭā* (DJBA 208), *baita* ‘house’ (MD 47).

HOW

- 1 Nud **dax**
 2 Mahm **muḡḡurra**
 2 VD Urmi **muḡḡa**, **muḡḡurra**, cf. *dax* ‘how (as in ‘How are you?’), as (Comparative)’
 1 U Urmi **dax**, **daxi** ‘how, as’ (UD 141); cf. *mujjur*, *mujjurra*, *mujja* (Armenia), *mú-jurra* (Canda) ‘how? what kind of?’ (UD 223)

Source:

#1 < **d-** relative particle ‘of’ & **ax** ‘as, like’ (UD 80):

d-: MEA: *d-* relative particle (SL 268); *d-*, *dī* conj. ‘that, so that, because’, relative pron. ‘who, which, the one who, the one that’, genitive marker ‘of, the one of’ (DJBA 307); *d-* relative particle (MD 491);

ax: MEA: *ʔayk* ‘as, like’ (SL 33), *ʔaykāw* ‘how’ (SL 34); *ʔek* conj. ‘that’ (introducing direct speech, in writs only (DJBA 113), *aiak* ‘like, as if, in manner that’ (MD 14);

#2 < *mu* ‘what’ (see WHAT) & *jur/jurra* ‘kind’ (UD 223) < Pers. *jur* ‘gender, variety, species’ (ODP); KJ. *cûre* ‘type, way, manner’ (KED 96).

TO HUNT

- 1 Nud +**sayduvva** (n) Øvd I, lit. ‘to do a hunt’
 2 Mahm +**avčuvva** (n) Øvd I, lit. ‘to do a hunt’
 2 VD Urmi +**avčuyta** (n) Øvd I, lit. ‘to do a hunt’
 1 U Urmi +**syd** I (+*sayəd*) ‘to hunt, to chase, to fish, to trap’ (UD 284)

Source:

#1 MEA: *šwd* pe. ‘to hunt, fish’ (SL 1277), *šayādūtā* ‘hunting, fishing’ (SL 1284); *šwd* pe. ‘to catch creatures’ (DJBA 953); *šud* pe. ‘to hunt, chase, fish’ (MD 390).

#2 < +**avči** ‘hunter’ (UD 91) & suffix **-uvva/-uyta**:

+*avči*: < Turkish *avcı*, Azer. *ovçu* ‘hunter’ < *av*, *ov* ‘to hunt’ (AIL 618);

-uvva/-uyta: < **-uṭa* (Talay 2008: 69).

HUSBAND

- 1 Nud **gavra** [g^hawra]
 1 Mahm +**gavra** [dʒavra]
 1 VD Urmi +**gora**
 1, 2 U Urmi +**jora** ‘man; husband’ (UD 169); **urza** (UD 86); both words mean ‘man (male); husband’

Source:

#1 MEA: *gabṛā* ‘man, person, husband’ (SL 202); *gabṛā* ‘man, husband’ (DJBA 258); *gabra* ‘man’ (MD 73). Reflexes of Aramaic *gabṛā* ‘man’ expanded their meaning to denote also ‘husband’ in many NENA (Mutzafi 2014: 113). The secondary differentiation between ‘man’ and ‘husband’, which originated from a single etymon (*gura* vs. *gavra*), occurred in Nud as well as in some other NENA varieties (NENA DB; Fox 2009: 146).

#2 see MAN (adult male)

I

1 Nud **ana**

1 Mahm **ana**

1 VD Urmi **ana**

1 U Urmi **’ana** ‘I’ (UD 77)

Source:

PNENA **ana* (Hoberman 1990: 85); MEA: *ʔenā* ‘I’ (SL 58); *ʔānā* ‘I’ (DJBA 143); *ana* ‘I’ (MD 24).

ICE

1 Nud **gdila**

1 Mahm **gdila** [dʒdila]

1 VD Urmi **gdila**

1 U Urmi **jdila** (UD 165)

Source:

MEA: *glīdā* ‘cold, ice’ (SL 236); *glīdā* ‘ice’ (DJBA 287).

IF

1 Nud **ən**

-1 Mahm **aga** [adʒa]

-1 VD Urmi **aga**

1, -1 U Urmi **’ən** ‘if, when’ (UD 82); **’ājar** ‘if’ (UD 76)

Source:

#1 MEA: *ʔen* ‘if’ (SL 58); *ʔi* ‘if, whether’ (DJBA 108); *hin* ‘if, if so, though, although’ (MD 146).

#-1 < Pers. *agar* (UD 76); Kj. *eger* ‘if’ (KED 173); Azer. *əgər* (AIL 264).

IN

1 Nud **gav** [gʲaw]

1 Mahm **gav** [dʒav]

1 VD Urmi **gav, gu**

1 U Urmi **jav, ju, ja, javət** ‘in’ (UD 164)

Source:

MEA: *gawwā* ‘inner part of the body, stomach’ (SL 210); *gawwāʔā* ‘inner (part)’ (DJBA 267); *gaua* ‘inside, interior’ (MD 74).

TO KILL

- 1 Nud **+qtl** I (+*qaṭəl* | +*bəqṭala* | +*qṭalle*)
 1 Mahm **+kṭl** I (+*kaṭəl* | +*bəḳṭala* | +*kṭalle*)
 1 VD Urmi **+kṭl** I
 1 U Urmi **+kṭl** I ‘to kill, to execute, to murder’ (UD 190)

Source:

MEA: *qṭl* pe. ‘to kill’ (SL 1352); *qṭl* pe. ‘to kill’ (DJBA 1006); *gṭl* pe. ‘to kill’ (MD 87).

KNEE

- 1 Nud **bərka**
 1 Mahm **bərca** [bərtʃ^ha]
 1 VD Urmi **bərca**
 1 U Urmi **bərca** ‘knee’ (UD 102)

Source:

MEA: *burkā* ‘knee’ (SL 131); *birkā* ‘knee’ (DJBA 206); *burka* ‘knee’ (MD 57).

TO KNOW

- 1 Nud **+yhh̄y, +dØy, +dȳy** I (+*yahhə*, +*da*, +*dayyi* | +*bəddá* | +*dile*)
 1 Mahm **+yhh̄y, +dØy, +dȳy** I (+*yahhini*, +*dayyini* | +*bəddá* | +*dile*)
 1 VD Urmi **+yṭØ, +yṭṭy, +dØy** I (+*yattə* | +*bəddaya* | +*dile*)
 1 U Urmi **+yṭØ, +yṭṭy, +dØy** I (+*yattə* | +*bəddaya*, +*bəddá* | +*dilun*) ‘to know’ (UD 343; Khan 2016.1: 361)

Source:

MEA: *yḏḏ* pe. ‘to know’ (SL 563); *yḏḏ* pe. ‘to know’ (DJBA 525); *yda*, *ḏda* pe. ‘to know, recognize’ (MD 188).

LAKE

There is no specific word for this meaning in Nud, Mahm, VD Urmi, the word **yama** (or occasionally *yama surta* ‘small sea’) is used instead.

- 1 U Urmi **yamta** ‘lake’ (UD 341); cf. *yama* ‘sea’ (UD 341)

Source:

Diminutive of *yama*, see SEA.

TO LAUGH

- 1 Nud **gxk** I (*gaxək* | *gəxka* | *gxəkle*)
 1 Mahm **gxc** [dʒ-] I (*gaxəc* | *bgəxca* | *gxəcle*)
 1 VD Urmi **gxc** I
 1 U Urmi **jxc** I ‘to laugh, to deride, to mock’ (UD 164)

Source:

MEA: *ghk* pe. ‘to laugh’ (SL 227); *ghk* pe. ‘to laugh’ (MD 81).

LEAF

1 Nud **+tarpa**1 Mahm **+tarpa**1 VD Urmi **+tarpa**1 U Urmi **+tarpa** ‘leaf; leaf of a book’ (UD 315)

Source:

MEA: *ṭarpā* ‘leaf’ (SL 555); *ātarpā* ‘leaf’ (DJBA 108); *aṭirpa* ‘leaf’ (MD 13).

LEFT

-1 Nud **+čapla**-1 Mahm **+čaplá, +čaplaya**-1 VD Urmi **+čapla**1, -1 U Urmi **səmmala** ‘left side’ (UD 277); **+čapla** ‘left side’ (UD 136)

Source:

#-1 < Kj. *čep* ‘left’; Sor. *čep* (KED 105a); Pers. *čap* ‘left’ (UD 136).#1 MEA: *semālā, semālāyā* ‘left’ (SL 1020); *smālā* ‘left side, left hand, left foot’ (DJBA 1188); *smal(a)* ‘left’ (MD 332).

LEG

There is no specific word in all varieties concerned, see FOOT.

TO LIE (IN BED)

1, 2 Nud **dmx** I (*daməx* | *bədməxa* | *dməxle*); **+šrṭx** QI (*+šarṭəx* | *+šarṭúx* | *+šarṭəxle*)1, 2 Mahm **dmx** I (*daməx* | *bədməxa* | *dməxle*); **+šrṭx** QI (*+šarṭəx* | *+šurṭuxa* | *+šurṭəxle*)1, 3 VD Urmi **dmx** I; **+ryx** I (*+rax* | *+bəryaxa* | *+rəxle*)1, 3 U Urmi **dmx** I ‘to lie down, to sleep’ (UD 138); **+ryx** I (*+rayəx, +rāx*) ‘to extend, to become long, to stretch, to stretch out (on the ground), to lie down’ (UD 270); cf. *+šṭx* I ‘to stretch out (on the floor to rest); to stretch so. out (on the ground)’ (UD 299)

Source:

#1 See TO SLEEP. The verb *dmx* ‘to sleep’ is gradually replacing other verbs for this meaning in non-specific contexts.#2 MEA: *šṭḥ* pe. ‘to spread out, make level’; pa. ‘to spread out’ (SL 1547); *šṭḥ* pe., pa. ‘to spread out’ (DJBA 1129); *šṭḥ* pe. ‘to stretch out, be prostrate, lie full length’ (MD 459). Infix *-r-* possibly originated from the dissimilation of the second consonant in D-stem (pa.) (Sabar 1982: 152, 159; Khan 2016.1: 449).#3 MEA: *ʔrk* pa. ‘to lengthen; to draw out’ (SL 100); *yrk* pe. ‘be long’ (SL 584); *ʔrk* af. ‘to be long, lengthen’ (DJBA 168); *ark* ‘to be long, prolonged, lengthy, last a long time’ (MD 37).

TO LIVE

1 Nud **xmy** I (*xam* | *bəxmá* | *xmīle*) ‘live, dwell’2 Mahm **xyØ, xØy** I (*xeni, xayyini* | *bəxxá* | *xīle*)2 VD Urmi **xyØ, xØy, Øxy** I (*xayyə* | *bəxxaya* | *xilə*)2 U Urmi **xyØ, xØy, xyy, Øxy** I (*xayyə* | *bəxyaya, bəxxaya* | *xilə*) ‘to live’ (UD 327; Khan 2016.1: 313)

Source:

- #1 Etymology is unknown. Cf. Zariwaw Neo-Aramaic *xmy* ‘to stay; to live’ (Napiorkowska 2015: 535); Urmi *xmy* II ‘to preserve; to keep; to look after’ (UD 322); MEA: *ḥmy* pe. ‘to see’ (SL 463); *ḥmy* pe. ‘to see’ (DJBA 468); *hma* af. ‘to guard, keep watch upon, protect’ (MD 149).
 #2 MEA: *ḥyy* pe. ‘to live’ (SL 446); *ḥyy*, *hyy* pe. ‘to live, recover, glow’ (DJBA 453), *hia* pe. ‘to live, exist’ (MD 140).

LIVER

1 Nud **kumta**, lit. ‘black (f.sg.)’

-1 Mahm **ḡigar cumta** [dʒidʒar tʃʰumta]

-1 VD Urmi **ḡigar**

-1, 1 U Urmi **+jijar cumta** (lit. ‘black +*jijar*’), **cumta** ‘liver’ (UD 174); cf. less frequent *coda* ‘liver; lungs’ (UD 122)

Source:

#1 See BLACK.

#-1 < Kj. *cegera* (rēš) ‘(black) liver’ (KED 85); cf. Azer. *ciyər*, *qara ciyər* ‘liver’ (ADW 184) or Pers. *jegar* ‘liver’ (ODP).

LONG

1 Nud **irixa**

1 Mahm **irixa**

1 VD Urmi **+yarixa**

1 U Urmi **+yarixa** ‘long’ (UD 343)

Source:

MEA: *arrīk* ‘long’ (SL 99); *arīk* ‘tall, long’ (DJBA 167); *arika* ‘long’ (MD 37); cf. *yrk* pe. ‘to be long’ (SL 584).

LOUSE

1 Nud **qalma**

1 Mahm **+kalma**

1 VD Urmi **kalma**

1 U Urmi **ḵalma** ‘louse’ (UD 177)

Source:

MEA: *qalmā* ‘louse’ (SL 1372); *qalmtā* ‘vermin’ (DJBA 1021).

MAN (ADULT MALE)

1, 2 Nud **gura** ‘man (male)’; **urza** ‘male, male animal’

2, 3 Mahm **urza** ‘male, male animal’; **ḡvanka** ‘man, young man’

3, 2 VD Urmi **ḡvanka** ‘man, young man’; **urza** ‘male, male animal’

2, 1 U Urmi **urza** (UD 86); **+jora** (UD 169); both words mean ‘man (male); husband’, cf. *jvanḵa* ‘young (unmarried) man, beautiful, comely’ (UD 173)

Source:

#1 See HUSBAND.

#2 The etymology is uncertain. Cf. Sor. *urz* ‘honor, purity, chastity’ (DKF 1755) and U Urmi *urzuŷta* ‘manliness; generosity’ (UD 86).

#3 < Kj. *ciwan* ‘young man’ (KED 93), Azer. *cavan* ‘youth, youngster’ (AIL 152).

MAN (HUMAN BEING) / PERSON

1 Nud **naša**

1, 2 Mahm **naša**, +**bar-naša**

1 VD Urmi **naša**

1 U Urmi **naša** ‘man, person; pl. family’ (UD 241); cf. +*bar-naša* ‘man, human, person, individual’ (UD 111)

Source:

#1 MEA: *nāšā* ‘man, human beings’ (SL 65); *ināšā* ‘man’ (DJBA 120); (*a*)*naša* ‘human being’ (MD 24).

#2 MEA: *barnāšā* ‘man’ (SL 178, 191); *bar ʔināš* ‘human being, person’ (DJBA 234); *br* (*a*)*naša*, *barnaša* ‘a man, an individual’ (MD 24, 68). Cf. *brā*, abs. *bar* ‘son’ (SL 177); *brā*, abs./cs. *bar* ‘son, offspring, descendant, having ability, being of age, resident of’ (DJBA 231); *br*, *bra*, abs./cs. *bar* ‘son’ (MD 68).

MANY

1 Nud **riba**

1 Mahm **riba**

1 VD Urmi +**raba**

1 U Urmi +**raba**, +**roba** ‘much, a lot of; many’ (UD 267)

Source:

MEA: *rabbā* ‘great, large’ (SL 1425); *rab* ‘great one, great thing’ (DJBA 1052); *rab*, *rba* ‘great, large, immense’ (MD 417, 422).

MEAT

1 Nud +**bəsra**

1 Mahm +**bəsra**

1 VD Urmi +**bəsra**

1 U Urmi +**busra** ‘meat, flesh’ (UD 115)

Source:

MEA: *besrā* ‘flesh, meat’ (SL 167); *bisrā* ‘flesh, meat’ (DJBA 207); *bisra* ‘flesh, meat’ (MD 62).

MOON

1 Nud +**sara**

1 Mahm +**sara**; **šəmši-+sara**

1 VD Urmi +**sara**

1 U Urmi +**sara** ‘moon’ (UD 283), cf. *šəmši-+sara* ‘moonlight’ (UD 63)

Source:

MEA: *sahrā* ‘moon’ (SL 974); *sehrā* ‘moon’ (DJBA 800); *sira* ‘moon’ (MD 329).

For etymology of *šəmši-* see SUN.

MOTHER

-1 Nud **da**

-1 Mahm **da**

1 VD Urmi **yəmma**

1 U Urmi **yəmma** ‘mother’ (UD 342); cf. *çoç* ‘mother’ (UD 127); *+mama* ‘mummy’ (child language) (UD 227)

Source:

#-1 < Kj. *dê, deya, diya, diya, dya* ‘mother’ (KED 147).

#1 MEA: *emmā* ‘mother’ (SL 52); *immā* ‘mother’ (DJBA 116); *ɛma* ‘mother’ (MD 352).

MOUNTAIN

1 Nud **+tura**

1 Mahm **+tura**

1 VD Urmi **+tuyra**

1 U Urmi **+tura** ‘mountain’ (UD 318)

Source:

MEA: *ṭūrā* ‘mountain’ (SL 521); *ṭūrā* ‘mountain’ (DJBA 498); *tura* ‘mountain, hill’ (MD 178).

MOUTH

1 Nud **pəmma**

1 Mahm **pumma**

1 VD Urmi **pumma**

1 U Urmi **pumma** ‘mouth’ (UD 257)

Source:

MEA: *pūmā* ‘mouth’ (SL 1165); *pūmā* ‘mouth’ (DJBA 889); *puma* ‘mouth’ (MD 368).

NAIL (FINGERNAIL)

1 Nud **+tuṣṣurta**

1 Mahm **+tuṣṣurta**

1 VD Urmi **+tuṣṣurta**

1 U Urmi **+tuṣṣurta** ‘fingernail, toenail, claw, hoof’ (UD 317)

Source:

MEA: *ṭeṣṣrā* ‘nail, claw, talon’ (SL 548); *ṭuṣṣrā* ‘fingernail, toenail’ (DJBA 498); *ṭupra* ‘claw, nail’ (MD 178). Cf. TAIL.

NAME

1 Nud **šəmma**

1 Mahm **šəmma**

1 VD Urmi **šəmma**

1 U Urmi **šəmma** ‘name; reputation, renown’ (UD 294)

Source:

MEA: *šmā* ‘name’ (SL 1569); *šmā* ‘name’ (DJBA 1153); *šuma* ‘name, reputation’ (MD 454).

NARROW

1 Nud **iqa**

1 Mahm **ika**

1 VD Urmi **ika**

1 U Urmi **ika** ‘narrow, cramped, of space’ (UD 84)

Source:

Resultative participle < *ʾayək* ‘to become narrow, to become tight’ (UD 81); MEA: *ʕwq* pa. ‘to oppress’ (SL 1084); *ʕwq* pa., af. ‘to cause trouble, distress’ (DJBA 848); *hwq*, *ʕwq* af. ‘to frighten, torment’ (MD 137).

NAVEL

0 Nud There is no special word, occasional borrowings from Russian are used.

1 Mahm **šurta**

1 VD Urmi **+šurta**

1 U Urmi **šurta** ‘navel; umbilical cord’ (UD 297)

Source:

MEA: *šerā* ‘navel, umbilical cord’ (SL 1599), *šurtā* ‘umbilical cord, navel’ (SL 1613); *šurrā* ‘umbilical cord’ (DJBA 1124); *šura* ‘navel’ (MD 456).

NEAR

1 Nud **qurba**

1 Mahm **kurba**

1 VD Urmi **kurba**

1 U Urmi **kurba** (adj.) ‘close, near, almost’; **kurbabət** (adv.) ‘in the proximity of, near, around’ (UD 185)

Source:

< a noun has come to be used as an adjective (Khan 2016.1: 556); cf. MEA: *qurbā* ‘nearness, proximity’ (SL 1342); *qurbā* ‘kinship; close distance’ (DJBA 1002); *qurba* ‘nearness, proximity’ (MD 409).

NECK

1 Nud **qdala**

1 Mahm **kdala**

1 VD Urmi **kdala**

1 U Urmi **kdala** ‘neck; collar, neckcloth’ (UD 182)

Source:

MEA: *qdālā* ‘neck’ (SL 1317); *qdālā* ‘neck’ (DJBA 984).

NEW

-1 Nud **taža**-1 Mahm **taza**-1 VD Urmi **taza**1, -1 U Urmi **xata** ‘new, fresh, recent’ (UD 325); **taza** ‘new, fresh’ (UD 307)

Source:

#-1 < Kj. *t’aze* [t^haze] ‘fresh, green; new’ (KED 599); Azer. *təzə* (UD 307).#1 MEA: *ḥaṭā* ‘new’ (SL 418); *ḥāḏaṭ* ‘new’ (DJBA 433); *hadta* ‘new’ (MD 116).

NIGHT

1 Nud **lil**1 Mahm **(b)lil**1 VD Urmi **(b)lel**1 U Urmi **lelə** ‘night’ (UD 198)

Source:

blil, *blel* < *b-lele* ‘at night’ < *b-* ‘at, in’ + **laylyā* ‘night’ (Napiorkowska 2015: 105);MEA: *lelyā* ‘night’ (SL 691); *lelyā* ‘night’ (DJBA 626); *lilia* ‘night’ (MD 236).

NOSE

1 Nud **naxira**1 Mahm **naxira**1 VD Urmi **naxira**1 U Urmi **naxira** ‘nose’ (UD 242)

Source:

MEA: *nḥīrē* (pl.) ‘nostrils’ (SL 907); *nḥīrā* ‘nostril’ (DJBA 741); *nhira* ‘nose’ (MD 291).

NOT

1 Nud **la** (Past), **le** (Fut)1 Mahm **la** (Past), **le** (Fut)1 VD Urmi **la** (Past), **le** (Fut)1 U Urmi **la** (Past), **le** (Fut) (UD 194)

Source:

MEA: *lā* ‘no’ (SL 665); *lā* ‘no’ (DJBA 613); *lā* ‘no, not’ (MD 227).

Comments:

“*le* <...> has arisen by a contraction of the /a/ of the negator *la* with a particle /i/, which can be identified with the /i/ in the particle *ci-* (§4.4.1.). It is used to negate both the *ci-patax* form and also the *bət-patax* form” (Khan 2016.2: 139).

OLD

1 Nud **itiqa**1 Mahm **itika**

1 VD Urmi atika

1 U Urmi 'atika 'old, ancient, antique' (UD 79)

Source:

MEA: *ʕattīqā* 'old' (SL 1147); *ʕattīq* 'old' (DJBA 885); *atiqa*, *hatiqa* 'old, ancient' (MD 43).

ONE

1 Nud xe 'one' (m.), **xda** 'one' (f.)

1 Mahm xe 'one' (m.), **xda** 'one' (f.)

1 VD Urmi xa

1 U Urmi xa, xa'a 'one; one time; the same; indefinite article' (UD 320)

Source:

MEA: *ḥaḍ* (m.), *ḥdā* (f.) 'one' (SL 413); *ḥaḍ* (m.), *ḥāḍā* (f.) 'one' (DJBA 430); *had* (m.), *hde* (f.) 'one, single' (MD 116).

OTHER (ANOTHER)

1 Nud xina

1 Mahm xina

1 VD Urmi xina

1 U Urmi xina 'other' (UD 330)

Source:

MEA: *ḥrēnā* 'other' (SL 29); *oḥōrēna* 'another' (DJBA 105); *hurina* 'other, another' (MD 137).

TO PLAY

1 Nud +tØl, tvl I (+*tēl* | +*bəṭol* | +*tolle*)

1 Mahm +tØl, tvl I (+*tēl* | +*bəṭola* | +*tolle*)

1 VD Urmi +tØl, tvl II (+*tal* | +*ṭavula* | +*ṭuvəlle*)

1 U Urmi +tØl, tvl II (+*ṭavəḷ*, +*ṭāl* | +*ṭavulə* | +*ṭuvəllun*) 'to play, to be amused, to gambol, to frolic' (UD 316; Khan 2016.1: 326)

Source:

< **ṭl* (Khan 2016.1: 326) < **ṭll*; MEA: *ṭll* pa., ethpa. 'to play, sport' (DJBA 505); *ṭll* pa., ethpa. 'to play, sport' (MD 180).

TO PULL

1 Nud grš I (*garəš* | *bəgraša* | *grəšle*)

1 Mahm grš [g- ~ dʒ-] I (*garəš* | *bəgraša* | *grəšle*)

1 VD Urmi grš I

1 U Urmi jrš I 'to pull, to haul, to drag, to draw (a weapon)' (UD 162)

Source:

MEA: *grš* pe. 'to pull, drag' (SL 264).

TO PUSH

1 Nud **xrbs** QI (*xərbəs* | *xərbús* | *xərbəsle*)

1 Mahm **xrbs** QI (*xərbəs* | *xarbusa* | *xurbəsse*)

1 VD Urmi **xrbs** QI

1 U Urmi **xrbs** QI ‘to push, to push aside, to cast out’ (UD 324); cf. *xrzp* QI ‘to push, to push away, to push over (a person)’ (UD 325)

Source:

#1 Etymology is unknown; cf. MEA *hbs* pe. ‘to crush’ (DJBA 427). Infix *-r-* possibly originated from the dissimilation of the second consonant in D-stem (Sabar 1982: 152, 159; Khan 2016.1: 449).

RAIN

1 Nud **+əryana**

1 Mahm **+aryana**

2 VD Urmi **+məṭra**

2 U Urmi **+muṭra** ‘rain’ (UD 238); cf. less frequent *+aryana* (UD 90)

Source:

#1 cf. MEA: *?aryā* ‘cold’ (SL 1139); *?ārītā* ‘chili, cold’ (DJBA 168). For semantic shift cf. Dobe *?ry* ‘to freeze’, Tyare *?ry*, *r?y* ‘to rain, snow’ (Mutzafi 2008: 332).

#2 MEA: *meṭrā* ‘rain’ (SL 749); *miṭrā* ‘rain’ (DJBA 665); *miṭra* ‘rain’ (MD 266).

RED

1 Nud **smoqa**

1 Mahm **smoka**

1 VD Urmi **smuka**

1 U Urmi **smuḵa** ‘red, ruddy, bay, of horse’ (UD 279)

Source:

MEA: *summāqā* ‘red’ (SL 981); *summāq* ‘red object, redness’ (DJBA 794); *s(u)maq(a)*, *sumqa* ‘red, ruddy’ (MD 322).

RIGHT (RIGHTSIDE)

-1 Nud **+rasta**

-1 Mahm **+rastə**

-1 VD Urmi **+rastə**

1 U Urmi **yammīna** ‘right’ (UD 341), cf. *+rast*, *+rastə* ‘correct; right side’ (UD 269) (attested only in texts from Armenia)

Source:

#-1 < Kj. *rast* ‘right (direction); correct’ (KED 503); *rāst* ‘straight, right’ (ODP); Azer. *rast* ‘straight; neatly, to the point’ (AIL 651; UD 269).

#1 MEA: *yammīnā* ‘right’ (SL 576); *yammīnā* ‘right side, right hand’ (DJBA 536); *iamina* ‘right, the right hand’ (MD 186).

RIVER

-1, 1 Nud **+čamma** ‘small river’; **nara** ‘big river’

1 Mahm **nara**

1 VD Urmi **nara**

1 U Urmi **nara** ‘river’ (UD 241)

Source:

#-1 < Kj. *ç'em* ‘stream, river’ (KED 103).

#1 MEA: *nahrā* ‘river’ (SL 894); *nahrā* ‘river’ (DJBA 734); *nahra* ‘river’ (MD 281)

ROAD / WAY

1 Nud **urxa**

1 Mahm **urxa**

1 VD Urmi **urxa**

1 U Urmi **urxa** ‘road, way, course, entrance, access’ (UD 85)

Source:

MEA: *ʔurhā* ‘road’ (SL 21); *ʔorhā* ‘road, journey, manner’ (DJBA 94); *ʔuhra* ‘road, way, journey’ (MD 343).

ROOT

1 Nud **+əqra**

1 Mahm **+akra**

1 VD Urmi **+akra**

-1 U Urmi **coca** ‘trunk of a tree, stock of a tree; root’ (UD 122); cf. *+əkra* ‘trunk of tree; stump, stub; stem; origin’ (UD 93)

Source:

#1 MEA: *ʔeqqārā* ‘root’ (SL 1132); *ʔiqqārā* ‘root, essential portion’ (DJBA 859); *aqara* ‘root’ (MD 33).

#-1 < Azer. *kök* ‘root’ (AIL 439); Kj. *k'ok* [kʰok] ‘root’ < Turkish *kök* ‘root’ (KED 330).

ROPE

1 Nud **gədda** ‘rope; thread’

2 Mahm **+xola** ‘rope’; cf. *gdala* [dʒdala] ‘thread’; *+čopər*, *+čaṭi* (both words are used occasionally by some speakers)

2, 1 VD Urmi **+xola** ‘rope, knitting thread’; **gdala** ‘thread, rope’

2 U Urmi **+xola** ‘rope’ (UD 338); cf. *tanap* ‘line, cord, used to hang out washing’ (UD 303); *ɟdala*, *ɟəddala* ‘thread; lace, strand’ (UD 164)

Source:

#1 MEA: *gdādā* ‘cutting off; thread’ (SL 204); *gdādā* ‘thread’ (DJBA 260); *gdada*, *gadada* ‘strip’ (MD 80).

#2 Etymology is unclear. Cf. Armenian *čopan* ‘rope’ (Aroutunian, Aroutunian 1993: 321).

#3 MEA: *ħablā* ‘rope’ (SL 408); *ħablā* ‘rope’ (DJBA 427); *habla* ‘rope’ (MD 115);

ROTTEN

1 Nud **spisa** ‘rotten’; cf. *xriva* ‘spoiled’

2, 1 Mahm **xmika** ‘rotten’; **spisa** ‘completely rotten’; cf. *xriva* ‘spoiled’

1 VD Urmi **spisa** ‘rotten, decayed’; cf. *xriva* ‘spoiled’

1 U Urmi **spisa** ‘rotten’ (UD 279); cf. *xriva* ‘spoiled, damaged’ (UD 231); *xmika* ‘decayed, rotten, putrefied, smelly’ (UD 330)

Source:

#1 Resultative participle < *saḫās* ‘to decay, to rot, to decompose’ (UD 275), probably < Greek *σάπιζω* ‘to rot’ (Sabar 2002: 242).

#2 Resultative participle < *xaməḵ* ‘to decay, to become rotten’ (UD 322); further etymology is unknown.

Comments:

Most speakers give *xriva* ‘spoiled’ as a first reaction instead of a more semantically specific word for ‘rotten’.

ROUND

1 Nud **glula** [glyla]

1 Mahm **glula** [glyla]

1 VD Urmi **gluyla**

1 U Urmi **jlula** ‘round; ball; circle’ (3d) (UD 166)

Source:

MEA: *glālā* ‘round’ (SL 238); *glālā* ‘stone-shaped object’ (DJBA 288); *glala* ‘stone, rock, hail-stone; something round, ball’ (MD 91).

TO RUB

1 Nud **prx** I (*parəx* | *bəpraxa* | *prəxle*)

1 Mahm **prx** I (*parəx* | *praxa* | *prəxle*)

1 VD Urmi **prx** I

1 U Urmi **prx** I ‘to rub, to scrub’ (UD 251)

Source:

MEA: *prk* pe. ‘pound, grind’, pa. ‘to rub, to pound’ (SL 1240); *prk* pe. ‘to break off, refute’ (DJBA 933).

TO RUN

1 Nud **+rxṭ** I (*+raxəṭ* | *+bərxəṭa* | *+rxəṭle*)

1 Mahm **+rxṭ** I (*+raxəṭ* | *+bərxəṭa* | *+rxəṭle*)

1 VD Urmi **+rxṭ** I; cf. *Ørk* I ‘to run away’

1 U Urmi **+rxṭ** I ‘to run, to rush (on foot)’ (UD 269); cf. *Ørk* I ‘to run, to run away, to escape, to flee, to evade, to avoid’ (UD 78)

Source:

MEA: *rhṭ* pe. ‘to run, go forth’ (SL 1440); *rhṭ* pe. ‘to run, come quickly, follow someone’s opinion, act in haste’ (DJBA 1060); *rhṭ* pe. ‘to run, hasten, move quickly, flow swiftly’ (MD 426).

SALT

- 1 Nud **məlxə**
 1 Mahm **məlxə**
 1 VD Urmi **məlxə**
 1 U Urmi **məlxə** ‘salt’ (UD 218)

Source:

MEA: *melḥā* ‘salt’ (SL 767); *milḥā* ‘salt’ (DJBA 667); *mihla* ‘salt’ (MD 266).

SAND

- 1 Nud **xiza**
 -2 Mahm **silə**
 -2 VD Urmi **silə**
 -2 U Urmi **silə** ‘sand’ (UD 278)

Source:

#-1 < Kj. *xîz* ‘sand; gravel, pebble’ (KED 665)

#-2 < Sor. *seylak* (*sêl*, *sêlak*, *sêlax*, *sêleh*, *sêlix*, *sîlewan*, *sîlik*) ‘sand’ (DKF 1495; 1498; 1524); *seylak* ‘sand’ (Qazzaz 2000: 364).

TO SAY

- 1 Nud **Ømr** I (*ṭamər* | *mara* | *mire*)
 1, 2 Mahm **Ømr** I; **tny** II (*tin(-ini)* | *mara*, *tunuṽṽə* | *mire*, *tunile*)
 1, 2 VD Urmi **Ømr** I; **tny** II (*tanə* | *mara*, *tanuṽṽə* | *mərre*, *tunile*)
 1, 2 U Urmi **Ømr** I ‘to say, to tell’ (UD 77); **tny** II ‘to repeat, to say again, to do again; to say, to tell’ (UD 303) (*tanə* | *mara*, *tanuyə* | *mərrun*, *tunilun*) (Khan 2016.1: 354)

Source:

#1 MEA: *ʔmr* pe. ‘to say’ (SL 57); *ʔmr* pe. ‘to say, recite, tell’ (DJBA 140); *amr* pe. ‘to say, speak, talk, narrate, discuss, command’ (MD 23).

#2 MEA: *tny* pe. ‘to repeat; to tell, relate; to say, speak; to recite’; pa. ‘to recall, learn; to do again; to tell’ (SL 1654); *tny* pe. ‘to repeat, do again; to learn a Tannaitic tradition; to recite a Tannaitic tradition; to explain, elucidate’; pa. ‘to relate, tell; to study the talmud’ (DJBA 1218), *tna* pe. ‘to do for a second time, repeat, reiterate’ (MD 488).

Comments:

In Urmi (which influenced Mahm), “the original verbal root ‘to say’ Ø-*m-r* I has become restricted to the past form, resultative participle and progressive. The verb *tanə* II (*t-n-y*), which originally meant ‘to recount (a tale)’, has now developed the synonymous sense ‘to say’. Unlike Ø-*m-r*, it is used in all inflections, so has replaced Ø-*m-r* in the present and imperative. It exists as an alternative to Ø-*m-r* in the past, resultative participle and progressive.” (Khan 2016.1: 354).

TO SCRATCH

- 1 Nud **+zrč** I (+*zarəč* | +*bəzrača* | +*zračle*)
 1 Mahm **+zrč** I (+*zarəč* | +*bəzrača* | +*zračle*)

1 VD Urmi +zrč̣ I

1 U Urmi +zrč̣ I, +zrzč̣ QI ‘to scratch (with nails and break skin), to claw’ (UD 350); cf. less frequent +*jrč̣* I ‘to scratch’ (UD 169)

Source:

MEA: *sr̥t* pe. ‘to scratch, prick; to write’ (SL 1045); *zr̥t* pa. ‘to torment’ > *zrāṭā* ‘scraping, shaving’ (SL 398); *sr̥t* pa. ‘to lacerate’ (DJBA 831). Cf. also: “The affrication of an original unvoiced /t/ in verbs with an emphatic suprasegmental setting has the effect of expressing a greater degree of force or hardness. <...> e.g. +*zarəṭ* I ‘to score (a board)’, +*zarəč̣* I ‘to scratch (with nails and break skin)’” (Khan 2016.1: 182).

SEA

1 Nud yama**1 Mahm yama****1 VD Urmi yama**

1 U Urmi yama ‘sea’ (UD 341); cf. *yamta* ‘lake’ (UD 341)

Source:

MEA: *yamā* ‘sea, lake’ (SL 575); *yammā*, *yōmā* ‘sea, ocean’ (DJBA 536); *iama* ‘sea, lake, stretch of water, lagoon, pool, river’ (MD 186).

TO SEE

1 Nud xzØ, xzy I (*xazzə* | *bəxzā* | *xzile*)**1 Mahm xzØ, xzy I** (*xaz(z)ini* | *bəxzā* | *xzile*)**1 VD Urmi xzØ, xzy I** (*xazzə* | *bəxzaya* | *xzile*)

1 U Urmi xzØ, xzy I (*xazzə*, *xazə*) ‘to see, to look, to look at; to consider, to notice’ (UD 327)

Source:

MEA: *ḥzy* pe. ‘to see’ (SL 438); *ḥzy* pe. ‘to see, look at’ (DJBA 444); *hza* pe. ‘to see, look’ (MD 138).

SEED

1 Nud +bərzarra**1 Mahm +barzarra****1 VD Urmi +bərzarra**

1 U Urmi +bar-+zarra, +bər-+zarra ‘seed, pit of fruit’ (UD 110)

Source:

MEA: *bar zarṯā* ‘seed’, lit. ‘son of seed’ (SL 180), *zarṯā* ‘seed’ (SL 400); *bazrā*, *bizrā* ‘seed’ (DJBA 195); *bazira*, *bazra* ‘seed’ (MD 46).

TO SEW

1 Nud +xyṭ I (*+xeṭ* | *+bəxyaṭa* | *+xəṭle*)**1 Mahm +xyṭ I** (*+xeṭni* | *+bəxyaṭa* | *+xəṭle*)**1 VD Urmi +xyṭ I** (*+xaṭni* | *+bəxyaṭa* | *+xəṭle*)

1 U Urmi +xyṭ I (*+xayəṭ*, *+xāṭ*) ‘to sew, to sew up’ (UD 337)

Source:

MEA: *hyt* pe. ‘to sew’ (SL 422); *hw̄t*, *h̄ty* pe. ‘to sew’ (DJBA 436); *hūt* pe. ‘to sew, stitch together’ (MD 135).

SHADOW

1 Nud **+t̄laniya**

1 Mahm **+t̄laniya**

1 VD Urmi **+t̄lanita**

1 U Urmi **+t̄lanita** ‘shadow, shade’ (UD 317)

Source:

MEA: *ṭellānītā* ‘shadow’ (SL 534); *ṭallānītā*, *ṭullānītā* ‘type of demon or shade’ (DJBA 505); *ṭulanita* ‘shade, shadow, spectre, phantasm, phantom, a demon’ (MD 177).

SHARP

1 Nud **xirupa**

1 Mahm **xirupa**

1 VD Urmi **xaruypa**

1 U Urmi **xarupa** ‘sharp; pungent; fervent’ (UD 324)

Source:

Adj. < *xarəp* ‘to become sharp (knife); to become pungent’ (UD 324). MEA: *hr̄p* ‘to be sharp’ (SL 494); *hr̄p* ‘to sharpen’ (DJBA 484); *hr̄p* ‘to be early, sharp’ (MD 153).

SHORT

-1, 1 Nud **qutta; kərya**

1 Mahm **cərya** [tʰərja]

1 VD Urmi **cərya**

1 U Urmi **cərya** ‘short’ (UD 120)

Source:

#-1 < Sor. or Kj. *qut* ‘short’ (DKF 1342).

#1 MEA: *karyā* ‘short’ (SL 651).

TO SING

1 Nud **zmr** I (*zamər* | *bəzmara* | *zmərre*)

1 Mahm **zmr** I (*zamər* | *bəzmara* | *zmərre*)

1 VD Urmi **zmr** I

1 U Urmi **zmr** I ‘to sing’ (UD 345)

Source:

MEA: *zmr* pe. ‘to sing; to pipe, play on pipe’; pa. ‘to sing, praise’ (SL 386); *zmr* pe., pa. ‘to sing’ (DJBA 417); *zmr* pe., pa. ‘to sing, make music, play (an instrument), make a sound, resound’ (MD 169).

TO SIT

- 1 Nud **y_{tv}**, **ty_v** I (*yatu* | *bətyava* | *tule*)
 1 Mahm **y_{tv}**, **ty_v** I (*yatəv* | *bətyava* | *təvle*)
 1 VD Urmi **y_{tv}**, **ty_v** I (*yatəv* | *bətyava* | *təvle*)
 1 U Urmi **y_{tv}**, **ty_v** I (*yatəv* | *bətyava* | *təvulun*) ‘to sit, to sit down’ (UD 341; Khan 2016.1: 323)

Source:

MEA: *y_{tb}* pe. ‘to sit’ (SL 587); *y_{tb}* pe. ‘to sit’ (DJBA 545); *y_{tb}* pe. ‘to sit, stay’ (MD 193).

SKIN (HUMAN)

- 1 Nud **gəlda**
 1 Mahm **gəlda** [dʒəlda]
 1 VD Urmi **gəlda**
 1 U Urmi **jəlda** ‘skin, fur’ (UD 165); cf. *jallaxta* ‘skin, hide; bark of tree’ (UD 170)

Source:

MEA: *geldā* ‘skin, leather’ (SL 233); *gildā* ‘scab, hide’ (DJBA 280); *gilda* ‘leather’ (MD 90).

SKY

- 1 Nud **šmayya**
 2 Mahm **dunya**
 2 VD Urmi **dənyə** ‘sky, world’
 1 U Urmi **šmayya** ‘sky, heaven’ (UD 296); cf. *dunya*, *dənyə* ‘world, universe’ (UD 145)

Source:

- #1 MEA: *šmayyā* ‘sky, heaven’ (SL 1572); *šmayyā* ‘sky, heaven’ (DJBA 1157); *šumya* ‘heavens’ (MD 455).
 #2 < K_j. *dinya*; *dunîya*, *dunya* ‘world’ (KED 155) or Azer. *dünya* ‘universe, world’ (AIL 240); < Arab. *dunyā* ‘world; earth’ (DMWA 295; UD 145). The meaning ‘sky’ is apparently developed within the NENA varieties of Verin Dvin from the common meaning ‘world’ so it’s not counted as loanword.

TO SLEEP

- 1 Nud **dmx** I (*daməx* | *bədməxa* | *dməxle*)
 1 Mahm **dmx** I (*daməx* | *bədməxa* | *dməxle*)
 1 VD Urmi **dmx** I
 1 U Urmi **dmx** I ‘to lie down, to sleep’ (UD 138)

Source:

MEA: *dmk* pe. ‘to sleep’ (SL 310); *dmk* pe. ‘to lie’ (DJBA 343).

SMALL

- 1 Nud **sora**
 1 Mahm **sura** [syra]
 1 VD Urmi **sura**

1 U Urmi **sura** ‘small, young’ (UD 280)

Source:

< *sōrā < *zōrā (Khan 2016.1: 191); MEA: *zāforā* ‘small’ (SL 390); *zfer* ‘small, young’ (DJBA 418).

TO SMELL

1 Nud **rixa** (n) **šql** I, lit. ‘to take the smell’

1 Mahm **rexa** (n) **grš** [g- ~ dʒ-] I / **šql** I, lit. ‘to pull / to take the smell’

1 VD Urmi **rexa** (n) **grš** I, lit. ‘to pull the smell’

1 U Urmi **rexa** (n) **škl** I ‘to smell’ (UD 68); cf. *mmØx* III (*mamməx*) ‘to smell, to sniff’ (UD 207)

Source:

MEA: *rēḥā* ‘smell; wind’ (SL 1461), cf. *ryḥ* af. ‘to smell’ (SL 1460); *rēḥā* ‘smell, odor’ (DJBA 1075), cf. *ryḥ* af. ‘to smell’ (DJBA 1075); *riha* ‘breath, smell, scent, perfume’ (MD 432), cf. *rha* af. ‘to inhale, smell, perceive by senses, breathe, cause to inhale or breathe in, enjoy’ (MD 425). Similar constructions are attested in Kurmanji, cf. *bihn standin* ‘to smell’ (DKF 169), *bîna ft-i k’işandin* ‘to breath in’ (KED 75), where *bihn/bîn* ‘smell’, *standin* ‘to take, receive’, *k’işandin* ‘to pull, draw’.

SMOKE

1 Nud **tənnə**

1 Mahm **tənnə**

1 VD Urmi **tənnə**

1 U Urmi **tənnə** ‘smoke’ (UD 307)

Source:

MEA: *tennānā* ‘smoke’ (SL 1656); *tnn?* ‘smoke’ (DJBA 1223); *tana* ‘a cosmic mist, vapour, or smoke’ (MD 479).

SMOOTH

-1 Nud **duz** [dyz]

-1 Mahm **duz** [dyz]

-1, 1 VD Urmi **duz**; **+mačču**

1 U Urmi **+mačču** ‘smooth’ (UD 225); cf. *dūz* ‘straight, flat; right, upright; correct, exact’ (UD 146)

Source:

#-1 See CORRECT.

#1 < *+maččuyə* ‘to smooth’ < **mašōyē*, where “an original *š has shifted to č with suprasegmental emphasis due to contact with a pharyngal, the historical root being *š-ʕ-ʕ” (Khan 2016.1: 175); MEA: *šaḥiḥā* ‘smooth’ (SL 1583), *šāḥā* ‘smooth’ (SL 1584), *šḥ* pe. ‘to smooth’, pa. ‘to make smooth’, af. ‘to polish, smooth’ (SL 1584); *šiaḥ* ‘smooth’ (DJBA 1137); *šaia* ‘lustrous, smooth, glorious’ (MD 440), *šaa*, *šua* ‘to be smooth, pleasant, even, harmonized, glorious’ (MD 438).

SNAKE

1 Nud **xuvva**

1 Mahm **xuvva**

1 VD Urmi **xuvva**

1 U Urmi **xuvva**, **xuvvə** ‘snake’ (UD 332)

Source:

MEA: *ḥewyā* ‘snake’ (SL 424); *ḥiwyā* ‘snake’ (DJBA 450); *hiuia* ‘serpent, snake’ (MD 142).

SNOW

1 Nud **təlga**

1 Mahm **təlga** [təldʒa]

1 VD Urmi **təlga**

1 U Urmi **talja** ‘snow’ (UD 303)

Source:

MEA: *talgā* ‘snow’ (SL 1647); *talgā* ‘snow’ (DJBA 1208); *talga* ‘snow’ (MD 478).

SOME (AN UNSPECIFIED QUANTITY OR NUMBER OF)

1, 2 Nud **xekma**; **xečča**

1, 2 Mahm **xacma** [xatʰma]; **xačča**

1, 2 VD Urmi **xacma**; **xačča**

1, 2 U Urmi **xacma** ‘some’ (UD 321); **xačča** (mod., adv.) ‘a little, a few, some; slightly’ (UD 321)

Source:

#1 < **xa* (see ONE) & **kma* (Khan 2016.1: 167): MEA: *kmā* ‘how, how much’ (SL 628); *kmā* ‘how much/many?’ (DJBA 585); *kma* ‘how, how much’ (MD 218).

#2 See FEW/LITTLE.

Comments:

xacma (*xekma*) is used with countable nouns; *xačča* (*xečča*) is used with uncountable nouns.

TO SPIT

1 Nud **ryq** I (*req* | *bəryaqa* | *raqle*)

1 Mahm **ryk**, **rØk** I (*rekni* | *bərraka* | *rakle*)

1 VD Urmi **ryk**, **rØk** I (*rakni* | *bərraka* | *rakle*)

1 U Urmi **ryk**, **rØk** I (*rayək*, *rāk*) ‘to spit’ (UD 265)

Source:

MEA: *rqq* pe. ‘to spit, spit upon’ (SL 1490); *rqq* pe. ‘to spit’ (DJBA 1094); *rqq* pe. ‘to spit’ (MD 437).

TO SPLIT (FIREWOOD)

1 Nud **+člp** I (+*čaləp* | +*bəčləpa* | +*čləple*)

1 Mahm **+člp** I (+*čaləp* | +*bəčləpa* | +*čləple*)

1 VD Urmi **+člp** I

1 U Urmi **+člp** I ‘to cleave, to chop, to cut off a branch; to split (tr. and intr.), to dissect, to burst (tr. and intr.)’ (UD 135)

Source:

MEA: *šlp* pe. ‘to split, cleave; to wound; to beat, strike’ (SL 1290); *šlp* pe. ‘to wound’ (DJBA 966); *šlp* pe., pa. ‘to bruise, injure, whip, beat, wound, flog’ (MD 395). “Affrication of stops

and sibilants in emphatic settings have developed in a number of ... verbs expressing forcefulness or hardness, e.g. +čaləp I ‘to split, to cleave’ < *šaləp (cf. Syr. šlap̄)” (Khan 2016.1: 183).

TO SQUEEZE

1 Nud **+xls** I (+xaləs | +bəxlasa | +xləsle)

1 Mahm **+xls** I (+xaləs | +bəxlasa | +xləsle)

1 VD Urmi **+xls** I

1 U Urmi **+xls** I ‘to press, to be pressed; to squeeze, to be squeezed; to squeeze out, to press down; to bind, to tighten; to gird, to be girded; to shut tight, to close’ (UD 334); cf. less frequent *ɟmš* I ‘to squeeze, to grab with five fingers; to wring; to hold tight in the hand; to take a handful’ (UD 160).

Source:

Etymology is unclear. Cf. MEA: *hlš* pe. ‘to show o.s. strong, valiant’, pa. ‘to pillage, steal; to draw out, extract’ (SL 460); *hlš* ‘to gird up the loins, brace oneself, be brave’ (MD 149). Probable semantic shift is as follows: ‘to gird up’ > ‘to press’ > ‘to squeeze’.

TO STAB

1 Nud **+t̥sØ, +t̥Øs** I (+t̥assə | +bətt̥asa | +t̥əsle)

1 Mahm **+t̥ys, +t̥Øs** I (+t̥as | +bətt̥asa | +t̥əsle); cf. +mØvr III ‘to insert’

1 VD Urmi **+t̥ys, +t̥Øs** I (+t̥as | +bətt̥asa | +t̥əsle); cf. +mØvr III ‘to insert’

1 U Urmi **+t̥ys, +t̥Øs** I (+t̥ayəs, +t̥ās) ‘to stick in, to thrust in, to indent, to prick, to pierce; to stick fast’ (UD 316); cf. *črz* I ‘to plunge, to knock in (sharp object), to thrust through; to stab (with a spear or a bayonet); to prod, to urge’ (UD 130); *cəš* I ‘to thrust (into sth.), to stuff (into sth.)’ (UD 119)

Source:

MEA: *d̥l̥š* pe. ‘to stick in, im plant’ (SL 315); *d̥l̥š, d̥h̥š* pe. ‘to stick, insert’ (DJBA 345); *dw̥š, d̥sy* pe. ‘to insert, stick in’ (DJBA 320); *daš, duš* pe. ‘to thrust in, insert, pierce, stick in’ (MD 100); see also Khan 2016.1: 291.

TO STAND

1 Nud **kly** I (*kal* | *bəklá* | *klile*)

1 Mahm **cl̥y** [k- ~ t̥ʰ-] I (*calni* | *bəclá* | *clile*)

1 VD Urmi **cl̥y** I (*cal* | *bəclaya* | *clile*)

1 U Urmi **cl̥y** I (*calə*) ‘to stand, to stop (intr.), to wait’ (UD 116)

Source:

MEA: *kly* pe. ‘to impede, prevent’ (SL 624); *kly* pe. ‘to be finished’ (DJBA 582); *kla, kll* pe. ‘to keep enclosed, hold back’ (MD 216).

STAR

1 Nud **kəxva**

1 Mahm **cəxva** [t̥ʰəxvə]

1 VD Urmi **cəxva**

1 U Urmi **cuxva, cəxva** ‘star’ (UD 124)

Source:

MEA: *kawkbā* ‘star’ (SL 606); *kokbā* ‘star’ (DJBA 558); *kukba* ‘star’ (MD 206).

STICK

1 Nud **qaysa**

1 Mahm **kaysa**

1 VD Urmi **kesa**

1 U Urmi **kesa** ‘wood, stick; shaft, of a spear; tree; gallows’ (UD 183)

Source:

MEA: *qaysā* ‘wood, piece of wood’ (SL 1364); *qinsā* ‘piece of wood, twig’, *qissā* ‘wood, twig’ (DJBA 1014); *qisa* ‘rough edge’ (MD 412).

STONE

1 Nud **kipa**

1 Mahm **cipa** [tʰipa]

1 VD Urmi **cipa**

1 U Urmi **cipa** ‘stone, rock’ (UD 121)

Source:

MEA: *kēpā* ‘stone’ (SL 594); *kēpā* ‘stone’ (DJBA 577).

STRAIGHT

1 Nud **duz**

1 Mahm **duz**

1 VD Urmi **duz**

1 U Urmi **dūz** ‘straight, flat; right, upright; correct, exact’ (UD 146)

Source:

See CORRECT.

TO SUCK (BREAST, MILK — ABOUT CHILD)

1 Nud **+mys, +Øms** I (+*mas* | +*bəmma* | +*məsle*)

1 Mahm **+mys** I (+*mesni* | +*myasa* | +*məsle*)

1 VD Urmi **+mys** I (+*masni* | +*myasa* | +*məsle*)

1 U Urmi **+mys** I (+*mayəs*, +*mās*) ‘to suck (milk from teat, also with sexual connotation)’ (UD 235)

Source:

< **m-š-š* (Khan 2016.1: 293); MEA: *mšš* pe. ‘to suck; squeeze out, press out; to drink’ (SL 818); *mšš* pe. ‘to suck, absorb’; pa. ‘to suck’ (DJBA 701); *mšš*, *muš* pe. ‘to press; to suck, suckle’ (MD 277). Cf. also *mšy* pe. ‘to wring, squeeze; to suck’ (SL 814); *mšy* pe. ‘to wring out, suck (of blood)’ (DJBA 699), *mša* pe. ‘to press; to suck out, suckle’ (MD 277).

SUN

1 Nud **šəmša**

2, 1 Mahm yuma, šəmša

1 VD Urmi šəmša

1 U Urmi šəmša ‘sun, heat’ (UD 294)

Source:

#1 MEA: *šemša* ‘sun’ (SL 1576); *šimša* ‘sun, sunlight’ (DJBA 1136); *šamša* ‘sun’ (MD 443).

#2 See DAY.

SWEET

1 Nud +xəlya

1 Mahm +xəlya

1 VD Urmi +xəlya

1 U Urmi +xəlya ‘sweet, delicious’ (UD 337)

Source:

MEA: *ḥalyā* ‘sweet’ (SL 455); *ḥālē*, det. *ḥalyā* ‘sweet’ (DJBA 462); *halia* ‘sweet, agreeable, pleasant, kind(ly)’ (MD 121).

TO SWELL

1 Nud zyr I (*zer* | *bəzyara* | *zire*)

1 Mahm zyr I (*zerni* | *bəzyara* | *zərre*)

1 VD Urmi zyr I (*zar* | *bəzyara* | *zərre*)

1 U Urmi zyr I (*zayər*, *zār*) ‘to swell (with food or liquid)’ (UD 347)

Source:

Etymology is unclear. Cf. MEA: *zwy* pe. ‘to swell, bloat’ (SL 372).

TO SWIM

1 Nud sxy, syx I (*sax* | *bəsyaxa* | *sxile*)

1 Mahm sxy I (*saxini* | *bəsxá* | *sxile*)

1 VD Urmi sxy I (*saxini* | *bəsxaya* | *sxile*)

1 U Urmi sxy I (*saxə*) ‘to swim, to bathe’ (UD 276)

Source:

MEA: *šhy* pe. ‘to wash o.s., bathe’ (SL 992); *šhy* pe. ‘to wash oneself, bathe’ (DJBA 797); *saa* pe. ‘to wash, perform ablutions’ (MD 308).

TAIL

1 Nud ʔəpra

1 Mahm ʔəpra

1 VD Urmi ʔəpra

1 U Urmi ʔəpra ‘tail’ (UD 313)

Source:

Etymology is unclear. Cf. MEA: *ʔəprā*, *ʔuprā* ‘nail, claw, talon’ (SL 548); *ʔuprā* ‘fingernail, toenail’ (DJBA 498); *ʔupra* ‘claw, nail’ (MD 178).

TO TAKE

- 1 Nud **šql** I (*šaqəl* | *bəšqala* | *šqalle*)
 1 Mahm **šql** I (*šaqəl* | *bəšqala* | *šqalle*)
 1 VD Urmi **škl** I
 1 U Urmi **škl** I ‘to take, to receive’ (UD 288)

Source:

- # MEA: *šql* pe. ‘to lift up, raise; to bear, carry, carry away; to take’ (SL 1595); *šql*, *šqy* pe. ‘to take; to remove, take out’ (DJBA 1174); *šql* pe. ‘to take away, remove, take up, put up, lift up, carry, raise, reject, repel’ (MD 473).

THAT

- 1 Nud **va**
 1 Mahm **+av**, **+avva**
 1 VD Urmi **+avva**
 1 U Urmi **+avva** ‘that’ (middle deixis demonstrative) (UD 93); **o** ‘that (attributive default demonstrative)’ (UD 85); cf. *+av*, *+avun* ‘he, it’ (UD 91)

Source:

- # *+avva* < **aw-ha* (Khan 2016.1: 240); *o* is a contraction of *aw* (Khan 201.1: 193). Cf. MEA: *haw* ‘that one’ (SL 333).

THERE

- 1 Nud **+tama**
 1 Mahm **+tama**
 1 VD Urmi **+tama**
 1 U Urmi **+tama** ‘there’ (UD 310)

Source:

- # MEA: *tammān* ‘there’ (SL 1653); *tammān* ‘there’ (DJBA 1213), *taman* ‘there’ (MD 479).

THEY

- 1 Nud **ani**
 1 Mahm **ani**
 1 VD Urmi **ani**
 1 U Urmi **ani** ‘those (independent default demonstrative)’ (UD 77)

Source:

- # PNENA **anhi(n)*, **ani* (Hoberman 1990: 85); MEA: *hennon* (m.), *hennēn* (f.) ‘they’ (SL 347); *ʔinnəhū* (m.), *ʔinnəhē* (f.) ‘they’ (DJBA 118, 119); *hinun* (m.), *hinin*, *hinən* (f.) ‘they, those’ (MD 146, 147).

THICK

- 1 Nud **+xlima**
 1 Mahm **+xlima**

1 VD Urmi **+xlima**

1 U Urmi **+xlima** ‘thick, fat (person)’ (UD 338)

Source:

MEA: *hlīmā* ‘sound, healthy person’ (SL 456).

THIGH

1 Nud **+əṭma**

1 Mahm **+əṭma**

1 VD Urmi **+əṭma**

2, 1 U Urmi **nica** ‘thigh’ (UD 243); **+əṭma**, **+uṭma** ‘thigh, lobe of buttocks’ (UD 14, 93)

Source:

#1 MEA: *ṣaṭmā* ‘flank, thigh’ (SL 1093); *ṣaṭmā*, *ṣaṭmā* ‘thigh, foundation wall’ (DJBA 107); *atma* ‘high, side, flank, hip, loin, upper part of the thigh’ (MD 13).

#2 The etymology is unknown. It is also present in Barwar (*neka*; Khan 2008: 1043) and Hamziye (*neka*; NENA DB).

THIN (1D & 2D)

-1, 1 Nud **nazik**, **niqida**; cf. *lavaza* ‘slim’ (about human)

-1 Mahm **nazic** [nazik]

-1 VD Urmi **nazic**

-2, 1 U Urmi **lavaza** ‘thin’ (UD 196); **naqida** ‘thin, slim’ (UD 239); cf. *nazuc* ‘slim, thin, fine, weak, feable’ (UD 243)

Source:

#-1 < Azer. *nazik* ‘thin, fine, delicate’ (AIL 595); Pers. *nāzok* ‘thin, light’ (ODP; UD 243).

#1 MEA: *naqḏā* ‘clean; (gramm.) tenuous’ (SL 944), *naqḏonā* ‘delicate’ (SL 945); *nquḏtā* ‘dot’ (DJBA 772); *nqīd* ‘clean, clear’ (DJBA 775).

#-2 < Kj. *lawaz* ‘thin, slim, slender, lean; puny, undersized’ (KED 346).

TO THINK

1 Nud **txmn** QI (*taxmən* | *taxmunə* | *taxmənle*)

1 Mahm **txmn** QI (*taxmən* | *taxmuna* | *tuxmənle*)

1 VD Urmi **txmn** QI

1 U Urmi **txmn** QI ‘to think, to deliberate, to ponder, to consider, to contemplate’ (UD 306)

Source:

< Kj. *t'exmîn* ‘supposition, guess, assumption’ (KED 613) < Arab. *taḵmîn* ‘appraisal, assessment, estimation’ (DMWA 263; UD 306).

THIS

1 Nud **+avva**

1 Mahm **a**

1 VD Urmi **+avva**

1 U Urmi **'aha**, **'a** ‘this’ (near deixis demonstrative) (UD 76); cf. *+avva* ‘that’ (middle deixis demonstrative) (UD 93).

Source:

< **hāḏā* (Khan 2016.1: 239); MEA: *hāḏ*, *hāḏe* ‘this’ (SL 330); *hāḏe* ‘pron.f. this, that’ (DJBA 358), *hāḏā* ‘pron. this’ (DJBA 361); *aha* ‘this’ (Neo-Mandaic form) (MD 8).

THREE

1 Nud **+ṭla**

1 Mahm **+ṭla**

1 VD Urmi **+ṭla**

1 U Urmi **+ṭla** ‘three’ (UD 317)

Source:

MEA: *tlāṭā* ‘three’ (SL 1651); *tlāṭā* ‘three’ (DJBA 1211); *tlata* ‘three’ (MD 487).

TO THROW

1 Nud **ropy** II (*ripi* | *rupúv* | *rupile*)

1 Mahm **ropy** II (*ripini* | *rupúv* | *rupile*)

1 VD Urmi **+rpØ**, **+rppy** II (*+rappə* | *+rappuvva* | *+ruppile*)

1 U Urmi **+rpØ**, **+rppy** II (*+rappə*) ‘to throw, to hurl’ (UD 268)

Source:

Cf. MEA: *rrp* pe. ‘to move (intr.)’ (SL 1485); *ropy* pe. ‘to be lax, slack’; pa. ‘to loosen, make lax; to let go, yield; to let pass, let go through’; af. ‘to weaken; to send hack, remove; to let go’ (SL 1483); *ropy* pe. ‘to be loose, soft, weak, undecided’; pa. ‘to loosen’ (DJBA 1091); *rpa* pe. ‘to loosen, relax, slacken, weaken, undo’ (MD 436). Cf. also Fassberg 2010: 259.

TO TIE

1 Nud **+Øsr**, **+syr** I (*+ṭasər* | *+bəsyara* | *+sərre*)

1 Mahm **+Øsr**, **+ysr** I (*+asər*, *+yasər* | *+bəssara* | *+sərre*)

1 VD Urmi **+ysr**, **+syr** I (*+yasər* | *+bəsyara* | *+ysərre*)

1 U Urmi **+ysr**, **+syr** I (*+yasər* | *+bəsyara* | *+sərrun*) ‘to tie, to bind, to tie up, to tie together; to twist, to entangle, to bandage, to harness’ (UD 343)

Source:

MEA: *šrr* pe. ‘to tie’ (SL 1304); *šrr*, *šry* pe. ‘to tie up’ (DJBA 974); *šur*, *srr* pe. ‘to surround, tie in’ (MD 391). Cf. MEA: *?sr* pe. ‘to bind, tie’ (SL 79); *?sr*, *?sy* pe. ‘to bind, imprison’ (DJBA 150); *asr*, *ʕsr* pe. ‘to bind, tie up’ (MD 29).

TONGUE

1 Nud **lišana**

1 Mahm **lišana**

1 VD Urmi **lišana**

1 U Urmi **lišana** ‘tongue; language’ (UD 198)

Source:

MEA: *leššānā* ‘tongue’ (SL 698); *liššānā* ‘tongue’ (DJBA 627); *lišana* ‘tongue’ (MD 237).

TOOTH

- 1 Nud **kika**
 1 Mahm **cica** [tʰitʰa]
 1 VD Urmi **cica**
 1 U Urmi **cica** ‘tooth’ (UD 121)

Source:

MEA: *kakkā* ‘molar tooth’ (SL 621); *kakkā* ‘molar tooth, (wooden) handle of a key, outer part of a letter’ (DJBA 580); *kaka* ‘(molar) tooth’ (MD 197).

Comments:

Semantic shift ‘molar tooth’ > ‘tooth’ has taken place in most NENA (Mutzafi 2014: 113; 2018: 296).

TREE

- 1 Nud **+qariya**
 2 Mahm **ilana**
 2 VD Urmi **ilana**
 2 U Urmi **ilana** ‘tree’ (UD 84)

Source:

#1 Etymology is unknown.

#2 MEA: *ʔilānā* ‘tree’ (SL 35); *ʔilānā* ‘tree’ (DJBA 116); *ilana* ‘tree’ (MD 351).

TO TURN (TR.)

- 1 Nud **prtl** QI (*pərtəl* | *pərtúl* | *pərtəlle*)
 1 Mahm **prtl** QI (*partəl* | *pərtula* | *purtəlle*)
 1 VD Urmi **prtl** QI
 1 U Urmi **prtl** QI ‘to spin, to twist (tr. and intr.), to whirl (tr. and intr.)’ (UD 252); cf. *mptl* III ‘to cause to turn, to cause to turn around’ (UD 209)

Source:

MEA: *ptl* pe. ‘to twist; turn aside’ (SL 1268); *ptl* ‘to twist’ (MD 385). Infix *-r-* possibly originated from the dissimilation of the second consonant in D-stem (Sabar 1982: 152, 159; Khan 2016.1: 449).

TWO

- 1 Nud **tre**
 1 Mahm **tre**
 1 VD Urmi **tre**
 1 U Urmi **tre** ‘two’ (UD 308)

Source:

MEA: *trēn* ‘two’ (SL 1666); *trē(n)* ‘two’ (DJBA 1233); *trin* ‘two’ (MD 490).

TO VOMIT

- 1, 2 Nud **gsy** I (*gasə* | *bəgsa* | *gsile*); **+mdØr** III (*+məddər* | *+məddúr* | *+məddərre*)
 2 Mahm **+mdØr** III (*+məddər* | *+muddura* | *+muddərre*)

2 VD Urmi **+mdØr** III (+*maddər* | +*maddura* | +*muddərre*); cf. archaic *gsy* I (*gasə* | *bəgyasa* | *gsile*)
1, 2 U Urmi **jsy** I (*jasə*) ‘to vomit, vomit up’ (UD 163); **+mdØr** III (+*maddər*) ‘to return (tr.); to vomit’ (UD 226)

Source:

- #1 MEA: *gsy* pe. ‘to spit out, belch, vomit; to flow, gush forth’; pa. ‘to disgorge, vomit; to give forth, spit out; to declare, tell’ (SL 250); *gsa* pe. ‘to throw up, vomit, belch, be replete, be glutted, to glut’ (MD 96).
 #2 Causative < +*dØr* ‘to return’ (UD 148); MEA: *dwr* pa. ‘to go from one place to another; to overtake, reach’; af. ‘to overtake’ (DJBA 322).

WARM

1 Nud **šaxina**
1 Mahm **šaxina**
1 VD Urmi **šaxina**
1 U Urmi **šaxina** ‘hot, warm’ (UD 292)

Source:

- # < *šaxən* ‘to become warm, to become hot’ (UD 292); MEA: *šħn* ‘to be inflamed’, pa. ‘to warm, heat’ (SL 1544); *šħn* ‘be inflamed, heat’ (DJBA 1128); *šħn* ‘to become hot’ (MD 451).

TO WASH

1 Nud **+xlll** QI (+*xalləl* | +*xallúl* | +*xalləlle*)
1 Mahm **+xlll** QI (+*xalləl* | +*xullula* | +*xulləlle*)
1 VD Urmi **+xlll** QI
1 U Urmi **1. +xlll** QI ‘to wash; to be washed; to wash (body surface, vessels); to wash face and hands’ (UD 334); cf. *xyp* I (*xayəp*, *xāp*) ‘to wash, to bathe, to take a shower’ (UD 326)

Source:

- # MEA: *ħll* pa. ‘to have permission for o.s., i.e. to enter; to cleanse, wash’ (SL 456); *ħll* pa. ‘to wash’ (DJBA 464); *ħll* pa. ‘to purify, cleanse, rinse, wash, make ritually clean, absolve’ (MD 148).

WATER

1 Nud **miyya**
1 Mahm **miyya**
1 VD Urmi **miyya**
1 U Urmi **miyya** ‘water’ (UD 221)

Source:

- # MEA: *mayyā* ‘water’ (SL 750); *mayyā* ‘water’ (DJBA 662); *mai* ‘water’ (MD 242).

Comments:

The words for ‘water’ are pluralia tantum in all concerned varieties as well as in many other Semitic languages.

WE

1 Nud **axnax**

1 Mahm **axnan**

1 VD Urmi **axnan**

1 U Urmi **'axnan** 'we' (UD 80)

Source:

PNENA **axnan*, *axni* (Hoberman 1990: 85); MEA: *ḥnan*, *ʔnḥnn* 'we' (SL 472, 60); *ʔānan*, *ʔānaḥnā* 'we' (DJBA 145, 144); *anin*, *anʕn* 'we' (MD 27).

TO WEEP

1 Nud **bx̣y** I (*bax* | *bəxya* | *bxile*)

1 Mahm **bx̣y** I (*baxini* | *bəxya* | *bxile*)

1 VD Urmi **bx̣y** I (*bax* | *bəxya* | *bxile*)

1 U Urmi **bx̣y** I (*baxə*) 'to cry, to weep' (UD 101)

Source:

MEA: *bky* pe. 'to weep, cry; to mourn'; pa. 'to lament, bewail' (SL 152); *bky* pe. 'to cry, weep' (DJBA 219), *bka* pe. 'to weep, cry, lament, mourn' (MD 64).

WET

1 Nud **tila**

1 Mahm **tayla**

1 VD Urmi **tayla**

1 U Urmi **tayla** 'wet, damp, moist' (UD 307)

Source:

< **tayila* < **taʔila* (cf. both forms *tayila* & *taʔila* in Marga (NENA DB)); cf. Khan 2016.1: 548; MEA: *talilā* 'adj. humid, moist' (SL 1648), *tll* pa. 'to soil, befoul', *etpa*. 'to be moistened' (SL 1650).

WHAT

1 Nud **mudi**, **mu**

1 Mahm **mu**

1 VD Urmi **mu**

1 U Urmi **mu**, **mudi**, **muy** 'what?' (UD 222)

Source:

mu (> *muy*) is contracted from *mudi* which is probably derived from **ma-məndi* 'what thing?' (Khan 2016.I: 246) < **ma** 'what' & **məndi** 'thing':

ma: MEA: *mā* 'what?' (SL 700); *mā* 'what?' (DJBA 643); *ma* 'what, that which, whatever' (MD 237); *məndi*: MEA: *meddem* 'something' (SL 715); *mīddā*, *mīddē* 'thing, object, pl. something, property, interrogative particle' (DJBA 663); *mindam* 'thing, something' (MD 267).

WHEN

1 Nud **iman**

1, 2 Mahm **iman**, **ə̌m +danta**

1, 2 VD Urmi **iman**, **ə̌m +dana**

1 U Urmi **'iman** 'when?' (UD 84); cf. *ə̌m-+dana* 'at what time?' (UD 81-82)

Source:

- #1 MEA: *ʔemmat* ‘when?’ (SL 58); *ʔimmat* ‘when’ (DJBA 117); *ʔmat* ‘when’ (MD 352).
 #2 < **əm** ‘which?’ (UD 81) & +**dana** ‘time’ (UD 147):
əm: MEA: *ʔaynā* ‘who? which one? what kind of a?’ (SL 36);
 +*dana*: MEA: *ʔedānā* ‘time’ (SL 1073); *ʔiddānā* ‘time’ (DJBA 851); *ʔdana* ‘time, moment, minute, hour, season, epoch’ (MD 341).

WHERE

- 1 Nud **ika**
 1 Mahm **ica** [itʰa]
 1 VD Urmi **ica**
 1 U Urmi **ica** ‘where?’ (UD 84)

Source:

- # MEA: *aykā* ‘where’ (SL 33); *hēkā* ‘where’ (DJBA 376); *haka, ʔka* ‘where’ (MD 120, 348).

WHITE

- 1 Nud +**xvara**
 1 Mahm +**xvara**
 1 VD Urmi +**xvara**
 1 U Urmi +**xvara** ‘white’ (UD 339)

Source:

- # MEA: *ḥewwārā* ‘white’ (SL 432); *ḥiwwār* ‘white’ (DJBA 450); *hiuara* ‘white’ (MD 142).

WHO

- 1 Nud **mini**
 1 Mahm **mini**
 1 VD Urmi **man**
 1 U Urmi **mani, mān** ‘who?’ (UD 207)

Source:

- # MEA: *man* ‘who?’ (SL 778); *man* ‘who’ (DJBA 636); *man* ‘who’ (MD 246).

WIDE

- 1 Nud **pətya**
 1 Mahm **pətya**
 1 VD Urmi **pətya**
 1 U Urmi **pətya** ‘wide, extensive, spacious’ (UD 255)

Source:

- # MEA: *paṭyā* ‘broad, wide’ (SL 1266); *pṭē* ‘wide’ (DJBA 947); *ptia* ‘open, wide’ (MD 384).

WIFE

- 1 Nud **baxta**

1 Mahm baxta

1 VD Urmi **baxta**; cf. *culpat* ‘spouse’

1 U Urmi **baxta** ‘woman; wife, spouse’ (UD 101); cf. *culpat* ‘nuclear family; wife’ (UD 119)

Source:

See WOMAN.

WIND

1 Nud pavxa

1 Mahm **pavxa**

1 VD Urmi **poxa**

1 U Urmi **poxa** ‘wind; breath’ (UD 256)

Source:

MEA: *pāwḥā* ‘odour, smell’ (SL 1161); *pwh* pe. ‘to blow, to breathe’ (SL 1160); *pwh* pe. ‘to breathe, blow up’ (DJBA 888).

WING

1 Nud gulpa

-1 Mahm **parra** ‘wing; feather’

-1 VD Urmi **parra** ‘wing; feather’

1 U Urmi **julpa** ‘wing; fin’ (UD 167)

Source:

#1 MEA: *gelpānā* ‘wing, feather’, *geppā* ‘wing’ (SL 239), 253; *gappā* ‘wing, feather’ (DJBA 297); *gadpa* ‘wing, limb’ (MD 74).

#-1 See FEATHER.

TO WIPE

1 Nud **mšy** II (*miši* | *mušúv* | *mušile*)

1 Mahm **mšy** I/II (*mašini* | *mša*, *mušuvvə* | *mšile*)

1 VD Urmi **mšy** I (*maš* | *mšaya* | *mšile*)

1 U Urmi **mšy** I (*mašə*) ‘to wipe, to wipe up, to swab’ (UD 213; Khan 2016.1: 303)

Source:

MEA: *mšy* pe. ‘to sweep up; to stroke, caress; to collect; to destroy, annihilate; to set aside, remove’; pa. ‘to stroke; to cleanse; to declare rubbish’ (SL 845); *mšy* pe. ‘to wash’; af. ‘to soak an object (in water)’ (DJBA 712), *mša* pe. ‘to wash; (rarely) to measure, stretch out’ (MD 279).

WITH (COMITATIVE)

1 Nud **+am**

2 Mahm **mən**

2 VD Urmi **mən**

2 U Urmi **mən** ‘from; with’ (UD 218; Khan 2016.1: 583); cf. instrumental *b-* ‘with, by, on, in’ (Khan 2016.1: 578). Prep. *ʾam* is used only by “educated people under the influence of the literary language” (Khan 2016.1: 583).

Source:

#1 MEA: *ʕam* ‘with’ (SL 1107); *ʕim* ‘with, together with’ (DJBA 869); *ʕm* ‘with’ (MD 351).

#2 The form *mən* is the result of merging two prepositions **min* ‘from’ and **ʕam* ‘with’ (see #1) (Khan 2016.1: 583). MEA: *men* ‘from, of, how, for, instead of’ (SL 779); *min*, *mi-* ‘from, out of’ (DJBA 683); *mn*, *m-* ‘from, out of, (some) of; with’ (MD 273, 237).

WOMAN

1 Nud **baxta**

1 Mahm **baxta**

1 VD Urmi **baxta**; cf. *culpat* ‘spouse’ at least by some speakers used to designate women in general

1 U Urmi **baxta** ‘woman; wife, spouse’ (UD 101)

Source:

The word is present in all NENA but its etymology is still unclear. Mutzafi (2005: 99; 2008: 337) suggests that *baxta* is a back formation of **bāktātā* (pl. of **bāketṭā*) ‘female spinners’ (cf. the plural form *baxtāṭa* ‘woman, wife’ in many conservative NENA dialects), “for spinning is one of the most important occupations of women throughout the distribution of NENA” (Mutzafi 2005: 99). Cf. Syr. *bāketṭā* ‘weaving woman; weft’ < *bkt* pe. ‘to spin’ (SL 153). Another possibility is a euphemism from the word *baxt* ‘luck; good fortune, fate, honour’ (borrowed from Kj. *bext* ‘honour; luck, chance, fate’ (KED 48)) in a society where direct reference to women was to be avoided (Khan 1999, 146–147) but such semantic shift is not known in other languages of the region.

WORM

1 Nud **+tvəlla**

1 Mahm **+tvəlla**

1 VD Urmi **+təvvəltā**

1 U Urmi **+təvvəltā**, **+tuvvūltā** ‘worm, a stomach worm, caterpillar’ (UD 312)

Source:

MEA: *tāwlʕā* ‘worm’ (SL 1630); *tōlaʕtā* ‘worm’ (DJBA 1197); *tulita* ‘worm, embryo’ (MD 483).

YEAR

1 Nud **šita**

1 Mahm **šita**

1 VD Urmi **šita**

1 U Urmi **šita** ‘year’ (UD 295)

Source:

MEA: *šattā* ‘year’ (SL 1581); *šattā* ‘year’ (DJBA 1183); *šita* ‘year’ (MD 464).

YELLOW

-1 Nud **zarda**

-1 Mahm **zarda**

-1 VD Urmi **zarda**

-1 U Urmi *zarda* ‘yellow’ (UD 346)

Source:

< Kj. *zar* (ESKJ 510), *zer*, *zerd* [zær, -d] (KRS 803) ‘yellow’.

YESTERDAY

1 Nud *qudmə*

1 Mahm *kudmə*

1 VD Urmi *kudmə*

1 U Urmi *ḵudmə* ‘yesterday; tomorrow’; for disambiguation the construction *ḵudmət +vərro* (lit. ‘yesterday that passed’) can be used (UD 184).

Source:

< **qadmayit* ‘formerly’ or < the plural form of the noun *qudmā* (Napiorkowska 2015: 105). MEA: *qadmayit* ‘at first’ (SL 1318); *qudmā* ‘principal part; front part; in front; early; before; in the presence of; from before; from all eternity’ (SL 1324); *qudam* ‘before, preceding, in front of, anterior’ (MD 406).

YOU (PL.)

1 Nud *axtu*

1 Mahm *axnoxun*

1 VD Urmi *axnoxun*

1 U Urmi *’axtun, ’axtoxun, ’axnoxun* ‘you (pl.), you (sg. polite)’ (UD 80)

Source:

PNENA **axtun* (Hoberman 1990: 85); MEA: *ʔatton* ‘you (pl. m.)’ (SL 66); *ʔattu* ‘you (pl. m.)’ (DJBA 175); *anatun* ‘you (pl.)’ (MD 25). The *x* in *’axtun* appeared by analogy with pronoun *’axnan* ~ *’axni* ‘we’ and/or with 2pl. pron. suffix *-oxun* (Hoberman 1990: 86; Mutzafi 2008: 329).

YOU (SG.)

1 Nud *at*

1 Mahm *atən*

1 VD Urmi *atən*

1 U Urmi *’at, ’atən* ‘you (sg.)’ (UD 79)

Source:

PNENA **at* (Hoberman 1990: 85); MEA: *ʔat* ‘you’ (m.sg.) (SL 66); *ʔant* ‘you’ (m.sg.) (DJBA 146); *anat, at* ‘pers. pron. thou’ (MD 24, 41).

5. Results and discussion

5.1. The cognate shares

Using the collected data, we calculated the cognate share for Swadesh-100 and for the extended wordlists (see 2.2). Out of 226 meanings in the compiled lists, six meanings (‘leg’,

‘lake’, ‘dull (not sharp)’, ‘berry’, ‘to float’, ‘to flow’) are not represented by separate lexemes in at least three varieties. Therefore, in the following, we use only 220 words (further referred to as “220 list”). On this stage, we also excluded the loanwords from the calculations since they do not indicate genealogical relationships and the substitution of original words by loans is regarded as chance mutations (Starostin 2000: 6; Starostin 2010: 85; Campbell 2003: 271). The results are provided in Table 3. The cognate shares are given in different columns: the first one is based on the Swadesh-100 list and the second one on the 220 list.

	Mahm		VD Urmi		U Urmi	
	<i>Sw-100</i>	<i>220 list</i>	<i>Sw-100</i>	<i>220 list</i>	<i>Sw-100</i>	<i>220 list</i>
Nud	96.8	93.4	96.8	93.4	95.7	94.6
Mahm			97.9	98.5	97.8	97.0
VD Urmi					98.9	97.0

Table 3. Cognate shares between the varieties, loans excluded (in percents)

As we can see from Table 3, all the varieties in question are close enough to be treated as dialects of a single language. Since Mahmudi and Nudiz belong to the Hakkari NENA cluster (Khan 2018a: 191), we suggested that they were linguistically closer to each other than to the Urmi varieties (see 2.1). However, based on Swadesh-100 (loans excluded), Nud, Mahm and VD Urmi are equally close to each other. Moreover, if we extend the lists (using the 220 list), Mahm and VD Urmi appear to be closer to each other than to Nud. It might indicate that similarity between Mahmudi and Urmi is more of a contact origin. As Starostin (2010: 92) argues, “testing relationship hypotheses on different chronological levels requires wordlists of different sizes”, with more extensive lists needed for closely related varieties.

Additionally, in order to estimate synchronic similarity (and not genealogical relationship), we calculated the cognate share for both lists with loanwords included. The results (see Table 4) show that Mahmudi and Verin Dvin Urmi are closer to each other than to Nudiz. This finding aligns with Thomason, Kaufman’s (1988) framework, wherein intensive contact creates secondary convergence zones that disrupt traditional genealogical classifications.

	Mahm		VD Urmi		U Urmi	
	<i>Sw-100</i>	<i>220 list</i>	<i>Sw-100</i>	<i>220 list</i>	<i>Sw-100</i>	<i>220 list</i>
Nud	94.0	91.3	94.0	90.4	92.0	91.8
Mahm			98.0	98.2	96.0	94.5
VD Urmi					97.0	95.5

Table 4. Cognate shares between the varieties, loans included (in percents)

Lexical patterns are considered in detail below (see 5.2). On the phonological and morphological differences see 2.1.

5.2. Lexical patterns

The lexical proximity of VD Urmi to Mahmudi is evident in 20 cases (out of 220 in our dataset). They can be divided into four patterns, presented in Table 5. The patterns are as follows:

- 1a. U Urmi, VD Urmi and Mahm have the same lexeme and Nud has a different one;
 1b. U Urmi, VD Urmi and Mahm have two lexemes and Nud has only one of them;
 2. VD Urmi and Mahm have the same lexeme, Nud has another and U Urmi has both;
 3. VD Urmi and Mahm have the same lexeme and Nud and U Urmi have a different but the same lexeme.

Pattern	Nud	Mahm	VD Urmi	U Urmi	Glosses
Swadesh -100					
1	<i>bəbla</i>	<i>varda</i>	<i>varda</i>	<i>varda</i>	flower
1	<i>mila</i>	<i>kina</i>	<i>kina</i>	<i>ḵina</i>	green
1	<i>qam</i>	<i>kavya</i>	<i>kuya</i>	<i>ḵuya</i>	hard
1	<i>xiza</i> L	<i>silā</i> L	<i>silā</i> L	<i>silā</i> L	sand
1	<i>+qariya</i>	<i>ilana</i>	<i>ilana</i>	<i>ʾilana</i>	tree
1	<i>+am</i>	<i>mən</i>	<i>mən</i>	<i>mən</i>	with
1b	<i>Ømr</i>	<i>Ømr; tny</i>	<i>Ømr; tny</i>	<i>Ømr; tny</i>	say, to
Other words					
1	<i>xmy</i>	<i>xyØ</i>	<i>xyØ</i>	<i>xyØ</i>	live, to
1	<i>gədda</i>	<i>+xola</i>	<i>+xola, gdala</i>	<i>+xola</i>	rope
1b	<i>+blç</i>	<i>+blç, bls; +mrç</i>	<i>+blç; +mrç</i>	<i>+blç, bls; +mrç</i>	crush, to
1b	<i>prm</i>	<i>prm; +kɬØ</i>	<i>+kɬØ</i>	<i>prm; +kɬØ</i>	cut, to
2	<i>xərba</i>	<i>pis</i> L	<i>pis</i> L	<i>xərba; pīs</i> L	bad
2	<i>+spay</i> L	<i>ğəns</i> L	<i>ğəns</i> L	<i>+spay</i> L; <i>jəns</i> L	good
2	<i>ən</i>	<i>aga</i> L	<i>aga</i> L	<i>ʾən; ʾǰar</i>	if
2	<i>gura</i>	<i>urza</i> L, <i>ğvanka</i>	<i>ğvanka, urza</i> L	<i>ʾurza; +jora</i>	man (adult male)
3	<i>dax</i>	<i>muğğurra</i>	<i>muğğā, muğğurra</i>	<i>dax, dāxi</i>	how
3	<i>+sayduvva Øvd</i>	<i>+avčuvva Øvd</i>	<i>+avčuyta Øvd</i>	<i>+syd</i>	hunt, to
3	<i>šmayya</i>	<i>dunya</i>	<i>dənyə</i>	<i>šmayya</i>	sky
3	<i>iman</i>	<i>iman, əm +danta</i>	<i>iman, əm +dana</i>	<i>ʾiman</i>	when
3	<i>gulpa</i>	<i>parra</i> L	<i>parra</i> L	<i>julpa</i>	wing

Table 5. Patterns of lexical proximity of VD Urmi to Mahmudi. The loanwords are marked with “L”

Pattern 3 is interesting because it presents cases where varieties spoken in Verin Dvin (Mahm and VD Urmi) share common innovations specific to them.

At the same time Mahmudi and Nudiz have only four shared lexemes (out of 220 words) that are different from Urmi varieties, see Table 6.

Nud	Mahm	VD Urmi	U Urmi	Glosses
<i>+šrɬx</i>	<i>+šrɬx</i>	<i>+ryx</i>	<i>+ryx</i>	to lie (in bed)
<i>+əryana</i>	<i>+aryana</i>	<i>+məɬra</i>	<i>+muɬra</i>	rain
<i>da</i> L	<i>da</i> L	<i>yəmma</i>	<i>yəmma</i>	mother
<i>Øzl</i>	<i>Øzl</i>	<i>Øzl; rxš</i>	<i>Øzl; rxš</i>	to go

Table 6. Patterns of lexical proximity of Nudiz to Mahmudi

5.3. Loanwords

NENA varieties exhibit numerous loanwords, both within the Swadesh-100 list and within the extended 220-word list. We evaluated the distribution of loanwords by source language (Table 7) and by target variety (Table 8).

Source of borrowing	Swadesh-100	Other words	Total words
Pers. or Kj./Sor. or Az.	feather/wing	if	2
Kj./Sor. or Az.	new	animal, bad, because, dust, forest, right, smooth/correct	8
Kj./Sor.	bark, good, liver, sand, yellow	fog, left, thin, mother	9
Pers.	green	because, to breathe	3
Az.	root	to hunt, thin	3
Arab.	breast (chest)		1
Total			26

Table 7. Loanwords distribution by source language

The exact source language for some borrowings is unclear, see 3.3. In general, the most common sources of borrowings are Kurdish (Kj/Sor) or Kurdish/Azerbaijani.

NENA variety	Swadesh-100	Other words	Total words
Nud	7	10	17
Mahm	6	14	20
VD Urmi	6	13	19
U Urmi	9	9	18

Table 8. Loanwords distribution by target NENA variety

All varieties demonstrate comparable shares of loanwords, ranging from 17 to 20 in total, as can be seen from Table 8. However, the data also suggest the possibility of interdialectal borrowing that may be obscured by the close genetic relationship between these varieties. The inconsistent phonological realization of original velars in Mahmudi (see Section 3.1) — particularly words displaying innovative affricate realizations — may indicate borrowed lexical items. While loanwords are traditionally considered problematic for genealogical classification due to their unsystematic nature, in this case they are likely to reflect the recent contact history of Mahmudi and Verin Dvin Urmi.

6. Conclusion

In this study, we have presented and analyzed lexical data from four Christian NENA varieties: Mahmudi, Nudiz, Verin Dvin Urmi, and Urmia Urmi, contributing to the documentation of the endangered and previously undescribed Mahmudi and Nudiz varieties. For all lexemes of the extended 220 list of basic vocabulary, we have established etymologies and provided cognates from earlier Aramaic languages. We also identified loanwords and traced their sources.

At the next stage, we calculated cognate shares for different subsets of our dataset. Our analysis reveals that all four varieties are quite close to each other and should be classified as dialects of a single language. However, what other dialects are to be included in this language and, consequently, what it should be called, remains to be found out in future studies.

We also recalculated cognate shares with loanwords included, trying to estimate the synchronic similarity and mutual intelligibility between lects. It appeared that two varieties spoken in one village for the last 67 years (Mahmudi and Verin Dvin Urmi) were closer to each other than to the varieties to which they are genealogically closer (Mahmudi to Nudiz and VD Urmi to U Urmi, respectively), even though Mahm and VD Urmi belong to different clusters and display several phonological and morphological dissimilarities.

Thus, the observed convergence between the two varieties demonstrates how intensive language contact can obscure genealogical patterns and complicate the subgrouping of closely related dialects based on the lexical distance. These results call for revised subgrouping criteria that account for both historical inheritance and contact-induced convergence, particularly when analyzing dialects with shared territorial history. A combination of comparative phonological, morphological analysis, and dialectometry would be the optimal direction for future research, yielding robust results for the internal classification of NENA.

Abbreviations

Languages:

Arab. = Arabic; Az. = Azerbaijani; Kj. = Kurmanji (Kurdish); Kurd. = Kurdish (Kurmanji or Sorani); Mahm = Mahmudi; MEA = Middle Eastern Aramaic; NENA = North-Eastern Neo-Aramaic; Nud = Nudiz; Pers. = Persian; PNENA = Proto-North-Eastern Neo-Aramaic; Sor. = Sorani (Kurdish); Syr. = Syriac; U Urmi = Urmia Urmi; VD Urmi = Verin Dvin Urmi

Grammatical:

2 = second person; ADD = additive particle; adj. = adjective; adv. = adverb; af. = *af̄el*, the causative verbal stem in Middle Aramaic; conj. = conjunction; ethpa. = *ethpaʿʿal*, one of Middle Aramaic detransitive verbal stems; ethpe. = *ethpeʿel*, one of Middle Aramaic detransitive verbal stems; ettaf. = *ettaʿʿal*, one of Middle Aramaic detransitive verbal stems; f. = feminine; I = the unmarked verbal pattern (stem) in Neo-Aramaic; II = one of the causative verbal pattern (stem) in Neo-Aramaic; III = one of the causative verbal pattern (stem) in Neo-Aramaic; intr. = intransitive verb; m. = masculine; mod. = modifier, modifying; n. = noun; onom. = onomatopoeic; pa. = *paʿʿel*, the intensive verbal stem in Middle Aramaic; pe. = *peʿal*, the unmarked verbal stem in Middle Aramaic; pl. = plural; prep. = preposition; pron. = pronoun; QI = quadriliteral verbal pattern (stem) in Neo-Aramaic; sg. = singular; tr. = transitive verb; t-stems = Middle Aramaic detransitive verbal stems

References

- ADW = Rahmati, Nemat. 1999. *Aserbajdschanisch-deutsches Wörterbuch: unter Berücksichtigung der Besonderheiten des Nord- und Südaserbajdschanischen*. Engelschoff: Verlag auf dem Ruffel.
- AIL = Hacıyev, E.İ. 2005. *Azərbaycanca-İngiliscə lüğət*. Bakı: “Şərq-Qərb”.
- AM = 2024. Ararat Marz. In: *Atlas the Republic of Armenia by the regions and Yerevan City*: 311–318. Yerevan: Statistical Committee Republic of Armenia.
- Aroutunian, Diana, Susanna Aroutunian. 1993. *Armenian-English / English-Armenian*. New York: Hippocrene Books.
- Badalyan, Gegham. 2015. *Atlas of Van-Vaspurakan*. Yerevan: Vaspurakan Society-Los Angeles Chapter (USA).
- Barsky, Eugene, Maksim Kalinin, Sergey Loesov. 2018a. The Basic Lexicon of Mlaḥsô: A Comparative Study. *Aula Orientalis* 36(2): 209–239.

- Barsky, Eugene, Yulia Furman, Sergey Loesov. 2018b. Two-Hundred-Word Swadesh List for a Modern Aramaic variety (Turoyo). *Aula Orientalis* 36(1): 75–110.
- Beyer, Klaus. 1986. *The Aramaic Language: Its Distribution and Subdivisions*. Göttingen: Vandenhoeck & Ruprecht.
- CAL = Kaufman, Stephen A. et al. *The Comprehensive Aramaic Lexicon*. Available online at: <https://cal.huc.edu> [Accessed 30.10.2025].
- Campbell, Lyle. 2003. How to show languages are related: methods for distant genetic relationship. In: B. D. Joseph, R. D. Janda. *Handbook of historical linguistics*: 262–282. Oxford: Blackwell.
- Cereteli, Konstantin. 1963. Obrazcy sovremennoj assirijskoj reči (Vanskije teksty) [Samples of the Modern Assyrian speech (Van texts)]. In: G. Š. Šarbatov (ed.). *Semitskie jazyki. Sbornik statej*. Vyp. 1: 202–218. Moscow: Izdatel'stvo vostočnoj literatury.
- CPED = Steingass, Francis J. 1892. *A Comprehensive Persian English Dictionary*. London: Routledge & Kegan Paul Limited.
- DJBA = Sokoloff, Michael. 2002. *A Dictionary of Jewish Babylonian Aramaic of the Talmudic and Geonic Periods*. Ramat-Gan, Israel: Bar Ilan University Press.
- DJPA = Sokoloff, Michael. 1992. *A Dictionary of Jewish Palestinian Aramaic of the Byzantine Period*. Ramat-Gan, Israel: Bar Ilan University Press.
- DKF = Bedir Khan, Kamuran A., Joséfa Bertolino, Kendal Nezan. 2017. *Dictionnaire Kurde-Français*. Paris: Riveneuve.
- DMWA = Wehr, Hans. 2011. *A Dictionary of Modern Written Arabic* (Edited by J. Milton Cowan). N.p.: Snowball Publishing.
- Dum-Tragut, Jasmine. 2013. Assyrians in Armenia: an interdisciplinary survey. In: L. Tang, D. W. Winkler. *From the Oxus river to the Chinese shores: studies on East Syriac Christianity in China and Central Asia*: 341–353. Zürich / Berlin: Lit Verlag.
- ESKJ = Cabolov, Ruslan Lazarevič. 2010. *Ètimologičeskij slovar' kurdsogo jazyka* [Etymological dictionary of the Kurdish language]. 2nd vol. Moscow: Vostočnaja literatura.
- Fassberg, Steven. 2010. *Jewish Neo-Aramaic dialect of Challa*. Leiden: Brill.
- Fox, Samuel E. 2009. *The Neo-Aramaic Dialect of Bohtan*. Piscataway, NJ: Gorgias Press.
- Fox, Samuel E. 2015. *a > e* raising in the Jilu dialect of Neo-Aramaic. *Aramaic Studies* 13(2): 161–173.
- Heggarty, Paul et al. 2023. Language trees with sampled ancestors support a hybrid model for the origin of Indo-European languages. *Science* 381/6656.
- Hoberman, Robert D. 1989. *The Syntax and Semantics of Verb Morphology in Modern Aramaic: A Jewish Dialect of Iraqi Kurdistan*. New Haven: American Oriental Society.
- Hoberman, Robert D. 1990. Reconstructing Pre-Modern Aramaic morphology: the independent pronouns. In: W. Heinrichs (ed.). *Studies in Neo-Aramaic*: 79–88. Atlanta, Georgia: Scholars Press.
- Hovhannisyan, Hasmik. 2007, October 29. Three Essential Issues Facing Assyrians in Armenia (Interview with Razmik Khosroev). *Hetq*. Available online at: <https://hetq.am/en/article/25486> [Accessed 02.05.2025].
- IE-CoR Meanings = Heggarty, Paul, Cormac Anderson, Matthew Scarborough. 2023. *Indo-European Cognate Relationships Database. Meanings*. Available online at: <https://iecor.cld.org/parameters> [Accessed 02.05.2025].
- Kassian, Alexei, George Starostin, Anna Dybo, Vasilij Chernov. 2010. The Swadesh wordlist. An attempt at semantic specification. *Journal of Language Relationship* 4: 46–89.
- KED = Chyet, Michael L. 2003. *Kurdish-English Dictionary*. New Haven / London: Yale University Press.
- Khan 2016.1 = Khan, Geoffrey. 2016. *The Neo-Aramaic dialect of the Assyrian Christians of Urmi. Volume 1. Grammar: phonology and morphology*. Leiden / Boston: Brill.
- Khan 2016.2 = Khan, Geoffrey. 2016. *The Neo-Aramaic dialect of the Assyrian Christians of Urmi. Volume 2. Grammar: syntax*. Leiden / Boston: Brill.
- Khan 2016.4 = Khan, Geoffrey. 2016. *The Neo-Aramaic dialect of the Assyrian Christians of Urmi. Volume 4. Texts*. Leiden / Boston: Brill.
- Khan, Geoffrey. 2007. The North-Eastern Neo-Aramaic Dialects. *Journal of Semitic Studies* 52 (1): 1–20.
- Khan, Geoffrey. 2011. North-Eastern Neo-Aramaic. In: S. Weninger (ed.). *Semitic Languages: An International Handbook*: 708–724. Berlin / Boston: De Gruyter Mouton.
- Khan, Geoffrey. 2016. *The Neo-Aramaic dialect of the Assyrian Christians of Urmi (4 vols)*. Leiden / Boston: Brill.
- Khan, Geoffrey. 2018a. The Neo-Aramaic dialects of eastern Anatolia and northwestern Iran. In: G. Haig, G. Khan (eds.). *The Languages and Linguistics of Western Asia: An Areal Perspective*: 190–236. Berlin / Boston: De Gruyter Mouton.

- Khan, Geoffrey. 2018b. The Neo-Aramaic dialects of northern Iraq. In: G. Haig, G. Khan (eds.). *The Languages and Linguistics of Western Asia: An Areal Perspective*: 305–353. Berlin / Boston: De Gruyter Mouton.
- Khan, Geoffrey. 2018c. The Neo-Aramaic dialects of western Iran. In: G. Haig, G. Khan (eds.). *The Languages and Linguistics of Western Asia: An Areal Perspective*: 481–532. Berlin / Boston: De Gruyter Mouton.
- Kim, Ronald. 2008. “Stammbaum” or continuum? The subgrouping of Modern Aramaic dialects reconsidered. *Journal of the American Oriental Society* 128(3): 505–531.
- Kogan, Leonid. 2015. *Genealogical Classification of Semitic: The Lexical Isoglosses*. Berlin / München / Boston: De Gruyter.
- Krotkoff, Georg. 1985. Studies in Neo-Aramaic Lexicology. In: Ann Kort, Scott Morschauer (eds.). *Biblical and Related Studies Presented to Samuel Iwry*: 123–134. Winona Lake, Indiana: Eisenbrauns.
- KRS = Kurdoev, Kanat Kalaševič. 1960. *Kurdsko-russkij slovar'* [Kurdish-Russian dictionary]. Moscow: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovar'ej.
- Lane, Edward W. 1863 [1968]. *An Arabic-English lexicon. In eight parts. Part 1*. Beirut: Librairie du Liban.
- Lyavdansky, Alexey. 2014. Zametki o leksičeskix novoarameizmax v klassičeskom sirijskom [Notes on Lexical Neo-Aramaicisms in Classical Syriac]. *Vestnik RGGU. Serija „Vostokovedenie. Afrikanistika“* 6: 51–65.
- Lyavdansky, Alexey. 2021. A Corpus-Based Swadesh Word List for Literary Christian Urmi (New Alphabet Texts). In: G. Khan, P. M. Noorlander (eds.). *Studies in the Grammar and Lexicon of Neo-Aramaic*: 415–442. Cambridge, UK: Open Book Publishers.
- MD = Drower, Ethel S., Rudolph Macuch. 1962. *A Mandaic dictionary*. Oxford: Clarendon.
- Murre-van den Berg, Heleen L. 1999. *From a spoken to a written language. The introduction and development of literary Urmia Aramaic in the nineteenth century*. Leiden: Nederlands Instituut voor het Nabije Oosten.
- Mutzafi, Hezy. 2005. Etymological notes on North-Eastern Neo-Aramaic. *Aramaic Studies* 3(1): 83–107.
- Mutzafi, Hezy. 2006. On the etymology of some enigmatic words in North-Eastern Neo-Aramaic. *Aramaic Studies* 4(1): 83–99.
- Mutzafi, Hezy. 2008. *The Jewish Neo-Aramaic dialect of Betanure (province of Dihok)*. Wiesbaden: Harrassowitz Verlag.
- Mutzafi, Hezy. 2014. *Comparative lexical studies in Neo-Mandaic*. Leiden / Boston: Brill.
- Napiorkowska, Lidia. 2015. *A Grammar of the Christian Neo-Aramaic Dialect of Diyana-Zariwaw*. Leiden / Boston: Brill.
- NENA DB = Khan, Geoffrey (ed.). 2016. *The North-Eastern Neo-Aramaic Database Project*. Available online at: <https://nena.ames.cam.ac.uk/> [Accessed 02.05.2025].
- ODP = Makarov, Yury, Dmitry Novokshanov. Online dictionary of Persian. Available online at: <https://iranica.space> [Accessed 30.10.2025].
- Ovsjannikova, Maria, Kirill Kozhanov, Natalia Logvinova, Alexey Lyavdansky, Alina Russkikh, Sergey Say, Varvara Shuvalova, Elena Shvedova, Elizaveta Zabelina. 2022–. *Corpus of NENA varieties spoken in Russia*. Available online at: <https://nenadict.iling.spb.ru/corpus> [Accessed 30.10.2025].
- Ovsjannikova, Maria, Elizaveta Zabelina, Kirill Kozhanov. Forthc. Non-Urmi Northeastern Neo-Aramaic varieties in the village of Urmiya, Russia. Part 1: Sociolinguistic overview and phonetic features. *Iran and the Caucasus*.
- Qazzaz, Shafiq. 2000. *The Sharezoor Kurdish-English Dictionary*. Erbil, Kurdistan: Aras Press and Publishers.
- Sabar, Yona. 1982. The quadriradical verb in Eastern Neo-Aramaic dialects. *Journal of Semitic Studies* 27: 149–175.
- Sabar, Yona. 2002. *A Jewish Neo-Aramaic dictionary. Dialects of Amidya, Dihok, Nerwa and Zakho, Northwestern Iraq*. Wiesbaden: Harrassowitz.
- Sado, Stephen. 2017. *Martirolog assirijtsev SSSR (1920–1950-e gody)* [Martyrology of the Assyrians of the USSR (1920–1950s)]. Saint-Petersburg: NP-Print.
- Sado, Stephen. 2021. *Biografičeskij slovar' assirijtsev v Rossii (XIX — seredina XX veka)* [Biographical dictionary of Assyrians in Russia: XIX — the middle of the XX centuries]. Saint-Petersburg: NP-Print.
- SL = Sokoloff, Michael. 2009. *A Syriac lexicon*. Winona Lake / Piscataway: Gorgias Press.
- Starostin, Sergei S. 2000. Comparative-historical Linguistics and Lexicostatistics. In: Colin Renfrew, April McMahon, R. L. Trask (eds.). *Time Depth in Historical Linguistics*: 223–259. Cambridge: McDonald Institute for Archaeological Research, Oxford Publishing Press.
- Starostin, George. 2010. Preliminary Lexicostatistics As A Basis For Language Classification: A New Approach. *Journal of Language Relationship* 3: 79–116.
- Swadesh, Morris. 1952. Lexicostatistic dating of prehistoric ethnic contacts. *Proceedings of the American Philological Society* 96: 452–463.
- Swadesh, Morris. 1955. Towards greater accuracy in lexicostatistic dating. *International Journal of American Linguistics* 21: 121–137.

- Tadmor, Uri. 2009. Loanwords in the world's languages: Findings and results. In: M. Haspelmath, U. Tadmor (eds.). *Loanwords in the World's Languages: A Comparative Handbook*: 55–75. Berlin / New York: De Gruyter Mouton.
- Talay, Shabo. 2008. *Die Neuaramäischen Dialekte Der Khabur-Assyrer in Nordostsyrien: Einführung, Phonologie und Morphologie*. Wiesbaden: Harrassowitz Verlag.
- Teule, Herman, Grigory Kessel. 2012. The Mikhail Sado Collection of Syriac Manuscripts in St. Petersburg. In: J. P. Monferrer-Sala, H. Teule, S. Torallas Tovar (eds.). *Eastern Christians and their written heritage: manuscripts, scribes and context*: 43–76. Leuven / Paris / Walpole, MA: Peeters.
- Thomason, Sara G., Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- TS = Payne Smith, Jessie (ed.). 1903 (1957). *A Compendious Syriac Dictionary founded upon the thesaurus syriacus of R. Payne Smith*. Oxford: Clarendon Press.
- UD = Khan, Geoffrey. 2016. *The Neo-Aramaic dialect of the Assyrian Christians of Urmi. Volume 3. Lexical studies and dictionary*. Leiden / Boston: Brill.
- Wilmshurst, David. 2000. *The ecclesiastical organisation of the Church of the East, 1318-1913*. Leuven: Peeters.
- Zabelina et al. 2022– = Zabelina, Elizaveta, Maria Ovsjannikova, Sergey Say. 2022–. *Sound dictionary of NENA varieties spoken in Russia*. Available online at: <https://nenadict.iling.spb.ru/dictionary> [Accessed 30.10.2025].
- Zabelina, Elizaveta. 2023. *The NENA variety of Nudiz* [Conference presentation abstract]. 5th Neo-Aramaic Languages Conference, October 26-27, 2023, Boğaziçi University, Istanbul, Türkiye. Available online at: <https://nalcon2023.bogazici.edu.tr/sites/nalcon2023.boun.edu.tr/files/zabelina.pdf>

Е. Е. Шведова, Ю. Б. Коряков, Е. А. Забелина. Количественная оценка лексических различий между нудезнайским, мамедайским и верин-двинским урмийским (северо-восточные новоарамейские языки)

В настоящем исследовании анализируются лексические данные четырех христианских северо-восточных новоарамейских идиомов: мамедайского, нудезнайского, урмийского, распространенного в с. Верин-Двин (Армения), и урмийского, распространенного в Иране. Мамедайский и нудезнайский диалекты ранее не были описаны в литературе. Приводятся соответствия между четырьмя идиомами на основе расширенного 226-словного списка базовой лексики, собранной для этого исследования, с указанием этимологии, родственных слов из более ранних арамейских языков и источников заимствований. Расчеты доли родственных слов показывают, что все четыре идиома довольно близки друг другу и могут считаться диалектами одного языка. Примечательно, что в результате почти семи десятилетий совместного бытования в условиях единого населенного пункта мамедайский и верин-двинский урмийский идиомы демонстрируют более сильную конвергенцию между собой, чем с исторически более близкими им идиомами.

Ключевые слова: семитские языки; арамейские языки; северо-восточные новоарамейские; диалектология; лексикостатистика; языковые контакты; лексическая семантика.

Hurrian and Ugaritic sibilants: Essay of phonetic interpretation

The phonetic values of the sibilant phonemes in Hurrian and Ugaritic must be revisited in the light of the evolving interpretation of sibilants in neighbouring languages, which is partly due to the evolution of the Affricate Hypothesis. We are arguing that the syllabic cuneiform orthography of Mitanni and Boğazköy Hurrian uses the Š-series to reflect the alveolar fricative phonemes /s/ and /z/, while the Z and S-series are deployed for the rendering of affricate phonemes. Furthermore, we are suggesting that the Ugaritic alphabetic grapheme (t) renders not the interdental /θ/, as was traditionally assumed, but the alveolar /s/. The revised sound values of the Hurrian simplify the scenario of how the Mesopotamian cuneiform script was adapted to writing Hittite.

Keywords: Hurrian language; Ugaritic language; sibilant phonemes; Affricate Hypothesis.

1. Sibilants in Egyptian, Hittite, and Akkadian

The phonetic interpretation of extinct languages recorded in ancient scripts requires a conjoined application of several research methods.¹ First, if a particular grapheme was used in a variety of languages, one can assess the mutual compatibility of its assumed phonetic values. Second, if a language family is associated with a variety of writing systems, one can investigate how historically related sounds are represented across various scripts. Third, one can study borrowings among languages recorded in different scripts with the aim of assessing matches among individual phonemes of loanwords and their sources. Although the correct phonetic interpretation is likely to maximize the number of exact correspondences in all the three cases, it is obvious that differences will persist. The graphemes sometimes alter their values across languages, historical sound changes do occur, while the adaptation of loanwords frequently implies phonetic changes of their own. In such cases, one strives for a solution that enhances the naturalness of the proposed correspondences. In addition to using the comparative method, one must assess the typological plausibility of a sound system (or its fragment) that is reconstructed for a particular extinct language.

¹ This paper combines the ideas in the presentation “Phonetic Interpretation of Hurrian Sibilants in the Light of Indo-European Evidence” offered by Ilya Yakubovich at the conference *The Sound of Indo-European: Phonetics, Phonemics, and Morphophonemics* (Copenhagen, April 2009) and the guest lecture “The evolution of the Semitic sibilants in Ugaritic and Phoenician” presented by Vera Tsukanova at the University of Naples *L’Orientale* in March 2024. We are grateful to the participants of both events for their constructive feedback. We are also much obliged to Sebastian Fischer (Berlin), who provided detailed comments on the preliminary version of this article and shared with us the relevant parts of his Habilitation thesis ahead of its publication. The substantial comments by Mauro Giorgieri (Pavia), Leonid Kogan (Moscow), and Roey Schneider (Jena) are also sincerely acknowledged, while Stepen Durnford (Brighton) and H. Craig Melchert (Carrboro, NC) helped to improve the style of this paper. Ilya Yakubovich contributed to the joint effort under the auspices of the project *Tive: Studies in the Hurrian Language and Texts* (YA 472/4-1), which is funded by the *Deutsche Forschungsgemeinschaft*. The final part of the work on the article was accomplished during the co-authors’ research visit to the University of Naples *L’Orientale* in October–November 2025, which was kindly organized by Michele Cammarosano.

These theoretical and somewhat abstract observations can be fleshed out with a look at the phonetic interpretation of some fricative and affricate coronal sounds (“sibilants”) in the most well-studied Ancient Near Eastern languages of the second millennium BCE.² Thus, it is widely believed that Ancient Egyptian of the Middle and New Kingdoms possessed two contrastive coronal fricative phonemes /s/ and /ʃ/, whose alveolar and palato-alveolar articulations were initially projected back from their Coptic reflexes (Peust 1999: 116, 125–126, Loprieno 2004: 171).³ In the Egyptian hieroglyphic script of the second millennium BCE, the phoneme /s/ could be written with two phonograms, conventionally transcribed ⟨s⟩ and ⟨ś⟩, or sometimes ⟨z⟩ and ⟨s⟩. These two phonograms had been contrastive in the third millennium BCE but came to function as allographs from the Middle Kingdom on (Peust 1999: 126). The phonogram used for the rendering of /ʃ/ is conventionally transcribed as ⟨š⟩. The distinction between /s/ and /ʃ/ in Egyptian appears to be fairly stable, and the analysis of loanwords in New Kingdom Egyptian is used as a logical starting point for studying the articulation of sibilants in the contemporary Ancient Near Eastern languages. In addition to the fricatives, Ancient Egyptian also featured sibilant affricates, which will not be addressed in this paper, since they are irrelevant to its main purpose.

In contrast, the Hittite coronal fricatives presumably had the same place of articulation but were distinguished through the opposition tense vs. lax.⁴ The syllabic signs ⟨ša⟩, ⟨še⟩, ⟨ši⟩, ⟨šu⟩, ⟨aš⟩, ⟨eš⟩, ⟨iš⟩, and ⟨uš⟩, hereafter the Š-series, were used for rendering the Hittite sibilant fricatives in the Anatolian cuneiform script, ultimately borrowed from Mesopotamian cuneiform. Accordingly, they are recorded as š, šš in the so-called bound transcription of Hittite cuneiform texts, which in turn prompted their conventional pronunciation as /ʃ/ in certain Hittitological circles. Yet, there are several arguments that plead for their alveolar character. First, the Hittite proper nouns in Egyptian transmission render them as ⟨s⟩ or ⟨ś⟩, with few exceptions, which all concern the tense fricative (Breyer 2010: 351, cf. Patri 2009: 95). Second, there is a consensus that the Hittite sibilants go back to the Proto-Indo-European phoneme */s/, whose reflexes are still rendered as /s/ in most Indo-European languages in use nowadays (Kimball 2017: 254). Third, if a language has just one place of articulation for coronal fricatives, it is most likely dental or alveolar (Maddieson 1984: 44). This is why a near-consensus about the pronunciation of the Hittite fricative sibilants as /s/ and /ss/ emerged among linguists in the recent years (Patri 2019: 218–221; Yakubovich 2020: 226; Rieken 2021: 110; Hoffner and Melchert 2024: 60). The only other Hittite sibilant phoneme is the affricate /ts/, which was rendered with the help of the Z-series of syllabograms, including ⟨za⟩, ⟨zi⟩, ⟨zu⟩, ⟨az⟩, ⟨iz⟩, and ⟨uz⟩ (Hoffner and Melchert 2024: 57–59).⁵

² The proposed definition of the sibilants admittedly represents one of several possible options. Some scholars assume that the sibilants constitute a subgroup of the fricative but not affricate sounds, while others exclude the interdental sounds from their definition. The way the sibilants are defined in the present paper is congruent with the approach of Schneider 2024, the latest monograph devoted to the interpretation and reconstruction of these phonemes in Near Eastern languages.

³ Even though this article is devoted to phonetic interpretation, all the segments addressed here typically function as contrastive phonemes. Therefore, we take license to use the markers of phonological transcription (/.../) throughout the article, typically reserving the square brackets for the philological notation of the restored parts of the text. The main exception are direct citations, where we retain the original formalism of the authors cited.

⁴ An alternative interpretation of Hittite phonology assumes just one coronal fricative, which was frequently geminated.

⁵ For the purposes of this paper, we ignore the additional Š-series, which was underdeveloped in syllabic cuneiform (e.g. ⟨ša⟩ = ⟨za⟩) and appears to be more relevant for the study of vocalic than consonantal system, at least in Anatolia (⟨ši⟩ = ⟨zé⟩).

Nevertheless, the selection of the cuneiform Š-series for rendering the Hittite alveolar fricatives requires explanation. Mesopotamian cuneiform features a systematic contrast among the syllabic signs of the Š-series, Z-series, and S-series. The signs of the latter group, whose use in Hittite texts is limited to heterograms, were traditionally interpreted as rendering combinations with the voiceless alveolar fricative. Yet, the readings of the signs under discussion must have varied in dialects of Akkadian and followed the sound changes affecting the respective sibilants, which contrasts with the stability of the opposition between /s/ and /ʃ/ in the history of Ancient Egyptian. More clarity in this issue has been reached within the framework of the Affricate Hypothesis, which slowly but surely acquires mainstream status in the field of Comparative Semitics (Kogan 2011: 61–71; Huehnergard 2019: 50; Schneider 2024: 51–52). While the full presentation of the Affricate Hypothesis goes beyond the scope of this paper, one of its results is the reinterpretation of Proto-Semitic *s and *š as */ts/ and */s/ respectively, and the same conclusion is true for the early stages of Akkadian, including Sargonic Akkadian and Old Assyrian (Hasselbach 2005: 96, 135; Kouwenberg 2017: 48).⁶ No phoneme /ʃ/ is reconstructed for either of these dialects. The interrelated sound shifts */ts/ > /s/ and */s/ > /ʃ/ were postulated for the Babylonian branch of Akkadian (Hasselbach-Andee 2021: 142) and can be demonstrated particularly well for its Late Babylonian stage (Kassian and Popova 2023).

The initial decipherment of Akkadian was accomplished with the help of the Bisitun trilingual inscription of the 6th century BCE, which features a Late Babylonian version. It is therefore historically logical that the signs hidden behind the S and Š-series of the Mesopotamian cuneiform syllabary were initially interpreted as the alveolar and postalveolar fricatives respectively, as befits their Late Babylonian pronunciation. It is also understandable that the same formal convention was secondarily extended to the Anatolian cuneiform syllabary after the discovery of Hittite cuneiform texts. Yet, a mechanical expansion of cuneiform conventions across time and space frequently obfuscates linguistic reality. According to the current consensus, the rulers of Hattusa have availed themselves of the early-second-millennium scribal tradition of north-western Syria for their cuneiform literacy. While the sound system of the relevant Akkadian dialect(s) has not been systematically investigated, the hypothesis that it resembled Old Akkadian and Old Assyrian in being not affected by the shifts */ts/ > /s/ and */s/ > /ʃ/ and therefore had no /ʃ/ in its phonological inventory would not be *a priori* surprising.⁷ This scenario could explain the use of the Z-series and Š-series for phonetic sequences containing /ts/ and /s/ in the relevant cuneiform texts from Syria, which would in turn motivate the retention of the same values in Hittite Anatolia.⁸

The results and hypotheses outlined above suggest that there is still room for refining the phonetic interpretation of the Ancient Near Eastern languages. With this consideration in mind, we shall approach the sibilants of Hurrian and Ugaritic, the neighbour languages of Egyptian, Hittite, and Akkadian in the second millennium BCE, whose phonetic interpretations continue to raise controversy and are mutually interdependent. In Sections 2 and 3 we are going to survey the research history pertaining to the analysis of the relevant phonemes. We shall see that the solutions anticipating some of the results of this paper had been ad-

⁶ Here and below, the representation of Proto-Semitic sibilants follows the middle version of the Affricate Hypothesis, the definition of which is provided in Kogan 2011: 61.

⁷ It is appropriate to observe at this point that the preservation of Proto-Semitic */s/ is postulated not only for the Old Assyrian but also for the Neo-Assyrian (Schneider 2024: 126), which would imply that its most likely pronunciation in Middle Assyrian was also /s/. For the preservation of Proto-Semitic */s/ throughout the history of Assyrian, see Kogan 2011: 88–89.

⁸ Such a formulation obviously obviates the question of the native language(s) of Syrian cuneiform scribes: this issue will be addressed in Section 6 of the present paper.

vanced at the dawn of Hurrian and Ugaritic studies but were then forgotten without being refuted. Sections 4 and 5 will be devoted to the presentation of our own views regarding the pronunciation of the Hurrian and Ugaritic sibilants, respectively. The concluding Section 6 strives to demonstrate that the revised interpretation of the Hurrian sibilants simplifies the account of how Mesopotamian cuneiform was adapted for writing Hittite.

2. Previous research on Hurrian sibilants

The Hurrian language was widespread in northern Mesopotamia, Syria, and probably south-eastern Anatolia in the second millennium BCE. It does not belong to either the Indo-European or the Semitic family, while its only attested close relative is the Urartian language, which was used in writing around Lake Van in the early first millennium BCE and mostly preserved through monumental cuneiform inscriptions. In the 16th century BCE, the Hurrian tribes united by a royal dynasty of Indo-Aryan origin founded the Kingdom of Mitanni, which acted as one of the major Near Eastern powers until its defeat by Hattusa in the mid-14th century BCE. The Hurrian texts are attested in both syllabic and alphabetic cuneiform systems; the latter developed in Ugarit in the 13th century BCE. A Hurrian composition featuring a unique and particularly elaborate syllabic orthography is the so-called Mitanni Letter, sent by Tushratta, king of Mitanni, to the Egyptian court and found in the diplomatic archive of El Amarna (Giorgieri 2000: 177–188, 181–182). The majority of the Hurrian syllabic texts available to us come from the archives of Boğazköy-Hattusa and are characterized by less rigid orthographic standards (Giorgieri 2000: 177, 181–182).⁹

Judging by the Mitanni orthography, the Hurrian phonetic system distinguished among at least three types of sibilants, corresponding to the Š, S, and Z-series, respectively (Diakonoff 1971: 48, Giorgieri 2000: 186, fn. 44). In contrast, the Hurrian texts from Boğazköy reflect only the contrast between the Š and Z-series, which is consistent with the general absence of a phonologically contrastive S-series in Anatolian cuneiform. For reasons stated in Section 4, it is likely that the Z-series were generalized at the expense of the S-series in Boğazköy orthography. In the intervocalic position, there is a contrast between the tense and lax fricatives of the Š and Z-series, which is reflected in the spellings *aš-ša* vs. *a-ša*, *az-za* vs. *a-za* etc. (Wegner 2007: 45–46). No such contrast is observed in the case of the S-series (see Table 3).

The phonemic character of the opposition tense vs. lax in Hurrian derives support from Ugaritic orthography, where the tense and lax sibilants marked by the signs of the Š series in syllabic cuneiform are recorded as ⟨t⟩ and ⟨d⟩, respectively. For example, the name of the Hurrian Storm-god can be spelled *te-eš-šu-up* in syllabic cuneiform and *tṭb* in Ugaritic script, while the name of the Hurrian Moon-god can be spelled *ku-šu-uḥ* in syllabic cuneiform and *kḏ(ḡ)* in Ugaritic orthography (Giorgieri 2000: 184). In the inherited Ugaritic lexicon, the signs ⟨t⟩ and ⟨d⟩ are deployed for the reflexes of Proto-Semitic interdental fricatives */θ/ and */ð/ respectively. In addition, there is limited evidence that the Ugaritic graphemes ⟨s⟩ and ⟨z⟩ were used for rendering the Hurrian sibilants that were expressed by the S and Z-series in Mitanni orthography (see section 5). The distribution between ⟨s⟩ and ⟨z⟩ was, however, not identical to that of the S and Z-series: presumably it had again to do with a tense/lax distinction. In most positions, the phonemic opposition between the tense and lax consonants is neutralized: the

⁹ Hattusa and its vicinity, or, for that matter, central Anatolia in general, never represented areas of compact settlement of Hurrian native speakers. For the sociolinguistic reasons that prompted the proliferation of Hurrian literacy at the court of Hattusa, see Yakubovich 2022a: 24–28.

tense consonants occur word-initially and in clusters with obstruents, while the lax consonants occur in word-final position and in clusters with sonorants (Giorgieri 2000: 184).¹⁰

Thus, one must postulate at least five sibilants for Hurrian: two corresponding to the Š-series, two corresponding to the Z-series, and one corresponding to the S-series (Wegner 2007: 46). There is, however, no agreement about the phonetic values of the phonemes under discussion. According to Giorgieri 2000, arguably the most authoritative introduction to the Hurrian language available to date, the tense and lax sounds corresponding to the Š-series had palato-alveolar fricative articulation, similar to the pronunciation of *sci* in Italian *sciocco* ‘foolish’. In the notation of the International Phonetic Alphabet, this would suggest their transcription as /ʃ/ and /ʒ/, respectively. The Hurrian sounds corresponding to the Z-series represented dental (=alveolar?) fricatives /s/ and /z/, while the S-series in the Mitanni orthography contained sequences with the affricate *c* ~ /ts/ (Giorgieri 2000: 185–186 with fn. 43, 44). The suggestions outlined above are generally based on the earlier proposal by Wilhelm (1988: 64, fn. 34). The interpretation of the Š-series by both scholars appears to derive inspiration from its traditional value in Akkadian cuneiform syllabaries.

The ideas of Wilhelm and Giorgieri underly the conventions of the analytical transcription that has been adopted by many scholars in the field of Hurrian studies. The following forms from the Mitanni Letter, cited after Giorgieri 2000: 268–273, would suffice to illustrate this convention: iii 90 *še-e-ni-íw-wu-tan* ~ *šēn(a)=iffu=dan* ‘from my brother’, iii 67 *ir-nu-u-ḫu-ši-a-a-ma* ~ *irn=ōg=ož=i=ā=ma* ‘he did not reciprocate’, iii 97 *za-la-am-ši* ~ *salamži* ‘statue’, iii 105 ^{KUR}*M[i]-zi-ir-re-e-we-né-e-wa* ~ *Mizir(i)=rē=ve=nē=va* ‘to Egyptian’, iii 72 *su-bi-a-a-maš-ti-en* ~ *cubi=am=ašt=i=en* ‘may he glorify’, iii 85 *ḫi-su-ú-ḫi-wa-a-en* ~ *ḫic=ūg=i=vā=en* ‘may he not upset’. The same phonetic treatment of Hurrian sibilants is retained in the most recent survey of the Hurrian grammar available to date (Hazenbos 2021: 168). At the same time, neither Wilhelm nor Giorgieri have ever regarded the proposed interpretation as cast in stone. Thus, the insistence of its conventional character is found in Giorgieri 2000: 186, fn. 43, while Wilhelm’s cautious stance toward his own hypothesis is clear from the following statement: “The affricate /ts/ (transcribed as *c*) is uncertain... The phonetic realization of /S/ and /Š/ is unknown” (Wilhelm 2004: 99).

A different interpretation of the Š-series in Hurrian was entertained in Bush 1964: 61–71, where the forms in Ugaritic alphabetic transmission were actively used for the study of Hurrian phonology. The signs of the Š-series are interpreted in this dissertation as standing for the phonetic sequences containing /θ/. Later on, the same view was upheld in (Diakonoff 1971: 46). In addition, Diakonoff (1971: 48) tentatively suggested that both the S and Z-series render combinations with voiceless and voiced affricate sounds in Hurrian respectively (the distinc-

¹⁰ Several additional orthographic conventions applicable to Hurrian deserve to be mentioned at this point. The earliest Hurrian composition published to date is known as the Tish-Atal inscription (see Wilhelm 1998 for the edition). The orthography of this text does not distinguish between the tense and lax consonants and probably shows confusion between different series of sibilants, a trait that is also typical of other syllabic texts from the UR III period. More interesting for our purposes is the rendering of Hurrian personal names in Old Babylonian orthography (cf. also Section 6). A number of Akkadian texts from Mari and Northern Mesopotamia use the Š and Z-series for the matching tense and lax sibilants of Hurrian, both of which were rendered with the Š-series in Mitanni and Boğazköy orthographies (Richter 1998: 131–133). This is obviously a point of convergence with the rendering of the same Hurrian sounds as ⟨t⟩ and ⟨d⟩ in Ugaritic script. A large number of Hurrian personal names and foreign words is attested in the Akkadian texts from Nuzi, while the distribution of the sibilant series in the Hurrian forms from Nuzi shows correlations with other Hurrian orthographies (see already Speiser 1938: 189–192). The systematic study of these data is contingent on the success of the project “Digital glossary of Nuzi Hurrian” (<https://cordis.europa.eu/project/id/101204566>), which starts on September 1, 2026.

tion between the tense and lax sibilants is not addressed in this book). A more agnostic stance on the value of the Š-series is found in Khachikian 1985: 38: the sibilant hidden behind the relevant signs is merely assumed to be phonetically distinct from /ʃ/, and probably also from /s/, again based on the Ugaritic evidence. Khachikian (1985: 38–39) concurs with Diakonoff on the affricate pronunciation of the sibilants marked by the S and Z-series in the Mitanni orthography. Both Diakonoff and Khachikian were among the early adepts of the Affricate Hypothesis for Akkadian and use the Akkadian loanwords in Hurrian in order to justify the existence of the Hurrian affricates (cf. Section 4). A compromise position is reflected in Campbell 2020: 209: the reconstructed pair of fricatives š(š), ž represents a point of convergence with the views of Wilhelm and Giorgieri, while the assumption that the remaining sibilants consist of a pair of affricates c, cc comes closer to the ideas of Diakonoff and Khachikian. The interpretation of the Hurrian Š-series as reflecting combinations with /θ/ obtained a recent endorsement in Schneider 2024: 111.¹¹

It is easy to see from the discussion above that the phonetic interpretation of Hurrian sibilants has not yet been settled. The table below summarizes the present state of our knowledge about the representation of sibilants in the most important Hurrian orthographies and the degree of disagreement on their interpretation in recent scholarly literature.¹²

Mitanni orthography	Boğazköy orthography	Ugaritic orthography	Interpretation by Wilhelm	Interpretation by Diakonoff
aš-ša, a-ša	aš-ša, a-ša	⟨t⟩, ⟨d⟩	ʃ, ʒ	θ
az-za, a-za	az-za, a-za (?)	⟨s⟩(?), ⟨z⟩(?)	s, z	dz
a-sa	a-za (?)	⟨s⟩(?), ⟨z⟩	ts	ts

Table 1: The sibilant phonemes of the Hurrian language and their interpretation

At the same time, the divide between scholars who advocate the interpretation of the Hurrian Š-series as sequences with /ʃ/ and /θ/ does not date back to the beginning of Hurrian Studies. According to the observation of Speiser 1938: 175, there was a “long-standing” view in the field that “in the Hurrian syllabary š represent s”. Such a view is advocated, for example in Ungnad 1924: 134, fn. 2, where Hurrian syllabaries are counted among the “western” cuneiform syllabaries, where the fricative series must reflect /s/ and not /ʃ/ based on the treatment of Ancient Egyptian loanwords. The turning point in this discussion appears to have been the decipherment of the Ugaritic script, which provided arguments for the existence of interdental fricatives in Hurrian. The apparent contradiction between cuneiform and Ugaritic orthographies was perceived as a puzzle that had not found a satisfactory solution. The first grammar of Hurrian used an abstract transcription $\bar{s}$ and $\bar{z}$ for the problematic sibilants. The author made an inconclusive comment on their pronunciation: “it is best to regard $\bar{s}$ as a sound patterned between [t] and [š] and $\bar{z}$ as the corresponding voiced sound. It is highly suggestive in this connection that Semitic [t] became [š] in Canaanite just as in Akkadian. A phonetic explanation for this type of change requires an intermediate [s], for which there is no evidence whatever in contemporary Semitic” (Speiser 1941: 35). We shall argue in this paper that Ephraim Speiser was just one step away from solving the issue.

¹¹ Cf. also the tentative interpretation of /θ/ as the dominant value of the Š-series in Northern Mesopotamia and Syria in the second millennium BCE, which is advanced in Weeden 2019: 138.

¹² Some of the conclusions in Table 1 anticipate the new descriptive evidence summarized in Sections 4 and 5. The selection of Igor Diakonoff and Gernot Wilhelm as scholars whose contrastive stances are reflected in Table 1 reflects the outstanding impact of both researchers on the linguistic study of Hurrian in the recent past.

The only explicit attempt to argue for the alveolar sibilant hidden behind the Hurrian Š-series that we are aware of is found in footnote 26 embedded in Tropper and Vita 2003. More specifically, the co-authors observe with regard to the traditional Hurrian transcription: “Die Verwendung der š-Symbole ist dabei konventionell und wahrscheinlich nicht korrekt. In Wirklichkeit dürften syllabische Š-Zeichen ebenso wie alphabetisches *ṭ* ... für hurritisches (stimmloses) [s] stehen”. It is remarkable that one of the authors of the paper discussion under is also a staunch supporter of the phonetic realization of Ugaritic ⟨ṭ⟩ as an interdental fricative (cf. the following section). The discrepancy between the implied realizations of the same grapheme in Hurrian and Ugaritic texts constitutes a possible reason why this suggestion has not been widely followed thus far. The joint treatment of the Hurrian and Ugaritic data in our paper is conceived as a step toward removing this obstacle.

3. Previous research on Ugaritic sibilants

The town of Ugarit, whose ruins constitute the archaeological site of Ras Shamra, was situated on the Mediterranean coast of Syria. The Ugaritic language belongs to the Northwest-Semitic group, and it was claimed on the grounds of lexical comparison that it is a Canaanite language, particularly closely related to Phoenician (Kogan 2015: 343–349). A unique feature of Ugaritic civilization is the use of alphabetic cuneiform, an innovative and visually distinct script that must have taken inspiration from one of the early Semitic alphabets. The Ugaritic alphabetic texts date back to the 13th and early 12th centuries BCE, when Ugarit constituted a client state of the Empire of Hattusa. Before the mid-14th century BCE, Ugarit maintained close ties with the Kingdom of Mitanni, even though its actual integration within the Hurrian state remains a matter of controversy (Singer 1999: 619–620 with ref.). What is beyond doubt is the profound influence of Hurrian culture on Ugarit, which manifests itself, among other things, through a high proportion of Hurrian personal names in Ugaritic onomastics (van Soldt 2003), the presence of Hurrian insertions in Ugaritic texts (Andrason and Vita 2016: 302–308) and the ancillary use of the Ugaritic cuneiform alphabet for writing Hurrian. Many local personal names are attested in both alphabetic and syllabic transliterations (Huehnergard 2012: 8–9, Watson 2016: 353–354). The sociolinguistic facts adduced in this paragraph correlate with the presence of both Hittite-Luwian and Hurrian loanwords in Ugaritic, although the latter are more numerous.

The sibilants of the Ugaritic language are conventionally transliterated as ⟨ṭ⟩, ⟨ḍ⟩, ⟨ẓ⟩, ⟨ṣ⟩, ⟨ẓ⟩, ⟨ṣ̌⟩, and ⟨ṣ̣̌⟩. It is commonly assumed that the triplets ⟨ṣ⟩, ⟨ẓ⟩, ⟨ṣ̌⟩ and ⟨ṭ⟩, ⟨ḍ⟩, ⟨ẓ⟩ comprise consonants that share the place of articulation and primarily differ through their laryngeal features. Each triplet consists of the voiceless, voiced and “emphatic” consonants, the last feature being usually understood as pharyngealization or ejective articulation. The traditional phonetic values ⟨ṣ⟩ = /s/, ⟨ẓ⟩ = /z/, and ⟨ṣ̌⟩ = /ṣ/ were presumably extrapolated from the pronunciation of their correspondences in other Central Semitic languages. The phonemes transliterated as ⟨ṭ⟩, ⟨ḍ⟩, and ⟨ẓ⟩ lack exclusive counterparts in Hebrew, where the system of sibilants has been simplified. Yet, they find historical correspondences in Standard Arabic, where they are pronounced as interdental fricatives. This brought into being the traditional phonetic values ⟨ṭ⟩ = /θ/, ⟨ḍ⟩ = /ð/, ⟨ẓ⟩ = /θ̣/. The pronunciation of ⟨ṣ̣̌⟩ as /ʃ/ is based on its typical Hebrew correspondence. The traditional phonetic interpretation of the Ugaritic sibilant graphemes is summarized, for example, in Segert 1984: 21, 30, while the Semitic historical correspondences that prompted such conclusions can be found in Segert 1984: 34.¹³ The Proto-Semitic lateral conso-

¹³ The discussion of the conventional phonetic values of Ugaritic sibilants can also be found in Huehnergard 2012: 23, where it is stated that their actual pronunciations are unknown.

nants were not preserved in Ugaritic, having merged with ⟨š⟩ and ⟨ś⟩. The historical correspondences between Ugaritic and its Northwest Semitic neighbours, Hebrew and Syriac, are offered in Table 2. The Proto-Semitic reconstruction is provided in two variants: the traditional one, found in the older Ugaritic handbooks, and its counterpart reflecting the middle version of the Affricate Hypothesis.

Proto-Semitic (traditional)	Proto-Semitic (Affr. Hyp.)	Ugaritic (alphabetic)	Biblical Hebrew	Syriac Aramaic
*s (s ₃)	*ts	⟨s⟩	⟨s⟩ (samek)	⟨s⟩ (semkat)
*z	*dz	⟨z⟩	⟨z⟩ (zayin)	⟨z⟩ (zayn)
*s ^ʕ	*ts ^ʕ	⟨ś⟩	⟨ś⟩ (tsade)	⟨ś⟩ (šāde)
*θ	*θ	⟨t⟩	⟨t⟩ (šīn)	⟨t⟩ (taw)
*ð	*ð	⟨d⟩	⟨z⟩ (zayin)	⟨d⟩ (dālād)
*θ ^ʕ	*θ ^ʕ	⟨z⟩	⟨ś⟩ (tsade)	⟨t⟩ (tet)
*ʃ (s ₁)	*s	⟨š⟩	⟨š⟩ (šīn)	⟨š⟩ (šīn)
*ʃ ^ʕ (s ₂)	*ʃ	⟨š⟩	⟨t⟩ (šīn)	⟨s⟩ (semkat)
*ʃ ^ʕ	*ʃ ^ʕ	⟨ś⟩	⟨ś⟩ (tsade)	⟨ś⟩ (šāde)

Table 2. Historical correspondences between Ugaritic and its neighbour languages

While the analysis of genetic correspondences between the Semitic sibilant phonemes is obviously important for their phonetic interpretation, a direct projection of phonetic values from randomly chosen related languages has its limitations. This was shown in Tropper 2012, a monograph that displays awareness of the Affricate Hypothesis and strives to draw consequences from the contacts between Ugaritic and the neighbouring languages, primarily Hittite and Akkadian.¹⁴ The important result of this investigation is the reinterpretation of the triplet ⟨s⟩, ⟨z⟩, ⟨ś⟩ as affricates /ts/, /dz/ and /tsʕ/ (Tropper 2012: 102–107). On the one hand, Josef Tropper observes that the inherited lexemes feature these phonemes as reflexes of Proto-Semitic sounds that are now reconstructed as */ts/, */dz/ and */tsʕ/. On the other hand, he points out that Ugar. ⟨s⟩ corresponds to the Hittite affricate /ts/ in the regional *Wanderwörter*, e.g., Ugar. *ḥswn* vs. Hitt. *ḥazzuwani*- ‘(a garden plant)’ or Ugar. *spsg*, *sbsg* vs. Hitt. *zapzagi*- ‘glaze (vel sim.)’.¹⁵ In syllabic cuneiform, Ugar. ⟨s⟩ was rendered with the Z-series or, more rarely, with the S-series (Tropper 2012: 104).

Tropper’s treatment of the remaining sibilants is equally thorough but less revolutionary. Thus, he endorses the postalveolar articulation of ⟨š⟩ in Ugaritic with the observation that the Canaanite borrowings in second-millennium Egyptian show ⟨š⟩, not ⟨s⟩ or ⟨ś⟩, as the Canaanite reflex of the same Proto-Semitic sibilants that yielded ⟨š⟩ in Ugaritic (Tropper 2012: 108). He also affirms the preservation of interdental fricatives in Ugaritic. Although he does not fail to

¹⁴ The first edition of Tropper’s monograph was published in the year 2000 and already contains essentially the same stances on the topics that are relevant to the present paper. The phoneme /θ/ is frequently misprinted as /m/ in Tropper 2012.

¹⁵ At the same time, he mentions a possibility that the Ugaritic phoneme /ts/ underwent deaffrication at least in some positions in the 13th century BCE. His argument for this claim is the innovative grapheme ⟨š̄⟩, which takes over the value /ts/ in some Ugaritic texts and is particularly common in borrowed lexemes (Tropper 2012: 40–50). For example, Ugar. *ḥswn* ‘(a garden plant)’ and *spsg*, *sbsg* ‘glaze’ have variants *ḥš̄wn* and *[š̄]p̄sg*, respectively (DUL: 406, 758). According to Wikander 2022: 439–440 with fn. 37, the result of the deaffrication of /ts/ may have been the laminal fricative /s̄/.

notice that ⟨t⟩ frequently renders alveolar sounds in borrowed words (cf. Section 5), he provides the following explanation for this phenomenon: “Assuming that in Ugaritic the series of sibilants /s-z-š/ was affricated, only /š/ and /t/ were available for the representation of apical voiceless sibilants. Since Ugaritic /š/ was apparently articulated as palatal [ʃ], it was unsuitable for this purpose. Thus, the Ugaritic phoneme /t/ remained the only candidate” (Tropper 2012: 113, translation ours).

Turning to the articulation of Ugar. ⟨d⟩, Tropper (2012: 101) stresses the instances of synchronic variation between this grapheme and the stop ⟨d⟩, as well as those instances where Ugar. ⟨d⟩ represents an etymological interdental fricative. For example, *ḏrʕ* ‘seed, seed-grain’ alternates with *drʕ* (see also DUL: 278), while *dbḥ* ‘sacrifice’ cannot be separated from *zbḥ*, attested with the same meaning in Hebrew, Phoenician, and Punic, as well as Arabic *dabḥ* ‘sacrifice’ (DUL: 259). Furthermore, as argued in Tropper 2012:113, there are instances of synchronic variation between the graphemes ⟨z⟩ and ⟨t⟩ in Ugaritic texts as well as lexemes where ⟨z⟩ goes back to the Proto-Semitic “emphatic” stop. Thus, Ugar. *ṯhrm* ‘pure (pl.)’ has the variant *zhrm* (see also DUL: 875), while Ugar. *ḥlmz*, if indeed meaning ‘snake’ or ‘dragon’, is cognate with Akkadian *ḥulmiṯtu* ‘dragon vel sim.’ and Syriac *ḥūlmaṯā* ‘large lizard, chameleon’ (see also DUL: 387). The facts mentioned in this paragraph support the preservation of the interdental articulation of ⟨d⟩ and ⟨z⟩. No parallel alternations are cited, however, in the instance of /t/, if we exclude the Ugaritic Short Alphabet, where the number of signs is generally reduced.

Josef Tropper’s views on the Ugaritic affricates found a broad following in modern scholarship, see, e.g., Kogan 2009: 208, Hawley 2020: 264, Schneider 2024: 128. At the same time, some scholars went further in the application of the Affricate Hypothesis and assumed that not only the deaffrication */ts/ > /s/ but also the shift */s/ > /ʃ/ failed to be implemented in Ugaritic. This claim is advanced in Hutton 2006: 78–81 in connection with the derivation of *mḥšt* ‘I smote’ from the root *mḥš* ‘to smite’. It is indeed logical to account for the morphophonemic alternation between ⟨š⟩ and ⟨ṣ⟩ in this case by assuming the conditioned sound change */ts/ > /s/ before the stops. At the same time, there is no compelling reason to believe that this sound change happened during the attested stage of Ugaritic, as opposed to occurring in its prehistory and feeding the subsequent shift */s/ > /ʃ/. Nevertheless, Hutton’s suggestion was followed in Kogan 2009: 208 and Hawley 2020: 264, perhaps because these scholars were reluctant to reconstruct a rich system of Ugaritic sibilants without a phoneme /s/ for typological reasons.

At this point, it behoves us to turn to another suggestion regarding the graphic rendering of /s/ in Ugaritic, which predate the rise of the Affricate Hypothesis. As stated in Cross 1962: 250: “It has long been recognized that etymological *t* at Ugarit [...] actually has already shifted in pronunciation, probably to something like /s/ by the time of our texts [...] The change from the sound *t* to *s* or the like can be shown by transcriptional materials, Hittite, Egyptian and Ugaritic”. The paper under discussion fails to provide references to earlier scholars that shared the same view or to explicate what constitutes the Ugaritic evidence. But the Egyptian and Hittite evidence are illustrated in Cross 1962: 250, fn. 115. On the one hand, the name of Suppiluliuma, king of Hattusa, whose initial syllable is rendered with ⟨šū⟩ in Hittite texts, is transliterated in Ugaritic as *ṯpllm* (DUL: 911–912). On the other hand, the island of Alasiya/Cyprus, recorded as *ʾa₂-ra-sa* in Egyptian group writing, is transliterated in Ugaritic as *ʾlty* (see also DUL: 65). These observations paraphrase parts of the argumentation that was advanced in Friedrich 1943: 7–9 in order to demonstrate that the phonetic realisation of Ugaritic ⟨t⟩ was /s/ or /ʃ/, but not /θ/.¹⁶ The hypothesis that Ugaritic ⟨t⟩ was phonetically close to the pronunciation

¹⁶ For the earlier papers, which advocated the same view in the wake of the decipherment of Ugaritic, see Friedrich 1943: 7, fn. 2. Naturally, the transcription signs used in these early publications differ from the modern phonological conventions adopted in the present paper.

of /s/ in the languages of Asia Minor was later endorsed in Degen 1967: 50 on the basis of similar contact-linguistic evidence.¹⁷

Yet, for whatever reason, the claims by Friedrich, Cross, and Degen made little impact on the subsequent development of Semitic historical phonology. It is remarkable that even the scholars addressing contacts between the Semitic languages and their neighbours frequently failed to cite these claims. This may be primarily due to the understandable background limitations of individual researchers. Thus, Hoch (1994: 403) assumes that the evidence for the alveolar pronunciation of Ugaritic ⟨t⟩ is based primarily on Levantine forms in Egyptian transmission, leaving out of consideration the Anatolian part of the argument. Conversely, no mention of Semitic borrowings in Egyptian is made in the discussion of rendering Hittite /s/ with Ugaritic ⟨t⟩ in Yakubovich 2010: 152b or Hoffner and Melchert 2024: 60, where the interpretation of Ugaritic data is entirely based on Tropper's views. At the same time, scholars focusing on the Semitic languages sometimes have a limited command of contact-linguistic issues involving their non-Semitic neighbours. Indicative of this state of affairs is the assertion: "The renderings of Hittite sibilants in Ugaritic texts have not yet been thoroughly studied" (Schneider 2024: 111, contrast, e.g., Patri 2009).

Given this history of research fragmentation, it is perhaps unfair to expect that impulses from Hurritology could affect the interpretation of Ugaritic sibilants to any significant extent. For the majority of Semiticists, drawing upon the Hurrian data in an attempt to elucidate Semitic historical phonology would probably represent a paragon example of *clarum per obscurius*. Objectively speaking, however, the intensity of contacts between the Hurrian and Ugaritic languages allots Hurrian a special place in the phonetic interpretation of Ugaritic, provided that the values of Hurrian phonemes can be assessed on independent grounds.

4. Phonetic interpretation of Hurrian sibilants.

The goal of this section is integrating various kinds of the available evidence in order to propose the most plausible phonetic values for the Hurrian sibilant phonemes. An exception is made for the Hurrian forms in Ugaritic transmission, whose treatment is relegated to the next section in order to avoid circular reasoning.

It seems appropriate to begin with the discussion of the Hurrian phoneme hidden behind the S-series, whose affricate interpretation represents one point on which modern scholars seem to agree. As mentioned in Section 2, its presence is limited to the Mitanni orthography. While the list of Mitanni Hurrian forms featuring the signs ⟨sa⟩, ⟨si⟩, or ⟨su⟩ may be extended with the definitive edition of the Mitanni Letter, at present we are aware of the following attestations:¹⁸

¹⁷ The most recent attempt to defend the phonetic realization of Ugaritic ⟨t⟩ as "something like /s/" is found in the papers by Ola Wikander (see in particular Wikander 2015: 388). One finds there a brief critical discussion of Tropper's views on the phonetic realization of ⟨t⟩, which resonates with the argumentation advanced in the hand-out "Phonetic Interpretation of Hurrian Sibilants in the Light of Indo-European Evidence" by Ilya Yakubovich (2009, see further fn.1). Without doubting the originality of Wikander's ideas, it is appropriate to treat his contribution as part of the intellectual dialogue that facilitated the gestation of the present paper.

¹⁸ Here and below, we continue to use the traditional analytical transcription of Hurrian forms, where the convention of rendering sibilants is based on the views of Wilhelm and Giorgieri (see Section 2). We differ, however, from the earlier practice of marking all the morpheme boundaries as ⟨=⟩ in the analytical transcription. In our transcription, the signs ⟨+⟩ and ⟨-⟩ mark the attachment of the derivational and inflectional morphemes, respectively, while ⟨=⟩ is reserved for clitic boundaries.

Transliteration	Conventional Transcription	Translation	Attestation
ap-su-u-ša	apc-ōž-a	‘he X-ed’	iv 63
ap-su-ša-a-ul-la	apc-ož-aw=lla	‘I X-ed them’	iv 63
¹ A-sa-a-li-	Acali	(PN)	iv 36, 37
ḫi-su-ú-ḫi-wa-‘a’-en	ḫic+ūğ-i-vā-en	‘may he not upset’	iii 85
ḫi-su-ú-ḫi-wa-a-en			iii 89
ḫi-su-ú-ḫ[i-wa]-a-en			III 95
ḫi-su-ú-hul-li-e-et-ta-	ḫic+ūğ-o(l)l-ē=tta	‘I would be upset’	iv 10
ḫi-su-ú-hu-lu-ú-en	ḫic+ūğ-ol-o-v(a)-en	‘may he not be upset’	i 110
ḫi-su-ú-hu-uš-x[	ḫic+ūğ-uš[tʰ...]	(unclear)	ii 52
ḫi-su-ú-hu-ši-úw-wu	ḫic+ūğ-ož-i-uffu	‘I did not upset’	iv 33
bi-su-un-ni-	pic+o+n(i)-ne-n	‘in a joyful way’	i 79
bi-su-u-u-ni-i-			iii 4
bi-su-uš-te-e-wa	pic-ošt(-i)-ēva	‘(I) would rejoice’	ii 55
bi-sa-an-ti-iš-ti(n)-	pic+and-išt-i	‘(one)? will rejoice’	iv 44
bi-sa-an-du-ši-i-it-ta	pic+and-ož-i-tta	‘I rejoiced’	iv 9
bi-su-uš-ta	pic-ošt-a	‘(I) rejoice’	ii 62
b[i-s]u-uš-ta-iš	pic-ost(-i)-ai-ž	‘that (we) may rejoice’	i 80
su-bi-a-a-maš-ti-en	cubi+ām-ašt-i-en	‘may he glorify’	iii 72
su-bi-a-a-maš-ti-e-‘ni’-			iii 88

Table 3. S-series in the Mitanni Letter

One can infer from the above examples that the attestation of the S-series is limited to one personal name and four verbal roots, *apc*- ‘?’, *ḫic*- ‘to upset’, *pic*- ‘to rejoice’, and *cubi*- ‘to glorify’. There are no instances of consonant replication involving the signs of this series. If we exclude the factor of chance (which is not so easy to do given the small number of the roots involved) this distribution speaks for the S-series corresponding to only one sibilant phoneme. The main etymological argument for its affricate realization is the comparison between Hurrian *pis*- ‘to rejoice’ and Urartian *pišuše* ‘joy’ (Diakonoff 1971: 76; Harouthisounian 2001: 458a). The affricate realization of the “emphatic” sibilant in Urartian finds direct support in the Urartian toponym ^{KUR}*Šu-pa-ni* corresponding to Armenian Տոփք/*Coph-kh* ‘Sophrène’ (Diakonoff 1971: 48; Harouthisounian 2001: 522).¹⁹

Luckily, the root *pis*- ‘to rejoice’ is also attested in Akkadian texts from Ugarit in the personal name *bi-zu-ni* and the forms *[b]i-zu-ni-la-an* and *bi-za-an-ti*- (van Soldt 1991: 357, fn. 224). As mentioned in Section 3, the cuneiform Z-series is otherwise used for sequences with Ugaritic /ts/. The confusion between the S and Z-series in Hurrian lexemes attested in syllabic cuneiform transmission in texts from Ugarit could find explanation in the phonetic similarities of the respective Hurrian sibilants. This brings us to the more general question of the distribution between the S and Z-series in Hurrian texts. The Boğazköy orthography does not use the S-series at all in Hurrian forms, but the Mitanni Letter features both groups of signs with lexi-

¹⁹ It is also probable that the Š-series was used for rendering combinations with affricates in Neo-Assyrian cuneiform, whence the Urartian cuneiform conventions were ultimately borrowed. A structural argument for the affricate realization of the sibilant associated with the Š-series in the Assyrian dialect of Akkadian, up to the Neo-Assyrian Period, is laid down in Kouwenberg 2003, see also Kogan 2011: 65 for evidence from the Neo-Assyrian loanwords in Egyptian.

cal distribution, as can be infer from the comparison between Table 3 and Table 4.²⁰ The principal distinction between the S and Z-series is that the latter one displays an opposition between spellings with and without consonant replication (contrast *ku-zu-u-šu* vs. *še-e-ni-íw-wu-ú-uz-zi*) and thus reflects a cluster of two phonemes, recorded as /s(s)/ and /z/ in the traditional analytical transcription.

Transliteration	Conv. Transcription	Translation	Attestation
an-za-a-an-nu-u-hu-ši	anz+ān+ōġ-ož-i	‘he honoured’	i 18
an-za-a-an-nu-u-[hi-e-n(i)-	anz+ān+ō[ġ-i-ēn]	‘may’ he honour’	iii 87
an-za-an-nu-u-hu-ša-a-ú	anz+ān+ōġ-ož-āw	‘I honoured’	iii 50, 51
ku-zu-u-ši-úw-wu-	koz-ōž-i-uffu	‘I did not detain’	iv 46
ku-zu-u-šu	koz-ōž-o	‘you (sg.) detained’	iv 45
ku-zu ¹ -uš-ti-wa-a-en (written {su}))	koz-ošt-i-vā-en	‘may he not detain’	iv 40
^K [UR]mi-zi-ir-ri-e-we	mizri-nē-ve	‘of Egypt’	i 62
^{KUR} m[i]-zi-ir-ri-e-we-ni-e-we	mizri-nē-ve-nē-va	‘to Egyptian’	iii 105
^{KUR} mi-zi-ir-ri-e-we-ni-eš	mizri-nē-ve-nē-ž	‘of Egypt’	i 85
še-e-ni-íw-wu-ú-uz-zi	šen-iffū-ssi	‘suitable for my brother’	iii 43
za-lam-ši	salm(i)+ži	‘statue’	iii 77, 90, 97, 99, 102
za ¹ -lam-ši (written {ḥa}))			iii 106
za-ar-ra-	sarr-a	‘booty’	i 89
za-a-zu-lu-u-ša[-a-ú]	sāz+ol-ōž-ā[w]	‘[I] nourished’	ii 117
za-a-zu-lu-u-š[a-a-ú]	sāz+ol-ōž-ā[w]	‘[I] nourished’	ii 124
zu-kán	sugan	(particle)	ii 11, iii 67, 68, 72
zu ¹ -kán (written {su}))			iii 16

Table 4. Z-series in the Mitanni Letter

The same lexemes that were written with the Z-series in the Mitanni Letter were also recorded with the Z-series in Boğazköy Hurrian. Thus, the form *za-a-zu-lu-u-ša[-a-ú]* ‘I nourished’ can be compared with KBo 32.15 vs. 27a’ *za-a-zu-li-li-ta* ‘he will be nourishing’,²¹ while *za-lam-ši* ‘statue’ is cognate with KUB 27.1 i 54 ^D*za-al-mi* ‘divine statue’, occurring in a Hittite context. Although the specific morphemes that are recorded with the S-series in the Mitanni Letter do not seem to occur in Boğazköy texts, it is impossible to attribute the lack of the S-series in Boğazköy orthography to a lexical gap. The corpus of Boğazköy Hurrian texts is far larger than that of Mitanni Hurrian, and statistical considerations would militate against an accidental absence of one of the Hurrian affricates in this corpus. One could rather assume that the confusion between the S and Z-series for rendering the Hurrian sounds, typical of Hurro-Akkadian, culminated in the generalization of the Z-series at the expense of the S-series typical

²⁰ Table 4 excludes fragmentary forms and lexemes of unknown meaning. The two instances of {su} written instead of {zu} in the forms collected in this table must be explained through the graphic similarity of both cuneiform signs, and in IV 40 one may even doubt which of the two signs was intended by the scribe.

²¹ The forms with the dissimilation of sibilants are also attested, e.g., KBo 32.19 vs. 6a, 8a *za-a-šu-lu-u-ša* ‘he nourished’. Compare also the alternating spelling *ka-a-šu-uš* and *ka-a-zu-u-š* ‘like a cup’ in the same Hurrian manuscript of the Song of Release KUB 32.19. In the last case we are probably dealing with different strategies of adapting a loanword.

of peripheral Akkadian. Once again, this graphic explanation would only make sense if the sibilants hidden behind the S and Z-series had a certain phonetic feature in common, which was not shared by the Š-series. As mentioned in Section 1, the Z-series is used for sequences with the affricate /ts/ in Hittite orthography, while the affricate pronunciation of the sibilant associated with the S-series was argued on independent grounds for Mitanni Hurrian earlier in this section (cf. also Section 5 for additional evidence from Ugarit). These facts are compatible with the conclusion that the shared phonetic feature of the Hurrian sibilants that were rendered with the help of the S and Z-series consisted in their affricate pronunciation.²²

Yet another argument for the affricate interpretation of the Z-series comes from the Mitanni Letter. The Hurrian stems for ‘statue’ and ‘Egypt’ listed in Table 4 represent transparent loanwords from Akkadian going back to *šalmu*- ‘statue’ and *mišru*- ‘frontier, Egypt’, respectively. The pronunciation of Proto-Semitic “emphatic” *š as */ts/ and its affricate reflexes in ancient Semitic languages constitute the most secure part of the Affricate Hypothesis (Kogan 2011: 62–65). There is no obvious rationale why this “emphatic” sound could be borrowed as /s/ in Hurrian, especially given independent evidence for the Hurrian affricate rendered by the S-series. In contrast, the borrowing of /ts/ as /ts/, with the loss of “emphatic” ejective articulation, would be phonologically natural, while the association of the Z-series with the affricate /ts/ would find a direct counterpart in Hittite orthography.²³ Admittedly, spellings such as *mi-zi-ir-ri-e* suggest the presence of a voiced/lax sibilant in the Hurrian word for ‘Egypt’, but this is no doubt a result of the voicing assimilation */mitsri/ > /midzri/. Such a process fully conforms to the rules of Hurrian phonology (Giorgieri 2000: 184) and must accordingly postdate the borrowing episode.²⁴

A further argument for the affricate articulation associated with the Z-series in Boğazköy Hurrian comes from a morphophonemic change. In Hurrian texts from Hattusa, the Z-series can be used in the morphemes that are otherwise spelled with the Š-series if the sibilant appears immediately after /n/ or /l/. For example, the standard plural marker of the Hurrian nouns and adjectives is -ž-, contrast the directive forms KBo 27.219+ rev. 5' *we-e-šu-ta* ~ *fē-ž-uda* ‘toward you (pl.)’ vs. KUB 47.2:12' *wi-ú-ta* ~ *fē-uda* ‘toward you (sg.)’. The plural marker, however, appears as -z- in the inflected forms of the third person plural pronoun after the nasal, e.g., DAAM 1.14 obv. 11' *ma-an-zu-u-uš* ~ *man-z-ōž* ‘they (ergative)’, KUB 31.3 obv. 6 *ma-an-zu-u-ra* ~ *man-z-ōra* ‘with them’ *ma-an-zu-u-ra* ~ *man-z-ōra* ‘with them’. The same phonological rule

²² The hypothesis outlined above does not go beyond positing a graphic merger between two sibilant series in Anatolia, which is sufficient for the essential claims of this paper. At the same time, if one reconstructs a community of Hurrian scribes whose native language was Hittite or Luwian, for example, in Kizzuwadna, such scribes may have easily lost phonetic distinctions between the affricates corresponding to the S and Z-series under the influence of their vernaculars.

²³ Cf. also the case of *za-ar-ra*- ‘booty’ in Table 4, which is sometimes compared with Classical Armenian *caṛay* ‘slave’. The correspondence between the initial affricates in both languages would be consistent with the interpretation offered in this section, but the suffix of Arm. *caṛay* rather points to the Aramaic origin of this noun (Nielsen 2023: 25–26). Perhaps we are dealing with a loanword from Hurrian or Urartian that passed through several languages on its way to Armenian.

²⁴ Additional lexical items borrowed from Akkadian can enhance our knowledge of the pronunciation of sibilants in those dialects of Akkadian that were in contact with Hurrian. Thus, Hurr. *a-(az)-zi-i-ri* ‘prisoner’, *ḥa-az-zi-zi* ‘wisdom’, *ka-a-(az)-zi* ‘cup, goblet’, and *pa-ri-iz-za-te* ‘a measure of capacity’, all attested in Boğazköy texts, are usually assumed to have been borrowed from Akkadian *asīru*- ‘prisoner of war’, *ḥasīsu*- ‘understanding’, *kāsu*- ‘cup, goblet’, and *parīsu* ‘(a measure of capacity)’ (BGH: 54, 141, 192, 301). If one accepts that the signs of the Z-series rendered combinations with affricates in Boğazköy Hurrian, the simplest solution would be to assume the affricate realization of the matching Akkadian sibilants in the contact zone. Such a conclusion fits well with the interpretation of Proto-Semitic *s as */ts/ within the framework of the Affricate Hypothesis (cf. Section 1).

underlies the allomorphic distribution of the Hurrian abstract suffix *+ze/+že*: contrast KUB 29.8+ iii 18 *it-kal-za-a-e* ~ *itk+al+z(e)-ae* ‘with purification’, KUB 32.15 rev. 2a *ki-ri-en-zi* ~ *kir+en+ze* ‘liberation’ vs. KUB 47.93:4 *ú-pu-kar-še-ni* ~ *ub+ugar+že-ne* ‘from the battle’.²⁵ This rule can be accounted for in terms of phonological naturalness through the assumption of a *t*-epenthesis between the nasal or lateral and the following sibilant driven by articulatory constraints, to be compared with the changes */ns/* > */nts/* and */ls/* > */lts/* in the geographically adjacent Luwian language (Yakubovich 2015, § 5.2.2, cf. Melchert 2020: 249).²⁶ The original place of articulation associated with the Š-series was presumably extended to the affricate resulting from the *t*-epenthesis.

One must now approach the phonetic realization of the Š-series, which is better represented in Hurrian texts than either the S or Z-series. There are two independent arguments that support the alveolar articulation */s(s)/*, */z/*, as opposed to the postalveolar articulation, */ʃ(ʃ)/*, */ʒ/*, for the phonemes under discussion. On the one hand, the few Hurrian names attested in Ancient Egyptian transmission, namely Agi-Tessob, Ini-Tessob, Tili-Tessob, and Tulbi-Sarri consistently use phonograms featuring ⟨ś⟩ in their transliteration for rendering the Hurrian sibilant in the morphemes *teššob* ‘Storm-god’ and *šarri* ‘divine king’ (Schneider 1992, ## 39, 79, 503, 510). As mentioned in Section 1, ⟨s⟩ and ⟨ś⟩, as opposed to ⟨š⟩, were used to transmit the alveolar fricative sounds in the second millennium BCE Egyptian. On the other hand, the Hurrian fricatives find a regular correspondence in the Uartian sibilant that was recorded with the help of the Š-series, e.g. Hurr. *ša-ú-ri* ‘weapon’ vs. Uart. *šú-ri* ‘id.’, Hurr. *ᵐte-eš-šu-up* ‘Storm-god’ vs. Uart. *ᵐte-i-še-e-ba* ‘id.’, Hurr. *tar-šu-wa-an-ni* ‘human’ vs. Uart. *ᵐta-ar-šú-a-ni* ‘id.’ (Harouthiounian 2001: 465b, 467b, 490a). Since Uartian cuneiform was derived from its Neo-Assyrian counterpart, it is reasonable to expect that the Š-series was used for sequences with the Uartian alveolar sibilant */s/*, as was also the case in Neo-Assyrian. Independent evidence for the alveolar articulation of the Uartian fricative hidden behind the Š-series comes from the Uartian borrowings in Classical Armenian, e.g., Uart. *šá-ni* ‘vessel’ vs. Arm. *san* ‘cauldron’, Uart. *ᵐtu-uš-pa* ‘(town name)’ vs. Arm. *Tosp* ‘(town name)’ (Harouthiounian 2001: 462b, 526a).²⁷ Since both Neo-Assyrian and Classical Armenian feature the phonemic distinction between */s/* and */ʃ/*, the proposed phonetic interpretation of the Uartian Š-series appears to be reasonably certain. The converging evidence of names in Egyptian transmission and Uartian cognates lends in turn considerable support to the association between the alveolar fricatives and the Hurrian Š-series. The fact that the fricative sibilants associated with this series occur more frequently than the affricates in Hurrian texts, of course, squares well with the general typological expectations.

²⁵ For the data that support the proposed distribution of the suffix allomorphs, see Giorgieri 2000: 2003, although the formulation of the phonological rule is our own. The same alternation pattern is attested in the collective suffix *+šari/+zari*: contrast, e.g., KBo 32.14 obv. 20a vs. *ti-ip-ša-a-ri* ‘matter(s)’ vs. KBo 32.14 obv. 36a *e-en-za-a-ri* ‘gods’. Sometimes one observes an optional application of the rule under discussion, e.g., KBo 32.214 i 8’ *al-zi-ia-al-li* vs. KUB 47.5+ iv 13 *a-al-ši-ia-al-li* ‘opponent (vel sim.)’.

²⁶ The same change is attested in the history of Hittite, as., e.g., in the oblique pronominal stem *anz-* < **ans-* < **ns-* ‘us’ (Melchert 1994: 97). Unlike Luwian, however, Hittite does not appear to feature this process as a synchronic morphophonemic rule.

²⁷ For the synchronic phonological system of Classical Armenian, rich in fricative and affricate sounds, see Schmitt 2007: 35–36. The scope of lexical contacts between Uartian and Classical Armenian continues to represent a debatable issue: one can contrast the liberal approach of Greppin 2011, a paper that specifically focuses on sibilants, with the more balanced treatment of Nielsen 2023: 2–37 and the skeptical attitude of Simon 2022. Yet, whichever approach one follows, there are enough data for establishing the basic correspondence between the Uartian Š-series and classical Armenian */s/*.

If one interprets the Hurrian fricatives as the tense alveolar fricative /*(s)s*/ and the matching lax alveolar fricative /*z*/, this is conducive to a hypothesis that the pair of tense and lax affricates rendered by the Hurrian Z-series had the same alveolar place of articulation. Indeed, we have seen earlier in this section that the Z-series is used in Akkadian loanwords, where it corresponds to the alveolar affricates interpreted as /*ts*/ and /*ts'*/ in the source language. Both sounds were presumably transmitted as the tense affricate /*ts*/ in Hurrian. A likely source for the matching lax affricate /*dz*/, which must have been additionally cued by voicing, was the epenthesis after /*n*/ or /*l*/, as in *kir+en+ze* ~ /*kirendze*/ ‘liberation’ or *itk+al+ze* ~ /*itkaldze*/ ‘purification’. The interpretation of the affricate in these clusters as lax follows from the general rule, according to which clusters with sonorants contain lax consonants, e.g., *mizri* ~ /*midzri*/ ‘Egypt’ or *ub+ugar+že* ~ /*ubugarze*/ ‘battle’.

The affricate corresponding to the S-series in Mitanni Hurrian must have contrasted with /*ts*/ and /*dz*/ . Pinpointing its precise phonetic realization appears to be impossible at the present state of our knowledge, but one can attempt to point out more and less probable interpretations. An assumption that the affricates differed from all the other Hurrian consonants in featuring additional oppositions expressed by larynx gestures, or reflecting them in writing, would hardly be a parsimonious solution. In addition, an affricate marked by a special laryngeal feature, such as aspiration or glottalization, would probably be written with consonant replication in the intervocalic position to mark the delayed release. The orthography reflected in Table 3 shows the opposite state of affairs: the signs of the S-series appear to be incompatible with consonant replication in Mitanni Hurrian. Against such a background, more credence accrues to a hypothesis that the S-series contrasted with the Z-series through the place of primary articulation of the relevant affricates. The use of the S-series for sequences with the postalveolar fricative /*j*/ in Neo-Assyrian would provide an approximate typological parallel for its potential use in sequences with the postalveolar affricate in Hurrian. In this connection, it is also worth pointing out the presence of postalveolar affricates in Old Indo-Aryan, even though none of the Hurrian lexemes reflected in Table 3 have a transparent Indo-Aryan etymology.

The revised interpretation of the Hurrian sibilants is reflected in Table 5 below, with the caveat that the proposed realization of the S-series merely represents one of several possibilities.

Mitanni orthography	Boğazköy orthography	Ugaritic orthography	Interpretation by Wilhelm	Interpretation by Diakonoff	New suggestion
aš-ša, a-ša	aš-ša, a-ša	⟨t⟩, ⟨d⟩	ʃ, ʒ	θ	s(s), z
az-za, a-za	az-za, a-za (?)	⟨s⟩(?), ⟨z⟩(?)	s, z	dz	ts, dz
a-sa	a-za (?)	⟨s⟩(?), ⟨z⟩	ts	ts	tʃ/dʒ (?) ²⁸

Table 5. The Hurrian sibilants and their revised interpretation

It is, however, to be noted that none of the considerations adduced in this chapter offers decisive arguments against the interpretation of the Hurrian fricatives as /*θ*(*θ*)/ and /*ð*/ . Neither the second millennium Egyptian nor Classical Armenian feature interdental fricatives as phonologically contrastive segments. Therefore, a devil’s advocate could argue that the Hurrian and Urartian interdental fricatives were systematically adapted as alveolar fricatives in these both languages. To be sure, it is possible to object to such an argumentation that the sys-

²⁸ It is tentatively assumed that the Hurrian phoneme corresponding to the S-series had tense and lax allophones. For the justification of this suggestion, see the discussion of the possible Ugaritic reflexes of this phoneme at the beginning of the next section.

tem featuring a pair of fricatives /θ(θ)/ and /ð/, the affricates /ts/ and /dz/, but no phoneme /s/, is typologically unusual. Yet, as far as phonetic interpretation goes, typological observations rarely outweigh direct evidence from the treatment of loanwords. The necessity to reconcile the conclusions made thus far with the extensive use of ⟨t⟩ in Hurrian forms attested in Ugaritic transmission prompts the renewed examination of Ugaritic sibilants.

5. Phonetic interpretation of Ugaritic sibilants

The Hurrian forms in Ugaritic alphabetic transmission are consistent with the affricate realization of the sibilants that are associated with the S-series in Mitanni orthography. The crucial match involves the Hurrian adverb *bi-su-un-ni* ‘in a joyful way’, (see Table 3). This adverb shares the root and probably the suffix with the second component of the compound personal name *lwr-pzn*, to be understood as “(Divine) lord is joy” or “(Divine) lord is joyful” (DUL: 129, Burlingame 2020: 201–202). Furthermore, the personal name *pzny* (DUL: 678) must represent a hypocoristic shortening of a compound personal name with an element *pzn* ‘joy(ful)’. The presence of the Hurrian female personal names ^h*bi-zu-ni*, ^h*bi-zu-ni-e*, and ^h*te-ḫu-ub-bi-zu-ni* in Nuzi onomastics yields further support to the proposed interpretation (Gelb et al. 1943: 246). If one accepts Tropper’s analysis of the Ugaritic ⟨z⟩, this means that the intervocalic sibilant rendered by the cuneiform signs of the S-series in Mitanni Hurrian was adapted as the affricate /dz/ in Ugaritic transmission.

In addition, there are some Hurrian lexemes featuring affricates in Ugaritic transmission, the counterparts of which do not occur in the Mitanni Letter. Thus, Ugaritic *ālz* cannot be separated from the West Hurrian goddess Allanzu, which is attested as both *al-la-zi* and *a-la-an-zu* in syllabic cuneiform transmission in Ugarit (BGH: 13, DUL: 66). The first morpheme of the Hurrian theophoric personal name *sl-ṭmg* (DUL: 751) must be considered together with the onomastic element *zil-*, which occurs with high frequency in Nuzi personal names, including theophoric names (Richter 2016: 506). Beyond these two examples, the matches become more controversial.²⁹ In any event, the affricates involved cannot be specifically equated with either S or Z-series.³⁰ The only thing one can affirm is that they do not contradict the affricate articulation associated with the Z-series, whose primary support comes from cuneiform evidence.³¹

The data become more explicit once we turn to the Ugaritic counterpart of the syllabic Š-series in Hurrian. Table 6 below features theophoric personal names with recurrent elements *ḏn/ṭn* ‘brother’, *ḏr/tr* ‘king’, *ḏrm/ṭrm* ‘Sarruma’, *ḥdm* ‘is shining’, *kḏ(ḡ)/kz* ‘Moon-god’, *šmg/ṭmg* ‘Sun-god’, and *ṭṭb* ‘Storm-god’. No attempt was made to collect all the names with the relevant

²⁹ For a tentative list of Hurrian forms in Ugaritic transmission with parallels in syllabic orthography, see BGH: 554–556, and many of the forms in this list feature ⟨s⟩ or ⟨z⟩. Unfortunately, some relevant lexemes were assigned meanings based on speculative etymologies in previous scholarship, while others may have been borrowed into both Ugaritic and Hurrian from a third source.

³⁰ It is possible to make tentative suggestions regarding the distribution of ⟨s⟩ and ⟨z⟩. In rendering the Hurrian affricates corresponding to the Z-series, it presumably reflected the phonological contrast between the tense and lax consonants, with the same neutralization positions as in the instance of the fricatives, for which see Giorgieri 2000: 184. In rendering the Hurrian affricate corresponding to the S-series, it was presumably automatic, with only ⟨z⟩ occurring intervocalically. This hypothesis admittedly reflects an extrapolation of the available Ugaritic evidence in an attempt at harmonizing it with the cuneiform orthography of the Mitanni Letter.

³¹ Cf. also the spelling of the Anatolian town Lawazzantiya as *lwsnd* and the Luwian personal name *Pirwa-ziti* as *prwsdy* in Ugaritic transmission (DUL: 503, 674). An irregular reflex of a Hurrian affricate is found in Ugar. *ḥndrt* ‘a fruit tree’ vs. Hurr. *ḫinzuri* ‘apple’ (DUL: 393, Kilani 2023: 38).

morphemes, but each of them occurs at least twice in Table 6, to ensure the reliability of the segmentation. All these elements are also attested in cuneiform transmission, where they are recorded with the signs of the Š-series: *šena* ‘brother’, *šarri* ‘king’, *šarruma* ‘Sarruma’, *ḥešmi* ‘is shining’, *kušuh* ‘Moon-god’, *šimiga* ‘Sun-god’, and *teššob* ‘Storm-god’.

Transliteration	Simplified transcription	Translation	Reference to DUL
ib/wrḏr	Evri-sarri	“(Divine) lord is king”	p. 12
āgb/pṭr	Agip-sarri	“(Divine) king raised”	p. 25
āg/kd/tṭb	Agi-Tessob	“Storm-god raised”	p. 25
āgpṭn	Agip-sena	“(Divine) brother raised”	p. 26
āḡlḏrm	Ehli-Sarruma	“Sarruma saved”	p. 29
īḡlkḏ / āḡlkz / ḡlkz	Ehli-Kuzuh	“Moon-god saved”	pp. 29, 314
āḡltṭb	Ehli-Tessob	“Storm-god saved”	p. 29
ā/irpṭr	Arip-sarri	“The (divine) king gave”	p. 101
irrṭrm	Evri-Sarruma	“The lord is Sarruma”	p. 102
aršmg	Ari-Simiga	“Sun-god gave”	p. 105
ārṭṭb	Ari-Tessob	“Storm-god gave”	p. 107
ḥḏmḏr	Hizmi-sarri	“(Divine) king is shining”	p. 383
ḥḏmtṭb	Hizmi-Tessob	“Storm-god is shining”	p. 383
ḥḏmyn, ḥzmyn	Hizmi-Ani	“Sky-god is shining”	p. 383
kḏḡḏl	Kuzuh-adal	“Moon-god is strong”	p. 428
pḏḡḏn	Fandi-sena	“(Divine) brother is right”	p. 664
ṭmḡḏl	Simiga-adal	“Sun-god (is) strong”	p. 901
ṭrtṭb	Sarri-Tessob	“King is the Storm-god”	p. 920

Table 6. Hurrian theophoric names in Ugaritic transmission

The distribution between the Hurrian tense consonant ⟨t⟩ and its lax counterpart ⟨ḏ⟩ in the table above does not offer a challenge. In agreement with the rules formulated in Giorgieri 2000: 184, the lax sound occurs in intervocalic position and in clusters with sonorants. The only problematic position is the consonant after a morpheme boundary, where the analogy with the free-standing form of the same morpheme can block the lenition. This happens, for example, in all the forms featuring *tṭb* ‘Storm-god’ as the second component, while in other compounds the lenition of the morpheme boundary appears to be sporadic: *āḡl-ḏrm* vs. *irr-ṭrm* (presumably, a scribal error for *īwr-ṭrm*). What represents a major challenge for the interpretation of Hurrian sibilants proposed thus far is the use of the signs for the etymological interdental fricatives where one expects alveolar fricatives. In only a few instances, all marked in bold script in Table 6, do we find different conventions, which will be addressed at the end of this section.

According to a widely followed explanation in Tropper 2012, this convention was introduced *faute de mieux*. The Ugaritic speakers allegedly processed foreign alveolar fricatives as interdental fricatives since they allegedly lacked alveolar fricatives in their phonological inventory (see Section 2). It is, however, to be noted that Tropper 2012 predicts a phonological system featuring /ts/ = ⟨s⟩, /θ/ = ⟨t⟩, /ʃ/ = ⟨š⟩, but no phoneme /s/. To our best knowledge, there is no living language with such a phonological configuration. In fact, the database Moran and McCloy

2019 lists only five languages featuring /θ/ and /ʃ/ but no /s/, namely Turkmen (Turkic), Shawnee (Algonquian), Mbukushu and Gweno (Bantu), and Burun (Nilotic). None of these five languages, however, has the voiceless alveolar affricate /ts/ in its phonological inventory.

An obvious alternative to Tropper's solution is to take Hurrian evidence at face value, at least in part. The Hurrian voiceless alveolar fricative /s/ was rendered as ⟨t⟩ in the Ugaritic alphabetic script because Ugaritic ⟨t⟩ was pronounced as /s/. The implied sound change */θ/ > /s/ is typologically common, occurring, for example, in Ethio-Semitic languages (Kogan 2011: 55). In the history of Ugaritic, this change was presumably realized as a pull shift, following the equally trivial change */s/ > /ʃ/. What remains to be seen is whether the new interpretation complicates the account of contacts between Ugaritic and its neighbour languages or certain aspects of Ugaritic grammar.

In fact, the opposite is true: the interpretation ⟨t⟩ = /s/ simplifies the description of Ugaritic phonology.³² Dwelling on the topic of language contact, the spelling of the Hittite king Suppiluliuma as *tpllm* no longer requires any additional assumptions: Hittite /s/ was adapted in Ugaritic as /s/. A recently emerged parallel involving a Luwian loanword in Ugaritic is *mlḡtm* 'offerings', matching *malḥašše* in Akkadian texts from Ugarit (Melchert 2024: 140 with reference to Pardee, forthcoming). The Luwian word is attested itself as *malḥašša-* and LIBARE-*hasa-* in cuneiform and hieroglyphic transmissions, respectively.³³ The geographic name *kṭky* 'Kaskaean' (DUL: 466) was presumably borrowed via Hittite or Luwian, since the Kaskaean constituted a northern Anatolian population group. In this case, we must also be dealing with an adaptation of /s/ in Ugaritic, whatever the original self-designation of the Kaskaean was. The geographic name *ātry* 'Assyrian' (DUL: 125) can likewise be explained via the adaptation of Assyrian /s/ (see Section 1 for the preservation of /s/ in the Assyrian dialect of Akkadian). The list of geographic names featuring ⟨t⟩ in Ugaritic transmission, despite the lack of independent evidence for the presence of /θ/ in these toponyms, can easily be extended (cf. the presentation of a similar argument in Friedrich 1943: 8–9).

A different argument for the interpretation of Ugar. ⟨t⟩ as /s/ is the high overall frequency of this phoneme, which distinguishes it from the reflexes of */θ/ in the other Semitic languages (Tropper 2012: 102). This suggests that, in addition to its primary origin, this phoneme must have had additional sources. According to Tropper (2012: 109–110) these are the irregular reflexes of Proto-Semitic *s₁ and, more rarely, *s₂, but most of the relevant etymologies are doubted in Kogan 2011: 94–95 on valid grounds. An alternative explanation would be a large number of loanwords with ⟨t⟩ absorbed by Ugaritic. Some of the relevant non-Semitic loanwords have Hurrian counterparts, e.g. *ūṭpt* 'quiver' vs. Hurr. *išpa(n)ti* 'id.' (DUL: 123), *ṭdr* 'waitress, girl' (?) vs. Hurr. *šiduri* 'maiden, girl' (DUL: 886), *ṭryn* '(suit of) armour, protective padding' vs. Hurr. *šar(i)yanni* 'cuirass' (DUL: 921). Given the demonstrably close contacts be-

³² In what follows, we forgo a detailed discussion of the data that can be easily reconciled with the phonetic interpretation of ⟨t⟩ as /s/ as well as /θ/. For example, the fact that the Canaanite reflexes of */θ/ were rendered with the phonograph ⟨s⟩ in Egyptian (Hoch 1994: 402) is of limited value for our purposes, since /s/ was the closest sound to /θ/ available in the Egyptian phonetic inventory. *Mutatis mutandis*, the same goes for the rendering of the Ugaritic reflexes of */θ/ by the Š-series in cuneiform orthography (Tropper 2012: 112) or for the possible borrowings of Babylonian /t/ as Ugaritic ⟨t⟩.

³³ A potential piece of counter-evidence against the adaptation of Luwian /s/ as ⟨t⟩ is the Luwian theophoric personal name attested in Ugaritic transmission as *trḡ(n)ds* (Patri 2009: 97; del Olmo Lete and Sanmartín 2015: 865). While the connection of this name with the Luwian Storm-god Tarhunt is undeniable, its final consonant need not reflect an etymological fricative. The abbreviated form of the Luwian name *Tarhunta-ziti*, lit. "man of the Storm-god", namely *Tarhunta-zi*, attested as such in Anatolian hieroglyphic transmission, provides a regular counterpart to the Ugaritic form, with the expected rendering of Luwian /ts/ as ⟨s⟩ (Yakubovich 2022b: 240).

tween the two languages, the Hurrian origin of these Ugaritic lexemes ranges from certain to probable. A different case is that of the Ugaritic nouns with Hittite matches, e.g., *mtyn* ‘a garment’ vs. Hitt. ^{TUG}*maššiya*- ‘id.’ (DUL: 598), *ītl* ‘saliva, spittle’ vs. Hitt. *iššalli*- ‘id.’ (DUL: 122), or *ḥtt* ‘silver’, which evokes the variant Hittite spellings ^{URU}*ḥattuša* and ^{URU}*KÛ.BABBAR-ša* ‘Hattusa’, where *KÛ.BABBAR* is the Sumerogram for ‘silver’ (DUL: 372). Since direct loanwords from Hittite to Ugaritic are few and far between, such cases may bear witness to a common borrowing from a third source. Since /s/ is the most common sibilant cross-linguistically, it is indeed probable that the loanwords containing this phoneme, many of which have not yet been identified as such, increased its overall frequency. Such an account would, however, be impossible under the traditional interpretation of ⟨t⟩ as /θ/.

Further support for the shift */θ/ > /s/ in the pre-history of Ugaritic comes from a morphophonemic change. To quote Schneider (2024: 87): “A truly remarkable case of sibilant grammatical rule in Semitic is the sibilant-interdental harmony rule found in Ugaritic. On the surface, the rule is a shift of the causative prefix */ʃ/ → /θ/ if the root has */θ/ as one of its radicals”. One must share the author’s surprise: the only other instances of sibilant harmony involving interdental fricatives that are familiar to us were described in the Tahtlan language and the Doig River dialect of the Beaver language, both belonging to the Athapaskan group. In these cases, however, the rule has a historical explanation: the opposition between /θ/ and /s/ in Tahltan and Beaver goes back to the opposition between */s/ and */ʃ/ in Proto-Athapaskan. (Hansson 2001: 69). In contrast, the harmony involving the alveolar vs. postalveolar distinction, that is /s, z, ts, dz/ vs. /ʃ, ʒ, tʃ, dʒ/, represents the most common type of sibilant harmony (Hansson 2001: 56). It is attested in the better-known Athapaskan languages, such as Navajo and Apache, but also occurs in many other languages across the globe, including Coptic dialects and Moroccan Arabic. The assumption of the anticipatory assimilation */ʃ/ → /s/ in the Ugaritic causative prefix, triggered by the presence of /s/ in the verbal root, immediately brings Ugaritic in line with this typologically common pattern.

Thus, there is convergent evidence for the interpretation of Ugaritic ⟨t⟩ as a voiceless alveolar fricative. At the same time, there is no need to extend the alveolar interpretation to the other etymological interdental sounds. As shown in Tropper 2012, Ugaritic ⟨d⟩ and ⟨z⟩ show variation with ⟨d⟩ and ⟨t⟩, respectively, in certain lexemes, while others bear witness to the historical changes */ð/ > ⟨d⟩ and */t/ > ⟨z⟩ (see Section 3). This pattern is consistent with the vestigial preservation of voiced and “emphatic” interdental fricatives in Ugaritic, despite their occasional confusion with coronal stops. The fact that no such confusion is attested in the instance of ⟨t⟩ underscores once again that it no longer belonged to the interdental series. The phonetic reality behind the change */ð/ > ⟨d⟩ is confirmed by syllabic cuneiform: sequences with the reflexes of */θ/ are always rendered by the Š-series, whereas both Z and D-series are used for the reflexes of */ð/ in Ugaritic (Tropper 2012: 112, 118).

The discrepancy between the development of */θ/ and the other interdental fricatives finds a motivation in the structure of the Ugaritic phonological system. Assuming that */θ/ > /s/ was indeed a pull shift triggered by */s/ > /ʃ/, there was nothing to “pull” */ð/ and */θ/ in the direction of the alveolar fricatives, given the absence of */z/ and */s/ in the Proto-Semitic phonemic inventory. Nor is the differential treatment of */θ/ and */ð/ unparalleled in Semitic. Schneider (2024: 92) provides the following information about two North-Eastern Neo-Aramaic dialects: “In the Jewish dialect of Challa, */θ/ merged with [s] but */ð/ merged with [d]. In the Jewish dialect of Amediya, /θ/ remained distinct, but /ð/ merged with [d], a situation reminiscent of that in Ugaritic”. Under our interpretation, the state of affairs in Ugaritic appears closer to that of the Challa dialect.

The preservation of */ð/ in at least some attested varieties of Ugaritic prompts the question why this phoneme was chosen to render /z/ in Hurrian personal names (see Table 6).³⁴ In this case one can legitimately argue that /ð/ was the closest phonetic approximation to Hurrian /z/, since there were no other voiced fricatives in Ugaritic for as long as /dz/ remained an affricate. Yet, this convention was not altogether stable. The spellings such as *āḡlkz* and *ḡlkz* alongside *īḡlkḏ* or *ḥḏmyn* alongside *ḥzmyn* suggest that, in some cases, the Ugaritic scribes opted for the faithful rendering of the place of articulation (alveolar) rather than the mode of articulation (fricative).³⁵ It is more difficult to account for the name *aršmg* instead of the expected **arθmg* = Ari-Simiga. In the absence of a plausible orthographic explanation, one is tempted to treat this example as reflecting the palatalization of /s/ between two front vowels. Mauro Giorgieri turns our attention to the variation between *āṯṯ* and *āṯš*, both rendering Hurrian *aššišša* ‘wash it!’ in parallel incantations in Ugaritic transmission (Giorgieri 2013: 174). In this case, too, the secondary š occurs in a position after the front vowel.³⁶

In view of the available converging evidence for the Ugaritic realization of Proto-Semitic */θ/ as /s/, it is worth asking what arguments could be advanced for the preservation of */θ/ in Ugaritic. The claim that the absence of a sound change vis-à-vis Proto-Semitic represents a more parsimonious solution would not qualify as a valid argument: the majority of the known Semitic languages did not preserve the interdental articulation of */θ/ (Schneider 2024: 28–29), as has been acknowledged in the field of Comparative Semitics since its early days. What arguably had the potential to sway the *consensus majorum* is the presumed parallelism in the treatment of */θ/ and */ð/. The seminal source that pleads against interpreting Ugaritic ⟨t⟩ as /θ/ is Friedrich 1943; the same paper claims that Ugaritic ⟨ḏ⟩ represents the voiced counterpart of ⟨t⟩, based on the analysis of the Hurrian names in Ugaritic transmission. The demonstration that Proto-Semitic */ð/ is frequently reflected as ⟨d⟩ in Ugaritic is found in Tropper 2012; the same monograph advocates the interpretation of Ugaritic ⟨t⟩ as /θ/.³⁷ Yet, formal similarities need not overshadow the differences in the evolution of the two interdental fricatives. On the one hand, ⟨ḏ⟩ was not used in the transmission of the Hittite and Luwian loanwords; on the other hand, Proto-Semitic */θ/ was not reflected in Ugaritic as ⟨t⟩. Furthermore, Hurrian /s/ was never reflected in Ugaritic as ⟨s⟩, contrast the occasional renderings of Hurrian /z/ as ⟨z⟩. The divergent phonetic development of */θ/ and */ð/ assumed in the present paper reflects a compromise position between the views of Friedrich 1943 and Tropper 2012. While such a solution may seem counter-intuitive at first glance, we have argued that it becomes a necessity once all the relevant data are taken into consideration.

³⁴ Cf. the following comment: “The fact that ⟨ð⟩, rather than ⟨z⟩, is the letter used for transcription of certain foreign phonemes represented by ⟨Z⟩ and ⟨Š⟩ in syllabic cuneiform, is perplexing” (Schneider 2024: 34).

³⁵ A difficult case is the loanword *brḏl* ‘iron’ (DUL: 234). Although this noun is ultimately borrowed from Luwian via Akkadian (Valério and Yakubovich 2010), one would expect Akkadian *parzillu* ‘iron’ to be reflected in Ugaritic as **przl*. Perhaps both the voicing of the initial consonant and the lenition */dz/ > /z/ reflect further mediation of an (unknown) Northwest Semitic dialect.

³⁶ The palatalization /s/ > /j/ in a variety of environments, including the position after /i/, was proposed for Late Luwian in Rieken 2010. A similar conditioned change is attested in a variety of ancient Indo-European languages, including the famous *ruki*-rule in Old Indic. If Ugar. *ḥšt* ‘sepulchre, mausoleum’ is borrowed from Hitt. *ḥešta-/ḥešti-*, the name of an installation in Hattusa that was related to the ancestor cult, this is a faint argument for Hittite being not immune to a conditioned palatalization of /s/. This is, however, not the only and perhaps not the most likely etymology of Ugar. *ḥšt* (cf. DUL: 407–408 and Schneider 2024: 217).

³⁷ The main argument advanced for the preservation of */θ/ in Ugaritic in Tropper 2012: 113 is the alleged parallelism of phonetic environments favouring the change **s*₁ > ⟨t⟩ and the preservation of ⟨ḏ⟩ in Ugaritic. Yet, the fact the bulk of examples adduced for *s*₁ > ⟨t⟩ is doubted in Kogan 2011: 94–95 undermines the formulation of specific licensing conditions for the sound change under discussion.

The conclusions of the present section are summarized from the historical perspective in Table 7, where the Proto-Semitic forms are reconstructed according to the middle version of the Affricate Hypothesis. It is easy to observe that the development of Ugaritic sibilants shares traits with the evolution of their counterparts in both Hebrew and Aramaic (Syriac) under the proposed analysis. The pull shift */θ/ > /s/ is assumed to represent a common Canaanite innovation in the second millennium BCE, which appears to be consistent with the evidence from Canaanite data in Egyptian transmission (Hoch 1994: 403–404). The positive reason for reconstructing this intermediate realization for the ancestor dialect of Hebrew is the same as for postulating it in the case of Ugaritic: a language featuring /ts/, /θ/ and /ʃ/ but no /s/ would be typologically unique.³⁸

Proto-Semitic	Hebrew	Syriac	Ugaritic (syllabic series)	Ugaritic (alphabetic)	Ugaritic (phonetic)
*ts	/s/	/s/	Z, more rarely S	⟨s⟩	/ts/, /s̥/
*t̥	/t̥/	/s/	Š	⟨s⟩	/ʃ/
*s	/ʃ/	/ʃ/	Š	⟨s⟩	/ʃ/
*θ	*s > /ʃ/	/t/	Š	⟨t⟩, rarely ⟨z⟩	/s/
*ð	/z/	/d/	D, Z	⟨d⟩ > ⟨d⟩	/ð/ > /d/

Table 7. Evolution of sibilants in Northwest Semitic

6. Sibilants in peripheral Akkadian

We have seen that the phonetic interpretation of Hurrian sibilants undertaken in Section 4 is compatible with the phonetic interpretation of Ugaritic sibilants undertaken in section 5, and indeed plays the crucial role in elucidating the Ugaritic data. The purpose of this section is to argue that this interpretation of the Hurrian data also facilitates the account of the values of the Š and Z-series that are assumed for Anatolian cuneiform.

According to a widely accepted hypothesis, the first cuneiform scribes at the service of the kingdom of Hattusa were deported or recruited from Alalakh after Hattusili I conquered this Syrian town (van den Hout 2020: 50).³⁹ Although Akkadian was the only attested written language in Alalakh in the relevant period, the individuals with Akkadian names demonstrably constituted a tiny minority of the local population. According to the statistics collected in Drafkorn 1959: 118, both the Amorite and the Hurrian names are attested about five times more frequently than the Akkadian names in the texts from the archaeological layer Alalakh VII, which reflect the ethnolinguistic composition of this town immediately before the expedition of Hattusili I. Significantly, some Alalakh VII scribes also had Hurrian names (Oliva 1999: 30), which opens the possibility that others were native speakers of Hurrian who assumed Akkadian names for reasons of prestige. Although the core of grammar of Alalakh VII texts fol-

³⁸ The preservation of Canaanite /ts/ in the second millennium BCE is borne out by Canaanite forms in Egyptian transmission (Hoch 1994: 407–408). For the possible development of the affricate /ts/ into the laminal fricative /s̥/ in the history of Ugaritic, see fn. 15.

³⁹ The Anatolian cuneiform did not inaugurate literacy in Asia Minor, since the Akkadian texts in Old Assyrian Script circulated there in the Karum period (20th–18th centuries BCE). Neither the ductus nor the sign inventories lend support to the continuity between the Old Assyrian and Anatolian cuneiform, even though there is a possibility that certain Anatolian cuneiform scribes were passively familiar with Old Assyrian (Rieken and Yakubovich 2023: 206–207).

lows the Akkadian norm (Weeden 2019: 132), they feature technical terms borrowed from Hurrian (Oliva 1999: 29). The Hurrian presence in Alalakh continues to mount after the expedition of Hattusili I, and by fifteenth century BCE the local tablets from the layer Alalakh IV are written in an imperfectly learned variety of Akkadian showing interference with Hurrian and commonly labelled Hurro-Akkadian. The use of Hurro-Akkadian in Late Bronze Age Syria must be distinguished from the contemporary use of Canaanite-Akkadian in southern Levant, where an advanced degree of interference between Akkadian and West Semitic contrasts with the lack of significant Hurrian influence (Andrason and Vita 2016).

A potential rationale for the distinction between cuneiform writing practices in the northern and southern Levant is that the latter never belonged to the Mitanni kingdom. Yet such a generalization would be deceptive. On the one hand, it is worth pointing out arguments for the Hurrian presence in the Levant south of the Mitanni borders. These include the Hurrian and Indo-Aryan ruler names associated with Canaan in mid-second-millennium BCE (Kilani 2020: 76) as well as some Hurrian loanwords in Biblical Hebrew (Noonan 2019: 309). On the other hand, the Hurrian influence on Akkadian predates the formation of Mitanni, as we have seen at the example of Alalakh VII. These considerations prompt a more restrictive and, in a sense, more obvious hypothesis: the Hurrian influence on peripheral Akkadian grew hand in hand with the presence of Hurrians in the respective epigraphic communities.⁴⁰

A prediction of this hypothesis is the presence of phonetic interference in contact-induced varieties of Akkadian reflecting the impact of Hurrian. Indeed, while the grammatical rules of Akkadian could in principle be learned based on the study of texts available in the school curriculum, the pronunciation of Akkadian forms was likely to be affected by the scribe's native language. It is reasonable to assume that the scribe of Hurrian origin spoke and read Akkadian with a Hurrian accent: after all, their job was to write good Akkadian and not to speak good Akkadian. Since the Hurrian phonological system features certain constraints that are absent in Akkadian phonology, it is worth investigating whether the choice of syllabic signs in Alalakh VII texts may have anything to do with following such constraints. If so, we would obtain an empirical confirmation of phonetic interference between Akkadian and Hurrian in the chanceries of Alalakh VII.

The conventions for rendering Akkadian stops in Alalakh VII, as described in Popova 2016, supply the requisite evidence. Contrary to the previous claims, this paper provides a rigorous statistical proof of the contention that the use of the syllabic signs in Alalakh VII correlated with the phonetic opposition between voiceless and voiced stops (and, to some extent, also emphatic stops). This is the innovation vis-à-vis the state of affairs in Old Akkadian and Old Assyrian, where the same signs can indiscriminately be used for the syllables beginning with voiceless, voiced, and emphatic plosives. Yet, in comparison with the contemporary Mari orthography, the texts of Alalakh VII render the laryngeal oppositions in a fairly inconsistent fashion, frequently confusing the signs of T and D, P and B, K and G-series, as had also been noticed in previous scholarship (see, e.g., Giacomakis 1970: 26).

Now, assuming that the graphic conventions for distinguishing voiceless, voiced and emphatic stops were known to the scribes, their haphazard implementation appears to be rather unlikely unless one assumes that the scribes experienced difficulties with identifying the re-

⁴⁰ This brief discussion is not meant to gloss over a complicated question of why Hurrians rather than Amorites played the key role in the spread of Akkadian literacy to northwestern Syria. The full answer to this question lies beyond the scope of the present paper and may even be impossible given the current state of our knowledge. It is worth, however, pointing out that the Tish-Atal inscription bears witness not only to the early exposure of the Hurrians to cuneiform literacy but also to their willingness to adopt other aspects of the Mesopotamian culture, which manifests itself through the use of Sumerograms for Hurrian gods.

spective sounds. At this point it is important to keep in mind that the Hurrian language features a different opposition between the tense and lax consonants, which are furthermore neutralized in many positions, in particular, word-initially and after another consonant (Giorgieri 2000: 184). This implies that the difference between the Akkadian voiced and voiceless consonants, not to mention their “emphatic” counterparts, must have been suppressed in Hurrian pronunciation in many cases. Therefore, its consistent implementation in writing would have represented a remarkable challenge for scribal education in Alalakh, which is consistent with the available evidence.

Thus, one can adopt a working hypothesis that the orthography of Alalakh VII Akkadian may reflect phonetic interference with Hurrian.⁴¹ This brings us to the graphic conventions for recording the Akkadian sibilants, which are likewise fuzzier in the Alalakh VII texts than in Old Babylonian. For example, the Old Babylonian orthography, as reflected in the texts emanating from Hammurapi’s chancery, systematically distinguishes between the Z, S, and Š-series. The signs of Z-series are reserved for sequences with the synchronic affricates, while the use of the S-series reflects secondary conditional deaffrication to /s/, whose phonetic outcome was, however, different from the reflexes of Proto-Semitic */s/ rendered by the Š-series (Streck 2006). In contrast, the Hurrian system of sibilants, as reconstructed in Section 4, did not feature secondary deaffrication. Accordingly, the S-series stays marginalized in the texts from Alalakh VII, while the contrast between the Z and Š-series in the local orthography correlates with the etymological distinction between the Semitic affricates and */s/ (Giacumakis 1970: 26–27).⁴² In this respect, it represents a logical starting point for the development of the Hittite orthography, where the S-series is marooned among the heterograms and no longer has a phonological application.

Yet, the similarities between Hittite orthography and the conventions governing the use of signs in the Alalakh VII corpus appear to go beyond the destiny of the S-series. If one follows the hypothesis that the pronunciation of Alalakh Akkadian was approximated to the phonological system of Hurrian, this implies the alveolar pronunciation of Akkadian fricatives. This finds, of course, a parallel in the alveolar pronunciation of the fricative sounds in Hittite (cf. Section 1). At the same time, many scholars assume that the Old Babylonian fricative hidden behind the Š-series had a different phonetic realization; according to Schneider 2024: 30, this was the lateral sound /ɬ/, later evolving into /ʃ/.⁴³ As for the adaptation of the Akkadian affricates, the conclusions of Section 4 are compatible with the assumption that they

⁴¹ An attempt to argue against the role of the Hurrians in the transmission of cuneiform from Syria to Anatolia is found in Giusfredi et al. 2023: 164–167. It seems, however, that the authors primarily target a substantially different claim from the one advocated here, namely that “Hurrian played a role in the diffusion of cuneiform to Anatolia”. There is certainly no evidence that Hurrian was used in writing in Alalakh VII, or for that matter in Hattusa before the 15th century BCE. There is, however, sufficient evidence that Hurrian was spoken in Alalakh VII.

⁴² The marginalization of the S-series finds a close parallel in Old Assyrian orthography and was probably ubiquitous in the history of cuneiform after the collapse of the Old Akkadian writing system. One must mention for fairness’ sake that the S-series found a new functional application in certain Syrian cuneiform systems of the second millennium BCE. Thus, the S and Z-series stay in a complementary distribution in the bulk of Hurrian texts found in Emar (Fischer 2025: 172–178); according to the phonetic interpretation of the present paper, they denote combinations with the tense and lax affricates, respectively. Such a convention is typical of the so-called “Middle Euphratic” orthography (cf. also the discussion below regarding the parallel “Middle Euphratic” convention for rendering fricatives). For the use of the “Middle Euphratic” orthography in Mari, see fn. 10.

⁴³ It is, of course, also possible to argue that the Hattusa scribes approximated the Š-series for rendering the Hittite alveolar fricatives merely because it was the only series associated with the fricatives in peripheral Akkadian, regarding of their specific place of articulation. Be it as it may, the match between Hittite and Alalakh Akkadian addressed immediately above simplifies the scenario of script adaptation.

came to be pronounced in Alalakh as either /ts/ or /dz/, the contrast between the two phonemes being restricted to the intervocalic position. This is a simplification vis-à-vis the phonology of the natively transmitted Akkadian, which maintained a distinction between the voiceless, voiced and emphatic affricates in the relevant time period (Schneider 2024: 30). It stands to reason that the reconstructed adaptation of peripheral Akkadian phonology simplifies the account of how Anatolian cuneiform came into being.⁴⁴

Another reason to entertain the hypothesis that the Hurrian speakers were among the scribes brought to Hattusa by Hattusili I is the similarity of conventions for rendering the contrast between the tense and lax stops in Hittite orthography on the one hand and in the normalized Hurrian orthography of the Mitanni Letter on the other hand. In both instances the use of VC signs (consonant replication), e.g., *ša-aš-ša* as opposed to *ša-ša*, marks the tense stops in intervocalic position (Hoffner and Melchert 2024: 53, Wegner 2007: 43).⁴⁵ Now, at least some Alalakh VII texts appear to share this convention with the Mitanni Letter and Boğazköy texts with regard to the spelling of Hurrian proper nouns.⁴⁶ Therefore, one may hypothesize that marking tense consonants as geminates originally arose in the context of rendering Hurrian proper nouns and lexical code-switches embedded in Akkadian texts before being extended to the orthographies of the Hurrian and Hittite texts proper. Such a hypothesis, however, would make more sense if the task of applying cuneiform to writing Hittite was entrusted to Hurrian and Hittite bilinguals, who appreciated the similarities between the phonological systems of the two languages.

The tentative comparison between the rendering of Hurrian forms in Alalakh VII texts and in the normalized orthography of the Mitanni Letter does not, of course, imply that the writing conventions were identical in both cases. For a list of graphic innovations of the Mitanni Letter, some of which generalize the earlier tendencies, one can consult Giorgieri 2000: 181. Of primary interest for our purposes is the innovative use of the S-series for sequences with a Hurrian affricate that can be rendered with the Z-series elsewhere (cf. Section 4). In our opinion, this innovation can be compared with the equally innovative use of the S-series in Old Babylonian cuneiform, where it came to denote sequences with de-affricated sounds (Streck 2006). In Mitanni as well as in Babylon, we are presumably dealing with creative efforts by well-established scribal schools, which strove to refine the inherited writing systems in the hope of gaining more phonological transparency for the Hurrian and Akkadian lan-

⁴⁴ It must be stressed that although the Alalakh VII corpus provides a logical starting point for our discussion of peripheral Akkadian, other contemporary peripheral Akkadian texts may have shared the same distribution and pronunciation of signs. This observation is important for purposes of the present paper, because there is a possibility that the first Hattusa scribes were imported not from Alalakh but from Aleppo (Halab), as Mirko Novák (Bern) is pointing out to us in a personal communication. Such a possibility is rarely discussed mainly because it is an *argumentum ex silentio*: the relevant layers of Bronze Age Aleppo largely remain unexcavated. Note also the objections of Popova 2016: 89 against the direct continuity between the orthographies of Alalakh VII and Hattusa.

⁴⁵ This convention is not so trivial as it may seem, because there was a competing “Middle Euphratic” strategy of rendering the intervocalic /(s)s/, /z/ in Hurrian forms by Š and Z series, respectively, which is attested, for example, in the Mari documents (see further Richter 1998: 131–133 and cf. fn. 10). The two competing strategies of rendering the same opposition, namely the contrast between the Š and Z series and consonant replication, are found distributed across Hurrian texts in Late Bronze Age Emar (Fischer 2025: 135). The second strategy is, however, limited to a handful of Emar tablets, all sharing additional epigraphic features that point out to their intrusive character.

⁴⁶ Contrast, e.g., the theophoric personal names *Ku-ša-aḥ-e-ri* (*40 rev. 1) and *A-ri-ip-te-eš-šu-pi* (*40 rev. 3) found on the same Alalakh VII tablet and featuring the names of the Moon-god Kuzuh and the Storm-god Tessob as their components (Wiseman 1953: 44).

guages respectively. The same S-series, which had been rendered redundant by the preceding evolution of the syllabic cuneiform script, was secondarily deployed to accomplish the stated goals in both cases.⁴⁷ Against the background of these reforms, the adaptation of the sibilant sign values to Hittite cuneiform emerges as fairly trivial.

References

- Andrason, Alexander, Juan Pablo Vita. 2016. Contact Languages of the Ancient Near East — Three more Case Studies (Ugaritic-Hurrian, Hurro-Akkadian and Canaanite-Akkadian). *Journal of Language Contact* 9: 293–334.
- BGH = Richter, Thomas. 2012. *Bibliographisches Glossar des Hurritischen*. Wiesbaden: Harrassowitz.
- Breyer, Francis. 2010. *Ägypten und Anatolien. Politische, kulturelle und sprachliche Kontakte zwischen dem Niltal und Kleinasien im 2. Jahrtausend v. Chr.* Vienna: Österreichische Akademie der Wissenschaften.
- Burlingame, Andrew R. 2020. New Evidence for Ugaritic and Hittite Onomastics and Prosopography at the End of the Late Bronze Age. *Zeitschrift für Assyriologie* 110: 196–211.
- Bush, Frederic W. 1964. *A Grammar of the Hurrian Language*. Ph.D. dissertation. Brandeis University.
- Campbell, Dennis R. M. 2020. Hurrian. In: Rebecca Hasselbach-Andee (ed.). *A Companion to Ancient Near Eastern Languages*: 203–219. Hoboken, NJ: Wiley Blackwell.
- Cross, Frank M. 1962. Yahweh and the God of the Patriarchs. *Harvard Theological Review* 55: 225–259.
- Degen, Rainer. 1967. Zur Schreibung des Kaška-Namens in ägyptischen, ugaritischen und altaramäischen Quellen: Kritische Anmerkungen zu einer Monographie über die Kaškäer. *Die Welt des Orients* 4: 48–60.
- Diakonoff, Igor M. 1971. *Hurrisch und Urartäisch. Transl. K. Sdrembek*. Munich: Kitzinger.
- Draffkorn, Ann. 1959. *Hurrian and Hurrians in Alalah: an ethno-linguistic analysis*. Ph.D. Thesis. University of Pennsylvania.
- DUL = del Olmo Lete, Gregorio, Joaquín Sanmartín. 2015. *A Dictionary of the Ugaritic Language in the Alphabetic Tradition. Third Revised Edition*. Leiden: Brill.
- Fischer, Sebastian. 2025. *Die hurritischen Texte aus Emar. Einblicke in die hurritischsprachige Omentradition*. Habilitation thesis. Free University of Berlin.
- Friedrich, Johannes. 1943. Kleinigkeiten zur ugaritischen Grammatik. *Orientalia* (NS) 12: 1–22.
- Gelb, Ignace G, Pierre M. Purves, Allan A. MacRae. 1943. *Nuzi Personal Names*. OIP 57. University of Chicago Press.
- Giacumakis, George. 1970. *The Akkadian of Alalah*. The Hague: Mouton.
- Giorgieri, Mauro. 2000. Schizzo grammaticale della lingua hurrica. *Parola del Passato* 55: 171–295.
- Giorgieri, Mauro. 2013. Diffusion et caractéristiques de la culture écrite d'origine hurrite dans le Proche-Orient asiatique et à Ougarit. In: Pierre Bordreuil et al. (eds). *Les écritures mises au jour sur le site antique d'Ougarit (Syrie) et leur déchiffrement*: 157–185. Paris: Académie des Inscriptions et Belles-Lettres.
- Giusfredi, Federico, Alvis Matessi, Valerio Pisaniello. 2023. Hittite Anatolia and the Cuneiform Koiné. In: Federico Giusfredi, Alvis Matessi, Valerio Pisaniello (eds). *Contacts of Languages and Peoples in the Hittite and Post-Hittite World*: 159–185. Leiden: Brill.
- Greppin, John A. C. 2011. Urartian Sibilants in Armenian. *Historische Sprachforschung* 124: 292–296.
- Hansson, Gunnar Ólafur. 2001. *Theoretical and Typological Issues in Consonant Harmony*. Ph.D. dissertation. University of California at Berkeley.
- Harouthiounian, Nikolaj. 2001. *Korpus urartskix klinoobraznyx nadpisej*. Yerevan: Gituthiun.
- Hasselbach, Rebecca. 2005. *Sargonic Akkadian: A Historical and Comparative Study of the Syllabic Texts*. Wiesbaden: Harrassowitz.
- Hasselbach-Andee, Rebecca. 2021. Classification of Akkadian within the Semitic Family. In: Juan-Pablo Vita (ed.). *History of the Akkadian Language*: 129–146. Leiden: Brill.
- Hawley, Robert. 2020. Ugaritic. In: Rebecca Hasselbach-Andee (ed.). *A Companion to Ancient Near Eastern Languages*: 257–278. Hoboken, NJ: Wiley Blackwell.
- Hazenbos, Joost. 2021. Hurritisch und Urartäisch. In: Michael P. Streck (ed.). *Sprachen des Alten Orients*: 166–191. Darmstadt: Wissenschaftliche Buchgesellschaft.

⁴⁷ For the use of the S-series in correlation with a set of phonemes (Proto-Semitic */s/ and */ʃ/) in Old Akkadian orthography, see Hasselbach 2005: 135–137.

- Hoch, E. James. 1994. *Semitic Words in Egyptian Texts of the New Kingdom and Third Intermediate Period*. Princeton: Princeton University.
- Hoffner, Harry A. Jr., H. Craig Melchert. 2024. *A Grammar of the Hittite Language. Part I: Reference Grammar* (2nd ed.). University Park, PA: Pennsylvania State University.
- Hutton, Jeremy M. 2006. Ugaritic */š/ and the Roots šbm and šm[d] in KTU 1.3.III.40. *Maarav* 13/1: 75–83.
- van den Hout, Theo. 2020. *A History of Hittite Literacy*. Cambridge University Press.
- Huehnergard, John. 2012. *An Introduction to Ugaritic*. Peabody, MA: Hendrickson.
- Huehnergard, John. 2019. Proto-Semitic. In: John Huehnergard, Na‘ama Pat-El (eds.). *The Semitic languages* (2nd ed.): 49–79. London: Routledge.
- Huehnergard, John, Na‘ama Pat-El. 2019. *Introduction to the Semitic languages and their history*. In: John Huehnergard, Na‘ama Pat-El (eds.). *The Semitic languages* (2nd ed.): 1–21. London: Routledge.
- Kassian, Alexei S., Olga V. Popova. 2023. Phonology of the Late Babylonian cuneiform signs for sibilants. *Ancient Near Eastern Studies* 60: 231–248.
- Khachikian, Margarita L. 1985. *Xurritskij i urartskij jazyki*. Yerevan: Armenian SSR, Academy of Sciences.
- Kilani, Marwan. 2020. Naming practices and identity in the early late bronze age Levant: A linguistic and geographical analysis of local rulers’ names attested in the Amarna letters. *Journal of Ancient Egyptian Interconnections* 27: 70–93.
- Kilani, Marwan. 2023. On the traces of apples, plums, and pears: Investigating a wanderword in ancient and modern Near Eastern languages. *Journal of Historical Linguistics* 15/1: 78–131.
- Kimball, Sarah. 2017. The phonology of Anatolian. In: Jared Klein, Brian Joseph, Mattias Fritz (eds.). *Handbook of Comparative and Historical Indo-European Linguistics, Vol. 1*: 249–256. Berlin: De Gruyter.
- Kogan, Leonid. 2009. Ugaritskij jazyk. In: Anna G. Belova et al. (eds.). *Jazyki mira: Semitskije jazyki*: 205–238. Moscow: Nauka.
- Kogan, Leonid. 2011. Proto-Semitic Phonetics and Phonology. In: Stefan Weninger (ed.). *The Semitic Languages: An International Handbook*: 54–151. Berlin: De Gruyter.
- Kogan, Leonid. 2015. *Genealogical Classification of Semitic*. Berlin: De Gruyter.
- Kouwenberg, N. J. C. 2003. Evidence for Post-Glottalized Consonants in Assyrian. *Journal of Cuneiform Studies* 55: 75–86.
- Kouwenberg, N. J. C. 2017. *A Grammar of Old Assyrian*. Leiden: Brill.
- Loprieno, Antonio. Ancient Egyptian and Coptic. In: Roger Woodard (ed.). *The Cambridge Encyclopedia of the World’s Ancient Languages*: 160–217. Cambridge University Press.
- Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge University Press.
- Melchert, H. Craig. 1994. *Anatolian Historical Phonology*. Atlanta: Rodopi.
- Melchert, H. Craig. 2020. Luwian. In: Rebecca Hasselbach-Andee (ed.). *A Companion to Ancient Near Eastern Languages*: 239–256. Hoboken, NJ: Wiley Blackwell.
- Melchert, H. Craig. 2024. *A Dictionary of Cuneiform Luwian*. Ann Arbor: Beech Stave.
- Moran, Steven, Daniel McCloy (eds.). 2019. *PHOIBLE 2.0*. Jena: Max Planck Institute for the Science of Human History [online database available at <http://phoible.org>, accessed on 2024-02-21].
- Nielsen, Rasmus. 2023. Prehistoric loanwords in Armenian: Hurro-Urartian, Kartvelian, and the unclassified substrate. Ph.D. dissertation. Leiden University.
- Noonan, Benjamin J. 2019. *Non-Semitic Loanwords in the Hebrew Bible: A Lexicon of Language Contact*. University Park, PA: Eisenbrauns.
- Oliva Mompeán, Juan Carlos. 1999. Consideraciones sobre el hibridismo hurro-semitico en el Alalakh VII. *Veleia* 16: 19–32.
- Oliva Mompeán, Juan Carlos. 2016. Hurro-Akkadian from Late Bronze Age Syria Reconsidered: Qatna TT1 and Al.T. 297. *Aula Orientalis* 34/1: 139–148.
- Pardee, Dennis. 2011. Ugaritic. In: Stefan Weninger et al. (eds.). *The Semitic languages: An international handbook*: 460–472. Berlin: De Gruyter.
- Pardee, Dennis. Forthcoming. *Les textes épistolaires en langue ougaritique de Ras Shamra et de Ras Ibn Hani*. Beirut: Institut Français du Proche-Orient.
- Patri, Sylvain. 2009. La perception des consonnes hittites dans les langues étrangères au XIII^e siècle. *Zeitschrift für Assyriologie* 99: 87–126.
- Patri, Sylvain. 2019. *Phonologie hittite. HdO 130*. Leiden: Brill.

- Peust, Carsten. 1999. *Egyptian Phonology: An Introduction to the Phonology of a Dead Language*. Goettingen: Peust and Gutschmidt.
- Popova, Olga V. 2016. Cuneiform Orthography of the Stops in Alalah VII Akkadian. *Zeitschrift für Assyriologie* 106/1: 62–90.
- Richter, Thomas. 1998. Anmerkungen zu den hurritischen Personennamen des *ḫapiru*-Prismas aus Tigunānu. In: David I. Owen, Gernot Wilhelm (eds.). *Studies on the Civilization and Culture of Nuzi and the Hurrians* 9: 125–134. CDL Press: Bethesda, MD.
- Richter, Thomas. 2016. *Vorarbeiten zu einem hurritischen Namenbuch. Erster Teil: Personennamen altbabylonischer Überlieferung vom Mittleren Euphrat und aus dem nördlichen Mesopotamien*. Wiesbaden: Harrassowitz.
- Rieken, Elisabeth. 2010. Das Zeichen (ṣa) im Hieroglyphen-Luwischen. In: Aygül Süel (ed.). *Acts of the VIIth International Congress of Hittitology*: 651–660. Ankara: Anıt.
- Rieken, Elisabeth. 2021. Hethitisch. In: Michael P. Streck (ed.). *Sprachen des Alten Orients*: 80–127. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Rieken, Elisabeth, Ilya Yakubovich. 2023. Encounters between Scripts in Bronze Age Asia Minor. In: Ilona Zsolnay (ed.). *Seen Not Heard: Composition, Iconicity, and the Classifier Systems of Logosyllabic Scripts*: 205–235. Chicago: ISAC.
- Schmitt, Rüdiger. 2007. *Grammatik des klassisch-Armenischen mit sprachenvergleichenden Erläuterungen* (2nd ed.). Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Schneider, Roey. 2024. *The Semitic Sibilants*. Wiesbaden: Harrassowitz.
- Schneider, Thomas. 1992. *Asiatische Personennamen in ägyptischen Quellen des Neuen Reiches*. Göttingen: Vandenhoeck and Ruprecht.
- Segert, Stanislav. 1984. *A Basic grammar of the Ugaritic Language*. Berkeley: University of California.
- Simon, Zsolt. 2022. The Hurro-Urartian loan contacts in Armenian: A revision. *Hungarian Assyriological Review* 3: 63–89.
- Singer, Itamar. 1999. A political history of Ugarit. In: Wilfred G.E. Watson, Nicholas Wyatt (eds.). *Handbook of Ugaritic Studies*: 603–733. Leiden: Brill.
- van Soldt, Winfred. 1991. *Studies in the Akkadian of Ugarit: Dating and Grammar*. Neukirchen-Vluyn: Kevelaer.
- van Soldt, Winfred. 2003. The Use of Hurrian Names at Ugarit. *Ugarit-Forschungen* 35: 681–707.
- Speiser, Ephraim A. 1938. Notes on Hurrian Phonology. *Journal of the American Oriental Society* 58/1: 173–201.
- Speiser, Ephraim A. 1941. *Introduction to Hurrian*. New Haven: American School of Oriental Research.
- Streck, Michael P. 2006. Sibilants in the Old Babylonian Texts of Hammurapi and the Governors in Qaṭṭunān. In: Guy Deutscher, N. J. C. Kouwenberg (eds.). *The Akkadian Language in its Semitic Context: Studies in the Akkadian of the Third and Second Millennium BC*: 215–251. Leiden: NINO.
- Tropper, Josef, Juan-Pablo Vita. 2004. Ugaritisch *il bt* und hurritisch *in tl(n)*. *Ugarit-Forschungen* 35: 673–680.
- Tropper, Josef. 2012. *Ugaritische Grammatik* (2nd ed.). Münster: Ugarit-Verlag.
- Ungnad, Arthur. 1924. Das hurritische Fragment des Gilgamesh-Epos. *Zeitschrift für Assyriologie* 35: 133–140.
- Valério, Miguel, Ilya Yakubovich. 2010. Semitic Word for Iron as Anatolian Loanword. In: Tatiana M. Nikolaeva (ed.). *Studies in Linguistics and Semiotics: A Collection of Articles for the Anniversary for Vyacheslav V. Ivanov*: 108–116. Moscow: Languages of Slavonic Culture.
- Watson, Wilfred G.E. 2016. Ugaritic Onomastics (9). *Aula Orientalis* 34/2: 335–358.
- Weeden, Mark. 2019. Remarks on Syllabaries at Alalah VII and IV: Arguments for an Archival Approach to the Study of Cuneiform Writing. In: Jörg Klinger, Sebastian Fischer (eds.). *Keilschriftliche Syllabare. Zur Methodik ihrer Erstellung*: 129–153. Berlin: PeWe-Verlag.
- Wegner, Ilse. 2007. *Hurritisch: Eine Einführung* (2nd ed.). Wiesbaden: Harrassowitz.
- Wikander, Ola. 2015. Emphatics, Sibilants and Interdentals in Hebrew and Ugaritic: an Interlocking Model. *Ugarit-Forschungen* 46: 373–397.
- Wikander, Ola. 2022. The Horse that said W: Northwest Semitic *ss(w)*, an Indo-European (Luwian) plural, (ṣ), and Poetic Motifs in Ugarit and the Hebrew Bible. *Ugarit-Forschungen* 53: 423–459.
- Wilhelm, Gernot. 1988. Gedanken zur Frühgeschichte der Hurriter und zum hurritisch-urartäischen Sprachvergleich. In: Volkert Haas (ed.). *Hurriter und Hurritisch. Vol. II*: 43–67. Konstanz: Universitätsverlag.
- Wilhelm, Gernot. 1998. Die Inschrift des Tišatal von Urkeš. In: Giorgio Buccelatti, Marilyn Kelly-Buccelatti (eds.). *Urkes and the Hurrians: Studies in Honour of Lloyd Cotser*: 117–143. Malibu, CA: Undena.
- Wilhelm, Gernot. 2004. Hurrian. In: Roger Woodard (ed.). *The Cambridge Encyclopedia of the World's Ancient Languages*: 95–118. Cambridge University Press.

- Wiseman, Donald J. 1953. *The Alalakh Tablets*. London: British Institute of Archaeology at Ankara.
- Yakubovich, Ilya. 2010. Review of Hoffner, Harry A Jr. and H. Craig Melchert, *A Grammar of the Hittite Language*, 1st ed. [Winona Lake, IN: Eisenbrauns, 2008]. *Bibliotheca Orientalis* 57: 148–154.
- Yakubovich, Ilya. 2015. The Luwian Language. In: *Oxford Handbooks Online*. Available at: <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199935345.001.0001/oxfordhb-9780199935345-e-18>
- Yakubovich, Ilya. 2020. Hittite. In: Rebecca Hasselbach-Andee (ed.). *A Companion to Ancient Near Eastern Languages*: 221–237. Hoboken, NJ: Wiley Blackwell.
- Yakubovich, Ilya. 2022a. Peoples and Languages. In: Stefano de Martino (ed.). *Handbook of Hittite Empire: Power Structures*: 3–43. Berlin: de Gruyter.
- Yakubovich, Ilya. 2022b. Review of Patri 2019. *Journal of American Oriental Society* 142/1: 239–242.

В. Л. Цуканова, И. С. Якубович. К фонетической интерпретации хурритских и угаритских сибилантов

Фонетическая интерпретация свистящих и шипящих фонем в хурритском и угаритском языках должна быть переосмыслена в свете новых интерпретаций сибилантов в ряде соседних языков, которые, в свою очередь, отражают развитие аффрикативной гипотезы. Мы утверждаем, что слоговая клинописная орфография митаннийского и богазкёйского хурритского использует серию Š для передачи альвеолярных фрикативных фонем /s/ и /z/, тогда как серии Z и S зарезервированы для передачи аффрикат. Кроме того, по нашему мнению, угаритская графема (ṭ) используется не для передачи межзубного /θ/, как это предполагалось в традиционной угаритологии, а для альвеолярного /s/. Предлагаемые фонетические значения хурритских сибилантов способствуют упрощению сценария адаптации месопотамской клинописи для письма по-хеттски.

Ключевые слова: хурритский язык; угаритский язык; сибиланты; аффрикативная гипотеза.

Issue 2 / Часть 2

Articles / Статьи

A phonological reconstruction of Proto-Dâw–Hup

In this paper, I examine the sound correspondences between Dâw, Yuhup, and Hup, which constitute a clade within the Naduhup language family of South America, and propose a phonological reconstruction of the ancestral language of this clade. The most common shape of morphemes in Proto-Dâw–Hup is */CVC/, but some */CV/-shaped morphemes are reconstructed as well, in addition to a few sesquisyllabic and disyllabic morphemes. I reconstruct 22 consonants for Proto-Dâw–Hup, including voiceless stops, glottalic stops, plain and glottalized glides, fricatives, nasals, and voiced stops. As for the vowels, I identify 45 nuclei, with contrastive length and nasality; some of these nuclei may have been diphthongs rather than monophthongs. Long vowels in Proto-Dâw–Hup presented a tonal distinction between rising and falling tones, though the opposition in question is not relevant with most voiceless codas.

Keywords: Naduhup languages; comparative method; phonological reconstruction.

This paper is concerned with the phonological reconstruction of Proto-Dâw–Hup, the ancestral language of Dâw, Yuhup, and Hup. These languages are spoken in northwestern Amazonia, in the Brazilian state of Amazonas and in the Colombian departments of Vaupés and Amazonas. They are universally recognized to constitute a clade within the Naduhup language family to the exclusion of a fourth Naduhup language, Nadëb (Martins 2005: 326; Epps & Bolaños 2017: 481). Yuhup and Hup likewise constitute a clade to the exclusion of Dâw. It has been hypothesized that the Naduhup languages are further related to the Kakua–Nikak language family and to the Puinave language, all spoken in Colombia (Martins 2005: 1), though this putative relationship has been contested (Epps & Bolaños 2017). Earlier labels for the Naduhup language family, now dispreferred, include “Eastern Maku” (derived from a local pejorative term for peoples who traditionally live in forest areas rather than along big rivers) and “Nadahup(an)” (a portmanteau that has been criticized by the speakers of these languages because of its associations with the Portuguese and Spanish word <nada> ‘nothing’).

Comparative work on the Naduhup languages is scarce. Martins (2005) is a pioneering attempt at the reconstruction of Proto-Naduhup phonology and lexicon, but it has been argued to present

“[...] methodological deviations from the Comparative Method; for example, Martins reconstructs words to “Proto-Eastern-Maku” that are attested only in closely related languages or even dialects of a single language (rather than across primary branches), including loanwords; proposes various uneconomical forms; and does not take into account the relevance of subgrouping and language contact in evaluating processes of change” (Epps & Bolaños 2017: 475, fn. 11).

Epps & Bolaños (2017) also include several dozen reconstructed Proto-Naduhup forms, with the proviso that their preliminary reconstructions “must be understood as extremely tentative and subject to revision in future work”. To the best of my knowledge, no published works deal specifically with the historical phonology of the Dâw–Hup branch (Barboza’s 2016 contribution is unpublished).

This paper is structured as follows. 1 presents some basic facts about Dâw, Yuhup, and Hup. 2 deals with the reconstruction of Proto-Dâw–Hup (PDH) root structure, where a dis-

inction is introduced between monosyllabic, sesquisyllabic, and disyllabic roots. In 3, a proposal is made regarding the reconstruction of PDH consonants, whereas 4 considers the vowels. 5 is dedicated to the reconstruction of the tonal oppositions and the vowel length. In 6, I discuss some additional etymologies. 7 concludes the paper.

1. Dâw, Yuhup, and Hup

All three languages considered in this paper have similar typological profiles. For the purposes of this paper, it suffices to mention their preference for CVC-shaped morphemes (see 2); presence of contrastive glottalization in consonants (see 3.2, 3.4); presence of contour oral-nasal segments after oral vowels in the coda position (see 3.6); rectangular vowel inventories with three contrastive heights, with a series of front unrounded vowels, back rounded vowels, and back/central unrounded vowels (see 4); lack of mid (close-mid) nasal vowels despite the existence of high and low (low-mid) nasal vowels (see 4); existence of tones (see 5). The languages in question are also typologically unusual for making more phonological distinctions among consonants in the morpheme-final position than morpheme-initially.

The **Dâw** [dɤw] language (Glottocode [daww1239], ISO 639-3 [kwa]) is spoken by the homonymous people. The Dâw people live in the Waruá community, situated on the right margin of the Rio Negro, just opposite the town of São Gabriel da Cachoeira (Amazonas state, Brazil). As of 2022, there were 142 speakers of the language (see Obert & Santos 2022 for details on the history and the sociolinguistic situation of the Dâw people). The oldest speakers of Dâw still have some proficiency in the Rio Negro dialect of Nheengatu (a Tupian *lingua franca* of the Upper Rio Negro region; Finbow 2020), but younger speakers rather speak Brazilian Portuguese as their second language. Martins (2004) analyzes the consonants of Dâw as /p b m m' w w' t d n n' l l' c ɟ n n' j j' k g x ŋ ʔ h/. Orthographically, these are spelt <p b m m' w w' t d n n' l l' ɟ j nh nh' y y' k g x ng' r>, respectively. Barboza (2017) adds /ŋ'/ <ng'>. The stops /c k/ are articulated as ejective stops in the onset position. Glottalized sonorants are preglottalized in the onset position, and postglottalized in the coda position. Nasal codas are preoralized when they follow oral vowels. The vowel system has three contrastive heights: /i ī e ē u ũ ɤ a ã u ũ o ɔ ɔ̃/ (orthographically <i ī ê e ē u ũ â a ã u ũ ô o õ>, respectively). Note that the nasal series includes only high and low vowels, but not mid vowels. There is a triple opposition between short vowels, long vowels with rising tone, and long vowels with falling tone, though only syllables with voiced codas allow for all three possibilities (voiceless codas are incompatible with the falling tone, and open syllable obligatorily carry falling tone). Orthographically, tones are not represented, whereas long vowels are spelt as double vowels without repeating the diacritics (e.g., <ũũ> *ũũ*, <êê> *e*).

The **Yuhup** [jùhúp] language (Glottocode [yuhu1238], ISO 639-3 [yab]) is spoken by the Yuhupdeh people. Their traditional territory spans from the middle and lower course of Tiquié River (Amazonas state, Brazil) to the lower course of the Apaporis River (Vaupés and Amazonas departments, Colombia). Their ethnic population was estimated to total 754 individuals in Brazil, as of 2010, and ca. 250 individuals in Colombia, as of 2000 (Silva & Silva 2012: 53). Many Yuhupdeh speak Tukano (Tukanoan) as a second language. A number of phonological analyses have been proposed for Yuhup. Here I largely follow Silva & Silva's (2012) analysis. The consonants are /p m m' w w' t n n' c c' ɟ n n' j j' k k' ŋ ŋ' ʔ h/, represented orthographically as <p m/b m'/b w w'/ t/r n/d/r n'/d/r s s ɟ j j y y'/ k k g g' ' h>. Of these, /ɟ n n' ŋ ŋ'/ are only used morpheme-finally. Note that the underlying nasals have fully oral allophones preceding oral vowels (of these, [b] and [d] have dedicated graphemes in the orthography), and preoral-

ized allophones following oral vowels (for these, the “nasal” graphemes <m n> are used). Glottalized segments are spelt with an apostrophe morpheme-finally, but morpheme-initially the grapheme representing the vowel is duplicated instead, since the [+constricted glottis] feature surfaces as a creaky voice on a rearticulated vowel, as in <koop> /k'ɔp/ (phonetically [kɔɔp]). The grapheme <r> represents flapped allophones of denti-alveolars. The vowel system has three contrastive heights: /i ɪ e æ ẽ i ĩ ə ɑ ã u ũ o ɔ ɔ̃/ (orthographically <i ɪ ẽ e ẽ i ĩ ä a ã u ũ ö o õ>, respectively). Note that the nasal series includes only high and low vowels, but not mid vowels. The tilde is not used when a nasal vowel follows <m> or <n>, as in <móh> /môh/; the sequences /m(̃)V/ and /n(̃)V/ are distinguished in spelling by using the graphemes <b d>, as in <bón> /môn/ (phonetically [bôdn]). There are four tones, of which two (rising and falling) are used in lexical morphemes; the falling tone is represented orthographically by means of an acute accent.

Finally, the **Hup** [húp] language (Glottocode [hupd1244], ISO 639-3 [jup]) is spoken by the Hupd'äh people in an area located between the Papuri and Tiquié Rivers, reaching the Vaupés River in the east. Epps & Obert (2022) estimate the speaker population at ca. 2500. Almost all Hupd'äh are bilingual in Hup and Tukano (Tukanoan). Epps (2005: 40) analyzes the consonants of Hup as /p b b' (p') w w' t d d' c ɟ ɟ' ɟ j j' k g g' ʔ h/ (orthographically <p m/b m'/b' p' w w' t/r n/d/r n'/d'/r' s j s'/j' ɟ y y' k g k'/g' ' h>), of which /ɟ g ɟ/ occur morpheme-finally only. /p'/ is found in only one loanword, and is restricted to the Tat-Dëh and Barreira dialects. The palatal series are articulated as denti-alveolar (/c ɟ ɟ' j j'/) or glottal (/ɟ/) with a palatal offglide and/or onglide. The nasals [m n ɲ ñ] are analyzed by Epps (2005: 52–53) as allophones of the voiced obstruents /b d ɟ g/ that occur in nasal morphemes (the former two allophones have dedicated graphemes, <m n>). In oral morphemes, Epps' /b d ɟ g/ surface as postnasalized morpheme-finally ([bm dn jdn gn]), and as prenasalized morpheme-initially ([mb nd]; /ɟ g/ do not occur in that position); these are spelt using the “oral” graphemes (<b d>). As for the glottalized consonants /b' d' ɟ' g'/, these surface morpheme-initially as [mb nd (t)ɟ k] in oral morphemes (<b' d' s' k'>), and as [m n (t)ɟ k] in nasal morphemes (<m' n' s' k'>), with laryngealization on the nucleus. Morpheme-finally, they surface as [p' t' jt' k'] in oral morphemes (<b' d' j' g'>), and as [mp' nt' jn' ñ] in nasal morphemes (<m' n' j' g'>). The graphemes <r r'> represent flapped allophones of denti-alveolars. The vowel system and its orthographic representation are identical to those of Yuhup, with three provisos. First, Epps (2005) does not posit nasal vowels as segments, but rather attributes their occurrence to a suprasegmental [+nasal] feature which is anchored to entire syllables or even morphemes (as also proposed for Yuhup by authors such as Ospina Bozzi 2002). Second, the low vowels of Hup are typically described as central rather than back (hence /a ã/ rather than /ɑ ã/ in Silva & Silva's treatment of Yuhup). Third, the tilde in nasal vowels is omitted in the orthography not only after <m n>, but also in morphemes that include the graphemes <m n> in **any** position, as in <pán> [pân] (compare Yuhup <pân> [pân]). There are two contrastive tones (rising and falling/high), which are only distinguished in stressed syllables; the falling/high tone has a high allotone preceding voiceless codas, and a falling allotone elsewhere. The rising tone is represented orthographically by means of a grave accent, and the falling tone is represented with an acute accent.

Linguistic data is given in two representations in this paper. The orthographic one, enclosed in chevrons (<>), is cited after Epps *et al.* (2018) for Dâw, Silva & Silva (2012) for Yuhup, and Ramirez (2005) for Hup. In italics, I give the respective broad transcriptions in the International Phonetic Alphabet. In these transcriptions, I represent major allophony patterns, such as preoralization of nasal codas and the allotones of the falling/high tone in Hup, but omit details such as the non-audible release diacritic. I do not use the symbols [ɑ] or [ã] for Yuhup and replace them with [a], [ã], since the distinction is questionable (authors such as Ospina Bozzi 2002 use [a], [ã]), and because the italic versions of the respective glyphs are difficult to distinguish.

2. Root structure

The Dâw–Hup languages have a strong preference for monosyllabic roots (Martins 2004: 13–14; Epps 2005: 36, 69), though there are some disyllabic roots in all three languages, typically with identical vowels in both syllables (Martins 2004: 62, 70–71; Silva & Silva 2012: 85; Epps 2005: 70–71). Longer roots are vanishingly rare (see Epps 2005: 70 for some examples of onomatopoeic words in Hup). CVC is the most common root type in the Dâw–Hup languages; CV is less frequent; CVCVC and CVCV are accordingly even less frequent. Roots must start with a consonant in all Dâw–Hup languages, but onsetless suffixes of the shape VC are allowed.

For the most part, cognate sets involve matching root structures in Dâw, Yuhup, and Hup, as shown below for CVC (1), CV (2), and CVCVC (3) roots.

- (1) PDH **pû:g* ‘porcupine’ > Dâw ⟨puug⟩ *pû:g* || Yuhup ⟨púg⟩ *pûgɨ* || Hup ⟨pùg⟩ *pûgɨ*
- (2) PDH **c’ô:* ‘flower’ > Dâw ⟨çoo⟩ *c’ô:* || Yuhup ⟨soo⟩ *tʃǒ* || Hup ⟨s’ó⟩ *tʃǒ:*
- (3) PDH **cəwəʃ:h* ‘thrush’ > Dâw ⟨sâwâar⟩ *ʃəwəʃ:h* || Yuhup ⟨sâwäh⟩ *ʃəwəh*¹ || Hup ⟨sâwäh⟩ *ʃəwəh*

However, in a number of etymologies Hup shows an additional word-initial syllable without any correspondence in Dâw or Yuhup. Two diachronic interpretations are conceivable. One could claim that Hup has historically fossilized erstwhile prefixes (or first elements of compounds), as contended by Epps (2005: 94–95) regarding the putative semantically opaque morpheme /cV-/. Alternatively, one could argue that Hup preserves an element lost in Dâw and Yuhup. The latter possibility is supported by comparative evidence from Nadëb, where one finds ⟨sawim⟩ *ʃawi:m* ‘abacaba palm’ (Barbosa 2005: 25), ⟨jaw’ääj⟩ *jawɔ:j* ‘capuchin monkey’, ⟨jawyk⟩ *jawik* ‘heavy’, ⟨majyyw⟩ *mājɨ:w* ‘blood’ (Martins 2005: 234, 279, 301). Since the quality of the word-initial consonant in Hup matches that of Nadëb, I conclude that Hup is in fact conservative in preserving a sesquisyllabic element lost elsewhere. In my reconstructions, given in 4–14, I use a period to separate the sesquisyllable from the stressed syllable, and do not specify a vowel that might have occurred in the sesquisyllables.²

- (4) PDH **c.ḅă:k* ‘blowgun’ > Dâw ⟨baak⟩ *bă:k* || Yuhup ⟨baák⟩ *băk* || Hup ⟨sab’āk⟩ *tʃabăk*
- (5) PDH **c.but* ‘electric eel’ > Dâw ⟨but⟩ *but* || Yuhup ⟨buut⟩ *büt* || Hup ⟨sub’út⟩ *tʃubút*
- (6) PDH **c.wĩ:m* ‘abacaba fruit (*Oenocarpus bacaba*)’ > Dâw ⟨wiim⟩ *wĩ:bm* || Yuhup ⟨wím⟩ *wĩbm* || Hup ⟨siwib⟩ *tʃiwĩbm*
- (7) PDH **c.wũ:k* ‘kapok cotton’ > Dâw ⟨wuuk⟩ *wũ:k* || Yuhup ⟨wúk⟩ *wûk* || Hup ⟨suwùk⟩ *tʃuwũk*
- (8) PDH **j.wă:ç* ‘capuchin monkey’ > Dâw ⟨waas⟩ *wă:f* || Yuhup ⟨wác⟩ *wâç* || Hup ⟨yawàç⟩ *djawăjh*
- (9) PDH **j.wak* ‘japurá fruit’ > Dâw ⟨wak⟩ *wak* || Yuhup ⟨wak⟩ *wăk* || Hup ⟨yawák⟩ *djawák*
- (10) PDH **j.wi₂:k* ~ **j.wi:k* ~ **j.wi₂:k* ‘heavy’ > Yuhup ⟨wik⟩ *wĩk* || Hup ⟨yiwík⟩ *djiwík*
- (11) PDH **k.jă:k* ‘manioc’ > Dâw ⟨yaak⟩ *jă:k* || Yuhup ⟨yak-⟩ *jak-* (in compounds) || Hup ⟨kayàk⟩ *kajăk*³
- (12) PDH **k.wǝ:ʔ* ‘to squeeze’ > Dâw ⟨’wôo’⟩ *ʔwǝ:ʔ* || Hup ⟨köw’ó’⟩ *korwǝʔ*
- (13) PDH **m.jĩ₂:w* ‘blood’ > Dâw ⟨yuuw⟩ *jũ:w* || Yuhup ⟨yíw⟩ *jĩw* || Hup ⟨biyìw⟩ *mbijĩw*⁴

¹ The expected falling tone is not attested in this form by Silva & Silva (2012).

² In Hup, the vowel of the sesquisyllable is usually the same as that of the main syllable (Epps 2005: 70), though in the Umari Norte dialect area the vowel [i] is sometimes found instead (Epps 2005: 88), as in *kijăk* ‘manioc’.

³ The form ⟨kehek-⟩ *kəhək-* is documented along the Vaupés and Japu Rivers (Epps 2005: 88).

⁴ The form ⟨bihiw⟩ *mbihĩw* is documented along the Vaupés and Japu Rivers (Epps 2005: 88).

- (14) PDH **m.jɔʔ* ‘spider’ > Dâw <yo> *jɔʔ* || Yuhup <yo> *jɔʔ* || Hup <boyó> *mbɔjɔʔ*

There are two situations when Yuhup patterns with Hup in preserving the sesquisyllable, however. This is the case in reduplicated forms (15–19), as well as in forms whose main syllable has a fricative onset (PDH **h* or **x* > Yuhup and Hup *h*; 20–28). Note that PDH **p* and **m* often yield Yuhup *w* in sesquisyllables (21, 27), though not always (23–26); the reason for this is unclear.

- (15) PDH **k.kɔ̃j* ‘crooked’ > Dâw <xôy> *xɔ̃jʔ* || Hup <kôkôy> *kɔ̃kɔ̃jʔ*
 (16) PDH **n.nă:p* ‘cockroach’ > Dâw <naap> *nă:p* || Yuhup <dadáp> *dadâp* || Hup <daràp> *ndarâp*
 (17) PDH **n.nud* ‘tadpole’ > Dâw <nud> *nud* || Yuhup <dudun> *dudũdn* || Hup <durùd> *ndurũdn*⁵
 (18) PDH **n.nũt* ‘moth’ > Dâw <nũt> *nũt* || Yuhup <nunut> *nũnũt* || Hup <nunút> *nũnũt* ‘sphinx moth’
 (19) PDH **x.xũ:j* ‘firefly’ > Dâw <xuuy> *xũ:j* || Yuhup <huhúy> *huhûj* ‘cicada’ || Hup <huhùy> *huhũj*
 (20) PDH **j.hũm* ‘avocado’ > Dâw <rũm> *hũm* || Yuhup <yũhum> *jũhũm* || Hup <yuhúm> *njũhũm*
 (21) PDH **m.hố:t* ‘wind’ > Dâw <rôot> *hố:t* || Yuhup <wöhót> *wohôt* || Hup <böhòt> *mbohót*
 (22) PDH **m.huh* ‘Amazon grape (*Pourouma cecropiaefolia*)’⁶ > Dâw <rur> *huh* || Yuhup <puhu> *puhũ* || Hup <buhúh> *mbuhúh*
 (23) PDH **m.hũʔ* ‘to play’ > Dâw <rũ> *hũʔ* || Yuhup <muhu> *mũhũʔ* || Hup <muhú> *mũhũʔ*
 (24) PDH **p.hãj*⁷ ‘sorva fruit (*Couma guianensis*)’ > Dâw <rãy> *hãj* || Yuhup <pāhay> *pāhãj* ‘fruit sp.’ || Hup <pāháy> *pāhãj*
 (25) PDH **p.hũ:t* ‘to blow’ > Dâw <rũut> *hũ:t* || Yuhup <pũhut> *pũhũt* || Hup <pũhút> *pũhũt*
 (26) PDH **p.xô:k* ‘gray hair’ > Dâw <xook> *xô:k* || Yuhup <pohok> *pchók* || Hup <pohòk> *pchók*
 (27) PDH **p.xot* ‘aracu fish’ > Dâw <xot> *xot* || Yuhup <wohot> *wchót* || Hup <pohót> *pchót*
 (28) PDH **w.hən* ‘old man’⁸ > Dâw <rân> *hɛdn* || Hup <wähád> *wəhâdn*

I am aware of two etymologies, shown in 29–30, that suggest the reconstruction of the sequence **j.ʔ*. However, one of these reconstructions must be wrong, since Dâw shows different onsets in each cognate set: *ʔ* and *ʔj*, respectively.

- (29) (?) PDH **j.ʔaw* ‘to chew’ > Dâw <aw> *ʔawʔ* || Hup <ya’áw> *djaʔáwʔ*
 (30) (?) PDH **j.ʔām* ‘jaguar’ > Dâw <yām> *ʔām* ‘dog’ || Yuhup <yāam> *jāǎm* || Hup <ya’ám> *njāǎm*

The diachronic loss of initial syllables, which may have had sesquisyllables as an intermediate stage, has resulted in synchronic alternations in Dâw, such as those seen in <ār> *ʔāh* ‘I’ → <rā> *h-āʔ* ‘I (focused form)’, <tuk> *tuk* ‘to want’ → <kēr> *k-ēh* (also <tukēr> *tuk-ēh*) ‘not to want’, <rām> *hām* ‘to go’ → <mōr> *m-ōh* ‘go!’ (Martins 2004: 106–111).

⁵ The rising tone, attested for Hup in Ramirez (2005: 67) and Epps (2005: 38), is unexpected.

⁶ The Yuhup form, if related, is irregular. The expected reflex in Yuhup would rather be *<wuhuh> or *<buhuh>.

⁷ This is an opaque derivative of a Proto-Naduhup term preserved in Nadëb as <pah> *pah* ‘sorva fruit’ (Weir 1984: 173).

⁸ This is an opaque derivative of PDH **wă:h* ‘old’ > Dâw <waar> *wă:h*, Yuhup <wáh> *wâh*.

3. Consonants

I reconstruct 22 consonants for PDH, listed in Table 1.

	labial	denti-alveolar	palatal	velar	glottal
voiceless stops	*p	*t	*c	*k	*ʔ
voiced stops	*b	*d	*j	*g	
glottalic stops	*ɸ	*ɖ	*c'	*k'	
nasals	*m	*n			
fricatives			*ç	*x	*h
glides	*w		*j		
glottalized glides	*w̥		*j̥		

Table 1. PDH consonantal inventory

3.1. Voiceless stops

Five voiceless obstruents are reconstructed for PDH: *p, *t, *c, *k, and *ʔ. They are largely preserved in Yuhup and Hup, except that *t may undergo intervocalic flapping in Yuhup and, dialectally, in Hup (Lopes 1995: 73–75; Ospina Bozzi 2002: 111–112; Epps 2005: 44–45); for the development *p > Yuhup w in sesquisyllables, see section 2. Even though the consonant /c/ is preserved as a phonemic unit in Yuhup and Hup, in both languages it shows a remarkable difference between the allophones that occur in onsets and codas, which are represented in my broad transcriptions in this paper. As for Dâw, the voiceless stops of PDH are preserved morpheme-finally, whereas morpheme-initially *c and *k are lenited to ʃ and x, respectively.⁹ This is summarized in Table 2.

PDH	Dâw	Yuhup	Hup
*p	⟨p⟩ p	⟨p⟩ p, ⟨w⟩ w ^A	⟨p⟩ p
*t	⟨t⟩ t	⟨t⟩ t, ⟨r⟩ r ^B	⟨t⟩ t, ⟨r⟩ r ^C
*c	⟨s⟩ ʃ, ⟨ç⟩ c ^D	⟨s⟩ tʃ, c ^{D10}	⟨s⟩ tʃ, j ^{tD}
*k	⟨x⟩ x, ⟨k⟩ k ^D	⟨k⟩ k	⟨k⟩ k
*ʔ	⟨ʔ⟩ ʔ ^E	⟨ʔ⟩ ʔ ^E	⟨ʔ⟩ ʔ ^E

^A = in some sesquisyllables; ^B = between vowels, possibly with an intervening glottal; ^C = between vowels in the Tat-Dëh and Umari Norte dialect areas; ^D = morpheme-finally; ^E = not represented orthographically in the word-initial position

Table 2. PDH voiceless stops and their reflexes

⁹ Since most Naduhup morphemes are monosyllabic, the morpheme-initial and morpheme-final positions typically coincide with the onset and coda positions in uninflected words, but this does not hold true when one considers words with vowel-initial suffixes. Notoriously, allophony in Naduhup is largely conditioned by the position of a segment within a morpheme rather than within a syllable (see Epps 2005: 41 for Hup). In this paper, I do not discuss allophony patterns found at morpheme boundaries.

¹⁰ The phonetic realizations of this phoneme in the onset position have been described as a palatal stop [c] (Ospina Bozzi 2002: 108), a postalveolar stop or fricative [ʃ] ~ [tʃ] (Silva & Silva 2012: 88), an alveopalatal affricate [tɕ] (Lopes 1995: 10), or an alveolar affricate [ts] varying with the fricatives [s], [ʃ], and [ʂ] (Fernandes 2017: 69). In this paper I use tʃ as an invariant representation. Its realizations in the coda position have been documented as [c'] (Silva & Silva 2012: 88) or [jt] (Lopes 1995); I use c as the invariant representation in this paper.

Some examples follow in 31–51.

- (31) PDH **pă:ç* ‘stone’ > Dâw <paas> *pă:f* || Yuhup <pác> *pâç* || Hup <pàç> *păjh*
 (32) PDH **pê:z* ‘to go upriver’ > Dâw <pee> *pê:* || Yuhup <pe> *pě* || Hup <pé> *pê:*
 (33) PDH **pûp* ‘insect sp.’ > Dâw <pûp> *pûp* || Hup <púp> *púp* ‘a kind of tick found in hammocks’
 (34) PDH **hõ:p* ‘to dive’ > Dâw <rôop> *hõ:p* || Yuhup <höp> *hõp* || Hup <hõp> *hóp*
 (35) PDH **tê:w* ‘club’ > Dâw <toow> *tê:w* || Yuhup <tów> *têw* || Hup <tòw> *tõw*
 (36) PDH **tăw* ‘to hit with a stick’ > Dâw <tāaw> *tă:w* ‘to beat to extract liquid’ || Yuhup <tāw> *tăw* || Hup <tăw> *tăw*
 (37) PDH **xă:t* ‘alligator’ > Dâw <xeet> *xě:t* || Yuhup <hát> *hât* || Hup <hàt> *hăt*
 (38) PDH **mă:t* ‘agouti’ > Dâw <mēet> *mě:t* || Yuhup <mét> *mât* || Hup <mèt> *măt*
 (39) PDH **că:k* ‘to climb’ > Dâw <saak> *śă:k* || Yuhup <sak> *śāk* || Hup <sák> *śāk*
 (40) PDH **cô:b* ‘finger’ > Dâw <sôob> *śô:b* ‘hand’ || Yuhup <söm> *śõbm* ‘to point with finger’ || Hup <sòb> *śõbm* ‘finger’, <sób> *śõbm* ‘to point with finger’
 (41) PDH **cũ:h* ‘non-venomous spider sp.’ > Dâw <sũuh> *śũ:h* || Yuhup <sũh> *śũh* || Hup <sũh> *śũh*
 (42) PDH **dũ:c* ‘parrot sp.’ > Dâw <duuç> *dũ:c* || Yuhup <duús> *duúc* ‘orange-cheeked parrot’ || Hup <d’ùs> *ndũjt* ‘blue-headed parrot’
 (43) PDH **ke₁c* ‘to rip or peel with one’s teeth’ > Dâw <xêç> *xec* || Yuhup <käs> *kăc* || Hup <kăs> *kăjt*
 (44) PDH **kub* ‘hungry’ > Dâw <xub> *xub* || Hup <kúb> *kûbm*
 (45) PDH **kaj* ‘to hug’ > Dâw <xay> *xaj*[?] || Yuhup <kay> *kăj*[?] || Hup <káy> *kâj*[?] ‘to hug with one arm’
 (46) PDH **tuk* ‘to want’ > Dâw <tuk> *tuk* || Yuhup <tuk> *túk* || Hup <túk> *túk*
 (47) PDH **hõk* ‘to saw’ > Dâw <rõk> *hõk* || Yuhup <hõk> *hõk* || Hup <hõk> *hõk*
 (48) PDH **?ũ:m* ‘to hit’ > Dâw <ũum> *?ũ:m* ‘to hit, to shoot’ || Yuhup <ũm> *?ũm* ‘to hit, to kill, to hunt’
 (49) PDH **?əg* ‘to drink’, **?ǵ₁g* ‘caxiri beverage’ > Dâw <âg> *?əg*, <âag> *?ǵ:g* || Yuhup <äg> *?ǵgŋ*, <ág> *?ǵgŋ* || Hup <ág> *?ǵgŋ*, <äg> *?ǵgŋ*
 (50) PDH **bu?* ‘termite’ > Dâw <bu> *bu?* || Yuhup <buu> *buũ?* || Hup <b’ú> *mbû?*
 (51) PDH **c’ăz₂?* ‘ray (fish)’ > Dâw <çee> *c’ě:?* || Yuhup <see> *tfææ?* || Hup <s’è> *tfæ?*

In Yuhup, some instances of morpheme-final stops have been documented as voiced by Ospina Bozzi (2002: 109), as in [hěb] ‘sweep!’, [úd] ‘thorn’, [mbõj] ‘dragonfly’, [wǵg] ‘leafcutter ant sp.’. Ospina Bozzi (2002) posits an opposition between /p t k/ and /b d g/ in that environment (however, /j/ is not reported to contrast with /c/). Martins (2005: 80–81, 83) likewise documents these segments as voiced, as in [?úd] ‘thorn’, [wǵg] ‘leafcutter ant sp.’, and posits an opposition between /p t c k/ and /b³ d³ c³ k³/.¹¹ Other sources on Yuhup, however, do not report voiced stops in the morpheme-final position (e.g., [?u:t] ‘thorn’ in Lopes 1995: 102), and Ospina Bozzi’s voiced stops generally correspond to voiceless stops in Dâw and Hup in that environment. Given the very limited evidence for the opposition in question in Yuhup, in this paper I follow Silva & Silva (2012) in representing such consonants as voiceless, and accordingly reconstruct voiceless stops in PDH, as shown in 52–54.

- (52) PDH **xăz₂p* ‘to sweep’ > Dâw <xēep> *xě:p* || Yuhup <hep> *hěp* || Hup <hép> *hěp*

¹¹ Martins’ (2005) /c/ corresponds to Lopes’ (1995) and Ospina Bozzi’s (2002) /jh/ and Silva & Silva’s (2002) and Fernandes’ (2017) /ç/.

(53) PDH *ʔú:t ‘thorn’ > Dâw <uut> ʔú:t || Yuhup <út> ʔút || Hup <ùt> ʔú:t

(54) PDH *wək ‘leafcutter ant sp.’ > Dâw <wâk> wək || Yuhup <wäk> wək || Hup <wák> wák

Given that Ospina Bozzi (2002) has worked only with speakers from the lower course of the Apaporis River (Colombia), whereas other works on Yuhup phonology are based on research carried out with speakers from Brazil, it is conceivable that the difference in voicing may turn out to be restricted to the southwestern fringe of the Yuhup-speaking area. Alternatively, it could be a mistranscription on Ospina Bozzi’s (2002) part. The status and origins of the putative voicing opposition in the morpheme-final position in Yuhup remain to be established.

A group of cognate sets shows a discrepancy between the presence of a final glottal stop in Dâw and its absence in Hup, or vice versa. These etymologies are also suspect for other reasons. In 55–56, Yuhup does show the expected glottal stop, but 55 fails to show laryngealization on the vowel; moreover, 56 is likely a Tukanoan loan (Ye’pâ-masa <yoó> ‘to carry in one’s hand or claws’; Ramirez 2019 [1997]). 57 shows a rare vowel correspondence (see 4.2). In 58, the Dâw verb is documented with rising tone, stated to be impossible in CV-shaped morphemes in Martins (2004: 69). In 59, Dâw points to a long vowel in PDH, whereas Yuhup and Hup point to a short vowel.

(55) (?) PDH *kʰĩ:ʔ ~ *kʰĩ: ‘hot’ > Dâw <kũũ> kʰũũ:ʔ ‘to heat (liquid)’ || Yuhup <ki> kĩʔ || Hup <kʰĩ> kʰĩ:

(56) (?) PDH *jô:ʔ ~ *jô: ‘to carry in one’s hands’ > Dâw <yôo> jô:ʔ || Yuhup <yö> jöʔ || Hup <yó> djô:

(57) (?) PDH *bA~ʔ ~ *bA~ ‘cold’ > Dâw <ba> baʔ || Hup <m’é> mǣ

(58) (?) PDH *tǎ:ʔ ‘to meet’ > Dâw <taa> tǎ: || Yuhup <ta> tǎʔ ‘to block’, <mih ta> mĩh-tǎʔ ‘to meet’ || Hup <tá> tǎʔ ‘to block’, <hitá> hi-tǎʔ ‘to meet’

(59) (?) PDH *nô: ~ *noʔ ‘child’ > Dâw <nô> nô: || Yuhup <dö> döʔ || Hup <dó> ndóʔ

3.2. Glottalic stops

PDH also had a series of glottalic stops, which I reconstruct as *b̥, *d̥, *c̥, and *k̥. The representation of the labial and denti-alveolar obstruents as implosives and of the palatal and velar ones as ejectives conforms to one’s typological expectations (Javkin 1977), and accounts neatly for the reflexes in the daughter languages.

I start by considering the reflexes of the glottalic stops in the morpheme-initial position. In Dâw, *c̥ and *k̥ are preserved as ejectives (Martins 2004: 26, 29). Note that [c̥] and [k̥] are synchronically analyzed as allophones of /c/ and /k/ in Dâw, since the language lacks plain [c] and [k] in onsets; recall that PDH *c and *k are reflected as fricatives in Dâw morpheme-initially. As for the implosives, Dâw reflects them as *b* and *d* before oral vowels (thus losing glottalization), and as ^h*m* and ^h*n* before nasal vowels. In Yuhup and Hup, the reflexes of *b̥, *d̥, *c̥, and *k̥ are articulated much like those of *m, *n, *c, and *k, but the following vowel surfaces with laryngealization, or creaky voice (see Epps 2005: 55–63; Ospina Bozzi 2002: 117–118). In Yuhup, but not Hup, the laryngealized vowel is often described as rearticulated ([V̥V̥] in Silva & Silva 2012: 85; [V̥ʔV̥] in Lopes 1995: 91; [V̥ʔV̥] ~ [V̥] in Epps 2005: 68). Epps (2005: 66–67) convincingly argues, based on distributional evidence, that Hup has no underlying laryngealized vowels; instead, the [+constricted glottis] feature is associated with the consonantal segments. Epps (2005) represents the glottalic stops of Hup as /b̥ d̥ ɟ̥ g̥/ (in addition to /p̥/, found dialectally in one loanword), but acknowledges that /ɟ̥ g̥/ surface as voiceless stops in onsets. This is summarized in Table 3.

PDH	Dâw	Yuhup	Hup
*b	⟨b⟩ <i>b</i> , ⟨m⟩ ⁱ <i>m</i> ^A	⟨bVV⟩ <i>b̥</i> , ⟨mVV⟩ <i>m̥</i> ^A	⟨b'⟩ <i>mb̥</i> , ⟨m'⟩ <i>m̥</i> ^A
*d	⟨d⟩ <i>d</i> , ⟨n⟩ ⁱ <i>n</i> ^A	⟨dVV⟩ <i>d̥</i> , ⟨nVV⟩ <i>n̥</i> ^A	⟨d'⟩ <i>nd̥</i> , ⟨n'⟩ <i>n̥</i> ^A
*c'	⟨ç⟩ <i>c'</i>	⟨sVV⟩ <i>t̥</i>	⟨s'⟩ <i>t̥</i>
*k'	⟨k⟩ <i>k'</i> ¹²	⟨kVV⟩ <i>k̥</i>	⟨k'⟩ <i>k̥</i>

^A = before nasal vowels / in nasal morphemes

Table 3. PDH glottalic stops and their reflexes (morpheme-initial position)

Some examples follow in 60–71.

- (60) PDH **bux* ‘to burst’ > Dâw ⟨bux⟩ *bux* || Yuhup ⟨buuh⟩ *bũh*
 (61) PDH **bæ?* ‘hard’ > Dâw ⟨be⟩ *bẽ?* || Yuhup ⟨(ta-)baa⟩ (ta-)bã? || Hup ⟨tab'á⟩ *tapbã?*¹³
 (62) PDH **bəw* ‘tree sp.’ > Dâw ⟨bâaw⟩ *bɔ:w* || Hup ⟨b'ăw⟩ *mbǝw* ‘escorrega-macaco tree (*Peltogyne paniculata*)’
 (63) PDH **bũ?* ‘earthworm’ > Dâw ⟨mũ⟩ ⁱ*mũ?* || Yuhup ⟨mii⟩ *mĩĩ?* ‘daracubi worm’ || Hup ⟨m'i⟩ *mĩ?*
 (64) PDH **bǝ:h* ‘star’ > Dâw ⟨mēer⟩ ⁱ*mē:h* || Yuhup ⟨wero-meéh⟩ *wærɔ-mǝǝh* || Hup ⟨wero-m'èh⟩ *wærɔ-mǝh*
 (65) PDH **dɔx* ‘rotten’ > Dâw ⟨dox⟩ *dɔx* || Yuhup ⟨dooh⟩ *dɔh* || Hup ⟨d'óh⟩ *ndɔh*
 (66) PDH **dum* ‘tail’ > Dâw ⟨dum⟩ *dubm* || Yuhup ⟨duum⟩ *duũbm* || Hup ⟨d'úb⟩ *nduũbm*
 (67) PDH **dǝk* ‘cross-eyed’ or ‘monocular-visioned’ > Dâw ⟨nǝak⟩ ⁱ*nǝ:k* ‘cross-eyed’ || Yuhup ⟨naak⟩ *nǝǝk* ‘monocular-visioned’
 (68) PDH **c'â:* ‘black; bitter’¹⁴ > Dâw ⟨çaa⟩ *c'â:* || Yuhup ⟨saa⟩ *tʃaǝ* || Hup ⟨s'á⟩ *tʃâ:*
 (69) PDH **c'ɔk* ‘eastern lowland olingo’ > Dâw ⟨çõk⟩ *c'ɔk* || Hup ⟨wõhsók⟩ *wõhtʃók*
 (70) PDH **k'əw* ‘cylindrical chunk; poisonous arthropod sp.’¹⁵ > Dâw ⟨kâw⟩ *k'əw* ‘cylindrical chunk; centipede’ || Yuhup ⟨kääw⟩ *kǝǝw* ‘cylindrical chunk’ || Hup ⟨k'ăw⟩ *kǝw* ‘cylindrical chunk; scorpion’
 (71) PDH **k'ê:m* ‘to flood’ > Dâw ⟨kũum⟩ *k'ũ:m* || Hup ⟨k'im⟩ *kĩm*

Morpheme-finally, examples of glottalic stops are less abundant. In Dâw, one typically finds glottalized nasals, except that the palatal glottalic stop is reflected as *j* after oral vowels (note that no *bona fide* examples are known for **d* and **k* after oral vowels). In Yuhup and Hup, one typically finds underlying glottalized segments, which surface as stops with no audible release (voiced in Yuhup, voiceless in Hup) word-finally; when preceded by nasal vowels, they are accompanied by a nasal transition, except that in Hup the palatal and the velar ones surface as [jn], [ŋ]. In Hup, the distinction between the reflexes of voiceless (**p*, **t*, **c*, **k*) and glottalic (**b*, **d*, **c'*, **k'*) stops is in fact neutralized after oral vowels word-finally, but is visible before vowel-initial suffixes (Epps 2005: 57–58). In Yuhup, a handful of words show

¹² I have considered positing a sound change **k'* > Dâw *c'* before the vowel **i* so as to account for the sound correspondence seen in Dâw ⟨çii⟩ *c'î:* ‘sour’ ~ Yuhup ⟨kii⟩ *kĩĩ*, Hup ⟨k'í⟩ *kî:* ‘sour’. However, no palatalization is seen in PDH **k'i₂?* ‘to split’ > Dâw ⟨ki⟩ *k'î?*, Yuhup ⟨kii⟩ *kĩĩ?*. It is therefore likely that the Dâw term for ‘sour’ is not after all cognate with the Yuhup and Hup material.

¹³ See Epps (2005: 94–95) on the ambisyllabic relativization of intervocalic consonants in Hup.

¹⁴ We are probably dealing with two homophonous roots here rather than with a true instance of polysemy. The homophony is found in all daughter languages.

¹⁵ We are probably dealing with two homophonous roots here rather than with a true instance of polysemy. The homophony is found in Dâw and Hup. Yuhup ⟨k'ăw⟩ *k'ǝw* means only ‘cylindrical chunk’.

unexpected creaky voice on the vowel; this could be regular in nasal vowels before *b (as in 76 and 87), and in oral vowels before *c' (as in 79). This is summarized in Table 4.

PDH	Dâw	Yuhup	Hup
*b	⟨m'⟩ <i>m'</i> ²	⟨m'⟩ <i>b</i> , ⟨VVm'⟩ <i>mb</i> ^A	⟨b'⟩ <i>p</i> , ⟨m'⟩ <i>mp</i> ^A
*d	⟨n'⟩ <i>n'</i> ^{2A}	⟨n'⟩ <i>d</i> ^B , <i>nd</i> ^A	⟨d'⟩ <i>t</i> ^B , ⟨n'⟩ <i>nt</i> ^A
*c'	⟨j'⟩ <i>j</i> , ⟨nh'⟩ <i>j</i> ^{2A}	⟨VVj'⟩ <i>j</i> , ⟨j'⟩ <i>j</i> ^A	⟨j'⟩ <i>jt</i> , ⟨j'⟩ <i>jn</i> ^A
*k'	⟨gn'⟩ <i>ŋ</i> ^{2A}	⟨g'⟩ <i>g</i> ^B , <i>ŋg</i> ^{AB}	⟨g'⟩ <i>k</i> ^B , ⟨g'⟩ <i>ŋ</i> ^A

^A = after nasal vowels / in nasal morphemes; ^B = hypothetical reflex predicted based on systemic considerations

Table 4. PDH glottalic stops and their reflexes (morpheme-final position)

Some examples follow in 72–82.

- (72) PDH **hê₂b* ‘to fan’, **hê₂b* ‘fan’ > Dâw ⟨hêem'⟩ *hê:m'*, ⟨hêem'⟩ *hê:m'* || Yuhup ⟨hêm'⟩ *hêb*, ⟨hêm'⟩ *hêb* || Hup ⟨hêb'⟩ *hép*, ⟨hêb'⟩ *hép*
- (73) PDH **pəb* ‘mushroom’ > Dâw ⟨pâm'⟩ *pɔm'* || Yuhup ⟨päm'⟩ *păb* || Hup ⟨páb'⟩ *páp* ‘mushroom sp. (white, edible, grows on wood)’
- (74) PDH **mă:b* ‘brother, companion’¹⁶ > Dâw ⟨maam'⟩ *mă:m'* || Yuhup ⟨bám'⟩ *bâb* || Hup ⟨báb'⟩ *mbáp*¹⁷
- (75) PDH **dî₂b* ‘to close’ > Dâw ⟨dim'⟩ *dim'* || Yuhup ⟨diim'⟩ *diïb* ‘to crumple’, ⟨diim'⟩ *diidïb* ‘curly’ || Hup ⟨d'id'ib'⟩ *didíp* ‘curly’
- (76) (?) PDH **mă:b* ~ **mă:b* ‘to grip’¹⁸ > Dâw ⟨mêm'⟩ *mêm'* ‘to grip under one’s arms’ || Yuhup ⟨meem'⟩ *mămb* || Hup ⟨mém'⟩ *mêm*
- (77) PDH **că:d* ‘horn’ > Dâw ⟨sāan'⟩ *ǰă:n'* || Yuhup ⟨sân'⟩ *tǰând* || Hup ⟨sàn'⟩ *tǰânt*
- (78) PDH **wac* ‘amphibian sp.’ > Dâw ⟨waj'⟩ *waɣ* ‘toad sp.’ || Hup ⟨wáj'⟩ *wájt* ‘frog sp. (lives in holes in trees)’
- (79) PDH **cî:c* ‘mottle-faced tamarin’ > Dâw ⟨siij'⟩ *fi:ɣ* || Yuhup ⟨noh-siij'⟩ *nǰh-tǰiɣ* ‘black monkey sp.’ || (?) Hup ⟨yom'oy-siɣ'⟩ *nǰmǰj-tǰiɣ*¹⁹
- (80) PDH **ôă:c* ‘mud’ > Dâw ⟨māanh'⟩ *ǰmă:n'* || Hup ⟨m'aj'⟩ *măjn* ‘earth, clay’
- (81) PDH **tûc* ‘to shrivel’ > Dâw ⟨tūunh'⟩ *tû:n'* || Yuhup ⟨tūj'⟩ *tūnɣ*
- (82) PDH **ǰă:k* ‘throat’ > Dâw ⟨ôogn'⟩ *ǰă:ŋ'* || Hup ⟨ôg'⟩ *ǰôŋ* ‘Adam’s apple’

In a handful of examples, Dâw and Hup/Yuhup show irregular correspondences involving reflexes of glottalic stops in one language and reflexes of other consonants in other languages. Such etymologies are, therefore, suspect and may result from horizontal transmission or chance. In 83, Dâw points to a voiceless stop and Hup to a glottalic stop. In 84, Dâw points to a glottalic stop and Hup to a nasal. In 85, Dâw points to a nasal and Hup to a glottalic stop.

¹⁶ This etymon is known to be a Tukanoan loanword (compare Ye'pâ-masa ⟨ba'pa⟩ ‘companion’; Ramirez 2019 [1997]), but it must have been present already in PDH given the existence of reflexes in all daughter languages which show regular correspondences (except for the tonal discrepancy commented on in footnote 17).

¹⁷ The high tone, attested for Hup in Ramirez (2005: 43) and Epps (2005: 59), is unexpected.

¹⁸ Dâw *ê* does not regularly correspond to Yuhup and Hup *ê*, but rather to *â*; Yuhup and Hup *ê* are expected to correspond to Dâw *ê*: instead. This is why I reconstruct two variants, **mă:b* (which accounts for the Dâw reflex) and **mă:b* (which accounts for the Yuhup and Hup reflexes). Note the long vowel in the Dâw verb ⟨mêm'⟩ *mêm* ‘to carry under one’s arm’, which could be etymologically related.

¹⁹ The Hup reflex is entirely irregular. One would expect *⟨yom'oy-siɣ'⟩ *nǰmǰj-tǰiɣ*. The PDH form is reconstructed based on the reflexes in Dâw and Yuhup.

In 86, Yuhup points to a voiced stop, Hup to a glottalic stop, and Dâw is ambiguous. In 87, Dâw points to a voiced stop, whereas Yuhup and Hup both point to a glottalic stop. In 88–89, Dâw points to a glottalic stop, and Hup points to a voiced stop or nasal. Furthermore, in a handful of etymologies a glottalic stop is clearly reconstructible based on evidence from Dâw and Hup, but Yuhup lacks the expected laryngealization (see 55, 165, 179, 259, 326).

- (83) (?) PDH **cēp* ~ **c'ēp* ‘to tie’ > Dâw <ṣūp> *ṣū.p* ‘to tie an *aturá* basket’ || Hup <s'íp> *tṣíp* ‘to tie (e.g. a canoe, a pole)’
- (84) (?) PDH **dī₂~k* ~ **nī₂~k* ‘asymmetrical’ > Dâw <ṭum-dik> *tum-dik* ‘monocular-visioned’ || Hup <ník> *ník* ‘crooked (arm, hand, bird leg)’
- (85) (?) PDH **t.xom* ~ **t.xob* ‘*piquiá* fruit (*Caryocar villosum*)’ > Dâw <xom> *xom* || Hup <tohób> *tóhóp*
- (86) (?) PDH **dī:c* ~ **dī:c* ‘to mash, to crush’ > Dâw <dij> *dī:c* ‘to smash’ || Yuhup <dij> *dījɲ* ‘to crush pepper’ || Hup <d'ij> *díjt* ‘to mash, to crush (food)’
- (87) (?) PDH **cī:b* ~ **cī:b* ‘to pinch, to dig one’s claws into’ > Dâw <siib> *ṣi:b* || Yuhup <sīim> *tṣīimb* || Hup <sím> *tṣimp*
- (88) (?) PDH **píd* ~ **píd* ~ **pín* ‘to have been used to’ > Dâw <pun> *puṇ*² || Hup <píd> *pīdn*
- (89) (?) PDH **hób* ~ **hób* ~ **hóm* ‘*abiu* fruit (*Pouteria caimito*)’ > Dâw <rôm> *hóm*² || Hup <hòb> *hó:bm* ‘plant similar to *abiu* with edible fruit’

3.3. Plain glides

The glides **w* and **j* are straightforwardly reflected as /w/ and /j/ in all daughter languages, except for Ospina Bozzi’s (2002) analysis of Yuhup, which treats morpheme-initial instances of [j] (before oral vowels) and [ɲ] (before nasal vowels) as allophones of the phoneme /j/. Other authors represent the respective sounds of Yuhup as [j] and [ɲ] (Fernandes 2017) or as [j] in oral and nasal environments alike (Lopes 1995; Silva & Silva 2012). The latter analysis is adopted in the broad transcriptions in this paper. Fortition of **j* to [j] and [ɲ] may turn out to be an innovation restricted to the southwestern fringe of the Yuhup-speaking area. In Hup, morpheme-initial instances of /j/ unpack to [dj] before oral vowels and to [nɲ] before nasal vowels (in this paper, I use *dj* and *nɲ* in my broad transcriptions). This is summarized in Table 5.

PDH	Dâw	Yuhup	Hup
*w	<w> <i>w</i>	<w> <i>w</i>	<w> <i>w</i>
*j	<y> <i>j</i>	<y> <i>j</i>	<y> <i>dj, nɲ^A, j^B</i>

^A = before nasal vowels / in nasal morphemes; ^B = morpheme-finally

Table 5. PDH plain glides and their reflexes

Some examples follow in 90–98.

- (90) PDH **wō:h* ‘speaker of a Tukanoan language’ > Dâw <woor> *wō:h* || Yuhup <wóh> *wôh* || Hup <wòh> *wôh*
- (91) PDH **wân* ‘machete’ > Dâw <wāan> *wā:n* || Yuhup <wān> *wān*²⁰ || Hup <wàn> *wān*
- (92) PDH **wī₂:w* ‘bullet ant’ > Dâw <wūw> *wū:w* || Yuhup <wíw> *wîw* || Hup <wìw> *wîw*
- (93) PDH **wāw* ‘crooked (of wood)’ > Dâw <wāw> *wāw* || Yuhup <wāw> *wāw*

²⁰ The rising tone in the Yuhup reflex, attested both in Silva & Silva (2012: 302) and Ospina Bozzi (2002: 204), is quite unexpected; based on the Dâw and Hup cognates, one would expect falling tone in Yuhup.

- (94) PDH **jǎ₂:ʔ* ‘to defecate’ > Dâw <yee> *jě:ʔ* || Yuhup <ye> *jǎ:ʔ* || Hup <yé> *djǎ:ʔ*
 (95) PDH **jũ:m* ‘to plant; cultivated plant’ > Dâw <yũum> *jũ:m* || Yuhup <yũm> *jũm* ‘to plant’; <yũm> *jũm* ‘cultivated plant’ || Hup <yũm> *njũm* ‘to plant’; <yũm> *njũm* ‘cultivated plant’
 (96) PDH **jô:j* ‘pineapple sp.’ > Dâw <yooy> *jô:j* || Yuhup <yóy> *jô:j* || Hup <yòy> *djô:j*
 (97) PDH **ʔâ₂:j* ‘to call’ > Dâw <âay> *ʔâ:j* || Yuhup <ëy> *ʔë:j* || Hup <éy> *ʔë:j*
 (98) PDH **mâ:j* ‘house’ > Dâw <māay> *mâ:j* || Yuhup <móy> *mâ:j* || Hup <mòy> *mâ:j*

3.4. Glottalized glides

Two glottalized glides can be reconstructed for PDH, **w̥* and **j̥*. In Dâw, their reflexes are articulated as preglottalized in onsets and postglottalized in codas (Martins 2004: 52–53). Postglottalization is also documented in their reflexes in the morpheme-final position in Yuhup and Hup. As for the morpheme-initial position in Yuhup and Hup, the reflexes of **w̥* and **j̥* are articulated much like those of **w* and **j*, respectively, but the following vowel surfaces with laryngealization, or creaky voice (see Epps 2005: 55–58, 64–65; Ospina Bozzi 2002: 117–118). In Yuhup, but not Hup, the laryngealized vowel is often described as rearticulated ([V̥V̥] in Silva & Silva 2012: 85; [V̥V̥] in Lopes 1995: 91; [V̥V̥] ~ [V̥] in Epps 2005: 68). As shown by Epps (2005: 66–67) for Hup, the laryngealization in such cases is best attributed to a [+constricted glottis] feature associated with the consonant in the morpheme-initial position (thus, underlying glottalized glides /w̥/ and /j̥/ are posited for Hup and, by some authors, for Yuhup). It is noteworthy that in some Hup words morpheme-initial *djV̥V̥* (/j̥V̥V̥/ in Epps’ analysis) varies with *tʃV̥V̥* (/tʃV̥V̥/ in Epps’ analysis), as documented in Ramirez (2005). This variation is in all likelihood dialectal. Note that there are no reliable examples featuring **j̥* in a nasal morpheme. These reflexes are summarized in Table 6.

PDH	Dâw	Yuhup	Hup
* <i>w̥</i>	<ʔw> <i>ʔw</i> , <wʔ> <i>wʔ^A</i>	<wVV> <i>w̥</i> , <wʔ> <i>wʔ^A</i>	<wʔ> <i>w̥</i> , <wʔ> <i>wʔ^A</i>
* <i>j̥</i>	<ʔy> <i>ʔj</i> , <yʔ> <i>jʔ^A</i>	<yVV> <i>j̥</i> , <jʔ> <i>jʔ^A</i>	<yʔ> <i>dj̥</i> ~ <sʔ> <i>tʃ̥^B</i> , <yʔ> <i>jʔ^A</i>

^A = morpheme-finally; ^B = the variation is possibly dialectal

Table 6. PDH glottalized glides and their reflexes

Some examples follow in 99–106.

- (99) PDH **w̥ot* ‘long’ > Dâw <ʔwât> *ʔw̥ot* || Yuhup <wäḁ̈t> *w̥ḁ̈t* || Hup <wʔât> *w̥ät*
 (100) PDH **w̥ô:b* ‘to put onto’²¹ > Dâw <ʔwôob> *ʔw̥ô:b* || Yuhup <wööm> *w̥ö̥bm* || Hup <wʔôb> *w̥ôbm*
 (101) PDH **d̥w̥* ‘to squeeze’ > Dâw <ʔnũw> *ʔnũ:w̥* ‘to mash’ || Yuhup <ni̥w> *n̥i̥w̥* ‘to squeeze one’s skin’
 (102) PDH **j̥æw* ‘to be crushed’, **j̥ǎ₁:w̥* ‘to crush’²² > Dâw <yew> *ʔj̥ew̥*, <yew> *ʔj̥ew̥* || Yuhup <yaw> *j̥aw̥* ‘to crush with one’s hand’ || Hup <yʔaw> *dj̥aw̥* ~ <yʔaw> *dj̥aw̥* ‘to crush with one’s hand’

²¹ This is an opaque causative of PDH **w̥ô:b* ‘to be on or over something’ > Dâw <wôob> *w̥ô:b*, Yuhup <wööm> *w̥ö̥bm*, Hup <wʔôb> *w̥ôbm* (cf. Epps 2005: 68 on this and other pairs with a similar alternation).

²² This cognate set presents several difficulties. Semantically, it is tempting to include Dâw <yew> *ʔj̥ew̥* ‘to mash (manioc), to knead’ here, but Dâw /e/ is not known to correspond to Yuhup and Hup /a/. As for the phonological issues, **j̥* shows an unexpected loss of glottalization in Yuhup, whereas in Hup the unexpected disyllabic

- (103) PDH **juʔ* ‘soft’ > Dâw <’yu’> *ʔjuʔ* || Yuhup <yuu’> *juʔʔ* || Hup <y’ú’> *djúʔ* ~ <s’ú’> *tʃúʔ*
 (104) PDH **j̥j̥ʔ* ‘to stretch’ > Dâw <’yoo’> *ʔj̥ʔ* || Yuhup <yoo’> *j̥j̥ʔ* || Hup <y’ó’> *dj̥ʔ* ~ <s’ó’> *tʃ̥ʔ*
 (105) PDH **kaj* ‘to grab, to hug’ > Dâw <xay’> *xajʔ* ‘to grab with arms’ || Yuhup <kay’> *kăjʔ* ‘to hug’ || Hup <káy’> *kâjʔ* ‘to hug (with one’s hand on another person’s shoulder)’
 (106) PDH **j̥j̥ʔ* ‘to sway, to shake’ > Dâw <yôoy’> *j̥j̥ʔ* || Yuhup <yöy’> *j̥j̥ʔ* || Hup <yöy’> *dj̥j̥ʔ*

In a handful of examples, Dâw and Hup/Yuhup show irregular correspondences involving reflexes of glottalized glides in one language and reflexes of other consonants in other languages. Such etymologies are, therefore, suspect and may result from horizontal transmission or chance. In 107–113, Dâw points to a glottalized glide, whereas Yuhup and Hup point to a plain glide. In 111, there is an additional mismatch between vowel length: Dâw points to a long vowel, whereas Yuhup and Hup point to a short vowel. 112 has further irregular variants in Yuhup (<yukuy> *jukúj*) and Hup (<kukúç> *kukújh*), and is likely to be a Wanderwort, since similar-sounding terms for monkeys are found in other languages (cf. Nikulin & Carvalho 2018: 557), including Proto-Goyaz (< Macro-Jê) **kúkôj* ‘monkey’; Yanomam (< Yanomaman) <kuukuumoxi> *ku:ku:-moxi* ‘three-striped night monkey’ (Emiri 1987: 41);²³ Yanomami (< Yanomaman) <kuukuumi> *ku:ku:-mi* ~ <kukumi> *kuku-mi* ‘three-striped night monkey’ (Mattei-Müller & Serowë 2007); and Venezuelan Spanish <cocu> *kɔ'kwi* ‘three-striped night monkey’ (Aguilar 2004: 3, 5), whose etymology I have been unable to verify.

- (107) (?) PDH **w̥ʃ:k* ~ **w̥ʃ:k* ‘vine sp.’ > Dâw <’wâak’> *w̥ʃ:k* || Yuhup <wêk’> *wêk* ‘uambé vine’ || Hup <wêk’> *wêk* ‘vine related to uambé (children use it to play)’
 (108) (?) PDH **pɔw* ~ **pɔw* ‘to float’ > Dâw <pow’> *pɔwʔ* || Yuhup <pow’> *pɔw*
 (109) (?) PDH **d.de₂w* ~ **d.de₂w* ‘knot’ > Dâw <lêw’> *lewʔ* ‘knot (on a thick rope or vine)’ || Yuhup <dëdëw’> *dedëw* ~ *derëw* || Hup <söb dërëw’> *tʃobm-nderëw* ‘interphalangeal joint’
 (110) (?) PDH **j̥j̥ʔ* ~ **j̥j̥ʔ* ‘wasp’ > Dâw <’yoo’> *ʔj̥ʔ* || Yuhup <yó’> *j̥ʔ* || Hup <yò’> *dj̥ʔ*
 (111) (?) PDH **j̥j̥:k* ~ **j̥j̥:k* ‘neotropical otter, giant otter’ > Dâw <’yook’> *ʔj̥:k* || Yuhup <yok’> *j̥k* || Hup <yók’> *dj̥k*
 (112) (?) PDH **k.kuj* ~ **k.kuj* ‘howler monkey’ > Dâw <xuy’> *xujʔ* || Yuhup <kukuy> *kukúj* ~ <yukuy> *jukúj* || Hup <kukúç> *kukúh* ~ <kukúç> *kukújh*
 (113) (?) PDH **m̥Āj* ~ **m̥Āj* ‘payment’ > Dâw <mây’> *mâjʔ* || Yuhup <mey’> *măj* ‘to avenge’, <méy’> *mâj* ‘payment’²⁴ || Hup <mey’> *mâj* ‘payment, price’, <méy’> *mâj* ‘to pay, to threaten’

A different kind of irregularity is seen in 114, where a plain glide in Dâw corresponds to a glottalized one in Yuhup and Hup. In 115, Dâw points to a glottalized glide, whereas Yuhup and Hup support the reconstruction of a voiced stop or a nasal. In 116, Dâw points to a glottalized glide, whereas Yuhup and Hup point to a glottalic stop. In 117, a reflex of a glottalized glide in Hup corresponds to zero in Dâw, which can go back to zero or to **h*. In 118, both Dâw and Hup support the reconstruction of **ʔ*, but Yuhup rather points to **w*. Finally, in 119, Dâw points to **ʔ*, Yuhup to **k*, and Hup to **w*.

variant <y’a’áw’> *djaʔâwʔ* is attested instead of the expected variant *<s’áw’> **tʃâwʔ*. Compare the semantically comparable Hup verb <y’á’> *djaʔ* ~ <s’á’> *tʃâʔ* ‘to mash (e.g., pepper)’.

²³ This form is from the Wakathau thëripë dialect. In the Maraxiu thëripë (Papiú) dialect, this is attested as <kukumoxi> (Perri Ferreira 2017: 119); note that the status of vowel length in Yanomam is unclear.

²⁴ The Yuhup noun <méy’> *mâj* is probably not a direct reflex of PDH **m̥Āj* ~ **m̥Āj* (otherwise a rising tone would be expected), but rather a back-formation from the verb ‘to avenge’, where the falling tone is a nominalizer described in Silva & Silva (2012: 101).

- (114) (?) PDH *c'ɔj ~ *c'ɔj̃ 'Amazon parrot' > Dâw <çoy> c'ɔj || Yuhup <sooy> tʃɔ̃j̃ʔ || Hup <s'óy> tʃɔ̃j̃ʔ
- (115) (?) PDH *poɰ ~ *pob ~ *pom 'to split wood' > Dâw <pôw> poɰʔ || Yuhup <pöm> pöbm || Hup <pób> pôbm
- (116) (?) PDH *c.ɰæ~j ~ *c.ɰæ~j̃ 'Large American opossum' > Dâw <'wey> ʔwej || Yuhup <maay> mǎǎj̃ || Hup <sam'áy> tʃāmǎǎj̃
- (117) (?) PDH *p.xɔ:(h) ~ *p.xɔ:wʔ 'to swell' > Dâw <xoo> xɔ: 'a bit swollen; abscess, tumor' || Hup <pohów> pɔhɔwʔ
- (118) (?) PDH *cɔʔ ~ *cɔɰ 'matapi fishing trap' > Dâw <sõ> ʃɔʔ || Yuhup <sõw> tʃɔ̃wʔ || Hup <sõ> tʃɔʔ
- (119) (?) PDH *nĩʔ ~ *nĩk' ~ *nĩɰ 'bait' > Dâw <nĩ> nĩʔ || Yuhup <nig> nĩŋg || Hup <níw> nĩwʔ

Advanced etymological research is needed in order to establish whether the etymologies in 114–119 involve actual cognates.

3.5. Fricatives

Three voiceless fricatives are reconstructed for PDH: *ç, *x, and *h. Of these, *ç appears in the morpheme-final position only, whereas *x and *h appear in any position. The fricatives are subject to the following innovations. In Dâw, *ç yields ʃ (120–125), and *x remains as x (126–131); therefore, morpheme-initially these segments merge with *c and *k, respectively. The glottal fricative *h is lost morpheme-finally in falling-toned morphemes (138–139); in rising-toned and toneless morphemes, it is preserved (134–137), giving rise to alternations between h and zero in Dâw (cf. Martins 2004: 41). In Yuhup and Hup, *ç remains as /ç/ (120–122), except that it yields h after *i, *ĩ, *ɣ₂ (123–125).²⁵ Note that in Hup /ç/ surfaces as [jh]. In Yuhup, /ç/ is documented as [ç] by Silva & Silva (2012) and Fernandes (2017), whereas Lopes (1995) and Ospina Bozzi (2002) document only [jh] (which is a phonemic sequence /jh/ in their analysis). The velar fricative *x merges with the glottal fricative *h as /h/ both in Yuhup and Hup (126–131). The aforementioned reflexes are summarized in Table 7.

PDH	Dâw	Yuhup	Hup
*ç	<s> ʃ	<ç> ç, <h> h ^A	<ç> jh
*x	<x> x	<h> h	<h> h
*h	<h> r, <∅> ∅ ^B	<h> h	<h> h

^A = after the vowels *i, *ĩ, *ɣ₂; ^B = morpheme-finally in falling-toned morphemes

Table 7. PDH voiceless fricatives and their reflexes

Some examples follow in 120–139.

- (120) PDH *k'ɣç 'to bite' > Dâw <kâs> k'ɣʃ || Yuhup <kääç> kǎǎç || Hup <k'áç> kǎjh
- (121) PDH *c'ɔ:ç 'to spit' > Dâw <çoos> c'ɔ:ʃ || Yuhup <sooç> tʃɔ̃ç || Hup <s'óc> tʃɔ̃jh
- (122) PDH *jê₂:ç 'guan' > Dâw <yêes> jêʃ || Yuhup <yěç> jêç || Hup <yèç> djêjh

²⁵ I have contemplated the possibility to reconstruct PDH *x for the correspondence Dâw ʃ ~ Yuhup/Hup h; in this case, Dâw would have undergone progressive palatalization of a velar, otherwise seen in *mĩŋ > mĩŋ 'crazy'. However, no palatalization is seen in PDH *nĩɰ:x 'water' > Dâw <nâax> nĩɰ:x, ruling out the reconstruction of *x in PDH *bĩɰ:ç 'to cross' > Dâw <bâas> bĩɰ:ç.

- (123) PDH **b̥ʒ₂:ç* ‘to cross’ > Dâw <bâas> *b̥ʒ:f* || Yuhup <bëeh> *b̥ɛ̃h* || Hup <b’êh> *mb̥ɛ̃h*
- (124) PDH **t̥i:ç* ‘root’ > Dâw <tiis> *t̥i:f* || Yuhup <tih> *t̥ih* || Hup <tih> *t̥ih*
- (125) PDH **m̥i:ç* ‘turtle’ > Dâw <m̥iis> *m̥i:f* || Yuhup <mih> *m̥ih* || Hup <mih> *m̥ih*
- (126) PDH **x̥a:t* ‘name’ > Dâw <xaat> *x̥a:t* || Yuhup <hát> *h̥at* || Hup <hàt> *h̥at*
- (127) PDH **x̥â:j* ‘forest, outside’ > Dâw <xaay> *x̥â:j* || Yuhup <háy> *h̥aj* ‘forest’, <hay-> *h̥aj-* ‘outside’²⁶ || Hup <hày> *h̥aj* ‘forest’, <háy> *h̥aj* ‘outside’
- (128) PDH **x̥i₂:* ‘to descend, to go downriver’ > Dâw <xuu> *x̥i:* || Yuhup <hi> *h̥i* || Hup <hi> *h̥i*
- (129) PDH **t̥a:x* ‘tapir’ > Dâw <taax> *t̥a:x* || Yuhup <táh> *t̥ah* || Hup <tàh> *t̥ah*
- (130) PDH **c̥o:x* ‘to walk with a walking stick’ > Dâw <sôox> *ʃo:x* || Yuhup <söh pēm> *tʃoh-p̥əm* ‘to squat’ || Hup <sòh> *tʃoh* ‘walking stick’
- (131) PDH **n̥õx* ‘to fall’ > Dâw <nõx> *n̥õx* || Yuhup <noh> *n̥õh* ‘to bump’ || Hup <nóh> *n̥õh*
- (132) PDH **h̥ɔ:k* ‘to drown’ > Dâw <râak> *h̥ɔ:k* || Hup <hák> *h̥ak*
- (133) PDH **h̥ɔ:p* ‘fish’ > Dâw <rãap> *h̥ɔ:p* || Yuhup <hóp> *h̥op* || Hup <hòp> *h̥op*
- (134) PDH **w̥i:h* ‘hawk’ > Dâw <wiir> *w̥i:h* || Yuhup <wih> *w̥ih* || Hup <wih> *w̥ih*
- (135) PDH **j̥õ:h* ‘medicine’ > Dâw <yoor> *j̥õ:h* || Yuhup <yóh> *j̥oh* || Hup <yòh> *nj̥oh*
- (136) PDH **ʔoh* ‘common woolly monkey’ > Dâw <ôr> *ʔoh* || Yuhup <öh> *ʔoh* || Hup <öh> *ʔoh*
- (137) PDH **n̥ũh* ‘head’ > Dâw <nũr> *n̥ũh* || Yuhup <nuh> *n̥ũh* || Hup <núh> *n̥ũh*
- (138) PDH **x̥ô:h* ‘canoe’ > Dâw <xoo> *x̥ô:* || Yuhup <hóh> *h̥oh* || Hup <hòh> *h̥oh*
- (139) PDH **m̥ê:h* ‘ucuqui fruit (*Pouteria ucuqui*)’ > Dâw <mũu> *m̥ê:* || Yuhup <mih> *m̥êh* || Hup <mih> *m̥êh*

3.6. Nasals and voiced stops

PDH had both nasals (**m*, **n*, **ŋ*) and voiced stops (**b*, **d*, **ɟ*, **g*), an opposition still present in Dâw. However, the contrast in question is robust only in the morpheme-final position. Morpheme-initially, by contrast, only **m*, **n*, and **d* are reconstructed. In Yuhup and Hup, nasals no longer contrast with voiced stops; in these languages, nasality is not a feature of individual segments, but rather of syllables due to Tukano influence (Epps 2005: 75).²⁷ Authors such as Martins (2005: 81–82) and Epps (2005: 52–53) analyze the nasal consonants of Yuhup and Hup, respectively, as allophones of voiced stops /*b d ɟ g*/. In the same vein, they treat the nasal–oral contour segments, such as [mb], as prenasalized stops, whereas the oral–nasal contour segments, such as [bm], are considered postnasalized stops in their analyses. However, there is evidence that all these segments are in fact realizations of underlying nasals /*m n ɲ ŋ*/. In Yuhup, the fact that we are dealing with /*m n ɲ ŋ*/ rather than /*b d ɟ g*/ is seen in inflected forms where vowel-initial suffixes attach to oral roots ending in [bm], [dn], [ɲn], [gɲ]: in such forms, the vowel assimilates the nasality of the root-final consonant: <tami> [tábmí:] /tám-í/ ‘tying’, <tamap> [tábm̥á:p] /tám-ŋp/ ‘tying’ (Silva & Silva 2012: 84, 86). As for Hup, the language has circumoralized contours (as in <ágáy> [ʔəgɲ.ŋgój] ‘drinking’; Epps 2005: 41), which are shown by Wetzels & Nevins (2018) to be possible only in languages with underlying nasals. Therefore, in Wetzels & Nevins’ (2018) terms, Yuhup and Hup have nasal shielding rather than nasal venting.

²⁶ This term must receive a suffix indicating the distance.

²⁷ In fact, both Ospina Bozzi (2002) and Epps (2005) consider that nasality is a property of entire morphemes rather than syllables, but at least in Hup there are several morphemes that combine oral and nasal syllables (Epps 2005: 74), even though they are exceedingly rare in the lexicon.

I start by considering the reflexes of PDH **m*, **n*, and **d* in the morpheme-initial position. In Dâw, the former two are preserved as nasals, whereas **d* lenites to *l*.²⁸ Recall that synchronically Dâw does have morpheme-initial voiced stops, but these come from erstwhile glottalic stops, as shown in 3.2. In Yuhup and Hup, **m*, **n*, and **d* have oralized ([b], [d]) or postalized ([mb], [nd]) reflexes in oral syllables, and fully nasal reflexes ([m], [n]) in nasal syllables; note that **n* and **d* merge in these languages at least in oral environments (no examples of **d* are reconstructed in nasal environments). The nasal phase in [mb] and [nd] is attested in Hup word-initially by all authors; in Yuhup, it is documented in Lopes (1995) and Ospina Bozzi (2002), but not in Martins (2005), Silva & Silva (2012), or Fernandes (2017). In my broad transcriptions, I use *b* and *d* for Yuhup, but *mb* and *nd* for Hup. In Yuhup and, dialectally, in Hup, the reflexes of **n* and **d* may undergo intervocalic flapping. This is summarized in Table 8.

PDH	Dâw	Yuhup	Hup
<i>*m</i>	⟨m⟩ <i>m</i>	⟨b⟩ <i>b</i> , ⟨m⟩ <i>m</i> ^A	⟨b⟩ <i>mb</i> , ⟨m⟩ <i>m</i> ^A
<i>*n</i>	⟨n⟩ <i>n</i>	⟨d⟩ <i>d</i> , ⟨n⟩ <i>n</i> ^A , ⟨ɾ⟩ <i>ɾ</i> ^B	⟨d⟩ <i>nd</i> , ⟨n⟩ <i>n</i> ^A , ⟨ɾ⟩ <i>ɾ</i> ^C
<i>*d</i>	⟨l⟩ <i>l</i>	⟨d⟩ <i>d</i> , ⟨n⟩ <i>n</i> ^{AD} , ⟨ɾ⟩ <i>ɾ</i> ^B	⟨d⟩ <i>nd</i> , ⟨n⟩ <i>n</i> ^{AD} , ⟨ɾ⟩ <i>ɾ</i> ^C

^A = before nasal vowels; ^B = between vowels, possibly with an intervening glottal; ^C = between vowels in the Tat-Dêh and Umari Norte dialect areas; ^D = hypothetical reflex predicted based on systemic considerations

Table 8. PDH nasals and voiced stop and their reflexes (morpheme-initial position)

Some examples follow in 140–149.

- (140) PDH **muj* ‘cold season’ > Dâw ⟨muy⟩ *muj* || Yuhup ⟨buy⟩ *bũj* || Hup ⟨búy⟩ *mbũj*
 (141) PDH **mən* ‘maripa palm fruit’ > Dâw ⟨mân⟩ *mən* || Yuhup ⟨băn⟩ *băn*
 (142) PDH **môh* ‘tinamou’ > Dâw ⟨môo⟩ *môh* || Yuhup ⟨móh⟩ *môh* || Hup ⟨mòh⟩ *môh*
 (143) PDH **môj* ‘Humboldt’s white-fronted capuchin’ > Dâw ⟨môoy⟩ *môj* || Yuhup ⟨móy⟩ *môj*
 (144) PDH **nid* ‘tree stump’ > Dâw ⟨bee nud⟩ *bê:-nuid* || Yuhup ⟨din⟩ *dĩdn* || Hup ⟨díd⟩ *ndĩdn*
 (145) PDH **nǝ₂:x* ‘water’ > Dâw ⟨nâax⟩ *nǝ:x* || Yuhup ⟨dêh⟩ *dêh* || Hup ⟨dêh⟩ *ndêh*
 (146) PDH **nâ:* ‘to say’ > Dâw ⟨nâa⟩ *nâ:* || Yuhup ⟨no⟩ *nǝ* || Hup ⟨nó⟩ *nâ:*
 (147) PDH **nê:m* ‘louse’ > Dâw ⟨nêem⟩ *nê:m* || Yuhup ⟨ném⟩ *nêem* || Hup ⟨nèm⟩ *nêem*
 (148) PDH **dâ:j* ‘fishhook’ > Dâw ⟨laay⟩ *lâ:j*? || Yuhup ⟨dáy⟩ *dâj*?
 (149) PDH **doj* ‘to crouch’ > Dâw ⟨lôy⟩ *loj* || Yuhup ⟨döy⟩ *děj* ‘to crawl, to creep’ || Hup ⟨döy⟩ *nděj* ‘to crouch, to crawl, to creep’

Morpheme-finally, all nasals and voiced stops are preserved in Dâw. Note that nasal codas following oral vowels surface with preoralization, thus instantiating the phenomenon of nasal shielding (Wetzels & Nevins 2018). In Yuhup and Hup, the nasal series merges with that of the voiced stops. Their reflexes are articulated as preoralized nasals after oral vowels (i.e., with nasal shielding), and as full nasals after nasal vowels. I believe that the nasal shielding phenomenon was already present in PDH, but I do not represent it in my reconstructions due to its subphonemic nature. The consonant **b* is not known to have occurred after nasal vowels.

²⁸ A reviewer suggests that “[a]nother possible option would be to reconstruct **l* instead of morpheme-initial **d*, with the typologically trivial change **l* > *n* in Yuhup and Hup”, which “would remove the major asymmetry between **d* and other voiced stops”. Such a possibility is rendered unlikely by the fact that a series of voiced stops, including **d*, is unequivocally reconstructed morpheme-finally (Table 9 and examples 160–163).

The velar nasal * η palatalizes after front high vowels. The relevant correspondences are summarized in Table 9.

PDH	Dâw	Yuhup	Hup
* m	⟨m⟩ <i>bm, m^A</i>	⟨m⟩ <i>bm, m^A</i>	⟨b⟩ <i>bm, ⟨m⟩ m^A</i>
* b	⟨b⟩ <i>b</i>	⟨m⟩ <i>bm, m^{AB}</i>	⟨b⟩ <i>bm, ⟨m⟩ m^{AB}</i>
* n	⟨n⟩ <i>dn, n^A</i>	⟨n⟩ <i>dn, n^A</i>	⟨d⟩ <i>dn, ⟨n⟩ n^A</i>
* d	⟨d⟩ <i>d</i>	⟨n⟩ <i>dn, n^A</i>	⟨d⟩ <i>dn, ⟨n⟩ n^A</i>
* j	⟨j⟩ <i>j</i>	⟨j⟩ <i>jn, j^A</i>	⟨j⟩ <i>jdn, jn^A</i>
* η	⟨gn⟩ <i>gη, η^{AB}, ⟨nh⟩ <i>jn^{BC}, j^{AC}</i></i>	⟨g⟩ <i>gη, η^{AB}</i>	⟨g⟩ <i>gη, η^{AB}</i>
* g	⟨g⟩ <i>g</i>	⟨g⟩ <i>gη, η^A</i>	⟨g⟩ <i>gη, η^A</i>

^A = before nasal vowels / in nasal morphemes; ^B = hypothetical reflex predicted based on systemic considerations; ^C = after front high vowels

Table 9. PDH nasals and voiced stop and their reflexes (morpheme-final position)

Some examples follow in 150–173.

- (150) PDH * $c\hat{t}_2:m$ ‘foot’ > Dâw ⟨çuəm⟩ *c’û:bm* || Yuhup ⟨siím⟩ *tʃiím* || Hup ⟨s’ib⟩ *tʃibm*
- (151) PDH * jum ‘vine’ > Dâw ⟨yum⟩ *jubm* || Yuhup ⟨yum⟩ *jǔbm* || Hup ⟨yúb⟩ *djúbm*
- (152) PDH * $h\hat{a}:m \sim *h\hat{a}m$ ‘to go’ > Dâw ⟨rāam⟩ *hā:m \sim ⟨rām⟩ hām* || Yuhup ⟨hām⟩ *hǎm* || Hup ⟨hám⟩ *hām*
- (153) PDH * $t\hat{õ}m$ ‘tree sp.’ > Dâw ⟨tôm⟩ *tõm* ‘embaúba tree (*Cecropia* spp.)’ || Hup ⟨tóm⟩ *tõm* ‘kind of *macucu* tree found in the caatinga’
- (154) PDH * $d\hat{a}:b$ ‘to weave’ > Dâw ⟨daab⟩ *dâ:b* ‘to weave (palm leaves)’ || Hup ⟨d’áb⟩ *dâbm* ‘to weave (hammock)’
- (155) PDH * $b\hat{o}b$ ‘loincloth; plant whose bast is used for making loincloths’ > Dâw ⟨bôb⟩ *bob* || Yuhup ⟨bööm⟩ *bqõbm* || Hup ⟨b’ób⟩ *mbôbm* ‘loincloth; tropical chestnut’
- (156) PDH * pun ‘to suck breast’, * $p\hat{u}:n$ ‘breast’ > Dâw ⟨pun⟩ *pun*, ⟨puun⟩ *pû:n* || Yuhup ⟨pun⟩ *pǔdn*, ⟨pún⟩ *pûdn* || Hup ⟨púd⟩ *pûdn*, ⟨pùn⟩ *pûdn*
- (157) PDH * $m.han$ ‘to appear’ > Dâw ⟨ran⟩ *hadn* || Yuhup ⟨wahan⟩ *wahǎdn* || Hup ⟨bahád⟩ *mbahâdn*
- (158) PDH * $c\hat{õ}n$ ‘to prick one’s foot on a stump’ > Dâw ⟨sôn⟩ *ʃn* || Yuhup ⟨sôn⟩ *tʃõn* ‘to step on a stump or in a puddle’ || Hup ⟨són⟩ *tʃõn* ‘to prick oneself’
- (159) PDH * $m.h\hat{a}:n$ ‘kinkajou’ > Dâw ⟨rēen⟩ *hē:n* || Yuhup ⟨wēhén⟩ *wǎhǎn* || Hup ⟨mehèn⟩ *mǎhǎn*
- (160) PDH * $d\hat{a}:d$ ‘to paint with genipap’ > Dâw ⟨daad⟩ *dâ:d* ‘to write, to study, to paint one’s body’ || Yuhup ⟨daan⟩ *dqǎdn* || Hup ⟨d’ád⟩ *ndâdn*
- (161) PDH * $k\hat{ə}d$ ‘to pass’ > Dâw ⟨xâd⟩ *xɔd* || Yuhup ⟨kän⟩ *kǎdn* || Hup ⟨kád⟩ *kâdn*
- (162) PDH * $n\hat{a}:d$ ‘to come’ > Dâw ⟨nēed⟩ *nē:d* || Yuhup ⟨nen⟩ *nǎn* || Hup ⟨nén⟩ *nân*
- (163) PDH * $m\hat{a}:d$ ‘downriver’ > Dâw ⟨mēed⟩ *mē:d* || Yuhup ⟨mer-⟩ *mǎr-* ~ ⟨men-⟩ *mân-* ~ *mǎn-*²⁹ || Hup ⟨mèr’ah⟩ *mǎr-ʔah* ~ ⟨mér’ah⟩ *mân-ʔah*³⁰

²⁹ This term must receive a suffix indicating the distance. Silva & Silva (2012: 216) document the root-final consonant as lenited to ⟨r⟩ *r*; while Ospina Bozzi (2002: 237) attests ⟨n⟩ *n*. The variation between the rising and falling tones, attested in Ospina Bozzi (2002: 128, 131, 236–238, 437, 448–449, 465), is unclear.

³⁰ The variant ⟨mír’ah⟩ *mîr-ʔah*, documented in Ramirez (2005), shows a sound change attested elsewhere in the Umari Norte dialect area (Epps 2005: 88), and may thus be representative of that variety. The variation between the rising and falling tones, attested in Epps (2005: 157, 302, 391, 440, 771), is unclear.

- (164) PDH **k'aj* 'green acouchi' > Dâw <kaj> *k'aj* || Yuhup <kaaj> *kaǎjɲ* || Hup <k'áj> *kâjɲn*
- (165) PDH **dɔj* 'to rain' > Dâw <dôj> *dɔj* || Yuhup <døj> *dǒjɲ*³¹ || Hup <d'ôj> *ndôjɲn*
- (166) PDH **k'ɔj* 'snail' > Dâw <kôj> *k'ɔj* || Yuhup <kōoj-tôn> *kǔjɲ-tǔdn* 'flute made of snail shell' || Hup <k'ôj> *kǔjɲ* 'snail sp. (not edible, mid-sized)'
- (167) (?) PDH **w.wɔ̃j* ~ **w.wɔ̃j* 'whirlpool'³² > Dâw <wōoj> *wǔj* || Yuhup <wōwoç> *wǔwǔç* 'to produce a whirlpool' || Hup <wōwôj> *wǔwǔjɲ*
- (168) PDH **minj* 'giant anteater' > Dâw <mugn> *muɣɲ* 'nickname for giant anteaters' || Yuhup <big> *bǐgɲ* || Hup <bíg> *mbǐgɲ*
- (169) PDH **mĩ₂ɲ* 'crazy' > Dâw <mĩnh> *mĩɲ* || Yuhup <mig> *mǐɲ* || Hup <míg> *mĩɲ*
- (170) PDH **tê₂g* 'firewood; tree (of a given species)' > Dâw <tâag> *tǎ:g* || Yuhup <tég> *têgɲ* || Hup <têg> *têgɲ* 'firewood', <-têg> *-tegɲ* 'tree (of a given species)'
- (171) PDH **jæg* 'hammock' > Dâw <yeg> *jɛg* || Yuhup <yag> *jǎgɲ* || Hup <yág> *djâgɲ*
- (172) PDH **cǎ:g* 'Brazilian tinamou' > Dâw <sōog> *ʃǎ:g* || Yuhup <sóg> *tʃǎ:ɲ* || Hup <sòg> *tʃǎ:ɲ*
- (173) PDH **nê:g* 'honey' > Dâw <nēeg> *nê:g* || Yuhup <nég> *nǎɲ* || Hup <nèg> *nǎɲ*

4. Vowels

The reconstruction of PDH vowels is associated with significant complexities. Although Dâw, Yuhup, and Hup have isomorphic vowel inventories, composed of nine oral vowels and six nasal vowels (in addition to their long counterparts), as shown in Table 10, the correspondences among them are not always straightforward. The proposal in this paper is preliminary and subject to improvements.

oral	front unrounded	non-front unrounded	non-front rounded	nasal	front unrounded	non-front unrounded	non-front rounded
high	<i>i</i>	D. <i>u</i> , Y./H. <i>ɨ</i>	<i>u</i>	high	<i>ĩ</i>	D. <i>ũ</i> , Y./H. <i>ĩ</i>	<i>ũ</i>
mid	<i>e</i>	D. <i>ɜ</i> , Y./H. <i>ə</i>	<i>o</i>	mid			
low	D. <i>ɛ</i> , Y./H. <i>æ</i>	<i>a</i>	<i>ɔ</i>	low	D. <i>ẽ</i> , Y./H. <i>ǣ</i>	<i>ã</i>	<i>õ</i>

Table 10. Synchronic vowel inventories of Dâw, Yuhup, and Hup

4.1. Correspondences with matching nasality values

Let us start by examining the correspondences with matching nasality values (that is, those involving oral vowels in Dâw, Yuhup, and Hup, or nasal vowels in Dâw, Yuhup, and Hup), and let us ignore the vowel length for the time being. There are 26 such correspondences with no obvious complementary distribution patterns. Roundedness and height are obviously stable features in the Dâw–Hup languages: 25 out of 26 correspondences have matching roundedness values, and 24 out of 26 involve vowels of the same height in all daughter languages

³¹ The absence of laryngealization in the Yuhup reflex, as attested in Silva & Silva (2012: 167), is unexpected. The expected form with laryngealization is, however, documented in Ospina Bozzi (2002: 403–404, 444).

³² This etymology is dubious because of two irregularities. First, the final consonant in the Yuhup reflex actually points to *ç rather than *j. Second, the falling tone in Dâw does not match the falling tone in Hup; Dâw points to a long vowel with falling tone, whereas Hup points to a short (toneless) vowel. The Yuhup cognate is a verb and is consequently useless for reconstructing the tone.

(if /æ/, /ɛ/, /a/, and /ɔ/ are all considered low). Backness is a less stable feature, however: only 18 out of 26 correspondences show matching backness values (if central and back vowels are lumped together). Table 11 lists the correspondences in question, and indicates whether each correspondence occurs in morphemes with short and long vowels.

Dâw	Yuhup/Hup	short	long
i	ĩ	+	–
ĩ	ĩ	+	+
u	ĩ	+	+
u	ĩ	–	+
u	u	+	+
e	ə	+	+
e	e	+	+
ɤ	ə	+	+
ɤ	e	–	+
o	o	+	+
ɛ	a	+	+
ɛ	æ	–	+
a	a	+	+

ɔ	ɔ	+	+
ɔ	o	+	+
e	æ	+	+

Dâw	Yuhup/Hup	short	long
ĩ	ĩ	(+)	–
ĩ	ĩ	+	+
ũ	ĩ	+	+

ũ	ũ	+	+
---	---	---	---

ẽ	ã	+	–
ẽ	ã	–	+
ã	ã	+	+
ã	ã	(+)	+
õ	õ	+	+

ã	õ	+	+
---	---	---	---

(+) = correspondence is attested only in cognate sets with additional irregularities

Table 11. Vowel correspondences between Dâw and Yuhup/Hup (matching nasality values only)

It can be seen from Table 11 that the rounded vowels of Dâw and Yuhup/Hup correspond neatly to each other; it is straightforward to posit the vowels **u*, **o*, **ɔ*, **ũ*, **õ* for PDH, all of which occur as short and long, with unchanged reflexes in the daughter languages, as exemplified in 174–188. These five vowels usually have trivial correspondences in Nadëb, as in <bung> *bun* ‘horsefly’, <-doo> *-do:* (non-indicative <-do> *-do*) ‘to take away (sg.)’, <-sóóp> *-fɔ:p* (plural <-s’óóp> *-fɔ:p*) ‘to go up (away from the river)’, <hũt> *hũ:t* ‘tobacco’, <-nooh> *-nɔ:h* ‘mouth’ (Weir 1984: 25, 141, 170; Barbosa 2005: 36, 37).

(174) PDH **bux* ‘horsefly’ > Dâw <bux> *bux* || Yuhup <buuh> *bũh* || Hup <b’úh> *mbúh*

(175) PDH **tú:g* ‘howler monkey’ > Dâw <tuug> *tú:g* || Yuhup <túg> *tũg* || Hup <tùg> *tũg*

(176) PDH **pũ:p* ‘*paxiúba* palm (*Socratea exorrhiza*)’ > Dâw <puup> *pũ:p* || Yuhup <púp> *púp* || Hup <pùp> *pũp*

- (177) PDH **dob* ‘to do gown (towards the river)’ > Dâw <dôb> *dob* || Yuhup <dööm> *dqǒbm* || Hup <d’ôb> *ndôbm*
- (178) PDH **nô*: ‘red, ripe’ > Dâw <dôo> *nô*: || Hup <d’ô> *ndô*:
- (179) PDH **dô*:? ‘to take out, to take away’ > Dâw <dôo> *dô*:? || Yuhup <dö> *dǒ?*³³ || Hup <d’ô> *ndǒ?*
- (180) PDH **c’ô*? ‘to untie’ > Dâw <çô> *c’ô*? || Yuhup <soo> *tʃǒ?* || Hup <s’ô> *tʃǒ?*
- (181) PDH **cǒ*:p ‘to go up (away from the river)’ > Dâw <soop> *ʃǒ:p* ‘to rise’ || Yuhup <sop> *tʃǒp* || Hup <sóp> *tʃǒp*
- (182) PDH **dǒ*:k ‘to go out, to be extinguished’ > Dâw <dook> *dǒ:k* || Yuhup <dook> *dǒ:k* || Hup <d’ók> *dǒk*
- (183) PDH **hũ*? ‘to end’ > Dâw <rũ> *hũ*? || Yuhup <hũ> *hũ?* || Hup <hú> *hũ?*
- (184) PDH **hũ*:t ‘tobacco’ > Dâw <rũut> *hũ:t* || Yuhup <hút> *hũt* || Hup <hũ> *hũt*
- (185) PDH **nũ*:h ‘manioc starch’ > Dâw <nũur> *nũ:h* || Yuhup <núh> *nũh* || Hup <nùh> *nũh*
- (186) PDH **nǒ*h ‘mouth’ > Dâw <nôr> *nǒh* || Yuhup <noh-köö> *nǒh-kǒǒdn* || Hup <noh-k’òd> *nǒh-kǒdn*
- (187) PDH **hǒ*k ‘to saw’ > Dâw <rôk> *hǒk* || Yuhup <hök> *hǒk* || Hup <hók> *hǒk*
- (188) PDH **hǒ*: ‘to burn (intr.)’ > Dâw <rôo> *hǒ*: || Yuhup <hō> *hǒ* || Hup <hó> *hǒ*:

The only complications involving rounded vowels are the sound correspondence between Dâw *ɔ*(:) and Yuhup/Hup *o*, found in two cognate sets only (189–190), and the correspondence between Dâw *ã*(:) and Yuhup/Hup *õ*, documented in at least seven cognate sets (e.g., 191–194). In the Roçado dialect of Nadëb, one finds the close-mid rounded vowel *o* in the former case (<poh> *poh* ‘sky’; Martins 2005: 288), and the high unrounded vowel *ĩ* in the latter case (<-’ỹh> *-?ẽh*, non-indicative <-’ỹh> *-?ĩh* ‘to sleep’, <myym> *mĩm* ‘ax’; Weir 1984: 31, 161). For the former correspondence, I use an *ad hoc* character **O*(:), which, judging by the evidence from Nadëb, may turn out to stand for plain **o*(:) with an irregular reflex in Dâw. Regarding the latter correspondence, I reconstruct an unrounded vowel **ĩ*(:), which must have been lowered in Dâw and rounded in Yuhup/Hup.

- (189) PDH **pOx* ‘up, above’ > Dâw <pox> *pox* (compare <poox> *pǒ:x* ‘sky’) || Yuhup <pöh> *pǒh* ‘tall’ (compare <pö> *pǒ* ‘up, above’) || Hup <pǒh> *póh* ‘up, above, sky’
- (190) PDH **kÔ*:ɿ ‘to scratch’ > Dâw <xooj> *xô:ɿ*³⁴ || Hup <kój> *kôjdn*
- (191) PDH **ĩ*ǒ:h ‘to sleep’ > Dâw <ãa> *ĩa*: || Yuhup <õh> *ĩh* || Hup <õh> *ĩh*
- (192) PDH **ĩ*ǒ:m ‘to be afraid’ > Dâw <ãam> *ĩa:m* || Yuhup <õm> *ĩm* || Hup <óm> *ĩm*
- (193) PDH **mĩ*:m ‘ax’ > Dâw <mãam> *mã:m* ‘stone ax’ || Yuhup <móm> *mĩm* ‘ax, metal’ || Hup <mòm> *mĩm* ‘metal, iron, metal ax’
- (194) PDH **cǒ*h ‘left (side)’ > Dâw <sôob sār> *ʃô:b ʃāh* ‘left hand’ || Yuhup <sōh> *tʃōh*³⁵ || Hup <sòh> *tʃōh*³⁶

Let us now consider the correspondences between unrounded vowels only. There are 19 such correspondences, of which 18 involve vowels of equal height in all daughter languages. The greatest complication here is the fact that the correspondences between non-front and

³³ The absence of laryngealization in the Yuhup reflex, as attested in Silva & Silva (2012: 165) and Ospina Bozzi (2002: 348), is unexpected.

³⁴ Martins (2005: 248) documents this as <xôoj> *xô:ɿ* instead. If the form with a low vowel, attested in Epps *et al.* (2018), is revealed to be a typo, one can simply reconstruct PDH **kô*:ɿ.

³⁵ This term must receive a suffix indicating the distance.

³⁶ The rising tone documented in the Hup reflex is unexpected.

front vowels are quite chaotic.³⁷ All four combinations exist. Dâw non-front vowels may correspond to non-front or front vowels of the same height in Yuhup/Hup. Similarly, Dâw front vowels may correspond to non-front or front vowels of the same height in Yuhup/Hup. In fact, only two combinations are not known to exist: Dâw *a(:)* is not known to correspond to Yuhup/Hup *æ*, and Dâw *ũ(:)* is not known to correspond to Yuhup/Hup *ĩ*. This raises the question of how to interpret these correspondences diachronically so as to avoid positing an implausibly large inventory of monophthongs. Did perhaps the vowel system of PDH include a typologically rare opposition between three degrees of backness in unrounded vowels? Or should we reconstruct an inventory of diphthongs for PDH? Or could the complexity result from splits conditioned by the consonantal environment, possibly including features such as velarization or palatalization lost in the contemporary languages?³⁸

Nadëb sheds little light on this complex issue. In most cases, oral high unrounded vowels of Dâw, Yuhup, and Hup, regardless of whether they are front or not, correspond to Nadëb *i* (as in <yb> *ʔib* ‘father’, <tyb> *tib* ‘egg’, <tym> *tim* ‘seed’; Weir 1984: 54, 71; Martins 2005: 339), and mid unrounded vowels of all three languages correspond to Nadëb *ə* (as in <-sët> *-ʃət* ‘to carry’, <atsëm> *ʔac’əbm* ‘at night’, <-gëët> *-k’ət* ‘to stand (sg.)’, <tëg> *tag* ‘tooth’; Weir 1984: 100, 141, 164; Barbosa 2005: 38). Compare the examples in 195–201.

- (195) PDH **ʔi:p* ‘father’ > Dâw <iip> *ʔi:p* || Yuhup <íp> *ʔip* || Hup <íp> *ʔip*³⁹
 (196) PDH **t̥i:p* ‘egg’ > Dâw <tuup> *tũ:p* || Yuhup <típ> *típ* || Hup <tip> *típ*
 (197) PDH **tim* ‘seed’ > Dâw <tum> *tum* || Hup <tīb> *tībm* ‘penis’
 (198) PDH **cě₂:t* ‘to carry on one’s back’ > Dâw <sêet> *ʃët* || Yuhup <sët> *tʃët* || Hup <sët> *tʃët*
 (199) PDH **c’e₁m* ‘night’ > Dâw <çem> *c’em* || Yuhup <sääm> *tʃäbm* || Hup <s’áb> *tʃābm*
 (200) PDH **k’ǝ₂:t* ‘to stand’ > Dâw <kâat> *k’ǝ:t* || Yuhup <këët> *këët* || Hup <k’ët> *kët*
 (201) PDH **təg* ‘tooth’ > Dâw <täg> *trg* || Yuhup <täg> *təgŋ* || Hup <täg> *təgŋ*

As for the oral low unrounded vowels of Dâw, Yuhup, and Hup, there are two possible correspondences in Nadëb: *ɪ* and *a*. The former is found when Dâw *a(:)* corresponds to Yuhup/Dâw *a*, as in <tsäng> *c’əŋ* ‘clay’ (Barbosa 2005: 39), <wäng> *wəŋ* ‘patauí fruit’ (Weir 1984: 230). The latter is found when Dâw *ɛ(:)* (rarely *e(:)*) corresponds to Yuhup/Hup *æ* or *a*, as in <nag’aad> *nāk’a:d* ‘tongue’, <ag> *ʔag* ‘fruit’, <-yat> *-jat* (non-indicative <-yad> *-jad*) ‘to lie on the ground (sg.)’ (Barbosa 2005: 55; Weir 1984: 66, 129, 141). Compare the examples in 202–206. The fact that Dâw patterns with Nadëb in distinguishing the vowels found in 202–203 and 205 suggests that Yuhup and Hup have merged at least some pairs of vowels, possibly by centralizing erstwhile front vowels.

- (202) PDH **c’ax* ‘earth’ > Dâw <çax> *c’ax* || Yuhup <saah> *tʃäħ* || Hup <s’áh> *tʃäh*
 (203) PDH **wax* ‘patauí fruit (*Oenocarpus bataua*)’ > Dâw <wax> *wax* || Yuhup <wəh> *wäh* || Hup <wáh> *wáh*

³⁷ I use the term “non-front” to refer to back and central vowels, since the distinction is not contrastive in any Naduhup language. Note that the non-front unrounded vowels of Dâw are usually described as back (/u/, /ũ/, /ɜ/) except for the low vowels, for which Martins (2004) uses the symbols /a/, /ã/ rather than /a/, /ã/. By contrast, in Yuhup and Hup the respective vowels are usually described as central (/i/, /ĩ/, /ə/, /a/, /ã/), except that Silva & Silva (2012) use the symbols /a/, /ã/ for Yuhup. Martins (2005) uses the symbols /u/, /ũ/, /ɜ/, /a/, /ã/ for all Naduhup languages.

³⁸ Martins (2005) attributes some of the sound correspondences considered here to vowel assimilation (umlaut), and reconstructs a plethora of disyllabic morphemes. While umlaut-like developments are known to have given rise to rich vowel inventories in multiple language groups (Permic, Khanty, Mansi, Germanic), the Dâw-Hup languages retain no traces of the putative second-syllable vowels, and their closest relatives (Nadëb, and possibly Kakua, Nikak, and Puinave) also show a clear preference for CVC-shaped morphemes, suggesting that the template in question must be quite old.

³⁹ The high tone documented in the Hup reflex (Epps 2005: 200) is unexpected.

- (204) PDH **nõh-k'æ₂:d* ‘tongue’ > Dâw <nõr keed> *nõh-k'ê:d* || Yuhup <noh-keén> *nõh-kæ̃ædn*
|| Hup <noh-k'èd> *nõh-kæ̃dn*
- (205) PDH **ʔæg* ‘fruit’ > Dâw <bee eg> *bê: ʔeg* || Yuhup <ag> *ʔagŋ* || Hup <ág> *ʔâgŋ*
- (206) PDH **jæ̃it* ‘to lie on the ground’ > Dâw <yêt> *jet* || Yuhup <yet> *jẵt* || Hup <yét> *djẵt*

There is another piece of evidence suggesting that Yuhup and Hup are less conservative than Dâw with regard to vowel backness. Recall from Table 11 that some correspondences are restricted to etyma with long vowels (for vowel length in PDH, see 5). These include Dâw *ẽ*: ~ Yuhup/Hup *æ̃*, Dâw *uu*: ~ Yuhup/Hup *i*, Dâw *ɾ*: ~ Yuhup/Hup *e*,⁴⁰ Dâw *ɛ*: ~ Yuhup/Hup *æ*, and possibly Dâw *e*: ~ Yuhup/Hup *e* (if 232 is shown to be a wrong etymology). Note that all these correspondences show a front vowel in Yuhup and Hup. In the case of *ẽ*: ~ *æ̃*, one can simply reconstruct PDH **æ̃*: (207–208), whose short counterpart (PDH **æ̃*) must have given rise to the correspondence *ẽ* ~ *ã* (209). The remaining correspondences (*uu*: ~ *i*, *ɾ*: ~ *e*, *ɛ*: ~ *æ*, *e*: ~ *e*) do have competing correspondences with a central vowel in Yuhup/Hup (*uu*: ~ *ɨ*, *ɾ*: ~ *ə*, *ɛ*: ~ *a*, *e*: ~ *ə*), but the fact that they lack short counterparts suggests that they may go back to erstwhile diphthongs (possibly *[iɨ], *[ɨi], *[æi], *[ei]), which would account both for vowel length in Dâw and for the fronting effect in Yuhup and Hup.⁴¹ In this paper, I notate the respective PDH vowels with a subscript ₂, as in **i₂*:, **ɨ₂*:, **æ₂*:, **e₂*: (210–217). These contrast with **i₁*:, **ɨ₁*:, **æ₁*:, **e₁*:, which have centralized reflexes in Yuhup and Hup (218–225).

- (207) PDH **hæ̃:h* ‘star’ > Dâw <’mēer> *ʔmē:h* || Yuhup <wero-meéh> *wæ̃rɔ-mæ̃æ̃h* || Hup <wer(h)o-m’èh> *wæ̃r(h)ɔ-mæ̃h*⁴²
- (208) PDH **jẫ*: ‘non-venomous snake’ > Dâw <yēe> *jễ*: ‘boa’ || Yuhup <yē> *jẵ* || Hup <yé> *njẵ*:
- (209) PDH **næ̃g* ‘fat, oil’ > Dâw <nēg> *nēg* || Yuhup <nag> *nẵŋ* || Hup <nág> *nẫŋ*
- (210) PDH **tî₂:w* ‘path’ > Dâw <tuuw> *tû:w* || Yuhup <tíw> *tîw* || Hup <tìw> *tîw*
- (211) PDH **pî₂:j* ‘cabari plant (*Clathrotropis macrocarpa*)’ > Dâw <puij> *pû:j* || Yuhup <píj> *pîjŋ* || Hup <pìj> *pîjdn*
- (212) PDH **c’ş₂:k* ‘to steal’ > Dâw <çâak> *c’ş:k* || Hup <s’ék> *tşék*
- (213) PDH **jş₂*: ‘to enter’ > Dâw <yâa> *jş*: ‘to come back’ || Yuhup <yē> *jě* || Hup <yé> *djê*:
- (214) PDH **wæ̃₂:d* ‘to eat’, **wæ̃₂:d* ‘food’ > Dâw <weed> *wê:d*, <weed> *wě:d* || Yuhup <wen> *wẵdn*, <wén> *wẵdn* || Hup <wéd> *wẵdn*, <wèd> *wẵdn*
- (215) PDH **k’æ̃₂:g* ‘bone; to choke on a fishbone’ > Dâw <keeg> *k’ê:g* ‘bone; to choke’ || Yuhup <keeg> *kẵgŋ* ‘to choke on a fishbone’ || Hup <k’ég> *kẵgŋ* ‘bone’; <k’ég> *kẵgŋ* ‘to choke’
- (216) PDH **kê₂*: ‘feather, wing’ > Dâw <xêe> *xê*: || Yuhup <kě> *kě* || Hup <ké> *kê*: ‘wing’
- (217) PDH **cě₂:m* ‘tick’ > Dâw <sêem> *şê:bm*⁴³ || Yuhup <táh-sém> *tâh-tşêbm* ‘large tick sp.’ || Hup <tah-sèb> *tah-tşêbm*

⁴⁰ In just one cognate set, Dâw *ɾ* (rather than *ɾ*:) corresponds to Yuhup/Hup *e*: Dâw <jâw> *jşw* ‘nine-banded armadillo’ ~ Yuhup <yéw> *jêw* ‘greater long-nosed armadillo’, Hup <yèw> *djêw* ‘greater long-nosed armadillo, nine-banded armadillo’. However, the falling tone in Yuhup and the rising tone in Hup both support the reconstruction of a long vowel in PDH. I reconstruct PDH **jş₂:w* (with an unknown tonal value) and posit irregular shortening in Dâw.

⁴¹ It is possible that the putative diphthongs are marginally preserved in some varieties of Yuhup. For example, Ospina Bozzi (2002: 104) documents the Yuhup reflex of **tî₂:g* ‘firewood; tree (of a given species)’ as [têjŋ]. However, other sources on the language attest just <tég> [têjŋ].

⁴² The loss of *h* is typical of the Western and Eastern dialect areas (Epps 2005: 87). The form <wirom’eh-têh> *wirom-mæ̃h-têh* is documented in the Umari Norte dialect area (Epps 2005: 297).

⁴³ In Epps *et al.* (2018), this is attested as <seem> *şê:bm*, which must be a typo (cf. Martins 2004: 17; Martins 2005: 333).

- (218) PDH **kĩ₁.t* ‘to cut’ > Dâw <xuut> *xu:t* ‘to cut, to tear’ || Hup <kít> *kít* ‘to cut (with an ax, a machete, etc.)’
- (219) PDH **wĩ₁.d* ‘to arrive, to reach’ > Dâw <wũud> *wũ:d* || Yuhup <win> *wĩdn* ‘to appear (of fish during spawning season)’, <yãh win> *jãh-wĩdn* ‘to reach’, <wit-> *wĩt-* ‘to arrive (in serial constructions)’ || Hup <wíd> *wĩdn* ‘to appear (of fish during spawning season); to arrive, to reach (in serial constructions)’
- (220) PDH **cĩ₁.k* ‘butt, buttocks’ > Dâw <sâak> *ĩ:k* || Hup <sàk> *tĩk*
- (221) PDH **bĩ₁.h* ‘to spill (liquid)’ > Dâw <bâa> *bĩ:* || Yuhup <bääh> *bĩh* || Hup <b’äh> *mbĩh*
- (222) PDH **c’ĩ₁.p* ‘to break by pulling (intr.)’ > Dâw <çee> *c’ĩ:p* || Hup <s’áp> *tĩp*
- (223) PDH **cĩ₁.?* ‘to carry on one’s hip’ > Dâw <see> *ĩ:?* || Yuhup <sa’> *tĩ?* || Hup <sá’> *tĩ?*
- (224) PDH **jĩ₁.n* ‘to hide’ > Dâw <yêen> *jĩ:dn* || Yuhup <yän> *jĩdn* || Hup <yád> *dĩdn*
- (225) PDH **tĩ₁.ç* ‘to cut (with an ax, a machete, etc.)’ > Dâw <têes> *tĩ:ç* || Yuhup <táč> *tĩç*

I extend the notation ₁ and ₂ to other pairs of correspondence sets where Dâw shows identical reflexes, whereas Yuhup and Hup show different vowels of the same height (central vowels in ₁-correspondences, and front vowels in ₂-correspondences). Therefore, I reconstruct PDH **i₁* for Dâw *i* ~ Yuhup/Hup *i* (226–227); **i₂* for Dâw *i* ~ Yuhup/Hup *i* (228–229); **e₁* for Dâw *e* ~ Yuhup/Hup *ə* (230–231); and **e₂* for Dâw *e* ~ Yuhup/Hup *e* (232).

- (226) PDH **bĩ₁.l* ‘common squirrel monkey’ > Dâw <bij> *bĩ* || Hup <b’ij> *mbĩjdn*
- (227) PDH **tĩ₁.t* ‘rope, cord’ > Dâw <tít> *tít* || Yuhup <tít> *tĩt* ‘cord-like, vine’ || Hup <tít> *tĩt* ‘cord-like, vine’
- (228) PDH **c’ĩ₂.d* ‘to wash’ > Dâw <çid> *c’id* || Yuhup <siin> *tĩĩdn* ‘to wash, to clean’ || Hup <s’id> *tĩdn*
- (229) PDH **wĩ₂.d* ‘to hug’ > Dâw <wid> *wid* ‘to hug strongly’ || (?) Yuhup <win> *wĩdn* ‘to wind (a thread, a fishing line)’ *wĩdn* || Hup <wíd> *wĩdn*
- (230) PDH **te₁.n* ‘to fish with *tĩmbó*’ > Dâw <tên> *tedn* || Yuhup <tän> *tĩdn* || Hup <tád> *tĩn*
- (231) PDH **c’e₁.m* ‘night’ > Dâw <çem> *c’em* || Yuhup <sääm> *tĩĩbm* || Hup <s’áb> *tĩbm*
- (232) PDH **d.de₂.m*, **n.ne₂.m*, or **n.ne₂.b* ‘round’ > Dâw <lêm> *lebm* ‘round and small (e.g., eye, turtle, mic head)’, <nêm> *nebm* ‘spherical (e.g., shell, pan, head, *açai* fruit stone)’, or <nêb> *neb* ‘round and small (e.g., seeds, fruit)’ || Hup <dërêb> *nderêbm*

I also use a subscript digit in PDH **ĩ₂* (Dâw *ĩ* ~ Yuhup/Hup *ĩ*), as in 169, but note that evidence for reconstructing **ĩ₁* (Dâw *ĩ* ~ Yuhup/Hup *ĩ*) is extremely weak, since the only known example of the latter correspondence (119) presents further irregularities. If the etymology in 119 is shown to be wrong, one can rewrite **ĩ₂* as **ĩ*.

The remaining vowel correspondences involving vowels of matching height and nasality can be straightforwardly reconstructed as PDH **i* (Dâw *iu* ~ Yuhup/Hup *i*), **ĩ* (Dâw *ĩu* ~ Yuhup/Hup *ĩ*), **ẽ* (Dâw *ĩu*: ~ Yuhup/Hup *ĩ*), **ə* (Dâw *ɣ* ~ Yuhup/Hup *ə*), **a* (Dâw *a* ~ Yuhup/Hup *a*), **a*: (Dâw *a*: ~ Yuhup/Hup *a*:), **ã* (Dâw *ã* ~ Yuhup/Hup *ã*), **ã*: (Dâw *ã*: ~ Yuhup/Hup *ã*:), **æ* (Dâw *ɛ* ~ Yuhup/Hup *a*), **i*: (Dâw *i*: ~ Yuhup/Hup *i*), **ĩ*: (Dâw *ĩ*: ~ Yuhup/Hup *ĩ*). No subscript digits are needed for these vowels, since Yuhup and Hup lack competing correspondences with a differing backness value (except for Dâw *ã* ~ Yuhup/Hup *ã*, on which see below). Some examples follow in 233–255.

- (233) PDH **p.xit* ‘wild banana’ > Dâw <xut> *xut* || Yuhup <wihit> *wihĩt* || Hup <pihít> *pihĩt*
- (234) PDH **c.bix* ‘bat’ > Dâw <bux> *bux* || Hup <sib’ih> *tĩb’ih*
- (235) PDH **hĩp* ‘to grate’ > Dâw <rũp> *hũp* || Yuhup <hĩp> *hĩp* || Hup <hĩp> *hĩp*
- (236) PDH **pĩn* ‘thick (of liquid)’ > Dâw <pũn> *pũn* || Hup <pĩn> *pĩn*

- (237) PDH **k'ê:m* 'to flood' > Dâw <kũum> *k'ũ:m* || Hup <k'ím> *k'ím*
- (238) PDH **mî:h* 'ucuqui fruit (*Pouteria ucuqui*)' > Dâw <mũu> *mũ:* || Yuhup <mîh> *mîh* || Hup <mîh> *mîh*
- (239) PDH **k'ɣɿ* 'to remove with one's fingernail' > Dâw <kâj> *k'ɣɿ* || Yuhup <kääj> *kääjɿ* || Hup <k'áj> *kájɿ*
- (240) PDH **cɔg* 'leftovers' > Dâw <sâg> *ɣɔg* || Hup <ság> *tʃɔgɿ*
- (241) PDH **wab* 'jirau platform' > Dâw <wab> *wab* || Yuhup <wam> *wăbm* || Hup <wáb> *wăbm*
- (242) PDH **paj* 'thing' > Dâw <pay> *paj* || Hup <páy> *páj* 'belongings, goods'
- (243) PDH **dă:k* 'to hang (intr.)' > Dâw <daak> *dă:k* (compare <dak> *dak* 'to hang (tr.)') || Yuhup <daak> *dăăk* || Hup <d'ák> *dák*
- (244) PDH **tâ:* 'to be cooked (ready), spicy' > Dâw <taa> *tâ:*⁴⁴ 'to be half-cooked, spicy' || Yuhup <ta> *tă* || Hup <tá> *tâ:* 'to be cooked (ready)'
- (245) PDH **câ:j* 'poisonous arthropod sp.' > Dâw <saay> *ſâ:j* 'scorpion; bee sp.' || Yuhup <sáy> *tʃâj* 'centipede' || Hup <sày> *tʃâj* 'centipede'
- (246) PDH **ʔâh* 'I' > Dâw <âr> *ʔâh* || Yuhup <âh> *ʔâh* || Hup <âh> *ʔâh*
- (247) PDH **ʔâm* 'you (sg.)' > Dâw <âm> *ʔâm* || Yuhup <âm> *ʔâm* || Hup <ám> *ʔâm*
- (248) PDH **că:h* ~ **câh* 'to think, to feel' > Dâw <sâar> *ſă:h* 'to feel, to think', <sâr> *ſâh* 'to think, to suppose, to teach' || Hup <sâh> *tʃâh* 'to accuse'
- (249) PDH **nă:m* 'curare poison for arrows' > Dâw <nâam> *nă:m* || Hup <nâm> *nă:m*
- (250) PDH **xæp* 'to scrape' > Dâw <xep> *xep* 'to peel, to scrape' || Hup <háp> *háp* 'to scrape (e.g., curare)'
- (251) PDH **dæp* 'meat' > Dâw <dep> *dɛp* || Yuhup <daap> *dăăp* || Hup <d'áp> *ndăp*
- (252) PDH **cî:ʔ* 'to urinate' > Dâw <sii> *ſî:ʔ* || Yuhup <si> *tʃî:ʔ* || Hup <sí> *tʃî:ʔ*
- (253) PDH **ʔî:p* 'father' > Dâw <iip> *ʔî:p* || Yuhup <íp> *ʔîp* || Hup <íp> *ʔîp*
- (254) PDH **mî:n* 'ingá fruit (*Inga spp.*)' > Dâw <mĩin> *mî:n* || Yuhup <mín> *mîn* || Hup <mìn> *mîn*
- (255) PDH **-mî:* 'branch' > Dâw <bee mĩi> *bê:-mî* || Yuhup <-mi> *-mî* 'water course' || Hup <-mi> *-mî:* 'water course'

Two residual correspondences with matching nasality values remain unassigned to any PDH vowel. Dâw *e(:)* ~ Yuhup/Hup *æ* is a rare example of a correspondence with unmatching height values, occurring in two verbal roots and their causatives (256–257). I reconstruct the diphthong **æ(:)i* to account for it. The correspondence between Dâw *ă(:)* and Yuhup/Hup *ẽ* is supported by just one unproblematic cognate set (258), in addition to a dubious etymology with an irregular correspondence in its final consonant (113); I use the *ad hoc* symbol **Ă(:)* for the respective PDH vowel. It is technically possible to reconstruct a distinction between **ă₁(:)* (> Dâw *ă(:)*, Yuhup/Hup *ă*) and **ă₂(:)* (> Dâw *ă(:)*, Yuhup/Hup *ẽ*), but I do not adopt such notation here, since there is otherwise no evidence for positing an opposition between two kinds of low non-front vowels (**a₁*/**a₂*, **a₁*/**a₂*), and because there is a similar correspondence between Dâw *a* and Yuhup/Hup *ã*, on which see below.

- (256) PDH **ʔæit* 'to lie on the ground', **ʔæ:it* 'to lay on the ground' > Dâw <yêt> *jet*, <yêet> *ʔjê:t* || Yuhup <yet> *jă:t*, <yeet> *jăă:t* || Hup <yét> *djă:t*, <y'et> *djă:t*
- (257) PDH **pæit* 'to break (intr.)', **pæ:it* 'to break (tr.)' > Dâw <pêt> *pet*, <pêet> *pê:t* 'to break (e.g., a bone or a stick)' || Hup <pét> *pă:t* 'to cut with one's teeth or mandibles'

⁴⁴ Epps *et al.* (2018) document this verb with a rising tone, but this is likely a typo, since CV-shaped roots are reported to obligatorily carry falling tone by Martins (2004: 91), who documents the root in question as <taa> *tâ:*.

- (258) PDH $*\hat{A}:m$ ‘wife’ > Dâw <ãam> $\hat{a}:m$ || Yuhup <ém> $\hat{e}m$ || Hup <te’ém> $t\tilde{e}-\hat{e}m$
 ‘daughter-in-law’ (literally ‘son’s wife’)

Table 12 lists the full inventory of PDH vowels reconstructed in this subsection, and Table 13 summarizes their reflexes in the daughter languages.

Dâw	front unrounded		back unrounded		back rounded	front unrounded		back unrounded		back rounded
Yuhup/Hup	central	front	central	front		central	front	central/back	front	
high	$*i_1$ —	$*i_2$ $*i:$	$*i$ $*i_1:$	— $*i_2:$	$*u$ $*u:$	$(*i_1)$ —	$*i_2$ $*i:$	$*i$ $*i:$	— —	$*ü$ $*ü:$
high-mid	$*e_1$ $*e_1:$	$*e_2$ $*e_2:$	$*e$ $*e_1:$	— $*e_2:$	$*o$ $*o:$					
low-mid	$*æ$ $*æ_1:$	— $*æ_2:$			$*ɔ$ $*ɔ:$	$*æ$ —	— $*æ:$	$*ɜ$ $*ɜ:$	— —	$*ɔ$ $*ɔ:$
low			$*a$ $*a:$					$*ā$ $*ā:$		
unclear		$*æ_1$ $*æ_1:$			$*O$ $*O:$					$(*\tilde{A})$ $*\tilde{A}:$

() = correspondence is attested only in cognate sets with additional irregularities

Table 12. PDH vocalic inventory

4.2. Unexpected nasality in Yuhup and Hup

In a couple dozen etymologies, Dâw oral vowels correspond to nasal vowels in Yuhup and Hup. Nadëb cognates show oral vowels in such cases, suggesting that the nasality in Yuhup and Hup is secondary: <-gä> - $k'a$ (non-indicative <-gää> - $k'a:$) ‘to lie in a hammock (sg.)’, <yón> $j\hat{o}dn$ ‘anteater’ (Weir 1984: 245; Barbosa 2005: 45). I have been unable to identify an environment that might have conditioned such a sound change. In my reconstructions, I represent the putative source of nasality in Yuhup and Hup by means of a tilde after the vowel, as in 259–276.

- (259) PDH $*k'a\sim?$ ‘to lie in a hammock’ > Dâw <ka> $k'a?$ || Yuhup <kā> $k\tilde{a}?$ ‘to lie in a hammock, to hang’ || Hup <k’ā> $k\tilde{a}?$ ‘to lie in a hammock, to hang’
 (260) PDH $*ja\sim h$ ‘uacu fruit (*Monopteryx uacuu*)’ > Dâw <yar> jah || Yuhup <yāh> $j\tilde{a}h$ || Hup <yāh> $nj\tilde{a}h$
 (261) PDH $*p\tilde{a}\sim h$ ‘to hear, to understand’ > Dâw <paar> $p\tilde{a}h$ || Yuhup <pāh> $p\tilde{a}h$ ‘to hear, to understand’ || Hup <hipáh> $hi-p\tilde{a}h$
 (262) PDH $*pa\sim t$ ‘hair’ > Dâw <pat> pat || Yuhup <pāt> $p\tilde{a}t \sim$ <pāhat> $p\tilde{a}h\tilde{a}t$ ⁴⁵ || Hup <pát> $p\tilde{a}t$ ‘hair, feather’
 (263) PDH $*ca\sim n$ ‘pubic hair’ > Dâw <san> $fadn$ || Yuhup <sān> $t\tilde{f}\tilde{a}n$ || Hup <sán> $t\tilde{f}\tilde{a}n$
 (264) PDH $*wa\sim?$ ‘vulture’ > Dâw <wa> $wa?$ || Yuhup <wā> $w\tilde{a}?$ || Hup <wá> $w\tilde{a}?$ ‘black vulture’
 (265) PDH $*ni\sim g$ ‘you (pl.)’ > Dâw <nug> nug || Yuhup <nig> $n\tilde{i}g$ || Hup <níg> $n\tilde{i}g$
 (266) PDH $*k.m\tilde{a}\sim n$ ‘to hug’ > Dâw <mūun> $m\tilde{u}dn$ || Yuhup <yāh min> $j\tilde{a}h-m\tilde{i}n$ ‘to grab in one’s arms’ || Hup <kimín> $k\tilde{i}m\tilde{i}n$ ‘to hug with both arms’
 (267) PDH $*\hat{A}\sim h$ ‘fire ant’ > Dâw <uu> $\hat{u}u:$ || Yuhup <íh> $\hat{i}h$ || Hup <ìh> $\hat{i}h$

⁴⁵ The variant <pāhat> $p\tilde{a}h\tilde{a}t$, attested in Yuhup by Silva & Silva (2012: 240), is of unclear origin.

PDH	Dâw	Yuhup and Hup
*a, *a:	a, a:	a
*ə, *ə₁:	ɤ, ɤ:	ə
*ə₂:	ɤ:	e
*i, *i₁:	u, u:	i
*i₂:	u:	i
*ɔ, *ɔ:	ɔ, ɔ:	ɔ
*o, *o:	o, o:	o
*u, *u:	u, u:	u
*æ, *æ₁:	ɛ, ɛ:	a
*æ₂:	ɛ:	æ
*æj, *æj₁:	e, e:	æ
*e₁, *e₁:	e, e:	ə
*e₂, *e₂:	e, e:	e
*i₁	i	i
*i₂, i:	i, i:	i
*O, *O:	ɔ, ɔ:	o

PDH	Dâw	Yuhup and Hup
*ã, *ã:	ã, ã:	ã
*ẽ, *ẽ:	ẽ, ẽ:	õ
*ĩ, *ĩ:	ũ, ũ:	ĩ
*õ, *õ:	õ, õ:	õ
*ũ, *ũ:	ũ, ũ:	ũ
*æ̃, *æ̃:	ẽ, ẽ:	ã, æ̃
(*ĩ₁)	(ĩ)	(ĩ)
*ĩ₂, *ĩ:	ĩ, ĩ:	ĩ
(*Ã), *Ã:	(ã), ã:	(æ̃), æ̃:

() = correspondence is attested only in cognate sets with additional irregularities

Table 13. PDH vowels and their reflexes

- (268) PDH *cʷ~m ‘to bathe’ > Dâw <çom> cʷbm || Yuhup <sõom> tʃõõm || Hup <s’óm> tʃõm
- (269) PDH *jɔ~n ‘anteater sp.’ > Dâw <yon> jɔdn ‘giant anteater’ || Yuhup <yõn> jõn ‘collared anteater’ || Hup <yón> njõn ‘anteater sp. (brown, no collar, lives side-by-side with collared anteater)’
- (270) PDH *tɔ~h ‘white-lipped peccary’, *tɔ~h-mě:t ‘collared peccary’ > Dâw <tor> tɔh, <tor mêt> tɔh-mêt || Yuhup <tõh> tõh, <tõh-mét> tõh-mæt || Hup <tóh> tóh ‘white-lipped peccary’
- (271) PDH *dɔ~n ‘to lick’ > Dâw <doon> dɔ:n || Yuhup <noon> nõõn
- (272) PDH *m.xu~p ‘right (side)’ > Dâw <sôob xup> jô:b xup ‘right hand’⁴⁶ || Yuhup <mu-huk> mũhũk-⁴⁷ || Hup <muhùp> mũhũp
- (273) PDH *pæ~n ‘sloth’ > Dâw <pen> pɛdn || Yuhup <pãn> pãn ‘Linnaeus’s two-toed sloth’ || Hup <pán> pãn
- (274) PDH *dæ~n ‘chigoe flea’ > Dâw <den> dɛdn || Yuhup <naan> nãñ ‘chigoe flea; flea’

⁴⁶ The form <xup> xup is attested in Martins (2005: 277) and Obert (2019: 96). Elsewhere, one finds the unexpected reflex <xub> xub (Epps *et al.* 2018; Obert 2019: 109).

⁴⁷ This term must receive a suffix indicating the distance. The root-final velar consonant is irregular; it may have emerged due to the influence of <muhúk> mũhũk ‘chief’.

- (275) PDH **pê:~m* ‘to sit’ > Dâw <peem> *pê:bm* || Yuhup <pēm> *pǣm* || Hup <pém> *pǣm*
 (276) PDH **cĭ:~p* ‘oriole sp. (*Icterus chryscephalus*)’ > Dâw <siip> *ſi:p* || Yuhup <śip> *tſip*
 ‘oriole’s song’

In a handful of cognate sets, one finds correspondences that cannot be equated with any oral vowels reconstructed in subsection 4.1, but that have close parallels among the nasal vowels. One such correspondence is Dâw *a:* ~ Yuhup and Hup *ǣ*, similar to the one derived from PDH **ǣ:*, but with an oral reflex in Dâw (277–280). Another correspondence is Dâw *a(:)~* Yuhup and Hup *ǣ*, similar to the one derived from PDH **ǣ(:)~*, but likewise with an oral reflex in Dâw (281–283). In this paper, I reconstruct **ǣ:*~ and **A(:)~*, respectively, even though no **ǣ:* or **A(:)* can be reconstructed. If such vowels were present in PDH, they must have merged with other vowels (possibly **a*) in all daughter languages. Example 280 instantiates what looks like a non-productive alternation in Yuhup.

- (277) PDH **wǣ:~m* ‘red squirrel sp.’ > Dâw <waam> *wâ:bm* || Yuhup <wǫm> *wǫm* ‘southern Amazon red squirrel’ || Hup <wòm> *wǫm* ‘northern Amazon red squirrel’
 (278) PDH **c.xǣ:~m* ‘crab’ > Dâw <xaam> *xâ:bm* || Yuhup <sōhóm> *tſhǫm* || Hup <sohòm> *tſhǫm*
 (279) PDH **wǣ:~n* ‘to chase, to follow’ > Dâw <waan> *wâ:dn* (compare <wan> *wadn* ‘to expel’) || Yuhup <wǫn> *wǫn* || Hup <wón> *wǫn*
 (280) PDH **xǣ:~n* ‘to vomit’ > Dâw <xaan> *xâ:dn* || Yuhup <hǫn> *hǫn* ‘to vomit’ (cf. <han> *hǫdn* ‘to regurgitate, to vomit an entire thing’) || Hup <hón> *hǫn*
 (281) PDH **jA~?* ~ **jǣ:~?* ‘to roast’ > Dâw <ya> *ja?* ~ <yaa> *jǣ?* || Yuhup <yē> *jǣ?* || Hup <yé> *njǣ?*
 (282) PDH **kǣ:~?* ‘to bury’ > Dâw <xaa> *xǣ?* || Yuhup <kē> *kǣ?*⁴⁸ || Hup <ké> *kǣ?*
 (283) (?) PDH **ǣA~?* ~ **ǣA~* ‘cold’ > Dâw <ba> *ba?* || Hup <m’é> *mǣ?*

In 284, Yuhup agrees with Dâw in showing an oral vowel, suggesting that Hup must have acquired nasality in an independent, irregular development.

- (284) PDH **kǣ:n* ‘to toast’ > Dâw <xeen> *xê:dn* || Yuhup <ken> *kǣdn* || Hup <kén> *kǣn*

4.3. Unexpected nasality in Dâw

A few cognate sets show a nasal vowel in Dâw and oral vowels in Yuhup and Hup (285–289). Once again, whenever a Nadëb cognate is available, it also shows an oral vowel, as in <-óót> -*ɔ:t* (non-indicative <-ód> -*ɔd*) ‘to cry’ (Weir 1984: 83, 102). I believe that these words have irregularly acquired nasality in Dâw, but for the time being I give alternative reconstructions with an oral vowel and a nasal vowel until more conclusive evidence is identified.

- (285) (?) PDH **hɔt* ~ **hǫt* ‘far’ > Dâw <rôt> *hǫt* || Yuhup <hot> *hǫt* || Hup <hót> *hót* ‘at some distance (neither near nor far)’
 (286) (?) PDH **ɔ:t* ~ **ɔ:t* ‘to cry’ > Dâw <ōot> *ɔ:t* || Yuhup <ot> *ɔ:t* || Hup <ót> *ɔ:t*
 (287) (?) PDH **ɔ:k* ~ **ɔ:k* ‘giant armadillo’ > Dâw <ōok> *ɔ:k* || Hup <òk> *ɔ:k*
 (288) (?) PDH **k’i₁η* ~ **k’i₁η* ‘to shoot with an arrow’ > Dâw <kính> *k’i₁η* || Yuhup <kiig> *k’i₁η* || Hup <k’ig> *k’i₁η*
 (289) (?) PDH **c’ip* ~ **c’ip* ‘tucum palm (*Astrocaryum vulgare*)’ > Dâw <çũp> *c’ũp* || Yuhup <siip> *tſiip*

⁴⁸ The absence of laryngealization in the Yuhup reflex, as attested in Silva & Silva (2012: 167) and Ospina Bozzi (2002: 363), is unexpected.

5. Tones and vowel length

Dâw, Yuhup, and Hup are tonal languages. In Dâw, only long vowels carry tone (either rising or falling), whereas short vowels are toneless; in fact, Martins (2004: 55–56) posits a triple opposition between toneless, rising-toned and falling-toned syllables, where vowel length is considered allophonic. In Yuhup and Hup, lexical morphemes typically carry falling or rising tone on their stressed (final) syllable; in Hup, the falling tone has a high allotone preceding voiceless codas (cf. Epps 2005: 81–82, who considers that the high tone is the underlying one and the falling tone is an allotone). Syllables with a contour tone are automatically lengthened in Yuhup (a fact not represented in my broad transcriptions). In Hup, open stressed syllables are lengthened, too (in this case, I opted to represent the lengthening in my transcriptions).

There is evidence suggesting that Proto-Yuhup–Hup may not necessarily have been tonal, since the tones of Yuhup and Hup do not correspond to Dâw tones, but rather to Dâw vowel length. Furthermore, the tonal opposition found in Dâw is also suspect, since, as shown by Martins (2004: 83), in most cases rising tone is associated with voiceless codas, and falling tone with voiced codas in Dâw. However, there are words with voiced codas and rising tone in Dâw. In this paper, I follow Martins (2004: 78–79) in positing a tonal distinction for Dâw and, consequently, for PDH, however low its functional load may turn out to be.

The basic correspondence pattern, shown in Table 14, is clearly seen in nouns and other lexical morphemes other than verbs. In this paper, I assume that Dâw is conservative regarding vowel length and tones, since the tones of Yuhup and Hup can be predicted based on those of the Dâw cognate.

	PDH	Dâw	Proto-Yuhup–Hup	Yuhup	Hup
*CVC	short (toneless)	short (toneless)	short (toneless)	long, rising	short, falling/high
*C $\check{V}$:C	long, rising tone	long, rising tone	long (toneless)	long, falling	short, rising
*C $\hat{V}$:C	long, falling tone	long, falling tone			
*C $\hat{V}$:	long, falling tone	long, falling tone	short (toneless)	long, rising	long, falling

Table 14. Tones

The absence of tone/vowel length in Dâw corresponds to rising tone in Yuhup, and falling (before voiced codas) or high (before voiceless codas) tone in Hup (Barboza 2016), as shown in 290–303.

(290) PDH *kɔw ‘pepper’ > Dâw <xow> xɔw || Yuhup <kow> kɔw || Hup <ków> kôw

(291) PDH *dom ‘acará fish’ > Dâw <dôm> dobm || Yuhup <dööm> doǫbm || Hup <d’ôb> ndôbm

(292) PDH *cug ‘hummingbird’ > Dâw <sug> fug || Yuhup <sug> tǫgɨ || Hup <súg> tǫgɨ

(293) PDH *xuj ‘fish sp.’ > Dâw <xuy> xuj || Yuhup <huy> ‘piaba fish sp.’ || Hup <húy> húj ‘piaba fish sp.’

(294) PDH *c’ɔp ‘fly’ > Dâw <çop> c’ɔp || Yuhup <soop> tǫǫp ‘maggot’ || Hup <s’óp> tǫǫp ‘fruit fly sp. (big and green)’

(295) PDH *tət ‘reptile or amphibian sp.’ > Dâw <tât> tɔt ‘salamander’ || (?) Yuhup <tât> tât ‘grasshopper sp.’ || Hup <tât> tât ‘lizard sp. (*Bachia spp.*); (?) insect sp. (similar to a bed bug)’

- (296) PDH *k'ɔt 'uncle, father-in-law' > Dâw <kot> k'ɔt || Yuhup <koot> kɔʔt 'maternal uncle, father-in-law' || Hup <k'ót> kɔʔt 'maternal uncle, father-in-law'
- (297) PDH *bɔk 'skin, bark, shell (of a turtle); pan' > Dâw <bok> bɔk 'shell, cover; pan' || Yuhup <book> bɔʔk 'skin, bark; pan' || Hup <b'ók> mbɔʔk
- (298) PDH *dɔk 'fish sp.' > Dâw <dok> dɔk 'fish sp. (found in creeks)' || Yuhup <dook> dɔʔk || Hup <d'ók> ndɔʔk 'acarapuru fish (*Erythrinus erythrinus*)'
- (299) PDH *jɔɪx 'frog sp.' > Dâw <yâx> jɔx || Yuhup <yäh> jəh || Hup <yáh> djəh 'frog sp. (large, edible, lives on river banks and on dry land)'
- (300) PDH *kɔx 'courbaril' > Dâw <xox> xɔx || Yuhup <koh> kəh || Hup <kóh> kəh
- (301) PDH *duʔ 'afternoon' > Dâw <du> duʔ 'afternoon, sunset' || Hup <d'ú> ndúʔ
- (302) PDH *woh 'frog sp.' > Dâw <wôx> woh || Yuhup <wöh> wəh 'toad sp.' || Hup <wóh> wəh 'frog sp. (small, terrestrial, and edible)'
- (303) PDH *duh 'stick' > Dâw <bee dur> bē-duh || Yuhup <tëg-duuh> teg-duuh ~ <tëguh> teg-uh⁴⁹ 'tree' || Hup <tëg-d'úh> teg-ndúh 'tree'

In CVC-shaped morphemes, the falling tone in Dâw corresponds to falling tone in Yuhup, and rising tone in Hup (Barboza 2016). In Dâw, all such morphemes have voiced codas, since voiceless codas are incompatible with falling tone in that language, with very few exceptions (Martins 2004: 83–84). In my reconstruction, falling tone was compatible with the voiceless coda *h in PDH, which is lost in Dâw. Some examples follow in 304–318.

- (304) PDH *xô:b 'woodpecker' > Dâw <xôob> xô:b || Yuhup <hóm> hôbm || Hup <hòb> hōbm
- (305) PDH *pê:d 'cunuri fruit (*Micrandra spruceana*)' > Dâw <pâad> pɛ:d || Yuhup <pén> pēdn || Hup <pèd> pēdn
- (306) PDH *pê:j 'umari fruit (*Poraqueiba* spp.)' > Dâw <peej> pɛ:j || Yuhup <péj> pēj || Hup <pèj> pēj
- (307) PDH *tô:j 'nose' > Dâw <tôoj> tô:j || Yuhup <tôj> tōj || Hup <tòj> tōj
- (308) PDH *nô:g 'uirapixuna/vapixuna/iuapixuna fruit (*Protium* spp.)' > Dâw <nôog> nô:g || Yuhup <dóg> dōg || Hup <dòg> ndōg
- (309) PDH *k'ô:g 'collared titi monkey' > Dâw <kôog> k'ô:g || Yuhup <köög> kōōg || Hup <k'ög> kōg
- (310) PDH *c'î:w 'peach palm fruit' > Dâw <çuw> c'û:w || Yuhup <siw> tɕiɰw || Hup <s'îw> tɕiɰw
- (311) PDH *pɔ:j 'catfish' > Dâw <pooy> pɔ:j || Yuhup <hóp-póy> hōp-pōj || Hup <hōp-pòy> hōp-pōj
- (312) PDH *bô:j 'traíra fish (*Hoplias* spp.)' > Dâw <bôoy> bô:j || Yuhup <bööy> bōōj || Hup <b'öy> mbōj
- (313) PDH *pê:j 'thunder' > Dâw <pēey> pɛ:j || Yuhup <pēy> pēj || Hup <péy> pēj
- (314) PDH *jû:n 'clothes' > Dâw <yuun> jû:dn || Yuhup <yún> jūdn || Hup <yùd> djūdn
- (315) PDH *j.xû:n 'giant anteater'⁵⁰ > Dâw <xuun> xû:dn
- (316) PDH *bê:h 'tree'⁵¹ > Dâw <bee> bē:

⁴⁹ The variant <tëguh> teg-uh, attested in Yuhup by Silva & Silva (2012: 284), is an irregular reduced form of <tëg-duuh> teg-duuh.

⁵⁰ No cognates in Yuhup or Hup are known, but compare Nadëb <yahuun> jahu:dn 'giant anteater' (Barbosa 2005: 39).

⁵¹ The consonant *h* resurfaces in Dâw when the tone of the noun is changed: <tum beer> tūm bē:h 'two trees'. No cognates in Yuhup or Hup are known, but compare Nadëb <baah> ba:h, pl. <b'aah> bā:h 'tree' (Epps & Obert 2021: 258).

- (317) PDH **c.6ô:h* ‘taylor’ > Dâw <’môo> *’mô:* || Yuhup <moóh> *môôh* || Hup <som’òh> *tjômôh*
 (318) PDH **n.wô:h* ~ **d.wô:h* ‘cheek’ > Dâw <woo> *wô:* || Yuhup <now-wóh> *nôw-wôh* ||
 Hup <dowòh> *ndôwôh*

The rising tone, found in CVC-shaped morphemes only, also corresponds to falling tone in Yuhup, and rising tone in Hup (Barboza 2016). In PDH, like in Dâw, such morphemes may have voiceless or voiced codas. Some examples follow in 319–337.

- (319) PDH **wô:p* ‘tumor’ > Dâw <woop> *wô:p* || Yuhup <wóp> *wôp* ‘hernia in the groin or testicles’ || Hup <wòp> *wôp* ‘a kind of cyst’
 (320) PDH **jô:p* ‘*anujá* fish (*Trachycorystes galeatus*)’ > Dâw <yooop> *yô:p* || Yuhup <yóp> *jôp* || Hup <yòp> *djôp* ‘fish sp. (~ 20 cm long; similar to the *daquiru* fish)’
 (321) PDH **k’ă:t* ‘leaf’ > Dâw <bee keet> *bê:k’ě:t* || Yuhup <su’-keet> *tʃuʔ-kææt* || Hup <suk’èt> *tʃukææt*, <k’èt> *kæt* ~ <k’ét> *kæt*⁵²
 (322) PDH **wô₂:t* ‘bird’ > Dâw <tâwâat> *tr-wô:t* || Yuhup <wét> *wêt* || Hup <wèt> *wêt* ‘plain-breasted ground dove’⁵³
 (323) PDH **bâ:k* ‘bunch, cluster’ > Dâw <baak> *bă:k* || Yuhup <baák> *baák* || Hup <b’âk> *mbăk*
 (324) PDH **tă:k* ‘resin, sap’ > Dâw <taak> *tă:k* ‘rubber’ || Yuhup <ták> *tâk* || Hup <tàk> *tâk*
 (325) PDH **wô₃:k* ‘*caatinga* (place with semi-arid vegetation)’ > Dâw <wâak> *wô:k* || Yuhup <wák> *wâk* || Hup <wàk> *wăk*
 (326) PDH **k’ũ:k* ‘bundle, beam’ > Dâw <kuuk> *k’ũ:k* || Yuhup <kúk> *kûk*⁵⁴ || Hup <k’uk> *kuk*⁵⁵
 (327) PDH **bă:?* ‘cassava bread’ > Dâw <baa’> *bă:?* || Hup <b’â’> *mbă?*
 (328) PDH **bô:?* ‘peacock bass’⁵⁶ > Dâw <bôo> *bô:?* || Yuhup <böö’> *bgô?* || Hup <b’ô’> *mbô?*
 (329) PDH **hô:h* ‘white-throated tinamou’ > Dâw <hôoh> *hô:h* || Hup <moh-hôh> *môh-hôh*
 (330) PDH **dũ:ç* ‘*timbo* vine’ > Dâw <duus> *dũ:f*⁵⁷ || Yuhup <duúç> *duúç* || Hup <d’ùç> *ndũjh*
 (331) PDH **ǎ:ç* ‘golden-backed uakari’ > Dâw <êes> *ǎ:f* || Yuhup <êç> *ǎç* || Hup <êç> *ǎjh*
 (332) PDH **tô:h* ‘larva’ > Dâw <tôor> *tô:h* ‘maggot’ || Yuhup <mín-tôh> *mîn-tôh* ‘caterpillar sp. (edible)’, <saak-tôh> *tʃaak-tôh* ‘caterpillar sp.’ || Hup <tôh> *tôh* ‘edible caterpillar’
 (333) PDH **hó:h* ‘white-throated tinamou’ > Dâw <rôor> *hó:h* || Hup <moh-hôh> *môh-hôh*
 (334) PDH **pũ:h* ‘king vulture’ > Dâw <puur> *pũ:h* || Hup <wă’-pũh> *wă’-pũh*
 (335) PDH **wă₂:d* ‘food’ > Dâw <weed> *wě:d* || Yuhup <wén> *wădn* || Hup <wèd> *wădn*
 (336) PDH **ǎ:g* ‘*caxiri* beverage’ > Dâw <âag> *ǎ:g* || Yuhup <ág> *ǎg* || Hup <ág> *ǎg*
 (337) PDH **nă:m* ‘*curare* poison for arrows’ > Dâw <nâam> *nă:m* || Hup <nàm> *năm*

At least in Dâw and Hup, inherited CV-shaped morphemes do not show length or tonal contrasts: in Dâw, they are long and receive falling tone (Martins 2004: 83); in Hup, they are invariably lengthened and take falling tone (Epps 2005: 81), though loanwords may take rising

⁵² Quite unexpectedly, Epps (2005: 210) documents a high tone in this noun: <tih k’êt> *tih=kææt* ‘its leaf’. The expected rising tone is documented in Ramirez (2005: 99, 171).

⁵³ In Hup, an unclear variant <wòt> *wôt* has also been documented (Ramirez 2005: 204).

⁵⁴ The absence of laryngealization in the Yuhup reflex, as attested in Silva & Silva (2012: 509), is unexpected.

⁵⁵ This noun is only ever attested in the unstressed position (see examples in Ramirez 2005: 109; Epps 2005: 610), hence its underlying tone is unknown. I predict that it has a rising tone when stressed (i.e., when preceded by a third-person singular possessor *tih=* or a numeral).

⁵⁶ This etymon is known to be a Tukanoan loanword (compare Ye’pâ-masa <bu’ú> ‘peacock bass’; Ramirez 2019 [1997]), but it must have been present already in PDH given the existence of reflexes in all daughter languages which show regular correspondences).

⁵⁷ This form is documented in Martins (2004: 40). Epps *et al.* (2018) unexpectedly attest this noun with falling tone, which is otherwise claimed to be impossible in native vocabulary before a voiceless stop (Martins 2004: 83).

tone. Their Yuhup cognates typically show rising tone. In other words, in Dâw such morphemes behave like other morphemes with falling tone, whereas in Yuhup and Hup they behave like the toneless morphemes of PDH (except for the automatic vowel length in Hup). I believe that Dâw is more conservative in this regard, because some CV-shaped morphemes display vowel correspondences that lack a short version (e.g., 32, 128, 213). Some examples follow in 338–343.

- (338) PDH *c'ô: 'flower' > Dâw <çoo> c'ô: || Yuhup <soo> tʃɔ̌ || Hup <s'ó> tʃɔ̌:
 (339) PDH *cû: 'coati' > Dâw <suu> fû: || Yuhup <su> tʃû || Hup <su> tʃû:
 (340) PDH *tû: 'below, on the ground' > Dâw <tuu> tû: || Yuhup <tu> tũ || Hup <tú> tû:
 (341) PDH *hû: 'game animal' > Dâw <rũ> hû:⁵⁸ || Yuhup <hũ> hũ || Hup <hú> hû:
 (342) PDH *jâ: 'non-venomous snake' > Dâw <yê> jê: 'boa' || Yuhup <yê> jă || Hup <yé> njă:
 (343) PDH *kê₂: 'feather, wing' > Dâw <xê> xê: || Yuhup <kê> kě || Hup <ké> kê: 'wing'

In verbs, only Dâw shows length/tonal distinctions, whereas in Yuhup and Hup have largely lost them, since tone in verbs almost always conveys grammatical information. One exception is the apprehensive mood in Hup, where verbs do show lexical tonal distinctions (Epps 2005: 84–85), but I have been unsuccessful to derive these distinctions from PDH, and at any rate only a small number of verbs are attested in the apprehensive mood in Hup. In this paper, verbal roots (including stative verbal roots, glossed in English as adjectives) are always given with rising tone in Yuhup and with falling/high tone in Hup.

In 344–350, I give examples of CVC-shaped toneless verb roots in PDH.

- (344) PDH *c₉₁k 'to peel, to whittle' > Dâw <sâk> ʃɔk 'to peel' || Yuhup <säk> tʃɔk 'to peel, to weed, to whittle' || Hup <sák> tʃɔk 'to whittle'
 (345) PDH *c'g₁k 'to jump' > Dâw <çâk> c'ɔk || Yuhup <sääk> tʃɔ̌k || Hup <s'ák> tʃɔk
 (346) PDH *ki? 'sticky, to stick' > Dâw <xu> xu? 'to stick' || Hup <kí> kí?
 (347) PDH *t.ɔx 'to run' > Dâw <ox> ɔx || Yuhup <tooh> tɔ̌h || Hup <to'óh> tɔ̌h
 (348) PDH *tok 'to pound' > Dâw <tôk> tok || Yuhup <tök> tók || Hup <tók> tók
 (349) PDH *pog 'big' > Dâw <pôg> pog || Yuhup <pög> pög || Hup <póg> pög
 (350) PDH *po? 'open, to open' > Dâw <pô> po? 'to open' || Yuhup <pö> pö? 'open' || Hup <pó> pó? 'to open'

In 351–355, I give examples of CVC-shaped verb roots in PDH with falling tone. All of them have a voiced coda or *h.

- (351) PDH *jâ:ɣ 'to peel' > Dâw <yooj> jâ:ɣ 'to skin' || Yuhup <yoy> jɔ̌ || Hup <yój> djɔ̌dn
 (352) PDH *tâ:ɣw 'angry' > Dâw <tâaw> tâ:ɣw || Yuhup <täw> tɔ̌w || Hup <tăw> tâw
 (353) PDH *tâ:ɣw 'to carry in one's arms' > Dâw <toow> tâ:ɣw || Yuhup <toow kâ> tɔ̌w-kă? || Hup <tów> tɔ̌w 'to carry on a stick distributed between two people's shoulders'
 (354) PDH *k'û:n 'to weave (a basket)' > Dâw <kuun> k'û:n 'to weave (aturá basket)' || Yuhup <kuun> kũdn 'to make baskets (from vine)' || Hup <k'úd> kũdn 'to weave (a basket)'
 (355) PDH *bâ:ɣh 'to spill (liquid)' > Dâw <bâa> bɔ̌: || Yuhup <bääh> bɔ̌h || Hup <b'áh> mbɔ̌h

In 356–365, I give examples of CVC-shaped verb roots in PDH with rising tone. These typically have voiceless codas, but some examples with voiced codas are also known.

⁵⁸ This form is documented in Martins (2005: 232). Epps *et al.* (2018) unexpectedly attest this noun with rising tone, which is otherwise claimed to be impossible in CV-shaped morphemes by Martins (2004: 69).

- (356) PDH **c'ǎ:n* 'to caress, to hug affectionately' > Dâw <çeen> *c'ǎ:n* 'to caress, to please' || Hup <s'én> *sǎn* 'to hug affectionately'
- (357) PDH **tõ:j* 'to carry on one's head or neck' > Dâw <tôoy> *tõ:j* 'to carry on one's neck' (compare <tôoy> *tõ:j* 'to carry on one's shoulder or head') || Yuhup <töy> *tõj* 'to heap' || Hup <hitöy> *hi-tõj* 'to carry on one's head'
- (358) PDH **cš₁:p* 'thin' > Dâw <sâap> *š:p* || Hup <sáp> *tšáp*
- (359) PDH **c'ũ:k* 'to itch' > Dâw <çuuk> *c'ũ:k* || Yuhup <suuk> *tšũk* || Hup <s'úk> *tšúk*
- (360) PDH **kš₁:k* 'to pull' > Dâw <xâak> *xš:k* || Yuhup <käk> *kšk* || Hup <kák> *kák*
- (361) PDH **wš₁:?* 'to hear' > Dâw <wâa> *wš:?* || Hup <wá> *wó?*
- (362) PDH **tš₁:?* 'to make fire, to put on fire' > Dâw <tâa> *tš:?* || Yuhup <tä> *tš?* 'to rekindle' || Hup <tá> *tó?*
- (363) PDH **nš₁:?* 'to give' > Dâw <nõo> *nš:?* || Yuhup <no> *nš?* 'to give, to sell' || Hup <nó> *nó?*
- (364) PDH **jũ:?* 'hot' > Dâw <yuu> *yũ:?* 'hot; year' || Hup <yú> *djú?* 'to burn (paper, litter)'
- (365) PDH **wă:h* 'unripe, semi-ripe' > Dâw <waar> *wă:h* 'unripe' || Hup <wàh> *wă:h* 'semi-ripe'

In 366–368, I give examples of CV-shaped verb roots in PDH. These invariably have vowel length and falling tone in Dâw, and I assume this was already the case in PDH.

- (366) PDH **tô:* 'dry (of wood)' > Dâw <too> *tô:* || Yuhup <to> *tš* 'to dry or to get burnt (in the sun)' || Hup <tó> *tô:*
- (367) PDH **cô:* 'happy' > Dâw <sôo> *šô:* || Yuhup <sö> *tšô* 'to have a rest' || Hup <hisösö> *hi-tšô-tšô:*
- (368) PDH **pû:* 'to get wet'⁵⁹ > Dâw <puu> *pû:* 'to soak (e.g., flour, rice, beans)' || Hup <pú> *pû:*

Tonal suprafixes are used to derive nouns from verbs in all Dâw–Hup languages. Dâw employs the rising tone for this purpose (Martins 2004: 516), Yuhup uses the falling tone (Silva & Silva 2012: 101), whereas Hup uses the rising tone (Epps 2005: 86), though exceptions do exist. In addition, Dâw uses tonal suprafixes for changing valency, and deriving verbs from nouns, processes that must have been present already in PDH. For this reason, some apparent irregular correspondences between tones can be plausibly ascribed to the derivational history of individual nouns. Consider the data in 369–372. The Yuhup noun for 'shaman' has an unexpected falling tone, which does not match the evidence from Dâw or Hup, revealing its deverbal origin. Similarly, the Yuhup and Hup noun for 'achiote' point to a long vowel in PDH, suggesting that they may have in fact been derived from the verb **hš₁:w*.

- (369) PDH **cšw* 'shaman' > Dâw <sâw> *šw* || Hup <sáw> *tšw*
- (370) PDH **cšw* 'to be a shaman, to practise shamanism' > Dâw <sâw> *šw* || Yuhup <säw> *tšw* 'to practise shamanism' → <sáw> *tšw* 'shaman' || Hup <sáw> *tšw*
- (371) PDH **hšw* 'achiote' > Dâw <râw> *hšw*
- (372) PDH **hš₁:w* 'to be painted with achiote' > Dâw <râaw> *hš:w* 'to be red' || Yuhup <hâw> *hšw* 'to paint oneself with achiote' → <hâw> *hšw* 'achiote' || Hup <hâw> *hšw* 'to be painted with achiote' → <hâw> *hšw* 'achiote'

Some exceptions to the general rule remain that cannot be plausibly attributed to tonal suprafixes. In a handful of etymologies, Dâw points to a short vowel in PDH, whereas Yuhup

⁵⁹ This etymon is known to be a Tukanoan loanword (compare Ye'pâ-masa <puû> 'to get wet'; Ramirez 2019 [1997]), but it must have been present already in PDH given the existence of reflexes in Dâw and Hup which show regular correspondences.

and Hup point to a long vowel, as shown in 373–378. At least in 378, the vowel length must have been irregularly lost in Dâw, since the vowel correspondence in question is not normally found in morphemes with short vowels.

- (373) (?) PDH **c'um* ~ **c'u:m* ‘euphonia sp.’ > Dâw <çum> *c'ubm* || Yuhup <suúm> *tʃuûbm* ‘short-tailed pygmy tyrant’ || Hup <s'ùb> *tʃûbm*
- (374) (?) PDH **bɔʔ* ~ **bɔ:ʔ* ‘gourd bowl’ > Dâw <bo'> *bɔʔ* || Yuhup <boó'> *bɔʔʔ* || Hup <b'ò'> *mbɔʔʔ*
- (375) (?) PDH **tɔç* ~ **tɔ:ç* ‘fart’ > Dâw <tos> *tɔf* || Yuhup <tóc> *tôç* || Hup <tóc> *tójh* ‘to fart’
- (376) (?) PDH **k'ɔh* ~ **k'ɔ:h* ‘insect egg, maggot’ > Dâw <kor> *k'ɔh* ‘egg of a bee or a wasp’ || Hup <k'òh> *kǎh* ‘maggot’
- (377) (?) PDH **pux* ~ **pũ:x* ‘foam’ > Dâw <pux> *pux* || Yuhup <púh> *pûh* || Hup <pùh> *pûh*
- (378) PDH **jə₂w* ‘nine-banded armadillo’ > Dâw <jàw> *jɔw* ‘nine-banded armadillo’ || Yuhup <yéw> *jêw* ‘greater long-nosed armadillo’ || Hup <yèw> *djêw* ‘greater long-nosed armadillo, nine-banded armadillo’

In several other etymologies, by contrast, Dâw points to a long vowel in PDH, whereas Yuhup and Hup point to a short vowel. This is exemplified in 379–382.

- (379) (?) PDH **p.pĩ:n* ~ **p.pi₁n* ‘great kiskadee’ > Dâw <piin> *pĩ:dn* || Yuhup <pipin> *pipĩdn*
- (380) (?) PDH **bâ:w* ~ **baw* ‘venomous snake sp.’ > Dâw <baaw> *bâ:w* ‘rattlesnake’ || Yuhup <baaw> *baǎw* || Hup <b'áw> *mbâw* ‘pit viper (*Bothrops spp.*)’
- (381) (?) PDH **ʔâ:j* ~ **ʔāj* ‘woman’ > Dâw <āay> *ʔâ:j* || Yuhup <āy> *ʔǎj* || Hup <áy> *ʔāj*
- (382) (?) PDH **m̃ô:h* ~ **môh* ‘lake, stagnant water’ > Dâw <mō> *m̃ô:* ‘lake’ || Yuhup <moh> *mǎh* ‘lake (away from the river bed)’ || Hup <móh> *mǎh*

In 383–384, Yuhup, but not Hup, agrees with Dâw, pointing to a PDH long vowel. Hup, however, shows falling/high tone. Note that these two nouns form a semantic field.

- (383) PDH **tô:g* ‘daughter’ > Dâw <tôog> *tô:g* || Yuhup <tóg> *tôgɨ* || Hup <tóg> *tôgɨ*⁶⁰
- (384) PDH **tâ:~h* ‘son’ > Dâw <tee> *tê:* || Yuhup <téh> *têh* || Hup <téh> *têh* ‘son, fraternal nephew, offspring’⁶¹

In 385–388, Hup agrees with Dâw, pointing to a PDH long vowel. Yuhup, however, unexpectedly shows rising tone, according to Silva & Silva’s (2012) attestations. Possible explanations include borrowing from Hup, incorrect transcriptions, or true irregular change.

- (385) PDH **dõ:p* ‘oropendola, cacique’ > Dâw <dôop> *dõ:p* || Yuhup <dööp> *dõǎp* || Hup <d'õp> *dǎp*
- (386) PDH **tũ:n* ‘nocturnal curassow’ > Dâw <tuun> *tũ:dn* ‘curassow sp.’ || Yuhup <moy-tun> *mǎj-tũdn* || Hup <moytùd> *mǎj-tũdn*
- (387) PDH **jũ:h* ‘frog sp.’ > Dâw <yuur> *jũ:h* || Yuhup <yuh> *jũh* || Hup <yùh> *djũh* ‘frog (large, gray, lives in lakes or on land, edible)’
- (388) PDH **bũ:k* ‘vine sp.’ > Dâw <buuk> *bũ:k* || (?) Yuhup <buuk> *bũǎk* ‘fruit sp.’ || Hup <b'ùk> *mbũk* ‘*apuí* vine (*Clusia spp.*)’

Only a few etymologies involve exceptional tonal development, however. The correspondence patterns identified in this section are quite robust. An important question for compara-

⁶⁰ The expected rising tone in this noun is in fact attested once in Epps (2005: 202), but elsewhere Epps (2005) documents the same noun with falling tone.

⁶¹ A possible explanation for the divergent tone in Hup is that it could be back-derived from the homonymous verb ‘to be pregnant (of animals)’ (Epps 2005: 181).

tive Naduhup studies is whether tone existed in Proto-Dâw–Hup at all, given that no tonal contrasts are reconstructed in CV-shaped morphemes as well as in morphemes with a voiceless coda (except **h*). In fact, Barboza (2016) posits a binary opposition between high-toned morphemes (my short vowels) and rising-toned morphemes (my long vowels), and claims that the tonal distinction found in Dâw is secondary. Considering that a contrast is reconstructed in morphemes with voiced codas, one is tempted to assume that it may have originally existed in morphemes with voiceless codas as well, but the combination of the falling tone and a voiceless coda must have been eliminated in some way (either by altering the tone or by voicing the coda). For the time being, the issue remains open.

6. Additional etymologies

In this section, I present some additional etymologies that were not discussed in 2–5. I start by presenting cognate sets that lack known reflexes either in Dâw or in Yuhup–Hup, but that must have existed in Proto-Dâw–Hup, since cognates have been identified in Nadëb.

The etymologies in 389–404 are only represented in Yuhup and/or Hup, but not in Dâw. They are securely reconstructed to PDH because there are external cognates in Nadëb. In 403, Yuhup points a long vowel, and Hup to a short one, just like in 383–384.

- (389) PDH **wa(:)t* ‘to visit, to walk’ > Yuhup <wat> *wăt* ‘to visit’ || Hup <wát> *wát* ~ Nadëb <-wät> *-wnt* (non-indicative <-wäd> *-wnd*) ‘to move’ (Weir 1984: 165)
- (390) PDH **k’ap* ‘fish sp.’ > Hup <k’áp> *k’áp* ‘*piaba* fish sp. (~ 5 cm long)’ ~ Nadëb <gäb> *k’ab* ‘cardinal tetra’ (Weir 1984: 190)
- (391) PDH **k’gh* ~ **k’g₁:h* ~ **ke₁(:)h* ‘sweet’ > Yuhup <kääh> *kəǎh* || Hup <k’áh> *kǎh* ~ Nadëb <-gëäh> *-k’ə:h* (Weir 1984: 66)
- (392) PDH **kǎ₁:?* ~ **kǎ₁:?* ‘bone, leg’ > Yuhup <ká’> *kâ?* ~ Nadëb <k’ëë> *kə:* ‘stem (of manioc)’ (Weir 1984: 154), ‘bone’ (Martins 2005: 278)
- (393) PDH **w.hg₁:h* ~ **w.he₁:h* ‘old woman’ > Yuhup <wäháh> *wəhâh*, <wähäh> *wəhǎh* ‘to become old (of women)’ ~ Nadëb <wahëh> *wahəh* ‘old (pl.)’ (Weir 1984: 152)
- (394) PDH **cɔ:w* ‘to shoot with a blowgun’ > Yuhup <sow> *tʃɔw* || Hup <sów> *tʃɔw* ~ Nadëb <-sóóm> *-fɔ:bm* (non-indicative <-sóów> *-fɔ:w*) ‘to shoot with a blowgun’ (Weir 1984: 240, 310)
- (395) PDH **ʔɔ:j* ‘pulp-like, ashes’ > (?) Yuhup <oy> *ʔj* ‘to chew’ || Hup <óy> *ʔj* ‘peach palm fruit pulp, grated manioc’, <tëg-óy> *teŋ-ʔj* ‘ashes’ ~ Nadëb <ooj> *ʔo:j* (Roçado dialect), <óój> *ʔo:j* (Rio Negro dialect) ‘ashes’ (Martins 2005: 321)
- (396) PDH **xô:* ~ **hô:* ‘liver’ > Yuhup <ho> *hǒ* || Hup <hó> *hô:* ~ Nadëb <hooh> *ho:h* ‘liver’ (Barbosa 2005: 27)
- (397) PDH **nǝh-cug* ‘beard’ > Yuhup <noh-sug> *nǝh-tʃŋ* ‘stalk of a fruit or a leaf’ || Hup <noh-súg> *nǝh-fŋ* ~ Nadëb <nasuuk> *nǝfu:k* ‘beard’ (Barbosa 2005: 44)
- (398) PDH **bu(:)j(-tɔk)* ‘ear’ > Yuhup <buy-tok> *bujtǔk* ~ <buy-rok> *bujrǔk* || Hup <b’otók> *mbɔtǔk* ~ <b’orók> *mbɔrǔk* ~ Nadëb <nabuu> *nǝbu:j* ‘ear’ (Barbosa 2005: 26)
- (399) PDH **xæ₂:g* ~ **hæ₂:g* ‘Southern American bushmaster’ > Yuhup <hé> *hǎŋ* || Hup <hè> *hǎŋ* ~ Nadëb <haak> *ha:k* ‘Southern American bushmaster’ (Barbosa 2005: 38)
- (400) PDH **k’î₂:* ~ **k’î₂:* ‘mosquito’ > Yuhup <kii> *kǐ* || Hup <k’í> *kǐ:* ~ Nadëb <giij> *k’iij* ‘mosquito’ (Weir 1984: 81)

- (401) PDH **tuʔ* ‘to be in water’ > Yuhup <tu> *tũʔ* || Hup <tú> *túʔ* ~ Nadëb <-tu> *-tu* ‘to be in water (sg.)’ (Weir 1984: 141)
- (402) PDH **ji₂ʔ* ~ **ji:ʔ* ‘man, male’ > Yuhup <yí> *jĩʔ* || Hup <tiyí> *ti-jĩʔ* ~ Nadëb <aj’yy> *ʔajjĩ* (pl. <ajyy> *ʔajjĩ*) ‘man’ (Epps & Obert 2021: 258)
- (403) PDH **ĩ:n* ‘mother’ > Yuhup <ín> *ĩn* || Hup <ín> *ĩn* ~ Nadëb <-’ỹyn> *ĩ:n* ‘mother’ (Weir 1984: 151)
- (404) PDH **c’.c’ib* ~ **c’.c’i₁b* ‘insect sp.’ > Yuhup <sisiim> *tʃitʃĩb* ‘mosquito sp.’ || Hup <s’is’ib> *tʃitʃĩp* ‘fly sp. (tiny, with red head)’ ~ Nadëb <watsyb> *wac’ib* ‘biting midge’ (Barbosa 2005: 35)

The etymologies in 405–409 are only represented in Dâw, but not in Yuhup or Hup. They are securely reconstructed to PDH because there are external cognates in Nadëb.

- (405) PDH **k’â:w* ‘to fell trees; cultivated field’ > Dâw <kaaw> *k’â:w* ~ Nadëb <-gääm> *k’â:bm* (non-indicative <-gääw> *-k’â:w*) ‘to fell trees’, <-gääw> *-k’â:w* ‘cultivated field’ (Weir 1984: 40)
- (406) PDH **c’â:j* ‘frog sp.’ > Dâw <çaay> *c’â:j* ~ Nadëb <ts’ääj> *c’â:j* ‘frog sp.’ (Martins 2005: 230)
- (407) PDH **tĩm* ‘eye’ > Dâw <tum> *tubm* ~ Nadëb <matym> *mātĩbm* (incorporated <ty> *tĩ*) ‘eye’ (Weir 1984: 298)
- (408) PDH **kɔm* ‘buttress root’ > Dâw <xom> *xɔm* ~ Nadëb <kóm> *kɔbm* ‘root’ (Martins 2005: 315)
- (409) PDH **coʔ* ‘to take out’ > Dâw <sô> *foʔ* ~ Nadëb <-sok> *-fok* ‘to take out (pl.)’ (Weir 1984: 289)

Some etymologies present irregularities of kinds not considered above, such as correspondences between a stop and a nasal (410), between a fricative and a stop (411), between unrounded and rounded vowels (412), between glottal and palatal fricatives in addition to an irregular vowel correspondence (413), between plain and glottalic/glottalized consonants in different positions (414), or involve multiple irregularities (415–416). They may involve irregular change, horizontal transmission, or simply turn out to be flawed comparisons. Note that 414 is similar to forms found in some Tupian languages of the upper basins of the Madeira and Tapajós Rivers, which do not show regular correspondences among themselves, and cannot be of Tupian origin: Sakurabiat <siokweet> *tsiokʷet* ‘toucan’ (Costa 2020: 135), Akuntsu *jōkʷet* ‘toucan’ (Aragon 2008: 35), Kawaiwete <sokwet> *sokʷet* ‘toucan’ (Weiss 2005: 101).

- (410) (?) PDH **dĩp* ~ **dĩm* ‘to disappear’ > Dâw <’nũp> *’nũp* || Yuhup <niim> *nĩm*
- (411) (?) PDH **bi₂ç* ~ **bi₂c* ‘bird sp.’ > Dâw <bis> || Yuhup <biis> *biĩç* ‘silver-beaked tanager’ || Hup <b’ís> *bĩjt* ‘tanager’
- (412) (?) PDH **mĩ:n* ~ **mũ:n* ‘caatinga (place with semi-arid vegetation)’ > Dâw <mũun> *mũ:n* || Hup <mùn> *mũn* (cf. also <mumún> *mũmũn* ‘id.’)
- (413) (?) PDH **wæh* ~ **wĀh* ~ **wĀç* ‘pigeon’ > Dâw <wēr> *wēh* || Yuhup <wēh> *wǎh* || Hup <wěç> *wǎjh*
- (414) (?) PDH **c’okwe₁t* ~ **cōke₁t* ~ **cōk’e₁t* ~ **cōkwe₁t* ‘toucan’ > Dâw <çôkwêt> *c’okwet* || Yuhup <sokät> *tʃokät* || Hup <säk’ât> *tʃokät* ~ <sokw’ât> *tʃokwät*
- (415) (?) PDH **cô:ŋ* ~ **j.jô:ŋ* ~ **mu-c’ô:ŋ* ~ **mu-jô:ŋ* ‘elbow’ > Dâw <sôogn> *ŋô:ŋ* || Yuhup <yöyôg> *jojôŋ* || Hup <bus’âg> *mbutʃôŋ* ~ <bus’ôg> *mbutʃôŋ* ~ <bös’ôg> *mbotʃôŋ* ~ <buyôg> *mbuyôŋ*
- (416) (?) PDH **c’ĩ:c* ~ **c’ĩ:d* ~ **c’ĩ:n* ~ **c’ĩ:j* ‘beetle sp.’ > Dâw <çĩnh> *c’ĩ:n* ‘*abacaba* palm beetle’ || Yuhup <siín> *tʃĩdn* ‘scarab beetle’ || Hup <s’ij> *tʃĩdn* ‘dung beetle’

7. Conclusion

In this paper, I have made an attempt at reconstructing the phonology and lexicon of Proto-Dâw-Hup, thus contributing to comparative studies of the Naduhup languages. I acknowledge that some decisions of mine need further reworking. In particular, a system of 45 nuclei is typologically surprising, and it is probable that future research will show that some of the vowel correspondences considered here are in complementary distribution, thus helping reduce the inventory; the optimal diachronic interpretation of what I symbolize by subscript digits and capital letters remains to be established. Another issue that requires further attention are the cases where Yuhup and Hup show nasality without any apparent source in PDH (the symbol ~ in my reconstruction is, of course, an *ad hoc* notation of a phenomenon whose nature is unclear). Finally, the fact that the tonal opposition of Dâw (and PDH) is only fully seen in morphemes with voiced codas and **h* is suspect. It remains to be determined whether the tones represent an innovation with regard to Proto-Nadahup (cf. Barboza 2016), or whether the tonal opposition had once been more robust.

Acknowledgement

I thank an anonymous *Journal of Language Relationship* reviewer for their careful reading of the manuscript. Needless to say, any remaining errors are my responsibility.

References

- Aguilar, Héctor F. 2004. Lista de los primates de la República Bolivariana de Venezuela. *Biodiversidad & Conservación Integral* 3(2): 13–20.
- Aragon, Carolina Coelho. 2008. *Fonologia e aspectos morfológicos e sintáticos da língua Akuntsú*. MA thesis. Universidade de Brasília.
- Barbosa, Jefferson Fernando. 2005. *Descrição fonológica da língua Nadêb*. MA thesis. Universidade de Brasília.
- Barboza, Lucas Cavalini. 2016. *Preliminary considerations on the development of glottalization in Nadahupan languages. Talk presented at Amazónicas VI. Coloquio Internacional Amazónicas (“La estructura de las lenguas amazónicas”)*. Leticia/Tabatinga: Universidad Nacional de Colombia; Universidade do Estado do Amazonas; Instituto Caro y Cuervo.
- Barboza, Lucas Cavalini. 2017. *Aspectos da glotalização na língua Dâw: um estudo de fonética experimental*. MA thesis. Universidade de São Paulo. doi: 10.11606/D.8.2017.tde-19092017-140841
- Costa, Carla Daniele Nascimento da. 2020. *Proposta de material didático para a língua Sakurabiat*. MA thesis. Universidade Federal do Pará.
- Emiri, Loretta. 1987. *Dicionário Yãnomamê-português (dialecto wakathautheri)*. Boa Vista: Comissão Prô-Índio de Roraima.
- Epps, Patience. 2005. *A grammar of Hup*. Ph.D. dissertation. University of Virginia.
- Epps, Patience, Katherine Bolaños. 2017. Reconsidering the “Makú” language family of Northwest Amazonia. *International Journal of American Linguistics* 83(3): 467–507. doi: 10.1086/691586
- Epps, Patience, Karolin Obert. 2021. Number marking in Nadëb. *Visitas al Patio* 15(2): 255–273. doi: 10.32997/RVP-vol.15-num.2-2021-3689
- Epps, Patience, Karolin Obert. 2022. Naduhup languages and the typology of nominal classification. *Cadernos de Etnolinguística* 10(1): e100104.
- Epps, Patience, Karolin Obert, Luciana Storto. 2018. Dictionary Dâw–Port–Eng v. 2018. In: Luciana Storto, Karolin Obert, Patience Epps (depositors). *Dâw Language Collection*, item 20051. Austin: Archive of the Indigenous Languages of Latin America (AILLA).
- Fernandes, Dacyo Cavalcante. 2017. *Estudo fonético e fonológico da língua Yuhup*. MA thesis. Universidade Estadual Paulista “Júlio de Mesquita Filho”.

- Finbow, Thomas. 2020. Nheengatu Dâw: a preliminary study of the phonetic, phonological and morpho-syntactic aspects of a case of Tupi-Guarani and Nadahup contact in the Upper Rio Negro. *Cadernos de Linguística* 1(3): 1–21. doi: 10.25189/2675-4916.2020.v1.n3.id236
- Javkin, Hector. 1977. Towards a phonetic explanation for universal preferences in implosives and ejectives. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 3: 559–565.
- Lopes, Aurise Brandão. 1995. *Fonologia da língua Yuhup: uma abordagem não-linear*. MA thesis. Universidade Federal de Santa Catarina.
- Martins, Silvana Andrade. 2004. *Fonologia e gramática Dâw*. Ph.D. dissertation. Vrije Universiteit Amsterdam.
- Martins, Valteir. 2005. *Reconstrução fonológica do Protomaku Oriental*. Ph.D. dissertation. Vrije Universiteit Amsterdam.
- Mattei-Müller, Marie-Claude; Serowë, Jacinto. 2007. *Lengua e cultura yanomami: diccionario ilustrado yanomami-español / español-yanomami*. Caracas: Epsilon libros.
- Nikulin, Andrey, Fernando O. de Carvalho. 2018. Prehistoria de las lenguas y familias lingüísticas del Gran Chaco, de la meseta brasileña y cercanías: propuesta de base de datos léxicos y resultados preliminares. In: María Alejandra Regúnaga, Silvia Andrea Spinelli, María Emilia Orden (eds.). *IV Encuentro de Lenguas Indígenas Americanas-ELIA: libro de Actas*: 545–560. Santa Rosa: Universidad Nacional de La Pampa.
- Obert, Karolin. 2019. *The linguistic encoding of space in Dâw*. Ph.D. dissertation. Universidade de São Paulo. doi: 10.11606/T.8.2019.tde-23102019-173847
- Obert, Karolin, João Vitor Fontanelli Santos. 2022. Dâw (Brazil) – language contexts. *Language Documentation and Description* 22(1). doi: 10.25894/ldd.323
- Ospina Bozzi, Ana María. 2002. *Les structures élémentaires du yuhup makú, langue de l'Amazonie colombienne : morphologie et syntaxe*. Ph.D. dissertation. Université Paris-Diderot (Paris-VII).
- Perri Ferreira, Helder. 2017. *Yanomama clause structure*. Ph.D. dissertation. Radboud Universiteit Nijmegen.
- Ramirez, Henri. 2005. *A língua dos Hupd'âh do alto Rio Negro. Dicionário e guia de conversação*. São Paulo: Saúde sem Limites.
- Ramirez, Henri. 2019 [1997]. *A fala Tukano dos Ye'pâ-Masa. Tomo II: dicionário*. Manaus: Inspetoria Salesiana Missionária da Amazônia / CEDEM.
- Silva, Cácio; Silva, Elisângela. 2012. *A língua dos Yuhupdeh: introdução etnolinguística, dicionário Yuhup-português e glossário semântico-gramatical*. São Gabriel da Cachoeira: Pró-Amazônia / AECIPY – Associação das Escolas e Comunidades Indígenas do Povo Yuhupdeh.
- Weir, E. M. Helen. 1984. *A negação e outros tópicos da gramática Nadëb*. MA thesis. Universidade Estadual de Campinas. doi: 10.47749/T/UNICAMP.1984.76698
- Weiss, Helga E. 2005. *Dicionário Kayabi-português com um glossário português-Kayabi*. Brasília: Sociedade Internacional de Linguística.
- Wetzels, W. Leo, Andrew Nevins. 2018. Prenasalized and postoralized consonants: The diverse functions of enhancement. *Language* 94(4): 834–866. doi: 10.1353/lan.2018.0055

A. В. Никулин. Фонологическая реконструкция праязыка дыу–хупской группы

В статье рассматриваются звуковые соответствия между языками дыу, юхуп и хуп, составляющими группу надухупской языковой семьи (Южная Америка), и предлагается фонологическая реконструкция праязыка этой группы. Морфемы в прадыу–хупском чаще всего имеют форму */CVC/, но восстанавливаются и морфемы структуры */CV/, а также некоторое количество полуторасложных и двусложных морфем. Прадыу–хупский консонантизм насчитывает 22 согласных, среди которых глухие и глоттализированные смычные, обычные и глоттализированные глайды, фрикативные, носовые, а также звонкие смычные. Что касается вокализма, обнаружено 45 рядов соответствий, противопоставленных в том числе по таким признакам, как долгота и назальность; некоторые из реконструированных гласных, возможно, были дифтонгами, а не монофтонгами. Для долгих гласных восстанавливается тональная оппозиция между восходящим и падающим тонами, однако оппозиция нейтрализуется перед большинством глухих финалей.

Ключевые слова: надухупские языки; сравнительно-исторический метод; фонологическая реконструкция.

Hokan III: Velar obstruents and glottal stop¹

The article continues the series of publications focused on the reassessment of the controversial Hokan hypothesis. In this paper, the traditional comparative method is applied to the reflexes of Proto-Hokan obstruents and glottal stop. The obtained results are compared to Terrence Kaufman's (1989) reconstruction of Proto-Hokan.

Keywords: comparative method; Hokan languages; Native American languages; historical phonology.

0. Introduction

The present article continues the series of publications dedicated to reassessing the controversial Hokan hypothesis (Zhivlov 2024, Zhivlov 2025). The previous installment treated sound correspondences for labial and coronal obstruents. In this article, I will survey the remaining correspondences for obstruents — those for velar obstruents and glottal stop. For the question of family membership, see Zhivlov (2025: 52). The survey of Hokan phonological systems can be found in Zhivlov (2025: 54–61). The list of default sources of language data is repeated below:

Chimariko — J. P. Harrington's fieldnotes (available online at <https://sova.si.edu/record/naa.1976-95/contents>). Harrington's fieldnotes are cited by volume, reel and frame, e.g., (2.20.0024) means Volume 2, Reel 20, Frame 0024;

Karuk — Bright & Gehr (n.d.);

Shasta — Silver (1964) and other works by the same author;

Achumawi — Nevin (2020);

Atsugewi — L. Talmy's fieldnotes (Talmy n.d., a; Talmy n.d., b);

Yana — Sapir & Swadesh (1960);

Pomoan — McLendon (1973);

Salinan — J. P. Harrington's fieldnotes (available online at <https://sova.si.edu/record/naa.1976-95/contents>);

Yuman — Mixco (1985) for Kiliwa, Crawford (1989) for Cocopa, Miller & Langdon (2008) for Jamul Tiipay (data from speakers Adolph Thing and Jane Thing Dumas) and Barona 'Iipay, Couro & Hutcheson (1973) for Mesa Grande 'Iipay, Munro et al. (1992) for Mojave, Salt River Pima-Maricopa Indian Community — O'Odham Piipaash Language Program (2000) for Maricopa, Halpern (1946a, 1946b, 1946c, 1947a, 1947b, 1947c) for Quechan, Joel (1966) for Paipai, Watahomigie et al. (2003) for Hualapai, Hinton (1984) for Havasupai, and Shaterian (1983) for Yavapai; Seri — Moser & Marlett (2010). On sources for 19th century Seri, see Marlett (2010).

Oaxaca Chontal — Turner & Turner (1971) for Highland Chontal, O'Connor (2013) for Lowland Chontal, and de Angulo & Freeland (1925) for Tequistlateco.

¹ The study was carried out with the assistance of the "Laboratory Of Unnecessary Things" of the Independent University of Moscow. I thank Drew Bennett, Fernando O. de Carvalho, Albert Davletshin, Stephen A. Marlett, Yoram Meroz, Andrey Nikulin, Juho Pystynen and two anonymous reviewers for their valuable advice. All remaining errors are mine alone.

1. Sound correspondences

1.0. Introduction

As in the previous paper (Zhivlov 2025), I present sound correspondences between obstruents in Hokan languages starting from their reflexes in Proto-Pomoan.

1.1.0. Correspondences involving Proto-Pomoan velar stops and fricatives

The following velar phonemes are reconstructed for Proto-Pomoan: *k, *k^h, *k̥ and *x (McLendon 1973: 23-24, 26).

1.1.1. Proto-Pomoan *k

The sound correspondence involving Proto-Pomoan *k is represented in Table 1. This correspondence reflects the Proto-Hokan plain labialized velar stop *k^w. There are no plausible comparisons where Proto-Pomoan *k corresponds to a non-labialized velar in Yuman or Chontal. The reflex in Achumawi may be connected with the position of the sound within the root (see below for Achumawi /t/ as a reflex of a root-final PH *-k), but the data are too sparse for any conclusion. Salinan /k̥/ in *ʔašák* < PH *k^hak^wá may represent a merger of expected /k/ with a suffix *-ʔ, on which see Zhivlov (2025: 63). Seri has zero (19th century Seri /β/).

PH	Chi	Kar	Sha	Ach	Ats	Yan	PPom	Sal	PYum	Seri	PCho
*k ^w	k	k	k	t		k	*k	k̥	*k ^w	β > Ø	*k ^w

Table 1. Proto-Hokan *k^w.

(k^w-01) PH *(a-)k^wá ‘mother’s father’

SHA *ʔák-wit* ‘mother’s father’ (Silver 1961: 2)

PPOM *-ká-čì ‘mother’s father’ > Pk -ca-š-, Ps -č-ač-, Pn -cáʔ, Pne tá-čì, Pse -ce[n], Pe ká-č (pPom *-čì ‘one’s own kinsman in generations above ego’)

PYUM *(n-)k^wa(·)-w ‘mother’s father’ > Ki -kuw² ‘be mother’s father (and siblings)’, Co n^y-k^waw ‘mother’s father’, JT nə-k^wa·w ‘grandfather’, MGI pa-nə-k^wa·w ‘his mother’s father’, Mo na-k^wew ‘mother’s father, grandfather’, Ma n-k^wo ‘grandfather (maternal)’, Hu k^wáw-a ‘mother’s father, grandfather’, Ha n-k^waw ‘grandfather (mother’s side)’, Ya k^wāw-a ‘mother’s father’ (*n- ‘kinship terms prefix’ [Miller 2001: 86])

SER aa-š ‘maternal grandfather; great-uncle (mother’s father’s brother); mother’s father’s sister’s husband’ (pl. aa-tk [sg. possessor], áa-tkox [pl. possessor], abs. ʔapé-ε-š, 1 sg. ʔε-aa-š)

NOTES. (1) This kinship term forms part of a mini-paradigm of terms for grandparents: PH *(a-)pá ‘father’s father’, PH *(a-)k^wá ‘mother’s father’ (k^w-01), PH *(a-)má ‘father’s mother’, PH *(a-)ká ‘mother’s mother’ (k-01). All four terms for grandparents are preserved in Pomoan, Yuman and Seri. (2) Shasta -wit, Yuman *-w and Seri -š are suffixes found in other terms for grandparents. (3) PH *k^w > Seri zero possibly by way of 19th century Seri /β/ (reflected as zero in modern Seri), but this particular word is not attested in 19th century wordlists. (4) Achumawi *wa-q^hu-wí* ‘mother’s father’ and Atsugewi *ʔaq^hon* ‘maternal grandfather’ (Talmy n.d., b: 182) are excluded from this set because of aspiration.

McLendon 1964: 131 || Gursky 1974: 191 (#aku-# ‘maternal grandfather’ (?)) || Webb 1981: 81 || Gursky 1989: 12 || Kaufman 1989: 145 (#k^wa:(wV) ‘mother’s father’ [N/S])

(k^w-02) PH *k^wá ‘to speak’YAN *ka-* ‘to speak, utter’PPOM *ka[hnów] ‘to talk’ > Pk *cahnów-* ‘to sound’, Ps *čáhnú* ‘to instruct young’, Pn *čanó-*, Pc *cá·now* ‘to voice’, Pne *táno·[-n]*, Pse *cnú* ‘word’, Pe *ka·Nú·* ‘speak, say, talk’, *ka·Nú* ‘word’PYUM *k^wa, *k^wa-*w* ‘to speak’ > Ki *k^wa¹* ‘imperfective aspectual auxiliary: speak, be audible’, *k^wa·w²* ‘shout; call out’ (-*w* directional), Pa *k^waw* ‘one speaks’, Hu *k^wa·w-k* ‘to talk, to speak’, Ha *k^wa·w-k* ‘to talk; to hold conversation’, Ya *k^wā·w-i*, *k^wāw-i* ‘speak’PCHO *·k^wa- ‘say’ > HCh *di-gwá-?ma* ‘he will talk, say, think’, LCh *kwa* ‘says, say, used often in stories for a general ‘he says’, ‘they say’NOTES. Cf. PH *i-k^wá ‘to say, talk’ (k^w-02).Sapir 1917: 7 || Sapir 1920b: 288 || McLendon 1964: 136 || Gursky 1965: 185 || Gursky 1968: 37 || Gursky 1974: 206 || J. G. Crawford 1976: 319 || Waterhouse 1976: 342 || Greenberg 1987: 252 || Kaufman 1989: 130 (#kwa [a ~ o] ‘to talk; say’ [S]; #k^ya ‘to speak, talk; by speech (incorp preposund)’ [N/S]) || Gursky 1990: 1, 9(k^w-03) PH *k^hak^wá ‘flint; arrowhead’ (= k^h-03)CHI *q^hako* <khako> ‘arrowhead’ (2.22.0211), <khákkò> ‘bullet, arrowhead’ (2.23.0991)KAR *sá·k* ‘obsidian, flint; arrowhead, bullet’ACH *sát* ‘arrowhead; knife; flint; metal; bullet; iron; obsidian’YAN CN *xaka*, Y *xahka* ‘flint; Flint (as a myth character); flint, a design element in basketry’, N *xaka-wi* ‘arrowheads’PPOM *q^hahká ‘flint’ > Pk *qahca*, Ps *kahča*, Pn *k^hačá*, Pc *k^hcá*, Pne *k^háhta*, Pse *qčá*, Pe *xa·ká*SAL A [*ʔa*]šák <‘ašák’> ‘arrow point of flint’ (2.85.0096), A [*ʔa*]šák <‘acák’>² ‘what they put on point of arrows’, M [*pá*]šák <‘p’acák’> ‘id.’ (2.86.0584)PYUM *(*ʔ*-)xk^wa ‘knife’ > Mo *ʔa-hk^we* ‘knife’, Hu *a-k^wá*, *ʔ-k^wa*, *k^wa* ‘knife’, Ha *k^wa-ʔa* ‘silver; knife; metal’, Ya *ʔ-k^wá* ‘metal, iron; knife’SER *ʔax* ‘arrow point’ (pl. *ʔax-lk*), 19th century Seri (Pinart) <vaX> ‘pedernal; flint’ (Hernández 1902: 253)NOTES. (1) The word may be a compound with PH *(a/i-)k^haʔa ‘stone’ (k^h-10) as the first part. (2) Seri reflects a metathesized form *k^wak^há.Sapir 1917: 8 || Sapir 1920b: 284 || Sapir 1921: 69 || Greenberg & Swadesh 1953: 219 || Jacobsen 1958: 200 || Haas 1964: 78 || McLendon 1964: 131 || Silver 1964: 173 || Gursky 1965: 169 || Gursky 1974: 188 (#q^hak^ya#; #sak^(y)#) || Langdon 1979: 637 || Greenberg 1987: 259 || Gursky 1989: 9 || Kaufman 1989: 138 (#x^yáHk^wa ‘flint; arrowpoint; knife’ [N/S]) || Leshchiner & Nikolaev 1992: 399 (*č^hákA) || Shaul 2020b: 302(k^w-04) PH *(i-)k^wá/i· ‘to hear’CHI *hike-ta* <hikéta> ‘he heard it’ (2.23.0335), *ʔike-niq* <‘ikēnik> ‘I heard it’ (2.22.0299)YAN *ka-* ‘to hear, listen to’PPOM *[xó·]ki- ‘to hear’ > Pk *šoc-* (Oswalt & Buckley 2018), Ps *šo·či-*, Pn *šó-*, Pc *šó·c-*, Pse *xko*, Pe *šó·k^h*PYUM *k^wa/i(·) ‘to hear’ > Ki *k^wi·²* ‘perceive, sense, hear’, Hu *vik^wá-k* ‘to hear a noise’, Ha *k^wa-k* ‘to hear; to feel’SER *k-i* ‘hear; understand’ (conjug.: *i-t-ii*, *i-t-ii-tim*; *i-t-iii*)PCHO *·k^we[x]- > HCh *di-gweh-?ma* ~ *di-gyeh-?ma* ‘he will understand, hear with understanding’ (Turner 1973: 100)NOTES. PH *k^w > Seri zero possibly by way of 19th century Seri /β/ (reflected as zero in modern Seri), but this particular word is not attested in 19th century wordlists.Sapir 1917: 8 || Sapir & Swadesh 1960: 16 (*ka-yi) || Gursky 1974: 193 (#kaya#) || J. G. Crawford 1976: 313 || Greenberg 1987: 137 || Gursky 1989: 16 || Kaufman 1989: 128 (#k^weyá ‘to hear’ [N/S])² Harrington notes: “á not very long but falling”.

(k^w-05) PH *a/i-k^wi ‘head louse’

PPOM *ʔahkí[n^y] ‘head louse’ > Pk ʔahci, Ps ʔahči, Pn čí, Pc cí, Pne tí[-ka·], Pse cin, Pe kí·

SER ii ‘head louse (*Pediculidae*)’ (pl. ii-t, abs. ʔai)

PCHO *-a-yk^wi ‘head louse’ > HCh t-á-ygwi ‘the black louse’, LCh a-ykwih? ‘head louse’

NOTES. (1) Cf. (č-05) PH *(a-)či ‘louse’. (2) PH *k^w > Seri zero possibly by way of 19th century

Seri /β/ (reflected as zero in modern Seri), but this particular word is not attested in 19th century wordlists.

Kaufman 1989: 152 (#aHk^wi ‘louse’ [N/S])

(k^w-06) PH *k^wumV ‘all’

CHI kumi[čin] <kúmítjin> ‘all’ (2.23.0761)

PPOM *ku(H)mú ‘all’ > Pe ku·mú (McLendon 1975: 30)

Gursky 1974: 178 || Kaufman 1989: 147 (#k^wu(m) ‘all’ [N])

1.1.2. Proto-Pomoan *k^h

The sound correspondence involving Proto-Pomoan *k^h is represented in Table 2. This correspondence reflects the Proto-Hokan aspirated labialized velar stop *k^{wh}. Karuk has /k/ before *i, but /p/ before *u, cf. below on PH *x^w > Karuk /f/ before *u.

PH	Chi	Kar	Sha	Ach	Ats	Yan	PPom	Sal	PYum	Ser	PCho
*k ^{wh}	x	k, p				k ^h	*k ^h	k	*k ^w		*k ^w

Table 2. Proto-Hokan *k^{wh}.

(k^{wh}-01) PH *a/i-k^{wh}i ‘rain’

PPOM *[k]ihk^hé(·) ~ *ihk^hé ‘rain’ > Pk ʔihc^he, Ps čehč^he, Pn [bi]šé-, Pc cé, Pe k^hi·ké·

PYUM *(ʔ-)k^wi[y] ‘cloud’ > Ki ʔ-k^wiy¹ ‘cloud’, Co k^wi· ‘cloud’, JT k^wi· ‘cloud, clouds’, MGI ʔə-k^wi·y ‘clouds’, Mo ʔ-i-k^we ‘cloud’, Ma k^we ‘cloud’, Qu ʔa-k^wé ‘cloud’, Pa ʔ-k^wi ‘cloud’, Hu k^wi ‘cloud’, Ha k^wi-ʔi ‘cloud’, Ya (ʔ)k^wi ‘cloud’

PCHO *-a-k^wi ‘rain’ > HCh t-á-gwi, LCh a-kwi?

Kroeber 1915: 282 || Greenberg & Swadesh 1953: 219 || Gursky 1974: 202 || J. G. Crawford 1976: 310 || Waterhouse 1976: 337 || Langdon 1979: 638 || Greenberg 1987: 139 || Kaufman 1989: 143 (#[a/i]k^{wh}ey ‘rain; cloud’ [N/S]) || Leshchiner & Nikolaev 1992: 399 (*ʔik^hi)

(k^{wh}-02) PH *(a-)k^{wh}il ‘to hang’

KAR [tan]kir / [tan]kiri- ‘to hang (things) over something, as a rack or a clothesline’

PPOM *hk^hil- > Pk hc^hil ‘hang, cling to a surface, swing’ (Oswalt & Buckley 2018), Pe k^hil- ‘be located, hanging or suspended’ (McLendon 1975: 185)

SAL M k-áke-l <kákē-l> ‘está colgado’, p-aké-l-o <pakē-lo khé> ‘yo lo voy a colgar’ (2.86.0628); A ake(·)l ‘to hang up’ in <paké-lo> ‘he hung it up’, <makél> ‘hang it up!’, <kakél> ‘hang it up (pl.)!’ (Jacobsen 1954: 159)

PYUM *k^wil ‘to hang’ > MGI a-k^wil-p ‘a long object is hanging down’, k^wil-k^wil ‘is dangling (of a large object)’

Gursky 1974: 192 || Gursky 1989: 15 || Kaufman 1989: 161 (#KiL [ʔasp] ‘to be hanging/hung; to hang’ [N])

(k^{wh}-03) PH *(a-)k^{wh}-Vʔiy ‘salt’

CHI ʔaki <ak^hi> ‘salt’ (2.20.0040)

PPOM *k^heʔé: ‘salt’ > Pn šeʔé, Pne č^héʔe, Pe k^he-ʔé

SAL A *ṭ-a·káj* <trākáʔ> ‘sal’ (2.85.0127), M *ṭ-a·káj* <trākáy> ‘salt’ (2.86.0609)

PYUM *k^wʔiy > Ki k^wʔiy¹ ‘salt’

NOTES. (1) PH *k^{wh}- here is possibly the Hokan ancestor of the Proto-Yuman relativizing prefix *k^w- ‘that which is X / does X’ (Miller 2001: 117; Mixco 1985: 52), in this case ‘that which is salty’. Cf. forms without *k^{wh}- under PH *(a-k^{wh}-)Vʔiy ‘salt’ (?-04). (2) Chimariko /k/ instead of expected /x/ is due to a change *k^{wh}iʔi > *k^{wh}ʔi > *k^wi.

Kroeber 1915: 282 || J. M. Crawford 1976: 185 || Langdon 1979: 638 || Greenberg 1987: 140 || Kaufman 1989: 167 (#(a:-)k^{yh}aʔi ‘salt’ [N]) || Gursky 1990: 1

(k^{wh}-04) PH *k^{wh}ú- ‘negative’

CHI x-..[-na] ‘negative verbal circumfix’ (Jany 2009: 113)

KAR *pu*- ‘no’, *pu*- (prefix) ‘negative; not’

YAN *k^hu*- ‘not; not to have’

PPOM *k^hó- ~ *ʔahk^hó- ‘negative’ > Pk *c^how*, Ps *ʔač^how* ‘be absent’, Pn *šó-* ‘not to be/have’, Pc *c^how* ‘not’, Pne *č^hó*[-t-on], Pe *k^hú*-

SAL *k*- ‘negative’ (a common component in all negative prefixes) (Turner 1987: 132–134)

NOTES. (1) In Chimariko, the initial vowels of verb stems change in the following way after the negative x-: /i/ > /u/, /e/ > /o/ (Jany 2009: 37). This reflects the PH labial vocalism of the negative morpheme. (2) McLendon reconstructs pPom *k^hów ~ *ʔak^hów, but -*w* in Kashaya and Southern Pomo is an absolutive/perfective suffix. The underlying Kashaya stem is *hc^ho*- (Oswalt & Buckley 2018).

Sapir 1917: 9 || Sapir 1920b: 288 || Sapir 1925: 423–424 (*k^u, *ku) || Greenberg & Swadesh 1953: 217 || Sapir & Swadesh 1960: 16 (*qhu-) || Haas 1964: 80 || McLendon 1964: 134 || Gursky 1965: 180–181 || Kaufman 1989: 124 (#k^{yh}u(wa) ‘negative’ [N/S]) || Gursky 1995: 147

1.1.3. Proto-Pomoan *k

The sound correspondence involving Proto-Pomoan *k is represented in Table 3. This correspondence can very tentatively be identified as reflecting the Proto-Hokan glottalized labialized velar stop *k^w.

PH	Chi	Kar	Sha	Ach	Ats	Yan	PPom	Sal	PYum	Ser	PCho
*k ^w	k̥		k̥			k̥	*k̥	k̥			

Table 3. Proto-Hokan *k^w.

(k^w-01) PH *k^wá ‘one’

YAN -*ká*[i]- ‘only one’ (applies to subject of intransitive or object of transitive)

PPOM *ká-, *ká- ‘one’ > Ps *čá*[-ʔa], Pn *čá*-, Pne *čá*- ~ *čá*[-ki], Pe *ká*[li]

McLendon 1964: 135 || Gursky 1974: 200 (#k^a(li)#) || Kaufman 1989: 124 (#k^ya ‘one’ [N]) || Leshchiner & Nikolaev 1992: 385 (*k^hĀHō)

This correspondence can be very tentatively complemented by the following comparisons without reflexes in Pomoan.

(k^w-02) PH *i-k^wa ‘to say, talk’

CHI *n-iķo* <nik^o> ‘talk!’ (2.23.0967)

SAL M *heka*(ʔ) <hék^a> ‘(el hombre) dijo’, <hék^ā> ‘(yo) dije’ (2.86.0621)

NOTES. Cf. PH *k^wá ‘to speak’ (k^w-02).

Gursky 1974: 206 || Kaufman 1989: 130 (#ik^ya [a ~ o] ‘to say; talk’ [N])

(K^w-03) PH *k^wá/iw ‘snow / hail’SHA *káw* ‘snow’ (Silver 1961: 7)SAL A *kew[nél]* ‘k’ewnél’ ‘granizo’ (2.85.0033), M *kaw[nél]* ‘k’awnél’ ‘granizo’, A *kew[nél]* ‘k’əñél’ ‘id.’ (2.86.0592)

Gursky 1974: 206 || Kaufman 1989: 143 (#K’aw ‘hail’ [N])

1.1.4. Proto-Pomoan *x

Proto-Pomoan *x is reconstructed by McLendon (1973: 26) for the correspondence between *x* in Southeastern Pomo and *š* in all other Pomoan languages. Oswalt (1972), Moshinsky (1976: 57) and Buckley (2024: 1228-1229) reconstruct Proto-Pomoan *š for this correspondence, although Buckley notes that “since the six other languages formed a single group after Southeastern split off, it is possible that */x/ > /š/ was an early shared change” (ibid.).

Proto-Pomoan *x forms part of two sound correspondences. The first correspondence is represented in Table 4. This correspondence reflects the Proto-Hokan labialized velar fricative *x^w. Karuk has /x/ before *a, but /f/ before *u, cf. above on PH **k^{wh} > Karuk /p/ before *u. Salinan has /h/ after a consonant. Proto-Pomoan has zero root-finally. Proto-Yuman has *x before *u, *x^w otherwise (*x^w before *u is synchronically impossible in Yuman). Seri has zero (19th century Seri /β/). Proto-Chontal has *x syllable-finally, *x^w otherwise.

PH	Chi	Kar	Sha	Ach	Ats	Yan	PPom	Sal	PYum	Ser	PCho
*x ^w	x ~ h	x, f	x	h	h	w	*x, Ø	x, h	*x ^w , *x	β > Ø	*x ^w , *x

Table 4. Proto-Hokan *x^w.(x^w-01) PH *(i-)x^wál ‘mountain’

PPOM *[wi]xál^y ‘ridge, mountain’ > Pk *wišali* ‘east’, Ps *wiš-a*, Pn *bišá*, Pne *wiša* ‘mountain’, Pe *-wi-šál* in place names

PCHO *-i-x^wal[a] ‘mountain’ > HCh *t-iWála* ‘the mountain, hill, pile’, LCh *i-Wala?* ~ *i-Wala* ‘highlands, mountains’

NOTES. PPom *wi- in *wixál^y is possibly related to PYum *wiy ‘stone; mountain’ (Miller 2018: A71). Leshchiner & Nikolaev 1992: 381 (*Hik_o’ola)

(x^w-02) PH *(u-)x^wí ‘to smell’

PPOM *mi-hxé ‘to stink’ > Ps *mehše-* ‘to smell, transitive or intransitive, good or bad’, Pn *mišé-*, Pse *mxe-*, Pe *mi-šé-* (pPom *mi- ‘protuberance near end of long obj., toe, nose, horn; reckon, read’ [Oswalt 1976: 16])

PYUM *x^wi ‘to smell’ > Co *x^wi* ‘to smell (trans.)’, JT *wə-x^wi* ‘he smells it’, MGI *wə-x^wi* ‘smells something’, Mo *a-h^we-k* ‘smell (something)’, *i-h^wi-v-k* ‘smell, have an odor; smell like; have a disagreeable odor, smell bad’, Qu *a-x^wé* ‘to be odorous’, Hu *h^wi-k*, *h^wi-k* ‘to smell something’, Ha *h^wi-k* ‘to smell something’, Ya *h^wi* ‘smell v.t.’

PCHO *-ux^we- ‘stink, smell’ > HCh *d-uWe-da* ‘it will have an odor’, LCh *We-loh-ma* ‘stink, smell’ Waterhouse 1976: 337 || Kaufman 1989: 165 (#x^we ‘to (have a) smell’ [S]) || Gursky 1990: 8

(x^w-03) PH *(i-)lax^w ‘head’

SHA *?i-nax* ‘hair’ (Silver 1961: 4), *-iraxi-* in *kíraxačúm-u* ‘greyheaded’, *kíwíraxiwakirik* ‘he washed his hair’, *kíwíraxapuk-ucwik* ‘he pushed it up with the top of his head’ (Silver 1966: 78), KON ‘kí’na ‘head’ (Dixon 1905: 216), NRSHA ‘kin’nux’ ‘head’ (ibid.), OKW ‘in’nux’ ‘head’ (ibid.). ACH *láh* ‘head’

ATS *náh[a]* ‘head’ (Talmy n.d., b: 98)

PPOM *[x/k]iná· ‘head’ > Pk *šina·* ‘head’ (elegant), Ps *šin·a*, Pn *šiná·*, Pc *šná·*, Pne *tína·*

PYUM *x^wal > Co *m-x^wat* ‘hair (of the head)’, BI *x^wat* ‘head’, MGI *ə-x^wat* ‘his forehead’

NOTES. (1) The nasal /n/ in Shastan, Atsugewi and Pomoan is due to the Proto-Hokan diminutive alternation *l ~ *n. (2) Konomihu, New River Shasta and Okwanuchu preserve the old meaning ‘head’, pointing to the innovative nature of the meaning ‘hair’ in Shasta. (3) Northeastern Pomo reflects initial *k-, other languages reflect *x-. These may be different fossilized prefixes. (4) McLendon (1973: 78) compares Eastern Pomo *ki·ná·l* ‘crown’, which supports the reconstruction *kiná·l^y for Proto-Pomo. This word may be unrelated to the rest of the set. (5) Yuman form reflects a metathesized form *x^wál.

Sapir 1917: 20 || Bright 1954: 64 || Gursky 1974: 193 (#l/na/ux#) || Greenberg 1987: 137 || Gursky 1989: 16 || Kaufman 1989: 127 (#(i-)La(x^w) or #(i-)La(xu) [l ~ n] ‘head’ [N])

This correspondence can be complemented by the following comparisons without a reflex in Pomoan.

(x^w-04) PH *x^wá ‘enemy / brave’

SAL A <(e)xai>, M <xai> ‘brave’ (Mason 1918: 151)

PYUM *(?-)x^wa ‘enemy’ > Ki *?-x^wa¹* ‘warrior; enemy, foreigner; principally Cocopa’, Co *x^wa* ‘army, soldier, fighter, warrior’, JT *x^wa* ‘enemy’, MGI *?ə-x^wa·* ‘stranger, enemy’, Mo *?a-h^we* ‘enemy’, *?a-h^way* ‘war, battle’, Qu *?a-x^wé* ‘enemy’, Hu *čih^wá* ‘enemy’, Ha *h^wa·č* ‘Apaches (San Carlos and White River)’, Ya *h^wá* ‘fighting, war; enemy’

J. G. Crawford 1976: 312 || Kaufman 1989: 161 (#x^wa(y) ‘enemy’ [N/S]) || Shaul 2020b: 301

(x^w-05) PH *a/i-x^wá ‘bear’

SHA *?é·xa·* ‘bear’ (Silver 1961: 8)

SAL A *ṭ-axá[ɣ]* <traqáʃ> ‘oso’ (2.85.1114)

PYUM *[m]x^wa(·) ‘badger’, *[m]x^wa(·)-tay ‘bear’ (lit. ‘big badger’) > Ki *mx^wa²* ‘swine (aboriginally badger)’, Co *mx^wa* ‘badger’, MGI *māx^wa·* ‘badger’, Mo *māh^wa* ‘badger’, *māh^wa-t* ‘bear’, Ma *mx^we-t* ‘bear’, Qu *max^wá·* ‘badger’, Hu *māh^wa·?* ‘badger’, Ha *?mh^wa·-?a* ‘badger, wild pig’, Ya *mh^wá·* ‘badger’, *mh^wā-t-a* ‘bear’

NOTES. (1) Possibly a tabooistic epithet of bear as ‘brave’, related to the previous root. (2) In Proto-Yuman, *mx^wa(·) means ‘badger’, while *mx^wa(·)-tay (literally ‘big badger’) means ‘bear’. This is exactly the situation found in neighboring Uto-Aztecan languages of the Cupan subgroup, where reflexes of Proto-Uto-Aztecan *hula mean ‘badger’, while reflexes of *hula-wit — the same root with an augmentative suffix — mean ‘bear’ (Hill 2020: 95–96)

Gursky 1988: 21

(x^w-06) PH *(i-)lax^wá ‘fingernail’

CHI *[po]loxol[t]* <poloqot> ‘fingernails’ (2.23.0719)

SHA *?é·raxa* ‘fingernail’ (Silver 1961: 5)

PYUM *(l^y)x^wa[w] ‘fingernail, claw’ > Ki *sal-xa·w²* ‘fingernail, claws’, Co *tax^wáw*, *ṭkax^wáw* ‘claw, nail (of finger or toe)’, MGI *sal^y-əx^wu·w* ‘his fingernails, claws’, Mo *i-sal^y kul^yoho* ‘fingernail, claw, hoof’, Hu *sirwo·?* ‘fingernail; claw (of an animal)’, Ha *srh^woʔo*, *srhvoʔo* ‘fingernail, claw’, Ya *sàl-slh^wō·* ‘fingernail’

PCHO *-[?na]ṭux ‘nail, claw’ > HCh *ṭ-a-?nátuk* ‘the nail (finger or toe); the claw’, LCh *a-ne-tuh* ‘fingernail, toenail, also a bird or animal claw’, Teq <nal> ‘nail’

NOTES. Syllable-finally, *x^w is not allowed in Proto-Chontal, but the labiality was apparently transferred onto the preceding vowel.

Haas 1963: 50–51 || Gursky 1974: 198 (#ilax^wa#) || Waterhouse 1976: 341 || Gursky 1989: 23 || Kaufman 1989: 138 ([i/a/u]-lax^w [a ~ o] 'nails; claw' [N/S])

(x^w-07) PH *a-x^wát 'blood / red' (= t-14)

KAR á·x 'blood; (as a postpound) red'

SHA ?á·xt[a] 'blood' (Silver 1964: 171); OKW <axta' > 'blood' (Dixon 1905: 216)

ACH aht[i] 'blood'

YAN CN wat?-tu-wi 'blood', C pa-wat?(ku)- 'red' (pa- 'in quality'), YN ta-wat?-sa- 'red' (ta-...-sa 'to be in appearance of form')

SAL A ?-á-[k]hat-a <ákhata> 'sangre', ?-é?-[k]hat <é'khat> 'mi sangre', M p-á-[k]hat-a <p'ákhata> 'blood (non-poss.)', p-é?-[k]hat <p'é'khat> 'mi sangre' (2.85.0726), A á-[k]hat-a <á·k^hatə> 'blood' (Jacobsen 1954: 16)

PYUM *x^wat 'blood' > Ki t-kh^wat¹ 'blood', Co n^y-x^wat 'blood', JT x^wat 'blood', MGI ?ə-x^wat 'is red, orange, brown; blood; ripe', Mo n^y-h^wat 'blood', Ma x^wet-m 'red', Qu -i-x^wet 'blood', Pa č-x^wat 'blood', Hu h^wat 'blood, (the color of) red', Ha h^wat 'red; blood', Ya (?-h^wat- 'red; blood'

SER ?aaít 'blood' (pl. ?aaít-lox), 19th century Seri B <avt> 'blood', Pt <a'vat> 'sangre' (Marlett 2010: 37)

PCHO *-x^wac 'blood' > HCh l-a-Wác, LCh a-Wác, Teq <fas>

NOTES. Salinan has <kh> in Harrington's transcription and <k^h> in Jacobsen's fieldnotes. Turner (1987: 67) phonologizes the word as /?ákata?/ because Salinan has no aspirated stops. However, this does not explain the persistent aspiration in Harrington's and Jacobsen's fieldnotes, and I prefer to postulate a biphonemic cluster /kh/ here, where /h/ reflects PH *x^w, and /k/ is a fossilized prefix, possibly the PH relativizing prefix *k^{wh}-, on which see note 1 under (k^{wh}-03)³. Note a similar development in the Kiliwa word for 'blood', which contains the Yuman relativizing prefix *k^w-, see Mixco (1977: 173).

Kroeber 1915: 282 || Sapir 1917: 16 || Sapir 1925: 405 (*axwati ?) || Greenberg & Swadesh 1953: 218 || Bright 1956: 46 || Jacobsen 1958: 198 || Silver 1964: 171 || Gursky 1965: 170 || Gursky 1974: 180 (#a/ix^yawaṭi#) || J. G. Crawford 1976: 309 || Waterhouse 1976: 337, 339 || Webb 1980: 75 || Greenberg 1987: 134 || Gursky 1988: 24–25 || Kaufman 1989: 156 (#(a-)x^wá(-t) 'blood; red' [N/S]) || Leščiner 1989: 163 (*x^waṭa) || Leshchiner & Nikolaev 1992: 376 (*x_oāṭa) || Shaul 2020b: 297

(x^w-08) PH *i-x^wáy '(to) smoke'

PYUM *x^way 'smoke' > Co xyay 'smoke, steam', MGI ?u·xuy 'the smoke', Mo ?a-h^wa· 'smoke (noun)', Hu o-?h^wáy 'smoke, smoke particles', Ha o-h^waya 'smoke', Ya ?ə-h^wáy 'smoke'

PCHO *-ix^we- > HCh d-iWe-?ma 'he will smoke' (cont. g-iWay)

Waterhouse 1976: 337 || Kaufman 1989: 142 (#ax^wa 'smoke' [N/S])

The following two comparisons also seem to reflect the same protophoneme, given that in Proto-Yuman, *x^w is not found before *u.

(x^w-09) PH *a-x^wú 'nose'

CHI hohu <hohu> 'nose' (2.20.0024), hoxu <hoqu> 'nose' (2.23.0349)

KAR [y]úf[iv] 'nose', [y]úf-ma·n 'split-open salmon head' (má·n 'skin')

PYUM *xu(·) 'nose' > Ki phi?¹ 'nose', Co ?i-xú 'nose, snout, beak', JT xu 'nose, nostrils', MGI ə-xu· 'his nose', Mo i-hu 'nose, beak', Ma i-xu· 'nose', Qu i-xú· 'nose', Pa xu· 'nose', Hu hu?, hu· 'head', Ha hu-?u 'head', Ya hū·, hū: 'nose; head'

³ With the implication that the original Proto-Hokan polysemy 'blood / red' was resolved in Salinan by creating a derivative *k^{wh}-x^wat 'blood = that which is red'.

NOTES. The vowel *u* in Karuk *yuf-* results from assimilation to the subsequently lost vowel of the following syllable.

Kroeber 1915: 282 || Sapir 1920b: 284 || Sapir 1925: 407-408 (*yaxu?) || Greenberg & Swadesh 1953: 219 || Jacobsen 1958: 201 || Gursky 1968: 34 || Gursky 1974: 199 (#xu#) || J. G. Crawford 1976: 316 || J. M. Crawford 1976: 183 || Greenberg 1987: 139 || Kaufman 1989: 127 (#(ya)xu [a ~ o] 'nose' [N/S], #fi or #pxi 'nose' [N/S])

(x^w-10) PH *x^wulu(k) 'into'

CHI *-xun / -xunok* 'in, into' (verbal directional suffix) (Jany 2009: 134)

KAR *-fúruk* 'into an enclosed space; indoors; inside a living house'

YAN *-wul-* 'in, entering', absolute form Y *-wu·lu* 'into'

NOTES. The nasal /n/ in Chimariko is apparently due to the Proto-Hokan diminutive alternation *l ~ *n.

Sapir 1917: 21 || Sapir 1925: 423 (*xwulu) || Kaufman 1989: 120 (#pxuLu [l ~ n] or #fuLu adv/dir 'into the house' [NC])

The second correspondence involving Proto-Pomoan *x is represented in Table 5. This correspondence, in my view, reflects the Proto-Hokan velar fricative *x. It is tempting to reconstruct here *š or *ʃ for Proto-Hokan, i.e., a “hushing” sibilant, distinct from a “hissing” *s. Indeed, Kaufman reconstructs Proto-Hokan *š for this correspondence. However, this would leave us with a system that has *x^w, but no non-labialized velar fricative (see below for the reconstruction of *k^h in place of Kaufman's *x^y). Therefore, I propose to reconstruct Proto-Hokan *x for this correspondence, despite the fact that the overwhelming majority of its reflexes are sibilants. The unusual reflex /t/ in Chimariko can be explained by a following sequence of changes: *x > *ʃ > /t/. While Karuk has /š/ in the only available example, *i-š* < PH *ixi, /s/ would be expected in a position not after /i/. Note that Seri reflects PH *s and *x as /š/ and /s/ respectively. If we were to reconstruct PH *x as *š, we would have to assume an unlikely flip-flop development in Seri.

PH	Chi	Kar	Sha	Ach	Ats	Yan	PPom	Sal	PYum	Ser	PCho
*x	t̚	[s,] š	s	s	s	s	*x	s ~ š	*ʃ	s	*š

Table 5. Proto-Hokan *x.

(x-01) PH *a/i-xál 'hand, arm'

CHI *hiṭa* <hitra> 'hand' (2.23.0200)

PPOM *ʔixál ~ *ʔixál^y 'arm' > Pk *ʔiša*, Ps *ʔi-šan*, Pn *šá*, Pc *šá*, Pse *xal*, Pe *šál* 'wing'

PYUM *šal^y 'arm, hand' > Ki *sal*¹ 'hand, etc.', Co *ʔi-šá-t^y* 'arm, hand, finger, wing (of a chicken)',

JT *šat^y* 'arm, hand', MGI *ə-sat^y* 'his hand, arm, fingers', Mo *i-sal^y* 'hand, arm, sleeve', Ma *i-ša-l^y*

'arm/hand', Qu *i-šá-l^y* 'hand', Pa *šal* 'hand', Hu *sal* 'hand', Ha *sal* 'hand, hands', Ya *sál* 'hand'

SER *i-seł* ~ *i-sil* 'shoulder joint; pectoral fin' (pl. *i-seł-ox* (sg. possessor), *i-seł-kox* (pl. possessor), abs. *ʔá-seł*)

PCHO *-(e)šal > HCh *l-a-šát* 'wing, feather, fin', *d-ešal-ʔma* 'he will grasp it with his hand (to carry it)', *d-ešal-af-ʔma* 'he will help someone to get up by giving them his hand', *d-ešal-fe-ʔma* 'he will put his hand (inside something) to remove (something)', *d-ešal-u-ʔma* 'he will lead by the hand with force', LCh *i-šal* 'wing'

NOTES. While the Chontal reflexes mean 'wing, feather, fin', traces of an earlier meaning 'hand' are preserved in the semantics of verbal derivatives: 'to grasp with hand', 'to help someone to get up by giving them his hand', 'to put one's hand (inside something) to remove (something)', 'to lead by the hand with force'.

Kroeber 1915: 282 || Greenberg & Swadesh 1953: 218 || Haas 1963: 50, 52 || Haas 1964: 79 || McLendon 1964: 137 || Gursky 1965: 178 || Gursky 1974: 178 (#išal#) || J. G. Crawford 1976: 313 || J. M. Crawford 1976: 181 || Water-

house 1976: 337, 341 || Langdon 1979: 636 || Greenberg 1987: 133 || Gursky 1988: 19-20 || Kaufman 1989: 136 (#i-šál' 'arm; wing; shoulder' [N/S]) || Leščiner 1989: 165 (*ʔešalΛ) || Leshchiner & Nikolaev 1992: 379 (*ʔešálΛ)

(x-02) PH *a-xí 'to name', *a-xí 'name'

PPOM *ʔahxí 'name' > Pk ší[hcima°č] 'to name', Ps ʔahši[·yaw], Pn ší, Pc ší, Pne [ʔawi]-šé[m] 'my name', Pse xi- 'to name', xin 'name', Pe ší

SAL A <a'se>, M <as> 'to name, call' (Mason 1918: 139)

PYUM *ši- 'to name' > Ki t-ʔ-si-² 'relate, (re)count, tell (something)', Co ší 'to name, call by name, give a name to', JT wə-ši-č, ši-č 'he names him, he calls him by name; he is named; he reads', MGI ču-xi- 'names, calls by his name', Mo ase-k 'call, name (verb)', Qu ašé 'to name', Hu si-k 'to name, to call', Ha si-k 'to give a name', Ya pà:ʔsī:ti 'I name them'

SER i-[t]áasi 'name' (pl. i-[t]áasi-tox, abs. ʔa-[t]áasi)

Haas 1964: 80 || Gursky 1974: 199 || J. G. Crawford 1976: 316 || Langdon 1979: 638 || Gursky 1989: 24 || Kaufman 1989: 130 (#(a:)si [s ~ š] 'name; to name' [N/S]) || Shaul 2020b: 304

(x-03) PH *(a/i-)xí 'meat, flesh'

CHI hiṭi[y] <hitrij> 'meat' (2.20.1049)

KAR í-š 'meat, flesh, body'

SHA [ʔúp]si 'meat' (Silver 1964: 172)

YAN N [pa]si, Y [pa]hsi 'flesh, meat, body'

PPOM *[bi]hxé 'deer, meat' > Pk bihše, Ps behše, Pn bišé, Pc pšé, Pne béhše, Pse bxé, Pe bi-šé

SAL A [ʔ]aš <ʔ> 'ciervo' (2.85.1138), M [p]aš <p'ác> 'elk', A [ʔ]aš <ʔ> 'id.' (2.87.0320)

PYUM *ši (or *si) > Ki siʔ¹ 'flesh, meat'

SER i-[pṣá]si 'meat, flesh, soft part of plant' (pl. i-pṣáš)

PCHO *-i-ši[k] 'meat, flesh' > HCh l-išik 'meat', LCh i-šik ~ a-šik 'flesh or meat, of human, animal, bird, fruit, whatever', Teq <çi> 'flesh, meat'

NOTES. (1) Shasta, Yana and Pomoan forms may reflect compounds with the first part PH *a/i- pá 'deer'. (2) Salinan p- looks like a widespread fossilized nominal prefix but may also be a fossilized first part of a compound.

Sapir 1917: 4-5 || Greenberg & Swadesh 1953: 219 || Haas 1964: 77 || McLendon 1964: 133 || Gursky 1965: 170 || Gursky 1968: 34 || Gursky 1974: 180 (#isi# ~ #isi + Vpa# ~ #Vpa + isi#) || Campbell 1976: 163 || J. G. Crawford 1976: 315 || Waterhouse 1976: 339, 341 || Webb 1980: 76 || Webb 1981: 82 || Greenberg 1987: 244 || Kaufman 1989: 126 (#(h)i:-ši 'meat, flesh; body' [N/S]; NC also 'person, Indian; man, male, husband') || Leshchiner & Nikolaev 1992: 394 (*^mbHéšö)

(x-04) PH *(a/i-)ximálk 'ear' (= k-06)

SHA ʔisa-k 'ear' (Silver 1961: 4); OKW <is'sawak> 'ear' (Dixon 1905: 216)

ACH issat 'ear'

ATS ʔasmak 'ear' (Talmy n.d., b: 88)

YAN CN malʔk[u], Y manʔk[u] 'ear'

PPOM *xi-mán^y 'ear' > Pk šima, Ps ši-mo[-mo] (mo 'hole'), Pn šimá, Pc šmá, Pne šima, Pse xmán[ca], Pe ši-má.

PYUM *šmal^yk 'ear' > Ki smaql¹ 'ear', Co šma-l 'ear', JT šmal^y, šəmal^y 'ears', MGI xəmat 'his ear(s)', Mo i-smal^yk 'ear; dried peach', Ma šmal^yk 'ears', Pa šmalk 'ear', Hu smalk 'large ear', smark 'small ear', Ha smalk 'ears (large ears, such as of deer, elephant)', smark 'ear, ears (small ears, such as of humans)', Ya smálk-a 'ear'

NOTES. This word with its nontrivial phonotactic structure may be an old compound.

Kroeber 1915: 282 || Sapir 1917: 10 || Dixon & Kroeber 1919: 106 || Sapir *apud* Dixon & Kroeber 1919: 109 || Sapir 1925: 406 (*isama, *isma(l)k- ?) || Greenberg & Swadesh 1953: 218 || Bright 1956: 46 || Olmsted 1957: 138 || Haas

1963: 43-46 || Haas 1964: 77-78 || McLendon 1964: 131 || Gursky 1974: 185 (#isamarak'u# ~ #isamak'ura#) || J. G. Crawford 1976: 312 || Langdon 1979: 637 || Webb 1980: 76 || Greenberg 1987: 233 || Gursky 1989: 4 || Kaufman 1989: 128-129 ([i/a]-ma-lʷ(-Ku) [lʷ ~ nʷ] 'ear; to hear', #i-sa-má 'ear' [N/S]) || Leščiner 1989: 162 (*šimal-(k'u)) || Leshchiner & Nikolaev 1992: 374 (*šimálΛ-(k'u))

(x-05) PH *(w)a-xú 'digging stick'

SHA [k]was- in *kʷás-hi-hú?* 'mountain mahogany' (Silver 1964: 172) (-*hi-hu?* 'tree, bush' [Silver 1964: 176])

YAN N *wasu*, Y *wahsu* 'digging stick'

PPOM *wa(H)xú > Pe *wa-šú* 'mountain mahogany', *wa-šú-xây* 'digging stick' (McLendon n.d.: 30)

SAL A *ṭ-é-wis* <tréwis> 'baston', *ṭ-m-é-wis* <tréméwis> 'tu ...', *ṭ-ewis-o* <trewiso> 'his...' (2.86.0042), M *ṭ-a-wás* <trāwás> 'baston', M *ṭ-é-wes* <tréwes> 'digging stick' (2.88.0780)

PCHO *-a-šu? 'hoe' > HCh *l-a-šú?* 'the hoe', LCh *a-šu?* 'type of agricultural tool for sowing seed, like a long-handled narrow spade or a pointed stick'

NOTES. Silver (1964: 172) notes that digging sticks were made of mountain mahogany.

Haas 1964: 77 || McLendon 1964: 130 || Silver 1964: 172 || Gursky 1974: 185 || Kaufman 1989: 160 (#wasu [s ~ š] 'digging stick' [N/?+Cho /aʃú7/])

(x-06) PH *a/i-xú 'raw'

YAN YN *-[x]su[i]-*, *-[ʔ]su[i]-*, *-su[i]-* 'unripe, uncooked'

PPOM *[q]a(h)xó- 'raw, alive' > Pk *qašo-q-* 'get well, be healed, be cured' (Oswalt & Buckley 2018), Ps *kahšo*, Pn *kašó*, Pne *kašóy*, Pe *qa-šóy* 'raw, alive, well'

PYUM *(x-p-)ši[w] ~ *(x-p-)šu[w] 'raw' > Ki *p-siw*¹ 'green, uncooked, raw', Co *xpsiw* 'be blue or green', JT *xəpši-w* 'it is green', MGI *xəpəši-w* 'is blue, green', Mo *i-vsū-k* 'be raw, uncooked', Ma *xvəšu-k* 'blue / green', Qu *a-vəšú* 'to be raw', Hu *suw-k* 'to be unripe; to be fresh, to be green/blue', *vasu-k* 'to be blue, to be green', Ha *suw-k* 'green (not ripe)', *vasuw-k* 'to be blue or purple; to be blue or green', Ya *hà-vsú(w)i* 'blue, green'

SER *k-is* 'raw' (conjug.: *t-is*)

McLendon 1964: 135 || Gursky 1974: 202 || J. G. Crawford 1976: 317 || Greenberg 1987: 139 || Gursky 1989: 29 || Kaufman 1989: 156 (#(P-)išiw [ʔglott/asp] 'raw; green; to heal, recover' [N/S], #(xʷ)əšúy 'raw, unripe; green/blue; alive' [N/S])

This correspondence can be complemented by the following comparisons without a reflex in Pomoan.

(x-07) PH *m-a-xá/úl 'yellow, red'

SER *k-másot* 'yellow; blond (hair)' (conjug.: *t-másot*; *t-mast*)

PCHO *um-šal-i > HCh *ʔun-šal-i* 'red'; pCho *a-šal-e > HCh *a-šale* 'tan'

NOTES. The word may contain stative prefix *m-.

Kaufman 1989: 156 (#masó(L) 'red, brown' [N/S]) || Gursky 1990: 22

(x-08) PH *i-xi 'man, person'

CHI *ʔiṭi* <ṭrṭ> 'man' (2.23.0124)

SHA *ʔis* 'Indian, person' (Silver 1964: 175); NRS HA *gè'ic* 'man' (Dixon 1905: 216)

ACH *ís* 'person; Indian'

YAN CN *hisi*, Y *hihsi* 'man; (Y) husband, husband's brother'

Sapir 1917: 9 || Sapir 1920b: 283 || Jacobsen 1958: 200 || Sapir & Swadesh 1960: 17 || Haas 1964: 77 || Gursky 1965: 180 || Gursky 1968: 33 || Gursky 1974: 197 (#isi#) || Webb 1981: 82 || Greenberg 1987: 242 || Kaufman 1989: 126 (#(h)i:-ši 'meat, flesh; body' [N/S]; NC also 'person, Indian; man, male, husband')

(x-09) PH *xú· ‘net’

YAN C *suu[wi]*- ‘to fish with a seine’PYUM *ʂu- > Mo *su·[ʔak]* ‘fishnet’

SER ʔá-so ‘fishing net’

J. G. Crawford 1976: 316 || Gursky 1989: 24

(x-10) PH *xúk ‘to be tired’ (= k-09)

CHI *ʔuk* ‘be tired, exhausted’ (Conathan 2002: 30)PCHO *-šox- ‘be tired’ > HCh *i-šog-i* ‘tiresome, tiring’, LCh *soh-ta* ‘be tired’

1.1.5. Comparison with Kaufman’s reconstruction

Kaufman reconstructs the following labialized velars for Proto-Hokan: *k^w, *k^w, *k^{wh} and *x^w (see Table 6).

PH	Chi	Kar	Sha	Ach	Ats	Yan	Wsh	PPom	Sal	PYum	Ser	PCho	Tol
*k ^w	k, q	k	k	k, q		k	k	*k ^y , *k, *q	k	*k ^w , [q ^w]	k	*k ^w	k
*k ^{wh}	kh, qh	k	k	k ^h , q ^h		k ^h	k ^h	*k ^{yh} , *k ^h , *q ^h	k	*k ^w , [q ^w]	k	*k ^w	k ^h
*k ^w	k', q'	k	k'	k', q'		k'	k'	*k ^{y'} , *k', *q'	k'	*k ^w , [q ^w]	k	*k ^w	k'
*x ^w	h, x	x	x	x		x	x	*h	x	*x ^w	x	*x ^w	h

Table 6. Correspondences of Hokan labialized velar obstruents according to Kaufman (1989).

Kaufman postulates that in Chimariko, Achumawi and Pomoan, labialized velars can yield uvulars, although he does not specify the precise conditions of such a development. I was unable to find reliable cases where labialized velars of Proto-Yuman or Proto-Chontal correspond to uvulars in Chimariko or Proto-Pomoan.

1.2.0. Correspondences involving Proto-Pomoan uvular stops

The following uvular stops are reconstructed for Proto-Pomoan: *q, *q^h and *q̣ (McLendon 1973: 24–25).

1.2.1. Proto-Pomoan *q

The sound correspondence involving Proto-Pomoan *q is represented in Table 7. This correspondence reflects the Proto-Hokan plain velar stop *k. The same phoneme in a root-final position is reflected as zero in Pomoan. Note that Proto-Yuman can have *k^y, *k and *q in reflexes of the same etymon: the alternation *k^y ~ *k ~ *q forms part of a complex Yuman system of diminutive/augmentative consonant symbolism. Glottalized /k/ found root-finally in Shasta and Yana probably reflects a merger of the expected /k/ with the suffix *-ʔ, on which see Zhivlov (2025: 63).

PH	Chi	Kar	Sha	Ach	Ats	Yan	PPom	Sal	PYum	Ser	PCho
*(-)k-		x	c	c	c	c	*q		*k ^y ~ *k ~ *q	k	*k
*-k	k	Ø	k ~ ḳ	t	k	k ~ ḳ	Ø	k	*k	k	*x

Table 7. Proto-Hokan *k.

(k-01) PH *(a-)ká ‘mother’s mother’

SHA *ʔac-wit* ~ *ʔac-it* ‘mother’s mother’ (Silver 1966: 64)

YAN CN *ʔacu[wi]*, Y *ʔahcu[wi]* ‘paternal grandmother (in C), woman’s daughter’s child, mother’s mother (in Y)’

PPOM *qá-·čì ‘mother’s mother’ > Pk *-qa-*, *-qa°s-*, Pn *-káʔ*, Pne *ká·čì-*, Pse *-qa*, Pe *qá·č*, *-qá·* (pPom *--·čì ‘one’s own kinsman in generations above ego’)

PYUM *(n-)ka(·)-w ‘mother’s mother’ > Ki *h-qha·w¹* ‘be maternal grandmother’, Co *n^y-ka* ‘mother’s mother’, JT *nə-ka·w* ‘grandmother’, Mo *na-kaw* ‘mother’s mother, grandmother’, Ma *n-k^{yew}* ‘grandmother (maternal)’, Hu *kór-a* ‘mother’s mother, grandmother’, Ha *nko-ʔo* ‘grandmother (mother’s side)’, Ya *kól-a* ‘mother’s mother’ (*n- ‘kinship terms prefix’ [Miller 2001: 86])

SER *aa-k[t]* ‘maternal grandmother; great-aunt (sister of mother’s mother); wife of mother’s mother’s brother’ (pl. *áa-ka[t]-x*, abs. *ʔapé-k[t]*)

NOTES. (1) This kinship term forms part of a mini-paradigm of terms for grandparents: PH *(a-)pá ‘father’s father’, PH *(a-)k^wá ‘mother’s father’ (k^w-01), PH *(a-)má ‘father’s mother’, PH *(a-)ká ‘mother’s mother’ (k-01). All four terms for grandparents are preserved in Pomoan, Yuman and Seri. (2) Shasta *-wit* and Yuman **-w* are suffixes found in other terms for grandparents.

Sapir 1917: 3 || McLendon 1964: 134 || Gursky 1974: 191 || J. G. Crawford 1976: 316 || Gursky 1989: 12 || Kaufman 1989: 145 (#k^ya:(wV) ‘mother’s mother’ [N/S])

(k-02) PH *(a/i-)ká ‘to bite’

KAR *ax* ‘(animal) to bite, to rend with the teeth; to kill’

ATS *icⁱ* in *p-icⁱ* ‘to bite’ (Talmy n.d., a: 111)

YAN *ca-* ‘to eat, bite’

PPOM *qa- ‘between two strongly opposed, sharp-pointed or gripping forces, such as the jaws and teeth’ > Pk *qa-*, Ps *ka-*, Pc *ka-*, Pn *ka-*, Pse *q-*, Pe *qa-*

PCHO *ka- > HCh *di-ga-ʔma* ‘he will bite into’

NOTES. Superscript *-i* in Atsugewi is a morphophonological symbol with unclear value used by Talmy.

Haas 1964: 81 || McLendon 1964: 137 || Gursky 1974: 180 || J. G. Crawford 1976: 308 || Kaufman 1989: 139 (#qa [Chi q] ‘to bite; by biting, with the teeth (incorp prepound)’ [N/S])

(k-03) PH *pakáy ‘manzanita berry’ (= p-04)

ACH *pacca* ‘black manzanita berry’, *pa-ca* ‘black manzanita’

YAN N *pa-cai-na* ‘manzanita berry’

PPOM *bahqáy ‘manzanita berry’ > Pk *bahqay*, Ps *bahkay*, Pn *bakáy*, Pc *ba-káy*, Pne *báhkay* ‘berry’, Pse *bqay*, Pe *bá·qay*

NOTES. Because of irregular sound correspondences, Karuk *fá·θ* ‘manzanita berry’ is either unrelated or results from diffusion.

Sapir 1917: 4 || Haas 1964: 79 || McLendon 1964: 133 || Gursky 1974: 197 || J. G. Crawford 1976: 315 || Webb 1979: 2 || Gursky 1989: 22 || Kaufman 1989: 154 (#fa:ca ‘manzanita’ [NC]) || Leshchiner & Nikolaev 1992: 390 (*bāč̣.á (~-ō))

(k-04) PH *(a-)kúl ‘long’

KAR *-xár[ah]* ‘long (of spatial measurement), tall’ (compounded with nouns)

YAN *cul-* ‘long, tall’, pl. *cauli-*

PPOM *ʔahqól ‘long’ > Pk *ʔahqol*, Ps *ʔahkon* (sg.), *ʔahko·n-ha* (pl.), Pn *kól* (sg.), *kó·l-ay* (pl.), Pc *kól*, Pne *kól-*, Pe *qól-a* (pl.)

PYUM *k^yul^y ~ *qul^y 'long' > Co *kuł^y* 'be long', JT *kuł* 'it is long', MGI *ʔa-qul* 'is long', Mo *ʔa-k^yu·l^y-m* 'be long', Ma *qol-m* 'be long', Qu *ʔa-qól* 'to be long', Hu *k^yul-k* 'to be long', Ha *k^yul-k* 'to be long', Ya *k^yú(-)l-i* 'long'; PYUM *kur 'far' > Co *kur* 'distance in space or time', JT *kur* (n.) 'far' /lit. a distant place/, MGI *ʔa-kur* 'is distant, far; long (time)', Qu *ʔa-kór* 'to be distant (in time or space)', Pa *ku-r* 'far', Hu *kur* 'long ago', Ha *kur* 'long ago; already; far away', Ya *kúr-a* 'long ago'

SER *k-áakox* 'large, big' (conjug.: *t-áakox*; *t-aak^wł*)

PCHO *kul[i] 'far' > HCh *a-gulíʔ* 'far', LCh *kuł^yiʔ* ~ *kuł^yi* 'far, far away from, distant'

NOTES. The first vowel *a* in Karuk *-xárah* apparently results from assimilation to the vowel of the following syllable.

Sapir 1917: 10 || McLendon 1964: 133 || Gursky 1974: 187, 196 || J. G. Crawford 1976: 315 || Waterhouse 1976: 337 || Langdon 1979: 638 || Greenberg 1987: 241 || Gursky 1989: 7 || Kaufman 1989: 146 (#Kur 'far, distant' [S], #aqol^y 'long' [N/S]) || Leshchiner & Nikolaev 1992: 381 (*qál(á))

(k-05) PH *(i-)ák 'bone'

SHA *ʔa·k* 'bone' (Silver 1961: 6)

PPOM *ʔihyá(·) 'bone' > Pk *ʔihya·*, Ps *ʔi·ha*, Pn *yá·*, Pc *yá·*, Pne *híya*, Pse *ya*, Pe *Ya*

SAL A [*ʔax*]á·k <áqāk> 'bone (non-poss.)', [*ʔáx*]a·k <áqāk> 'my bone' (2.85.0729), M [*páx*]ak <p'áqak> 'mi hueso' (2.85.0730)

PYUM *ak 'bone' > Ki *haq¹* 'bone', *n^y-haq¹* 'a bone, someone's bone; skull', Co *ya·k* 'bone (removed from the body), skeleton', *n^y-ya·k* 'bone', JT *ak* 'bone', MGI *aq* 'his bone', Mo *n^ya-sak* 'bone', Hu *čiya·k* 'skeleton, bone', Ha *čya·k* 'bones', Ya *čyá·k-a* 'bone'

SER *í-[t]ak* 'bone; rib, petiole, limb, branch, stem; skeleton or rib (of boat)' (pl. *í-[t]axk* [sg. possessor], *í-[t]ax-ox* [pl. possessor])

NOTES. Pomoan does not allow two vowels in hiatus, so *y in *ʔihyá(·) is a predictable hiatus-filler, while the *h is a laryngeal increment, added when *y was already in place.

Kroeber 1915: 282 || Sapir 1917: 8-9 || Sapir 1925: 405 (*ihyaka?) || Greenberg & Swadesh 1953: 218 || Gursky 1974: 181 (#i/aʔyak#) || J. G. Crawford 1976: 309 || Langdon 1979: 637 || Greenberg 1987: 197 || Kaufman 1989: 138 (#i(H)θa 'bone' [N/S], #a(:)K [K ~ K'] 'bone' [N/S]) || Leščiner 1989: 164 (*ʔöja) || Leshchiner & Nikolaev 1992: 378 (*ʔöjá) || Shaul 2020b: 297-298

(k-06) PH *(a/i-)ximálk 'ear' (= x-04)

SHA *ʔisa·k* 'ear' (Silver 1961: 4); OKW <is'sawak> 'ear' (Dixon 1905: 216)

ACH *issat* 'ear'

ATS *ʔasmak* 'ear' (Talmy n.d., b: 88)

YAN CN *malʔk[u]*, Y *manʔk[u]* 'ear'

PPOM *xi-mán^y 'ear' > Pk *šima*, Ps *ši-mo[-mo]* (*mo* 'hole'), Pn *šimá*, Pc *šmá*, Pne *šima*, Pse *xmán[ca]*, Pe *ši-má·*

PYUM *šmal^yk 'ear' > Ki *smaql¹* 'ear', Co *šma·l* 'ear', JT *šmał^y*, *šəmał^y* 'ears', MGI *xəmał* 'his ear(s)', Mo *ismal^yk* 'ear; dried peach', Ma *šmal^yk* 'ears', Pa *šmalk* 'ear', Hu *smalk* 'large ear', *smark* 'small ear', Ha *smalk* 'ears (large ears, such as of deer, elephant)', *smark* 'ear, ears (small ears, such as of humans)', Ya *smálka* 'ear'

NOTES. This word with its nontrivial phonotactic structure may be an old compound.

Kroeber 1915: 282 || Sapir 1917: 10 || Dixon & Kroeber 1919: 106 || Sapir *apud* Dixon & Kroeber 1919: 109 || Sapir 1925: 406 (*isama, *isma(l)k-?) || Greenberg & Swadesh 1953: 218 || Bright 1956: 46 || Olmsted 1957: 138 || Haas 1963: 43-46 || Haas 1964: 77-78 || McLendon 1964: 131 || Gursky 1974: 185 (#isamarak'u# ~ #isamak'ura#) || J. G. Crawford 1976: 312 || Langdon 1979: 637 || Webb 1980: 76 || Greenberg 1987: 233 || Gursky 1989: 4 || Kaufman 1989: 128-129 (#[i/a]-ma-l^y(-Ku) [l^y ~ n^y] 'ear; to hear', #i-sa-má 'ear' [N/S]) || Leščiner 1989: 162 (*šimal-(k'u)) || Leshchiner & Nikolaev 1992: 374 (*šimálA-(k'u))

This correspondence can be complemented by the following comparisons without a reflex in Pomoan.

(k-07) PH *mák ‘back’

YAN *mak[i]* ‘back, backbone’

PYUM *mak ‘back; behind’ > Ki *h-maq¹* ‘give one’s back, turn backward’, Co *mak* ‘(space or time) back of, behind, in the rear of’, JT *mak* ‘area in back or behind, in or to an area in back or behind’, MGI *ə-mak* ‘is behind, in back’, Mo *i-mak* ‘back (of the body)’, Ma *mak* ‘back, behind (something)’, Qu *a-mák* ‘behind’, Pa *māk* ‘back; behind; beyond’, Hu *mák-o* ‘back (of the body)’, Ha *mak-o* ‘back (noun; bodypart)’, Ya *mák-a* ‘back, backward’

Sapir 1917: 10 || Greenberg & Swadesh 1953: 218 || Jacobsen 1958: 198 || Haas 1964: 77 || McLendon 1964: 129 || Gursky 1965: 169 || Gursky 1974: 179 (#asilmak#) || J. G. Crawford 1976: 308 || Gursky 1988: 20-21 || Kaufman 1989: 126 (#(i)mak^y ‘back’ [N/S]) || Leščiner 1989: 164 (*mbak^ə) || Leščiner 1990: 36 (*mbāk^ə) || Leshchiner & Nikolaev 1992: 377 (*^mbāk^ó)

(k-08) PH *a/i-púk ‘neck; nape of neck’ (= p-20)

KAR [*v*]ú·p ‘neck’

SHA *ʔé·pk[a]* ‘nape of the neck’ (Silver 1964: 174)

ACH [*w*]ápt[*i*] ‘neck’

ATS *ʔúpuk[i]* ‘neck’ (Talmy n.d., b: 107)

PYUM *puk ‘nape; base’ > Ki *puq¹* ‘under, beneath, nether; inside, under cover’, *n^y-puq¹* ‘a spine, someone’s spine’, Co *mpuk* ‘nape of neck’, JT *i-puk* ‘nape of neck, back of neck’, MGI *əyəpuk*, *i-puk* ‘his neck, nape’, Mo *hipuk* ‘base, bottom, foundation; root’, Ma *mi-puk* ‘neck’, Pa *ipu·k* ‘neck’, Hu *yip^húk* ‘back of the neck’, Ha *puk-l* ‘in the corner; at the bottom, down below’, Ya *púk* ‘bottom, foot (figurative)’, *i-púk* ‘neck’

PCHO *-epux ‘nape’ > LCh *-epúh* ‘nape’ (Waterhouse 1976: 337)

NOTES. The vowel *u* in Karuk *vú·p* results from assimilation to the subsequently lost vowel of the following syllable.

Dixon & Kroeber 1919: 106 || Olmsted 1956: 75 || Jacobsen 1958: 201 || Haas 1963: 48-49 || Silver 1964: 174 || Gursky 1974: 199 (#i^upaki#) || Waterhouse 1976: 337, 341 || Gursky 1989: 24 || Kaufman 1989: 129 (#[i:/u]-pu(K) ‘nape; neck; shoulder; arm’) || Leščiner 1989: 163 (*wəp^húk^hi) || Leshchiner & Nikolaev 1992: 375 (*wāp^húk^hi)

(k-09) PH *xúk ‘to be tired’ (= x-10)

CHI *ʔuk* ‘be tired, exhausted’ (Conathan 2002: 30)

PCHO *-šox- ‘be tired’ > HCh *i-šog-i* ‘tiresome, tiring’, LCh *soh-ta* ‘be tired’

(k-10) PH *-k ‘hither’

CHI *-ku* ‘hither’ (Jany 2009: 40), *-k^h <-k^s* ‘motion towards here’ (Berman 2001: 1050-1051)

KAR *-uk* ‘to here, hither’

SHA *-k* ‘hither’ (Silver 1964: 173)

ACH *k* ‘hither’

ATS *-ik* ‘hither’ (Talmy 1972: 40)

PYUM *-k ‘hither’ > Ki *-k*, Co *-k*, JT *-k* (Miller 2001: 70), MGI *-k* (Langdon 1970: 96-97), Ma *-k* (Gordon 1986: 46), Qu *-k* (Halpern 1947a: 28), Pa *-k* (Joel 1966: 35), Ya *-k* (Shaterian 1983: 104)

NOTES. Judging by the Karuk and Achumawi reflexes, the developments expected for root-final position did not apply to suffixes.

Sapir 1917: 20 || Dixon & Kroeber 1919: 107 || Sapir *apud* Dixon & Kroeber 1919: 111 || Sapir 1925: 414 (*-k^h(w)i) || Jacobsen 1958: 201 || Olmsted 1959: 639 || Silver 1964: 173 || Gursky 1965: 178 || J. M. Crawford 1976: 182 || Kaufman 1989: 119 (#-uk^y ‘hither’ [N/S]) || Gursky 1995: 143 || Zhivlov 2018: 154-157

1.2.2. Proto-Pomoan *q^h

The sound correspondence involving Proto-Pomoan *q^h is represented in Table 8. This correspondence, in my view, reflects the Proto-Hokan aspirated velar stop *k^h. This is the well-known correspondence symbolized as *x^y by Sapir, Gursky and Kaufman. Reconstructing *x^y, however, requires postulation of a repeated typologically rare shift “fricative > affricate” in Shasta and Atsugewi and “fricative > stop” in Proto-Pomoan and Chimariko. Since there are no other candidates for a Proto-Hokan *k^h, I propose to reconstruct *k^h for this correspondence. Karuk has *x* before **u* and *s/š* otherwise (the choice between *s* and *š* follows a synchronic rule: *š* is found after *i*). Yana has *x* before **a*, N *s*, Y *c^h* before **i*, and *k^h* before **u*. Salinan has *š* or *šx* before **a* and **i*, but *k* before **u*. Yuman has **x* before **a* and **s* before **i*. Seri has *χ* before **a* and **u*, but *s* before **i*.

PH	Chi	Kar	Sha	Ach	Ats	Yan	PPom	Sal	PYum	Ser	PCho
*k ^h	q ^h	s, š, x	c	s	c ^h	x, c ^h , k ^h	*q ^h	š(x), k	*χ, *s	χ, s	*χ

Table 8. Proto-Hokan *k^h.(k^h-01) PH *a-k^há ‘water’

CHI *ʔaq^ha* <ʔákhà> ‘water’ (2.23.0328), *aq^ha* <äqǎ> ‘water’ (Berman 2001: 1071)

KAR *á·s* ‘water; juice’

SHA *ʔác·a* ‘water’ (Silver 1964: 176); NRSHA <gaʔatsʔ> ‘water’ (Dixon 1905: 216); OKW <atʔsa> ‘water’ (ibid.)

ACH *as* ‘water’

ATS *ʔác^hi* ‘water; watering place’ (Talmy n.d., b: 61)

YAN *xa-na* ‘water, liquid’

PPOM *ʔahq^há ‘water’ > Pk *ʔahq^ha*, Ps *ʔahk^ha*, Pn *k^há*, Pc *k^há*, Pne *k^há*, Pse *χá*, Pe *xá*

SAL A *t-šaʔ* <ʔfǎ>, M *ʔ-šaʔ* <ʔrǎʔǎ> ‘water’, A M *ʔ-é-[t]aš* <ʔréʔaš> ‘mi agua’ (2.85.0084)

PYUM *(ʔ-)*xa* ‘water’ > Ki *ʔ-xaʔⁱ* ‘water’, Co *xa* ‘water’, JT *xa* ‘water’, MGI *ʔa-xa·* ‘water’, Mo *ʔa-ha* ‘water’, Ma *xa* ‘water’, Qu *ʔa-xá* ‘water’, Pa *ʔ-xa* ‘water’, Hu *haʔ*, *ʔ-haʔ* ‘water’, Ha *ha-ʔa*, *ʔ-ha-ʔa* ‘water’, Ya *ʔ-há* ‘water’

SER *ʔa-χ* ‘fresh water; superficial well’ (pl. *ʔá-χa-xam*), *i-χ* ‘liquid, sap, juice’

PCHO *-*xa-ʔ* ‘water’ (HCh *l-a-haʔ*, LCh *a-haʔ*, Teq *xa*)

NOTES. This word is connected by ablaut with PH *i/a-k^hí· ‘to drink’ (k^h-13).

Dixon & Kroeber 1913: 651 || Kroeber 1915: 282 || Sapir 1917: 8 || Sapir 1920b: 285 || Sapir 1921: 68 || Greenberg & Swadesh 1953: 220 || Haas 1954 || Bright 1956: 45-46 || Jacobsen 1958: 203 || Sapir & Swadesh 1960: 16 || Haas 1964: 81 || McLendon 1964: 137 || Silver 1964: 176 || Gursky 1965: 188, 193 || Gursky 1974: 210 (#ax^ya#) || J. G. Crawford 1976: 320 || J. M. Crawford 1976: 187 || Silver 1976: 202 || Waterhouse 1976: 337 || Langdon 1979: 639 || Webb 1980: 76 || Greenberg 1987: 213 || Kaufman 1989: 131 (#a:-x^yáʔ ‘water’ [N/S]) || Leščiner 1990: 36 (*ʔǎč^hǎ) || Leshchiner & Nikolaev 1992: 398 (*ʔǎč^hǎ) || Shaul 2020b: 309–310

(k^h-02) PH *a/uk^há ‘belly’

YAN [*y*]o·*xai-* ‘pregnant’

PPOM *ʔuhq^há, *ʔuhq^há· ‘belly’⁴ > Pk *ʔuhq^ha*, Ps *ʔuk^ha*, Pn *k^ha[šé]*, Pc [*m*]k^há, Pse *χá* ‘stomach, entrails, paunch, tripe’, Pe *yu·xá* ‘paunch, tripe, stomach, guts’

SER [*y*]aχ ‘belly, abdomen’ (pl.: [*y*]áχ-kox, abs.: ʔa-[*y*]áχ)

⁴ Oswalt (1972: 27) gives Ps *ʔuhk^ha*, Pc *wq^há*, *mq^há*. If these forms are more correct than those given by McLendon, the variant *ʔuhq^há· with a long vowel can be eliminated.

Gursky 1974: 179 || Greenberg 1987: 133-134 || Gursky 1988: 22 || Kaufman 1989: 133 (#[i/u]q^há 'belly' [N]) || Gursky 1990: 23

(k^h-03) PH *k^hak^wá 'flint; arrowhead' (= k^w-03)

CHI q^hako <khako> 'arrowhead' (2.22.0211), <khákkò> 'bullet, arrowhead' (2.23.0991)

KAR sá·k 'obsidian, flint; arrowhead, bullet'

ACH sāt 'arrowhead; knife; flint; metal; bullet; iron; obsidian'

YAN CN xaka, Y xahka 'flint; Flint (as a myth character); flint, a design element in basketry', N xaka-wi 'arrowheads'

PPOM *q^hahká 'flint' > Pk qahca, Ps kahča, Pn k^hačá, Pc k^hcá, Pne k^háhta, Pse qcá, Pe xa·ká

SAL A [ʔa]šák <ʔák> 'arrow point of flint' (2.85.0096), A [ʔa]šák <ʔacák>⁵ 'what they put on point of arrows', M [pa]šák <p'acák> 'id.' (2.86.0584)

PYUM *(ʔ-)xk^wa 'knife' > Mo ʔ-ahk^we 'knife', Hu a-k^wá, ʔ-k^wa, k^wa 'knife', Ha k^wa-ʔa 'silver; knife; metal', Ya ʔ-k^wa 'metal, iron; knife'

SER ʔax 'arrow point' (pl. ʔax-lk), 19th century Seri (Pinart) <vaX> 'pedernal; flint' (Hernández 1902: 253)

NOTES. (1) The word may be a compound with PH *(a/i-)k^haʔa 'stone' (k^h-10) as the first part. (2) Seri reflects a metathesized form *k^wak^há.

Sapir 1917: 8 || Sapir 1920b: 284 || Sapir 1921: 69 || Greenberg & Swadesh 1953: 219 || Jacobsen 1958: 200 || Haas 1964: 78 || McLendon 1964: 131 || Silver 1964: 173 || Gursky 1965: 169 || Gursky 1974: 188 (#q^hak^wa#; #sak^(ʔ)#) || Langdon 1979: 637 || Greenberg 1987: 259 || Gursky 1989: 9 || Kaufman 1989: 138 (#x^yáHk^wa 'flint; arrowpoint; knife' [N/S]) || Leshchiner & Nikolaev 1992: 399 (*č^hákA) || Shaul 2020b: 302

(k^h-04) PH *k^hapá 'cloud' (= p-03)

PPOM *q^haʔbá(·) 'cloud' > Pk q^haba·, Ps k^habʔa, Pn k^habá·, Pc k^hbá·, Pne k^háʔba· 'fog', Pse xba 'fog', Pe xa·bá 'fog'

PYUM *xpa 'cloud' > Ki p-xpaʔ¹ 'small puffs of smoke, cloud'

NOTES. This word may be a compound with PH *a-k^ha 'water' (k^h-01) as a first part. The second part may be compared with Sal A pay <páj> 'nube' (2.85.0024).

(k^h-05) PH *k^hapí 'metate, grinding stone' (= p-08)

CHI q^hapi[na] <khabina> 'mortar' (2.20.1080), q^hapi[na] <khap'ina> 'mortar' (2.22.0249)

PPOM *q^haʔbé 'rock' > Pk q^haʔbe, Ps k^haʔbe, Pn k^habé, Pc k^hbé, Pne k^háʔbe, Pse xbe, Pe xa·bé

SAL A t-šxaʔ <tʃqá> 'a rock' (2.85.0081), A t-šxáp-ahnel <tʃqápahnel> 'stones' (2.85.0082), M t-šxáp-elat <tʃqáp'elat> 'stones' (2.85.0083), A t-í-šxaʔ <tríʃqá> 'mi dinero, mi piedra', t-mí-šxaʔ <trímíʃqá> 'tu dinero ó piedra', t-í-šxáʔ-wo <tríʃqá'wə> 'his dinero' (2.85.0083)

PYUM *(ʔ-)xpi 'metate' > Ki xpiʔ³ 'metate, millstone', Co xpi 'metate', JT xpi 'metate', MGI ʔə-xpi 'metate', Mo ʔa-hpe 'metate; name of a cat's cradle design', Pa ʔ-xpi 'metate', Hu p^hi, p^hiʔ, ʔ-p^hiʔ 'metate, grinding stone', Ha p^hi-ʔi 'flat grinding rock', Ya ʔ-hpi 'metate, grinding stone'

SER xap 'grinding stone' (pl. xáp-kox)

NOTES. This word may be a compound with the first part being a shortened form of PH *(a/i-)k^haʔa 'stone' (k^h-10).

Sapir 1925: 411 (*ixa'-pi) || Greenberg & Swadesh 1953: 219 || Gursky 1974: 207 (#ax^yaʔpi#) || Campbell 1976: 162 || J. G. Crawford 1976: 315, 319 || J. M. Crawford 1976: 183 || Greenberg 1987: 141 || Kaufman 1989: 138 (#x^yáʔ=PI(y) [x^y ~ x] compound 'mortar, metate; stone' (<"stone pound" or "stone stone")) || Shaul 2020b: 306

(k^h-06) PH *(a/i-)k^haʔá(y) 'morning' (= ʔ-01)

SHA ʔač·áy 'day' (Silver 1966: 208)

⁵ Harrington notes: "á not very long but falling".

YAN *xa[ʔlai]*- ‘to be daylight, dawn’

PPOM *q^haʔá ‘dawn, morning’ > Ps *k^haʔ-a-*, Pn *k^haʔá*, Pne *k^haʔá-[wi]*, Pse *ʔá-*, Pe *xa-ʔá-*

SAL A [*k*]ešxáy ‘k’eqáj’ ‘la mañana’ (2.85.0014), A [*k*]ešxáy ‘k’ecxáy’ ‘ya ameneció’, M [*p*]ešxay[*léʔ*] ‘p’ecxaylé’ ‘id.’ (2.86.0589)

NOTES. Shasta *č* instead of expected *c* is due to the change *k^haʔáy > *k^hʔáy > *káy.

Sapir 1917: 8 || Sapir 1925: 409 || McLendon 1964: 134 || Gursky 1974: 197 || Gursky 1989: 22 || Kaufman 1989: 143 (#a)x^ya7 ‘morning, daylight’ [N])

(k^h-07) PH *a/i-k^hiy ‘dew’

PPOM *ʔiq^he ‘dew’ > Pk *ʔik^he*⁶, Ps *k^he-ʔe*, Pc *k^hé*, Pn *k^hé-* (Oswalt 1972: 32)

PCHO *-iʔ]axe ‘dew’ > HCh *t-iʔahe* ‘dew’

(k^h-08) PH *puk^há/íla ‘a kind of basket’ (= p-12)

CHI *puq^hela* ‘púk’éla’ ‘small basket’ (2.20.0038)

PPOM *buhq^hál ‘open-woven pack basket’ > Pk *buhq^hal*, Pn *bok^hál*, Pc *p^hk^hál*, Pse [*fáli*] *bɣàl*, Pe *bu-xál* ‘fish trap basket (open-woven and elongated)’

NOTES. (1) The word may be a compound with the first component *(a)pú(-p) ‘to weave baskets’ (p-11). (2) This word has a limited distribution and belongs to cultural vocabulary, so borrowing is not excluded. I tentatively retain this comparison because Chimariko and Pomoan are not immediate neighbors.

Gursky 1974: 179 || Gursky 1988: 21 || Kaufman 1989: 158 (#puq^ha(l^y) [l^y ~ n^y] ‘basket’ [N])

(k^h-09) PH *k^ha/učV ‘cold’ (= č-01)

CHI *q^hača* ‘q’át’sa’ ‘ice’ (Berman 2001: 1063)

YAN CN *xa-č[iʔ]*- ‘to feel cold’

PPOM *q^hu(ʔ)čá- ‘winter’ > Pk *q^hoša-*, Pn *k^hečá[ma-]*, Pne *k^huʔčá[ti]*, Pe *xu-čá[y]*

PYUM *xčur[r] ‘to be cold’ > Co *xsu-r* ‘be cold’, JT *xəču-r* ‘it is cold’, MGI *xəču-r* ‘is cold’, Mo *haču-r-k* ‘be cold’, Ma *xčur-k* ‘be cold’, Qu *xacú-r* ‘to be cold’, Hu *č^hur* ~ *č^hu-r* ‘winter’, Ha *č^hu-r* ‘year; winter’, Ya *ʔ-č^hú-ri* ‘winter’

NOTES. Yana *xa-č[iʔ]*- ‘to feel cold’ is derived from *xa-* Y ‘cold’. This morpheme is also found in Yan Y *xa-ai-na* ‘cold’. If the Yana form really belongs to this set, the Proto-Hokan form must also be considered a compound with the first element PH *k^ha- ‘cold’.

Sapir 1917: 8 || Greenberg & Swadesh 1953: 218 || Gursky 1965: 172 || Gursky 1974: 183 || Greenberg 1987: 207 || Gursky 1988: 32 || Kaufman 1989: 143 (#x^yač [a ~ o; x^y ~ ɣ] ‘cold [*a]; winter [*u]’ [N/S]) || Leshchiner & Nikolaev 1992: 401 (*q^háččó)

This correspondence can be complemented by the following comparisons without a reflex in Pomoan.

(k^h-10) PH *(a/i-)k^haʔa ‘stone’ (= ʔ-07)

CHI *q^haʔa* ‘khá’á’ ‘rock’ (2.23.0606)

KAR *as* / *asa-* ‘rock; stone; pit of fruit; in some compounds, earth, dirt’

SHA *ʔíc-aʔ* ‘rock’ (Silver 1964: 175)

YAN *kā[i]-na* ‘stone’

NOTES. Yana *k* instead of expected *x* is due to the change *k^haʔa > *k^hʔa > *ka.

Dixon & Kroeber 1913: 651 || Sapir 1917: 9 || Sapir 1925: 411 (*ixa) || Haas 1964: 81 || Silver 1964: 175 || Gursky 1974: 207 || Kaufman 1989: 138 ([i/a]-x^ya7 ‘stone’ [N/S])

⁶ Oswalt & Buckley (2018) give Pk *ʔik^he-y*.

(k^h-11) PH *k^hán ‘good’

PYUM *xan ‘good’ > Co *xan*^y ‘be new, fine, genuine, real, good, first-class’, JT *xan* ‘he or it is good, fine, well, comfortable, all right, okay, nice’, MGI *ʔa-xan* ‘is good, correct, right, tasty, true, sure’, Mo *ʔa-xan* ‘right (direction); good, proper, original’ (old word used in fixed expressions), Ma *m-xan-k* ‘like, want’, Qu *ʔa-xán* ‘to be genuine’, *m-axán* ‘to admire’, Pa *xan* ‘be good, fine’, *xan...i* ‘one likes, cares for’, Hu *han-k* ‘to be good, to be well’, *hán-o-k* ‘to like something’, Ha *han-k* ‘good’, *han-o-k* ‘to like’, Ya (?)*hān-i* ‘good, handsome, right, first, real, perfect’

PCHO *-xan[a]- > HCh *di-hanah-ʔma* ‘he will like it, love it’, LCh *hanah-ma* ‘like, be pleased by’, *hanáʔ* ‘all right’ (Waterhouse 1976: 337), Teq <xana> ‘to taste good’

Waterhouse 1976: 337 || Gursky 1989: 12 || Kaufman 1989: 161 (#Xan ‘good; pleasant’ [S])

(k^h-12) PH *(i-)k^há/íp ‘to extinguish’ (= p-23)

CHI *n-iq^hap* <ʔu nikháp> ‘put the fire out!’ (2.23.0840), *n-eq^hap* <nekháp> ‘put it out!’ (2.22.0113) (*n-* ‘imperative’ [Jany 2009: 118])

KAR *ışip* ‘to extinguish a fire’, *im-şıp* ‘to cool off; (fire) to be extinguished’ (*im-* ‘by heat or fire’ [Haas 1980: 69])

ATS -*cp* in *p-cí-cp* ‘to extinguish a fire’, *ma-cp* ‘to put out a fire with one’s foot’, *phú-cp* ‘to blow out a fire’ (Talmy n.d., a: 41)

SAL A <cap>, M <cap> ‘to extinguish, put out’ (Mason 1918: 139), M *šap* <cap> ‘ya está apagado’ (2.88.0068)

PYUM *sip ‘to be extinguished’ > JT *wəsip* ‘it goes out (said of a fire)’, MGI *əsip* ‘goes out, is extinguished’

NOTES. (1) Chimariko and Salinan reflect PH *k^háp, while Karuk and Yuman reflect PH *k^híp. These two protoforms are connected by the Proto-Hokan ablaut *a/i. (2) Atsugewi *c* instead of expected *c^h* is due to synchronic neutralization of plain and aspirated consonants in coda position (Good 2004: 8).

J. G. Crawford 1976: 312 || Kaufman 1989: 160 (#šip [i ~ a] [ʔasp] ‘(to die [fire];) to douse (fire)’ [N/S]) || Shaul 2020b: 301

(k^h-13) PH *i/a-k^hí ‘to drink’

KAR *ış* ‘to drink’

SHA -*ic-* ‘to drink’ (Silver 1964: 172)

ACH *is* ‘to drink’

ATS -*c^{hi}* in *i-c^{hi}* ‘to drink’ (Talmy n.d., a: 61)

YAN N *si-*, Y *c^{hi}i-* ‘to drink’

SAL A <icim> ‘to drink’ (Mason 1918: 141); A *eše[m]* ‘to drink’ in <mé~šem> ‘drink!’, <kéšem hék> ‘I’m drinking’ (Jacobsen 1954: 106)

PYUM *si ‘to drink’ > Ki *čhi*² ‘drink’, Co *ši* ‘drink’, JT *wə-si* ‘he drinks something’, MGI *wə-si* ‘drinks’, Mo *i-θi-m* ‘drink; eat soup’, Ma *si-m* ‘drink’, Qu *a-sí/a-sé* ‘to drink’, Pa *si* ‘one drinks’, Hu *θi-k* ‘to drink’, Ha *θi-k* ‘to drink’, Ya *θi* ‘drink’

SER *k-i-si* ‘to drink; to bite and suck (e.g. mosquito)’ (conjug.: *i-t-ási*, *i-t-ásy-at*; *i-t-ásy-ox*, *i-t-ásy-ołka*)

NOTES. (1) This word is connected by ablaut with PH *a-k^há ‘water’ (k^h-01). (2) Superscript -*i* in Atsugewi is a morphophonological symbol with unclear value used by Talmy.

Kroeber 1915: 282 || Sapir 1917: 14 || Greenberg & Swadesh 1953: 218 || Haas 1964: 77 || McLendon 1964: 131 || Silver 1964: 172 || Gursky 1968: 39 || Gursky 1974: 185 || J. G. Crawford 1976: 311 || Silver 1976: 202 || Webb 1980: 76 || Greenberg 1987: 135 || Gursky 1989: 3 || Kaufman 1989: 131 (#i/a-sí [s ~ š] ‘to drink’ [N/S]) || Shaul 2020b: 300

(k^h-14) PH *(i/a/u-)k^hí ‘to sing’

SHA -e·c[ni]- ‘to sing’ (Silver 1966: 74)

ACH as ‘to sing’

ATS -c^{hi} in ye·c^{hi} ‘to sing’ (Talmy n.d., a: 62)PPOM *HV(h)q^hé ‘dance, song’ > Pn k^hé, Pc k^he, Pne k^hé, Pse xé, Pe xe

SER k-ooš ‘to sing’ (conjug.: t-ooš, t-ooš-tim; t-óosi)

NOTES. Superscript -i in Atsugewi is a morphophonological symbol with unclear value used by Talmy.

Sapir 1917: 7 || Gursky 1974: 205 || Greenberg 1987: 255 || Kaufman 1989: 159 (#(e:)x^ye ‘to sing; song; dance’ [N])

|| Leshchiner & Nikolaev 1992: 385 (*ʔōčō)

(k^h-15) PH *k^hu- ‘2nd person plural’CHI q^ho-, q^hu- ‘2nd person plural’ (verbal prefix) (Jany 2009: 34)SAL ko- ‘2nd person plural’ (verbal prefix)⁷ (Turner 1987: 129).

NOTES. This set forms part of the paradigmatic comparison of Chimariko and Salinan personal prefixes by Sapir (1920a).

Sapir 1920a || Gursky 1965: 182 || Kaufman 1989: 116 (#q^hV ‘second person plural pronoun marker’ (same as #KV?) [N]) || Zhivlov 2018: 144(k^h-16) PH *a-ták^hu ‘many’ (= t-06)

SER k-átxo ‘many, much, abundant’ (conjug.: t-átxo; t-átxo-x [t-átxo-xam])

PCHO *ataxu > LCh atahu ‘many, a lot of’

Gursky 1974: 197 || Waterhouse 1976: 339 || Greenberg 1987: 237 || Kaufman 1989: 147 (#aTaXu ‘many’ [S])

(k^h-17) PH *(a/i-)k^húl ‘dry’

KAR axr[ah] ‘dead tree’, iv-áxr[ah] ‘to be dry’ (im- ‘denominative’)

SHA k-ic-úr ‘dry’ (Silver 1961: 12) (k- ‘deverbative nominalizing prefix’ [Silver 1966: 210])

ACH [w]asú·ll[i] ‘it dries’

YAN YN k^hul[si]- ‘to feel dry, to dry’ (erroneously placed by Sapir & Swadesh [1960: 104] under the root k^hul- ‘to be, feel’)

SER k-íxox ‘to dry (columnar cactus fruit)’ (conjug.: i-t-íxox; i-t-íxox)

PCHO *-xuł ‘to dry’ > HCh di-húł-ʔma ‘it will dry out’, i-húł ‘dry’, LCh huł-ma ‘dry, dry up’, Teq <hur> ‘to dry’

Waterhouse 1976: 341 || Gursky 1989: 4 || Kaufman 1989: 142 (#Xulʔ ‘dry’ [S]) || Gursky 1990: 24

1.2.3. Proto-Pomoan *q̣

The sound correspondence involving Proto-Pomoan *q̣ is represented in Table 9. This correspondence reflects the Proto-Hokan glottalized velar stop *ḳ.

PH	Chi	Kar	Sha	Ach	Ats	Yan	PPom	Sal	PYum	Ser	PCho
*ḳ	q̣	x	č̣	č̣		ḳ	*q̣	ḳ	*x	x̣	*ḳ

Table 9. Proto-Hokan *ḳ.

(ḳ-01) PH *ká/ís ‘to scrape, to scratch’ (= s-05)

KAR ta-xíš-xiš ‘to scrape’ (ta- ‘with tool or implement’), [θiv]xiš ‘to plane (a board)’

⁷ Turner gives this as k-, but remarks: “Whatever the first vowel of the verb stem, it is replaced by a in the first person plural and by o in the second person plural” (Turner 1987: 129).

PPOM *-q̄a·s > Pk -q̄a·s (verb root) ‘scratch and make a streak’, e.g. in *da-q̄a·s* ‘scratch and make a streak, with the hand or paw’ (*da-* ‘with the hand, especially the palm, with the paw of animals’) (Oswalt & Buckley 2018), Pe *da-q̄a·s* ‘scratch with nails; harrow’ (*da-* ‘with or affecting the hand’) (McLendon 1975: 66)

PYUM *xas ‘to scrape’ > Co *č-xaš* ‘scrape, scrub, scour, whittle, shave, sharpen (a pencil)’, Mo *a-haθ-* ‘peel (something) with a peeler or knife’, *a-ha-θ-* ‘trim down a stick, whittle’, Ha *haθ-k* ‘to scrape (using a knife)’, Ya *háθ-v-i* ‘scrape, shave’

SER *k-[ta]xš* ‘scratch; take away forcibly’ (conjug.: *i-t-[tá]xš*, *i-t-[tá]x-ka*; *i-t-[tá]x-kox*), *k-[šá]xaš* ‘scratch’ (*i-t-[šá]xaš*; *i-t-[šá]xaš-ox*), *k-[pa]xš* ‘scratchy, rough; broke, without money’ (conjug.: *t-[pa]xš*; *t-[pa]x-ka*)

NOTES. Seri *ta-*, *ša-* and *pa-* look like fossilized instrumental prefixes/prepounds; *ta-* may be the same prepound that is reflected in Karuk *ta-*, Pomoan **da-* and Cocopa *č-* (pYum **t-*)⁸.

J. G. Crawford 1976: 317 || Kaufman 1989: 139 (#(T)aXas [s ~ š; a₂ ~ i] ‘to scrape; scratch’ [SW/?+Kar] (?causative prepound pHok *ta=))

(k-02) PH *(a-)k̄aw ‘fox’

KAR *apra·x* ‘gray fox (*Urocyon cinereoargenteus*)’

PPOM *haq̄aw ‘fox’ > Pk *haq̄aw*, Pn *k̄aw*, Pc *k̄aw*, Pne *k̄aw-*, Pe [*qa-*]q̄aw

PYUM *prxa·w ‘fox’ > Ki *pnxa·w*² ‘fox (probably kit fox; *Vulpes macrotis*)’, Co *prxa·w* ‘kit fox (*Vulpes macrotis*)’, JT *p̄rxa·w* ‘fox (gray fox, *Urocyon cinereoargenteus*)’, MGI *parxa·w* ‘fox’, Mo *marho* ‘grey fox’

NOTES. Cf. PH *a-pVI-k̄aw ‘fox’ (p-21).

Langdon 1979: 637 || Kaufman 1989: 148 (#(a)pVrax ‘fox’ [Kar:NC +Yum]) || Leshchiner & Nikolaev 1992: 394 (*h̄eq’ówá)

(k-03) PH *(a-)ká/i ‘to finish’

CHI *q̄e[h]-ta* <κ’ēhta> ‘she is dead’ (2.20.1114), *q̄e[?]-men* <κ’ē’men> ‘I pretty near died’ (2.22.0145)

YAN N [*ti*]ka[u]- ‘used up; to finish’ (cf. *tinti-* ‘to stop’), Y [*te-tu*]ka[u]..[?a] ‘to finish’, -[*tu*(·)]ka[m]- ‘to finish’

PPOM *-(?)q̄á[l] > Pe *q̄ál* ‘to terminate, finish an activity’ (McLendon 1975: 53)

SER *k-i-xi* ‘finish; finish making’ (conjug.: *i-t-áxi*, *i-t-áxi-tim*; *i-t-áx-t* [*i-t-áx*]), *k-óoxi* ‘die (especially animal); fade’ (*t-óoxi*; *t-óoxya-t*, *t-óoxya-tołka*)

McLendon 1964: 131 || Kaufman 1989: 163 (#q’a ‘to finish’ [N])

(k-04) PH *(a-)k̄úy ‘neck’

PPOM *q̄óy[u] ‘neck’ > Pk *kuyu* ‘windpipe’, Pse *q̄oy*, Pe *q̄óy*

SAL A [*l*]á·koy <lák’oj> ‘mi nuca’, applied to back of neck only (2.85.0478)

Gursky 1974: 199 || J. G. Crawford 1976: 319 || Gursky 1989: 24 || Kaufman 1989: 129 (#q’o(y) ‘neck; throat; to swallow’ [N/S]) || Shaul 2020b: 305

This correspondence can be complemented by the following comparisons without a reflex in Pomoan.

(k-05) PH *kal ‘to scrape, to scratch’

KAR *ak-xár[ap]* ‘to scratch’ (?*ak-* ‘with the hands’ [Haas 1980: 70]), [*vim*]xár[ap]-sur / [*vim*]xár[ap]-suru- ‘to scratch off’ (-sur/-suru- ‘off’)

⁸ According to S. A. Marlett (pers. comm.), Mary Moser noticed that a lot of verbs that have *t-* at the beginning have something to do with the hand.

YAN N *kal-čau-* ‘to scratch with sharp claws’ (*-čau-* ‘to get stuck in cleft, claws, teeth’)

PCHO **-kał-* ‘to scrape’ > HCh *di-kát-?ma* ‘he will whittle or trim it; he will chip it away’, LCh
kał-u-ma ‘scrape, of the skin’

Haas 1964: 81 || Gursky 1968: 39 || Gursky 1990: 2

(k-06) PH *(i-)kay ‘bitter, sour’

CHI *qoy[oʔ]-it* <‘ojo’it> ‘it is sour’ (2.20.0045)

SHA *ʔičáy* ‘bitter’ (Silver 1961: 11)

ACH *i-ča-yiʔ* ‘strong tasting; bitter; sour’

YAN *kai-* N ‘bitter, strong (liquor)’ Y ‘bad-tasting’

Sapir 1917: 9 || Sapir 1925: 418 (*ikay(a)?) || Haas 1964: 77 || McLendon 1964: 129 || Silver 1964: 171 || Gursky 1965: 189 || Gursky 1974: 180 || Greenberg 1987: 258 || Kaufman 1989: 134 (#ikʷa(y) [a ~ o] ‘bitter, sour’ [N/S] OR: #qʷay [a ~ o] ‘bitter, sour’ [NC])

(k-07) PH *(a-)kít ‘unripe, young’ (= t-15)

KAR *xí-t* ‘green, unripe’

SER *k-axt* ‘tender, delicate, young [plant, vegetable]; young [person]’ (conjug.: *t-axt*; *t-axt*)

NOTES. Notions ‘unripe’ and ‘young’ are frequently colexified in world’s languages, see
<https://clics.clld.org/edges/1207-1958>.

Gursky 1990: 19

(k-08) PH *kūn ‘to want’

YAN YC *kun-*, N *kut-* ‘to want, like’

SAL A <-cu*nip> ‘desear, apetecer, &c.’: <comcu*nip> ‘no desear’, <cozamcu*nip> ‘no desearas, no apeteceras’ (Sitjar 1861: 21)

NOTES. (1) The reconstruction *kʷun is also possible. (2) Mason writes the Salinan verb taken from Sitjar as <k’unip> ‘desire’ (Mason 1918: 143). According to him, in Sitjar’s vocabulary “[t]he fortis stop was generally expressed by an asterisk after the vowel” (Mason 1918: 121).

Sapir 1921: 71 || Haas 1964: 81 || Gursky 1965: 187 || Gursky 1974: 210 || Greenberg 1987: 142 || Kaufman 1989: 167 (#KʷuN ‘to want, desire’ [N])

1.2.4. Comparison with Kaufman’s reconstruction

Kaufman reconstructs the following nonlabialized velar and uvular consonants: *kʷ, *kʷʷ, *kʷʰ, *xʷ, *q, *qʷ, *qʰ and *x. In addition, he reconstructs the cluster *qx. See Table 10 for Kaufman’s correspondences.

PH	Chi	Kar	Sha	Ach	Ats	Yan	Wsh	PPom	Sal	PYum	Ser	PCho	Tol
*kʷ	k	k	k	k	k	k	k	*kʷ	k	*kʷ, *k	k	*k	k
*kʷʰ	kh	k	k	kʰ	?	kʰ	kʰ	*kʷʰ	k	*kʷ, *k	k	*k	kʰ
*kʷʷ	kʷ	k	kʷ	kʷ	kʷ	kʷ	kʷ	*kʷʷ	kʷ	*kʷ, *k	k	*kʷ	kʷ
*xʷ	qʰ	s	ʕ	s	c	x	š	*qʰ	š	*x	x	*x	h
*q	q	k	k	q	[q]	k	k	*q	k	*k, *q	k	*k	k
*qʰ	qh	k	k	qʰ	?	kʰ	kʰ	*qʰ	k	*k, *q	k	*k	kʰ
*qʷ	qʷ	k	kʷ	qʷ	[qʷ]	kʷ	kʷ	*qʷ	kʷ	*k, *q	k	*kʷ	kʷ
*qx	qh	x, k	k, x	k, x	[q], x	kʰ	kʰ	*qʰ	k, x	*q(x), *k	k	?	kʰ
*x	x	x	x	x		x	x	*h	x	*x	x	*x	h

Table 10. Correspondences of Hokan nonlabialized velar and uvular obstruents according to Kaufman (1989).

The main difference with my reconstruction is that Kaufman postulates a triple opposition between front velar, labialized velar and uvular stops. However, the only languages where the reflexes of Kaufman's front velars and labialized velars are consistently different⁹ are Yuman, Chontal and Comecrudo¹⁰, where the labialized velars preserve their labialization. This means that any word reconstructed with a front velar without reflexes in Yuman or Chontal can be reconstructed with a labialized velar instead. Note that while Kaufman uses the cover symbols *K for *q or *k^y and *K! for *q, *k^y, or *k^w (Kaufman 1989: 109-110) he has no dedicated cover symbol for “*k^y or *k^w”, although this must be a frequent situation given the overlap between correspondences for *k^y and *k^w. Unambiguous cases of front velars in Kaufman's reconstructions are quite rare, and none of them is reliable enough to support the reconstruction of separate phonemes. Therefore, I propose to reconstruct labialized velars in place of both Kaufman's front velars and labialized velars.

As for Kaufman's uvulars, these supposedly give uvulars in Chimariko, Achumawi, Atsugewi, Pomoan and (in some cases) Yuman. The Yuman evidence is ambiguous, since Proto-Yuman *q was used as an augmentative grade of *k in consonant gradation. Kaufman himself admits that “[i]n fact /k/ may be the normal Yuman reflex of pHok *q” (Kaufman 1989: 101). While there is evidence that Chimariko and Pomoan uvulars correspond to each other, in Achumawi and Atsugewi they correspond to sibilant affricates and fricatives. For this type of correspondences, I reconstruct velar stops. Thus, the origin of uvular stops in Achumawi and Atsugewi remains so far unknown.

Kaufman also reconstructs three velar/postvelar fricatives: front velar *x^y, labialized velar *x^w and uvular *χ. Kaufman's *x^y has such reflexes as Chimariko /q^h/, Shasta /c/, Atsugewi /c^h/ and Proto-Pomoan *q^h. This makes it rather unlikely that we deal here with an original fricative, because it would require several independent typologically rare shifts of fricative to stop or affricate. Since I do not find sufficient evidence for Kaufman's *q^h, which would be *k^h in my system, I reconstruct Proto-Hokan *k^h in place of Kaufman's *x^y. Instead, I reconstruct as *x the phoneme reconstructed by Kaufman as *š (see above).

Reflexes of Kaufman's labialized velar *x^w and uvular *χ differ only in Yuman, Chontal and Comecrudo¹¹. This means that any word reconstructed with *χ without reflexes in Yuman or Chontal can be reconstructed with *x^w instead. Note that while Kaufman uses the cover symbols *X for *χ or *x^y and *X! for *χ, *x^y, or *x^w (Kaufman 1989: 109–110) he has no dedicated cover symbol for “*χ or *x^w”, although this must be a frequent situation given the overlap between correspondences for *χ and *x^w. In such cases, Kaufman usually reconstructs *χ. Unambiguous cases of *χ in Kaufman's reconstructions are quite rare and none of them is reliable enough to support the reconstruction of a separate phoneme. Therefore, I propose to reconstruct *x^w in place of both Kaufman's *χ and *x^w.

Finally, Kaufman's correspondence for *qx features too many unexplained reflexes to be considered as regular: it can give /k/ or /x/ without complementary distribution in Karuk, Shasta, Achumawi and Salinan.

⁹ According to Kaufman, labialized velars can yield uvulars in Chimariko, Achumawi and Proto-Pomoan, although he does not specify the precise conditions of such a development. I was unable to find reliable cases where labialized velars of Proto-Yuman or Proto-Chontal correspond to uvulars in Chimariko or Proto-Pomoan.

¹⁰ Comecrudo data are scarce; they are best left out of consideration at the present stage.

¹¹ Kaufman also postulates partly different reflexes in Chimariko: while *χ yields Chimariko /x/, *x^w can yield both /h/ and /x/ under unknown conditions.

1.3.0. Correspondence involving Proto-Pomoan glottal stop

Glottal stop *ʔ is reconstructed for Proto-Pomoan in a large number of words (McLendon 1973: 25). However, most of these cases belong to one of the two following categories: 1) word-initial glottal stop, and 2) glottal stop as a “laryngeal increment” before consonants. There are no vowel-initial words in Proto-Pomoan, and word-initial *ʔ and *h are in near complementary distribution: *ʔ if the next consonant is voiceless and unglottalized, but *h if the next consonant is a glottalized stop, voiced stop or ʔ. Before a voiced sonorant, both *ʔ and *h are possible (Oswalt 1976a: 3). Because word-initial glottal stop is almost always predictable, such cases will not be considered below¹². Laryngeal increments, i.e. laryngeal consonants immediately preceding the initial consonant of the tonic syllable (McLendon 1973: 50-51), will be discussed in one of the next papers in this series. What remains is the glottal stop in Proto-Pomoan stems of the shape *CVʔV.

1.3.1. Proto-Pomoan *ʔ

The sound correspondence involving Proto-Pomoan *ʔ in *CVʔV stems is represented in Table 11. This correspondence reflects the Proto-Hokan glottal stop *ʔ. See below on the reflex *h in Pomoan.

PH	Chi	Kar	Sha	Ach	Ats	Yan	PPom	Sal	PYum	Ser	PCho
*ʔ	ʔ	Ø	ʔ	Ø		ʔ	*ʔ, *h	ʔ ~ Ø	*ʔ	ʔ	*ʔ

Table 11. Proto-Hokan *ʔ.

(ʔ-01) PH *(a/i-)k^haʔá(y) ‘morning’ (= k^h-06)

SHA ʔač·áy ‘day’ (Silver 1966: 208)

YAN xa[ʔlai]- ‘to be daylight, dawn’

PPOM *q^haʔá ‘dawn, morning’ > Ps k^haʔ-a-, Pn k^haʔá, Pne k^haʔá·[wi], Pse ʔá-, Pe xa·ʔá-

SAL A [k]ešxáy <k’eqáj> ‘la mañana’ (2.85.0014), A [k]ešxáy <k’ecxáy> ‘ya ameneció’, M [p]ešxay[léʔ] <p’ecxaylé> ‘id.’ (2.86.0589)

NOTES. Shasta č instead of expected c is due to the change *k^haʔáy > *k^hʔáy > *káy.

Sapir 1917: 8 || Sapir 1925: 409 || McLendon 1964: 134 || Gursky 1974: 197 || Gursky 1989: 22 || Kaufman 1989: 143 (#(a)xʔa7 ‘morning, daylight’ [N])

(ʔ-02) PH *ʔáy ‘father’

PPOM *(me-)ʔé ‘father’ > Pk meʔe ‘your father’, Ps meʔe ‘your father’ (O), Pn -ʔé ‘father’, Pne -meʔé-, Pse mʔe ‘father’, Pe meʔé ‘your father’

PYUM *n-ʔa(·)y ‘father (of woman)’ > Co n^y-ʔa· ‘father (f.s.)’, Mo na-ʔay ‘father (of a woman)’, Ma n-ʔay ‘father (female’s)’ (*n- ‘kinship terms prefix’ [Miller 2001: 86])

PCHO *-ʔay[i] ‘father’ > HCh ʔ-a-ʔáyíʔ ‘father; priest’, Teq ‘ai’, ‘yayi’ ‘father’

NOTES. According to McLendon (1973: 74), “Pk, Ps, Pn, and Pe have all apparently reshaped the proto morpheme by segmenting the initial syllable, identifying it with the second person singular prefix *mi-, as in OLDER BROTHER”. I prefer a simpler scenario, in which

¹² Seri and Proto-Yuman contrast glottal stop and absence of a consonant word-initially. In particular, the absolutive prefix can be reconstructed as *ʔa- for the last common ancestor of Seri and Yuman (Zhivlov 2018: 141). At the present stage of reconstruction it is hard to say whether this contrast goes back to Proto-Hokan.

*me- is the second person prefix (going back to Proto-Hokan *m-, see Zhivlov 2018: 144), while the Northern Pomo form reflects the unprefixated stem.

Gursky 1974: 187 || Waterhouse 1976: 337 || Langdon 1979: 637 || Gursky 1989: 7-8 || Kaufman 1989: 145 (#7ay [ʔi ~ a] 'father' [N/S])

(ʔ-03) PH *i-ʔíy 'hair of head'

PPOM *heʔʔé, *heʔey(?) 'hair of head' > Pk *heʔe*, Ps *heʔ·e*, Pn *ʔé·*, Pc *ʔé·*, Pne *héle (?)*, Pse *ʔéy*

PYUM *ʔiy 'hair of head' > Ki *ʔiy¹* 'head', Mo *i-ʔe*, *n^y-i-ʔe* 'hair of the head', Ma *i-ʔe* 'hair', Qu *i-ʔé* 'hair'

Gursky 1974: 192 || Langdon 1979: 637 || Gursky 1989: 14 || Kaufman 1989: 128 (#i-ʔíy 'head; hair; fur' [N/S])

(ʔ-04) PH *(a-k^{wh}-)Vʔíy 'salt'

CHI *ʔakí* <'akí> 'salt' (2.20.0040)

PPOM *k^heʔé: 'salt' > Pn *šeʔé·*, Pne *č^héʔe·*, Pe *k^he·ʔé*

SAL A *ʔ-a·káy* <trākáy> 'sal' (2.85.0127), M *ʔ-a·káy* <trākáy> 'salt' (2.86.0609)

PYUM *k^w-/s-ʔiy 'salt' > Ki *k^wʔiy¹* 'salt', JT *seʔil^y* 'salt', MGI *ʔa-sil^y* 'salt', Mo *ʔa-ʔi·* 'salt', Ma *sʔi·* 'salt', Qu *ʔa-sʔi·* 'salt', Pa *ʔ-si·* 'salt', Hu *i-ʔi·* 'salt', Ha *ʔi-ʔi* 'salt', Ya *ʔ-ʔi·* 'salt'

PCHO *-uʔwe 'salt' > HCh *ʔ-óʔwe* 'salt', LCh *uwe* 'salt', Teq <weh> 'salt'

NOTES. (1) Cf. PH *(a-)k^{wh}-Vʔíy 'salt' (k^{wh}-03), where only forms with the prefix *k^{wh}- are ad-duced. (2) PCho *-uʔwe < *-uʔe (sequences like *uʔV, where V is a non-labial vowel, are prohibited in Proto-Chontal).

Kroeber 1915: 282 || J. M. Crawford 1976: 185 || Langdon 1979: 638 || Greenberg 1987: 140 || Kaufman 1989: 167 (#(a-)k^{yh}aʔi 'salt' [N]) || Gursky 1990: 1

The following two cases have unexpected *h instead of *ʔ in Pomoan. One possible explanation is a dissimilation *ʔVʔV > *ʔVhV (the word-initial glottal stop possibly belonged to the PH absolutive prefix *ʔa-, see note 10 above). While synchronically in Proto-Pomo the choice between word-initial *ʔ and *h depends on the following consonant, historically the direction of dissimilation might have been the reverse in this particular case.

(ʔ-05) PH *(a-)ʔáw 'fire; wood'

SHA *ʔá·w[a]* 'wood' (Silver 1964: 176); NRS HA <ga'au'> 'wood' (Dixon 1905: 216)

ACH *áw* 'wood; sticks'

YAN *ʔau-na* 'fire'

PPOM *ʔohhó 'fire' > Pk *ʔoho*, Ps *ʔoh·o*, Pn *hó*, Pc *hó*, Pne *ʔóho*, Pse *xó*, Pe *xó*

SAL A *ʔ-aʔáw* <tra'aw> 'fire' (2.86.0021), M *ʔ-aʔáw* <tra'aw> 'lumbre' (2.86.0022)

PYUM *(ʔ-)ʔa(·)w > Ki *ʔ-ʔa·w²* 'fire', Co *ʔa-ʔá·* 'fire, embers, electricity', JT *ʔa·w ~ ʔa-ʔa·w* 'a fire', MGI *ʔa·w* 'fire', Mo *ʔa-ʔaw* 'fire', Qu *ʔa-ʔáw* 'fire', Pa *ʔo-ʔo ~ o-ʔo* 'flame; fire', Hu *o-ʔó* 'fire', Ha *ʔo-ʔo-ʔo* 'fire', Ya *ʔ-ʔó* 'fire'

NOTES: (1) McLendon (1973: 75) reconstructs pPom *ʔohxó, while Oswalt (1972: 35) reconstructs pPom *ʔoho. I assume that the Proto-Pomoan form was *ʔohho, where the first *h is a laryngeal increment. Geminated *hh was strengthened in Southeastern Pomo and Eastern Pomo and simplified in all other languages save Southern Pomo. (2) Notions 'fire' and 'wood' are frequently colexified in world's languages, see <https://clics.clld.org/edges/221-1803>.

Kroeber 1915: 282 || Sapir 1917: 4 || Sapir 1920b: 284 || Sapir 1925: 410 (*(a)awa) || Greenberg & Swadesh 1953: 218 || Haas 1964: 78 || McLendon 1964: 131 || Gursky 1965: 175 || Gursky 1968: 30 || Gursky 1974: 188 (#(ʔa)ʔahw#), 211 || Langdon 1979: 637 || Greenberg 1987: 136 || Kaufman 1989: 141 (#a-Hów or #a-Háw [a ~ o] 'fire; (fire)wood; to make a fire' [N/S]) || Leshchiner & Nikolaev 1992: 400 (*ʔáŋHa) || Shaul 2020b: 301

(?-06) PH *(a-)ʔiy ‘tree, wood’

YAN CN ʔi-na, Y ʔi-wi ‘stick, wood, tree’

PPOM *ʔahháy ‘wood’ > Pk ʔahay, Ps ʔah-ay, Pn háy, Pc háy, Pne ʔahá, Pse xay, Pe xáy

PYUM *(?-)ʔiy ‘wood’ > Co ʔi· ‘wood’, JT ʔi· ‘wood, stick, lumber, tree’, MGI ʔilʷ ‘wood, stick, tree’, Mo ʔa-ʔi· ‘wood, firewood’, Ma ʔi· ‘wood’, Qu ʔa-ʔi· ‘wood, tree’, Pa ʔi-ʔi· ~ i-ʔi· ‘wood; stick’, Hu i-ʔi ~ i-ʔi· ~ ʔi ~ i· ‘wood, stick’, Ha i-ʔi ~ i-ʔi-ʔi ‘wood, stick, log’, Ya ʔ-ʔi· ‘wood, tree’

SER ʔéʔé ‘plant, bush, tree; wood; stick, club; meter’ (pl.: ʔéʔé-t)

PCHO *ʔe[x] ‘tree’ > HCh al-ʔek ‘tree, wood’, LCh ʔeh ‘tree, branch’, Teq <eh> ‘tree, wood, plant’

NOTES. (1) McLendon (1973: 75) reconstructs pPom *ʔahxáy, while Oswalt (1972: 35) reconstructs pPom *ʔahay. I assume that the Proto-Pomoan form was *ʔahháy, where the first *h is a laryngeal increment. Geminated *hh was strengthened in Southeastern Pomo and Eastern Pomo and simplified in all other languages save Southern Pomo. (2) The vowel *a instead of expected *e (< PH *i) in Pomoan may be due either to a translaryngeal assimilation *ʔa-ʔiy > *ʔa-ʔáy or to a contamination with PH *(a-)ʔáw ‘fire; wood’ before *aʷ in the latter has yielded pPom *o.

Kroeber 1915: 282 || Sapir 1917: 9 || Sapir 1925: 411 (*ixai?) || Greenberg & Swadesh 1953: 220 || Gursky 1965: 187 || Gursky 1968: 39 || Gursky 1974: 209 (#ʔi#) || J. G. Crawford 1976: 320 || Waterhouse 1976: 341 || Langdon 1979: 639 || Greenberg 1987: 268 || Kaufman 1989: 141 (#i-ʔly ‘tree, wood’ [N/S]) || Leshchiner & Nikolaev 1992: 388 (*ʔähö)

(?-07) PH *(a/i-)k^haʔa ‘stone’ (= k^h-10)CHI q^haʔa <kháʔ> ‘rock’ (2.23.0606)

KAR as / asa- ‘rock; stone; pit of fruit; in some compounds, earth, dirt’

SHA ʔíc-aʔ ‘rock’ (Silver 1964: 175)

YAN ka[i]-na ‘stone’

NOTES. Yana k̄ instead of expected x is due to the change *k^haʔa > *k^hʔa > *ka.

Dixon & Kroeber 1913: 651 || Sapir 1917: 9 || Sapir 1925: 411 (*ixa) || Haas 1964: 81 || Silver 1964: 175 || Gursky 1974: 207 || Kaufman 1989: 138 (#[i/a]-xʔáʔ ‘stone’ [N/S])

(?-08) PH *(a-)ʔil ‘flea’

SAL A ʔ-a-yél <trājél> ‘pulga’ (2.85.0859)

PYUM *ʔilʷ ‘louse’ > Ki ʔilʷ ‘louse, vermin; worm’, Co ʔi-ʔilʷ ‘head lice’, JT ma-ʔilʷ ‘lice (*Pediculus humanus spp.*)’, MGI ʔ-m-ilʷ ‘his head louse’, Mo nʷ-i-ʔi-lʷ ‘head louse’, Pa ʔil ‘worm’, Hu ʔil ‘worm’, ʔel ‘a louse’, Ha il ‘worm’, ʔel ‘louse’, Ya ʔ-ʔi-l-a ‘worm’, h-ʔél ‘louse’

PCHO *ʔilʷ ‘flea’ > HCh gál-ʔilʷ ‘flea’, LCh ilʷ ‘flea; the common flea that dogs would have’

Gursky 1974: 197 || Waterhouse 1976: 337 || Greenberg 1987: 136 || Kaufman 1989: 152 (#ʔilʷ ‘louse; flea’ [N/S])

1.3.2. Comparison with Kaufman’s reconstruction

Kaufman reconstructs two laryngeals for Proto-Hokan: *ʔ and *h (see Table 12).

PH	Chi	Kar	Sha	Ach	Ats	Yan	Wsh	PPom	Sal	PYum	Ser	PCho	Tol
*ʔ	Ø	Ø	ʔ	ʔ	ʔ	Ø		Ø	ʔ	*ʔ		*ʔ	Ø
*h	h	h	h	h	h	h		*h	ʔ	*ʔ, *y, *w		ʔ	

Table 12. Correspondences of Hokan laryngeals according to Kaufman (1989).

Kaufman’s correspondences for *ʔ do not differ significantly from the correspondences postulated here. Cases where Kaufman reconstructs *h will be discussed in one of the next papers in this series under vowel-initial roots.

The next paper in this series will deal with the reconstruction of sonorants.

Language abbreviations

Ach — Achumawi; Ats — Atsugewi; BI — Barona 'Iipay; Chi — Chimariko; Co — Cocopa; Ha — Havasupai; HCh — Highland Chontal; Hu — Hualapai; JT — Jamul Tiipay; Kar — Karuk; Ki — Kiliwa; Kon — Konomihu; LCh — Lowland Chontal; Ma — Maricopa; MGI — Mesa Grande 'Iipay; Mo — Mohave; NRSha — New River Shasta; Okw — Okwanuchu; Pa — Paipai; Pc — Central Pomo; PCho — Proto-Chontal; Pe — Eastern Pomo; Pk — Kashaya; Pn — Northern Pomo; Pne — Northeastern Pomo; PPom — Proto-Pomoan; Ps — Southern Pomo; Pse — Southeastern Pomo; PYum — Proto-Yuman; Qu — Quechan; Sal — Salinan; Sal A — Antoniano Salinan; Sal M — Migueleño Salinan; Ser — Seri; Sha — Shasta; Teq — Tequistlateco; Wsh — Washo, Ya — Yavapai; Yan — Yana; Yan C — Central Yana; Yan N — Northern Yana; Yan Y — Yahi.

Literature

- 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 Hokaltec material. *Language* 32(1): 42–48. doi:10.2307/410651.
- Bright, William & Susan Gehr. Karuk dictionary. Available online at: <http://linguistics.berkeley.edu/~karuk/karuk-dictionary.php>. [Accessed 20 June, 2019].
- Buckley, Eugene. 2024. Pomoan. In: Carmen Dagostino, Marianne Mithun, Keren Rice (eds.). *The languages and linguistics of Indigenous North America: A comprehensive guide* (The World of Linguistics 13.2): vol. 2, 1223–1246. Berlin / Boston: De Gruyter Mouton. doi:10.1515/9783110712742-049.
- 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.
- 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>.
- Couro, Ted & Christina Hutcheson. 1973. *Dictionary of Mesa Grande Diegueño*. Banning, CA: Malki Museum Press.
- 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, James M. 1989. *Cocopa dictionary* (University of California Publications in Linguistics 114). Berkeley / Los Angeles / London: University of California Press.
- 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.
- de Angulo, J. & L. S. Freeland. 1925. The Chontal language (dialect of Tequixistlan). *Anthropos* 20(5/6): 1032–1052. Available online at: <https://www.jstor.org/stable/40444929>.
- 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. & Alfred L. Kroeber. 1913. 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. 1919. Linguistic families of California. *University of California Publications in American Archaeology and Ethnology* 16(3): 47–118.
- Good, Jeff. 2004. A sketch of Atsugewi phonology. Presented at the SSILA Annual Meeting, Boston. Available online at: <https://www.researchgate.net/publication/267809944>. [Accessed 4 August, 2024].

- Gordon, Lynn. 1986. *Maricopa morphology and syntax* (University of California Publications in Linguistics 108). Berkeley, CA: University of California Press.
- Greenberg, Joseph H. 1987. *Language in the Americas*. Stanford, CA: 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. 1965. Ein lexikalischer Vergleich der Algonkin-Golf- und Hoka-Subtiaba-Sprachen (1). *Orbis* 14: 160–215.
- 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. 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. 1980. Notes on Karok internal reconstruction. 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): 67–76. The Hague / Paris / New York: Mouton Publishers. doi:10.1515/9783110808681.67.
- Halpern, A. M. 1946a. Yuma I: Phonemics. *International Journal of American Linguistics* 12(1): 25–33. doi:10.1086/463884.
- Halpern, A. M. 1946b. Yuma II: Morphophonemics. *International Journal of American Linguistics* 12(3): 147–151. doi:10.1086/463905.
- Halpern, A. M. 1946c. Yuma III: Grammatical processes and the noun. *International Journal of American Linguistics* 12(4): 204–212. doi:10.1086/463915.
- Halpern, A. M. 1947a. Yuma IV: Verb themes. *International Journal of American Linguistics* 13(1): 18–30. doi:10.1086/463926.
- Halpern, A. M. 1947b. Yuma V: Conjugation of the verb theme. *International Journal of American Linguistics* 13(2): 92–107. doi:10.1086/463933.
- Halpern, A. M. 1947c. Yuma VI: Miscellaneous morphemes. *International Journal of American Linguistics* 13(3): 147–166. doi:10.1086/463945.
- Hernández, Fortunato. 1902. *Laz razas indígenas de Sonora y la guerra del Yaqui*. Mexico: J. de Elizalde.
- Hill, Kenneth C. 2020. *Wick Miller’s Uto-Aztecan cognate sets* (Survey of California and Other Indian Languages, Report 18). Vol. I. Berkeley: University of California, Berkeley. Available online at: <https://escholarship.org/uc/item/9px6p8h8>.
- Hinton, Leanne. 1984. *A dictionary of the Havasupai language*. The Havasupai Tribe.
- Jacobsen, Jr., William H. 1954. [Binder of field notes on Antoniaño Salinan], 2014-21.001.061, in “William H. Jacobsen Materials on Indigenous Languages of North America”, California Language Archive, Survey of California and Other Indian Languages, University of California, Berkeley. Available online at: <http://cla.berkeley.edu/item/22959>.
- 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.
- Jany, Carmen. 2009. *Chimariko grammar: areal and typological perspective* (University of California Publications in Linguistics 142). Berkeley / Los Angeles / London: University of California Press.
- Joel, Dina Judith. 1966. *Paipai phonology and morphology*. PhD Thesis. University of California, Los Angeles.

- 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.
- Kroeber, Alfred L. 1915. Serian, Tequistlatecan, and Hokan. *University of California Publications in American Archaeology and Ethnology* 11(4): 279–290.
- Langdon, Margaret. 1970. *A grammar of Diegueño: the Mesa Grande dialect* (University of California Publications in Linguistics 66). Berkeley / Los Angeles / London: University of California Press.
- 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.
- 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 etape. 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. 2010. *The nineteenth century Seri word lists: comparison and analysis* (SIL-Mexico Branch Electronic Working Papers 008). SIL International. Available online at: <http://www.mexico.sil.org/resources/archives/56726>. [Accessed 7 March, 2019].
- Mason, J. Alden. 1918. The language of the Salinan Indians. *University of California Publications in American Archaeology and Ethnology* 14(1): 1–154.
- 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 / Los Angeles / London: 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.
- McLendon, Sally. n.d. A brief word list of Eastern Pomo, McLendon.001, in "Miscellaneous papers from the Survey of California and Other Indian Languages", California Language Archive, Survey of California and Other Indian Languages, University of California, Berkeley. doi:10.7297/X2RV0KNJ.
- 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.
- Miller, Amy & Margaret Langdon. 2008. *Barona inter-tribal dictionary: 'Tipay Aa Tiipay Aa Uumall*. Lakeside, CA: Barona Museum Press.
- Mixco, Mauricio J. 1977. The innovation of /h, hw/ in Kiliwa. *International Journal of American Linguistics* 43(3): 167–175. doi:10.1086/465478.
- Mixco, Mauricio J. 1985. *Kiliwa dictionary* (University of Utah Anthropological Papers 109). Salt Lake City, UT: University of Utah Press.
- Moser, Mary Beck & Stephen A. Marlett. 2010. *Comcaac quih yaza quih hant ihiip hac: cmiiue 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.
- Moshinsky, Julius. 1976. Historical Pomo phonology. 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): 55–75. The Hague / Paris: Mouton. doi:10.1515/9783110819113-004.
- Munro, Pamela, Nellie Brown & Judith G. Crawford. 1992. *A Mojave dictionary* (UCLA Occasional Papers in Linguistics 10). Los Angeles: Department of Linguistics, University of California, Los Angeles.

- Nevin, Bruce. 2020. Achumawi dictionary. Webonary.org. SIL International. Available online at: <https://www.webonary.org/odissi/>. [Accessed 29 August, 2025].
- 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.
- Oswalt, Robert & Eugene Buckley. 2018. Kashaya — English dictionary. Webonary.org. SIL International. Available online at: <https://www.webonary.org/kashaya/>. [Accessed 20 June, 2019].
- Oswalt, Robert L. 1972. ProtoPomo reconstructions, Oswalt.002.004, in “Robert Louis Oswalt Papers on Pomoan Languages”, California Language Archive, Survey of California and Other Indian Languages, University of California, Berkeley. Available online at: <http://cla.berkeley.edu/item/2133>.
- Oswalt, Robert L. 1976a. Baby talk and the genesis of some basic Pomo words. *International Journal of American Linguistics* 42(1): 1–13. doi:10.1086/465380.
- Oswalt, Robert L. 1976b. Comparative verb morphology of Pomo. 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): 13–28. The Hague / Paris: Mouton. doi:10.1515/9783110819113-002.
- Salt River Pima-Maricopa Indian Community — O'Odham Piipaash Language Program. 2000. *Piipayk M'iim: A Maricopa vocabulary and phrase book*. Scottsdale, AZ: Salt River Pima-Maricopa Indian Community Cultural Resources Department O'Odham Piipaash Language Program. Available online at: <https://www.saltriverschools.org/common/pages/DisplayFile.aspx?itemId=8926463>. [Accessed 22 September, 2024].
- 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. 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. 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.
- Shaterian, Alan William. 1983. *Phonology and dictionary of Yavapai*. PhD Thesis. University of California, Berkeley.
- Shaul, David Leedom. 2020. *Salinan language studies* (LINCOM Studies in Native American Linguistics 85). Muenchen: LINCOM GmbH.
- Silver, Shirley. 1961. [Shasta vocabulary], Silver.002, in “Shirley Silver Papers on California Indian Languages”, California Language Archive, Survey of California and Other Indian Languages, University of California, Berkeley. doi:10.7297/X2JM27K0.
- 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. 1966. *The Shasta language*. PhD Thesis. University of California, Berkeley.
- 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.
- Sitjar, Bonaventure. 1861. *Vocabulary of the language of San Antonio mission, California. By Father Bonaventure Sitjar, of the order of St. Francis*. (Shea's Library of American Linguistics VII). New York: Cramoisy Press.
- Talmy, Leonard. 1972. *Semantic structures in English and Atsugewi*. PhD Thesis. University of California at Berkeley.
- Talmy, Leonard. n.d. [Atsugewi vocabulary], Talmy.002.001, in “Leonard Talmy Papers on the Atsugewi Language”, California Language Archive, Survey of California and Other Indian Languages, University of California, Berkeley. Available online at: <http://cla.berkeley.edu/item/1744>.

- Talmy, Leonard. n.d. [Atsugewi vocabulary], Talmy.002.003, in “Leonard Talmy Papers on the Atsugewi Language”, California Language Archive, Survey of California and Other Indian Languages, University of California, Berkeley. Available online at: <http://cla.berkeley.edu/item/1746>.
- Turner, Katherine. 1987. *Aspects of Salinan grammar*. PhD Thesis. University of California, Berkeley.
- Turner, Paul R. & Shirley Turner. 1971. *Dictionary: Chontal to Spanish-English, Spanish to Chontal*. Tucson, AZ: The University of Arizona Press.
- Watahomigie, Lucille J., Jorigine Bender, Malinda Powskey, Josie Steele, Philbert Watahomigie Sr. & Yamamoto, Akira Y. 2003. *A dictionary of the Hualapai language*. Osaka: Osaka Gakuin University, Japan: Endangered Languages of the Pacific Rim, Faculty of Informatics.
- 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, IL: 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, IL: 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, IL: 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. doi:10.31826/jlr-2018-161-211.
- Zhivlov, Mikhail. 2024. Hokan I: A review of comparative studies. *Journal of Language Relationship* / *Вопросы языкового родства* 22(1–2): 162–188. doi:10.31826/jlr-2024-221-209.
- Zhivlov, Mikhail. 2025. Hokan II: Labial and coronal obstruents. *Journal of Language Relationship* / *Вопросы языкового родства* 23(1–2): 51–93.

М. А. Живлов. Хока III: велярные шумные и гортанная смычка

Настоящая статья продолжает цикл работ, посвященных переоценке остающейся по сей день спорной гипотезы о родстве хоканских языков. В данной статье классический сравнительно-исторический метод применяется к рефлексам прахоканских велярных шумных и гортанной смычки. Полученные результаты сравниваются с хоканской реконструкцией Терренса Кауфмана (1989).

Ключевые слова: сравнительный метод; языки хока; индейские языки; историческая фонетика.

Relative chronology of a series of phonetic changes in Old Indo-Aryan clusters based on data from the Kullui language¹

In this study I attempt to reconstruct the relative chronology of phonetic changes affecting Old Indo-Aryan consonant clusters in the Kullui language (Himachali). Analysis confirms the preservation of archaic clusters like those with *r* and *kṣ*, the latter shifting to *tʃʰ* independently from the Southern Indo-Aryan languages. The relative sequence of changes including *-mr-* > *-mb-*, NT > ND, and (limited) ND > NN is examined. Bloch's replication (anticipation) theory for *r*-metathesis is supported over Morgenstierne's; the rule of assimilation of the second non-homorganic cluster is derived. Two potential chronological models for these shifts are proposed based on Kullui evidence.

Keywords: Kullui language; Indo-Aryan languages; relative chronology; metathesis.

1. Introduction

In a previous paper (Krylova 2024) I have addressed the problem of accent shift in the Kullui language, while also providing basic information about Kullui and the historical phonetics of the New Indo-Aryan (NIA) languages. In the present study, I would like to follow this research up by examining a series of changes in Old Indo-Aryan (OIA) consonant combinations, based on Kullui data. Since we will be dealing with phonetic changes dating to the Middle Indo-Aryan (MIA) period, we should first dwell in more detail on a number of specific archaisms in the historical development of the ancestor language of Kullui (an unattested hypothetical "Kullui Prakrit") during this period. It may be assumed that this "Kullui Prakrit" still preserved a number of OIA clusters, especially of the type "labial / velar / non-aspirated dental + *r*", as well as *rṣ*, *rś*, *ml*, *kṣ*. Furthermore, Kullui has preserved the distinction between two sibilants, since *ś*, *ṣ* > *ʃ*, while *s* > *s*. Let us consider the most important of these archaisms.

2. Preservation of *kṣ*

During the Middle Indo-Aryan period, the normal development of OIA *kṣ* in Central and Eastern Indo-Aryan languages was *kh-*, *-kkh-* (Masica 1991: 173; Oberlies 2001: 106), while the Western dialectal development became *-ch-*. In most Himachali dialects (Hendriksen 1986: 192) OIA *kṣ* yielded *kh*. However, in Kullui *kṣ* developed into an aspirated affricate, which, nevertheless, did not merge with Middle Indo-Aryan *ch*. Unfortunately, the size of our dictionary does not allow us to examine initial clusters separately, but the behavior of medial clusters suggests

¹ This research was supported by the Russian Science Foundation grant No 24-78-10052 (<https://rscf.ru/project/24-78-10052>).

that initial ones behaved similarly, with the exception that medial clusters became geminates, and initial ones became single consonants.² OIA *kṣ* in Kullui generally corresponds to *tʰ*:

Kullui	OIA
<i>tʰʷŋ</i> ‘moment’	<i>kṣaṇa-</i> ‘moment’ ³
<i>mɔtʰi</i> ‘fly’	<i>makṣikā-</i> ‘fly, bee’
<i>tʰar</i> ‘ashes’	<i>kṣāra-</i> ‘alkali’
<i>ɔtʰi</i> ‘eye’	<i>akṣi-</i> ‘eye’
<i>tʰɔpŋa</i> ‘to boil away’	<i>kṣapyate</i> ‘is destroyed’
<i>tʰura</i> ‘knife’	<i>kṣura-</i> ‘razor’
<i>tʰa</i> ‘sharp’	<i>tikṣṇa-</i> ‘sharp’
<i>antʰa</i> ‘raspberry, blackberry’	<i>ākṣika-</i> ‘Morinda tinctoria’
<i>kutʰ</i> ‘menstrual blood, belly of a pregnant woman’	<i>kukṣi-</i> ‘belly’

At the same time, OIA *ch*, both word-initially and medially, systematically corresponds to *tsʰ* in Kullui:

Kullui	OIA
<i>putsʰŋa</i> ‘ask’	<i>pr̥cchati</i> ‘asks’
<i>tsʰaũ</i> ‘shadow’	<i>chāyā-</i> ‘shadow’
<i>mutsʰŋa</i> ‘knead’	<i>mūrchayati</i> ‘makes hard’
<i>bətsʰaŋ</i> ‘bed’	<i>*vicchādāna-</i> ‘covering’ ⁴
<i>kɔtsʰi</i> ‘ladle’	<i>*kaṭacchu-, kaḍacchaka-</i> ‘ladle’ ⁵
<i>tsʰelu</i> ‘kid’	<i>*chagalla-</i> ‘goat’ ⁶
<i>tsʰati</i> ‘chest’	<i>*chātti-</i> ‘chest’ ⁷
<i>mutsʰa</i> ‘moustache’	<i>*mucchā-</i> ‘moustache’ ⁸

² Hereafter, in cases where initial and medial clusters develop similarly, I combine the few available examples of both.

³ Here and below, unless otherwise indicated, Old Indo-Aryan protoforms are given according to Turner 1966. Kullui lexicon is almost absent from Turner’s data, but etymologies of Kullui forms, for the most part, are quite transparent. Accent markers as well as superfluous length markers above the vowels *e* and *o* have been removed. All the abbreviations in notes (language names, grammatical terms, etc.) are quoted directly from Turner’s dictionary.

⁴ Cf. Pa. *vicchādānā-* f. ‘covering’. In R. L. Turner’s dictionary, a significant number of entries are represented by asterisked forms, indicating reconstructions for which no Sanskrit or Vedic equivalents are recorded. When using such data, it is important to recognize that sometimes forms have only been reconstructed to the MIA level, especially those containing geminates. Geminated consonants are not very common in Sanskrit, but in the MIA period, most OIA consonant clusters shifted to geminates, and later were reduced to single consonants in the NIA languages. Because of this, it is generally impossible to determine, based solely on data from the NIA languages, which specific OIA consonant cluster serves as the origin of a given consonant. In such cases, when the reconstructed form obviously does not represent the OIA level, it would be more accurate to call it Common MIA, despite the fact that at the MIA stage, Common IA had already disintegrated. To mark these reconstructions as OIA entries would be a simplification. To enable the reader to assess the phonetic and semantic accuracy of the reconstructed forms, I hereafter provide in footnotes some Indo-Aryan cognates cited in Turner’s dictionary.

⁵ Pa. *kaṭacchu-* m. ‘ladle, spoon’; Pk. *kaḍacchu-* f. ‘ladle’.

⁶ Pa. *chakala-* m. ‘goat’, NiDoc. *chaḡalaḡa*, Pk. *chagala-*.

⁷ S. P. *chāti* f., Ku. N. B. Or. Mth. *chāti*; H. *chāti* f. ‘chest’.

⁸ S. *mucha* f. ‘moustache’, L. *mucch* f., P. *mucch* m., Ku.gng. *mūch*, Or. *mucha*, H. *mūch* f., G. *mūch*, *mūch* f.

<i>ots^ha</i> ‘small’	<i>*occha-</i> ‘small’ ⁹
<i>ts^hinda</i> ‘hole in ear’	<i>*chindati</i> ‘cuts, splits’ ¹⁰
<i>ts^hilṛa</i> ‘type of flatbread’	<i>*chilla-</i> ‘skin, hide’ ¹¹
<i>tsifla</i> ‘slippery’ (with metathesis)	<i>picchala-</i> ‘slippery’
<i>ts^hidzṇa</i> ‘crack’	<i>chidyate</i> ‘is cut, split’
<i>ts^hiṛṇkṇa</i> ‘sprinkle’	<i>*chiṭ-</i> ‘sudden movement, flash, splash’ ¹²

Moreover, other OIA clusters that yielded MIA *ch*, *cch* also yielded *ts^h* in Kullui, i.e., merged with OIA *ch*:

Kullui	OIA
<i>pits^he</i> ‘behind’	<i>*paśca-</i> ‘hinder part’ (contamination with <i>prṣṭha-</i> ‘back’) ¹³
<i>bits^hu</i> ‘scorpion’	<i>vṛścika-</i> ‘scorpion’
<i>mṛts^hi</i> ‘fish’	<i>matsya-</i> ‘fish’
<i>bṛts^hu</i> ‘calf’	<i>vatsa-</i> ‘calf’

The shift of *kṣ* to an affricate which does not merge with the descendants of MIA *ch* unites Kullui only with Dardic languages, but not with such Southern and Western Indo-Aryan languages as Gujarati, Marathi, and Sinhalese, where such a merger with *ch* did occur. Thus, in Northern languages, the shift *kṣ* > *tʃ^h* occurred independently, as noted in Masica 1991: 460, Southworth 2005: 158, and, judging by my data, at quite a late date. It can be assumed that in Kullui, the combination *kṣ* was preserved as an archaism during the MIA period.

3. Preservation of a series of clusters with *r*

A specific archaism of the Himachali group and some other Northwestern Indo-Aryan languages is the preservation of OIA clusters of the type “velar or labial stop + *r*” (Hendriksen 1986: 196–197), and also, apparently, “dental non-aspirated stop + *r*”. The latter group of clusters later underwent a shift to palatal affricates in most Himachali idioms, including Kullui, but remained unchanged in Mandiyali. Such combinations are also partially preserved in a number of other Northwestern idioms, such as Punjabi, Lahnda, and Sindhi, as well as in some Dardic idioms (Hendriksen 1986: 201) and in Romani (Turner 1985a: 183). At the same time, over the “mainstream” course of Indo-Aryan language development, *r* in these sequences underwent complete assimilation. However, in the modern Kullui lexicon, this archaism is not very consistently traceable; there are also quite numerous examples where *r* in such clusters was completely assimilated, and the source of borrowing is not always obvious.

⁹ Pk. *uccha-* ‘low, mean’; S. *ocho* ‘paltry, mean’, L. (Shahpur) *ochā*, awāṇ. *hochā*; P. *ochā* ‘light, vain, absurd’; WPah. bhal. *hoccho* ‘mean, low’; Ku. *occho* ‘mean, contemptible, arrogant’; Or. *ocha* ‘small, mean, exhausted’; Bhoj. *ōch* ‘mean’; H. *ochā* ‘empty, small, silly’.

¹⁰ *chinātti*, pl. *chindānti* ‘cuts off, splits’ RV. Pa. *chindati*, NiDoc. *chinṇati*, pp. *chinṇida*, Pk. *chindāi*.

¹¹ S. *chila*, f. ‘peel, bark, skin, shell’, *chiluru* m. ‘fish scale, crust of bread’; L. *chill* f. ‘skin, bark, rind’, awāṇ. *chillur* ‘skin, peel’, P. *chill*, *chillak* f., *chilkā* m.; WPah. bhal. *chil* f. ‘bark’; N. *chilko* ‘crust, rind, skin of milk’.

¹² K. *chirkāwun* ‘to sprinkle’, P. *chirkaṇā*; Mth. *chirab* ‘to be scattered’; OAw. *chirakāi* ‘sprinkles’; H. *chiraknā*, M. *śidakṇē*.

¹³ Pk. *paccha-* n. ‘hinder part’.

Kullui	OIA
<i>dzagra</i> ‘night vigil by the sacred fire’	<i>jāgrat-</i> ‘vigil’
<i>grā</i> ‘village’	<i>grāma-</i> ‘village’
<i>grō</i> ‘thin bamboo’	<i>*grava-</i> ‘branch’ ¹⁴
<i>krifŋa</i> ‘to smooth wood with a tool’	<i>*kriśati, *kriśnāti</i> ‘presses’ ¹⁵
<i>tsukri</i> ‘sty (on eye)’	<i>cukra-</i> ‘sour’
<i>prehi</i> ‘spleen’	<i>*prihan-</i> ‘spleen’ ¹⁶
<i>porfu</i> ‘sweat’	<i>*praśūta-</i> (past participle of <i>*praśūyate</i> ‘swells’) ¹⁷
<i>b^hrahu-dzi</i> ‘younger brother’s wife’	<i>bhrātṛvadhū-</i> ‘brother’s wife’
<i>b^hrūa</i> ‘eyebrows’	<i>*bhrumu-</i> ‘eyebrow’ ¹⁸
<i>b^hraru</i> ‘siblings’	<i>bhrātaraḥ</i> ‘brothers’
<i>dzifa</i> ‘morning’	<i>dr̥śika-</i> ‘shining’ (preservation of the syllabic sonorant as <i>r</i> indicates an early Sanskritism)
<i>ād̥z</i> ‘guts’	<i>āntra-</i> ‘intestines’
<i>gatfi</i> ‘female belt’	<i>gātrikā-</i> ‘belt’
<i>datfi</i> ‘sickle’	<i>*dātrī-</i> ‘knife, sickle’ ¹⁹
<i>d̥zatf</i> ‘festival, fair’	<i>yātrā-</i> ‘journey, pilgrimage’
<i>tfuŋa</i> ‘break (vi)’	<i>truṭyati</i> ‘breaks (vi)’
<i>t^hatf</i> ‘sheep pasture in the mountains’	<i>sthātra-</i> ‘place’
<i>gontf</i> ‘cow urine’	<i>gomūtra-</i> ‘cow urine’
<i>gojntf</i> ‘cow urine’	<i>gomitra-</i> ‘cow urine’
<i>tforna</i> ‘break (vt)’	<i>troṭayati</i> ‘breaks (vt)’
<i>tf^hōtf</i> ‘mushroom’	<i>chattra-</i> ‘parasol’
<i>d̥zəŋkaŋa</i> ‘threaten’	<i>*draṭ-</i> ‘press’ ²⁰
<i>həwd̥z</i> ‘turmeric’	<i>haridrā-</i> ‘turmeric’
<i>k^hatf</i> ‘pit’	<i>khātra-</i> ‘hole’
<i>mutŋa</i> ‘urinate’	<i>mūtrayati</i> ‘urinates’
<i>poŋfa</i> ‘leaf’	<i>pattra-</i> ‘leaf’
<i>tfeka</i> ‘loin’	<i>*trikka-</i> ‘hips’ ²¹

In the NIA period, common MIA affricates shifted to the dental series in Kullui (*c*, *ch*, *j*, *jh* > *ts*, *ts^h*, *dz*, *dz^h*), while the clusters *tr*, *dr*, *kṣ* gave rise to a new series of palatal affricates: *tr* > *tf*, *dr* > *d̥z*, *kṣ* > *tf^h*.

Furthermore, it can be noted that *r* in the combinations *rṣ*, *rś* was also preserved in Kullui, although the number of examples is small:

¹⁴ Wg. *grō* ‘branch’, Paš. *lawā*.

¹⁵ *kliśnāti* ‘torments’ MBh., *sāmkliśnāti* ‘presses together’ ŚBr. WPah. roh. *kriśṇō* ‘to comb’, Kal. rumb. *kriṣnim* ‘I knead, crush’.

¹⁶ *plīhān* (later *plīhan-*) m. ‘spleen’ AV., *plīhā-* f. lex., *plīhan-* m. Yājñ., Pr. *pṣṭigē* ‘spleen’.

¹⁷ Ku. *pasujaṇo* ‘to steam, get steamed, convert into vapour’.

¹⁸ BHSk. *bhramu-* ‘eyebrow’, Pa. *bhamu-* f., Pk. *bhumā-*, °*mayā-*, *bhamayā-* f.; — Pa. *bhamuka-* m., *bhamukha-* (BHSk. *bhrumukha-*: infl. by *mūkha-* ?), Pk. *bhamuha-* n. (cf. *muha-* n.), *bhamuhā-*, *bhaiūmhā-*, *bhaiūhā-* f., Ap. *bhōhā-* f.

¹⁹ Pk. *dattiyā-* f. ‘little sickle’; Paš.ar. *drā^tt*, chil. *lāit* ‘sickle’, K. *dröc^u* f., L. *dātrī* f., awāṇ. *dātrī* f. P. *dātrī*, *dātrī*, *darātī*, *darātī*, *dāt(t)ī* f., WPah.cur. cam. *drāṭī*, bhad. bhal. *ḍlāṭī*, khaś. šeu. *lāṭī*, roh. *dace*.

²⁰ M. *daḍṇē* ‘to press down’, *daḍaṇ* f. n. ‘hiding place’, *daḍā* m. ‘plug’, *daḍakṇē*, *daḍapṇē* ‘to press down’.

²¹ L. *tirik* ‘loins, hips’; P. *tikk* m. ‘waist, loins, body below waist’; — ext. -*la-*: Paš.ar. *trikaló*, dar. weg. *lekali*, chil. *lexeló* ‘tripod’, Gaw. *lakala*; L. *trikkal* f. ‘back’, awāṇ. *trikkul* ‘neck’.

Kullui	OIA
<i>arfu</i> ‘mirror’	<i>ādarśa-</i> ‘mirror’
<i>bərʃ</i> ‘year’	<i>varṣa-</i> ‘год’
<i>bərʃoṛi</i> ‘generation’	* <i>varṣakoṭi-</i> ‘10 million years’ ²²
<i>prʃa</i> ‘rib’	<i>parśu-</i> ‘rib’

Although the historical phonetics of Indo-Aryan languages has been relatively well studied (primarily by Ralph L. Turner in his numerous publications), fairly little attention has been paid to the relative chronology of individual phonetic changes. Below, I shall attempt to fill this gap by analyzing the relative chronology of changes in a number of OIA clusters during the MIA period. First of all, let us consider these phonetic changes separately.

4. Shift of dental + *y* clusters to affricates

As a result of the development of clusters *-ty-*, *-dy-*, *-dhy-*, new palatal affricates were formed in the ancestor language of Kullui, as in all MIA languages. In the NIA period, they shifted to the respective dental affricates in Kullui and a number of other Himachali languages:

Kullui	OIA
<i>betsŋa</i> ‘sell’	* <i>vetyayati</i> ‘sets a price’ ²³
<i>nɔtsŋa</i> ‘dance’	<i>nṛtyati</i> ‘dances’
<i>bidzʃi</i> ‘lightning’	<i>vidyullatā-</i> ‘forked lightning’
<i>badzŋa</i> ‘sound (of a musical instrument)’	<i>vādyate</i> ‘sounds’
<i>mindzu</i> ‘brain’	<i>medya-</i> ‘fatty’
<i>ɔdz</i> ‘today’	<i>adya-</i> ‘today’
<i>daz</i> ‘dowry’	<i>dāyādyā-</i> ‘inheritance’ (<i>dz > z</i> due to contamination with Class. Persian <i>jahāez</i> ‘dowry’)
<i>budzʰŋa</i> ‘know, care for’	<i>budhyate</i> ‘observes, understands’
<i>edzŋa</i> ‘come’	* <i>adhiyeti/adhyeti</i> ‘approaches’ ²⁴ (vowel changed due to contamination with <i>eŋa</i> ‘come’)
<i>mɔdzʰe</i> ‘between’	<i>madhya-</i> ‘middle’, <i>madhyena</i> (l.sg. from <i>madhya-</i>)
<i>sɔndzʰ</i> ‘evening’	<i>samdhya-</i> ‘twilight’

5. Middle Indo-Aryan shift *-mr-* > *-mb-*

Although many phonetic changes are attested in MIA written records, it would be wrong to claim that the historical phonetics of this period is fully reflected in them. Thus, the shift *-mr-* > *-mb-* is a frequent (though not entirely regular) MIA phenomenon that requires additional scrutiny. Its result is already recorded in Pali (Masica 1991: 175–176, Oberlies 2001: 108–109):

²² The protoform (cf. also Hindi *barsaurī* ‘annual, annual rent’) is not attested in Turner 1966, and is proposed by the author of this article.

²³ Denom. from **vetyā-* ‘hire, price’ in (MIA.) *vecā-* (v.l. *vetā-*) f. ‘hire, wages’ lex. < **veccā-*, cf. *vetana-* n. ‘hire, wages’ MBh. (*vetanena krītaḥ* ‘hired’ Pāṇ.com.), ‘price’ Rājat. (*vikrīṇānō 'lpavetanaiḥ*); cf. *veti* ‘procures’ RV., √*vī* (J. C. W.).

²⁴ Cf. *ādhiyeti* ‘turns the mind to’ RV. and *ādhiyacchatī* ‘approaches’.

OIA	Pali
<i>tāmra</i> - ‘copper’	<i>tamba</i> - ‘copper’
<i>amra</i> - ‘mango’	<i>amba</i> - ‘mango’

At the same time, both Turner (1985b: 367) and Morgenstierne (1973: 238) assume an intermediate stage for this shift: *-mr-* > *-mbr-* > *-mb-* (the loss of *r* in triconsonantal sequences is a common MIA development, see Masica 1991: 179). In this case, the chronology of the shift implies two stages preceding the period of MIA written records. An epenthetic *b* is also observed in Kullui in similar contexts before the suffix *-ŋ-*, for example, *tsɔmbɾa* ~ *tsɔmɾa* ‘skin’ from OIA *carman*- ‘skin’, *p^hembɾa* ‘kind of soup’, apparently from OIA *phena*- ‘foam’ with dissimilation *ŋr* > *mɾ* (cf. a similar dissimilation in **buŋŋi* ‘knitting’ > *bumɾi* ‘quilt pin’).

Based on Kullui data, the shift *-mr-* > *-mb-* is represented by the following examples:

Kullui	OIA
<i>tamba</i> ‘copper’	<i>tāmra</i> - ‘copper’
<i>tsambɾa</i> ‘cauldron’	<i>tāmra</i> - ‘copper’
<i>amb</i> ‘mango’	<i>āmra</i> - ‘mango’
<i>tambu</i> ‘bamboo’	<i>tamra</i> - ‘darkening’ ²⁵

Although the hypothesis of a two-stage shift *-mr-* > *-mbr-* > *-mb-* sounds quite plausible, one cannot forget that combinations of labials with *r* in Kullui and a number of other languages have been preserved in many cases. However, in the case of triconsonantal sequences of the type *CPr*, no such archaisms have been found. Apparently, in these cases, *r* was lost in the ancestor language of Kullui according to common MIA rules.

6. Combinations of nasals with stops (NT > ND, ND > NN)

It is rather difficult to date innovations for a language without a continuous written history. As indicated in Turner 1985c: 270 and Masica 1991: 203, the voicing of combinations of nasals with voiceless stops is already observed in Kharoṣṭhī documents, i.e., it belongs to the MIA period, but it is impossible to say exactly when it occurred in the ancestor language of Kullui. Currently, this isogloss widely covers the northwestern area – Punjabi, Sindhi, Himachali, Gōjri, Romani, and some Dardic languages, including Kashmiri (Masica 1991: 203). In Kullui, it is represented quite extensively:

Kullui	OIA
<i>ādʒ</i> ‘guts’	<i>āntra</i> - ‘intestines’
<i>dɔnd</i> ‘tooth’	<i>danta</i> - ‘tooth’
<i>hʃjünd</i> ‘winter’	<i>hemanta</i> - ‘winter’
<i>bɔŋdɾa</i> ‘share’	<i>vaŋṭati</i> ‘shares’
<i>ɔŋdɾa</i> ‘walk’	<i>*haŋṭ-</i> ‘move’ ²⁶

²⁵ The protoform (cf. Skt. *tamāla* ‘having dark bark’, ‘bamboo bark’) is not attested in Turner 1966, and is proposed by the author of this article.

²⁶ D. *hanina* ‘to go about, proceed’; P.bhaṭ. *haṇḍṇā* ‘to walk’; WPah.pāḍ. *haṇḍaṇ* ‘to flow’, jaun. *hāḍṇō* ‘to walk’; B. *hāṭā* ‘to walk’; Or. *hāṇṭibā* ‘to travel’.

<i>g^hṇḍi</i> ‘Adam’s apple’	<i>*ghaṇṭa-</i> ‘throat’ ²⁷
<i>g^huṇḍi</i> ‘heel’	<i>ghuṇṭa-</i> ‘ankle’
<i>g^huṇḍu</i> ‘scarf’	<i>*ghuṇṭa-</i> ‘knot, tag’ ²⁸
<i>kṇḍa</i> ‘thorn’	<i>kaṇṭa-</i> ‘thorn’
<i>k^huṇḍi</i> ‘place for tethering cows’	<i>*khuṇṭa-</i> ‘peg’ ²⁹
<i>ṣuṇḍ^h</i> ‘dry ginger’	<i>śuṇṭhī-</i> ‘dry ginger’
<i>tsuṇḍna</i> ‘sway’	<i>cuṇṭati</i> ‘cuts, strikes’
<i>tuṇḍla</i> ‘armless’	<i>*tuṇṭa-</i> ‘defective’ ³⁰
<i>ṭuṇḍu</i> ‘hand’	<i>*tuṇṭa-</i> ‘defective’ ³¹
<i>kundzi</i> ‘key’	<i>kuñcikā-</i> ‘key’
<i>pṇḍz</i> ‘five’	<i>pañca-</i> ‘five’
<i>pundza</i> ‘claw’	<i>*pahuñca-</i> ‘wrist, forearm’ ³²
<i>tsundz</i> ‘beak’	<i>*cuñca-</i> ‘beak’ ³³
<i>kring</i> ‘shout’	pejorative prefix <i>ku-</i> + <i>*riñk-</i> ‘roar, cry’ ³⁴
<i>sungər</i> ‘pig’	<i>*sūñkara-</i> ‘boar’ ³⁵
<i>-nd-</i> imperfective suffix	<i>-nt-</i> present participle suffix ³⁶

While absolute dating is hard to come by in this case, relative dating is possible: there is another phonetic shift, the order of which relative to NT > ND can be established. This is the shift of OIA clusters of the type "nasal + voiced stop" to a nasal geminate (ND > NN), which has been mentioned in previous publications (Hendriksen 1986: 67; Kogan 2019: 216) alongside the aforementioned NT > ND shift for the same extensive group of languages (the latter author includes Iranian as well). It is important, however, to note that these are two different shifts, of which ND > NN must have preceded NT > ND, otherwise their outcomes would have coincided. In Kullui, there are two examples of this kind, and they are apparently explained by early (before ND > NN in Kullui) borrowing from a language where NT > ND occurred earlier than ND > NN in the ancestor language of Kullui: *kṇmṇa* ‘tremble’ < OIA *kampate* ‘trembles’ and *ṣṭṣ* ‘centipede’ < **śalañka-* ‘grasshopper’³⁷. At the same time, ND > NN is very limited in Kullui; it is absent in most consonant groups of the corresponding type:

²⁷ L. *ghaṇḍī* f. ‘Adam’s apple’, awāṇ. also ‘soft palate’; P. *ghaṇḍ* m., °ḍī f. ‘Adam’s apple’; A. *ghāṭ* ‘protuberance on snout of crocodile’; H. *ghāṭī* f. ‘throat, Adam’s apple, uvula, soft palate’ (→ N. *ghāṭī* ‘throat’); G. *ghāṭṭ* m. ‘throat’, °ṭī f. ‘Adam’s apple’; M. *ghāṭī* f. ‘throat, Adam’s apple, larynx’.

²⁸ S. *ghuṇḍu* m. ‘frown’, *ghuṇḍī* f. ‘knot of thread, tangle’; P. *ghuṇḍī* f. ‘knot, knot of wheat chaff, button’ (→ H. *ghuṇḍī* f. ‘tag, button’); N. *ghūṇī* ‘tag or button to catch in a loop’; B. Or. *ghuṇṭi* ‘cloth button’; G. *ghūṭī* f. ‘entanglement’. WPah.kṭg. *ghòṇḍi* f. ‘doll’ (altern. < **guḍḍa-*), *ghuṇḍu* m. ‘veil carried by the goddess’; J. *ghūḍ* m. ‘veil’.

²⁹ Pk. *khuṇṭa-*, *khoṇṭaya-* m. ‘peg, post’.

³⁰ Pk. *tuṇṭa-* ‘having the hands cut off’.

³¹ P. *tuṇḍ* m. ‘handless arm, withered hand, bare trunk of tree’.

³² L. *pôcā* m. ‘paw’, (Shahpur) *paucā* m. ‘paw, claw’; P. *pahūcā* m. ‘wrist, paw’; N. *pañjā* ‘paw’ OAw. *pahuñcihi* obl. sg. f. ‘wrist’; H. *pahūcā* m. ‘forearm, wrist’; G. *pōhōc* m. ‘wrist’, M. *pohācī* f.

³³ Pk. *cañcu-* f. *cuñculi-* ‘beak’, P. *cuñj* f., WPah. bhal. *cunc* f.

³⁴ Wot. Bshk. *riñg-* ‘to weep, cry’, Tor. *žriñg-*, Sv. *riñg-*; S. *riñgha* f. ‘crying, fretting (of child)’; L. *riñgaṇ* ‘to bellow, roar (of buffalo or she-camel)’; P. *riñṇā* ‘to low’.

³⁵ WPah.jaun. *sūgar* ‘pig’, Ku. *sūgar*; N. *sūgar*, *sūgur* ‘domesticated pig’, WPah.kṭg. *sūngər*, kc. *sunṅur* m. ‘pig’, kṭg. *sūngṭu* m. ‘pig’ (*gṭ* < *grṭ* Him.I 215).

³⁶ On this etymology, see Masica 1991: 203.

³⁷ *śalaka-* m. ‘spider’ lex., Paš.weg. *salāñ* ‘grasshopper’.

Kullui	OIA
<i>dzangli</i> ‘lower (outer) edge of a terraced field’	<i>jāṅgala-</i> ‘arid, desert, wild’
<i>dzɔŋg^h</i> ‘leg’	<i>jaṅghā-</i> ‘shank’
<i>gar</i> ‘coal’	<i>aṅgāra-</i> ‘glowing coal’
<i>gu^hʔa</i> ‘thumb’	<i>aṅguṣṭha-</i> ‘thumb’
<i>lingta</i> ‘tail’	<i>liṅga-</i> ‘characteristic mark, penis’
<i>mɔŋgəl</i> ‘Tuesday’	<i>maṅgala-</i> ‘auspicious sign, Mars’
<i>rɔŋŋa</i> ‘dye, paint’	<i>*raṅgayati</i> ‘dyes’ ³⁸
<i>fiŋg</i> ‘horn’	<i>śṛṅga-</i> ‘horn’
<i>ʔ^hunga</i> ‘dull’	<i>*ṭhuṅga-</i> ‘defective’ ³⁹
<i>ts^hriŋg</i> ‘spark’	<i>sphuliṅga-</i> ‘spark’
<i>aṇḍa</i> ‘egg’	<i>āṇḍa-</i> ‘egg’
<i>b^haṇḍa</i> ‘dishes’	<i>bhāṇḍa-</i> ‘dishes’
<i>rɔŋḍ</i> ‘widow’	<i>raṇḍa-</i> ‘widow’
<i>giṇdu</i> ‘ball’	<i>giṇduka-</i> ‘ball’
<i>ts^hinda</i> ‘hole in ear’	<i>*chindati</i> ‘cuts, splits’
<i>sɔndz^h</i> ‘evening’	<i>saṁdhya-</i> ‘twilight’
<i>t^hɔmba</i> ‘pillar’	<i>stambha-</i> ‘pillar’
<i>tombṛa</i> ‘light green squash’	<i>tumba-</i> ‘bottle gourd’
<i>nimbu</i> ‘lime, lemon’	<i>nimbū-</i> ‘lime’

The following list of examples shows that this shift in Kullui has affected only the clusters *ndh/ṁdh* and partially *mb*, *mbh*:

Kullui	OIA
<i>bɔŋŋa</i> ‘bind’	<i>bandhati</i> ‘binds’
<i>binŋa</i> ‘pierce (ears, nose)’	<i>*vindhati</i> ‘pierces’ ⁴⁰
<i>nhjara</i> ‘darkness’	<i>*andhīkāra-</i> ‘darkness’ ⁴¹
<i>sɔnh</i> ‘evening’	<i>saṁdhā-</i> ‘twilight’
<i>kɔnh</i> ‘shoulder’	<i>skandha-</i> ‘shoulder’
<i>dzmai</i> ‘yawn (noun)’	<i>*jṛmbhāyita-</i> ‘yawned, yawning’ ⁴²
<i>lɔma</i> ‘long’	<i>lamba-</i> ‘hanging, long’
<i>nim</i> ‘neem tree’	<i>nimba-</i> ‘neem tree’
<i>t^hɔmŋa</i> ‘be established’	<i>stambhate-</i> ‘supports, restrains’
<i>k^hmar</i> ‘potter’	<i>kumbhakāra-</i> ‘potter’

In the case of *mb* > *mm*, the assimilation is not entirely regular; there are three examples (*t^hɔmba* ‘pillar’, *tombṛa* ‘light green squash’, and *nimbu* ‘lime, lemon’) in which, despite their general historical regularity, the shift *mb* > *mm* is not recorded. If we also consider examples

³⁸ Pk. *raṅgai* ‘dyes’.

³⁹ Or. *ṭhuṅgā* ‘lopped off, pollard, naked, solitary’; H. *ṭhūgnā* ‘dwarfish’.

⁴⁰ Pk. *vimdhai* ‘pierces’.

⁴¹ S. *ādheru* m. ‘tyranny’, L. *anhērā*, awāṇ. *hanērā*; P. *anher* ‘dark’, m. ‘darkness’, EP. *nherā*; bhaṭ. *nhērī* ‘storm’, WPah. *pañ. nēār*; sod. *nyāro* ‘darkness’.

⁴² Pk. *jaṁbhāia-* n. ‘yawn’.

with other deviations from regular development, Kullui material provides two more examples of the *mb* > *mm* shift (*ama* ‘mother’ < OIA *ambā-* ‘mother’ with irregular compensatory lengthening of the first vowel and *t^hima* ‘mountain’ < OIA **stumba-* ‘mound’⁴³ with an unexplained vowel change and cerebralization), as well as one example of the absence of such a shift (*t^hab^hṛa* ‘stuttering’ < OIA *stambhate* ‘restrains’ with irregular compensatory lengthening and loss of *m*). Thus, the shift *mb* > *mm* in Kullui must still be considered more regular than its absence. It can be noted that none of the examples on the shift *mr* > (*mbr* >) *mb* (namely, *tamba* ‘copper’, *tṣambṛa* ‘cauldron’, *amb* ‘mango’, *tṣambu* ‘bamboo’) underwent the shift *mb* > *mm*, so it can be assumed that *mr* > (*mbr* >) *mb* occurred later than *mb* > *mm*.

Furthermore, judging by the correspondence of Kullui *sṇdz^h* ‘evening’ to OIA *saṇdhyā-* ‘twilight’, it can also be assumed that the MIA shift *dhy* > *jh* occurred earlier than *ndh* > *nnh*, provided we can rely on this single example.

Obviously, for these four listed phonetic shifts in the historical phonetics of Kullui, two variants of relative chronology are possible:

- 1) *dhy* > *jh*; *mb*, *ndh* > *mm*, *nnh*; *NT* > *ND*; *mr* > *mb*
- 2) *dhy* > *jh*; *mb*, *ndh* > *mm*, *nnh*; *mr* > *mb*; *NT* > *ND*

7. Middle Indo-Aryan metathesis of *r*

Another example of a MIA shift not fully reflected in written sources is the so-called metathesis of *r* (Turner 1985c: 271, 1985d; Hendriksen 1986: 66). It is described as the disappearance of the consonant *r* in a non-initial consonant cluster and its reappearance after the initial consonant or consonant cluster. This phenomenon is examined in detail in a paper by G. Morgenstierne (1973, reprinted from 1947). Its first examples are noted already in Ashokan inscriptions: *pruva* < *pūrovaḥ*, *grabhagara* < *garbhāgāram*, *dhramma* < *dharmah*, *draśana* < *darśanam* (Turner 1985c: 271). However, in considering this phonetic change, it is important to track at least two possible interpretations by researchers. The first one is associated with the works of J. Bloch (Bloch 1965: 89; first published in French in 1934) and Turner (1985d: 216, first published in 1924), suggesting that it is not metathesis in the proper sense. If *r* simply moved forward, the adjacent consonant (often a stop) would end up in an intervocalic position and would subsequently be subject to loss, which never happens in cases of this “metathesis”. The corresponding consonants are later preserved just as they would have been preserved over the normal course of Indo-Aryan phonetic development. Both authors assume a process divided in several stages, beginning with a regressive *r*-replication (“anticipation”, in their own terms), after which the second cluster assimilates according to the general rule: OIA *dīrgha-* > **drirgha* > MIA *driggha* > Sindhi *ḍrighō* (instead of **ḍrihō*). However, the actual stage of *r*-replication remains unrecorded in monuments, although Bloch cites two examples with doubled *r* from Kati: *drgr* and *trūtr* (from OIA *tantra*).

At the same time, Morgenstierne (1973: 239) and Turner, in a slightly later work (Turner 1985c: 271), propose a different interpretation of this phenomenon. In his paper, Turner interprets the chain of shifts precisely as a metathesis after the gemination of the adjacent consonant: *dīrghaḥ* > *dīrghgaḥ* > **drīgha-*. He restricts this innovation to the Northwestern area – Sindhi, Lahnda, Dardic, and Western Pahari (Himachali). G. Morgenstierne, whose work mainly concerns Dardic languages, divides the metathesis into two types: type I refers to the transfer

⁴³ N. *thum* ‘ridge of a mountain, space round a hilltop’; H. *thūbā* m. ‘lump of earth’; G. *thumṛū* n. ‘a collection of ears of corn’.

of *r* to the left from combinations of the type *rC*, while type II refers to *Cr* clusters. According to his assumptions, type I is earlier than type II. At the same time, he assumes that there was no stage of *r*-replication; actual metathesis did occur, and the further behavior of the second cluster is explained by the fact that consonants after *r* underwent gemination in the MIA period, as indicated in the *Prāṭisākhya*s. Thus, his interpretation also assumes the chain OIA *ḍir-gha* > MIA *dirggha*- > MIA *driggha*. Triconsonantal sequences including *r* in the initial or final position lost this consonant according to common MIA rules, therefore such geminate clusters with *r* subsequently developed like ordinary geminates (Masica 1991: 179):

Sanskrit	Pali/Prakrit
<i>loptrā</i> - ‘loot’	<i>lotta-/lutta-</i> (Pkt)
<i>utkramati</i> ‘ascends’	<i>ukkamati</i>
<i>udgrahaṇa</i> - ‘taking out’	<i>uggahaṇa-</i>
<i>ardra</i> - ‘wet’	<i>adda-</i>
<i>vardhra</i> - ‘thong’	<i>vaddha-</i>
<i>rāṣṭra</i> - ‘country’	<i>raṭṭha-</i>
<i>vastra</i> - ‘clothes’	<i>vattha-</i>
<i>pārśva</i> - ‘side’	<i>passa-</i>
<i>vartman</i> - ‘wheel track’	<i>vaṭṭa-</i> (Pkt)
<i>kartya</i> - ‘to be cut off’	<i>kacca-</i> (Pkt)

The advantage of this interpretation of *r*-metathesis is that we do not have to assume a special stage of *r*-replication unrecorded in written monuments; its disadvantage is that it does not explain the supposedly later type II. According to Morgenstierne, type I is regular for Dardic languages while type II is not.

This shift is recorded in the Northwestern languages of the region as an irregular phenomenon, and the list of such languages is not limited to Dardic and Nuristani, which Morgenstierne focuses on. It includes Sindhi, Lahnda, Himachali and Gujarati; Turner observes similar reflexes in Romani. Historically, this change may have been more widespread, but in modern languages its traces can only be observed in languages that have preserved OIA *Cr*-type clusters.

Taking into account the NIA shift *tr*, *dr* > *tʃ*, *dʒ*, this phenomenon is represented in Kullui by the following words:

Kullui	OIA
<i>brag^h</i> ‘leopard’	<i>vyāghra</i> - ‘tiger’
<i>prat</i> ‘dish’	<i>pātra</i> - ‘dish’
<i>tʃambra</i> ‘cauldron’	<i>tāmra</i> - ‘copper’
<i>tʃokṇa</i> ‘rot’	<i>takra</i> - ‘bitter’
<i>dʒugli</i> ‘cave’	<i>durga</i> - ‘impassable place’
<i>dʒuba</i> ‘sacred durva grass’	<i>dūrvā</i> - ‘sacred durva grass’
<i>prɔʃa</i> ‘rib’	<i>parśu</i> - ‘rib’

Let us consider both interpretations (actual metathesis and *r*-replication) based on Kullui data.

The version of Morgenstierne (and the later version of Turner) explains why none of the Kullui words that underwent metathesis include traces of the second cluster with *r*, even though many such clusters were preserved in Kullui:

Kullui	OIA
<i>brag^h</i> ‘leopard’ (instead of <i>**brag^hr</i>)	<i>vyāghra-</i> ‘tiger’
<i>prat</i> ‘dish’ (instead of <i>**pratʃ</i>)	<i>pātra-</i> ‘dish’
<i>tʃəkṇa</i> ‘rot’ (instead of <i>**tʃəkṇa</i>)	<i>takra-</i> ‘bitter’
<i>prɔʃa</i> ‘rib’ (instead of <i>**prɔʃa</i>)	<i>parśu-</i> ‘rib’

However, this version relies on the assumption that in “Kullui Prakrit”, triconsonantal OIA clusters with *r* developed with the loss of *r*, similar to ordinary geminates, as demonstrated by general MIA development. Meanwhile, Kullui has a number of examples contradicting this presumption, showing that *r* in combinations with MIA geminates developed in the same way as they did in combinations with single consonants:

Kullui	OIA
<i>tʃ^hɔtʃi</i> ‘mushroom’	<i>chattrikā-</i> ‘umbrella’
<i>pɔʃa</i> ‘leaf’	<i>pātra-</i> ‘leaf’
<i>dʒotʃiṇa</i> ‘yoke’	<i>yoktrayati</i> ‘ties’

The version of Bloch and the early version of Turner, without resorting to the assumption of consonant gemination after or before *r*, could explain why the medial consonants did not undergo weakening, as intervocalic ones would have:

Kullui	OIA
<i>brag^h</i> ‘leopard’ (instead of <i>**brah</i>)	<i>vyāghra-</i> ‘tiger’
<i>prat</i> ‘dish’ (instead of <i>**pra</i>)	<i>pātra-</i> ‘dish’
<i>tʃambɾa</i> ‘cauldron’ (instead of <i>**tʃaūṛa</i>)	<i>tāmra-</i> ‘copper’
<i>tʃəkṇa</i> ‘rot’ (instead of <i>**tʃṇa</i>)	<i>takra-</i> ‘bitter’
<i>dʒugli</i> ‘cave’ (instead of <i>**dʒuli</i>)	<i>durga-</i> ‘impassable place’
<i>dʒuba</i> ‘sacred durva grass’ (instead of <i>**dʒuwa</i>)	<i>dūrvā-</i> ‘sacred durva grass’

However, it does not explain why there are no cases of *r*-replication in which the medial cluster with *r* managed to survive, instead of developing similarly to ordinary MIA geminates. Based on the limited number of examples, one could explain this by the prohibition of affricates and clusters at the absolute end of words in Kullui, similar to Old Indo-Aryan. However, numerous examples listed above (*āḍʒ* ‘guts’; *dʒatʃ* ‘festival, fair’; *t^hatʃ* ‘sheep pasture in the mountains’; *kutʃ^h* ‘menstrual blood, belly of a pregnant woman’; *bɔʃ* ‘year’, etc.) contradict this assumption. Consequently, we can postulate for the ancestor language of Kullui a rule concerning the period when most clusters had already undergone MIA assimilation:

- within one word, two clusters of non-homorganic consonants are impossible; in such a case, the second cluster undergoes complete assimilation according to common MIA rules.

Such a rule, in addition to the cases of “metathesis”, also explains the phonetic shape of the descendant of OIA *kṣetra* ‘field’ (let us recall that the sequence *kṣ* was apparently preserved until the late NIA period, when the old palatal affricates, together with the palatal affricates in Persian borrowings, shifted to dental affricates; Krylova 2024: 211):

Kullui	OIA
<i>brag^h</i> ‘leopard’ (instead of <i>**brag^hr</i>)	<i>vyāghra-</i> ‘tiger’
<i>prat</i> ‘dish’ (instead of <i>**pratf</i>)	<i>pātra-</i> ‘dish’
<i>tʃəkṇa</i> ‘rot’ (instead of <i>**tʃəkṇa</i>)	<i>takra-</i> ‘bitter’
<i>prɔʃa</i> ‘rib’ (instead of <i>**prɔʃa</i>)	<i>parśu-</i> ‘rib’
<i>tʃ^het</i> ‘field’ (instead of <i>**tʃ^hetf</i>)	<i>kṣetra-</i> ‘field’

Thus, based on Kullui data, the interpretation of this phenomenon in terms of *r*-replication is still more likely. In this case, the shift implies two stages, the first of which (the appearance of *r* after the initial consonant) preceded the shift *mr* > *mb*, as demonstrated by the example *tʃambṛa* ‘cauldron’ < *tāmra-* ‘copper’. In the case of type I metathesis according to Morgenstierne (*dʒugli* ‘cave’ < *durga-* ‘impassable place’, *dʒuba* ‘sacred durva grass’ < *dūrvā-* ‘sacred durva grass’, *prɔʃa* ‘rib’ < *parśu-* ‘rib’) the first stage also occurred before the MIA shift of clusters with initial *r* to geminates. At the same time, based on Kullui data, no conclusions can be drawn about a fundamental difference between types I and II – both *r* preceding and following the medial consonant underwent metathesis irregularly (cf., *tɔkḷi* ‘spindle’ < *tarku-* ‘spindle’ and *datʃi* ‘sickle’ < *dātrī-* ‘knife, sickle’, not subjected to metathesis in “Kullui Prakrit” but subjected to it in other languages). Therefore, it is quite possible that these types of changes occurred simultaneously.

The matter is more complicated with the second stage. If the shift *mr* > *mb* and the shifts of *rC* clusters to geminates took place in the MIA period, then the clusters *g^hr*, *kr*, *tr* (and possibly also *rś*, see above) in Himachali and a number of other languages were preserved even in the NIA period; however, in available examples, all non-initial consonant clusters underwent assimilation.

To explain this point, I introduce a rule of assimilation for the second of two non-homorganic consonant clusters in a word. This rule obviously took place before the shifts of *kʃ*, *tr*, *dr* to palatal affricates, but it is not entirely clear whether it should be dated to the MIA period (and then the clusters turned into geminates, which later underwent further simplification) or in the NIA period (and then the clusters turned into single consonants, but according to a rule coinciding with the common MIA rules of cluster assimilation in other languages). Given the scarce likelihood of such a coincidence, as well as the fact that the stage with doubled *r* is not recorded in MIA texts, the first option seems to be more credible.

8. Conclusion

Analysis of the Kullui material allows us to reach the following conclusions with varying degrees of confidence:

1. The version of the shift *mr* > *mb* via an intermediate stage of *mbr* is typologically plausible but requires the assumption that the cluster *mbr* in the ancestor language of Kullui, unlike biconsonantal clusters with labials (represented in Kullui by cases of *b^hr*, *pr*), regularly underwent assimilation of *r*.
2. Bloch’s version of *r*-replication / metathesis of *r* is more plausible for Kullui data than Morgenstierne’s version, since *r* in three-element clusters of the type *CTr* in the ancestor language of Kullui did not undergo assimilation, but it requires us to assume assimilation of the second of two non-homorganic consonant clusters.
3. The *ND* > *NN* shift in the ancestor language of Kullui affected only the combinations *ndh*, *mb*, with the latter not being entirely regular.

4. For a series of cluster transformations (with a correction for the irregularity of $mb > mm$) the following variants of relative chronology are probable:
 - a) $dhy > jh$; $mb, ndh > mm, nnh$; $NT > ND$; $mr > mb$
 - b) $dhy > jh$; $mb, ndh > mm, nnh$; $mr > mb$; $NT > ND$
5. The first stage of r -metathesis / r -replication preceded the shift $mr > mb$, and the second stage most likely occurred in the MIA period parallel to the assimilation of other r -clusters.
6. It is interesting to compare the noted cluster changes with the MIA hierarchy of consonant dominance described in Oberlies 2001: 99: non-palatal occlusives (excluding palatalization of dentals before y) $>$ nasals $>$ palatals $>$ sibilants $>$ $l > v > y > r$. Although most clusters in Kullui Prakrit were modified according to MIA rules, the assimilation of stops by nasals in the corresponding clusters, as well as the change $kṣ > tʃ^h$, clearly contradicts the hierarchy of Pali. Also, retention of r in a number of clusters indicates its significant power of resistance, whereas the above hierarchy would place r in the last place.

References

- Hendriksen, Hans. 1986. *Himachali studies III. Grammar*. København: Det Kongelige Danske Videnskabernes Selskab.
- Kogan, A. I. 2019. *Problemy izuchenija jazyka kashmiri v sravnitel'no-istoricheskom aspekte (Problems in the Study of the Kashmiri Language from a Comparative-Historical Perspective)*. Doctoral Dissertation: Institute of Oriental Studies, Moscow.
- Krylova, Anastasia. 2024. Accent in the Indo-Aryan language Kullui and the relative chronology of phonetic changes. *Journal of Language Relationship* 22(3-4): 191–215.
- Masica, Colin P. 1991. *The Indo-Aryan languages*. Cambridge University Press.
- Morgenstierne, Georg. Metathesis of liquids in Dardic. 1973 (1947). In: Georg Morgenstierne: *Irano-Dardica*: 231–240. Wiesbaden: Reichert.
- Oberlies, Thomas. 2001. *Pāli – A grammar of the language of the Theravāda Tipiṭaka. With a concordance to Pischel's Grammatik der Prakrit-Sprachen [Indian philology and South Asian studies 3]*. Berlin: Mouton de Gruyter.
- Southworth, Franklin C. 2005. *Linguistic archaeology of South Asia*. London: Routledge.
- Turner, Ralph L. 1966. *A comparative dictionary of Indo-Aryan languages*. Oxford University Press.
- Turner, Ralph L. 1985a. Romani *traṣ* 'to frighten'. In: Ralph L. Turner. *Indo-Aryan Linguistics. Collected Papers, 1912–1973*: 181–185. Delhi: Disha Publications.
- Turner, Ralph L. 1985b. Anticipation of normal sound-changes in Indo-Aryan. In: Ralph L. Turner. *Indo-Aryan Linguistics. Collected Papers, 1912–1973*: 357–367. Delhi: Disha Publications.
- Turner, Ralph L. 1985c. The Position of Romani in Indo-Aryan. In: Ralph L. Turner. *Indo-Aryan Linguistics. Collected Papers, 1912–1973*: 251–290. Delhi: Disha Publications.
- Turner, Ralph L. 1985d. Cerebralization in Sindhi. In: Ralph L. Turner. *Indo-Aryan Linguistics. Collected Papers, 1912–1973*: 206–227. Delhi: Disha Publications.

А. С. Крылова. Относительная хронология ряда фонетических изменений древнеиндийских кластеров на материале языка куллуи

В данном исследовании устанавливается относительная хронология фонетических изменений древнеиндийских кластеров в языке куллуи (химачали). Анализ подтверждает сохранение в среднеиндийский период ряда сочетаний с r , а также кластера $kṣ$, причем последний переходит в $tʃ^h$ независимо от южных индоарийских языков. Рассматривается относительная последовательность изменений, в том числе $-mr- > -mb-$, $NT > ND$ и $ND > NN$. Теория Блока о метатезе r как результате репликации (антисипации)

подтверждается в противовес теории Моргенштерне. При этом требуется постулировать правило ассимиляции второго негоморганного кластера в слове. На основе данных по языку куллуи предлагаются две возможные хронологические последовательности этих сдвигов.

Ключевые слова: куллуи язык; индоарийские языки; относительная хронология; метатеза.

Оценка достоверности исходных данных при лексикостатистических расчётах

В статье рассматривается проблема надёжности исходных данных, используемых в лексикостатистике для получения генеалогических классификаций и лингвистических датировок. Такие данные, представленные таблицей совпадений между основными списками языков, могут содержать ошибки или неточности, которые обусловлены сложностью и субъективностью процедуры этимологического анализа и существенно снижают достоверность лексикостатистических расчётов. Для решения этой проблемы предложена формальная методика, основанная на критерии взаимной согласованности (транзитивности) долей совпадений в лексикостатистической таблице. Использование этого критерия позволяет выявить недостоверные значения в исходных данных, а также численно оценить величину ошибки в каждом случае. Преимуществами предложенного подхода являются его простота и универсальность, объективность результатов, а также удобство реализации в виде компьютерного приложения. Апробация методики на материале романских и тюркских языков подтверждает её применимость и практическую эффективность при рассмотрении как небольших, так и крупных языковых групп.

Ключевые слова: лексикостатистика; глоттохронология; матрица расстояний; критерий согласованности.

Одним из главных факторов, влияющих на точность и достоверность генеалогических классификаций, а также лингвистических датировок, полученных с помощью лексикостатистики, является качество исходных данных. Первичными данными для большинства лексикостатистических исследований служат доли совпадений, которые рассчитываются путём сопоставления основных списков (ОС) рассматриваемых языков и установления количества этимологических соответствий (когнатов) между ними. Найденные значения долей совпадений, представленные в табличном виде, называют исходной лексикостатистической таблицей (или матрицей)¹.

Наиболее критичными и ответственными при сборе данных являются подготовительные этапы, а именно: составление списков базисной лексики и проведение сравнительного этимологического анализа. Поскольку обе эти процедуры плохо поддаются формализации, в большинстве случаев они проводятся вручную, что неизбежно приводит к субъективному («экспертному») характеру решений, как при отборе лексики, так и при определении этимологии. Как неоднократно было показано на разнообразном языковом материале, именно от качества собранных списков и их этимологического анализа в первую очередь зависит как точность лексикостатистических расчётов, так и в целом адекватность полученных результатов².

¹ По сути данная таблица является аналогом матрицы расстояний, используемой для кластеризации и анализа данных во многих научных областях (математической статистике, информатике, экономике, социологии, биологии и т. д.).

² В частности, работы Starostin 2000: 228; Милитарёв 2000; Бурлак & Старостин 2005: 148; Vydrin 2009: 112–114; дискуссия Соловьёв, Дыбо & Старостин 2010: 194; Starostin 2010: 84 убедительно демонстрируют,

Проблема неточности экспертных оценок и возможные пути её решения (или смягчения) широко обсуждаются в работах по компаративистике³, однако большинство авторов сходятся во мнении, что процесс составления и этимологизации основных списков не может быть полностью формализован и в любом случае подразумевает некоторую субъективность. В связи с этим наиболее очевидным способом нивелировать возникающие ошибки представляется тщательная и всесторонняя проверка полученных данных — как самим исследователем, так и с привлечением других лингвистов.

Данный содержательный подход, при всех неоспоримых достоинствах, обладает также рядом очевидных недостатков. Во-первых, он не гарантирует выявление всех ошибок в исходных данных, а лишь повышает надёжность экспертных оценок за счёт увеличения их количества⁴. Во-вторых, результат проверки по-прежнему остаётся субъективным и в значительной степени зависит от возможностей компаративиста и/или его владения анализируемым материалом. В-третьих, ручное сравнение основных списков позволяет определить количество когнатов в базисной лексике каждой пары языков, однако не обеспечивает согласованность⁵ полученных этимологических оценок. Наконец, в-четвёртых, процесс верификации списков и этимологий по своей трудоёмкости часто не уступает первоначальному сбору и анализу данных.

Перечисленные недостатки существенно ограничивают применимость такого подхода, а также значительно снижают его эффективность во многих случаях. При этом большинство указанных ограничений обусловлено качественным и экспертным характером процедуры верификации.

Возможной альтернативой этой процедуре могла бы стать некая формальная методика, позволяющая быстро и объективно оценить достоверность исходных данных, а также по возможности выявить заведомо недостоверные значения. Такую методику можно предложить на основе совокупного количественного анализа долей совпадений в исходной лексикостатистической таблице.

Представим два языка А и В, которые в равной степени близкородственны с некоторым третьим языком С. Очевидно, что в этом случае А и В не могут оказаться очень далёкими родственниками, так как это будет противоречить их близкому родству с идиомом С. Из этого следует, что доли совпадений, полученные в ходе попарного сравнения ОС этих языков, обладают взаимозависимостью, а значит — в какой-то степени определяют друг друга. Учитывая это обстоятельство, естественно предположить, что мы можем оценить достоверность долей совпадений для любой выбранной пары языков, опираясь на доли совпадений этих идиомов с остальными языками. При этом критерием достоверности будет согласованность (или непротиворечивость) этой оценки по отношению к остальным значениям лексикостатистической таблицы.

Поясним смысл согласованности оценок для группы из трёх языков А, В и С, доли совпадений между ОС которых представлены в табл. 1.

как систематические экспертные ошибки, допущенные на этом этапе, не только существенно сказываются на точности топологии генетических деревьев или глоттохронологических датировках, но и могут приводить к совершенно абсурдным результатам.

³ См., например, Старостин 2007: 416–417; Militarev 2005: 345; Starostin 2010: 79–116.

⁴ В скобках отметим, что и это не всегда осуществимо. Например, в случае с малоизученными языками (для которых применение лексикостатистики особенно актуально) верификация собранного материала другими специалистами зачастую невозможна из-за их отсутствия.

⁵ Смысл согласованности будет подробно обсуждаться ниже в тексте статьи.

Согласно таблице, доля общей лексики в списках языков А и В составляет 98%, что соответствует разнице («расстоянию») в $100-98=2$ слова⁶ между ними. Аналогично для языков В и С доля совпадений равна 96% (расстояние в 4 слова), а для языков А и С — 97% (расстояние в 3 слова). Для наглядности обозначим доли совпадений и расстояния между языками на диаграмме (рис. 1).

Язык	Доли совпадений		
	А	В	С
А	1	0,98	0,97
В	0,98	1	0,96
С	0,97	0,96	1

Таблица 1. Доли совпадений ОС языков А, В и С

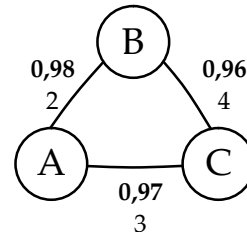


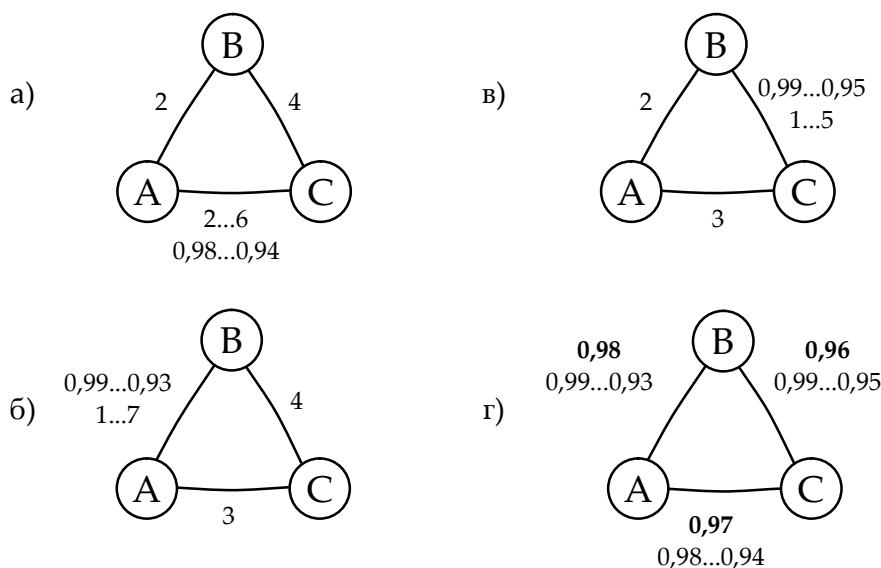
Рис. 1. Доли совпадений и расстояния между языками А, В и С

Если мы предположим, что 2 значения, различающие языки А и В, не совпадают с 4 значениями, отличающимися в языках В и С, то количество несовпадающих слов (расстояние) между языками А и С будет равно сумме расстояний А–В и В–С: $2+4=6$.

Если же допустить, что оба эти значения входят в число четырёх различающихся значений идиомов В и С, то основные списки языков А и С будут отличаться всего 2 словами, что соответствует разнице тех же расстояний: $4-2=2$.

Следовательно, расстояние между языками А и С может составлять от 2 до 6 слов, а доля совпадений между ними, соответственно, должна находиться в диапазоне $0,98...0,94$ (рис. 2а). Очевидно, что значение $N_{AC}=0,97$, заданное в табл. 1, входит в этот интервал, а значит — согласуется с известными долями совпадений каждого из языков А и С с языком В (N_{AB} и N_{BC}).

Аналогичные рассуждения и расчёты, проведённые для остальных пар языков (АВ и ВС — рис. 2б, в), показывают, что значения N для любой пары идиомов согласованы относительно долей совпадений с оставшимся третьим языком.

Рис. 2. Проверка взаимной согласованности значений N_{AB} , N_{BC} , N_{AC}

⁶ При использовании полных 100-словных списков.

Сопоставляя найденные допустимые диапазоны и известные значения N_{AB} , N_{BC} , N_{AC} (рис. 2г), нетрудно убедиться, что все доли совпадения для триады языков А, В и С являются *взаимно согласованными*.

Используя тот же пример, смоделируем теперь ситуацию несогласованности исходных данных. Для этого изменим долю совпадений для языков В и С в исходной матрице (табл. 1) на 0,94. Полученные значения представлены в табл. 2 и на рис. 3.

Язык	Доли совпадений		
	А	В	С
А	1	0,98	0,97
В	0,98	1	0,94
С	0,97	0,94	1

Таблица 2. Доли совпадений ОС языков А, В и С (несогласованные значения)

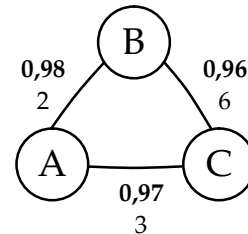


Рис. 3. Доли совпадений и расстояния между языками А, В и С (несогласованные значения)

В данном случае расстояние между языками А и С может принимать значения от 8 (6+2) до 4 (6–2) слов, а соответствующая доля значений N_{AC} должна находиться в диапазоне 0,92...0,96 (рис. 4а). При этом табличное значение $N_{AC}=0,97$ выходит за пределы этого интервала и не согласуется с двумя остальными долями совпадений. Более того, анализируя другие пары языков (рис. 4б, в), мы видим, что теперь ни одно значение N_{AB} , N_{BC} и N_{AC} не укладывается в диапазоны согласованности (рис. 4г). Таким образом, изменение всего одного значения N_{BC} в исходной таблице привело к нарушению согласованности одновременно для всех пар внутри рассматриваемой триады языков.

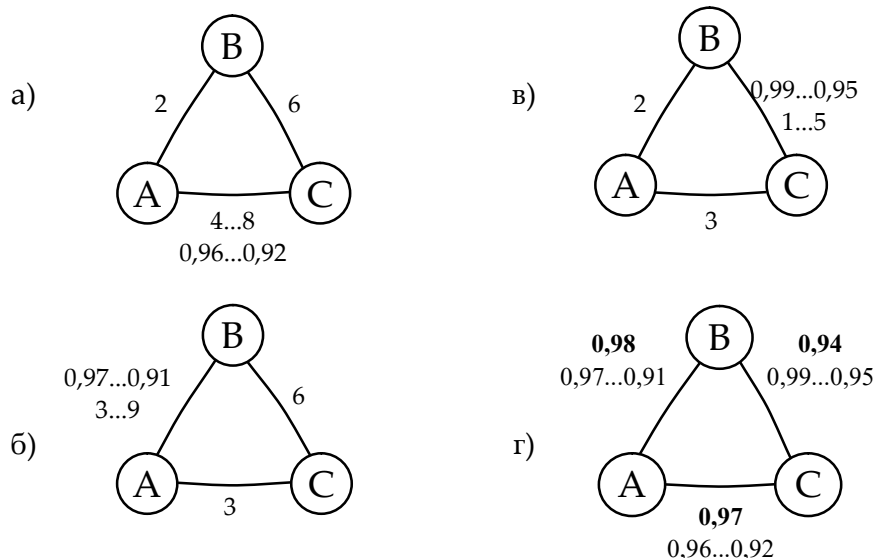


Рис. 4. Иллюстрация несогласованности значений N_{AB} , N_{BC} , N_{AC}

Обратим внимание, что при этом ошибка согласованности (E_c) для любой пары языков одинакова и составляет одно слово (см. рис. 4г). Так, для пары языков АВ величина ошибки равна разнице между их долей совпадений ($N_{AB}=0,98$) и ближайшим значением из допустимого диапазона (0,97...0,91): $E_{AB}=|0,98-0,97|=0,01^7$.

⁷ Аналогично для пары ВС получаем $E_{BC}=|0,94-0,95|=0,01$, а для пары АС – $E_{AC}=|0,97-0,96|=0,01$.

Из рассмотренных примеров можно сделать вывод, что для проверки согласованности (или наоборот — выявления несогласованности) долей совпадений между тремя языками достаточно проверить согласованность любой пары из этих языков относительно третьего. Иначе говоря, согласованность значений N для любых двух идиомов в триаде языков является необходимым и достаточным условием согласованности всей рассматриваемой триады⁸.

Как было показано на примерах выше, это условие выполняется в случае, когда доли совпадений для любых родственных языков A , B и C удовлетворяют двум следующим критериям:

- 1) доля несовпадающих значений (расстояние) между языками A и C не может быть больше суммы расстояний между языками $A-B$ и $B-C$:

$$1 - N_{AC} \leq (1 - N_{AB}) + (1 - N_{BC})$$

или, после упрощения,

$$N_{AC} \geq N_{AB} + N_{BC} - 1; \quad (1)$$

- 2) расстояние между идиомами A и C не может быть меньше разницы расстояний между идиомами $A-B$ и $B-C$, взятой по модулю:

$$\begin{aligned} 1 - N_{AC} &\geq |(1 - N_{AB}) - (1 - N_{BC})|, \\ N_{AC} &\leq 1 - |N_{AB} - N_{BC}|. \end{aligned} \quad (2)$$

Объединив неравенства (1) и (2), получим общее выражение условия согласованности для языков A , B и C :

$$N_{AB} + N_{BC} - 1 \leq N_{AC} \leq 1 - |N_{AB} - N_{BC}|. \quad (3)$$

Данное условие, вслед за С. А. Бурлак и С. А. Старостиным, можно также назвать *критерием транзитивности*⁹.

Проверим справедливость критерия транзитивности в его формальном выражении на уже известных примерах. Для этого поочерёдно подставим в полученное неравенство (3) сначала согласованные (табл. 1), а затем — несогласованные данные (табл. 2):

$$\begin{aligned} 0,98 + 0,96 - 1 &= 0,94 \leq 0,97 \leq 1 - |0,98 - 0,96| = 0,98 \text{ (по данным табл. 1),} \\ 0,98 + 0,94 - 1 &= 0,92 \leq 0,97 \leq 1 - |0,98 - 0,94| = 0,96 \text{ (по данным табл. 2).} \end{aligned}$$

Очевидно, что при подстановке согласованных долей совпадений (табл. 1) мы получаем верное неравенство ($0,94 \leq 0,97 \leq 0,98$) — т. е. критерий транзитивности соблюдается. Тогда как при расчёте по несогласованным значениям (табл. 2) правая часть неравенства ($0,92 \leq 0,97 \leq 0,96$) не выполняется и, следовательно, исходные данные не удовлетворяют условию согласованности.

Перейдём к обсуждению применения критерия транзитивности для анализа реальных лексикостатистических данных. Очевидно, что при рассмотрении группы языков, число которых превышает три, мы должны проанализировать на согласованность все возможные триады идиомов в этой группе. Например, в группе, состоящей из 4-х языков A , B , C и D , нам потребуется проверить всего 4 триады: $A-B-C$, $A-B-D$, $A-C-D$ и $B-C-D$.

⁸ Данное утверждение можно доказать аналитически для общего случая. См. полное доказательство в Приложении 1.

⁹ Это название, хорошо отражающее смысл условия согласованности, предложено в работе Бурлак & Старостин 2005: 103, в которой аналогичный ход рассуждений применяется для выявления скрытых заимствований в основных списках сравниваемых языков.

В общем случае, для группы из n языков, число таких триад определяется как количество сочетаний из n по три:

$$C_n^3 = \frac{n!}{(n-3)!3!},$$

например, для 5 языков ($n=5$) получим уже 10 триад, а для 9 языков ($n=9$) — 84 триады:

$$C_5^3 = \frac{5!}{(5-3)!3!} = 10,$$

$$C_9^3 = \frac{9!}{(9-3)!3!} = 84.$$

Используя критерий транзитивности, описанный выше, проанализируем согласованность лексикостатистических данных для небольшой группы романских языков (табл. 3).

№	Название языка	Доли совпадений основных список языков								
		LAT	ITA	FRA	PRT	ESP	GAL	VAL	PRV	ROM
LAT	классическая латынь	1	0,88	0,83	0,87	0,85	0,89	0,87	0,83	0,82
ITA	итальянский	0,88	1	0,93	0,90	0,89	0,92	0,93	0,90	0,88
FRA	французский	0,83	0,93	1	0,83	0,85	0,85	0,90	0,97	0,86
PRT	португальский	0,87	0,9	0,83	1	0,96	0,97	0,95	0,86	0,84
ESP	испанский	0,85	0,89	0,85	0,96	1	0,95	0,96	0,87	0,84
GAL	галисийский	0,89	0,92	0,85	0,97	0,95	1	0,95	0,87	0,87
VAL	вален. наречие каталанского	0,87	0,93	0,90	0,95	0,96	0,95	1	0,93	0,87
PRV	провансальский	0,83	0,90	0,97	0,86	0,87	0,87	0,93	1	0,84
ROM	румынский	0,82	0,88	0,86	0,84	0,84	0,87	0,87	0,84	1

Таблица 3. Доли совпадений основных списков романских языков¹⁰

В результате анализа были выявлены всего 5 триад с несогласованными значениями долей совпадений. Список этих триад, а также соответствующие им ошибки несогласованности представлены в табл. 4. Среди них 3 триады с ошибкой в два слова и 2 триады с ошибкой в одно слово.

Обозначения языков			Доли совпадений			Ошибка согласованности E_c
A	B	C	N_{AB}	N_{BC}	N_{AC}	
ESP	VAL	PRV	0,96	0,93	0,87	0,02
FRA	PRT	VAL	0,83	0,95	0,90	0,02
PRT	VAL	PRV	0,95	0,93	0,86	0,02
FRA	ESP	VAL	0,85	0,96	0,90	0,01
GAL	VAL	PRV	0,95	0,93	0,87	0,01

Таблица 4. Результаты анализа согласованности исходных данных для романских языков (указаны только несогласованные триады)

¹⁰ Доли совпадений приводятся по данным из архива проекта «Глобальная лексикостатистическая база данных» (starlingdb.org/new100/main.htm).

С учётом того, что доля несогласованных триад составляет всего 6% от их общего количества (5 из 84), а величина ошибки не превышает 2 слова, исходную лексикостатистическую матрицу романских языков (табл. 3) можно считать хорошо согласованной. Тем не менее, попробуем установить причину несогласованности в указанных пяти триадах, опираясь на результаты проверки, приведённые в табл. 4.

В силу того, что взаимная согласованность всех долей совпадений в триаде языков нарушается при изменении любого из трёх значений, мы не можем установить, какое значение является недостоверным в отдельно взятой триаде. Однако, если проанализировать частотность вхождения отдельных языков (или пар языков) в несогласованные триады, можно косвенно определить, в каких ОС с наибольшей вероятностью содержатся недостоверные этимологические оценки.

В частности, при рассмотрении списка триад в табл. 4 мы обнаруживаем, что во все несогласованные триады входит валенсийское наречие каталанского (VAL), а вторым по частотности является провансальский (PRV), который встречается трижды. Это означает, что несогласованность триад, вероятнее всего, обусловлена недостоверными долями совпадений этих языков между собой (в триадах, где они встречаются вместе) или с другими идиомами (в остальных случаях). Следовательно, для устранения несогласованности исходных данных нам следует в первую очередь искать неточности, допущенные при сравнительном анализе валенсийского и провансальского основных списков¹¹. При этом значение ошибки согласованности (в правом столбце табл. 4) показывает, на сколько слов (или процентов) мы должны скорректировать доли совпадений между идиомами, чтобы они стали согласованными.

Рассмотрим теперь более сложную ситуацию на примере группы тюркских языков, включающей 31 идиом. Проверка исходной таблицы долей совпадений (см. табл. 8 в Приложении 2), показывает, что согласованность данной группы заметно ниже, чем в случае с романскими языками¹². Об этом свидетельствует как больший процент несогласованных триад (10% от общего количества), так и — что более существенно — повышение абсолютной величины ошибок согласованности.

Как видно из распределения в табл. 5, большинство несогласованных триад (343 из 467) имеют минимальные ошибки в 1–2 слова. Однако мы также фиксируем большое число ошибок в 3 слова (70 триад), 4 слова (34 триады), 5 слов (14 триад) и даже 6 слов (4 триады). При этом максимальное значение несогласованности достигает 7 и 8 слов! Такие значительные ошибки могут не только снизить точность лексикостатистических расчётов, но и существенно повлиять на их качественные результаты — например, при построении генеалогического дерева и определении принадлежности языка к той или иной языковой подгруппе¹³.

¹¹ Отметим, что анализ частотности позволяет установить именно наиболее вероятных «виновников» несогласованности, однако нельзя исключать, что в отдельных случаях причиной рассогласования могут быть доли совпадений идиомов, которые встречаются всего в одной или нескольких несогласованных триадах.

¹² Можно предположить, что в любой лексикостатистической таблице, составленной по реальным языковым данным, всегда присутствует некоторое количество несогласованных значений, и чем больше группа рассматриваемых языков, тем больше будет её несогласованность. Однако это предположение ошибочно: например, при проверке таблицы совпадений для 16 дардских языков все 560 триад оказались согласованными (Васильев & Коган 2013: 158–159).

¹³ В частности, хорошо известно, что метод «ближайших соседей» и его модификации, которые широко применяются для построения генетических классификаций, очень чувствительны к завышенным значениям долей совпадений. По этой причине даже единичные ошибочные значения в таблице исходных данных зачастую приводят к масштабным изменениям всей структуры дерева. См., например, Васильев 2010; Васильев & Коган 2013; Васильев & Саенко 2020.

Ошибка согласованности E_c слов	Количество триад с ошибкой E_c
8	1
7	1
6	4
5	14
4	34
3	70
2	122
1	221
Всего	467

Таблица 5. Распределение триад по величине ошибки согласованности (тюркские языки)

Чтобы установить наиболее вероятные источники несогласованности в исходных данных, проанализируем список несогласованных триад (табл. 9 в Приложении 3) и выделим языки, которые встречаются в них чаще всего. Для удобства представим полученные результаты в виде двух разных таблиц: для отдельных языков (табл. 6) и для пар языков (табл. 7).

Язык	Число вхождений	Язык	Число вхождений	Язык	Число вхождений	Язык	Число вхождений
UZB	188	AZB	47	ALT	33	SJG	22
KLP	109	TRM	47	TUB	29	CHV	21
QUM	79	BLK	46	SHR	29	UIG	20
NOG	78	GAG	40	SAL	29	TOF	19
BAS	78	KRX	36	KMD	27	HAK	18
TAT	73	XAL	35	TUV	26	KRM	18
KAZ	71	TRK	35	KRQ	22	JAK	15
KRG	54	KUU	35	ATU	22		

Таблица 6. Список тюркских языков, входящих в несогласованные триады, упорядоченный по числу вхождений

При рассмотрении табл. 6 видно, что абсолютным лидером индивидуального списка является узбекский язык (UZB), который входит почти в 40% всех несогласованных триад (188 из 467), в том числе в триады с наибольшими ошибками согласованности 8, 7 и 6 слов (см. табл. 9). Вторым по частотности стал каракалпакский (KLP), входящий в состав 109 триад (около четверти от общего количества). Эти данные указывают на необходимость проверки и уточнения долей совпадения узбекского и каракалпакского ОС с остальными списками.

Для выявления конкретных пар языков с недостоверными значениями N целесообразно использовать табл. 7. При этом в первую очередь следует обращать внимание на частотные пары языков, входящие в триады с большой ошибкой согласованности (3, 4, 5 и более слов). К таким парам, например, относятся: KLP–KAZ, UZB–AZB, UZB–KAZ, UZB–SAL, BAS–KLP, UZB–KUU, UZB–GAG¹⁴.

¹⁴ Показательно, что во все эти пары входят узбекский либо каракалпакский языки, чаще всего встречающиеся в несогласованных триадах (см. табл. 6). Также отметим, что высокая частотность пар KLP–KAZ и

Как уже говорилось выше, с помощью критерия транзитивности невозможно однозначно определить, какое из трёх значений N является причиной несогласованности в триаде языков. Более того, недостоверными могут оказаться сразу два или все три значения. Поэтому данные о количестве вхождений отдельных языков или пар идиомов в несогласованные триады являются не строгой, а вероятностной оценкой. Тем не менее, они позволяют сделать процесс поиска недостоверных значений в исходных данных целенаправленным и системным, что многократно упрощает и ускоряет работу исследователя.

Обозначения пары языков		N_{AB}	Число вхождений	Распределение общего числа вхождений по значению ошибки E_c							
A	B			0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08
KLP	KAZ	1.00	27	10	4	10	2	0	1	0	0
NOG	QUM	1.00	22	11	8	3	0	0	0	0	0
UZB	AZB	0.95	21	0	4	4	2	8	1	1	1
UZB	KAZ	0.99	20	4	4	4	5	1	1	0	1
BLK	QUM	0.98	20	10	7	1	2	0	0	0	0
UZB	SAL	0.95	19	3	8	2	5	1	0	0	0
UZB	QUM	0.98	18	6	7	5	0	0	0	0	0
BAS	KLP	0.98	18	4	5	3	4	2	0	0	0
UZB	KUU	0.93	17	4	7	3	2	1	0	0	0
TAT	BAS	0.99	16	11	3	2	0	0	0	0	0
UZB	TRM	0.97	16	5	6	4	1	0	0	0	0
NOG	KLP	0.99	16	9	4	2	1	0	0	0	0
UZB	BLK	0.95	16	6	6	2	1	1	0	0	0
UZB	GAG	0.90	16	4	6	1	4	1	0	0	0
UZB	ATU	0.95	15	6	4	4	1	0	0	0	0
UZB	SHR	0.91	15	3	5	5	1	0	1	0	0
UZB	KRX	0.96	15	2	8	3	2	0	0	0	0
...											

Таблица 7. Список пар тюркских языков, входящих в несогласованные триады, упорядоченный по числу вхождений (фрагмент)¹⁵

Рассмотренные выше примеры хорошо демонстрируют основные преимущества описанной методики, обусловленные её формальным характером. К ним относится простота и универсальность¹⁶ применения, объективность получаемых результатов, а также возможность не только качественной, но и количественной оценки достоверности исходных данных. При этом все необходимые вычисления могут быть легко реализованы в виде компьютерной программы и выполняться полностью автоматически, что открыва-

NOG-QUM объясняется полным совпадением их ОС ($N=1$, табл. 7). В этом случае критерий транзитивности будет выполняться только тогда, когда доли совпадений каждого из двух языков с любым третьим языком попарно равны.

¹⁵ Таблица с полным списком всех пар приводится в сопровождающих материалах, доступных на сайте журнала.

¹⁶ Методика может применяться к исходным лексикостатистическим данным, полученным любым способом, и не влияет на выбор моделей и методов, используемых в ходе дальнейшего исследования.

ет широкие возможности для её практического использования как отдельно, так и в составе уже существующих лексикостатистических приложений.

В то же время, вместе с достоинствами такого формального подхода следует также отметить его ограничения. Прежде всего, нужно учитывать, что согласованность долей совпадений не гарантирует их достоверности¹⁷. Использование критерия транзитивности позволяет нам выявить только те значения, которые привели к ошибкам согласованности исходных данных, однако это не означает, что в них не содержится других неточностей или ошибок.

Во-вторых, предлагаемая методика является по сути «диагностическим» инструментом, с помощью которого можно быстро оценить согласованность исходной таблицы и установить недостоверные значения, нуждающиеся в пересмотре. Однако выяснение причин несогласованности и собственно корректировка долей совпадений невозможны без детального этимологического анализа основных списков.

Таким образом, на практике наиболее эффективным и целесообразным представляется комбинированное применение формального и содержательного подходов, при котором сначала проводится автоматизированная проверка исходных данных на согласованность, а затем несогласованные доли совпадений, выявленные в ходе проверки, анализируются и корректируются вручную специалистом¹⁸. При необходимости эта процедура может быть повторена несколько раз до достижения полной согласованности исходных данных, либо уменьшения количества и величины ошибок согласованности до приемлемых значений.

Приложение 1

Доказательство необходимого и достаточного условия согласованности оценок долей совпадений ОС для трёх языков

Докажем, что необходимым и достаточным условием согласованности долей совпадений между ОС трёх языков является согласованность долей совпадений для любой одной пары языков (AB, BC или AC) по отношению к третьему языку (C, A или B).

1. Условие согласованности оценок N_{AB} , N_{BC} , N_{AC} для трёх языков A, B и C не выполняется по определению, если это условие не выполняется хотя бы для одной пары языков по отношению к соответствующему третьему языку.

Необходимое условие доказано.

2. Допустим, что условие согласованности выполняется для пары языков A и C:

$$N_{AB} + N_{BC} - 1 \leq N_{AC} \leq 1 - |N_{AB} - N_{BC}|. \quad (1)$$

Для компактного изложения рассуждений перепишем (1) в виде:

$$a + b - 1 \leq c \leq 1 - |a - b|,$$

где обозначения определяются почленным сопоставлением (1) и (2).

¹⁷ В строгом смысле согласованность долей совпадений является необходимым, но не достаточным условием их достоверности.

¹⁸ Либо соответствующие идиомы целиком исключаются из рассмотрения, если причину несогласованности установить не удалось.

Для случая $a \geq b$ (2) примет вид:

$$a + b - 1 \leq c \leq 1 - a + b,$$

от которого можно перейти к двум неравенствам:

$$a - c - 1 \leq -b; \quad (3)$$

$$-b \leq 1 - a - c. \quad (4)$$

Из (4) получаем

$$a + c - 1 \leq b, \quad (5)$$

и (3) и (4) можно переписать в виде:

$$a \leq 1 - b + c;$$

$$a \leq 1 + b - c,$$

который соответствует выражению

$$a \leq 1 - |b - c| \quad (6)$$

для произвольных b и c .

Раскрывая (2) для случая $a \leq b$, получим аналогично:

$$a + b - 1 \leq c \leq 1 - b + a, \quad (7)$$

$$b - c - 1 \leq -a;$$

$$-a \leq 1 - b - c.$$

и, далее,

$$b + c - 1 \leq a, \quad (8)$$

а также

$$b \leq 1 - a + c;$$

$$b \leq 1 + a - c,$$

что соответствует

$$b \leq 1 - |a - c| \quad (9)$$

для произвольных a и c .

Таким образом, при выполнении неравенства (2) для пары языков A и C :

$$a + b - 1 \leq c \leq 1 - |a - b|, \quad (2)$$

выполняются условия (5), (6), (8), (9):

$$b + c - 1 \leq a, \quad (8)$$

$$a \leq 1 - |b - c| \quad (6)$$

$$a + c - 1 \leq b, \quad (5)$$

$$b \leq 1 - |a - c|, \quad (9)$$

которые соответствуют искомым условиям согласованности оценок для пар языков A и B , B и C по отношению к третьему языку C и A :

$$N_{BC} + N_{AC} - 1 \leq N_{AB} \leq 1 - |N_{BC} - N_{AC}|;$$

$$N_{AB} + N_{AC} - 1 \leq N_{BC} \leq 1 - |N_{AB} - N_{AC}|.$$

Достаточное условие доказано.

Приложение 2

№	LANG	TRK	TAT	UZB	UIG	SJG	AZB	TRM	HAK	CHV	JAK	TUV	KRG	NOG	BAS
1	TRK	1	0,86	0,9	0,85	0,8	0,94	0,92	0,81	0,73	0,74	0,77	0,88	0,88	0,82
2	TAT	0,86	1	0,98	0,97	0,85	0,88	0,91	0,89	0,77	0,77	0,8	0,98	0,99	0,99
3	UZB	0,9	0,98	1	0,96	0,9	0,95	0,97	0,91	0,79	0,79	0,84	0,99	0,98	0,96
4	UIG	0,85	0,97	0,96	1	0,9	0,88	0,91	0,9	0,73	0,77	0,82	0,95	0,96	0,94
5	SJG	0,8	0,85	0,9	0,9	1	0,82	0,86	0,86	0,71	0,78	0,83	0,86	0,87	0,83
6	AZB	0,94	0,88	0,95	0,88	0,82	1	0,96	0,82	0,78	0,74	0,76	0,89	0,88	0,84
7	TRM	0,92	0,91	0,97	0,91	0,86	0,96	1	0,88	0,77	0,77	0,81	0,94	0,95	0,91
8	HAK	0,81	0,89	0,91	0,9	0,86	0,82	0,88	1	0,71	0,79	0,88	0,9	0,88	0,88
9	CHV	0,73	0,77	0,79	0,73	0,71	0,78	0,77	0,71	1	0,69	0,67	0,77	0,77	0,75
10	JAK	0,74	0,77	0,79	0,77	0,78	0,74	0,77	0,79	0,69	1	0,76	0,79	0,78	0,78
11	TUV	0,77	0,8	0,84	0,82	0,83	0,76	0,81	0,88	0,67	0,76	1	0,85	0,81	0,79
12	KRG	0,88	0,98	0,99	0,95	0,86	0,89	0,94	0,9	0,77	0,79	0,85	1	0,97	0,96
13	NOG	0,88	0,99	0,98	0,96	0,87	0,88	0,95	0,88	0,77	0,78	0,81	0,97	1	0,97
14	BAS	0,82	0,99	0,96	0,94	0,83	0,84	0,91	0,88	0,75	0,78	0,79	0,96	0,97	1
15	BLK	0,83	0,93	0,95	0,9	0,83	0,85	0,89	0,86	0,73	0,8	0,8	0,93	0,96	0,93
16	GAG	0,98	0,84	0,9	0,86	0,78	0,92	0,9	0,79	0,73	0,75	0,74	0,88	0,86	0,81
17	KRM	0,86	0,94	0,95	0,92	0,84	0,88	0,92	0,86	0,72	0,75	0,79	0,94	0,92	0,93
18	KLP	0,89	0,98	0,99	0,96	0,87	0,89	0,95	0,93	0,77	0,78	0,85	0,99	0,99	0,98
19	ATU	0,87	0,9	0,95	0,92	0,88	0,88	0,89	0,87	0,76	0,84	0,79	0,91	0,9	0,89
20	SAL	0,85	0,89	0,95	0,92	0,84	0,9	0,92	0,84	0,76	0,74	0,75	0,9	0,91	0,87
21	SHR	0,79	0,86	0,91	0,9	0,85	0,8	0,85	0,95	0,67	0,76	0,85	0,89	0,87	0,86
22	TOF	0,76	0,8	0,83	0,82	0,81	0,75	0,79	0,87	0,68	0,77	0,97	0,84	0,82	0,79
23	KAZ	0,86	0,97	0,99	0,95	0,85	0,86	0,94	0,9	0,75	0,75	0,79	0,99	0,98	0,95
24	QUM	0,87	0,96	0,98	0,95	0,86	0,91	0,93	0,88	0,79	0,77	0,79	0,94	1	0,97
25	KRX	0,88	0,92	0,96	0,93	0,9	0,89	0,92	0,91	0,78	0,86	0,86	0,92	0,9	0,9
26	KRQ	0,92	0,95	0,94	0,94	0,85	0,92	0,91	0,88	0,72	0,74	0,78	0,93	0,95	0,93
27	ALT	0,82	0,92	0,95	0,94	0,86	0,85	0,89	0,93	0,72	0,76	0,88	0,94	0,92	0,91
28	XAL	0,86	0,82	0,9	0,86	0,8	0,9	0,92	0,8	0,77	0,76	0,73	0,85	0,85	0,81
29	KUU	0,81	0,89	0,93	0,91	0,85	0,83	0,88	0,9	0,69	0,74	0,84	0,92	0,9	0,86
30	KMD	0,77	0,86	0,88	0,89	0,87	0,78	0,86	0,88	0,69	0,73	0,84	0,89	0,87	0,88
31	TUB	0,81	0,91	0,92	0,94	0,86	0,83	0,88	0,91	0,71	0,77	0,86	0,93	0,91	0,93

BLK	GAG	KRM	KLP	ATU	SAL	SHR	TOF	KAZ	QUM	KRX	KRQ	ALT	XAL	KUU	KMD	TUB
0,83	0,98	0,86	0,89	0,87	0,85	0,79	0,76	0,86	0,87	0,88	0,92	0,82	0,86	0,81	0,77	0,81
0,93	0,84	0,94	0,98	0,9	0,89	0,86	0,8	0,97	0,96	0,92	0,95	0,92	0,82	0,89	0,86	0,91
0,95	0,9	0,95	0,99	0,95	0,95	0,91	0,83	0,99	0,98	0,96	0,94	0,95	0,9	0,93	0,88	0,92
0,9	0,86	0,92	0,96	0,92	0,92	0,9	0,82	0,95	0,95	0,93	0,94	0,94	0,86	0,91	0,89	0,94
0,83	0,78	0,84	0,87	0,88	0,84	0,85	0,81	0,85	0,86	0,9	0,85	0,86	0,8	0,85	0,87	0,86
0,85	0,92	0,88	0,89	0,88	0,9	0,8	0,75	0,86	0,91	0,89	0,92	0,85	0,9	0,83	0,78	0,83
0,89	0,9	0,92	0,95	0,89	0,92	0,85	0,79	0,94	0,93	0,92	0,91	0,89	0,92	0,88	0,86	0,88
0,86	0,79	0,86	0,93	0,87	0,84	0,95	0,87	0,9	0,88	0,91	0,88	0,93	0,8	0,9	0,88	0,91
0,73	0,73	0,72	0,77	0,76	0,76	0,67	0,68	0,75	0,79	0,78	0,72	0,72	0,77	0,69	0,69	0,71
0,8	0,75	0,75	0,78	0,84	0,74	0,76	0,77	0,75	0,77	0,86	0,74	0,76	0,76	0,74	0,73	0,77
0,8	0,74	0,79	0,85	0,79	0,75	0,85	0,97	0,79	0,79	0,86	0,78	0,88	0,73	0,84	0,84	0,86
0,93	0,88	0,94	0,99	0,91	0,9	0,89	0,84	0,99	0,94	0,92	0,93	0,94	0,85	0,92	0,89	0,93
0,96	0,86	0,92	0,99	0,9	0,91	0,87	0,82	0,98	1	0,9	0,95	0,92	0,85	0,9	0,87	0,91
0,93	0,81	0,93	0,98	0,89	0,87	0,86	0,79	0,95	0,97	0,9	0,93	0,91	0,81	0,86	0,88	0,93
1	0,83	0,93	0,94	0,9	0,87	0,84	0,8	0,93	0,98	0,89	0,9	0,89	0,81	0,86	0,83	0,88
0,83	1	0,84	0,88	0,85	0,84	0,77	0,74	0,85	0,86	0,88	0,9	0,82	0,84	0,79	0,78	0,8
0,93	0,84	1	0,93	0,89	0,9	0,87	0,79	0,92	0,94	0,89	0,95	0,9	0,86	0,87	0,86	0,91
0,94	0,88	0,93	1	0,92	0,9	0,9	0,85	1	0,97	0,94	0,94	0,95	0,87	0,92	0,89	0,93
0,9	0,85	0,89	0,92	1	0,89	0,85	0,81	0,9	0,91	0,97	0,88	0,89	0,88	0,88	0,84	0,86
0,87	0,84	0,9	0,9	0,89	1	0,84	0,76	0,89	0,91	0,89	0,89	0,88	0,92	0,85	0,83	0,85
0,84	0,77	0,87	0,9	0,85	0,84	1	0,82	0,87	0,86	0,9	0,86	0,91	0,79	0,9	0,88	0,9
0,8	0,74	0,79	0,85	0,81	0,76	0,82	1	0,81	0,81	0,83	0,77	0,88	0,72	0,84	0,84	0,86
0,93	0,85	0,92	1	0,9	0,89	0,87	0,81	1	0,96	0,91	0,93	0,92	0,83	0,9	0,88	0,9
0,98	0,86	0,94	0,97	0,91	0,91	0,86	0,81	0,96	1	0,91	0,94	0,92	0,85	0,89	0,86	0,89
0,89	0,88	0,89	0,94	0,97	0,89	0,9	0,83	0,91	0,91	1	0,91	0,89	0,88	0,87	0,85	0,88
0,9	0,9	0,95	0,94	0,88	0,89	0,86	0,77	0,93	0,94	0,91	1	0,89	0,88	0,86	0,84	0,89
0,89	0,82	0,9	0,95	0,89	0,88	0,91	0,88	0,92	0,92	0,89	0,89	1	0,82	0,96	0,95	0,96
0,81	0,84	0,86	0,87	0,88	0,92	0,79	0,72	0,83	0,85	0,88	0,88	0,82	1	0,79	0,78	0,8
0,86	0,79	0,87	0,92	0,88	0,85	0,9	0,84	0,9	0,89	0,87	0,86	0,96	0,79	1	0,96	0,96
0,83	0,78	0,86	0,89	0,84	0,83	0,88	0,84	0,88	0,86	0,85	0,84	0,95	0,78	0,96	1	0,96
0,88	0,8	0,91	0,93	0,86	0,85	0,9	0,86	0,9	0,89	0,88	0,89	0,96	0,8	0,96	0,96	1

Таблица 8. Доли совпадений основных списков тюркских языков¹⁹

¹⁹ Доли совпадений приводятся по данным из архива проекта «Глобальная лексикостатистическая база данных» (starlingdb.org/new100/main.htm).

Приложение 3

Обозначения языков			Доли совпадений			Ошибка согласованности E_c
A	B	C	N_{AB}	N_{BC}	N_{AC}	
UZB	AZB	KAZ	0.95	0.86	0.99	0.08
UZB	AZB	BAS	0.95	0.84	0.96	0.07
UZB	KAZ	XAL	0.99	0.83	0.90	0.06
UZB	AZB	SHR	0.95	0.80	0.91	0.06
TUV	KLP	KAZ	0.85	1,00	0.79	0.06
TAT	UZB	XAL	0.98	0.90	0.82	0.06
UZB	AZB	KUU	0.95	0.83	0.93	0.05
TRK	BAS	KLP	0.82	0.98	0.89	0.05
UZB	AZB	ALT	0.95	0.85	0.95	0.05
UZB	AZB	KMD	0.95	0.78	0.88	0.05
TAT	UZB	AZB	0.98	0.95	0.88	0.05
UZB	BAS	GAG	0.96	0.81	0.90	0.05
UZB	BAS	XAL	0.96	0.81	0.90	0.05
BAS	GAG	KLP	0.81	0.88	0.98	0.05
UZB	SAL	KAZ	0.95	0.89	0.99	0.05
UZB	AZB	KRG	0.95	0.89	0.99	0.05
UZB	AZB	NOG	0.95	0.88	0.98	0.05
...						

Таблица 9. Результаты анализа согласованности исходных данных для тюркских языков (фрагмент таблицы)²⁰

Обозначения языков

TRK — турецкий, TAT — татарский, UZB — узбекский, UIG — уйгурский, SJG — сарыюгурский, AZB — азербайджанский, TRM — туркменский, HAK — хакаский, CHV — чувашский, JAK — якутский, TUV — тувинский, KRG — киргизский, NOG — ногайский, BAS — башкирский, BLK — балкарский, GAG — гагаузский, KRM — караимский, KLP — каракалпакский, ATU — древнетюркский, SAL — саларский, SHR — шорский, TOF — тофаларский, KAZ — казахский, QUM — кумыкский, KRX — караханидский, KRQ — крымский караимский, ALT — алтайский, XAL — халаджский, KUU — челканский (куу), KMD — кумандинский, TUB — тубаларский.

Литература

- Бурлак, С. А., С. А. Старостин. 2005. *Сравнительно-историческое языкознание*. Москва: «Академия».
- Васильев, М. Е. 2010. Об использовании лексического критерия для построения генеалогической классификации. *Бюллетень Общества востоковедов РАН* 17: 530–572.
- Васильев, М. Е., А. И. Коган. 2013. К вопросу о восточнодардской языковой общности. *Вопросы языкового родства* 16(117): 149–177.

²⁰ Таблица с полным списком несогласованных триад приводится в сопровождающих материалах, доступных на сайте журнала.

- Васильев, М. Е., М. Н. Саенко. 2020. Анализ топологии и оценка точности лексикостатистических классификаций (на примере славянских языков). *Вопросы языкового родства* 18/4: 320–347.
- Милитарёв, А. Ю. 2000. О дохлой лошади, оказавшейся скакуном. *Знание — сила* 4: 20–31.
- Соловьёв, В. Д., А. В. Дыбо, Г. С. Старостин. 2010. Типологическая схожесть языков как метод изучения языковой эволюции. Дискуссия. *Вопросы языкового родства* 4: 177–198.
- Старостин, С. А. 2007. Сравнительно-историческое языкознание и лексикостатистика. В: С. А. Старостин. *Труды по языкознанию*: 407–447. Москва: «Языки славянских культур».

References

- Burlak, S. A., S. A. Starostin. 2005. *Sravnitel'no-istoricheskoe jazykoznanie*. Moskva: «Akademija».
- Militarev, Alexander. 2005. Once more about glottochronology and the comparative method: the Omotic-Afrasian case. In: Ilya Smirnov (ed.). *Orientalia et Classica. Trudy Instituta vostochnyx kultur i antichnosti*: 339–408. Moscow: RSUH Publishers.
- Militarev, A. Ju. 2000. O doxloj loshadi, okazavshejs'a skakunom. *Znanie — sila* 4: 20–31.
- Solovjev, V. D., A. V. Dybo, G. S. Starostin. 2010. Tipologicheskaja sxozhest' jazykov kak metod izuchenija jazykoj evol'utsii. Diskussija. *Journal of Language Relationship* 4: 177–198.
- Starostin, George. 2010. Preliminary lexicostatistics as a basis for language classification: A new approach. *Journal of Language Relationship* 3: 79–116.
- Starostin, S. A. 2007. Sravnitel'no-istoricheskoe jazykoznanie i leksikostatistika. In: S. Starostin. *Trudy po jazykoznaniju*: 407–447. Moskva: «Jazyki slav'anskix kul'tur».
- Starostin, Sergey. 2000. Comparative-historical linguistics and lexicostatistics. In: Colin Renfrew, April McMahon, Larry Trask (eds.). *Time Depth in Historical Linguistics*, Vol. 1: 223–259. Cambridge: The McDonald Institute for Archaeological Research Publications.
- Vasilyev, M. E. 2010. Ob ispol'zovanii leksicheskogo kriterija dl'a postrojenija genealogicheskoy klassifikatsii. *B'ulleten' Obshchestva vostokovedov RAN* 17: 530–572.
- Vasilyev, M. E., A. I. Kogan. 2013. K voprosu o vostochnodardskoj jazykovoj obshchnosti. *Journal of Language Relationship* 16(117): 149–177.
- Vasilyev, M. E., M. N. Saenko. 2020. Analiz topologii i otsenka tochnosti leksikostatisticheskix klassifikatsij (na primere slav'anskix jazykov). *Journal of Language Relationship* 18/4: 320–347.
- Vydrin, Valentin. 2009. On the Problem of the Proto-Mande Homeland. *Journal of Language Relationship* 1: 107–142.

Mikhail Vasilyev. Evaluating the reliability of source data in lexicostatistical analysis

The article addresses the reliability of the source data used in comparative-historical linguistics for lexicostatistical purposes (to obtain genealogical classifications and linguistic dating). Such data, usually represented by a table of cognacy percentages between the basic wordlists of languages, may contain errors or inaccuracies, stemming from the complexity and subjectivity of the etymological analysis procedure, which may significantly reduce the reliability of lexicostatistical calculations. To solve this issue, a formal methodology is proposed based on the criterion of consistency (or transitivity) of percentage values in the initial lexicostatistical table. Applying this criterion enables the identification of unreliable values in the source data, as well as the numerical estimation of data inconsistency in each case. The advantages of the proposed approach include its simplicity and versatility, the objectivity of the results, and ease of implementation as a computer application. Testing the methodology on the lexicostatistical data of Romance and Turkic languages proves its applicability and practical efficiency while examining both small and large language groups.

Keywords: lexicostatistics; glottochronology; distance matrix; consistency criterion.