

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

№ 21, issue 3—4



Moscow 2023

Российский государственный гуманитарный университет  
Институт языкознания Российской Академии наук

# Вопросы языкового родства

Международный научный журнал

№21, часть 3—4

**ВЯР**

Москва 2023

*Advisory Board:*

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

*Editorial Staff:*

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

Founded by Kirill BABAEV

*Редакционный совет:*

Х. АЙХНЕР (Вена) / председатель

В. БЛАЖЕК (Брно)

У. БЭКСТЕР (Анн Арбор)

В. Ф. ВЫДРИН (Париж)

Ф. КОРТЛАНДТ (Лейден)

А. ЛУБОЦКИЙ (Лейден)

А. Ю. МИЛИТАРЕВ (Москва)

Дж. МЭЛЛОРИ (Белфаст)

Л. ХАЙМАН (Беркли)

*Редакционная коллегия:*

Г. С. СТАРОСТИН (главный редактор)

А. В. ДЫБО (заместитель главного редактора)

Т. А. МИХАЙЛОВА

М. Н. САЕНКО

И. С. ЯКУБОВИЧ

А. Г. РОМАНЦЕВСКИ (ответственный секретарь)

Журнал основан К. В. БАБАЕВЫМ

Вопросы языкового родства: Международный научный журнал / Рос. гос. гуманитар. ун-т; Рос. акад. наук. Ин-т языкознания; под ред. Г. С. Старостина. — М., 2023. — № 21/3—4. — x + 176 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, 2023. — № 21/3—4. — x + 176 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, Москва, Миусская пл., 6

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

Подписано в печать 11.03.2024. Выход в свет 18.03.2024  
Формат 60×90/8. Уч.-изд. л. 14,0. Усл. печ. л. 23,3  
Тираж 1050 экз. Свободная цена  
Заказ № 1925

Отпечатано в типографии Издательского центра  
Российского государственного гуманитарного университета  
125047, Москва, Миусская пл., 6  
[www.rsuh.ru](http://www.rsuh.ru)

# Table of Contents / Содержание

Contributors / Сведения об авторах . . . . .	ix
Note for Contributors / Будущим авторам . . . . .	x

## Issue 3 / Часть 3

### Articles / Статьи

<i>Iliia Afanasev</i> . Cipher, transform, get lost: an anti-transparent system for distance measurement in East Slavic lects . . . . .	159
[И. Афанасьев. Потенциал систем — «чёрных ящиков» для автоматического измерения расстояния между восточнославянскими лектами]	
<i>Alexander Mankov</i> . Proto-Indo-European Nom.-Acc. Dual and the Germanic Dual of nouns . . . . .	178
[А. Е. Маньков. Праиндоевропейский им.-вин. падеж двойственного числа существительных и двойственное число существительных в германских языках]	
<i>Andrey Nikulin</i> . Nasal harmony in consonants in Chiquitano and its origins . . . . .	184
[А. В. Никулин. Консонантная гармония по назальности в чикитано и её истоки]	
<i>Anna Smirnitskaya</i> . Nominations for siblings: Proto-Dravidian reconstruction and borrowability . . . . .	201
[А. Е. Маньков. Праиндоевропейский им.-вин. падеж двойственного числа существительных и двойственное число существительных в германских языках]	
<i>A. A. Trofimov</i> . О правилах ударения в пушту . . . . .	224
[Artem Trofimov. On the rules of stress in Pashto nominal stems]	

## Issue 4 / Часть 4

### Articles / Статьи

<i>Jesse P. Gates</i> . The uvularization of <i>g-/d-</i> in Tibetic languages . . . . .	245
[Дж. Гейтс. О происхождении увулярных согласных в тибетических языках]	
<i>Florian Réveilhas</i> . * <i>Tubar(i)-</i> : a divine epithet reflected in Luwic onomastics . . . . .	259
[Ф. Ревейяк. Божественный эпитет * <i>tubar(i)-</i> в лувийской ономастике]	
<i>С. К. Белых</i> . Пермско-иранские языковые контакты по данным фонетики и морфологии удмуртского и коми языков . . . . .	282
[Sergei Belykh. Permic-Iranian language contacts according to phonetics and morphology of the Udmurt and Komi languages]	
<i>И. В. Кузнецов</i> . Из истории изучения дальнего родства: шесть фил в Северной Америке . . . . .	300
[Igor Kuznetsov. Toward a history of macro-comparative linguistics: on the six phyla of North America]	

## Contributors

*Iliia Afanasev* — postgraduate student, School of Linguistics of the Higher School of Economics (Moscow), [ilia.afanasev.1997@gmail.com](mailto:ilia.afanasev.1997@gmail.com)

*Sergei Belykh* — researcher, School for Advanced Studies in the Humanities, RANEPA (Moscow), [bibulus@e-izhevsk.ru](mailto:bibulus@e-izhevsk.ru)

*Jesse P. Gates* — Ph.D., post-doctoral fellow, Nankai University, School of Literature (Tianjin, China), [stauskad@gmail.com](mailto:stauskad@gmail.com)

*Igor Kuznetsov* — candidate of sciences (History), senior researcher, Institute of Linguistics of the Russian Academy of Sciences (Moscow), [i.kuznetsov@iling-ran.ru](mailto:i.kuznetsov@iling-ran.ru)

*Alexander Mankov* — candidate of sciences (Philology), senior researcher, Institute of Linguistics of the Russian Academy of Sciences (Moscow), [mankov@iling-ran.ru](mailto:mankov@iling-ran.ru)

*Andrey Nikulin* — Ph.D., faculty member, Núcleo Takinahakỹ de Formação Superior Indígena, Federal University of Goiás, [andre.n.guzman@gmail.com](mailto:andre.n.guzman@gmail.com)

*Florian Réveillac* — Ph.D., post-doctoral fellow, University of Marburg (Germany), [florian.reveillac@gmail.com](mailto:florian.reveillac@gmail.com)

*Anna Smirnitckaya* — candidate of sciences (Philology), researcher, Institute of Linguistics of the Russian Academy of Sciences (Moscow) / Institute of Oriental Studies of the Russian Academy of Sciences (Moscow), [nyushas@gmail.com](mailto:nyushas@gmail.com)

*Artem Trofimov* — candidate of sciences (Philology), researcher, School for Advanced Studies in the Humanities, RANEPA (Moscow), [artemii.trofimov@gmail.com](mailto:artemii.trofimov@gmail.com)

## Сведения об авторах

*Афанасьев, Илья Андреевич* — аспирант Школы лингвистики НИУ ВШЭ (Москва), [ilia.afanasev.1997@gmail.com](mailto:ilia.afanasev.1997@gmail.com)

*Белых, Сергей Константинович* — научный сотрудник Школы актуальных гуманитарных исследований РАНХиГС (Москва), [bibulus@e-izhevsk.ru](mailto:bibulus@e-izhevsk.ru)

*Гейтс, Джесси П.* — Ph.D., постдокторант Школы литературы Нанькайского университета (Тяньцзинь, КНР), [stauskad@gmail.com](mailto:stauskad@gmail.com)

*Кузнецов, Игорь Валерьевич* — канд. исторических наук, старший научный сотрудник Института языкознания РАН (Москва), [i.kuznetsov@iling-ran.ru](mailto:i.kuznetsov@iling-ran.ru)

*Маньков, Александр Евгеньевич* — канд. филол. наук, старший научный сотрудник Института языкознания РАН (Москва), [mankov@iling-ran.ru](mailto:mankov@iling-ran.ru)

*Никулин, Андрей Владимирович* — Ph.D., сотрудник федерального университета штата Гояс, [andre.n.guzman@gmail.com](mailto:andre.n.guzman@gmail.com)

*Ревьяк, Флориан* — Ph.D., постдокторант Марбургского университета (Германия), [florian.reveillac@gmail.com](mailto:florian.reveillac@gmail.com)

*Смирнитская, Анна Александровна* — канд. филол. наук, научный сотрудник Института языкознания РАН / Института востоковедения РАН (Москва), [nyushas@gmail.com](mailto:nyushas@gmail.com)

*Трофимов, Артем Александрович* — канд. филол. наук, научный сотрудник Лаборатории востоковедения и сравнительно-исторического языкознания ШАГИ РАНХиГС (Москва), [artemii.trofimov@gmail.com](mailto:artemii.trofimov@gmail.com)



## Note for Contributors

*Journal of Language Relationship* welcomes submissions from everyone specializing in comparative-historical linguistics and related disciplines, in the form of original articles as well as reviews of recent publications. Articles are published preferably in English or Russian, although publication of texts in other major European languages (such as French or German) may be considered by the Editorial Board.

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

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

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

## Будущим авторам

Журнал *Вопросы языкового родства* принимает заявки на публикацию оригинальных научных статей, а также рецензий от всех, кто специализируется в области сравнительно-исторического языкознания и смежных дисциплин. Предпочтительные языки публикации — английский или русский, хотя по согласованию с редакционной коллегией возможна также публикация статей на других крупных европейских языках (французский, немецкий и т. п.).

Подача материалов для публикации (в форматах MS Word и PDF) осуществляется через *Электронную форму подачи рукописей* на официальном сайте журнала (<http://jolr.ru>). К каждой статье обязательно прикладываются сведения об авторах (имена, аффилиации, контактная информация), краткое резюме (не более 300 слов) и список подходящих ключевых слов.

Подробнее о требованиях к оформлению рукописи, редакционной политике журнала и т. п. вы можете узнать на нашем сайте по адресу: <http://www.jolr.ru> или непосредственно от редакционной коллегии по электронной почте ([gstarst1@gmail.com](mailto:gstarst1@gmail.com)). По различным вопросам с редакцией журнала можно также связаться по обычной почте:

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



**Issue 3 / Часть 3**

**Articles / Статьи**



## **Cipher, transform, get lost: an anti-transparent system for distance measurement in East Slavic lects**

Recent advances in computational historical linguistics have inspired a discussion on newly implemented quantitative methods — mainly, it is about their lack of transparency, and the ways to overcome it. This paper aims to demonstrate the advantages of transparency for such tools.

The study compares two types of language distance measurement systems used in classification. Black-box systems transform the input data (such as the Swadesh list) into output data (language distance) with human- and machine-unexplainable decision-making. Language-agnostic systems (such as string similarity measures) analyse the input data and produce output data transparently, but do not consider the specifics of each language. For a proper comparison, I propose a new anti-transparent system based on hashing algorithms, vectorisation and language contact emulation.

For my purposes, I use material from two test groups — East Slavic and Taa, both lexical and grammatical. East Slavic data are extracted from the corpora of Belogornoje, Megra, and Khislavichi and feature lists of Mokshenskaja, Kritskovschina and Pestschanka. Taa material consists of previously published Swadesh lists for the closely related !Xóõ (!Xoong), Kafia (Masarwa) and N|u|len. An important new contribution of this work is the publication of new Swadesh wordlists for three East Slavic dialects.

**Keywords:** black-box methods; East Slavic languages; South Khoisan languages; Taa languages; language-agnostic methods; automatic language distance measurement; automatic classification; string similarity measures; basic lexicon.

### **Background**

Recent innovations in neural network methods (Sutskever et al. 2013; Cho et al. 2014; Vaswani et al. 2017; Lewis et al. 2019) and training data have critically facilitated a plethora of tasks from grammatical error correction (Syvokon et al. 2023) to sentiment analysis (Barić et al. 2023). However, these results came at the disturbing cost of interpretability (Bastings et al. 2022). The need for explanatory techniques for machine learning (ML) systems has become more viable than ever (Munn & Pitman 2022).

Historical comparative linguistics was among the first fields to raise suspicions about these cutting-edge technologies (Jäger 2019). It was among the first to adopt computational phylogenetic methods, right after dialectometry (Nerbonne and Heeringa 1997; Wichmann et al. 2011; Snoek 2013; List 2014; Rama & Borin 2015; Wichmann & Rama 2018). At the same time, it remains rightfully aware of the new methods becoming less and less interpretable (Carvalho 2020).

In this fashion, Prokić & Moran 2013 require the methods of automatic language distance measurement to be transparent and linguistically explainable. Their otherwise valid point, however, depends on the Levenshtein distance being a black box method (Prokić & Moran 2013: 442), which it is not. It is a transparent method, even if it is language-agnostic.

Black box and language-agnostic approaches are distinctly different in the following aspects: transparency, implementation, reproducibility, overall efficiency, and application scope.

The most important one is explainability, the possibility for a human researcher to trace the inner workings of a model, either manually or with the help of automatic tools (Munn & Pitman 2022). Table 1 demonstrates the key differences between these two types of systems.

System property	Black-box model	Language-agnostic
Transparency	As non-transparent as possible	As transparent as possible
Explainability	Non-explainable by definition	Explainable, often inherently
Implementation	Implemented based on the researcher's idea of what may efficiently transform input to output, with little attention to the inherent properties of the studied object	Implemented based on the inherent properties of the studied object (sequentiality of the string, whether a genetic code sequence or a word in a list)
Reproducibility	Almost impossible to reproduce reliably	Easily reproducible
Transferability	Designed for a specific task, attempts at applying it to other tasks lead to unpredictable fluctuations in results	Designed for a specific type of task
Efficiency	Vary in efficiency	Achieve high enough score

Table 1. The crucial differences in system properties between black-box and language-agnostic models.

This paper intends to illustrate these differences by building a system that matches the definition of a black box as closest to completely nonsensical decision-making mechanisms.

I hypothesise the following:

**H1.** A true black-box method heavily differs from language-agnostic methods in the degree to which one may linguistically explain its functionality.

**H2.** Using a true black-box method leads to building a less reliable classification than an application of a language-agnostic method.

**H3.** A black-box method is much more sensitive to farther degrees of relationship than a language-agnostic method.

I compare a black-box method and a language-agnostic method with different human classifications of specific lects under consideration as well as to a set of language-aware computational historical linguistics methods.

### Necessity of cross-evaluation

The introduction of a completely new method requires some necessary steps from a researcher, lack of which may undermine the whole intent of the process. I unite these actions under an umbrella of cross-evaluation, a set of practices that help to highlight the possible restrictions of the introduced method. Cross-evaluation may be implemented for both method and data.

Data cross-evaluation shows the advantages and disadvantages of an introduced method by using different datasets.

Method cross-evaluation demonstrates the advantages and disadvantages of an introduced method by comparing its results to the results acquired by other previously established, and proven efficient, methods. In the case of genetic classification, this includes a comparison between automatic methods and human classification. There are cases, when one may not rely on human classification as a gold standard, due to the lack of consensus between the researchers.

Mostly these are either poorly studied families, not to mention macrofamilies (see examples in Starostin 2011, Vajda 2012, Zhivlov 2021). However, internal reconstruction of some well-established families, such as Slavic, also may lead to this issue (for instance, Feld and Maxwell 2019; Ryko and Spiricheva 2022). It does not mean that automatic classification is more objective or better, yet human data cannot be employed as an indisputable gold standard in the automatic classification task in numerous cases that demonstrate a lack of consensus on the genetic classification even by the human researchers. I propose an approach akin to the Natural Language Generation (NLG). This approach requires us to be cautious of the human data, especially in vague cases of long-distance relationship, or close lects classification, and demanding of the metrics (Novikova et al. 2017).

## Data

I utilise three datasets, two in the form of a Swadesh list, and one in the form of a lect feature listing.

The main dataset consists of two corpora of East Slavic small territorial lects. These are the Saratov dialect corpus (Kryuchkova & Goldin 2011) and the Khislavichi corpus (Ryko & Spiricheva 2020). The Saratov dialect corpus represents (among others) two Russian lects: Megra and Belogornoje. The Khislavichi corpus represents the single East Slavic lect of disputable genetic attribution between the Belarusian and Russian continua (Ryko & Spiricheva 2022; Afanasev 2023).

Megra is a northern (Kryuchkova & Goldin 2011) Russian dialect, spoken in the Vologda Region. The Megra corpus consists of transcribed interviews (mostly slice-of-life stories) from Saratov State University field trips (1980–2019). Belogornoje is a central (Barannikova 2005) Russian dialect. The Belogornoje corpus consists of transcribed folklore tales and interviews from Saratov State University field trips (1980–2019). Khislavichi is spoken in the Smolensk Region (Russia). Ryko and Spiricheva (2022) treat it as a Northern Belarusian dialect, intensely Russified during the XXth century. Khislavichi material is a collection of slice-of-life stories, gathered during a 2019 field trip.

There is no ready-made Swadesh list for any of these lects and there are no dictionaries. I collect a basic 40-word list (Holman et al. 2008), generally following the guidelines for East Slavic languages (Kassian et al. 2010). Within classical lexicostatistics framework, all the items would be clear matches and comparison would be meaningless. However, as the metrics applied in this article are more sensitive, using a 40-word list may yield meaningful results, as proven earlier in computational dialectometry (Nerbonne & Heeringa 1997; Gooskens & Heeringa 2004) and computational phylogenetic linguistics (Holman et al. 2008).

I present the concepts and their realisations for each lect in the form of phonetic transcriptions. The symbolic representations of sounds are taken from IPA. The transcriptions are phonematic and preserve key features of the lects, such as *okanje*, distinguishing between [o] and [a] allophones of phoneme <o> in the first unstressed position; but not the individual features of speakers. The transcription also preserves all the irregularities (such as *okanje* in *voda* ‘water’ ‘fire’ in Khislavichi, a dialect more prone to *akanje*, a coincidence of [o] and [a] allophones of phoneme <o> in the first unstressed position). As word stress does not contribute to the differences between the given lects and the distance measurement methods, discussed in the article, do not utilise it (Holman et al. 2008), the transcriptions do not represent it.

Concept	Megra (Northern Russian, Vologda Region, Russia)	Belogornoje (Central Russian, Saratov Region, Russia)	Khislavichi (Northern Belarusian, Smolensk Region, Russia)
eye	glʲas	glʲas	ɣlʲas
ear	uxo	uxo	vuxa
nose	nos	nos	nos
tongue	jazik	jazik	jazik
tooth	zup	zup	zup
hand	ruka	ruka	ruka
knee	kolʲeno	kolʲeno*	kalʲena
blood	krof	krofʲ	krow
bone	kosʲtʲ*	kosʲtʲ*	kosʲtʲ*
breast (woman's)	grutʲ	grutʲ	ɣrutʲsʲ
liver	pʲetʲʃenʲ	pʲetʲʃenʲ	NA
skin	koza	koza	koza
louse	voʃ	blʲoxa*	blʲaxa*
dog	sobaka	sobaka	sabaka
fish (noun)	riba	riba	riba
horn (animal part)	rok*	rok*	rox*
tree	dʲerʲevo	dʲerʲeva	dʲzʲerʲeva
leaf	lʲist	lʲist	lʲist
person	fʲjelovek	fʲjelovek	fʲjelavek
name (noun)	imʲa	imʲa	imʲa
sun	solʲniʃko**	sonʲtse	sonʲtse
star	zvʲozdoʲfka*/**	zvʲezda	zvʲezda
water	voda	voda	voda
fire	ogonʲ	ogonʲ	aʲonʲ
stone	kamenʲ	kamenʲ	kamenʲ
path	doroga	doroga	daroya
mountain	gora	gora	ɣara
night (dark time)	noʲtʲ	noʲtʲ	noʲʃ
drink (verb)	pʲitʲ	pʲitʲ	pʲitʲ
die	umʲiratʲ	umʲiratʲ	umʲiratsʲ
see	vʲidʲetʲ	vʲidʲetʲ	vʲidʲetʲ
hear	slʲyʃatʲ	slʲyʃatʲ	slʲyʃatʲ
come	prʲijtʲi	prʲijtʲi	prʲijtʲitʲ
new	novij	novij	novij
full	polʲnij	polʲnij	pownij
one	odʲin	odʲin	adʲin
two	dva	dva	dva
I	ja	ja	ja
you	tʲi	tʲi	tʲi
we	mi	mi	mi

Table 2. Swadesh lists for Megra, Belogornoje and Khislavichi lects. The lists are separately published in an open-source repository<sup>1</sup>.

<sup>1</sup> <https://huggingface.co/datasets/djulian13/east-slavic-swadesh-lists>



I gather these lists from raw corpora material. The words for which I was not able to find a word in its base form are marked with asterisk (\*) in the table. In such cases, I manually transformed the word into its base form. I surmise its paradigm from the data I have gathered during my study of the corpus.

Gathering lists from raw corpora material, without possibility to access the lexicographical data, adds other complications, such as impossibility to reliably check whether a particular word is indeed the best match for a particular concept within the lect from historical point of view, or only a contextual one. This is the case of *bl'axa/bl'oxa* 'louse (lit. flea)' in Khislavichi and Belogornoje, cf. example from Khislavichi: *Блохи . Кланы , меня уже что ?* 'Lice (lit. fleas). Bedbugs, [that bit] me already, what?' (Ryko & Spiricheva 2020). However, since both the Khislavichi and Belogornoje corpora contain the word in this meaning, rather than *\*voš* 'louse', as in Megra, the final dataset has no other choice than to include it.

The Megra Swadesh list includes diminutives (marked with a double asterisk) as the names of astronomical entities (*sun* and *star*). In the Megra lect, they are more frequent than their historically non-diminutive counterparts, more stylistically neutral (used in different context types) and semantically narrow, a relatively frequent phenomenon in Slavic languages<sup>2</sup>. One might also argue that they may be not as historically stable as non-diminutives. However, there is no sufficient data to support this claim. There are almost no data prior to 1980 on the Megra lect, and data from 1980 to 2019 consistently favour the hypothesis I present.

There is no attested word for *liver* in the Khislavichi corpus, henceforth *NA* value in the table.

The purpose of the second Swadesh list dataset is cross-evaluation on a set of lects that would be completely different both genetically and typologically. For these goals, I use three Tuu (Taa subgroup) lects from the Khoisan linguistic area: !Xóõ, N|u||en and Kakia (Masarwa). For Taa, there is an existing wordlist, compiled and annotated by G. Starostin (2021; 2022) from previously published sources. Table 3 reproduces the list. Since Tuu is not the focus of this study, I give the wordlist itself for reference, but do not thoroughly discuss it, or the relationship between the lects themselves.

The last dataset I use is a set of phonetic features across East Slavic dialects taken from Marchenko & al. 2023. I treat features as sets  $\{Realisation1, Realisation2, \dots, RealisationN\}_{feature}$  (Archangeli & Pulleybank 2022: 32). Each realisation receives a specific letter (I use A for the realisation I found first in the first analysed lect, B for the realisation I found first in the second analysed lect, and so forth). If there is no realisation of a feature in a lect, initially I insert an absence sign. After the lists of features for each lect are ready, the pre-processing turns them into a string (cf., CCCC-----CCCCC--CCCCCC-AC-C--CC-CC-CC-C-C-CCCC-----A---C-BA--B--A-AA----AACCC-ACCCCC-A-AA-----AC-C-CBBBB--A-AAB--BC----CCC-----).

There is no information of such kind for Khislavichi, Megra and Belogornoje. Yet there is information for the lects that seems to be of the same East Slavic continuum areas. I take Mokshenskaja (Northern Russian, Arkhangelsk Region, close relative of Megra), Piestchanka (Southern Russian, Saratov Region, close relative of Belogornoje), and Kritskovschina (Western Russian, Smolensk Region, close relative of Khislavichi). Picture 1 shows the geographical distribution of all lects.

<sup>2</sup> <https://starlingdb.org/cgi-bin/response.cgi?root=new100&basename=new100\ier\slv>.

Concept	!Xóǒ	Kakia (Masarwa)	N u len
eye	!ʔũ	x'wĩ	ʔũ
ear	ʔũā <sup>h</sup>	ĩwa:	ʔu-ša
nose	ĩ <sup>h</sup> -ja	ĩu-ča	ĩu-ša
tongue	ʔnà'n	ĩa:n	!a:ni
tooth	q <sup>h</sup> ǎ:	xũ	ʔn-te
hand	x'à:	x'a	x'a
knee	xú:-ĩàn	ǒ-ĩaŋ	ũ  i
blood	ĩá. <sup>f</sup>	ĩá. <sup>f</sup> a	ĩá. <sup>f</sup> a
bone	ʔǎ:	a:	ʔǎ
breast (woman's)	ú:	am	u #
liver	ám	NA	ʔm
skin	tú <sup>f</sup> m	t'üm	t'üm
louse	ǒú. <sup>f</sup>	NA	NA
dog	ʔq <sup>h</sup> ǎi	ʔxai	ʔ <sup>h</sup> i
fish (noun)	NA	NA	NA
horn (animal part)	ǎē	ʔn-ša	ǎ
tree	ʔθnàye	ǒoe:	ǒ'a:
leaf	ána	a:na	ĩabu
person	tâ:	tu	tu
name (noun)	ǎũ	x'ǎũ	ǎ
sun	ǎn	ʔn	ʔn
star	óna	wana-te	ana-te
water	!q <sup>h</sup> ǎ:	! <sup>h</sup> ǎ	! <sup>h</sup> a
fire	!ǎ:	!ǎ:	!ǎ
stone	ĩũ-le	ĩũ-le	!um
path	ʔólo	dau #	dau #
mountain	!ù <sup>h</sup> m	ĩu:-n	!um
night (dark time)	ĩúe <sup>f</sup>	ĩǒe <sup>f</sup>	ĩǒe
drink (verb)	x'ǎ. <sup>h</sup>	x'ǎ	x'a-a
die	!ǎ:	!a	!a:
see	ĩǎ	ĩa	ĩe:
hear	tá. <sup>f</sup>	tāa	sa
come	sî:	si	si
new	qu <sup>f</sup> V	xwe	NA
full	!ù <sup>h</sup> m	!úm	!um
one	ʔũā	!k'we	!oe
two	ʔũm	ĩum	ĩum
I	ñ	n	ŋ
you	NA	NA	NA
we	ĩ <sup>h</sup>	i	i

Table 3: Swadesh lists for !Xóǒ, N|u|len and Kakia (Masarwa). NA denotes concepts absent in the material.



For each pair of lects, I swap some concepts to see how this distortion contributes to the classification result. I either swap a restricted number of random concepts between two lects or replace words from one lect with words from another lect, emulating an intense stream of borrowings. The first type of simulation results in a transfer like  $a, b \rightarrow b, a$ , the second – in a transfer like  $a, b \rightarrow a, a$ . Then I retrain and re-evaluate the classifier. I acquire a squared error between the original and the resulting F1-score. I repeat the process for a certain number of runs to reduce the element of randomness. I use UPGMA (Sokal & Michener 1958) to build genealogical trees based on acquired data, a tree for a run (for 100 runs I get 100 trees).

The groups under consideration are closely related, providing a similar evolution rate, and monophyletic, having a single ancestor (Ryko & Spiricheva 2020; Starostin 2021; Starostin 2022). Thus, UPGMA is preferable to NJ, which in rare cases gives a negative distance value between groups.

The resulting method is as close to a black box as possible. Table 4 compares Cipher-RF with a Levenshtein distance method by the criteria described in Section 1.

Criterion	Cipher-RF	Levenshtein distance
Transparency	Opaque	Transparent
Explainability	Unexplainable	Explainable
Object-awareness	Object-unaware	Object-aware
Reliability of transfer	Unreliably transferrable	Reliably transferrable
Predictability	Unpredictable	Predictable

Table 4. Comparison between Cipher-RF and Levenshtein distance.

I use two measurements for evaluation, the probability of a correct tree and the average split distance difference. As the datasets I use consist of only three lects, these metrics transform into correct outgroup identification probability and average inner split distance error.

Correct outgroup identification probability is a measure defined by the division of several runs when a model successfully predicted which lect was the most distant from the other two in the group, by an overall number of runs. Correct outgroup identification probability relies on existing classifications, and one should carefully apply it when the relationship between lects is debatable.

Average inner split distance loss is a measure that computes the difference between some pre-existing data on how early the two remaining lects split, and the results of the method under consideration. I score an average inner split distance only when the outgroup identification is correct. Average inner split distance loss requires another automatic language distance measurement method. This may cause some issues for the data that are traditionally problematic for automatic language distance measurement methods (Wichmann & Rama 2018), so one should apply the metric with extreme caution.

Cross-evaluation of the method includes three string similarity measures: two fully language-agnostic and one language-aware.

The first language-agnostic method I use is the Levenshtein distance normalised divided (LDND), a classic method for language distance measurement (Nerbonne & Heeringa 1997; Gooskens & Heeringa 2004; Holman et al. 2008). It consists of three steps. First, one scores additions, deletions, and substitutions (Levenshtein distance measurement) between two sequences. Then, one normalises them, dividing them by the size of the longest string in comparison. These two steps repeat for each concept in the Swadesh list, thus yielding a list of LDNs (Levenshtein distance normalised scores). After this, I score their mean value and thus acquire LDND.

The second implemented language-agnostic method is the weighted Jaro-Winkler distance normalised divided (WJWDND). Jaro-Winkler distance is a string similarity measure that resembles the Levenshtein distance, though it prioritises strings that have similar beginnings (Jaro 1989; Winkler 1990). The weighted Jaro-Winkler distance is a multiplication of the Jaro-Winkler distance by the Levenshtein distance, to get the best of two metrics (Gueddah et al. 2015). The normalisation and division parts remain the same as for LDND.

The language-aware method is the phonetics-aware Damerau-Levenshtein distance normalised divided (PADLDND). It works similar to LDND and WJWDND, with two key differences. It scores transpositions, which may be useful for metathesis detection, like Russian *vsjakij* – Croatian *svaki* ‘everyone among set’. PADLDND also multiplies each Damerau-Levenshtein distance score by a difference in the vectors of phonetic features of the symbols under consideration. There are different implementations of this metric with proven efficiency, but they are generally closed-source (Normanskaya 2020). I use the open-source implementation<sup>3</sup>.

Data cross-evaluation tactics include using two different datasets of Swadesh list items, one of them consisting of understudied closely related lects and the other one of somewhat more distantly related, but relatively better studied ones.

The former, consisting of East Slavic Swadesh wordlists, is the main one. For this dataset, I use every method that I have mentioned in this section up to this point: Cipher-RF, LDND, WJWDND, and PADLDND.

The latter dataset consists of Tuu (Taa subgroup) Swadesh wordlists. This dataset allows me to test whether the efficiency of language-agnostic methods gets closer to the one of black-box methods with the language distance increase. Here I use only Cipher-RF, LDND, and WJWDND. It feels safe – to some degree – to compare the automatic classification results for Taa with the current state-of-the-art human classification provided by Starostin (2022).

I also use a combined method and data cross-evaluation. This is a phonetics-aware Hamming distance (PAHD), a full-fledged alternative for a Swadesh list-based classification, based on calculating the Hamming distance between strings of phonetic features. PAHD does not analyse language units directly but deals with the results of human analysis.

The code that implements all the methods is available on GitHub<sup>4</sup>.

## Experiments and Analysis

The experiments run in four stages. I start by testing the black-box method via the correct out-group identification measurement on the East Slavic material. The method is a Random Forest classifier of vectorised ciphered data (Cipher-RF). The next stage is to cross-evaluate via language-agnostic and language-aware string similarity measures. I also combine method and data cross-evaluation with the phonetics-aware Hamming distance (PAHD). The basis for comparison is the average inner split distance error. For the final stage, we repeat the first and the second stages on Taa lects.

### Black-box method and East Slavic lects

As described in the section dedicated to methodology, using Cipher-RF includes 10 subsequent instances of lects re-classification after swapping concepts, which differ by the number of

<sup>3</sup> <https://github.com/Stoneberry/phonetic-algorithmIPA>.

<sup>4</sup> <https://github.com/The-One-Who-Speaks-and-Depicts/black-box>

swapped concepts (3, 5, 8, 11, 14), and the presence of borrowing processes emulation (when the concepts from one lect replace their counterparts in the second lect). Each setup, for statistical correctness, goes through 100 runs of random swaps. The results of the experiments are in Tables 5 and 6.

Number of swaps	3	5	8	11	14
Present borrowing	0.67	0.62	0.6	0.62	0.63
Non-present borrowing	0.69	0.62	0.52	0.5	0.48

Table 5. Correct outgroup identification probability for the East Slavic lects by Cipher-RF

Number of swaps	3	5	8	11	14
Present borrowing	0.001	0.001	0.002	0.002	0.002
Non-present borrowing	0.001	0.002	0.002	0.002	0.002

Table 6. Average correct inner split distance for East Slavic lects, Cipher-RF (only runs with correctly identified outgroup)

The correct outgroup identification here means that metric joins Belogornoje and Megra into a single group, while leaving Khislavichi as an outgroup, as presented in figure 2.

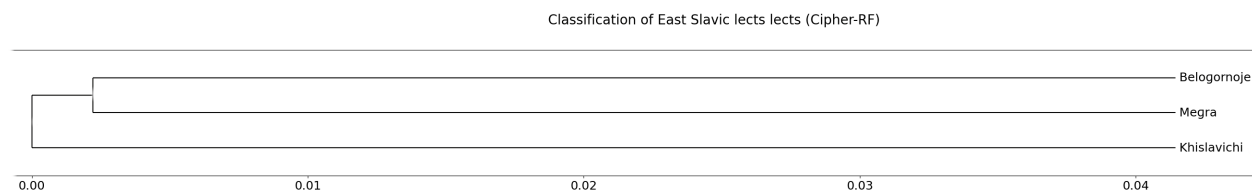


Figure 2. Correct classification of Belogornoje, Megra and Khislavichi with Cipher-RF.

While one may evaluate average inner split distance results only later, when compared to the other metrics, correct outgroup identification probability shows an overall low efficiency of Cipher-RF. Only seven setups out of 10 lead to significantly higher than a 50% chance of correct classification. There is no correlation between either the number of the swapped concepts or the presence of borrowings and the quality of Cipher-RF performance. It can be seen that the presence of borrowing makes the results more stable, even though lower for the lesser number of swaps. However, there is no way to explain it, as Cipher-RF is a black box. One just does not know what drives it and which data transformations are the most crucial.

### String similarity measures and East Slavic lects

The next step is to test string similarity measures against the same dataset. The Levenshtein distance normalised divided (LDND)-based classification and the weighted Jaro-Winkler distance normalised divided (WJWDND)-based classification are in figures 3 and 4 respectively.

Both language-agnostic string similarity measures demonstrate the correct outgroup prediction, placing Khislavichi quite far from the last common ancestor (LCA; Brower & Schuh 2021) of Megra and Belogornoje. One can also see the difference in the inner split: it is 0.039 to 0.04 for LDND and 0.031 to 0.032 for WJWDND. For Cipher-RF this value is generally almost ten times lower. The comparison is shown in Tables 7 and 8.

Levenshtein distance normalised divided classification of East Slavic lects

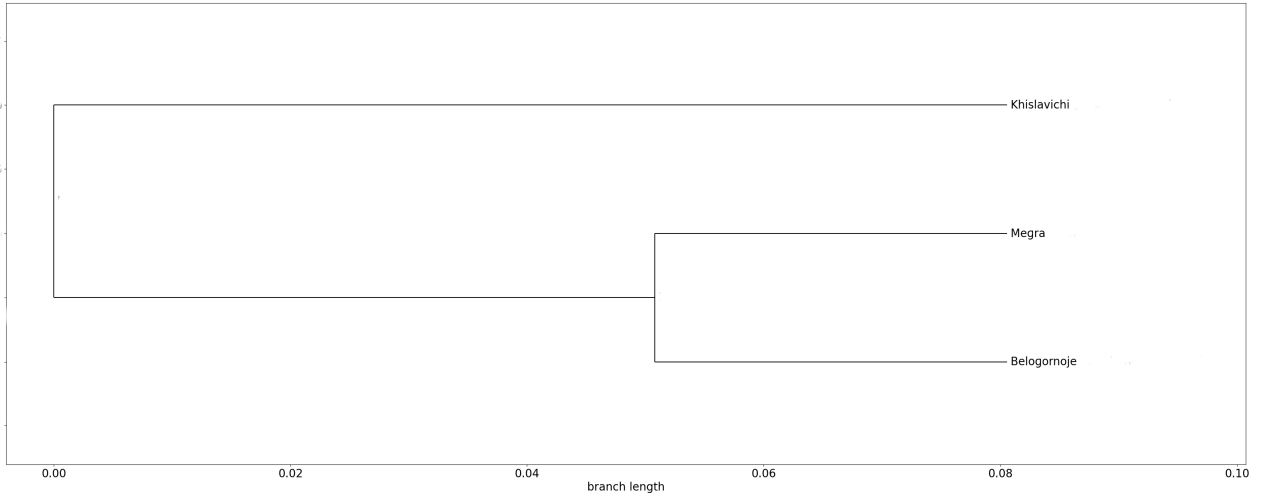


Figure 3. Classification of Belogornoje, Megra and Khislavichi with LDND.

Weighted Jaro-Winkler distance normalised divided classification of East Slavic lects

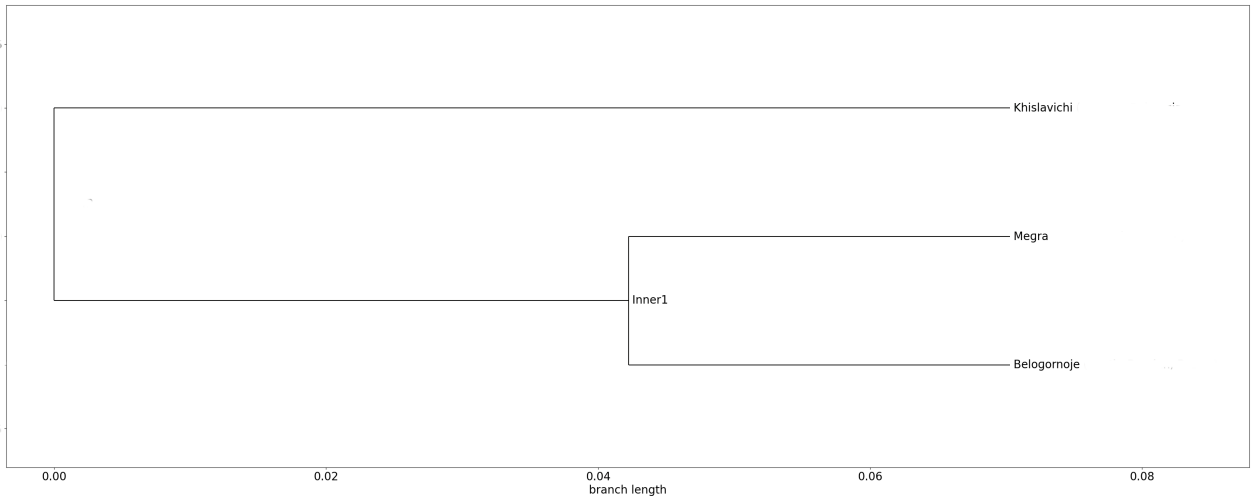


Figure 4. Classification of Belogornoje, Megra and Khislavichi with WJWDND.

Number of swaps	3	5	8	11	14
Present borrowing	0.04	0.04	0.039	0.039	0.039
Non-present borrowing	0.04	0.039	0.039	0.039	0.039

Table 7. Average inner split distance loss between Cipher-RF (only runs with correctly identified outgroup) and LDND.

Number of swaps	3	5	8	11	14
Present borrowing	0.032	0.032	0.032	0.031	0.032
Non-present borrowing	0.032	0.032	0.032	0.031	0.032

Table 8. Average inner split distance loss between Cipher-RF (only runs with correctly identified outgroup) and WJWDND.

For the difference in scores between Cipher-RF and string similarity measures, there are two different reasons. The first is the difference of scale: LDND and WJWDND may have values of 0 to 1, while Cipher-RF possesses a restriction of the squared difference between the gold score and the set up score (in this series of experiments, it is approximately 0.45). The second reason is that the distance between the LCA of Khislavichi, Belogornoje and Megra, and the LCA of Belogornoje and Megra is lesser for Cipher-RF classification than for LDND and WJWDND classification.

**Method cross-evaluation: language-aware string similarity measure and East Slavic lects**

Figure 5 shows the results of experiments based on a language-aware string similarity measure.

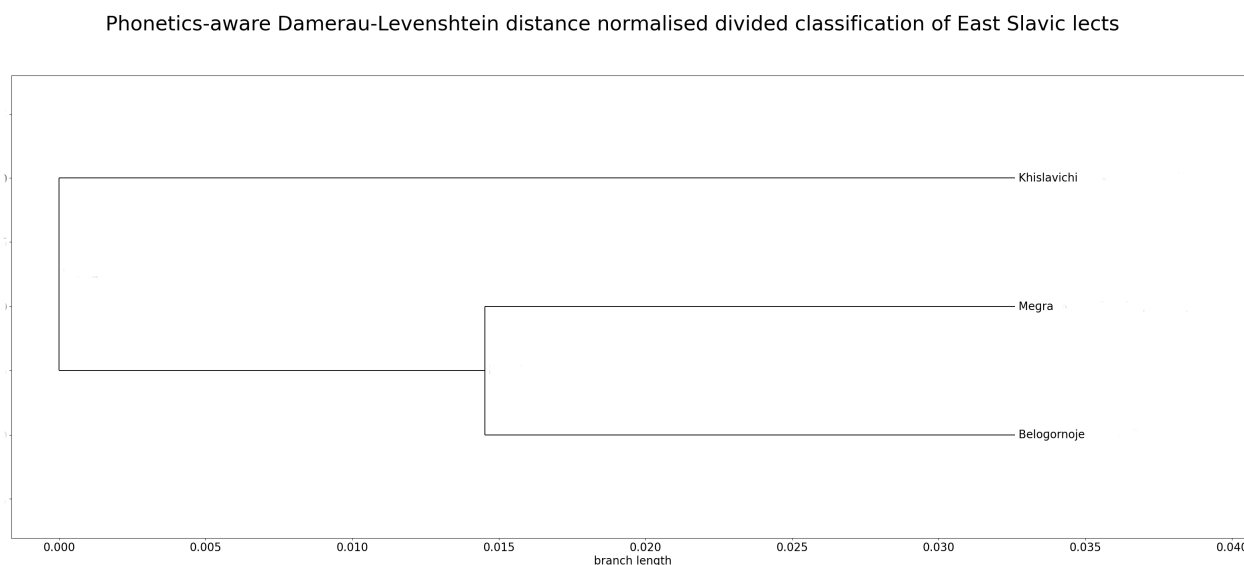


Figure 5. Classification of Belogornoje, Megra and Khislavichi with PADLDND.

The phonetics-aware string similarity measure demonstrates correct outgroup prediction in the same manner as the language-agnostic string similarity measures do. However, there are two differences. The first one is branch length: it is approximately two times smaller in LDND and WJWDND than in PADLDND. The second difference is more linguistically explainable. PADLDND considers Megra and Belogornoje to be much more closely related to each other than LDND and WJWDND both do. This correlates with coincidences and disagreements of some East Slavic phonetic phenomena manifestations between lects (strong presence of *okanje* in both Megra and Belogornoje, contrasted with *akanje* in Khislavichi), and proves that language awareness indeed helps string similarity measures to be more precise. However, the average inner split distance loss between PADLDND (0.011) and Cipher-RF is still significant, as one may see in Table 9.

Number of swaps	3	5	8	11	14
Present borrowing	0.012	0.011	0.011	0.011	0.011
Non-present borrowing	0.011	0.011	0.011	0.011	0.011

Table 9. Average inner split distance loss between Cipher-RF (only runs with correctly identified outgroup) and PADLDND.



Overall, these experiments show that there is a crucial difference between black-box and language-agnostic methods. Language-agnostic methods are much closer in their results to the language-aware methods of the same type, though they lack some necessary linguistic insight. However, their advantages and shortcomings are easily explainable, and any researcher with sufficient skills may attempt to maximise the former and minimise the latter by introducing language-aware features.

### Data cross-evaluation: the Taa lects

The question, however, remains: does this difference prevail on a bigger diachronic scale? I took the Taa lects as a representative of somewhat more distantly related lects. Figure 6 reproduces a classification of Taa lects by Starostin (2021; 2022)<sup>5</sup>.

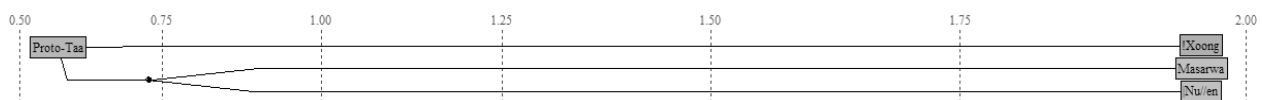


Figure 6. Classification of Taa lects.

The results of measuring language distance between Taa lects with Cipher-RF are in Tables 9 and 10.

Number of swaps	3	5	8	11	14
Present borrowing	0.44	0.43	0.32	0.41	0.36
Non-present borrowing	0.34	0.35	0.31	0.39	0.26

Table 9. Correct outgroup identification probability for the Taa lects by Cipher-RF.

Number of swaps	3	5	8	11	14
Present borrowing	0.005	0.006	0.006	0.005	0.006
Non-present borrowing	0.005	0.006	0.006	0.006	0.007

Table 10. Average correct inner split distance for the Taa lects, Cipher-RF (only runs with correctly identified outgroup).

Thus, Cipher-RF results are becoming worse with a language distance increase between the analysed lects. There are at least some (though unknown) pieces of linguistic information that influence the model, even after all the data transformations.

The average inner split distance is the same, which means that Cipher-RF implementation blocks our attempts to transfer language distance information into a precise absolute timing (i.e., how many years ago) of the split. As Figure 7 shows, the branch length for Taa is the same as for East Slavic. It may seem that the black-box method-produced graph is closer to the gold one than the language-agnostic string similarity measures-produced one. Cipher-RF, however, detects much fewer differences between lects than any other method.

Language-agnostic string similarity measures predictions remain correct. The length of branches also grows (0.087 and 0.078 for LDND and WJWDND), depending on the time passed, as is visible in Figures 8 and 9. It makes these methods more suitable for diachronic studies of language variation.

<sup>5</sup> Image taken from <https://starlingdb.org/images/xoo.png>.

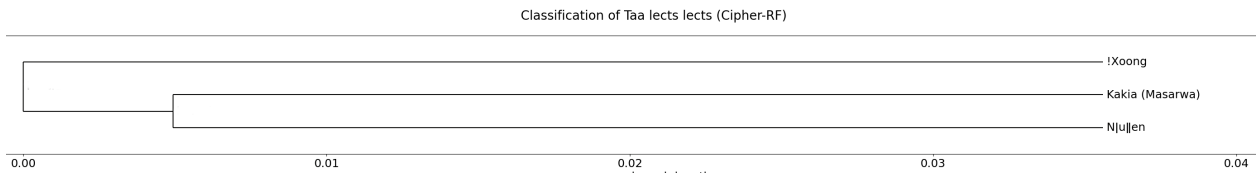


Figure 7. Correct classification of !Xóǒ, Kakia (Masarwa) and N|u||en with Cipher-RF.

The bigger distance also highlights the difference between LDND and WJWDND. LDND augments the differences between the lects under consideration, while WJWDND tends to smooth it by normalisation. It helps to distinguish the possible spheres of application for both methods: LDND for comparing closely related lects and WJWDND – for the distantly related ones.

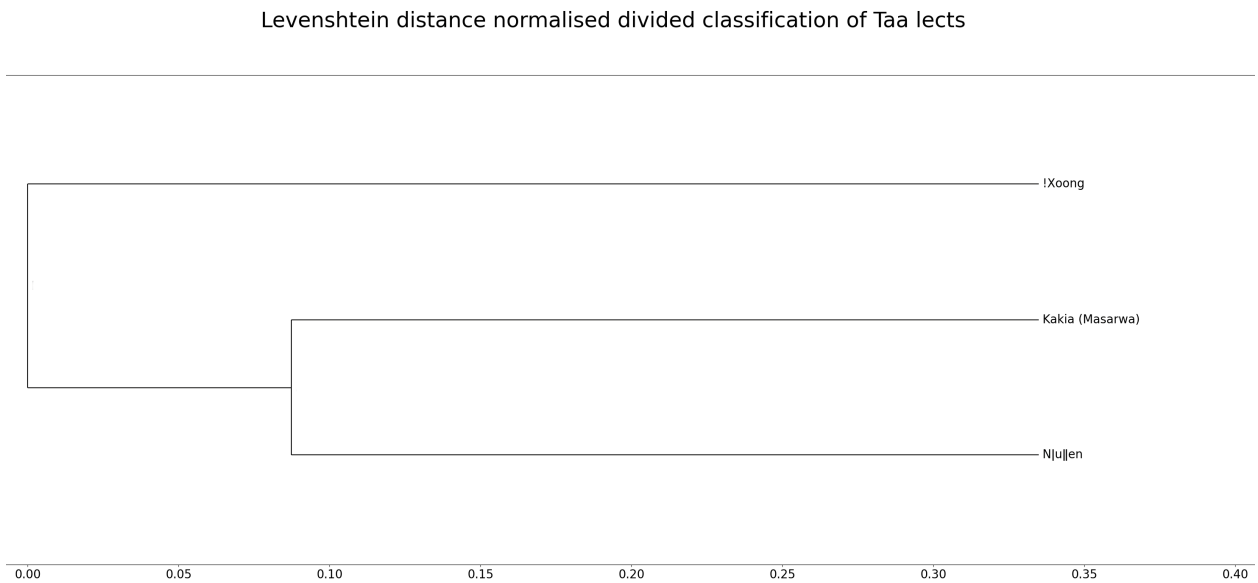


Figure 8. Classification of !Xóǒ, Kakia (Masarwa) and N|u||en with LDND.

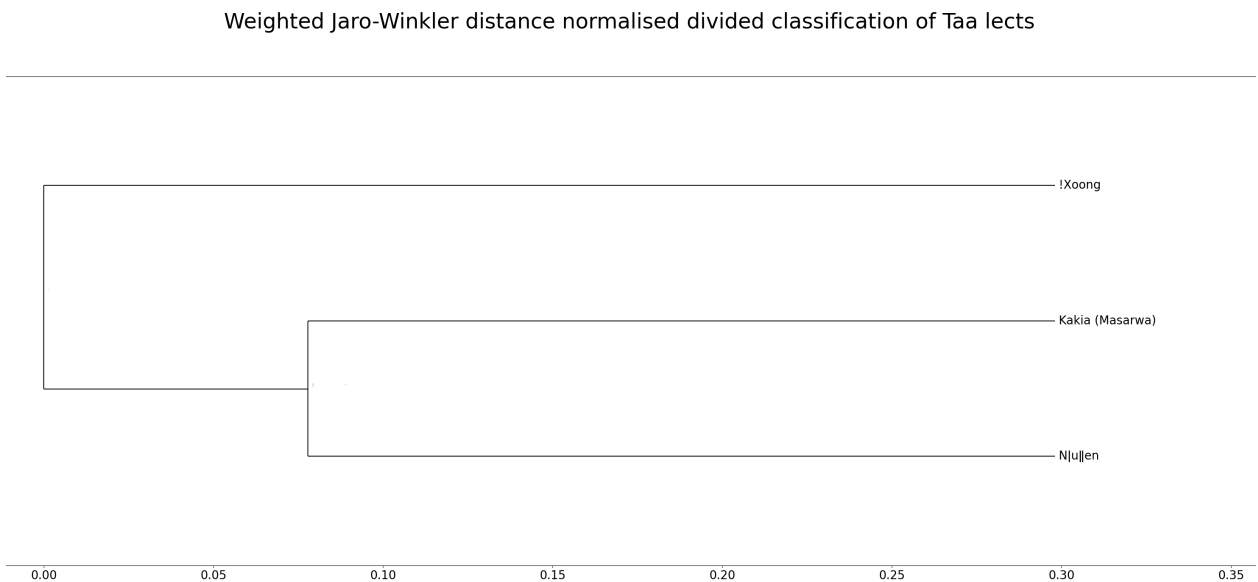


Figure 9. Classification of !Xóǒ, Kakia (Masarwa) and N|u||en with WJWDND.

Language-agnostic string similarity measures, yet again, demonstrate a greater ability to deal with language distance than Cipher-RF. The average inner split distance loss becomes more visible, as seen in Tables 11 and 12.

Number of swaps	3	5	8	11	14
Present borrowing	0.082	0.082	0.082	0.082	0.082
Non-present borrowing	0.082	0.081	0.081	0.081	0.081

Table 11. Average inner split distance loss between Cipher-RF (only runs with correctly identified outgroup) and LDND for Taa lects classification.

Number of swaps	3	5	8	11	14
Present borrowing	0.072	0.072	0.072	0.072	0.072
Non-present borrowing	0.072	0.072	0.072	0.072	0.071

Table 12. Average inner split distance loss between Cipher-RF (only runs with correctly identified outgroup) and WJWDND for the Taa lects classification.

Thus, the behaviour and efficiency of black-box and language-agnostic methods differ. While language-agnostic methods are still sensitive to the degree of language distance under consideration, black-box methods do not provide a researcher with any additional information on diachronic scope.

### Method and data cross-evaluation: phonetically aware Hamming distance

PAHD experiments (Figure 10) include scoring the Hamming distance on strings of phonetic features for Kritskovschina, Mokshenskaja and Piestchanka East Slavic lects. It is important to note that these lects are not direct representations of respective Khislavichi, Megra and Belogornoje phonetic states (Kritskovschina and Khislavichi belong to different, though phonetically close, dialect continua). Thus, the comparison is approximate.



Figure 10. Classification of Kritskovschina, Mokshenskaja and Piestchanka with PAHD.

The difference here is probably the most drastic one, as table 13 shows.

Number of swaps	3	5	8	11	14
Present borrowing	0.191	0.19	0.19	0.19	0.19
Non-present borrowing	0.19	0.19	0.19	0.19	0.19

Table 13. Average inner split distance loss between Cipher-RF (only runs with correctly identified outgroup) and PAHD for East Slavic lects classification.

PAHD separates the lects more strictly by a huge margin, the distances it calculates are larger than the ones of LDND and WJWDND classifications for Taa. Using only phonetic features, thus, magnifies the scale of differences by almost ten times.

This experiment again proves that Cipher-RF is the method that is unlikely to provide us with linguistic insight (or any insight, for that matter) on its decisions. It also shows that restricting the data to historically stable vocabulary smoothes out phonetic differences between lects, making methods more sustainable on a large scale while somewhat harming their sensitivity on a small scale.

PAHD agrees with LDND, WJWDND, and PADLDND on the division between Northern Belarusian/Western Russian and Northern/Southern Russian lects. It confirms the original gold presupposition of Khislavichi being an outgroup for Belogornoje and Megra, and thus reiterates the correctness of Cipher-RF evaluation.

## Conclusion

This research presents a new black-box method for historical comparative linguistics, Cipher-RF, based on a combination of hashing algorithms and a Random Forest Classifier. I have applied this method to East Slavic and Taa lects and showed that language-agnostic methods and black-box methods significantly differ in their behaviour and that it is crucial to distinguish between them, contrary to past papers on the topic, such as Prokić and Moran 2013. Black-box methods are less efficient for the purposes of historical comparative linguistics, but they provide a good baseline for other automatic methods to beat while being compared to an ideal classification. One should still use language-agnostic methods with a high degree of caution and always try to interpret their results linguistically. Their transparent structure allows for that.

The paper introduces a new corpus-based dataset of Swadesh-type lists for East Slavic lects of Khislavichi, Megra, and Belogornoje. The dataset consists of lexical units gathered from both open and unpublished corpora. The actual type of Swadesh list is ASJP; in the future, it is going to be expanded into a more classical 110-item one (Kassian et al. 2010).

The next step could be the introduction of new lects and new classification methods, especially when the material is presented by raw corpora. One more possible expansion is the automatic search for Swadesh list items in raw corpora.

## Acknowledgements

The author owes his reviewers for their insightful comments and discussion of the material. All the remaining errata are author's. The author is also grateful to the compilers of Saratov dialectological corpus, Khislavichi corpus, Russian dialectology atlas, and Global Lexicostatistical Database, without whom access to the research material would be impossible.

## References

- Afanasev, Ilija. 2023. The Use of Khislavichi Lect Morphological Tagging to Determine its Position in the East Slavic Group. In: *Tenth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2023)*: 174–186.
- Archangeli, Diana, Douglas Pulleyblank. 2022. *Emergent phonology*. Berlin: Language Science Press.
- Barannikova, Lidija. 2005. Govory territorij pozdnego zaselenija i problema ih klassifikacii. In: Valentin Goldin, Olga Kryuchkova (eds.). *Barannikova L. I. Obshhee jazykoznanie: izbrannye raboty*: 192–203. Moscow: KomKniga.

- Bastings, Jasmijn, Yonatan Belinkov, Yanai Elazar, Dieuwke Hupkes, Naomi Saphra, Sarah Wiegrefe. 2022. In: *Proceedings of the Fifth BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP*. Association for Computational Linguistics, Abu Dhabi, United Arab Emirates (Hybrid).
- Barić, Ana, Laura Majer, David Dukić, Marijana Grbeša-Zenzerović, Jan Snajder. 2023. Target Two Birds With One SToNe: Entity-Level Sentiment and Tone Analysis in Croatian News Headlines. In: *Proceedings of the 9th Workshop on Slavic Natural Language Processing 2023 (SlavicNLP 2023)*: 78–85.
- Brower, Andrew V. Z., Randall T. Schuh. 2021. *Biological Systematics: Principles and Applications*. Ithaca, NY: Cornell University Press.
- Burlak, Svetlana. 2021. Ustoichivost' i chastonost': jest' li korrel'acija? [Stability and frequency: is there a correlation?] *Journal of Language Relationship* 19(3-4): 293–307.
- Carvalho, Fernando O. de. 2020. Evaluation of cognation judgments undermines computational phylogeny of the Arawakan language family. *Journal of Language Relationship* 18(1-2): 87–110.
- Cho, Kyunghyun, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, Yoshua Bengio. 2014. Learning phrase representations using RNN encoder–decoder for statistical machine translation. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*: 1724–1734.
- Gage, Philip. 1994. A New Algorithm for Data Compression. *The C User Journal* 12(2): 23–38.
- Gooskens, Charlotte, Wilbert Heeringa. 2004. Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. *Language Variation and Change* 16: 189–207.
- Gueddah, Hicham, Abdellah Yousfi, Mostafa Belkasm. 2015. The filtered combination of the weighted edit distance and the Jaro-Winkler distance to improve spellchecking Arabic texts. In: *2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA)*: 1–6.
- Feld, Jan, Alexander Maxwell. 2019. Sampling error in lexicostatistical measurements: A Slavic case study. *Diachronica* 36(1): 100–120.
- Ho, Tin Kam. 1995. Random decision forests. In: *Proceedings of 3rd international conference on document analysis and recognition*: 278–282.
- Holman, Eric, Søren Wichmann, Cecil Brown, Viveka Velupillai, André Müller, Dik Bakker. 2008. Explorations in automated language classification. *Folia Linguistica* 42: 331–354.
- Jaro, Matthew A. 1989. Advances in record linkage methodology as applied to the 1985 census of Tampa Florida. *Journal of the American Statistical Association* 84: 414–420.
- Jäger, Gerhard. 2019. Computational historical linguistics. *Theoretical Linguistics* 45(3-4): 151–182.
- Kanjirangat, Vani, Tanja Samardžić, Ljiljana Dolamic, Fabio Rinaldi. 2023. Optimizing the Size of Subword Vocabularies in Dialect Classification. In: *Tenth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2023)*: 14–30.
- Kassian, Alexei, George Starostin, Anna Dybo, Vasiliy Chernov. 2010. The Swadesh wordlist. An attempt at semantic specification. *Journal of Language Relationship* 16: 46–89.
- Kryuchkova, Olga, Valentin Goldin. 2011. Corpus of Russian dialect speech: concept and parameters of evaluation. In: *Computational Linguistics and Intellectual Technologies. Proceedings of International Conference “Dialog–2011”*: 359–367.
- Kuparinen, Olli, Yves Scherrer. 2023. Dialect Representation Learning with Neural Dialect-to-Standard Normalization. In: *Tenth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2023)*: 200–212.
- Levenshtein, Vladimir I. 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady* 10(8): 707–710.
- Lewis, Mike, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, Luke Zettlemoyer. 2020. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*: 7871–7880.
- List, Johann-Mattis. 2014. *Sequence Comparison in Historical Linguistics*. Walter de Gruyter GmbH & Co KG.
- Marchenko, I. A., O. N. Dolgov, A. S. Azanova, M. S. Zambrzhitskaya, E. A. Zalivina, S. A. Zemlyanskaya, D. I. Mochul'skij, E. I. Tsejtina, D. G. Chistyakova, R. V. Ron'ko. Database of the Dialectological Atlas of the Russian Language. Available online at: <https://da.ruslang.ru/> [accessed: 18.10.2023].
- Munn, Michael, David Pitman. 2022. *Explainable AI for Practitioners*. O'Reilly Media, Inc.
- Nerbonne, John, Wilbert Heeringa. 1997. Measuring Dialect Distance Phonetically. *Proceedings of the EACL 1997*: 1–18.

- Normanskaya, Julija V. 2020. Komi-jaz'vinskij — dialekt komi-permjackogo ili otdel'nyj jazyk? *Ezhegodnik finno-ugorskih issledovanij* 4: 628–641.
- Novikova, Jekaterina, Ondřej Dušek, Amanda Cercas Curry, Verena Rieser. 2017. Why We Need New Evaluation Metrics for NLG. In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*: 2241–2252.
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel et al. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12: 2825–2830.
- Radford, A., J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever. 2019. *Language Models are Unsupervised Multitask Learners*. Available at: [https://d4mucfpksywv.cloudfront.net/better-language-models/language\\_models\\_are\\_unsupervised\\_multitask\\_learners.pdf](https://d4mucfpksywv.cloudfront.net/better-language-models/language_models_are_unsupervised_multitask_learners.pdf) [Accessed: 18.10.2023].
- Rama, Taraka, Lars Borin. 2015. Comparative Evaluation of String Similarity Measures for Automatic Language Classification. In: G. K. Mikros, J. Macutek (eds.). *Sequences in Language and Text*: 171–200. Berlin / München / Boston: De Gruyter Mouton.
- Ryko, Anastasiya I., Margarita S. Spiricheva. 2020. Corpus of the Russian dialect spoken in Khislavichi district. Available online at: <http://lingconlab.ru/khislavichi/> [accessed: 18.10.2023].
- Ryko, Anastasiya I., Margarita S. Spiricheva. 2022. The Degree of Preservation of Dialectal Features in Different Generations (Khislavichi District of the Smolensk Region). *RSUH/RGGU Bulletin. "Literary Theory. Linguistics. Cultural Studies" Series* 5: 121–141.
- Snoek, Connor. 2013. Using semantically restricted word-lists to investigate relationships among Athapaskan languages. In: Lars Borin, Aniu Saxena (eds.). *Approaches to Measuring Linguistic Differences*: 231–248. Boston/Berlin: Walter De Gruyter GmbH.
- Sokal, Robert Reuven, Charles Duncan Michener. 1958. A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin* 38:1409–1438.
- Starostin, George. 2011. On Mimi. *Journal of Language Relationship* 6(1): 115–140.
- Starostin, George. 2021. Lexicostatistical studies in Khoisan II/1: How to make a Swadesh wordlist for Proto-Tuu. *Journal of Language Relationship* 19(1-2): 99–135.
- Starostin, George. 2022. Lexicostatistical studies in Khoisan II/2: Towards a more precise phylogeny for the Tuu family. *Journal of Language Relationship* 20(1-2): 25–70.
- Sutskever, Ilya, Oriol Vinyals, Quoc V Le. 2014. Sequence to sequence learning with neural networks. *Advances in neural information processing systems*: 3104–3112.
- Syvokon, Oleksiy, Olena Nahorna, Pavlo Kuchmiichuk, Nastasiia Osidach. 2023. UA-GEC: Grammatical Error Correction and Fluency Corpus for the Ukrainian Language. In: *Proceedings of the Second Ukrainian Natural Language Processing Workshop (UNLP)*: 96–102. 1 Linguistics.
- Vajda, Edward. 2012. The Dene-Yeniseian connection: a reply to G. Starostin. *Journal of Language Relationship* 8(1): 138–152.
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin. 2017. Attention Is All You Need. *Advances in neural information processing systems* 30: 1–11.
- Wichmann, Søren, Eric W. Holman, Dietrich Stauffer, Cecil H. Brown. 2011. Similarities among languages of the Americas: An exploration of the WALS evidence. *Journal of Language Relationship* 5: 130–134.
- Wichmann, Søren, Taraka Rama. 2018. Jackknifing the Black Sheep: ASJP Classification Performance and Austronesian. In: Ritsuko Kikusawa, Lawrence A. Reid (eds.). *Let's Talk about Trees: Genetic Relationships of Languages and Their Phylogenetic Representation*: 39–58. Osaka: National Museum of Ethnology, Japan.
- Winkler, William E. 1990. String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage. In: *Proceedings of the Section on Survey Research Methods*: 354–359.
- Zhivlov, Mikhail. 2021. Does the comparative method work in New Caledonia? *Proceedings of the 15th International Conference on Austronesian Linguistics, Olomouc*. Available online at: <https://www.hse.ru/data/2021/07/04/432013048/Comparative%20method.pdf> [accessed: 01.09.2023].

И. Афанасьев. Потенциал систем — «чёрных ящиков» для автоматического измерения расстояния между восточнославянскими лектами.

Активное развитие новых количественных методов в современной исторической лингвистике в 2000-е — 2010-е годы актуализировало проблему невозможности адек-

ватной интерпретации подобных методов и поиск путей преодоления данной проблемы. Задачей данной статьи является демонстрация преимуществ систем, которые по умолчанию обладают прозрачностью для исследователя.

В работе сравниваются два типа систем, измеряющих языковое расстояние и используемых для задач внутренней генетической классификации. Механизм действия систем — «чёрных ящиков» предполагает обработку исходных данных и представление результата максимально непрозрачным как для исследователя, так и для автоматических методов анализа. Напротив, независимые от частных языковых свойств методы (к примеру, меры сходства строк) анализируют данные прозрачным образом, но не учитывают особенностей конкретных языков. Для сравнения систем — «чёрных ящиков» с существующими независимыми от частных языковых свойств методами в данной статье предлагается новая непрозрачная система, основанная на хешировании, векторизации и имитации языкового контакта.

В статье использован восточнославянский материал (лексический и грамматический), а также материал группы таа (койсанский языковой ареал Южной Африки). Восточнославянские данные состоят из корпусов говоров с. Белогорное, д. Мегра и с. Хиславичи, а также списков фонетических особенностей говоров д. Мокшенская, д. Крицковщина и с. Песчанка. Данные таа представлены списками Сводеша для къхонг, масарва и н|у||ен. Важным вкладом работы является публикация новых списков Сводеша для ряда восточнославянских диалектов.

*Ключевые слова:* «чёрный ящик»; восточнославянские языки; южнокойсанские языки; языки туу; независимые от частных языковых свойств методы; автоматическое измерение языковой дистанции; автоматическая классификация; меры сходства строк; базисная лексика.

## Proto-Indo-European Nom.-Acc. Dual and the Germanic Dual of nouns

In this article, I present a brief outline of Proto-Indo-European endings of nouns in the nom.-acc. dual and discuss the Germanic noun for “breast”, which supposedly had dual forms in Proto-Germanic. OIcel. *brjóst* n. and OE *brēost* n. may reflect the dual *\*breust-ō* (< PIE *\*-ō*, the dual of thematic nouns). Because of the homonymy with the neuter nom.-acc. plural in *-ō* (< PIE *\*-ā*), this form was reanalysed as n. pl. and became a thematic neuter (like Goth. *daur* ‘door’). Goth. *brusts* directly reflects PGerm. *\*brust-iz*, the athematic plural (like OIcel. *dyrr*). OSw. *bryst* n. is ambiguous because it can reflect PGerm. *\*breust-ō* (dual and pl.), *\*brust-ī* (dual), *\*brust-iz* (pl.). The Proto-Germanic “breast” could originally be a proterokinetic noun with an ablauting root, *\*breust-/\*brust-*. As the inflexional expression of duality is extinct, nouns which denote paired organs can develop alternative ways to express it. The article describes such nouns in Swedish dialects of Estonia.

*Keywords:* nominative and accusative dual; Proto-Indo-European; Proto-Germanic.

### Indo-European Nom. and Acc. Dual

The specific ending of nom.-acc. dual in Proto-Indo-European is reconstructed as *\*-h<sub>1</sub>* (Fortson 2010: 117; cf. Malzahn 1999). It is found in athematic masculine and feminine nouns in Greek, e.g. *πόδες, μητέρες*, with *-ε* from the vocalised *\*-h<sub>1</sub>*. As for athematic masculine and feminine nouns in Sanskrit, they have *-ā* (and *-āu*), which was probably taken from the thematic declension (Brugmann 1911: § 201). Athematic neuter nouns have *-ī* (e.g. Skt. *čakṣuṣī*, of *čakṣus-* n. ‘eye’), for which PIE *\*-īh<sub>1</sub>* is reconstructed (Fortson: *ibid.*). The laryngeal *\*-h<sub>1</sub>* in *\*-īh<sub>1</sub>* looks the same as in masculines and feminines, but the element *ī* is unclear. As there are very few neuter root nouns (Schindler 1972: 8), it is difficult to reconstruct their original dual form. It is therefore tempting to explain *-ī* in athematic neuters as non-original and taken from the dual of *i*-stems. These have *-ī* < *\*-īh<sub>1</sub>* (*\*h<sub>1</sub>* does not colour *i*), e.g. Skt. *munī, rātrī* of *muni-* m. ‘sage’, *rātri-* f. ‘night’. *ū*-stems are parallel, with *-ū* < *\*-ūh<sub>1</sub>*, e.g. Skt. *tarū, dhenū* of *taru-* m. ‘tree’, *dhenu-* f. ‘cow’, though the neuter has *-ī*: Ved. *urū-ī*, of *urū-* ‘wide’ (Macdonnell 1910: 297).

Thematic masculine nouns add *\*-h<sub>1</sub>* to the thematic vowel *\*-o-* (Fortson 2020: 126, 128), hence Gr. *-ω*, e.g. *λύκω* of *λύκος* ‘wolf’, Skt. (Vedic) *-ā*, e.g. *priyā* of *priyā-* ‘dear’, OCS *-a*, e.g. *raba* of *rabŭ* ‘slave’. Thematic neuter nouns have the same ending in Greek: *τέκνω* (of *τέκνον* ‘child’). Sanskrit and Old Church Slavonic reflect a different ending, *\*-oi*, probably from *\*-o-ī*, with *\*-ī* that may ultimately have been taken from *i*-stems: Skt. *priyé*, OCS *selě*, of *selo* ‘village’. It is difficult to judge whether Gr. *-ω* in thematic neuters is original (PIE *\*-o-h<sub>1</sub>*) or was taken from thematic masculines having replaced some other, specifically neuter, dual ending; cf. Chantraine 1984: § 22.

*ā*-stems behave identically to thematic neuters. In Greek they have *-ā*, e.g. *ἡμέρᾱ*, *Ἀτρείδᾱ*, (of *ἡμέρᾱ* ‘day’, *Ἀτρείδης*), which may go back to *\*-ā-h<sub>1</sub>*. In Sanskrit and Old Church Slavonic they reflect *-ī* added to the stem vowel *ā*, i.e. *\*-ā-ī* > *\*-ai*: Skt. *dve*, OCS *ženě*, of *žena* ‘wife’. An alternative explanation is that *-ā* in Greek first appeared in masculine *ā*-stems by



analogy with the dual in  $-\omega$  of masculine  $o$ -stems and then replaced the original dual ending  $*-\alpha\iota$  (Brugmann 1911: § 197). The latter came to be used as the nom. pl., having replaced the old  $*-\bar{\alpha}\varsigma$  (Chantraine 1984: § 33). This explanation presumes that the ending  $*-ai$  in the feminine nom.-acc. dual is original, though this is not granted.

The ending  $-au$  (i.e.  $-\bar{a}u$ ) of masculine nouns in Sanskrit does not fit into this system of  $*-h_1$  combined with the preceding stem vowels.<sup>1</sup> It can be explained as the regular dual ending  $-\bar{a}$  ( $< *-\bar{o}-h_1$ ) with the dual personal pronoun  $*\bar{u}\bar{e}$ - ‘we two’<sup>2</sup> used enclitically, i.e.  $*-\bar{o}-h_1-\bar{u}\bar{e} > *\bar{o}\bar{u}$ .

### Remnants of the inflexional expression of duality in Germanic

In historical Germanic languages, the dual of nouns is extinct and is reconstructed only through indirect evidence of several forms, one of which is probably the word for “breast”, e.g. Goth. *brusts* τὸ στῆθος, a feminine root noun, attested only in acc. pl. (plurale tantum?): *motareis... sloh in brusts seinos* ἔτυπτεν τὸ στῆθος αὐτοῦ (Luke 18:13), OHG *brust* f. (consonant stem and  $i$ -stem), OE *brēost* n., OIcel. *brjóst* n., OSw. *bryst* n., Sw. *bröst*; for a full list of cognates see EWAhd II: 399. Gothic and Old High German reflect the PGerm. zero grade, Old English and Icelandic reflect  $*eu$ , see Kroonen 2013: 76, 80. The Old Swedish form is ambiguous and allows for the following explanations. First, it may reflect  $*br\bar{y}st < *bri\bar{u}st$  with  $\bar{y}$  shortened before  $st$  (Kock 1906: 467)  $< *breust-$ ; in this case it fully corresponds to Icel. *brjóst* (PGerm.  $*eu$  after  $r$  gives  $j\acute{o}$  in Icelandic and  $i\bar{u} > \bar{y}$  in Old Swedish, cf. OSw. *br\bar{y}ta* ‘to break’ and Icel. *brj\acute{o}ta* ‘id.’). Alternatively, *bryst* could appear through  $i$ -umlaut in the plural of an athematic feminine noun with a zero-grade root (PScand.  $*brust-iR$ , cf. Goth. *brusts*),<sup>3</sup> then  $y$  spread onto the whole paradigm; OSw.  $\bar{y}$  is regularly reflected as Sw.  $\ddot{o}$ , hence *bröst*. The neuter gender is in this case secondary. In Swedish, there is a number of originally feminine root nouns with  $i$ -umlaut in the root which in the contemporary language became neuter, e.g. OSw. *stup* ‘prop; pole’, *m\bar{u}s* ‘mouse’ (fem. root nouns) and Sw. *st\ddot{o}d*, *m\ddot{o}ss*, which are neuter.<sup>4</sup> Finally, OSw. *bryst* may reflect a Proto-Germanic dual form. An idea about the dual of this noun was originally proposed by Fr. Kluge (1882: 510). He reconstructs PGerm.  $*breust\bar{o}$ , where  $-\bar{o} =$  Vedic  $-\bar{a}$ ; this hypothesis was supported by H. Wagner (1956). They assume that PIE  $*-e$  in the dual of athematic masculines and feminines was replaced with  $*-\bar{o}$  (from thematic stems) not only in Sanskrit, but also in Proto-Germanic. B. Kahle (1887: 38) claims that Kluge’s idea about the dual is untenable because the ending should have been  $*-e$  rather than  $*-\bar{o}$ . This objection is not convincing because we can hardly expect that Proto-Germanic should strictly correspond to Proto-Indo-European in the distribution of dual endings, which to a certain extent was reshaped in many IE languages, including Sanskrit.

<sup>1</sup> In Rigveda,  $-\bar{a}$  and  $-au$  came to be in complementary distribution which is determined phonetically:  $-\bar{a}$  occurs before consonants, in pausa at the end of a pāda, or within a pāda in coalescence with a following vowel,  $-au$  occurs only before vowels as  $-\bar{a}v$ , which removes a hiatus; the ending  $-\bar{a}$  is much more frequent (Macdonnell 1910: 258).

<sup>2</sup> OCS  $\bar{v}\bar{e}$  ‘we two’, Lith. *v\bar{e}du* ‘id.’  $< *ue-du\bar{o}$  (Pokorny 1959: 1114).

<sup>3</sup> This means that the zero grade of the root of this word possibly could exist not only in Gothic and West Germanic, but also in Scandinavian; cf. Schwarz 1951: 131.

<sup>4</sup> Such root nouns with roots terminating in  $n$  and  $s$  assimilated the plural ending, e.g. OSw. *m\bar{u}s* ‘mouse’, pl. *myss*  $< *m\bar{y}ss < *m\bar{y}s-R < *m\bar{u}s-iR$  (Noreen 1904: §§ 433, 238.5). Such plural forms as *myss* were perceived as having a zero ending; since the zero ending in the plural is typical of neuter nouns in Scandinavian languages, root nouns with an assimilation in the plural could become neuter. Sw. *m\ddot{o}ss* ( $< OSw. myss$ ) occurs as a neuter form in the singular, but is less common than the non-neuter *mus* (SAOB: M 1625). Sw. *st\ddot{o}d* is only neuter in the contemporary language. For more examples see Wessén 1965: 101.

The dual ending  $*-\bar{o}$  in Proto-Germanic may be evidenced by OE *duru* ‘door’ (as an object with **two** sides) and *nosu* ‘nose’, feminine *u*-stems, which were probably dual forms with  $*-\bar{o}$  in Proto-Germanic (Kluge 1882: 506ff.; Griepentrog 1995). As the dual ending  $*-\bar{o}$  coincided with the neuter nom.-acc. pl.  $*-\bar{o}$  (PIE  $*-\bar{a}$ ), the noun “door” together with “breast” became neuter, cf. Goth. *daur* n., Icel. *brjóst* n. The dual ending  $*-\bar{i}$  in Proto-Germanic cannot be excluded either; Hultman (1894: § 9.13) explains OSw. *bryst* as an originally neuter *i*-stem with the dual  $-\bar{i}$ , i.e.  $*brust-\bar{i}$ . Though it cannot be proved that it used to be a neuter *i*-stem, the root vowel *y* can really be due to the *i*-umlaut caused by  $*\bar{i}$ .<sup>5</sup> In favour of the ending  $*-\bar{i}$  which was used alongside  $*-\bar{o}$  speaks the fact that “door” became an *i*-stem in Old High German, *turi* f. (Braune, Reiffenstein 2004: § 220). On the PGerm. level, the ending  $*-\bar{i}$  was probably more convenient than  $*-\bar{o}$  because it excluded the homonymy with  $*-\bar{o}$  (< PIE  $*-\bar{a}$ ) in the plural of neuter nouns.

The morphological relationship between Goth. *brusts* f. pl. and Icel. *brjóst* n., OE *brēost* n. looks identical to Olcel. *dyrr* f. pl. ‘door’ (also n. pl., Noreen 1923: § 416, anm. 4) and Goth. *daur* n. (Wagner 1956: 178), which is an argument in favour of the originally dual form of  $*breust-$ .

As for the phonetic relationship between the stems  $*breust-$  and  $*brust-$ , it looks parallel to Ved. *dvārah* ( $*dhu\acute{o}r-$ , nom. pl. with stress on the root) and acc. pl. *durāh* ( $*dhur-$  with stress on the ending), see Wagner 1956: 178. PGerm.  $*breust-$  can therefore be reconstructed as an originally proterokinetic root noun with ablaut in the root, i.e.  $*bréust-$  in the nominative (> Icel. *brjóst* etc.) and  $*brust-$  in oblique cases (> Goth. *brusts*); cf. Ringe 2006: 198.

Difficulties in the reconstruction of the Proto-Germanic paradigm of this noun are also due to the unclear element *t*, which in all likelihood did not originally belong to the root. PGerm.  $*breust-/brust-$  is regarded to be a derivative of PIE  $*bhreus-$  ‘schellen; sprießen’, cf. Rus. *брюхо* ‘belly’ (IEW 170; Orel 2003: 56, 59; Kroonen 2013: 76). O. Trubachev compares PGerm.  $*breust-$  with Polish dial. *bżusc* ‘calf (of the leg)’ (Pslav.  $*br’ustv$ ), Upper Sorbian *bristej* f., gen. *-stwje* ‘id.’, *bristwja* f. ‘id.’, *bristw(j)o* f. ‘id.’, Pol. dial. *brzustwa* ‘id.’ ( $*br’ustv\bar{v}$ ), Upper Sorbian *briščo* n. ‘id.’, Pol. dial. *bżusce* ‘id.’ ( $*br’ustvje$ ), Czech *bříštec* m. ‘finger pad’, Pol. dial. *brześć* ‘id.’ ( $*br’ustv\bar{c}$ ; ЭССЯ 3: 34–35). The zero-grade form  $*brust-$  is compared with Serb. *бѣцм* m. ‘young shoots’, Ukr. *бpocм* m. ‘bud’ ( $*br\bar{v}stv$ ), Bulg. dial. *бpвc* f. ‘tender shoots used to feed goats and sheep in winter’, Rus. dial. *бpocмb* f. ‘unfolded buds of bushes’ ( $*br\bar{v}stv$ ; ЭССЯ 3: 57–58), though it is also possible that these forms are related to  $*br\bar{v}sati$ ,  $*br\bar{v}sn\bar{v}ti$  (e.g. Church-Sl. *br\bar{v}sn\bar{v}ti* ‘to shave’), with the development ‘something picked, plucked’ > ‘buds, shoots, green leaves’ (ЭССЯ 3: 58).

Alternatively,  $*breust-/brust-$  may be compared with the word for “heart”, which is also a proterokinetic noun, PIE  $*k’erd$  nom.,  $*k’rd-ós$  gen. (PGerm.  $*hert\bar{v}n-$  n.). Morphological and semantic affinity of these words could be the reason for the analogical appearance of  $t < *d$  in PGerm.  $*breus-$ .

### Examples of non-inflexional expression of duality in Germanic dialects

Instead of the inflexional expression of duality, Germanic dialects have developed alternative ways to express it. For example, in Swedish dialects of Estonia a contamination with the numeral “two” and a reduplication of the root took place in nouns that designate paired organs. The noun for “temple” (part of the head) is known in the following forms in these dialects: *tinnij* in the dialect of the island of Runö (Vendell 1882–1887: 141); *finijng* in the dialect of the island of Ormsö (Tiberg Ob.;  $f < tv-$ , cf. *f\bar{o}* ‘two’ in this dialect); *twining* or *twiniing* in the village

<sup>5</sup> There could, however, be an alternative cause of *i*-umlaut in this form, namely the plural ending  $*-iR$ .

of Vippal (Rußwurm 1855: 323; Freudenthal, Vendell 1886: 243); *tvining* or *tvinniŋ* in the dialect of the islands of Rågöarna (Freudenthal, Vendell 1886: 243; Tiberg Ob.); *tfiniŋ* or *tvinniŋ* in the dialect of village of Gammalsvenskby (Freudenthal, Vendell: 231; Tiberg Ob.).

The form *tinniŋ* recorded in the dialect of Runö directly corresponds to Sw. *tinning* (OSw. *thynning*, *thinning*),<sup>6</sup> while the forms with *tv-* and *tf-* are due to a contamination with *tfō*, *tvō* ‘two’ (Hellquist 1922: 975).<sup>7</sup> I have also recorded *tfinniŋ* during my fieldwork in the village of Gammalsvenskby. In the dialect of the island of Dagö this noun was transformed into *filing* (*f* < *tv-*), thus becoming homonymous with the word for “twin”, cf. Sw. *tvilling* (Tiberg Ob.).

A contamination with “two” is also found in the present-day dialect of Gammalsvenskby in *tfinn* ‘fin’ (of fishes), which occurs alongside *finn* (cf. MLG *vinne* f. ‘id.’ and OSw. *finā* ‘id.’). A similar influence of the numeral “two” is believed to have caused *d-* instead of the phonetically regular *dh-* in Skt. *dvār-* ‘door’ (Mayrhofer 1992: 765).

As for reduplication, I suppose that it took place in the word for “gill” (of fish; Sw. *gäl*) in the dialect of Gammalsvenskby. I have recorded *gäigöl*, pl. *gäiglar*, m. (*l* is a retroflex flap). The same form is found in previous descriptions: *gäigäl*, pl. *gäiglar*, m. (Freudenthal, Vendell 1886: 75); *gäigöl* m. (Karlgrén 1964: 65); *gäigəl* or *gäigäl* (Tiberg Ob.). Other Swedish dialects of Estonia have forms that directly correspond to Sw. *gäl*: cf. *gäil* in the dialects of Dagö, Rågöarna, Vippal (Freudenthal, Vendell 1886: 64), *gail* in the dialects of Nuckö and Ormsö (Danell 1905–1934: 151). A. Karlgrén (ibid.) suggested that *gäigöl* was borrowed from MLG *gegel* n. and m. ‘palate; gum’ (Schiller & Lübben 2: 5), but in this case it is strange that this borrowing was recorded only in Gammalsvenskby. However, it cannot be excluded that a contamination with MLG *gegel* took place, which was triggered by the reduplicative shape of this form. It should be noted that *gäiglar* resembles the reduplication in PIE *\*k<sup>w</sup>e-k<sup>w</sup>l-o-* ‘wheel’, where it may express the iterativity of turning (Fortson 2010: 130). However, both contamination with “two” and reduplication are by no means regular in Swedish dialects of Estonia and are only confined to separate forms.

## References

- ЭССЯ = Trubachev, O. N. 1976. *Etimologičeskij slovar’ slav’anskix jazykov*. Vol. 3. Moskva: Nauka.
- Braune, Wilhelm. 2004. *Althochdeutsche Grammatik. I. Laut- und Formenlehre*. 15. Aufl., bearb. von Ingo Reiffenstein. Tübingen: Max Niemeyer.
- Brugmann, Karl. 1911. *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*. Vol. 2, pt. 1. 2<sup>nd</sup> ed. Strassburg: Karl J. Trübner.
- Chantraine, Pierre. 1984. *Morphologie historique du grec*. 2<sup>nd</sup> ed. Paris: Klincksieck.
- Danell, Gideon. 1905–1934. *Nuckömålet*. Stockholm: Kungl. boktryckeriet P.A. Norstedt & söner.
- EWAhd = Lloyd, Albert L., Rosemarie Lühr, Otto Springer (eds). 1998. *Etymologisches Wörterbuch des Althochdeutschen*. Vol. II. Göttingen / Zürich: Vandenhoeck & Ruprecht.
- Fortson, Benjamin W. 2004. *Indo-European Language and Culture: An Introduction*. 2<sup>nd</sup> ed. Oxford: Blackwell.

<sup>6</sup> This noun is originally a compound, cf. OIcel. *þunn-vangi* m. ‘temple’, lit. “thin cheek”. OSw. *y* in *thynning* cannot be due to *i*-umlaut (because *i* in the unstressed syllable reflects *\*a* here, see Noreen 1904: § 62.3); it probably goes back to OSw. *þynder*, a doublet of *þunder* ‘thin’, with *y* from the comparative degree (ibid.: § 468). As for *a* > *i* in the second element of the compound, cf. Sw. *köping* ‘market town’ < *\*kaup-angra-* “trade meadow” (Olson 1916: 259). The vowel *i* instead of *y* in *thinning* is due to folk etymology, namely the association with *tinne* ‘pinnacle’; an attempt to explain the change *y* > *i* by phonetic reasons was made by Kock (1906: 444, § 525).

<sup>7</sup> Another explanation of *tf-*, *tv-* in this word in Swedish dialects of Estonia is that it is due to a metathesis of *v* in *\*þunn-vang-* (Danell 1905–1934: 171); this is improbable. It should be noted that forms with *tv-* are also attested in Old Swedish, e.g. *tvinningen* acc. sg. with a suffixed article *-en* (Söderwall II: 757).

- Freudenthal, Axel, Herman Vendell. 1886. *Ordbok öfver estländsk-svenska dialekterna*. Helsinki: Tidnings- & tryckeri-aktiebolagets tryckeri. [Reprinted in 1982 by the Society *Svenska Odlingens Vänner*.]
- Griepentrog, Wolfgang. 1995. *Die Wurzelnomina des Germanischen und ihre Vorgeschichte*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Hellquist, Elof. 1922. *Svensk etymologisk ordbok*. Lund: C.W.K. Gleerups förlag.
- Hultman, Oskar Fredrik. 1894. *De östsvenska dialekterna*. Helsinki: Tidnings- & Tryckeri-Aktiebolagets tryckeri.
- IEW = Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch*. Bern: Francke.
- Kahle, Bernhard. 1887. *Zur Entwicklung der consonantischen Declination im Germanischen*. Berlin: Haude- & Spener-sche Buchhandlung (F. Weidling).
- Karlgren, Anton. 1964. *Ordbok öfver Gammalsvenskbymålet*. Manuscript kept at the Swedish Institute for Language and Folklore (Institutet för språk och folkminnen), Uppsala (№ 25555).
- Kluge, Friedrich. 1882. Sprachhistorische Miscellen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 1882: 506–539.
- Kock, Axel. 1906. *Svensk ljudhistoria, I*. Lund: C.W.K. Gleerup; Leipzig: Otto Harrassowitz.
- Kroonen, Guus. 2013. *Etymological Dictionary of Proto-Germanic*. Leiden / Boston: Brill.
- Macdonnel, Arthur Anthony. 1910. *Vedic Grammar*. Strassburg: Karl J. Trübner.
- Malzahn, Melanie. 1999. Die nominalen Flexionsendungen des idg. Duals. *Historische Sprachforschung* 112/2: 204–226.
- Mayrhofer, Manfred. 1992. *Etymologisches Wörterbuch des Altindoarischen. I*. Heidelberg: Carl Winter.
- Noreen, Adolf. 1904. *Altschwedische Grammatik mit Einschluss des Altgutnischen*. Halle: Niemeyer.
- Noreen, Adolf. 1923. *Altisländische und altnorwegische Grammatik*. Halle: Niemeyer.
- Olson, Emil. 1916. *De appellativa substantivens bildning i fornsvenskan. Bidrag till den fornsvenska ordbildningsläran*. Lund: C.W.K. Gleerup.
- Orel, Vladimir. 2003. *A Handbook of Germanic Etymology*. Leiden / Boston: Brill.
- Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch*. Bern / München: Francke.
- Ringe, Don. 2006. *A Linguistic History of English. Vol. 1. From Proto-Indo-European to Proto-Germanic*. Oxford University Press.
- Rußwurm, Carl. 1855. *Eibofolke, oder die Schweden an den Küsten Ehtlands und auf Runö, I–II*. Reval / Leipzig: Fr. Fleischer.
- SAOB = *Svenska Akademiens ordbok över svenska språket, 1–39*. 1898–2023. Lund: Svenska Akademien. Available online at: <https://www.saob.se/>.
- Schiller, Karl, August Lübben. 1875–1881. *Mittelniederdeutsches Wörterbuch, 1–6*. Bremen: Verlag von J. Kühtmann's Buchhandlung; Verlag von Hinricus Fischer.
- Schindler, Hans Jochem. 1972. *Das Wurzelnomen im Arischen und Griechischen*. Inaugural-Dissertation. Würzburg.
- Schwarz, Ernst. 1951. *Goten, Nordgermanen, Angelsachsen. Studien zur Ausgliederung der germanischen Sprachen*. Bern: A. Francke AG Verlag; München: Leo Lehnen Verlag.
- Söderwall, Knut Fredrik. 1884–1918. *Ordbok över svenska medeltids-språket, I–II*. Lund: Berlingska boktryckeri- och stilgjuteri-aktiebolaget.
- Tiberg Ob. = Tiberg, Nils. n.d. *Estlandssvensk kortordbok*. Ms. card dictionary. Uppsala: Swedish Institute for Language and Folklore (Institutet för språk och folkminnen).
- Vendell, Herman. 1882–1887. *Runömålet. Ljud- ock formlära samt ordbok*. Stockholm: Kongl. boktryckeriet P. A. Norstedt & söner.
- Wagner, Heinrich. 1956. Zum Dual im Germanischen. *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 74(3/4): 177–184.
- Wessén, Elias. 1965. *Svensk språkhistoria. I. Ljudlära och ordböjningslära*. Stockholm: Almqvist & Wiksell.

A. E. Маньков. Праиндоевропейский им.-вин. падеж двойственного числа существительных и двойственное число существительных в германских языках

В статье рассматриваются индоевропейские окончания дв. ч. существительных, затем остатки двойственного числа существительных в германских языках и, наконец, нефлективные способы выражения двойственности в германских диалектах. Особое внимание уделяется обозначению груди: прагерм. *\*breust-*, *\*brust-*, которое, возможно, имело

окончание дв. ч. в прагерманском. Как предполагается, др.-исл. *brjóst* n., др.-англ. *brēost* n. отражают прагерм. *\*breust-ō* (< пие. *\*-ō*, им.-вин. дв. ч. тематических существительных). Вследствие омонимии с им.-вин. п. мн. ч. среднего рода, которое также имело окончание *\*-ō*, данная форма была переосмыслена как мн. ч. среднего рода и стала склоняться как тематическое существительное (аналогом является гот. *daur* 'дверь'). Гот. *brusts* отражает прагерм. *\*brust-iz*, форму атематического мн. ч. (аналогичную др.-исл. *dyrr*). Др.-шв. *bryst* n. может отражать прагерм. *\*breust-ō* (и дв., и мн. ч.), *\*brust-ī* (дв. ч.), *\*brust-iz* (мн.). Мы предполагаем, что прагерманское обозначение груди являлось протерокинетическим существительным с аблаутом в корне, *\*breust-/\*brust-*, где *-t-* < *-d-* могло появиться по аналогии с «сердцем», которое также являлось протерокинетическим существительным с аблаутом. Поскольку выражение двойственности на уровне словоизменения исчезло в германских языках, существительные, обозначающие парные органы (напр., виски, плавники, жабры), иногда развивают альтернативные способы её выражения, а именно контаминацию с числительным «два» и редупликацию. Эти явления рассмотрены на материале шведских диалектов Эстонии.

*Ключевые слова:* двойственное число в индоевропейских языках; именительный и винительный падеж двойственного числа; праиндоевропейская морфология; прагерманская морфология.

## Nasal harmony in consonants in Chiquitano and its origins

This article deals with the origins of the so-called consonant nasal harmony in Chiquitano (Bolivia/Brazil, Macro-Jê family), in which the consonants /β ɾ j ɥ/ change to /m n ɲ ɳ/, usually when a nasal segment is present elsewhere in the word. The exact rules vary from dialect to dialect and are not fully described in the literature. Based on published works and my own field recordings, I provide a description of nasal harmony in contemporary varieties of Chiquitano. I argue that nasal harmony had vowels as its primary targets in Proto-Chiquitano, whereas consonants were indirectly affected by the process due to tautosyllabic assimilation. I also provide evidence that nasal harmony in consonants arose when nasal vowels underwent massive denasalization, thus phonologizing the erstwhile nasal and non-nasal allophones of the sonorant series. The present hypothesis explains why morphemes without a single nasal segment can have a floating feature [+nasal] in the contemporary Chiquitano varieties under examination and accounts for the phonological adaptation of certain loanwords from Spanish and Guaraní.

**Keywords:** Chiquitano language; Macro-Jê languages; nasal harmony; consonant harmony.

In Chiquitano (< Macro-Jê; Bolivia, Brazil), consonants assimilate in nasality under certain circumstances, whereas intervening vowels do not necessarily nasalize. This phenomenon is known as NASAL HARMONY, and it is already noted in the earliest surviving descriptions of Chiquitano (1).

(1) 18th-century Javeriano Chiquitano (Anonymous n/d *apud* Adam & Henry 1880: 3)

- |    |             |               |
|----|-------------|---------------|
| a. | <obobos>    | <omomoma>     |
|    | /o-βoβó-ʃi/ | → /o-momó-ma/ |
|    | NHA-toad-X  | NHA-toad-DIM  |
|    | ‘toad’      | ‘little toad’ |

18<sup>th</sup> century Santiagueño Chiquitano (Pellejà n/d: 3)

- |    |           |               |
|----|-----------|---------------|
| b. | <turus>   | <tunumaâ>     |
|    | /turu-ʃi/ | → /tunú-maʔã/ |
|    | door-X    | door-DIM      |
|    | ‘door’    | ‘little door’ |

Similar facts are reported and discussed for all known Chiquitano varieties, including Bésiro (Sans 2011), the San Rafael subdialect of Eastern Chiquitano (Girard 2014), the Ignaciano subdialect of Eastern Chiquitano (Ciucci & Macoñó Tomichá 2017: 38–39), and Migueleño (Nikulin 2020a: 4–7).

While nasal harmony targeting vowels is cross-linguistically well-attested (Botma 2024), Chiquitano is unusual in showing a kind of nasal harmony where consonants are the main targets of the process. A textbook example of nasal harmony in consonants is Hyman’s (1995) account of KiYaka (< Bantu < Niger–Congo; Democratic Republic of Congo). It has been argued that consonant harmony is fundamentally different from vowel harmony in that it involves featural agreement as opposed to spreading (Hansson 2001; Rose & Walker 2001, among others), though this has been contested (Jurgec 2011). In any case, nasal harmony re-

stricted to consonants is entirely unattested in South American languages other than Chiquitano, in stark contrast with nasal harmony in vowels. The latter phenomenon is widely attested in geographically adjacent languages, such as Gwarayu (< Tupi–Guaranian < Tupian; Bolivia; Danielsen 2019), Eastern Bolivian Guaraní (< Tupi–Guaranian < Tupian; Bolivia, Argentina, Paraguay; Dietrich 1986: 60–63), Ayoreo (< Core Zamucoan < Zamucoan; Bolivia, Paraguay),<sup>1</sup> and Mosestén (Mosestenan; Bolivia; Sakel 2004: 40–41), as well as in the genetically related language Maxakalí (< Maxakalian < Macro-Jê; Brazil; Silva 2020: 139–145). In languages that show nasal harmony in vowels, consonants are also often affected by the process, but they are not the primary target. Therefore, Chiquitano is typologically and areally salient regarding its nasal harmony pattern.

The goal of this article is twofold. After a brief presentation of the dialectal diversity of Chiquitano (Section 1), I make an attempt at elucidating the synchronic workings of the nasal harmony in the attested Chiquitano varieties (Section 2). I show that there are at least two processes involved, with different directionalities and different domains. I also identify a phonotactic tendency and a loanword adaptation pattern that provide a clue to the diachronic origins of the nasal harmony in consonants. In Section 3, I propose that Chiquitano at an earlier stage displayed a classic nasal harmony, whereby nasality spread bidirectionally from nasal vowels. At that stage, nasal consonants occurred as allophones of approximants due to local assimilation. Later, vowels underwent denasalization in most environments, and nasal consonants became phonemic. Section 4 concludes the paper.

## 1. Chiquitano and its dialects

Chiquitano is a macrolanguage in the sense that it is usually referred to as a single language but is in fact a dialect continuum consisting of varieties perhaps best viewed as distinct languages by the criterion of mutual intelligibility. It is spoken in the Chiquitanía (locally *Chiquitania*) region, spanning eastern Santa Cruz (Bolivia) and southwestern Mato Grosso (Brazil). A significant part of the respective ethnic group identifies as *Monkóxi* (often spelt *Monkox*), pluralized as *Monkóka*, and refers to their language as *Bésiro*. These endonyms are, however, not accepted in parts of the Chiquitanía. In this article, the term *Bésiro* refers to a specific Chiquitano variety (see 1.1 below).

The 2012 Bolivian census reports a total of 148,736 individuals nationwide to have declared themselves to be ethnically Chiquitano, Monkox, or (speakers of) *Bésiro*. However, in the same census, only 6,709 respondents aged 4 or above declared Chiquitano/*Bésiro* as their L1, and only 2,401 individuals aged 6 or above claimed Chiquitano/*Bésiro* to be the language they speak the best (INE 2015). The 2022 Brazilian census reports a total of 197 Indigenous residents in the Portal do Encantado Indigenous Land that falls within the Mato Grosso portion of the Chiquitanía region described above, as opposed to 1,046 in the 2010 Census (IBGE 2023: 139). No demographic statistics are currently available on the ethnic Chiquitano or Camba (Kamba) population in communities of Mato Grosso that lie outside Portal do Encantado, such as Vila Nova Barbecho, and within the urban limits of towns such as Porto Esperidião, Cáceres, Pontes e Lacerda, Vila Bela da Santíssima Trindade, Várzea Grande, or Corumbá (Pacini 2012: 276; Silva 2009). In any case, the Chiquitano language in Brazil is se-

<sup>1</sup> Nasal harmony in Ayoreo is described as affecting consonants in published works (e.g. Bertinetto 2009: 11–2), but Pier Marco Bertinetto (p. c., 2023) and Luca Ciucci (p. c., 2023) confirm that unstressed vowels are also weakly nasalized under nasal spread.

verely endangered; Santana (2014: 27) reported as few as 4 or 5 speakers, though revitalization efforts are underway (Santana 2014; Rupé 2022). In total, Chiquitano is spoken by fewer than 7,000 individuals (probably around 2,000 if only fluent speakers are considered) out of an estimated ethnic population of 150,000.

Chiquitano is classified as a divergent member of the Macro-Jê family (Adelaar 2008). Historically, the Chiquitanía was a multilingual region, with Chiquitano spoken alongside an undetermined number of Chapacuran, Arawakan, Tupian, Bororoan, and Zamucoan languages. From the 16<sup>th</sup> century onward, Spanish and Portuguese were added to this list. In the late 17<sup>th</sup> and 18<sup>th</sup> centuries, however, Jesuit missionaries imposed Chiquitano as the *lingua franca* of the Jesuit missions, resulting in a massive language shift to Chiquitano from other indigenous languages; of these, only the Arawakan language Paunaka survives to this day in use in areas corresponding to the old mission towns. The mid-20<sup>th</sup> century saw another language shift from Chiquitano to Eastern Bolivian Spanish in Bolivia and to Brazilian Portuguese in Brazil. The linguistic effects of these situations of language contact are discussed in Nikulin (2019).

### 1.1. Dialectology

Regarding the internal dialectal diversity of Chiquitano, at least three dialect zones, or maybe three languages, can be identified. I refer to these as **Bésiro**, **Eastern Chiquitano**, and **Migueleño**.

The **Bésiro** variety, currently the most vital, is recognized as one of the official languages of the Plurinational State of Bolivia. It is the only Chiquitano variety that has a codified orthographic norm (Parapaino Castro 2008). It is spoken in the Lomerío area and in the town of Concepción, in the province of Ñuflo de Chávez (Santa Cruz department, Bolivia), by descendants of refugees who fled from the Jesuit missions. Tomichá Chuvé (2023) is the most recent sociolinguistic study on Bésiro. The Bésiro forms in this article come from Galeote Tormo (1993), Parapaino Castro (2008), and Sans (2010).

**Eastern Chiquitano** is a provisional umbrella term for the varieties spoken in northern, northeastern, and southeastern Chiquitanía, which are notable for using the root *şoşiji-* ‘rainbow’ as opposed to *içi-*, found in other dialects. Subdialects include Ignaciano (spoken in and around the town of San Ignacio de Velasco, in the province of José Miguel de Velasco, Santa Cruz, Bolivia), Brazilian (remembered by some elders in the villages of Acorizal, Fazendinha, and Vila Nova Barbecho, in the Brazilian state of Mato Grosso), Santiagueño (spoken in Santiago de Chiquitos, in the province of Chiquitos, Santa Cruz, Bolivia), and possibly others.

The **Migueleño** variety is spoken by no more than 30 individuals in the province of San Miguel de Velasco (Santa Cruz, Bolivia). This is the variety that I have most first-hand experience with. I studied it in six fieldwork trips to the village of San Juan de Lomerío and the town of San Miguel de Velasco between 2017 and 2023.

An unresolved issue in Chiquitano dialectology is how to classify the Javeriano variety, spoken in San Javier de Chiquitos (Ñuflo de Chávez province, Santa Cruz, Bolivia). Galeote Tormo (2014: 269) offers a very limited amount of data, representative of a transitional variety combining both Migueleño and Eastern Chiquitano features, which is quite unexpected given the location of San Javier de Chiquitos. The descriptions of Anonymous (1718) and Anonymous (n/d), which are also most likely based on the variety spoken in San Javier three centuries ago (henceforth **18<sup>th</sup>-century Javeriano Chiquitano**), represent a lect that differs significantly from all known contemporary varieties of Chiquitano. Given the complex sociolinguistic situation of the Chiquitanía and the high incidence of migrations between different mis-



sions, it appears unlikely that 18th century Javeriano Chiquitano is a direct ancestor of modern Javeriano, or of any other modern Chiquitano lect.

## 1.2. Phonology

This subsection provides a pandialectal outline of Chiquitano phonology. The preferred syllable structures in Chiquitano are CV and V; consonant clusters or codas are rarely tolerated. Nevertheless, they may occur as a result of vowel elision or in loanwords.

Table 1 lists the consonants that have phonemic status in at least one Chiquitano variety. Chiquitano shows progressive palatalization of consonants, and the palatalized consonants (analyzed either as allophones of plain consonants or as independent phonemes, depending on the variety and on each author's analytical decision) are indicated below the respective phonemes in Table 1 without the slash marks.

/p/ p <sup>j</sup> ~ ɸ	/t/ t <sup>j</sup>	(/ts/)	/tʃ/	/k/ k <sup>j</sup> ~ ɸ	/ʔ/
	/ɕ/ ɸ	/s/		(/x/) ɸ	/h/
/β/ β <sup>j</sup> ~ j		/ɾ/	/j/	(/uɥ/)	
/m/ m <sup>j</sup> ~ ɲ		/n/	/ɲ/	(/ŋ/)	

**Table 1.** Chiquitano consonants

The phoneme /ts/ is present in all varieties except Bésiro, where it has merged with /s/ (2). The phoneme /x/ is present only in Migueleño, which shows a chain shift /ɾ/ (before a consonant) > /ɕ/ > /x/ (3). The phoneme /h/ occurs marginally in all varieties except Bésiro (4). The marginal phonemes /uɥ/ and /ŋ/ occur in all dialects except Bésiro, which has /β/ and /m/ instead (5). The aforementioned progressive palatalization process operates in slightly different ways depending on the specific variety. For example, Bésiro palatalizes /p β m t ɕ k/ to [p<sup>j</sup> β<sup>j</sup> m<sup>j</sup> t<sup>j</sup> ɸ k<sup>j</sup>], respectively, whereas Migueleño palatalizes /p β m t x k/ to [ɸ<sup>j</sup> j<sup>j</sup> n<sup>j</sup> t<sup>j</sup> ɸ ɸ] (6).

- (2) Bésiro *n-o-sokoré-s* ~ Migueleño *o-tsokoré-s* ‘seriema bird.♂’
- (3) Bésiro *ɕáñ-ɸ* ~ Migueleño *xaʔí-ɸ* ‘faeces’
- (4) Bésiro *aémo* ~ Migueleño *haémo* ‘for you’
- (5) Bésiro *n-o-iβó-ɕ* ~ Migueleño *o-iuó-x* ‘deer.♂’
- (6) Bésiro /n-i-po-ɕi/ *nip<sup>j</sup>óɕ* ~ Migueleño /i-poó-ɕi/ *coóx* ‘her house’

Table 2 shows the vowel inventory of the Chiquitano varieties. In this article, I follow the mainstream practice of representing the non-high front vowels as /e/ and /ẽ/, though I find their most typical realizations to be closer to [ɛ] and [ɛ̃], respectively. The status of vowel length as a distinguishing feature is disputed (cf. Ciucci & Macoñó Tomichá 2017: 40).

/i/ (/ĩ/)	/i/ (/ĩ/)	/u/ (/ũ/)
/e/ (/ẽ/)	/a/ (/ã/)	/o/ (/õ/)

**Table 2.** Chiquitano vowels

Nasal vowels are present in all varieties, though their status is contested in Bésiro (cf. Sans 2010: 90–1). In varieties other than Bésiro, they almost always occur next to another vowel, possibly separated by a glottal stop (7).<sup>2</sup> In other environments, nasal vowel phonemes arguably surface as [VÑ] or [VN] sequences (8), though it remains debatable whether the segments resulting in these sequences are in fact underlying nasal vowels or /VN/ sequences. [VÑ] and [VN] sequences are rare in native vocabulary (8a–d), though they are common in borrowings from Spanish (8e–f). The vowel in such sequences is audibly nasal before palatal or velar consonants (8a, c–d), but less so before labial or dentalveolar ones (8b, e–f).

- (7) a. Bésiro *a-taitʃomé-ka* ~ Migueleño *a-taitʃomé-ka* ‘you squeeze out’  
 b. Bésiro *kimúintʃa* ~ Migueleño *çimóitʃa* ‘in the middle’  
 c. Bésiro *çoéŋ-s* ~ *çuéŋ-s* ~ Migueleño *çoʔé-s* ‘pampa field’  
 d. Bésiro *kán-ŋ* ~ Migueleño *kãʔã-x* ‘stone’  
 e. Bésiro *n-u-táu-ma* ~ Migueleño *u-tãt-ma*? ‘bird.♂’
- (8) a. Migueleño *ĩ-ca-ti* [ĩŋcat̪i] ‘I go away’  
 b. Migueleño *tãtó-x* [tan'tox] ‘güembé vine’  
 c. Migueleño *ũka* [ũŋka] ‘who.♀’  
 d. Migueleño *ĩkamáʔa* [ĩŋka'maʔa] ‘there’  
 e. Migueleño *kãpo-x* [kãpox] ‘space’ (< Spanish *campo*)  
 f. Migueleño *βéta-x* [βɛntax] ‘shop’ (< Spanish *venta*)

Stress is contrastive and mobile (Nikulin 2022), and is indicated by means of an acute accent in this article. The circumflex accent symbolizes the so-called “strong stress”, which lengthens the vowel and removes floating accents to its right (Nikulin 2022: 15–18).

More detailed phonological descriptions are available for the Bésiro (Krüsi & Krüsi 1978; Sans 2010) and Migueleño (Nikulin 2020a, 2021) varieties.

## 2. Nasal harmony: a synchronic view

This section identifies several nasality-related facts that hold synchronically in Chiquitano. 2.1 discusses a regressive nasal harmony process. 2.2 and 2.3 state important static restrictions regarding the occurrence of oral and nasal sonorants as well as nasal vowels. 2.4 provides an interim summary of the section.

### 2.1. Regressive nasal harmony

Regressive nasal harmony in Chiquitano is a process whereby nasality spreads from a nasal segment (any of /m n ɲ ŋ ã õ ã̃ ã̃̃/) leftwards. In 9–11, the sonorants /β r j uɣ/ are nasalized to /m n ɲ ŋ/. (We will later see that the high vowels /i ĩ u/ may also be targeted by this process, at least in the Migueleño variety.) The triggers are given in boldface. The domain of this process includes the root with all inflectional and derivational prefixes, as well as certain suffixes. No segments are known to block the leftward nasal spread.

<sup>2</sup> In 7c–d, Bésiro [ŋ] is clearly not phonemic, since that variety lacks the phoneme /ŋ/. Krüsi and Krüsi (1978: 60–1) analyze it as a transition between a nasal vowel and a velar stop or sibilant. Similarly, Sans (2010: 94) identifies [ŋ̃s], [ŋ̃ʃ], [ŋ̃j], and [ŋ̃k] as allophones of /s/, /ʃ/, and /k/ after nasal vowels. Those authors that do not recognize the existence of underlying nasal vowels apparently consider [ŋ] to be an allophone of /ŋ/.

## (9) Migualeño

- |                  |   |                           |
|------------------|---|---------------------------|
| a. /kijará-xi/   | → | /kijaná- <b>ma</b> ?a/    |
| fox-X            |   | fox-DIM                   |
| ‘fox (♀)’        |   | ‘little fox (♀)’          |
| b. /kuruβasí-xi/ | → | /kunumasí- <b>ma</b> ?a/  |
| chicken-X        |   | chicken-DIM               |
| ‘chicken (♀)’    |   | ‘little chicken (♀)’      |
| c. /kurusí-xi/   | → | /∅-çunusi- <b>ná</b> -ta/ |
| cross-X          |   | INV-cross-CAUS-F.3SGP     |
| ‘cross’          |   | ‘she is blessing her’     |
| d. /tʃáβ-o/      | → | /ma-tʃám- <b>an</b> -a/   |
| drink-F.3        |   | ANTP-drink-CAUS-F.3       |
| ‘she drinks’     |   | ‘she offers a drink’      |
| e. /táriβ-o/     | → | /ma-taním- <b>an</b> -a/  |
| break-F.3        |   | ANTP-break-CAUS-F.3       |
| ‘it is broken’   |   | ‘she breaks’              |

## (10) Eastern Chiquitano, Rafaeleño subdialect (Girard 2014)

- |                      |   |                              |
|----------------------|---|------------------------------|
| a. /oseuó-ʃi/        | → | /osenó- <b>ma</b> ?/         |
| maize-X              |   | maize-DIM                    |
| ‘maize’              |   | ‘little maize’               |
| b. /jo:ripʼaki?ó-ʃi/ | → | /jo:nipʼaki?ó- <b>ma</b> ?/  |
| courbaril-X          |   | courbaril-DIM                |
| ‘courbaril tree’     |   | ‘little courbaril tree’      |
| c. /∅-βa-topí-kʼa/   | → | /∅-ma-topi- <b>ɲaká</b> -ka/ |
| 1+2-ANTP-bathe-F.N3  |   | 1+2-ANTP-bathe-CAUS-F.N3     |
| ‘we bathe’           |   | ‘we bathe something’         |
| d. /u-βá-ka/         | → | /u-ma-t- <b>é</b> ?/         |
| 1+2-eat-F.N3         |   | 1+2-eat-F.3SGP-OBLINV        |
| ‘we (incl.) eat’     |   | ‘we (incl.) eat it’          |
| e. /jo-/             | → | /jo-toki?á:ʃi/               |
| 3PL-                 |   | 3PL-navel-X                  |
| ‘their’              |   | ‘their navels’               |

(11) 18th-century Javeriano Chiquitano (Anonymous n/d *apud* Adam & Henry 1880: 3)

- |                 |   |                               |
|-----------------|---|-------------------------------|
| a. <ibobica>    |   | <imomicoê>                    |
| /i-βoβí-ka/     | → | /i-momi-k-o?é/                |
| 1SG-invite-F.N3 |   | 1SG-invite-F.3SGP-OBLINV      |
| ‘I invite’      |   | ‘I invite her’                |
| b. <yaçaborica> |   | <ñaçamoniteê>                 |
| /j-asaβorí-ka/  | → | /ɲ-asamoni-t-e?é/             |
| 1SG-look-F.N3   |   | 1SG-look-F.3SGP-OBLINV        |
| ‘I look’        |   | ‘I look at it’                |
| c. <iziborica>  |   | <iquìmonimacaca>              |
| /i-tsiβorí-ka/  | → | /∅-i-kimoni- <b>maká</b> -ka/ |
| 1SG-spank-F.N3  |   | 1SG-INV-spank-CAUS-F.N3       |
| ‘I get spanked’ |   | ‘I spank’                     |

d. <obobos>		<omomoma>
/o-βoβó-çi/	→	/o-momó-ma/
NHA-toad-x		NHA-toad-DIM
'toad'		'little toad'

The leftward nasal spread is productive in Chiquitano, and applies to older borrowings from Spanish (in Nikulin's 2019 classification, all those that belong to stratum A and some of those that belong to stratum B), as shown in 12a–b. However, newer borrowings (12c–d) are not affected by nasal harmony, and constitute lexical exceptions. (12d shows an effect of rightward nasal spreading, on which see 2.2.)

## (12) Migueleño

a. /içpiniṭú-maʔa/	<	older Spanish	<i>espíritu</i> [eʃ'pɪritu] + DIM
hummingbird-DIM			spirit
'hummingbird'			'spirit'
b. /motôni-çi/	<	older Spanish	<i>botón</i> [bo'ton] + x
button-x			button
'button (clothing)'			'button'
c. /ariβiróne/ <sup>3</sup>	<	older Spanish	<i>almidón</i> [almi'ðon]
starch			starch
'starch'			'starch'
d. /rominú-xi/	<	older Spanish	<i>lomillo</i> [lo'miʎo] + x
saddle-x			saddle
'saddle'			'saddle'

Furthermore, some suffixes lie outside the nasal spreading domain. In 13, nasality does not spread from the suffixes /-ŋĩĩ/ and /-iño/; forms such as \*/op-i-ṭʰasun-ó-ŋĩĩ/ or \*/maʔ-iño/ are, therefore, unattested.

## (13) Migueleño

a. /op-i-ṭʰasur-ó-ŋĩĩ/	b. /βaʔ-iño/
3PL-INV-call-F.3-1SG <sub>P</sub>	DEM.PL-PL
'they call me'	'these (♀)'

A handful of so-called Trojan morphemes, which carry a floating [+nasal] feature, exist in Chiquitano. These trigger nasal spreading despite lacking any of /m n ɲ ŋ ɿ ũ ẽ õ ã/. In 14–16, the Trojan morphemes are given in boldface.

## (14) Migueleño /ij-/ 1SG.♂

a. /ɲ-Ø-á-ka/	b. /ɲ-ápa/	c. /ɲ-ótu/
1SG.♂-ANTP-put-F.N3	1SG.♂-louse	1SG.♂-tongue
'I put (♂)'	'my lice (♂)'	'my tongue (♂)'

## (15) Eastern Chiquitano, Rafaeleño subdialect /aβ-/ 2PL (Girard 2014)

a. /am-asi-ká-ti/	b. /am-óʔo/	c. /am-ótu/
2PL-look-F.N3-CTPT	2PL-tooth	2PL-tongue
'you (pl.) look'	'your (pl.) teeth'	'your (pl.) tongues'

<sup>3</sup> In this example, Spanish [m] is unexpectedly adapted as Migueleño /β/. I surmise this is due to the necessity to preserve the oral quality of /ɾ/ (< Spanish [ð]), which would otherwise nasalize to /n/ as per the progressive nasal harmony process (2.2).

(16) 18th-century Javeriano Chiquitano (Anonymous n/d *apud* Adam & Henry 1880: 3)

⟨yebabaca⟩		⟨iñemamacaca⟩
/j-eβaβá-ka/	→	/∅-ij-emama- <b>ká</b> -ka/
1SG-move_closer-F.N3		1SG-INV-move_closer-CAUS-F.N3
‘I move closer’		‘I move it closer’

Finally, at least in the Migueleño variety, high vowels are also targeted by nasal spread, but only when preceded by another vowel (possibly with an intervening glottal stop), as shown in 17a’-e’.

## (17) Migueleño

a. /tsoi-tfokó-ka/	a’. /tsoĩ-tf <i>im</i> ó-ka/
1+3-dance-F.N3	1+3-sit-F.N3
‘we (excl.) dance’	‘we (excl.) sit’
b. /∅-a-ij-axúβi/	b’. /∅-a-ĩ-t <i>á</i> ʔã/
NF-2SG-INV-wash	NF-2SG-INV-bring
‘for you to wash it’	‘for you to bring it’
c. /jaʔi-xi/	c’. /jaʔi- <b>ma</b> ʔa/
young_man-X	young_man-DIM
‘young man’	‘boy’
d. /taβaí-xi/	d’. /tamaí- <b>ma</b> ʔa/
chicha-X	chicha-DIM
‘chicha’	‘chicha (dim.)’
e. /a-u-tfáʔa/	e’. /á-ũ-timo/
NF-3PL-drink	NF-3PL-sit
‘for them to drink’	‘for them to sit’

## 2.2. Progressive nasal harmony

Sans (2011) posits a rightward nasal spread process for Bésiro based on negative evidence. Indeed, there is a static restriction in all Chiquitano varieties, whereby /mV/, /nV/, /ɲV/, /ŋV/, and nasal vowels may not be immediately followed by any of /β ɾ j u/. By contrast, sequences of syllables with nasal onsets are common in the lexicon of Chiquitano (18).

## (18) Migueleño

a. /anené-si/	b. /ɾominú-xi/	c. / <b>m</b> omené-si/
day-X	saddle-X	comb-X
‘day’	‘saddle’	‘comb’

Examples such as 18b (borrowed from Spanish *lomillo* [lo'miʎo]) are particularly revealing. Since Spanish /ʎ/ is normally adapted as Chiquitano /ɾ/ or /j/ in borrowings (Nikulin 2019: 12), one could expect the resulting form to be \*/romirú-xi/, which is, however, unattested. The illicit sequence \*/mir/ is instead replaced with /min/.

Another piece of evidence supporting the rightward nasal spread is the distribution of stem-final consonants in verbs. In general, verbal roots commonly end with /β/, /ɾ/, rarely with /j/, /u/ (19a–e). However, roots that include a nasal segment can only end with one of those segments if a voiceless consonant intervenes (as in 19b–c). Otherwise, the root-final consonant of the stem can only be /m/ or /n/ (rarely /ɲ/), as shown in 19f–h. All examples in 19 contain the F.3 suffix /-o/ (or /-a/, as per vowel harmony); in addition, 19c–d and 19g contain the antipassive prefix /βa-/ (or /ma-/ as per regressive nasal harmony).

## (19) Migueleño

- |                                                 |                                            |                                                      |                                                |
|-------------------------------------------------|--------------------------------------------|------------------------------------------------------|------------------------------------------------|
| a. /tʃáβ-o/<br>drink-F.3<br>'she drinks'        | b. /aétsor-o/<br>lose-F.3<br>'she is lost' | c. /ma-kátar-a/<br>ANTP-sing-F.3<br>'she sings'      | d. /βá-tʃej-o/<br>ANTP-give-F.3<br>'she gives' |
| e. /síuq-o/<br>be_dry-F.3<br>'it (wood) is dry' | f. /mánom-o/<br>sleep-F.3<br>'she sleeps'  | g. /ma-kitjónon-o/<br>ANTP-snore-F.3<br>'she snores' | h. /tomóen-o/<br>tie-F.3<br>'it is tied'       |

Rightward nasal spread is clearly blocked by voiceless consonants (Sans 2011). Any voiceless segment that intervenes between a potential trigger (underlying nasal segment) and a potential target (oral sonorant) results in the failure of the sonorant to nasalize (20–21).

## (20) Migueleño

- |                                                                            |                                                                          |
|----------------------------------------------------------------------------|--------------------------------------------------------------------------|
| a. /aétsor-o/ (*aétson-o/)<br>lose-F.3<br>'she is lost'                    | b. /ma-kátar-a/ (*ma-kátan-a/)<br>ANTP-sing-F.3<br>'she sings'           |
| c. /kātaβó-xi/ (*kātamó-xi/)<br>lock-X<br>'lock'                           | d. /matorí-xi/ (*matoní-xi/)<br>parrot-X<br>'parrot (♀)'                 |
| e. /materá-xi/ (*matená-xi/)<br>flag-X<br>'flag'                           | f. /omenetíβo/ (*omenetímo/)<br>how_many<br>'how many'                   |
| g. /maíçtiru/ (*maíçtinu/)<br>teacher<br>'teacher (♀)'                     | h. /ma-kásar-a/ (*ma-kásar-a/)<br>ANTP-have_rest-F.3<br>'she has a rest' |
| i. /ma-ematakúruts-o/ (*ma-ematakúnuts-o/)<br>ANTP-wait-F.3<br>'she waits' |                                                                          |

## (21) Bésiro

- |                                                                                                   |                                                                                                      |
|---------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------|
| a. <i>najkiré-s</i> (* <i>najkiné-s</i> )<br>dragonfly-X<br>'dragonfly'                           | b. <i>n-o-tijkirí-ç</i> (* <i>n-o-tijkiní-ç</i> )<br>L-NHA-rufous_hornero-X<br>'rufous hornero bird' |
| c. <i>n-u-manturé-s</i> (* <i>n-u-mantuné-s</i> )<br>L-NHA-crab_eating_fox-X<br>'crab-eating fox' | d. <i>metúura</i> (* <i>metúuna</i> )<br>Ventura<br>'Ventura (♀)'                                    |

One exception to the rightward nasal spread rule is the Bésiro linking consonant *n-*, which is added to noun forms that would otherwise be word-initial and does not trigger nasal spread, as in *n-aró-ɕ* (\**n-anó-ɕ*) 'rice'. In other Chiquitano varieties, the cognate prefix has the form /r-/ (unless regressive nasal harmony applies), as in Migueleño /r-aró-xi/ 'rice', suggesting that its form in Bésiro is the result of a recent sound change \*r > n in the word-initial position, which counterfeeds the progressive nasal harmony.

Some loanwords from Spanish are also exempt from progressive nasal harmony (22), though others – such as the example in 18b – do undergo it.

## (22) Migueleño

- |                                                    |                                                                                              |
|----------------------------------------------------|----------------------------------------------------------------------------------------------|
| a. /sajnorá-xi/ (*sajnoná-xi/)<br>lady-X<br>'lady' | b. /miêrkuri-çi/ (*miênkuni-çi/)<br>Wednesday-X<br>'Wednesday' (< Spanish <i>miércoles</i> ) |
|----------------------------------------------------|----------------------------------------------------------------------------------------------|

- |                                                                       |                                                                                 |
|-----------------------------------------------------------------------|---------------------------------------------------------------------------------|
| c. / <b>ma</b> rijá-xi/ (*/ <b>mani</b> já-xi/)<br>María-x<br>'María' | d. /∅-tʰ <b>ani</b> nére/ (*/tʰ <b>ani</b> néne/)<br>∅-M-Daniel<br>'Daniel (♂)' |
|-----------------------------------------------------------------------|---------------------------------------------------------------------------------|

### 2.3. Static restrictions on nasal vowels

An important fact about the nasal vowels in Chiquitano varieties other than Bésiro is that they are frequent only as part of /VV/ and /V?V/ sequences in native vocabulary (that is, only when they are adjacent to another vowel, possibly with an intervening glottal stop), as briefly discussed in 1.2 (examples 7–8). Additional examples are given in 23 (the /VV/ and /V?V/ sequences are in boldface).

#### (23) Migueleño

- |                                                                                  |                                                                                  |                                                                       |
|----------------------------------------------------------------------------------|----------------------------------------------------------------------------------|-----------------------------------------------------------------------|
| a. / <b>co</b> ótoʔo/<br>soon<br>'soon'                                          | b. /i-tʃ <b>óki</b> ʔã/<br>1SG-navel<br>'my navel'                               | c. / <b>má</b> -ĩki-ʔo/<br>ANTP-ask-F.3<br>'she asks, she greets'     |
| d. / <b>ma</b> -ũxókon-o/<br>ANTP-be_ill-F.3<br>'she is ill'                     | e. / <b>a</b> -ĩ-cá-ʔi/<br>2SG-INV-bring-IMP<br>'bring it!'                      | f. /po <b>no</b> éto-xi/<br>belt-x<br>'belt'                          |
| g. /∅-tʰ <b>okĩ</b> -ʔokó-ta/<br>INV-nude-CAUS-F.3SGP<br>'she is undressing her' | h. /∅-çakã <b>ʔá</b> -xi/<br>3SG-liver-x<br>'her liver'                          | i. /ç- <b>õ</b> ʔõ-k-óʔi/<br>1SG.♀-listen-F.N3-PSTV<br>'I listen (♀)' |
| j. / <b>pá</b> ĩ-xi/<br>month-x<br>'month'                                       | k. / <b>ma</b> -ētoními-ʔo/<br>ANTP-wash_hands-F.3<br>'she is washing her hands' |                                                                       |

The Bésiro cognates of these forms contain VN sequences, which are variably analyzed as underlying nasal vowels or as bisegmental /VN/ sequences: *kónto* 'soon', *ma-unʃókon-o* 'she is ill', *poñoentó-ʃ* 'belt'.

In other environments, nasal vowels are rare in native vocabulary (24), and at least in the Migueleño variety they are pronounced with a clearly audible consonantal phase, raising doubts on whether they should be really analyzed as underlying nasal vowels (cf. Nikulin 2021: 24–26) or as bisegmental sequences. Most occurrences of nasal vowels (or maybe bisegmental sequences) in this environment are found in transparent loans from Spanish (25).

#### (24) Migueleño

- |                                                                  |                                                    |
|------------------------------------------------------------------|----------------------------------------------------|
| a. /tátó-xi/ [tan'tox]<br>güembé-x<br>'güembé vine'              | b. /í-çá-ti/ [ĩçati]<br>1SG.go-F.N3-CTFG<br>'I go' |
| c. /sãkijoré-si/ [saŋkijo'res]<br>blackbird-x<br>'blackbird (♀)' | d. /ĩkamáʔa/ [ĩŋka'maʔa]<br>there<br>'there'       |

#### (25) Migueleño

- |                                                                              |                                                                                           |
|------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|
| a. /tomíço/ [to'mĩŋço]<br>Sunday<br>'Sunday' (< Spanish <i>domingo</i> )     | b. /ma-pésar-a/ [ma'pẽŋsarə]<br>ANTP-think-F.3<br>'she thinks' (< Spanish <i>pensar</i> ) |
| c. /tʃôpa-xi/ [tʃompax]<br>sweater-x<br>'sweater' (< Spanish <i>chompa</i> ) | d. /tʰuβété-si/ [tʰuwen'tes]<br>duende-x<br>'duende'                                      |

## 2.4. Interim summary

So far, we have seen that Chiquitano exhibits two kinds of nasal spread with differing directionalities: the regressive (right-to-left) type of spread (2.1) is not blocked by any segments, whereas the progressive (left-to-right) type of spread (2.2) is blocked by voiceless segments. Both processes target the sonorants /β ɾ j ɥ/ and the high vowels /i i u/ (the latter are only affected when they occur in the environment  $V(?)\_)$ , with the outcomes being, respectively, /m n ɲ ŋ/ and /ĩ ã õ/. The usual triggers are the nasal segments /m n ɲ ŋ ã õ ã õ ã õ/, but nasal spread is also triggered by Trojan morphemes, i.e. lexically specified morphemes that do not contain any of the normally expected segments. Loanwords are often, but not always, exempt from nasal spread. An additional important fact is that nasal vowels are frequent only in /VV/ and /V?V/ sequences in native vocabulary, except in the Bésiro variety, whereas sequences of the type /CVC/ are mostly found in recent loans.

## 3. A diachronic account

I propose that nasal harmony in Chiquitano originally involved vowels and not consonants. This is plausible from a typological point of view: nasal harmony processes are common in Eastern South America, and mostly affect vowels (in addition to triggering allophony in sonorant segments), as proposed by Picanço (2010) for the Mundurukuan branch of the Tupian language family and by Silva (2020) for the Macro-Jê language Maxakalí.

I reconstruct bidirectional nasal spread for Proto-Chiquitano. It was triggered by underlying nasal vowels (/ĩ ã õ ã õ ã õ/) and targeted vowels rather than consonants. The right-to-left nasal spread was not blocked by any segment, whereas the left-to-right nasal spread was blocked by voiceless segments. In my proposal, Proto-Chiquitano lacked the nasal phonemes \*/m n ɲ ŋ/. Instead, I reconstruct the sonorants \*/β ɾ j ɥ/, which surfaced as \*[m n ɲ ŋ] before nasal(ized) vowels, as proposed by Sans (2011) for the Bésiro variety. I further propose that all nasal vowels were **denasalized** in the contemporary Chiquitano varieties, except when they were part of /VV/ or /V?V/ sequences. The segments \*[m n ɲ ŋ] – originally positionally conditioned allophones of \*/β ɾ j ɥ/ – have thus acquired phonemic status, as shown in 26.

(26)	Proto-Chiquitano	Migueleno	
	a. */tākōré-si/ [tākō'nēsɪ]	/takoné-si/ [tako'nɛs]	'sugarcane'
	b. */āṛēré-si/ [ānē'nēsɪ]	/anené-si/ [anɛ'nɛs]	'day'
	c. */ōβ-étsoṛ-o/ [ō'mētsoṛo]	/om-étsoṛ-o/ [o'mɛtsoro]	'they are lost'
	d. */βárōβ-ō/ [mānōmō]	/mánom-o/ [manomɔ]	'she sleeps'

The erstwhile nasal harmony in vowels then gave rise to a long-distance (consonantal) nasal harmony (27).

(27)	Proto-Chiquitano	Migueleno	
	a. */kijará-ʃi/	/kijará-xi/	'fox'
	a'. */kījārá-βāʔā/	/kijanā-maʔa/	'little fox'
	b. */tjáf-o/	/tjáf-o/	'she drinks'
	b'. */βā-tjáf-ār-ā/	/ma-tjám-an-a/	'she offers a drink'

This hypothesis handily accounts for the existence of Trojan morphemes in contemporary varieties of Chiquitano, tracing them back to morphemes with an erstwhile nasal vowel that is not part of a /VV/ or /V?V/ sequence (28).



(28)	Proto-Chiquitano	Migueleño	
a.	*/-ótu/ [-ótu]	/-ótu/ [-'ot̩]	'tongue'
a'	*/ij-/ + */-ótu/ → */ij-ótu/ [i'jnot̩]	/jnot̩/ [jnot̩]	'my tongue (♂)'
b.	*/-ápa/ [-'ápa] ~ */-áp̃a/ [-'áp̃a]	/-ápa/ [-'apa]	'louse, lice'
b'	*/ij-/ + */-áp̃a/ ~ */-áp̃a/ → */ij-áp̃a/ [i'jnáp̃a] ~ */ij-áp̃a/ [i'jnáp̃a]	/jnáp̃a/ [jnáp̃a]	'my lice (♂)'

This scenario is corroborated by comparative evidence. The Chiquitano root for 'tongue' (28a–a) has Macro-Jê cognates with a nasal vowel in the initial syllable (Adelaar 2008: 24), and the respective Proto-Macro-Jê etymon is reconstructed as *\*-ñũ<sub>1</sub>ctôk* (Nikulin 2020b: 386).

Further support comes from the adaptation patterns of loanwords from Spanish. In the early colonial period, Spanish /mV/, /nV/, /ŋV/ were borrowed as Chiquitano /βṼ/ [mṼ], /rṼ/ [nṼ], /jṼ/ [jṼ], whereas Spanish /VNC/ sequences were borrowed as /ṼC/, since underlying codas were not allowed (29). Example 29c shows that at this point only progressive harmony applied, whereas the regressive was already largely inactive.

- (29) a. Spanish *ventana* /bentána/ 'window' → Chiquitano \*/βētarã-/ + progressive harmony → \*/βētarã-/ \* [mētanã-] > Migueleño /metaná-xi/ 'window'  
 b. Spanish *bandera* /bandéra/ 'flag' → Chiquitano \*/βãtera-/ \* [mãtera-] > Migueleño /materá-xi/ 'flag'  
 c. Spanish *lomillo* /lomílo/ 'saddle' → Chiquitano \*/roβĩru-/ + progressive harmony → \*/roβĩrũ-/ \* [romĩnũ-] > Migueleño /rominú-xi/ 'saddle'  
 d. Spanish *Ventura* /bentúra/ 'Ventura' → Chiquitano \*/βētûra/ \* [mē'tu.ra] > Bésiro /metûra/ 'Ventura (♀)'

Loans from Guaraní (which has nasal spread) preserve nasal(ized) vowels as such in earlier Chiquitano (30).

- (30) Guaraní /tak<sup>w</sup>ar-e?ẽ/ [tāk<sup>w</sup>ārē'ẽ] 'sugarcane' → Chiquitano /tākōrẽ-/ [tākōnẽ-] > Migueleño /takoné-si/ 'sugarcane'

In /VV/ and /V?V/ sequences, denasalization did not take place, and nasality was preserved in all varieties except Bésiro, on which see below (31).

- (31) a. Proto-Chiquitano \*/∅-ō?ōpakí-ʃi/ \* [ō?ōpa'kiʃi] ~ \*/∅-ō?ōpākí-ʃi/ \* [ō?ōpā'kiʃi] > Migueleño /∅-ō?ōpakí-xi/ 'her shoulder'  
 b. Proto-Chiquitano \*/iʃō?é-si/ \* [iʃō?ēs̩i] > Migueleño /ʃo?é-si/ 'pampa field'  
 c. Proto-Chiquitano \*/kã?á-ʃi/ \* [kã?ãʃi] > Migueleño /kã?á-xi/ 'stone'  
 d. Proto-Chiquitano \*/jã?í-βã?ã/ \* [jã?imã?ã] > Migueleño /jã?í-ma?a/ 'boy'  
 e. Proto-Chiquitano \*/ã-ĩ-tá?ã/ \* [ã?ĩtã?ã] > Migueleño /a-ĩ-tã?ã/ 'for you to bring'  
 f. Proto-Chiquitano \*/tsõ-tʃĩβó-ka/ \* [tsõ?itʃimóka] > Migueleño /tsoĩ-tʃimó-ka/ 'we (excl.) sit'  
 g. Proto-Chiquitano \*/á-ũ-tĩβō/ \* [ã?ĩtĩmō] > Migueleño /á-ũ-timo/ 'for you (pl.) or them to sit'  
 h. Proto-Chiquitano \*/ã-ĩ-tá?ã/ \* [ã?ĩtã?ã] > Migueleño /a-ĩ-tã?ã/ 'for you to bring'  
 i. Proto-Chiquitano \*/tãβãí-βã?ã/ \* [tã?imã?ã] > Migueleño /tamaí-ma?a/ 'chicha.DIM'

At least in the Migueleño variety, nasality is not always clearly audible on vowels immediately following nasal consonants, as in [tama?a] 'one', [maa'tax] 'fishhook', [naa'kiç] 'peanut'. The respective phonological representations are perhaps /tamã?á/, /m-ã(?)ã-tá-xi/, /nã(?)ãkí-çi/ (from Proto-Chiquitano \*/tãβã?á/, \*/βã?ãtá-ʃi/, \*/rã?ãkí-ʃi/).

In the Bésiro variety, \*/Ṽ?Ṽ/ sequences yielded VVN (if the vowels are different) or VN (if the vowels are identical), where N is realized as [m] before /p/ (32a), [n] before /t tʃ/ (32b),

[ŋ] before /k s ʃ/ as well as word-finally in the numeral ‘one’ (32c–g), and zero before /m n ɲ/ or word-finally in most instances (32h).

- (32) a. Proto-Chiquitano \*/ $\emptyset$ - $\tilde{o}$ - $\tilde{o}$ pakí-ʃi/ \* $[\tilde{o}\tilde{o}pa'kiʃi]$  ~ \*/ $\emptyset$ - $\tilde{o}$ - $\tilde{o}$ pákí-ʃi/ \* $[\tilde{o}\tilde{o}pã'kiʃi]$  > Bésiro *n-ompakí-ʃ* ‘her shoulder’  
 b. Proto-Chiquitano \*/ $\tilde{u}$ - $\beta\tilde{a}\tilde{a}\tilde{t}ur\acute{e}$ -si/ \* $[\tilde{u}mã\tilde{a}\tilde{t}u'resi]$  > Bésiro *n-u-manturé-s* ‘crab-eating fox’  
 c. Proto-Chiquitano \*/ $r\tilde{a}\tilde{a}\tilde{a}k\acute{i}$ -ʃi/ \* $[\tilde{n}\tilde{a}\tilde{a}\tilde{a}kiʃi]$  ~ \*/ $r\tilde{a}\tilde{a}\tilde{a}k\acute{i}$ -ʃi/ \* $[\tilde{n}\tilde{a}\tilde{a}\tilde{a}kiʃi]$  > Bésiro *nanjá-ç* ‘peanut’  
 d. Proto-Chiquitano \*/ $\beta\tilde{a}\tilde{a}\tilde{a}k\acute{a}$ -ʃi/ \* $[\tilde{m}\tilde{a}\tilde{a}\tilde{a}kaʃi]$  ~ \*/ $\beta\tilde{a}\tilde{a}\tilde{a}k\acute{a}$ -ʃi/ \* $[\tilde{m}\tilde{a}\tilde{a}\tilde{a}kaʃi]$  > Bésiro *manjá-ʃ* ‘south wind’  
 e. Proto-Chiquitano \*/ $\tilde{i}\tilde{s}\tilde{o}\tilde{t}\acute{e}$ -si/ \* $[\tilde{i}\tilde{s}\tilde{o}'\tilde{t}\acute{e}si]$  > Bésiro *çoéŋ-s* ‘pampa field’  
 f. Proto-Chiquitano \*/ $k\tilde{a}\tilde{a}\tilde{a}$ -ʃi/ \* $[\tilde{k}\tilde{a}\tilde{a}\tilde{a}ʃi]$  > Bésiro *kán-ʃ* ‘stone’  
 g. Proto-Chiquitano \*/ $t\tilde{a}\tilde{\beta}\tilde{a}\tilde{a}$ -ʃi/ \* $[\tilde{t}\tilde{a}\tilde{m}\tilde{a}\tilde{a}'\tilde{a}]$  > Bésiro *tamáŋ* ‘one’  
 h. Proto-Chiquitano \*/ $j\tilde{a}\tilde{a}\tilde{a}$ - $\beta\tilde{a}\tilde{a}$ -ʃi/ \* $[\tilde{j}\tilde{n}\tilde{a}\tilde{a}'\tilde{m}\tilde{a}\tilde{a}]$  > Bésiro *jáima* ‘boy’

Word-finally, reconstructed \*/ $\tilde{V}\tilde{V}$ / sequences do not show traces of nasalization in Bésiro, at least according to the published sources. The Bésiro reflexes in 33 are from Parapaino Castro (2008).

- (33) a. Proto-Chiquitano \*/ $\tilde{a}$ - $\tilde{i}$ - $\tilde{s}\tilde{a}\tilde{\beta}\tilde{u}$ -ts- $\acute{e}\tilde{t}\tilde{e}$ / \* $[\tilde{a}\tilde{i}\tilde{s}\tilde{a}\tilde{m}\tilde{u}'ts\acute{e}\tilde{t}\tilde{e}]$  > Bésiro *a-i-samu-séte* ‘make it!’  
 (compare Migueleño /a- $\tilde{i}$ -samu-ts- $\acute{e}\tilde{t}\tilde{e}$ /)  
 b. Proto-Chiquitano \*/ $k\tilde{a}\tilde{i}\tilde{\beta}\tilde{a}\tilde{a}$ -ʃi/ \* $[\tilde{k}\tilde{a}\tilde{i}\tilde{m}\tilde{a}\tilde{a}'\tilde{a}]$  > Bésiro *kaimá* ‘now’  
 (compare Migueleño /ka $\tilde{i}$ maʃi/)  
 c. Proto-Chiquitano \*/ $\tilde{a}$ - $\tilde{i}$ - $\tilde{t}\tilde{o}\tilde{\beta}\tilde{o}\tilde{t}\tilde{e}$ / \* $[\tilde{a}\tilde{i}\tilde{t}\tilde{o}'\tilde{m}\tilde{o}\tilde{t}\tilde{e}]$  > Bésiro *a-i-t<sup>h</sup>omóte* ‘for you to tie it’  
 (compare Migueleño /a- $\tilde{i}$ -t<sup>h</sup>omóte/)

As a result of differentiated evolution of \*/ $\tilde{V}(\tilde{?})\tilde{V}$ / sequences (which did not denasalize completely) and simplex nasal vowels (which did denasalize), contemporary Chiquitano varieties display synchronically active alternations between oral and nasal segments. In Migueleño, for example, one finds multiple morphemes where morpheme-final \*/ $\tilde{V}\tilde{V}$ / sequences alternate with short oral vowels before certain affixes (34). These mostly go back to Proto-Chiquitano alternations between \*/ $\tilde{V}$ / and \*/ $\tilde{V}\tilde{V}$ / (cf. Nikulin 2020a: 5, fn. 6). In Bésiro, alternations between *V* and *VN* are common (35), which go back to \*/ $\tilde{V}\tilde{V}$ / sequences word-finally or before an obstruent, respectively.

- |         |                      |                                                                                                             |                                                                   |
|---------|----------------------|-------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------|
| (34) a. | Proto-Chiquitano     | */ $\tilde{u}$ - $\beta\tilde{a}\tilde{a}$ - $\tilde{a}\tilde{r}\tilde{i}\tilde{t}\tilde{a}$ -ka/           | * $[\tilde{u}mã:n\tilde{i}'\tilde{t}\tilde{a}ka]$                 |
|         | Migueleño            | / $\emptyset$ -ma-ni <sup>h</sup> á-ka/<br>1+2-ANTP-speak-F.N3<br>‘we (incl.) speak’                        |                                                                   |
|         | a'. Proto-Chiquitano | */h- $\tilde{u}$ - $\beta\tilde{a}\tilde{a}$ - $\tilde{a}\tilde{r}\tilde{i}\tilde{t}\tilde{a}\tilde{a}$ /   | * $[\tilde{h}\tilde{u}mã:n\tilde{i}'\tilde{t}\tilde{a}\tilde{a}]$ |
|         | Migueleño            | / $\emptyset$ - $\emptyset$ -ma-ni <sup>h</sup> áʔa/<br>NF-1+2-ANTP-speak-F.N3<br>‘for us (incl.) to speak’ |                                                                   |
|         | b. Proto-Chiquitano  | */ $\tilde{i}$ - $\tilde{k}\tilde{a}$ -ta/                                                                  | * $[\tilde{i}'k\tilde{a}ta]$                                      |
|         | Migueleño            | / $\emptyset$ -çá-ta/<br>INV-carry-F.3SGP<br>‘she carries it’                                               |                                                                   |
|         | b'. Proto-Chiquitano | */h- $\tilde{a}\tilde{p}$ - $\tilde{a}$ - $\tilde{k}\tilde{a}\tilde{a}$ /                                   | * $[\tilde{h}\tilde{a}pã'kãʔa]$                                   |
|         | Migueleño            | /(h)-ap-a-káʔa/<br>NF-2PL-ANTP-carry<br>‘for you guys to carry’                                             |                                                                   |

(35) a. Proto-Chiquitano	<i>*/tsĩβãʔá-βãʔã/</i>	<i>*[tsĩm'ã'ʔãmãʔã]</i>
Bésiro	<i>tʃim'áma</i> small-DIM 'it is small'	
a'. Proto-Chiquitano	<i>*/tsĩβãʔá-βãʔã=tai/</i>	<i>*[tsĩm'ã'ʔãmãʔãtai]</i>
Bésiro	<i>tʃim'ámantai</i> small-DIM=just 'it is just small'	
b. Proto-Chiquitano	<i>*/ã-ĩ-tõβóʔẽ/</i>	<i>*[ãĩt'õ'mõʔẽ]</i>
Bésiro	<i>ait'omóʔe</i> 2SG-INV-tie 'tie it!'	
b'. Proto-Chiquitano	<i>*/ã-ĩ-tõβóʔẽ-ta/</i>	<i>*[ãĩt'õmõ'ʔẽta]</i>
Bésiro	<i>ait'omoénta</i> 2SG-INV-tie-F.3SG <sub>P</sub> 'you tie it'	

The scenario proposed in this section accounts for a range of facts. First of all, it derives a cross-linguistically unusual consonant harmony pattern from a typologically plausible vowel harmony pattern (26–27). Second, it provides a diachronic explanation for the existence of Trojan morphemes (28). Third, it accounts for the adaptation patterns of a handful of early loanwords from Spanish and Guaraní (29–30). It also accounts for the reflexes and alternations found in daughter varieties (31–35).

Under this proposal, the consonantal inventory of Proto-Chiquitano can be reduced to 12 contrastive segments (as opposed to 21 phonemic consonants in Migueleño), as shown in Table 3. The allophones *\*[pʲ βʲ mʲ tʲ ɸ kʲ]* must have existed in Proto-Chiquitano in palatalizing environments (*\*/i\_ V/, \*/ĩ\_ V/*, where *V ≠ /i\_ ĩ/*, except for */s/*, which palatalizes even when the following vowel is one of */i\_ ĩ/*). The allophones *\*[m mʲ n ɲ ɳ]* must have occurred in the protolanguage preceding nasal or nasalized vowels. The consonants *[uɥ ɳ]* are not mapped to any phoneme; instead, they are considered here to be inserted automatically in the environments *\*/i\_ V/, \*/ĩ\_ Ṽ/*, respectively (cf. Nikulin 2021:20–1).

<i>*/p/</i> <i>*[p pʲ]</i>	<i>*/t/</i> <i>*[t tʲ]</i>	<i>*/ts/</i> <i>*[ts]</i>	<i>*/tʃ/</i> <i>*[tʃ]</i>	<i>*/k/</i> <i>*[k kʲ]</i>	<i>*/ʔ/</i> <i>*[ʔ]</i>
	<i>*/s/</i> <i>*[s ɸ]</i>	<i>*/s/</i> <i>*[s]</i>			
<i>*/β/</i> <i>*[β βʲ m mʲ]</i>		<i>*/ɾ/</i> <i>*[ɾ n]</i>	<i>*/j/</i> <i>*[j ɲ]</i>	epenthetic <i>*[uɥ ɳ]</i>	<i>*/h/</i> <i>*[h]</i>

Table 3. Proto-Chiquitano consonants

The vowel inventory of Proto-Chiquitano is shown in Table 4.

<i>/i_ ĩ/</i>	<i>/ĩ_ ĩ/</i>	<i>/u_ ã/</i>
<i>/e_ ě/</i>	<i>/a_ ã/</i>	<i>/o_ õ/</i>

Table 4. Proto-Chiquitano vowels

#### 4. Conclusion

In this article, a diachronic explanation was proposed for the existence of an unusual nasal harmony pattern in Chiquitano, whereby consonants, rather than vowels, are the main targets of the process. In my account, nasal consonants \*[m n ɲ] and sonorants \*[β ɾ j] were allophones of one single series of phonemes in Proto-Chiquitano, with the choice determined by the nasality of the following vowel (\*[ɲ] and \*[ɥ] were epenthetic segments and probably were not phonemic at all). Nasality spread from all underlying nasal vowels leftwards (with no blocking segments) and rightwards (blocked by voiceless segments), in a way quite similar to the neighboring Tupi–Guaranian languages. Contemporary Chiquitano varieties mostly substituted oral vowels for their nasal counterparts through extensive denasalization, except where nasals were found in \*/VV/ or \*/VʔV/ sequences. Early loans from Spanish and Guaraní entered the language before the denasalization phase took place. When denasalization was complete, nasal consonants became phonemic, and nasal harmony started to be associated mostly with consonants.

The reconstruction proposed in this article is compatible with external data from other Macro-Jê languages. The Proto-Macro-Jê etyma of Proto-Chiquitano forms with a nasal vowel are also reconstructed with a nasal vowel: compare Proto-Chiquitano \*/-áʔã/ ‘feces’, \*/-óʔõ/ ‘food’, \*/-ótu/ ‘tongue’, \*/-βáɾõβ-/ ‘to sleep’, \*/-ʔijã/ ‘nose’, \*/-éʔẽ/ ‘hand’, \*/-ãjẽ-/ ‘meat’ and Proto-Macro-Jê \*-ñṼt° ‘feces’, \*-ñũ<sub>2</sub>(C) ‘food’, \*-ñũ<sub>1</sub>ctôk ‘tongue’, \*nũ<sub>2</sub>p ‘to lie (down)’, \*-ñija ‘nose’, \*-ñim° ‘hand’, \*-ñit ‘meat’.

Nasality in contemporary Chiquitano varieties has not yet been studied instrumentally. Future research will need to address the degree to which phonetic nasalization persists in environments such as /C\_C/, /\_?\_/, as well as next to nasal consonants.

#### Abbreviations

1/2/3	first/second/third person	NF	nonfinite
ANTP	antipassive	NHA	non-human animate gender
CAUS	causative	N3	non-third person
CTFG	centrifugal	OBLINV	oblique inverse voice
CTPT	centripetal	P	patient
DEM	demonstrative	PL	plural
DIM	diminutive	PSTV	postverb
F	finite	SG	singular
INV	inverse voice	X	singular, non-diminutive, with no referential possessor
L	linking consonant	♂	male speech
M	masculine gender	♀	female speech

#### Acknowledgments

I am grateful to the Chiquitano speakers who have welcomed me in their homes and taught me their language with great patience: Ignacia Tomichá Yopíé, Antonia Socoré Masáí, Victoriano Julián Laverán Ramos, Micaela Ribera Montero, and Miguel Putaré Tapanaché. Elmar Socoré Casupá, Rubén Pitigá Socoré, and their families have made my stays in San Juan de Lomerío possible. The present article would not exist without the support of my fellow Bobikíxh colleagues, especially Brittany Hause — who has also proofread the text and made a number of valuable suggestions — and Luca Ciucci. This paper was originally presented at the 18<sup>th</sup> Sergei Starostin Memorial Conference on Comparative-Historical Linguistics in March 2023. I thank the audience for their comments and suggestions.

## References

- Adam, Lucien, Victor Henry. 1880. *Arte y vocabulario de la lengua chiquita*. Con algunos textos traducidos y explicados compuestos sobre manuscritos inéditos del XVIII siglo. Paris: Maisonneuve y Cia.
- Adelaar, Willem F. H. 2008. *Relações externas do Macro-Jê: o caso do Chiquitano*. In Stella Telles & Aldir Santos de Paula (orgs.). *Topicalizando Macro-Jê*: 9–28. Recife: Nectar.
- Anonymous. 1718. *Arte de la lengua Chiquita*. San Xavier de Chiquitos. Manuscript.
- Anonymous. No date. *Grammatica de la lengua Chiquita*. Manuscript.
- Bertinetto, Pier Marco. 2009. *Ayoreo (Zamuco)*. A grammatical sketch. *Quaderni del Laboratorio di Linguistica* 8: 1–59.
- Botma, Bert. 2024. Nasal harmony. In: Nancy Ritter and Harry van der Hulst (eds.). *The Oxford handbook of vowel harmony*. Oxford University Press.
- Danielsen, Swintha. 2019. Las ortografías del guarayu: historia y actualidad — Araka'e ikwachiapri iyavei kuritëi ikwachiapri. *Página & Signos* 15: 89–131.
- Dietrich, Wolf. 1986. *El idioma chiriguano*. Gramática, textos, vocabulario. Madrid: Instituto de Cooperación Iberoamericana.
- Galeote Tormo, Jesús. 1993. *Manitana auqui besüro: gramática moderna de la lengua chiquitana y vocabulario básico*. Santa Cruz de la Sierra: Imprenta “Los Huérfanos”.
- Galeote Tormo, Jesús. 2014. Chiquitano (Besiro). In: Pieter Muysken and Mily Crevels (eds.), *Lenguas de Bolivia*, vol. III: Oriente: 259–301. La Paz: Plural Editores.
- Girard, Raphaël. 2014. Agreement without similarity: the case of Chiquitano nasal (consonant) harmony. Poster presented on May 18–19, 2014 at Agreement by Correspondence–ABC Conference, University of California–Berkeley, Berkeley. Retrieved from <<http://linguistics.berkeley.edu/~inkelas/PosterAbstracts/Girard.pdf>>.
- Hansson, Gunnar Ólafur. 2001. *Theoretical and typological issues in consonant harmony*. Ph.D. dissertation, University of California, Berkeley.
- Hyman, Larry M. 1995. Nasal consonant harmony at a distance: The case of Yaka. *Studies in African Linguistics* 24(1): 5–30. doi:10.32473/sal.v24i1.107408
- IBGE = Instituto Brasileiro de Geografia e Estatística. 2023. *Censo Demográfico 2022 — Indígenas: primeiros resultados do universo*. Rio de Janeiro: 2023.
- INE = Instituto Nacional de Estadística. 2015. *Censo de población y vivienda 2012*. Bolivia. Características de la población. La Paz: Instituto Nacional de Estadística.
- Jurjec, Peter. 2011. *Feature spreading 2.0: a unified theory of assimilation*. Ph.D. dissertation, University of Tromsø.
- Krüsi, Martin, Dorothee Krüsi. 1978. Phonology of Chiquitano. In *Work papers of the Summer Institute of Linguistics*. Riberalta, Bolivia. 1972–1976: 53–93. Riberalta: Summer Institute of Linguistics.
- Nikulin, Andrey. 2019. Contacto de lenguas en la Chiquitanía. *Revista Brasileira de Línguas Indígenas* 2(2): 5–30. doi:10.18468/rbli.2019v2n2.p05-30
- Nikulin, Andrey. 2020a. Elementos de la morfofonología del chiquitano migueleño. *LIAMES* 20:e020015. doi:10.20396/liames.v20i0.8660822
- Nikulin, Andrey. 2020b. *Proto-Macro-Jê: um estudo reconstrutivo*. Ph.D. dissertation, Universidade de Brasília.
- Nikulin, Andrey. 2021. Fonología segmental del chiquitano migueleño. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas* 16(2): e20200054. doi:10.1590/2178-2547-BGOELDI-2020-0054
- Nikulin, Andrey. 2022. La fonología del acento en el chiquitano migueleño. *Cadernos de Etnolingüística* 10(1): e100110.
- Pacini, Aloir. 2012. *Identidade étnica e território chiquitano na fronteira (Brasil–Bolívia)*. Ph.D. dissertation, Universidade Federal do Rio Grande do Sul.
- Parapaino Castro, Pablino. 2008. *Isiukiché nikorokó Bésiro*. Guía de escritura del idioma Bésiro. Santa Cruz de la Sierra: Unión de Artesanos de la Tierra–UNIARTE.
- Pellejà, Josep (?). No date. *Gramatica de la lengua de los indios llamados chiquitos pertenecientes al Gobierno de Chuquisaca en el Reyno del Perú doctrinados por los PP. de la Extinta Compañia de Jesus de la Provincia del Paraguay*. Manuscript.
- Picanço, Gessiane. 2010. A reconstruction of nasal harmony in Proto-Mundurukú (Tupi). *International Journal of American Linguistics* 76(4): 411–38. doi:10.1086/656313

- Rose, Sharon, Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80(3): 475–531.
- Rupé, Maria Síría. 2022. Ensino da língua Chiquitano e o registro de gêneros textuais orais. MA thesis, Universidade do Estado de Mato Grosso.
- Sakel, Jeanette. 2004. A grammar of Mosestén. Berlin/New York: Mouton de Gruyter.
- Sans, Pierric. 2010. Éléments de sociolinguistique et de phonologie du bésiro (chiquitano). *Langue en danger des basses terres de Bolivie*. MA thesis, Université Lumière Lyon 2.
- Sans, Pierric. 2011. Is nasality an autosegmental feature in Bésiro (a.k.a. Chiquitano)? In: Ana Suely Arruda Câmara Cabral, Aryon Dall’Igna Rodrigues, Jorge Domingues Lopes, Maria Risolêta Silva Julião (orgs.). *Línguas e culturas Tupí*, vol. 2: 237–248. Campinas: Curt Nimuendajú.
- Santana, Áurea Cavalcante. 2014. Línguas cruzadas, histórias que se mesclam: ações de documentação, valorização e fortalecimento da língua Chiquitano no Brasil. Jundiaí: Paco Editorial.
- Silva, Giovani José da. 2009. A presença Camba-Chiquitano na fronteira Brasil–Bolívia (1938–1987): identidades, migrações e práticas culturais. Ph.D. dissertation, Universidade Federal de Goiás.
- Silva, Mário André Coelho da. 2020. Tikmũn yĩ ax tinã xohi xi xahĩnãg — Sons e pedaços da língua Maxakalí: descrição da fonologia e morfologia de uma língua Macro-Jê. Ph.D. dissertation, Universidade Federal de Minas Gerais.
- Tomichá Chuvé, Ignacio. 2023. Situación del idioma bésiro, base para la autonomía indígena Monkox de Lomerío. Talk presented on July 27–28, 2023 at III Congreso MOVA: Memoria Oral y Voces de América, Universidad de Salamanca & Universidad Católica de la Santísima Concepción de Chile, Salamanca. Available online at: [www.youtube.com/watch?v=xgdtTX\\_aNZk](http://www.youtube.com/watch?v=xgdtTX_aNZk).

*A. В. Никулин. К вопросу о происхождении консонантной гармонии по назальности в языке чикитано*

В статье рассматривается происхождение так называемой консонантной гармонии по назальности в чикитано (Боливия/Бразилия, семья макро-же) — процесса, в рамках которого согласные /β ɾ j ɟ/ ассимилируются в /m n ɲ ɳ/, как правило, при наличии носового сегмента в том же слове. Точное правило варьируется от диалекта к диалекту, причём в литературе оно описано недостаточно полно. В статье предлагается описание гармонии по назальности в современных диалектах чикитано на материале опубликованных данных, а также полевых данных автора. Предполагается, что в прачикитано основными объектами гармонии по назальности были гласные, тогда как согласные затрагивались этим процессом опосредованно, в результате внутрислоговой ассимиляции. Консонантная гармония по назальности возникла как следствие массовой деназализации носовых гласных, приведшей к фонологизации бывших носовых и неносовых аллофонов сонорных согласных. Представляемая гипотеза позволяет объяснить, почему некоторые морфемы без единого носового сегмента имеют плавающий признак назальности в современных диалектах чикитано, а также описать фонологическую адаптацию заимствований из испанского и гуарани.

**Ключевые слова:** чикитано язык; макро-же языки; гармония по назальности; консонантная гармония.

## Nominations for siblings: Proto-Dravidian reconstruction and borrowability

The paper represents a historical analysis of the lexical terms for siblings in Dravidian languages, with special emphasis on the factor of borrowability. It is concluded that borrowings within this lexical subsystem of Dravidian languages must have started primarily from terms for elder siblings (more precisely, from the meaning ‘elder brother’). The preferred scenario eventually resulted in borrowing the entire subsystem, with its following coexistence along with the original subsystem. However, words denoting younger siblings retained more resistant to borrowing; these were the ones that could, in particular, preserve within themselves some of the most archaic linguistic features, such as prefixes of inalienable possession. The evolution of the system of those prefixes is also explored: it is proposed to consider reconstructions *\*y-* 1SG / *\*n-* 2SG / *\*t-* 3SG relevant for the Proto-Dravidian level, while in the North Dravidian subgroup of Kurukh-Malto these were replaced with the more productive oblique forms of personal pronouns (*eng-* 1SG / *ning-* 2SG / *tang-* 3SG).

*Keywords:* kinship terms; sibling terms; inalienable possession; borrowability; Dravidian languages; etymology; lexical borrowings.

### 1.1. Introduction: Kinship terms in Dravidian languages <sup>1</sup>

In this paper, I make an attempt at a reconstruction of how the Dravidian system for sibling terms changed since the time of Proto-Dravidian language (approximately the end of the 3rd millennium BC, according to Starostin 2000), including both family-internal changes and replacements through borrowing. Such an attempt implies focusing not only on the individual protoforms and their development, but on the kinship system as a whole, with its inherent characteristics. Thus, the meanings of Proto-Dravidian nominations for siblings have to be considered along with other data on Dravidian siblings and kinship terms.

A well-known attempt to visualize the Proto-Dravidian kinship system, not from a linguistic, but from an anthropological point of view was made by T. Trautmann in his “Dravidian Kinship” (Trautmann 1995). According to his conclusions, the basic structure of the cross-cousin kinship system, similar to the modern one, already existed at the Proto-Dravidian time, i.e. eB, yB, eZ and yZ<sup>2</sup> (see Figure 1 below). He argued that the question is no longer about the existence of the ancestral Dravidian kinship system with cross-cousin marriage rule, but rather about “what precise form that ancestral rule took” (ibid.: 236).

We find in his paper a description of Dravidian-speaking communities with a system of kinship terms of the so-called “Dravidian” bifurcative-merging type. These are (as it appears

---

<sup>1</sup> The study was supported by the Russian Science Foundation grant 22-28-00072 «Strategies for nomination in the field of basic zoo- and anthroponymic vocabulary in the languages of Eurasia». The author wishes to express her deep and sincere gratitude to George Starostin for his comments on the preliminary version of the article.

<sup>2</sup> Kinship terms for siblings are referred to in the article as eB ‘elder brother’, yB ‘younger brother’, eZ ‘elder sister’ and yZ ‘younger sister’ respectively.

		♂			♀		
		X	II				X
G <sup>2</sup>	e	MF	FFeB	FF	MM	MMeZ	FM
	y		FFyB			MMyZ	
G <sup>1</sup>	e	MB	FeB	F	M	MeZ	FZ
	y		FyB			MyZ	
G <sup>0</sup>	e	e(MBS)	eB		eZ		e(MBD)
	y	y(MBS)	yB		yZ		y(MBD)
G <sup>-1</sup>	e	♂ZS	♂eBS	S	D	♂eBD	♂ZD
	y		♂yBS			♂yBD	
G <sup>-2</sup>		♂DS, ♀SS	♂SS, ♀DS		♂SD, ♀DD		♂DD, ♀SD

Figure 1. Proto-Dravidian kinship reconstruction in anthropological terms, according to Trautmann 1995: 232.

from his description) communities speaking Tamil, Malayalam, Kannada, Tulu, Toda, Kodagu, Koya, Telugu, Gondi, Malto and Brahui languages, see Trautmann 1995: 41, etc. He explains the typical features of kinship terms system of this type – cross-cousin marriage and grouping of male relatives in the +1 generation.

Indeed, the grouping of male relatives of collateral lines of kinship in the parents' generation is one of the essential parameters of the typology of kinship term systems in anthropology. This feature became the basis of the four-member typology of kinship terms systems, introduced back in the 19th century by L. G. Morgan (see Morgan 1997 /1871/), further developed by other researchers (cf. Olderogge 1960, Dziebel 2001) and still in use today (cf., e.g., Popov 2015).

The author draws the border and the frontier zone of the typically “Dravidian” kinship term system structure along the line of the Central India contact area of Dravidian and Indo-Aryan languages (Trautmann 1995: 111). According to him, further to the north such a system is not preserved; this area is occupied by a typical Indo-Aryan family structure with no place for cross-cousin marriage. An example is the sibling term system of Hindi, for which distinctions between cross- and parallel cousins and cross-cousin marriage are not relevant, and neither is relative age distinction, see *bhāi* ‘brother’ B; *bahen* ‘sister’ Z (the meaning includes cousins from both sides; Trautmann 1995: 93). In relation to the Dravidian situation this means that we should not be expecting any traces of such a system in Northern Dravidian languages (i.e. to the north of the aforementioned line).

An earlier attempt to reconstruct the kinship system for the Proto-Dravidian period has been conducted by Bh. Krishnamurti in his fundamental work “Dravidian languages” (Krishnamurti 2003). He also believes it to have been a bifurcative-merging system with preservation of the cross-cousin marriage rule (ibid.: 10). In one of his papers, M. B. Emeneau (1953) examined the general system of Proto-Dravidian kinship, with evidence drawn from Old Tamil and modern Dravidian languages. He drew attention to the general markers of inalienable possession, as in Old Tamil *entai* ‘my or our father’, *nuntai* ‘your father’, *tantai* ‘his, her or their father’ (see discussion in section 4 below).

In the present paper I will focus specifically on one section of the Proto-Dravidian kinship system, namely, the subsystem for “siblings”. The natural questions to be asked are: (a) what did it look like? (b) which of the systems in modern descendent languages have retained the original proto stems? (c) how did this system develop, what happened to it later in different lineages of Dravidian languages, and which elements were replaced by borrowings?



One should keep in mind that kinship terms (or, at least, certain subsets of kinship terms) are sometimes considered to belong to the “basic lexicon” of the language and thus not highly prone to being borrowed. Even so, contact-induced changes can also include changes in kinship terminology, as Metsäranta et al. (2023) have shown in their recent work; the degree of their resistance to borrowings provides us with “a lens with which to evaluate the nature and intensity of contact situations” (Metsäranta et al. 2023: 141).

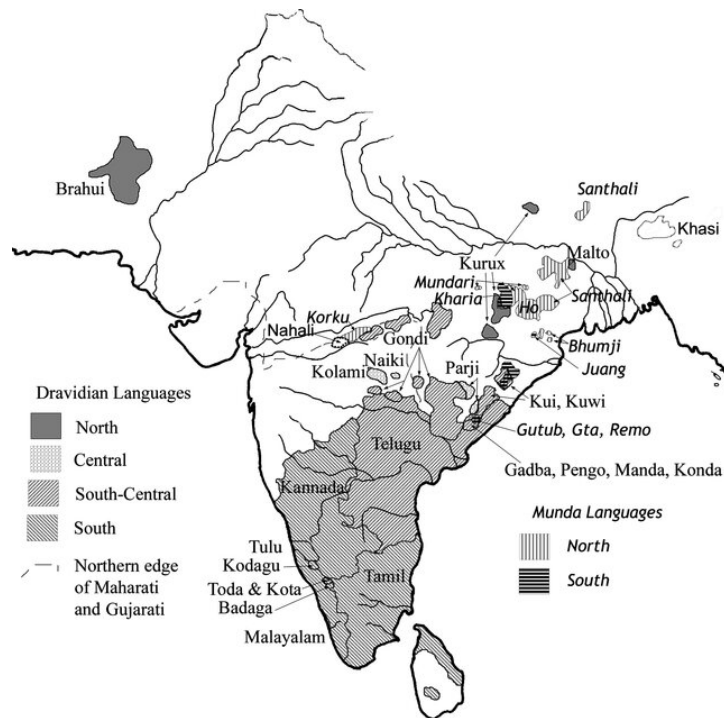


Figure 2. Modern distribution of Dravidian languages (also including Munda).

Since any reconstruction of the evolution of kinship terms in the history of Dravidian languages must necessarily be dependent on a specific model of Dravidian classification, it is useful to at least briefly describe the situation here. According to a widespread (though not uncontested) classification scheme, Brahui was the first language to split from the Proto-Dravidian community (see Figure 2). This was followed by the separation of the northeastern Dravidian languages Kurukh and Malto, and later the bulk of Central Dravidian; all these groups of languages separated from the Proto-Dravidian community relatively early. The process of linguistic development may have coincided with the supposed gradual advancement of the Dravidians in the direction from the northwest to the south (Gurov 2013: 23). In his dissertation, Georgiy Starostin dates the separation of Southern and Central Dravidian to approximately 1200 BC, based on glottochronological calculations (Starostin 2000: 13). The splitting of the South Dravidian group occurs at a later date, some time during the 1st millennium B.C., around the same time as the split of Kurukh and Malto in the North.

The methodology of our research generally rests on the standard comparative method as described, e.g., in Burlak & Starostin 2005 and in Starostin 2013. We also take into consideration accumulated data on semantic shifts and semantic reconstruction (see Zalizniak 2018; Zalizniak et al. 2023), as well as studies of borrowability, such as Matras 2009 and others.

The data were collected from dictionaries of the languages under investigation, primarily from the classic etymological dictionary by T. Barrow and M. B. Emeneau (hereafter DEDR 1984) that collects data on a significant number of small and endangered Dravidian languages, as well as dictionaries of concrete languages (listed below in the “Sources of linguistic data”

section). Reconstructions of Proto-Dravidian stems are given according to the dissertation of G. S. Starostin (Starostin 2000), supported by the online etymological database (Starostin 1998–2005), further referred to as StarLing 2023; in some cases the etymologies of B. Krishnamurti are given, cf. Krishnamurti 2003. Two more etymological dictionaries (Monier-Williams 1899; Turner 1966) were used in the analysis of Indo-Aryan loans.

## 1.2. Notation convention

In the tables below we indicate the Dravidian protoform for the lexeme if such a protoform can be extracted from existing etymological databases and dictionaries. If a protoform cannot be reconstructed due to the word in question most likely being of borrowed origin, we mark it as a borrowing (LOAN). If the source of borrowing is relatively transparent, we specify the notation further, e.g. LOAN-OIA for loans from Old Indo-Aryan, LOAN-Mar for borrowings from Marathi, and so on. When the source cannot be ascertained with clarity, but there are still strong arguments in favor of borrowing, we mark the item as LOAN-Un (short for Borrowing-Unknown). Hopefully, further research on Dravidian contacts with languages from other families will eventually help identify the concrete source of borrowings in such cases.

## 2. Kinship terms for siblings in Dravidian languages<sup>3</sup>

The kinship system of the Dravidian peoples was first described by the American scientist L. G. Morgan, one of the founders of evolutionism in the social sciences and cultural anthropology (Morgan 1997 (1871)). He noticed the similarity between the kinship term systems of the Dravidians (his material was based on Tamil data) and the Iroquois (Seneca and Ojibwe) of North America, and identified them as a special type. The 1950s saw the publication of several studies by L. Dumont, which included specific descriptions of Dravidian family terms. This is what Dumont wrote about the Dravidian system: “Dravidian kinship terminology, and with it other terminologies of the same type, can be considered in its broad features as springing from the combination in precise configurations of four principles of opposition: distinction of generation (qualified as an ordered scale), distinction of sex, distinction of kin identical with alliance relationship, and distinction of age” (Dumont 1953: 39).

In the kinship term system of this type, parallel cousins (children of the father’s brother or mother’s sister) are considered brothers and sisters and are equated to siblings, while children of opposite-sex siblings (cross-cousins) are considered more distant relatives, among whom it is generally encouraged to search for a spouse (so-called cross-cousin marriage), cf. Trautmann 1995.

For example, there is a special term in Telugu connected with one of the kinship lines — the line through the mother’s brother, i.e. connecting Ego with a group of cross-cousins from whom he can choose a marriage partner: *mēnarikamu* ‘(for a man) marrying the daughter of his maternal uncle’; ‘(for a woman) marriage with the son of her paternal aunt’ (Brown 1903: 1035), also *mēnarikam* ‘(of a man) marriage with his maternal uncle’s daughter; (of a woman) marriage with her paternal aunt’s son’ (Gwynn 1991: 437); *mēnamāma* ‘maternal uncle, mother’s brother’, *mēnamaradalu* ‘father’s sister’s daughter’ etc. These intra-family relationships and connections with “intended marriage partners” are very important for the family structure as a whole (Trawick 1992), and in particular for the sibling subsystem.

---

<sup>3</sup> The author would like to express her gratitude to Veronika Milanova and Niklas Metsäranta for the discussion of borrowability issues.

According to Krishnamurti 2003, the Proto-Dravidian sibling system also reflected the idea of elder and younger siblings of Ego (although, as he points out, the basis for his reconstruction rests primarily on the South Dravidian kinship systems).

He underlines that “separate labels existed for the elder and younger in ego’s generation”, according to him, these terms were *\*akka-* ‘elder sister’; *\*tam-kay*, *\*cēlāl* ‘younger sister’; *\*aṇṇa-* ‘elder brother’, *\*tamp-V-* ‘younger brother’ (Krishnamurti 2003: 10).

These protoforms are almost identical to the ones currently found in the online source *Dravidian Etymological Database*, published as part of the Tower of Babel resource (Starostin 1998–2005 = StarLing 2023). Starostin reconstructs the Proto-Dravidian sibling system as *\*āṇa-* ‘elder brother’; *\*ak-* ‘an elder relative’; *\*aṇ-* ‘a k. of relative (brother/sister)’; *\*t-amp-* ‘younger brother’; *\*cel-* ‘companion, sister’. In general, forms in StarLing 2023 aim at reconstruction on a deeper historical level compared to those in Krishnamurti 2003. In addition, one other protoform is reconstructed for the sibling subsystem: PDR *\*tōz-* ‘friend’; ‘younger brother’, ‘assistance, help’ (StarLing 2023).

In both sources, the meaning ‘younger brother’ is given with the prepositive *t-*, which, in our opinion, can be doubted, as we will show below.

In addition to the sibling subsystem including lexemes expressing the meanings of ‘brothers’ and ‘sisters’ of Ego, we have to take into account the importance of the parameter of “relative age” in the majority of Dravidian languages, respectively dividing these into ‘elder’ and ‘younger’. The full set of resulting meanings (List A) is as follows:

- *elder brother*
- *younger brother*
- *elder sister*
- *younger sister*.

For subsystems not distinguishing relative age, only the meanings *brother* and *sister* are relevant.

Keeping in mind the importance of cross-cousin marriage, we should consider as one of its consequences the distinction between cross- and parallel lineage of cousins, and the merging of consanguineous brothers and sisters with parallel cousins of the same sex and relative age. The extended set (List A+), therefore, must be as follows: *elder brother or parallel male cousin, younger brother or parallel male cousin, elder sister or parallel female cousin, younger sister or parallel female cousin; elder male cross-cousin; younger male cross-cousin; elder female cross-cousin; younger female cross-cousin*. However, since for many languages, especially small ones, there is no detailed information about cousins and the branch to which they belong, in the further presentation we adhere to List A.

Kinship terms are sometimes considered to belong to the «basic vocabulary» and thus to remain largely immune to borrowing. However, they are not included in the Swadesh 100-item list of the most “stable” part of the lexicon (Swadesh 1971); only in the larger 200-item Swadesh list several kinship terms are included, such as ‘father’, ‘mother’, ‘child’, ‘husband’, ‘wife’. They are also not included in the Leipzig-Jakarta list of most borrowing-resistant meanings (Tadmor, Haspelmath, Taylor 2010: 233).

Recently, several studies have appeared indicating that borrowings form a significant part of such systems. In particular, work by Milanova et al. (2020) includes a comparative analysis of borrowings in Indo-European kinship and social term systems; the study by Metsäranta et al. (2023) explores the borrowability of kinship terms in Uralic languages. When bilingual speakers make a choice within the complex repertoire of kinship system elements from two contact languages (as described in Matras 2009), they may choose the kinship term from the

other language (as they may do with other lexical items), which opens a possibility for borrowing within the kinship system. It was argued that while affinal kin categories such as those denoting spouses, spouse's siblings, and sibling's spouses have the largest number of loanwords, "among the kin categories with the largest number of loanwords were also consanguineal categories such as those of 'mother' and 'father'" (Metsäranta et al. 2023: 141).

Since it is clear that the influence of Indo-Aryan languages on Dravidian in this region goes at least as far back as the 2nd millennium BC and has been continuous ever since then, in different forms and on different levels, it is not surprising that many sibling terms in these languages turn out to be of borrowed origin. Such lexical changes most frequently occur in contact languages, so there are borrowings from Marathi in Kolami, from Prakrits in Telugu, etc.

### 3. Linguistic data on siblings: from South to North Dravidian languages

In this section, we shall consider the relevant available data of the languages of different Dravidian family subgroups and their current state. On the basis of this lexical data we will try to figure out what we can say about the current state of the sibling subsystem, and make assumptions about how the development of the subsystem looked like since the collapse of the Proto-Dravidian community.

The classification model of Dravidian languages adopted for the subsequent presentation shall follow Starostin 2000; although, where necessary, we also take into account the lexical reconstructions from Krishnamurti 2003. We do not adopt his classification (which, among other things, groups Telugu together with the Gondwana languages), rather preferring to rely on the scheme based on the lexicostatistical calculations of M. S. Andronov and G. Starostin.

#### 3.1. Tamil and other South Dravidian languages

Tracing the language data from South to North, we consider first the data of Dravidian languages from the subgroup "South Dravidian I" according to Krishnamurti 2003 (referred to as simply the South Dravidian subgroup in Andronov 1978: 8). Here, we take into consideration the data from Tamil, Malayalam, Kannada, Tulu, Irula and Badaga.

**Elder sister.** The equivalents for this meaning in South Dravidian are quite similar: Tam. *akkā*, Mal. *akka*, Kan. *akka*, Bad. *akka* 'sister (classificatory)', Irula *akka*, Tulu. *akka*, *akkè*. Thus, the supposed Proto-South-Dravidian equivalent is *\*akka* (cf. also StarLing 2023).

**Younger sister.** Tamil *taṅkai*, Mal. *taṅka*, *taṅkacci*, Kan. *taṅgi*, Bad. *tange*, Irula *tange*, *tammakke*, Tulu *taṅgi*, with the Proto-South-Dravidian equivalent reconstructible as *\*taṅg-*. This variant is the closest to the general Proto-Dravidian state, compared to the forms of other languages.

The meaning **elder brother** in South Dravidian is expressed by Tam. *aṅṅaṅ*, Mal. *aṅṅan*, Kan. *aṅṅa*, Bad. *aṅṅa* 'elder brother', 'mother's sister's son', Tulu *aṅṅe* 'elder brother', 'maternal uncle' (a rare case; in languages of Eurasia we often find the polysemy 'elder brother' = 'paternal uncle', but not 'maternal uncle'. It seems that these meanings are connected as different manifestations of an 'elder person'), 'term of address to an elderly man', Irula *aṅe*, cf. DEDR 1984: 14. The protoform is reconstructed as *\*aṅṅa* 'elder brother' both in Krishnamurti 2003: 131 and in StarLing 2023.

**Younger brother.** Tam. *tampi*, Mal. *tampi*, *tampān*, Kan. *tamma*, Irula *tambi*, *tamma*, Tulu *tammaiya*, Badaga *tamma*. The reconstructed form is as follows: *\*tamp-* 'younger brother' (StarLing 2023) and *\*tamp-V-* 'younger brother' (Krishnamurti 2003: 10). More specific meanings

recorded for this term in South Dravidian languages include ‘younger brother or male parallel cousin (younger than the speaker)’, ‘term of endearment’, ‘term of address for any boy or young man’, ‘title of some temple-priests’ (DEDR: 3085).

The symmetrical subsystem for sibling nominations in Tamil (I propose to adopt the term “South Dravidian square” for this structure), compared with its Proto-Dravidian sources, is given in Table 1.

	Modern Tamil		Reconstruction	
	<i>female</i> <sup>4</sup>	<i>male</i>	<i>female</i>	<i>male</i>
<i>elder</i>	akkāḷ	aṇṇan	PDR *akk-	PDR *āṇa-
<i>younger</i>	taṅkai	tampi	PDR *aṅ-	PDR *t-amp-

Table 1. Main sibling terms in Modern Tamil and their ancestry.

Some of these words clearly share certain phonetic properties and could be classified as so-called *Lallwörter* or “children’s words” – lexemes consisting of the most easily mastered (in reference to articulation) phonemes, which also happen to be the earliest items in a child’s language acquisition. Such stems, usually with a plosive consonant (labial, less often dental) and a sonorant, with a short mid vowel and repeating syllables, are mentioned in Jakobson 1985: 107, cf. English *mom, mummy, dad, daddy*; Russian *baba, deda* etc. Such words often occupy a separate niche in kinship system; they are used as appellatives in everyday conversation and can often be synonymously paired with another kinship term with a reference meaning. This “Lallwort” vocabulary tends to retain its articulation over the course of history; phonetic laws in comparative linguistics frequently do not apply to its properties. It has also been observed that they may be avoided in poetic speech and not be attested in ancient written sources (cf. Smirniškaia 2022: 185).

Although both the Modern Tamil sibling terms and their Proto-Dravidian sources can be seen to generally agree with the above-mentioned phonetic properties, the one important difference is that in this case, they do not behave as appellatives and there are no correlated synonyms for them from a different linguistic register; therefore, applying the term “Lallwörter” to this group of terms would be technically incorrect, even if historically it is quite probable that they do go back to actual “Lallwörter”.

### 3.2. Telugu

Telugu, the largest existing Dravidian language, belongs to the South Dravidian II group (according to Krishnamurti 2003) or to the South-Eastern group (according to Andronov 1978). While Krishnamurti unites it with Gondwana languages, Andronov in his turn assigns Telugu to its own special group. Both of these decisions indicate that, one way or the other, this language stands out from the rest of the Dravidian idioms in that region. G. Starostin also believes that Telugu represents a special subgroup within the Dravidian family (Starostin 2000: 13).

Throughout its history, Telugu has been in heavy contact with Indo-Aryan languages, especially with Sanskrit. As some researchers point out, “the number of words borrowed from Sanskrit and Prakrit accounts for about half of the total vocabulary of the Telugu language” (Gurov et al. 2013: 360). Kinship terms are not an exception.

<sup>4</sup> No evidence of distinction between kinship terms for male and female Ego is observed in our Dravidian data. Here the designation of gender refers to the gender of the sibling under discussion.

We can observe the preservation in Telugu of the Proto-Dravidian vocabulary in the expression of meanings indicating the relative age of a sibling: *annayya* < PDr \**āṇa*-<sup>5</sup> ‘elder brother’ along with *tammuḍu* < PDr \**t-amp-* ‘younger brother’; *akka* ‘elder sister’ < PDr \**ak-* ‘elder relative’ along with *celli* ‘younger sister’ < PDr \**cel-* ‘companion’, ‘sister’. Meanwhile, Indo-Aryan borrowings appear in cases where the parameter of relative age is not expressed: *sōdari* ‘sister’ < Skt. ‘sister’ with the derivative from the same root *sōdaruḍu* ‘brother’, and *bhrāta* ‘brother’ < Skt. ‘brother’, but their usage does not go beyond the bounds of high literary language.

Since Telugu, according to Andronov 1978 and Starostin 2000, forms a separate subgroup all by itself, we do not discuss the reconstructions of Proto-Telugu here (they more or less coincide with the forms of Classical Telugu), but simply list the modern forms together with their Proto-Dravidian ancestors in Table 2.

The Telugu form *tammuḍu* ‘younger brother’, extended by means of the productive masculine suffix *-ḍu*, corresponds to the Tamil form *tampi* ‘younger brother’: Burrow and Emeneau place both in the same entry (DEDR: 3085), while G. Starostin points that “although the labial series is considered one of the most stable in the Dravidian languages, [p] is dropped here in \**t-amp-* in this case”, see Starostin 2000: 99, 106.

	MODERN		RECONSTRUCTION	
	<i>female</i>	<i>male</i>	<i>female</i>	<i>male</i>
<i>elder</i>	akka	annayya	PDR * <i>akk-</i>	PDR * <i>āṇa-</i>
<i>younger</i>	celli	tammuḍu	PDR * <i>cel-</i>	PDR * <i>t-amp-</i>

Table 2. Main sibling terms in modern Telugu with reconstructions.

Another interesting fact is that the form *cellu* ‘younger sister’ < PDr \**cel-* ‘companion’, ‘sister’ in Telugu seems completely disconnected from the system observed in South Dravidian languages such as Tamil, where we usually find represented reflexes of PDr \**aṅ*: Tamil *taṅkai*, Mal. *taṅka*, *taṅkacci*, Kan. *taṅgi*, Bad. *tange*, Irula *tange*, *tammakke*, Tulu *taṅgi* etc. Though the term *cellu* is evidently of Proto-Dravidian origin (Krishnamurti 2003: 10), it is different in two ways. First, phonetically it does not fit in with the “South Dravidian square” (see the *Lallwörter* discussion in section 3.1 above) like the other terms of this group; second, the attested meanings of words formed with this root in cognate languages reflect a semantic connection between kinship and society relationships (see the detailed list of such meanings in DEDR #2783). This suggests that this form may have been inherited from a different system of oppositions. A similar pattern is described in the work Zhivlov et al. 2023, where, based on cross-linguistic comparison, several models of nomination of the concept “husband” in the languages of Eurasia are identified: of these, the main opposition is based on the “social” idea of a husband as an “outstanding member of society”, and the individual model of “husband” as “my man (*the* man)”. In a similar way, Tel. *cellu* may be implementing the “social” idea of “sister as an assistant, person who helps, a companion”, as opposed to the Proto-South-Dravidian \**taṅ-* (>Tamil *taṅkai* etc.), which does not have such a meaning and stays strictly within the semantic field of kinship terminology proper.

Other expressions denoting **sister**, mostly borrowed from Indo-Aryan, are: *anu-jāta* ‘younger sister’ (from Sanskrit ‘born together’ < Sanskrit *jāti* ‘birth’); *kanīyasi* ‘younger sister’;

<sup>5</sup> Proto-Dravidian accent is provisionally set up by G. Starostin to explain the distinction between geminated and non-geminated consonants in \*(C)VC(C)V-type structures (see Starostin 2000 for more details); more traditional reconstructions, e.g. Krishnamurti’s, simply reconstruct geminated consonants, e.g. \**aṅṅa-* instead of \**āṇa-*.

*kaniṣṭha* ‘younger sister’, *kaniṣṭhurālu* ‘youngest sister’; *pūrvaja* ‘elder sister’ (cf. Sanskrit *pūrvamu* ‘first’, ‘former’, ‘ancient’); *bhagini* ‘sister’; *sahaja*, *svasa* ‘sister’. These terms are not basic and not so frequent, mostly restricted to the literary language.

For the meaning **brother**, dictionaries also indicate additional forms: *kanīyudu* ‘younger brother’; *kaniṣṭhudu* ‘youngest brother’; *agrajanmudu* ‘elder brother’ (from Sanskrit *agramu* ‘end’, ‘front’, ‘top’, ‘first’, ‘chief’, from the same stem comes *agrani* ‘leader’); *jaghanyujudu* ‘younger brother’; *pūrvajudu* ‘elder brother’; *bhrāta* (from Sanskrit) ‘brother’ (also related to the compound form *bhrātri-snēhamu* ‘brotherly love’); *venukaṭi-vādu* lit. ‘the one who [appeared] after’ is often used in the sense of ‘younger brother’ (cf. *venuka* ‘before’, ‘behind’). As a result of the fight against Sanskritisms in Telugu, we also have the compound form *tōda butṭuwu*, lit. ‘born together’, produced by analogy with *sahōdarudu*. It may have the meanings of either ‘brother’ or ‘sister’ depending on the context, but its use is limited to literary language.

### 3.3. Gondwana Dravidian languages

The “Gondwana” subgroup includes Dravidian languages spoken in Central India, in the states of Maharashtra, Madhya Pradesh, Chhattisgarh, Andhra Pradesh and Orissa. They belong to the “Gondwana” subgroup under the classification by M. S. Andronov (Andronov 1978: 8), while Bh. Krishnamurti refers to them as the Central-South (or South Dravidian II) group (Krishnamurti 2003: 19).

The ancient supercontinent Gondwana was named after this region by archaeologists; its name goes back to Sanskrit *gondavana* “Gondwana Forest”, with the root *gonda-* in turn believed to come from Proto-Dravidian *\*kō* “mountain”, see Gurov et al. 2013. The Gondwana subgroup consists of Gondi language (with numerous dialects, including Koya), Pengo, Konda, Manda, Kui, Kuwi. The majority of speakers of these languages are bilingual; the second language is, depending on the place of residence, either Telugu (the official language of Andhra Pradesh) or Oriya (the official language of Orissa). Most of these languages have been rather poorly studied. There are lexicographic sources and some grammatical descriptions for most of them, but almost no comprehensive dictionaries. The sibling terminology for Gondwana languages is shown in Table 3.

The meaning **younger sister** is expressed in Kui as *angi*, *tangi* ‘younger sister’; Kuwi *angi*, *tangi*, *bōpi* ‘younger sister’; Konda *taṇi(si)* ‘younger sister’; Manda: *e(ē)mi* ‘younger sister’; Gondi: *sēlaḍ* ‘younger sister’. Pengo has no recorded term for ‘a younger sister’, only for ‘sister’ in general.

We can trace the remains of the PDR form *\*ang-* in some, but not in all the languages. In Pengo there is no distinction of younger vs. elder siblings altogether; the curious Manda form *e(ē)mi* ‘younger sister’ stays without a plausible etymology. Gondi *sēlaḍ* ‘younger sister’ could be construed as a borrowing from Telugu *celli* id., but this is not highly likely, given that cognates of the same root are also found in Gadaba *cellel* ‘younger sister’ and Parji *cālal* ‘sister’ (cf. also Tamil *cilati* ‘female servant or companion’). A more likely scenario than borrowing would imply the antiquity of the usage of this root to express the concept of “social sister”, perhaps going all the way back to Proto-Dravidian. Such semantics could then function as the basis for parallel independent innovations towards the meaning ‘sister’ (proper kinship term) in different languages. Of course, the geographic proximity to and cultural influence of Telugu here could have also influenced the semantic development of this word in Gondi, Gadaba and other languages.

The meaning **elder sister** is expressed by Kui *bāi* (*bāiḍi*) ‘elder sister, cousin’; Kuwi *nāna* ‘elder sister’; Manda *nana* ‘elder sister’, *boyni* ‘sister’; Pengo *nāna* ‘elder sister’, *torndel* ‘sister’;

Konda *bībi* / *bībsi* ‘elder sister’, ‘brother’s wife’, ‘mother’s mother’ (Persian loan through Indo-Aryan, cf. also the term *buba* ‘father’ in this language, used for 2<sup>nd</sup> person as an address with suppletion *apo(si)* for 3<sup>rd</sup> person); Gondi *takka* ‘elder sister’, *bāyi* ‘elder sister’.

The meaning **elder brother** is expressed by Kui and Kuwi *dāda* ‘elder brother’; Manda *dāda* ‘elder brother’, *bāyi* ‘brother’; Pengo *dāda* ‘elder brother’, *bay* ‘brother’, *ṭonḍen* ‘brother’; Gondi *dādal* ‘elder brother, father, grandfather’, *tādāl* ‘elder brother’, *taṇṇē* ‘elder brother’; Konda *ana(si)* ‘elder brother’ and *dāda* ‘elder brother’, ‘grandfather’.

The meaning **younger brother** is expressed by Kui *ambesa* (*tambesa*), *jūsu*, *āu* ‘younger brother’; Kuwi *tayi* and *bōva* ‘younger brother’; Manda *buḍa* ‘younger brother’; Gondi *kōko* ‘younger brother’, *tammūr* ‘younger brother’; Konda *tamberi* ‘younger brother’. No words for younger relatives have been found in the lexicographic sources on Pengo that we have examined. In Konda there is also a form *bāy* ‘brother’ without expression of the idea of relative age, apparently of Indo-Aryan origin.

		Modern Languages		
		<i>female</i>		<i>male</i>
<i>elder</i>	Kui: <i>bāi</i> Kuwi: <i>nāna</i> Manda: <i>nana</i> Pengo: <i>nāna</i> Gondi: <i>bāyi</i> , <i>takka</i> Konda: <i>bībi</i> , <i>bīb(si)</i>	Manda: <i>boyni</i> , Pengo: <i>toṛndel</i>	Kui: <i>dāda</i> Kuwi: <i>dāda</i> Manda: <i>dāda</i> Pengo: <i>dāda</i> Gondi: <i>taṇṇē</i> Konda: <i>dāda</i> , <i>ana(si)</i>	Manda: <i>bāyi</i> Pengo: <i>bay</i> Pengo: <i>ṭonḍen</i>
<i>younger</i>	Kui: <i>angi</i> , <i>tangi</i> ; Kuwi: <i>angi</i> , <i>tangi</i> , <i>bōpi</i> Manda: <i>e(ē)mi</i> Gondi: <i>sēlaḍ</i> Konda: <i>taṇi(si)</i>		Kui: ( <i>t</i> ) <i>ambesa</i> , <i>jūsu</i> , <i>āu</i> Kuwi: <i>tayi</i> , <i>bōva</i> Manda: <i>buḍa</i> Gondi: <i>kōko</i> , <i>tammūr</i> ; Konda: <i>tamberi(si)</i>	

Table 3. Main sibling terms in languages of the Gondwana subgroup.

To summarize, all the sibling forms of Gondwana Dravidian languages with their possible origin are shown in Table 4:

		Lexical sources		
		<i>female</i>		<i>male</i>
<i>elder</i>	LOAN-OIA: <i>bhai</i> , <i>nana</i> , <i>bībi</i> PDr * <i>ak-</i>	LOAN-OIA: <i>boyni</i> ; PDr * <i>tōz-</i> .	LOAN-OIA: OIA: * <i>dāda</i> PDr * <i>āṇa-</i>	LOAN-OIA: <i>bāyi</i> ; PDr * <i>tōz-</i>
<i>younger</i>	PDr * <i>aṅ-</i> PDr * <i>cel-</i> LOAN-UN: Kuwi: <i>bōpi</i> Manda: <i>e(ē)mi</i>		PDr * <i>t-amp-</i> I-A Gondi: <i>kōko</i>  LOAN-UN: Kui: <i>jūsu</i> , <i>āu</i> Kuwi: <i>tayi</i> , <i>bōva</i> Gondi: <i>tammūr</i> Manda: <i>buḍa</i>	

Table 4. Lexical sources for sibling terms in Gondwana languages.



As we can see in Gondwana languages, it is a possible and even widespread situation when there is more than one term denoting a sibling of the same relational category with Ego. In some cases the terms belong to different systems: one that is inherited and one that is represented by borrowings from one or another contact language, cf. in Pengo *tonḍen* 'brother' < PDr \**tōz-* while *bay* 'brother' < Oriya *bhāi* 'brother' < Sanskrit *bhrātṛ* 'brother' (Turner 1966: #9661). These terms do not show the relative age distinction, and the same situation is seen in Gondi where *takka* 'elder sister' < PDr \**akk-* 'elder relative', while at the same time *bāyi* 'elder sister' is borrowed from Indo-Aryan.

We can suggest that, according to a typical contact-induced scenario, as described, for example, in Matras 2009, the "minor" Dravidian languages borrowed the sibling terms from their larger and more prestigious Indo-Aryan neighbors. Such borrowing always happens in stages: first, when the item is borrowed, it co-exists along with the earlier inherited term, and only after some time it ends up fully replacing the older term. Moreover, we observe that kinship terms are not always borrowed individually; the system, or at least a large part of it, is often borrowed as a whole. For example, if there is a borrowed term for 'elder sister', e.g. Pengo *nāna*, we may expect to simultaneously find a borrowed term for the meaning 'elder brother', e.g. Pengo *dāda*, with both loanwords coming from the same source (in this case, Oriya).

Another obvious source of borrowing in this region might have been the neighboring Munda languages, but almost no compatible terms for Munda siblings have been found (see Parkin 1985 on Munda kinship terms). The only case that could potentially be traced back to a Munda source is Manda (Dravidian) *buḍa* 'younger brother'; it has a possible parallel in Gutob (Munda) *buḍi* 'younger sister, sister-in-law (husband's younger brother's wife), stepsister'. Even so, Norman Zide<sup>6</sup> suggests that this word was, in its turn, borrowed from the Indo-Aryan language Desiya, the *lingua franca* of that region, where it has the meaning 'small girl' (Zide and Das 1963); since it is further comparable to Prakrit *bōḍa-* 'young' (Turner 1966: 524, #9268), its ultimate origin once again goes back to Indo-Aryan languages rather than Munda.

Still, although the majority of borrowed sibling terms in Gondwana (as well as other Dravidian) languages can be attributed to Indo-Aryan sources, the origin of certain borrowings remains unclear, such as Kuwi *bōpi*, Manda *e(ē)mi* 'younger sister'; Kui *jūsu*, *āu*; Kuwi *tayi*, *bōva*; Manda *buḍa* 'younger brother'. With neither Dravidian, nor Munda, nor Indo-Aryan etymologies readily available for these words, one could provisionally assume the influence of certain old substrates — for example, a subset of terminology remaining from the pre-Dravidian languages of the corresponding regions, which were home to many tribes who switched to Dravidian languages after the arrival of the Dravidian migration, but may have left behind some substratum elements. Given that these lexical elements more frequently apply to younger siblings — an area of kinship terminology that is more resistant to borrowing, which will be shown in section 5 — this is an additional argument for their earlier, substrate-related origin. However, this is a rather speculative assumption, warranting further search for any potential ancient cognates of these items.

### 3.4. Central Dravidian languages: Kolami, Parji, Gadaba

Central Dravidian is a small subgroup of Dravidian languages spoken in the central regions of India, in the states of Maharashtra, Andhra Pradesh, Chhattisgarh and Orissa. According to contemporary historical linguists, this branch separated from the Proto-Dravidian community

<sup>6</sup> The author would like to express her sincere gratitude to Evgenia Renkovskaya for discussing the examples from Munda languages.

relatively early, approximately around 1200 BC (Starostin 2000: 13). The group includes Parji, Gadaba, Kolami, Naikri and Naiki languages (the last two shall not be covered in the present paper due to scarceness of information).

Native speakers of these languages may also be bilingual in certain Indo-Aryan languages, such as Marathi for Kolami, Halbi and Bhatari for Parji, Oriya for Gadaba. As for contacting languages of Dravidian origin, speakers of Kolami and Gadaba living in Andhra Pradesh can also speak Telugu, and many of the people in this group also have some knowledge of certain dialects of Gondi. Table 5 provides the list of attested forms for siblings together with their origins.

The meaning **elder sister** finds the following equivalents in Central Dravidian: Kolami *akkābai*, *dō bai*; Gadaba *kāko*. The meaning **younger sister** is expressed in Kolami as *tōrndal*, *sinnam bai*; in Gadaba as *cellel*. In Parji there is no distinction.

The meaning **elder brother** corresponds to Kolami *dādak*, *dō tōren*; Parji *toled*, *pāva*; Gadaba *dāda*, *tōḍṇ*; the meaning **younger brother** is expressed as Kolami *bāyi*, *tōren*; Parji *toled*; Gadaba *bāyi*, *tōḍṇ*. As we can see below, in Parji no relative age distinction for the meaning ‘sister’ is observed; known lexicographic sources only adduce the lexeme *cālal* ‘sister’.

	Modern		Reconstruction and lexical sources	
	<i>female</i>	<i>male</i>	<i>female</i>	<i>male</i>
<i>elder</i>	Kol. <i>akkābai</i> , <i>dō bai</i> Gad. <i>kāko</i>	Kol. <i>dādak</i> , <i>dō tōren</i> Gad. <i>dāda</i> , <i>tōḍṇ</i> Parj. <i>toled</i> , <i>pāva</i>	PDR *ak — only Kol. PDR *cel- Parj.  Also, borrowed: LOAN-OIA*kaka: Gad. <i>kāko</i>	PDR *āṇa — no reflexes found PDR: *tōz- — Kol, Gad., Parj.  Also, a parallel existing system — LOAN–Mar <i>dada</i> or OIA * <i>dada</i> (Turner 1966: #6261): Kol. <i>dādak</i> , gad. <i>dāda</i> LOAN-OIA * <i>bāppa</i> ‘father’. (Turner 1966: #9209): Parj. <i>pāva</i>
<i>younger</i>	Kol. <i>tōrndal</i> , <i>sinnam bai</i> ; Gad. <i>cellel</i>	Kol. <i>bāyi</i> , <i>tōren</i> ; Gad. <i>bāyi</i> , <i>tōḍṇ</i> ; Parji <i>toled</i>	PDR *aṅ- — no reflexes PDR *cel- Parji  PDR *čīn- together with borrowed adjective <i>sinnam bai</i> (< OIA <i>bhāi</i> )	PDR *t-amp- — no reflexes PDR: *tōz — Kol. <i>tōren</i> ; gad. <i>tōḍṇ</i> ; parji <i>toled</i>  LOAN — Mar <i>bhaginī</i> ‘sister’ LOAN — OIA Kol. <i>bāyi</i> from OIA <i>bhāi</i>
no relative age distinction	Parj. <i>cālal</i> ‘sister’.	no term		

Table 5. Sibling terms in Central Dravidian languages with their sources

We see in these examples many borrowings, especially from Indo-Aryan languages, such as Kolami *dādak*, Gadaba *dāda* from Marathi *dada* (or, perhaps, directly from Old Indo-Aryan \**dada* (Turner 1966: #6261). They form the core of the sibling system, occupying most of the relevant slots. A few meanings still retain inherited Dravidian equivalents, such as PDR: \**tōz-* → Kolami *tōren*; Gadaba *tōḍṇ*; Parji *toled*. However, reflexes of this stem seem to appear in all instances only as an additional term, complementary to those expressed by Indo-Aryan borrowings. Since they always express different meanings (‘younger sister’, ‘younger brother’, ‘elder brother’), we may assume that its earlier stage meaning was closer to a general ‘sibling’, ‘kin’.

Borrowed lexemes are sometimes intertwined with original Dravidian ones, forming combinations such as Kolami *sinnam bai* ‘younger sister’, consisting of *sinnam* ‘younger’ < PDR \**č̣in-* ‘small’ and *bai* ‘sister’ < OIA *bhāi* ‘sister’. Another interesting case is Kolami *akkābai*, comprised of two iterations of one and the same meaning, one of Dravidian and one of Indo-Aryan origin: *akka* < PDR \**akk-* ‘elder relative, elder sister’ (the only remnant of this item across all Central Dravidian systems) and *bai* ‘sister’ < OIA *bhāi* ‘sister’.

Reflexes of PDR \**aŋ-* ‘younger relative’ (brother/sister) did not remain in the sibling system, having been relocated to lexical periphery. However, we still find their traces in the kinship domain: according to Setumadhava Rao, descendants of this stem are still found in Kolami dialects, e.g. *taŋgod* ‘husband's elder brother’; cf. also Naikri *taŋgon*, *taŋgol* ‘wife's elder brother’; Naiki *taŋgon* ‘husband's or wife's elder brother’ (StarLing 2023). Derived forms also include Kolami (dialectal) *taŋgodā* ‘wife's elder sister’; Naikri *taŋgodal* ‘wife's elder sister’; Naiki *taŋgoda* ‘husband's elder sister’ and possibly Gadaba (dialectal) *naggal* ‘wife's younger sister’.

The opposition “elder ~ younger” is occasionally neutralized, but only for the meaning ‘sister’, as manifested by Parji *cālal* < PDR \**cel-* ‘companion, sister’. This may be the result of the influence of neighboring Indo-Aryan kinship terms system in which there is no such opposition.

### 3.5. North Dravidian: Malto

Malto is one of the two languages belonging to the North Dravidian (Kurukh-Malto) group. According to the classification model of G. Starostin, separation of the Kurukh-Malto subgroup is dated back to around the 3rd millennium BC, with Malto emerging as an independent language already in the 1st millennium AD (Starostin 2000: 15).

Malto is spoken in the Rajmahal Hills in the far Northeastern Indian state of Jharkhand, with speakers also living in West Bengal, the states of Tripura and Orissa. Currently, there are about 340 thousand speakers of Malto including different dialects; many of them also speak Bengali (an Indo-Aryan branch of the Indo-European language family) or Santali (Munda language family).

The semantics of **brother** in Malto is usually expressed by the lexemes *baya* ‘elder brother’ and a lexical set denoting possession, such as *taŋḍo* ‘his or her younger brother or sister’ and others; also by the meanings *undḡle* ‘brother from the same parents’; *nuna* ‘affectionate name for a son or younger brother’, possibly related to Oriya *nunu* ‘son’ and *nuni* ‘eldest daughter’, ‘a small non-Brāhmaṇa girl’; *dúbáy* ‘sisters and brothers from the same parents’.

Lexemes used to express the semantics of **sister** are *báyi* ‘elder sister’ of Indo-Aryan origin, *undḡl̥ni* ‘sister from the same parents’, *nuni* ‘affectionate name for a younger sister or daughter’ and other terms cited above denoting ‘younger sister or brother’ together, as *engḍo* ‘my younger brother or sister’ etc. The forms with their reconstruction are shown in Table 6.

An important feature of this system is the consistent fused designation of younger siblings with terms that include earlier prefixes of inalienable possession: *taŋḍo* ‘his or her younger brother or sister’, *niŋḍo* ‘your younger brother or sister’. For some reason, such prefixes were preserved only in designations of younger siblings. Typically, this could be regarded as an archaic element, hinting at an inherited Dravidian origin for the root *ḍo* ‘younger brother or sister’; but no parallels are found in other Dravidian languages for this root which, furthermore, also violates the commonly accepted phonotactics for Proto-Dravidian (featuring a retroflex consonant in root-initial position). This is an interesting and difficult historical issue that warrants further investigation.

	Modern		Reconstruction	
	<i>female</i>	<i>male</i>	<i>female</i>	<i>male</i>
<i>elder</i>	<i>bāyi</i>	<i>baya</i>	LOAN-OIA <i>bhaginī</i> 'sister'	LOAN-OIA <i>bhrāṭṛ</i> 'brother' (Turner 1966: #552) through Beng. <i>bhāi</i> 'brother'
<i>younger</i>	<i>nuni</i> 'affectionate name for a younger sister or daughter'	<i>nuna</i> 'affectionate name for a son or younger brother'	?LOAN-Or: <i>nunu</i> 'son'	LOAN-Or <i>nunu</i> 'son'
	- <i>ḍo</i> 'younger brother or sister' (with the prepositive marker of possession)	- <i>ḍo</i> 'younger brother or sister' (with the prepositive marker of possession)	??LOAN – source unknown	

Table 6. Main sibling terms in Malto.

### 3.6. North Dravidian: Kurukh

Kurukh (Kurux, Oraon) is a language of the North Dravidian group, spoken in the territory of northeastern India (the states of Jharkhand, Bihar, Orissa, small communities also in Madhya Pradesh, Assam, Tripura, Chhattisgarh, West Bengal), as well as in the Republic of Bangladesh and southern Nepal. Most speakers are bilingual in Indo-Aryan: Hindi, Bengali, Bhojpuri. The terms for sibling in Kurukh with their reconstruction are given in Table 7.

The meaning **sister** is expressed in Kurukh with the lexemes *bhaiyā* 'younger sister' (also *bhiyā*, *hiyā*, *biā*), *-dī* 'younger sister', only in combination with possessive prepositional markers: *ingrī* 'my or our younger sister', *ningdī* 'your younger sister', *tangrī* 'his, her, their younger sister'; *-dāi* 'elder sister': *engdāi*, *ningdāi*.

The meaning **brother** is expressed by the lexemes *bhāis* 'brother', 'cousin', 'respectful treatment between equals' (also *bhaiyos*), *bhaiyas* 'younger brother', 'benevolent address to the younger', 'jagirdar, land owner'; also *bias*, as well as *dadas*, *dās* 'elder brother', 'elder brother of wife, husband': *engdadas* 'my elder brother', *ningdadas* 'your elder brother', etc.; *nimhai dās* 'your elder brother'; *ningri(s)* 'your younger brother', *tangri(s)* 'his younger brother'.

	Modern		Reconstruction	
	<i>female</i>	<i>male</i>	<i>female</i>	<i>male</i>
<i>elder</i>	<i>-dāi</i> , with possessive: <i>engdāi</i> etc.	<i>dadas</i> , <i>dās</i> with inalienable possessive <i>engdadas</i> etc.	LOAN – OIA unknown for <i>-dāi</i> , or maybe Dravidian source	LOAN-OIA * <i>dādda</i> 'father or other elderly relative' (Turner 1966: #6261) through Hindi <i>dādā</i> 'father, elder brother'
<i>younger</i>	<i>bhaiyā</i> ; <i>-dī</i> , only in combination with inalienable possessive: <i>ningdī</i> 'your younger sister'	<i>bhaiyas</i> 'younger brother'; <i>ningri(s)</i> 'your younger brother'	LOAN-OIA cf. Beng. <i>bhaginī</i> 'sister'  LOAN-Un for * <i>-dī</i> , or maybe Dravidian source	LOAN-OIA cf. Beng <i>bāyi</i>  LOAN-Un for * <i>-dī(s)</i> , or maybe Dravidian source

Table 7. Main sibling terms in Kurukh.

In this language we can see many forms of Indo-Aryan descent: *bhaiyā* 'younger sister' < OIA; *bhāis* 'brother', 'cousin', 'respectful treatment between equals' < OIA. Even the form with

prepositive inalienable possession marker appears to be of Indo-Aryan origin: *engdadas* ‘my elder brother’ < OIA, cf. *\*dādda* ‘father or other elderly relative’.

No traces of reflexes of the old Dravidian forms for ‘elder brother’ and ‘elder sister’ were found. Neither dictionaries of Kurukh, nor the etymological material in DEDR 1984 give any evidence for any remnants of PDR *\*āṇa* ‘elder brother’ or *\*ak-* ‘elder relative, elder sister’.

In Kurukh, the markers of inalienable possession are regular and, unlike other languages in which such forms are found, are used with almost all the elements of the kinship term system: *embas* ‘my father’, *nimbās* ‘his father’, *tambās* ‘their father’. This regularity may be explained by assuming that at the time of the separation of Kurukh-Malto from Common Dravidian, this prefixal system was still relatively productive, and remained so throughout the more recent history of Kurukh, with the prefixes commonly added even to new terms of borrowed origin.

If so, another phonetic decomposition of the forms *tangrī* ‘his younger sister’ and *tangrī(s)* ‘his younger brother’ is possible. We propose that the original marker of inalienable possession for the 3rd person for Kurukh was not *\*t-* but rather *\*taŋ-*. According to Starostin 2000, referring to Kurukh, “here /d/ and /r/ may be reflexes of an earlier /d/”, so the forms *tangrī* and *tangrī(s)* more probably consist of the earlier prefix *\*taŋ-* + a corresponding sibling term. Thus, at the level of Proto-North-Dravidian Kurukh *tangrīs* ‘younger brother’ can be decomposed into *\*taŋ-* + *\*di* ‘younger brother’, *tangrī* ‘younger sister’ into *\*taŋ-* + *\*dī*. Since there is not much evidence to reconstruct a morphological difference between protoforms depending on gender (cf., however, notes on final *-i* as a feminine suffix in the Classical Tamil period in Wilden 2018: 31), with the exact same form for both genders in Malto, this opposition in Kurukh is probably innovative and can hardly be traced back to Proto-Kurukh-Malto, let alone Proto-Dravidian.

Such an analysis, however, cannot be applied to other groups of Dravidian languages, in particular, the South and Central branches. Thus, Tamil *tampi* ‘younger brother’ goes back to PDR *\*t-amp-*, consisting of the inalienable possession prefix *\*t-* and the stem *\*-amp-* ‘younger brother’; this form cannot reflect an earlier prefix *taŋ-*. The existence of Tamil forms with and without this prefix can be considered as an argument for this hypothesis: *ampi* ‘younger brother’ (Brahman dialect or colloquial), *n-ampi* ‘the elite among men’, ‘a term of endearment’, etc.; outside of South Dravidian, cf. the prefixless variant in Kui *ambesa* / *t-ambesa* ‘younger brother’. On the whole, all those forms agree with the reconstruction of a simple system of prefixes of inalienable possession for Proto-Dravidian: *\*t-* (3rd person), *\*n-* (2nd person), *\*y-/0-* (1st person).

If so, the longer forms of prefixes in Kurukh-Malto must be understood as a secondary development, during which the old system was analogically restructured and the old prefixes replaced by the full forms of indirect (oblique) pronouns: in Kurukh these are *eŋg-* 1st p. sg., *niŋg-* 2nd p. sg. and *taŋg-* 3rd p. sg., see discussion in Kobayashi & Tirkey 2017: 85; for the Malto forms, see section 3.5 above.

The stems for ‘younger siblings’ in Kurukh and Malto can hardly be inherited from Proto-Dravidian, even if we hypothetically suggest that they have been lost in all other branches of the family. First, Kurukh *-dī* / *-di* and Malto *-ḍo* cannot be reflexes of the same protoform. Second, the Malto stem *-ḍo* is atypical for Dravidian languages which usually do not allow retroflex consonants in root-initial position. Finally, assuming these roots’ antiquity in the required meanings of ‘younger brother’ and ‘younger sister’ requires an explanation for the appearance of *\*t-aŋ-* ‘younger sister’ and *\*t-amp-* ‘younger brother’ in the other branches — roots that are far more compatible with the phonetic typology of Proto-Dravidian than the ones in Kurukh-Malto. Given all these considerations, even if we cannot pinpoint an exact source of provenance for the Kurukh-Malto forms, it is far more likely that they are of contact than inherited origin.

## 3.7. North Dravidian: Brahui

Brahui is the northernmost of all the Dravidian languages, spoken in Pakistan, Iran and Afghanistan. Its separation from the rest of Dravidian community, according to lexicostatistical data, was the earliest split from Proto-Dravidian, taking place at the end of the 3rd millennium BC (Starostin 2000). Today, Brahui is the only language spoken so far to the North, being surrounded by languages of other families. Most of the speakers are bilingual with Balochi (of Indo-Iranian origin). The sibling terms of Brahui and their possible sources are shown in Table 8.

	Modern		Reconstruction	
	<i>female</i>	<i>male</i>	<i>female</i>	<i>male</i>
<i>no relative age distinction</i>	<i>iṛ</i>  <i>addī</i>	<i>īlum</i> ;  <i>adā</i>	PDR *īl-[aj-] ‘young’ LOAN-Bal. <i>aḍḍī</i> ‘sister’	PDR *īl-[aj-] ‘young’ LOAN-Bal. <i>aḍḍā</i> ‘brother’
<i>younger</i>	no special term	<i>īlumkō</i> ‘younger or youngest brother’	no term	PDR *īl-[aj-] ‘young’

Table 8: Main sibling terms in Brahui

In Brahui the factor of relative age is not taken into consideration; this parameter does not distinguish the meanings of ‘younger sister’ and ‘elder sister’ as in the majority of other Dravidian languages. Interestingly, the subsystem of siblings as a whole has two types of expression — one of Dravidian and the other of borrowed Iranian origin. There is a pair of sibling terms that form a group of borrowed origin: *adā* ‘brother’; (rare) ‘son’; ‘respectful address to a younger man’ ~ cf. Balochi (Iranian family) *aḍḍā* ‘brother (familiar)’. Also *addī* ‘sister’, (rare) ‘daughter’; ‘respectful address to a woman’ ~ cf. Balochi *aḍḍī* ‘sister’. Judging by the fact that the borrowings from Balochi have hardly undergone any phonetic change, they must have entered the Brahui vocabulary quite recently.

Parallel Dravidian forms for siblings are: *īlum* ‘brother’, *īlumkō* ‘younger or youngest brother’ and *iṛ* ‘sister’ < PDR \*īl-[aj-] ‘young’ (StarLing 2023). Possibly connected with these forms is also Brahui *lumma* ‘mother’ which, according to (Bray 1934), also derives from the stem \*īl-[aj-] ‘young’; however, this etymology is questionable for phonetic and morphological reasons.

Moreover, this etymology seems doubtful from the point of view of semantics and the data of semantic typology. The semantic derivation of *īlum* ‘brother’ from PDR \*īl-[aj-] ‘young’ can be explained as a result of possible shift *young* → *younger sibling*, with subsequent expansion of the meaning to other kinship terms (compare for example with the existing shifts #8334 ‘little, small’ → ‘child’ and #0159 ‘young’ → ‘husband’ in Zalizniak et al. 2023).

The reconstruction proposed in Bray 1934 for Brahui *lumma* ‘mother’ appears doubtful from this point of view, since no direct semantic derivation *young* → *mother* seems to be possible. We can propose only one hypothetical scheme for the emergence of such a semantic development, with a transitional stage during which a compound was formed from PDR \*īl + PDR \*āma- ‘mother’ = initially ‘young mother’ or something of the kind (‘young mother who recently gave birth to a child’, ‘young mother with small children’, etc., or simply ‘a mother of really young age’, said about girls that were married in their early years and gave birth to a child while still being children themselves). Phonetic contraction of the compound could then result in a form like *lumma*; however, this is still a hypothesis waiting for additional arguments.

#### 4. Inalienable possession

One of the most important features of the Dravidian kinship system, already mentioned several times in relation to specific branches of this family, since its traces are found everywhere from the Northern Dravidian languages Kurukh and Malto to Old Tamil, is the feature of inalienable possession, expressed in kinship terms as a kind of prefix (even though grammatical meanings are usually expressed in these languages by suffixes, not prefixes). There is a clear etymological connection between such prefixes, e.g., in Old Tamil *entai* ‘my father’, *nuntai* ‘your father’ and *tantai* ‘his, her or their father’ and in Malto *engḍo* ‘my younger brother or sister’, *ningḍo* ‘your younger brother or sister’, *tangḍo* ‘his or her younger brother or sister’ (see section 3.5).

In modern Tamil, as well as in the majority of Dravidian languages, this feature has not been preserved, except for the forms that originally referred to the 3rd: cf. Tam. *tantai* ‘father’, *tāy* ‘mother’, or Gondi *taṇṇē* ‘elder brother’ etc. — in all these forms, the original prefix has fused with the root and became desemantized. However, in modern Kurukh and Malto reflexes of this system still seem to remain quite transparent and even somewhat productive. According to M. B. Emeneau’s conclusions, this feature was a specific property of the Proto-Dravidian language (Emeneau 1953: 346). The loss of forms of kinship with inalienable possession by the majority of Dravidian languages is sometimes considered an important marker of the development of Dravidian languages “from active towards a nominative type system” (Gurov 2013: 45). Note, however, that quite a few spoken dialects of Tamil still preserve traces of the old system, as shown in Table 9.

In some cases, differentiated marking has been preserved only for 2nd person, as in the Ceylon dialect and Tamil dialects of Kanyakumari and Ramanadhapuram. This similarity is of an areal nature, since these districts are located on the coast of the Indian Ocean, on the Southern tip of India (Kanyakumari) and on the Southeastern coast of India (Ramanadhapuram), in the areas closest to Ceylon.

It may be seen from this table that the Proto-Kurukh-Malto system and the systems in other Dravidian languages, most notably Tamil, cannot be derived from each other. The Kurukh-Malto system is basically identical with the indirect stems of Kurukh-Malto personal pronouns; the South Dravidian system (occasional relics of which can be found in some Central Dravidian languages as well) consists of simpler prefixes *\*y-* ‘1st p.’ (in Tamil *\*y-a-* → *e-*, *\*y-ā-* → *yā-*), *\*n-* ‘2nd p.’ (it is not quite clear if the Tamil variant *nu-*, unquestionably related to the Tamil oblique stem *un-*, is archaic or innovative), *\*t-* ‘3rd p.’. The most logical scenario is that the system in Kurukh-Malto was rebuilt (relatively recently) in accordance with its more current stems for personal pronouns.

#### 5. Conclusion

In this paper we have presented an analysis of the system of sibling terms in Dravidian languages. According to our data, the terminology system that exists today in South Dravidian languages with its distinction of relative age has every reason to be the reflection of an ancient Proto-Dravidian system of sibling terms. The main system most probably consisted of the same stems that continue to be used in South Dravidian languages, organized in what we earlier called the “South Dravidian square”: PDR *\*ak-* ‘elder sister’; PDR *\*áṇa-* ‘elder brother’; PDR *\*(t)-aṇ-* ‘younger sister’ and PDR *\*(t)-amp-* ‘younger brother’. The two latter forms (but probably not the two first ones) could, at the PDR stage, freely combine with the 3rd person

	1 my, our	2 your	3 his, her, their
Malto	<i>engḍo</i> (younger brother or sister); <i>engade</i> (son); <i>engadi</i> (daughter); <i>ayya</i> (mother)	<i>ningḍo</i> (younger brother or sister); <i>ningade</i> (son); <i>ningadi</i> (daughter); <i>ijjo</i> (mother); <i>ijjo-abbor</i> (father and mother, parents)	<i>tangḍo</i> (younger brother or sister); <i>tangade</i> (son); <i>tangadi</i> (daughter); <i>teho</i> (mother); <i>teho-tamba-kor</i> (father and mother, parents)
Kurukh	<i>embas</i> (father); <i>ingyō</i> (mother); <i>ingdī</i> (younger sister); <i>engdāi</i> (elder sister); <i>engri(s)</i> (younger brother); <i>engdā</i> (daughter)	<i>nimbās</i> (father); <i>ningyō</i> (mother); <i>ningdī</i> (younger sister); <i>ningdāi</i> (elder sister); <i>ningri(s)</i> (younger brother); <i>ningdā</i> (daughter)	<i>tambās</i> (father); <i>tangyō</i> (mother); <i>tangri</i> (younger sister); <i>tangdāi</i> (elder sister); <i>tangri(s)</i> (younger brother); <i>tangdā</i> (daughter)
Old Tamil	<i>entai</i> (father); <i>yāy</i> (mother); <i>enkai</i> (younger sister); <i>evvai</i> (cecrpa); <i>empi</i> (younger brother); <i>emmun</i> (elder brother)	<i>nuntai</i> (father); <i>nāy</i> (mother); <i>nunkai</i> (younger sister); <i>nuvvai</i> (cecrpa); <i>numpi</i> , <i>umpi</i> (younger brother); <i>nummun</i> (elder brother)	<i>tantai</i> (father); <i>tāy</i> (mother); <i>tankai</i> (younger sister); <i>tampi</i> (younger brother); <i>tammun</i> (elder brother)
Ceylon dialect of Tamil	<i>ayyalapu</i> (father); <i>ammā/āci</i> (mother); <i>akkā</i> (sister); <i>aṇṇan</i> (elder brother)	<i>koyyā/kopu</i> (father); <i>kommā/kōci/kōcci</i> (mother); <i>akkā</i> (your elder sister); <i>koṇṇan</i> (your elder brother)	
Tamil dialect of Kanyakumari	<i>appan</i> (father in general); <i>am-mān</i> (uncle, mother's brother in general.); <i>aṇṇan</i> (elder brother in general)	<i>koppān</i> (your father); <i>kommān</i> (your uncle, mother's brother), <i>koṇṇan</i> (your elder brother)	
Tamil dialect of Ramanadhapuram		<i>ṇoppān</i> (your father), <i>ṇommā</i> (your mother), <i>ṇokkā</i> (your elder sister), <i>ṇoṇṇān</i> (your elder brother)	
Suggested reconstruction for Proto-Kurukh-Malto level	*eng- (< *en-k-)	*ning- (< *nin-k-)	*tang- (< *tan-k-)
Suggested reconstruction for Proto-Dravidian	*y-	*n- ~ *nu-	*t-

Table 9. The preserved forms with markers of inalienable possession in Dravidian

prefix *\*t-* (data on their possible combination with 1st and 2nd p. prefixes are lacking, which does not formally allow to reconstruct them on the PDR level).

In Kurukh-Malto, this system was significantly changed, particularly in regard to terms for younger siblings, cf. Kurukh *tangri* ‘younger sister’ < PNorthD *\*tanḡ- + \*dī* ‘younger sister’, *tangris* ‘younger brother’ < PNorthD *\*tanḡ- + \*di* ‘younger brother’; the root here is of unknown origin, while the possessive prefix has been reconfigured by analogy with the more current shape of the pronominal oblique stem in Kurukh (and Malto).

The alternate hypothesis — namely, that the kinship terms of Proto-Dravidian were classificatory and referred to entire groups of people connected by similar relations to Ego — is much less likely, given that the only form that expresses such a fused meaning (in Malto) is probably not of Dravidian origin.

The other two stems with occasional “sibling semantics” that we can also reconstruct for Proto-Dravidian refer to “social background”: these are the PDR roots *\*cel-* ‘companion; sister’ which, it could be said, show some gender ambiguity with inclination towards female gender



(it is the source both of forms meaning ‘sister’, as in Parji *cālal*, and masculine forms such as Tamil *cilatan* ‘male companion, associate, servant’) and PDR *\*tōz-* ‘friend’, ‘younger brother’, ‘assistance, help’, usually denoting masculine gender (the source of forms meaning mainly ‘companion’, ‘brother’, as in Konda *toran* ‘younger brother’ or Kui *tōrenju* ‘male friend’), though cf. Pengo *torndel* ‘sister’. Their meanings are connected with the idea of “help”; see semantic shifts ‘brother’ – ‘friend’ and ‘friend’ → ‘husband’ / ‘wife’ in the *Database of semantic shifts of languages of the world* (Zalizniak et al. 2023). They seem to belong to a completely different subsystem of nomination, hinting at certain structural details of the ancient society that have yet to be explored.

The other terms are of borrowed origin. The sources of borrowings are different, depending on the historical processes and contact languages, mainly from different Indo-Aryan and, more rarely, from Indo-Iranian languages. We found that the way of adaptation of borrowings in the sibling subsystems of Dravidian languages proceed in three stages.

1. New term appears as a borrowing, usually first for ‘elder siblings’ and even more precisely, for ‘elder brother’. This may be connected with ‘elder brother’s prestigious position and preferred association with more prestigious language (though this consideration is rather speculative). Terms for younger siblings are less prone to borrowing; they retain the old shape and the old features. The most ancient features of the system still remain in the subsystem of ‘younger siblings’, such as the distinction of relative age in Brahui *īlumkō* ‘younger or youngest brother’ and the prefixes of inalienable possession in the majority of languages.
2. Two terms co-exist, as in Brahui, Telugu, Central Dravidian, Gondwana languages.
  - a. Sometimes they form two ‘parallel’ systems of sibling terms along with the authentic one, as in Brahui and Gondwana languages;
  - b. Sometimes they can appear within one and the same compound word, as in Kolumbi *akkabai* < *akka* (of Dravidian origin) + *bai* (of Indo-Aryan origin) and possibly Malto *dúbáy* ‘sisters and brothers from the same parents’ (this form is not so clear; *báy* clearly stems from *báyi* ‘elder sister’, *baya* ‘elder brother’ of Indo-Aryan origin, but the component *dú*, possibly connected with *ḍo* in *tangḍo* ‘his younger brother or sister’, is not very likely to be of Dravidian origin, either – perhaps an older borrowing from an unknown source).
3. The borrowed term gains a more prominent position than its predecessor, cf. the situation with the terms for ‘elder siblings’ in Central Dravidian languages, and in Malto (although the phonetic peculiarities in terms for younger siblings there remains an issue for further research). The original terms may remain at the periphery of the system, as a lexical item with a different meaning or restricted usage, or vanish completely<sup>7</sup>.

The resulting configuration of the sibling term system is quite diverse in Dravidian languages. In most South Dravidian languages all the main terms are still of authentic origin, so they represent the most ancient type of the system. In Telugu all the main terms remain authentic but some are replaced with terms of “social background”. In Central Dravidian lan-

<sup>7</sup> It is useful to list some semantically adjacent examples from close semantic kinship areas: thus, in Konda the original form *apo(si)* ‘father’ has only preserved the meaning of vocative, as opposed to the borrowed main term *buba* ‘father’ (Persian loan through Indo-Aryan, in the same subsystem with *bībi* ~ *bībsi* ‘elder sister’). A similar situation is observed in Brahui: the inherited Dravidian form *māmā* ‘uncle’ has only retained the meaning ‘form of address from son-in-law to father-in-law’, although the custom of real cross-cousin marriage that the form originally reflected did not survive and gave way to Muslim traditions.

guages terms of “social background” remain in the younger sibling subsystem along with borrowings, while the distinction of relative age is eroding. In Gondwana languages the distinction between ‘younger’ and ‘elder’ is erased for ‘brothers’; the authentic system remains as an additional parallel to the borrowed one (the origin of terms for “younger siblings” in Gondwana languages still stays a subject for further investigation, though it is definitely not from Munda languages, but possibly a remnant of pre-Dravidian and pre-Indo-Aryan substrate, see section 3.3).

### Sources of linguistic data

Unreferenced examples come from the following sources: Brahui — Bray 1909, 1934; Badaga — Hockings & Pilot-Raichoor 1992; Gadaba — Bhaskararao 1980; Balochi — Dames 1881; Gondi — Subrahmanyam 1968, Rafath et al. 2009; Gutob — Zide 1963; Indo-Aryan loanwords — Turner 1966–1985; Irula — Periyalwar 1979, Zvelebil 1979; Kannada — Učida & Rajapurohit 2013; Kolami — Emeneau 1955; Kui — Winfield 1929; Kurukh — Grignard 1924, Kobayashi & Tirkey 2017; Kuvi — Israel 1979, Burrow & Bhattacharya 1963; Malayalam — Gundert 1872, Andronov & Makarenko 1971; Malto — Droese 1884, Das 1973, Mahapatra 1979; Manda — Ramakrishna Reddy 2009; Marathi — Molesworth 1857; Parji — Burrow & Bhattacharya 1953; Pengo — Burrow & Bhattacharya 1970; Sanskrit — Monier-Williams 1899; Tamil — TL, Fabricius 1972, Winslow 1862; Ceylon Tamil dialect — Smiritskaya 2013a; Tamil dialects — Smiritskaya 2013b; Telugu — Gwynn 1872, Brown 1903; Tulu — Männer 1888, Brigel 1982.

### References

- Andronov, Mikhail. 1978. *Sravnitel'naja grammatika dravidijskix jazykov* [Comparative grammar of Dravidian languages]. Moscow: Nauka.
- Andronov, Mikhail, Vladimir Makarenko. 1971. *Malayalam-russkij slovar'* [Malayalam–Russian dictionary]. Moscow: Sovetskaja enciklopedija.
- Burlak, Svetlana, Sergei Starostin. 2005. *Sravnitel'no-istoricheskoje jazykoznanije* [Comparative historical linguistics]. Moscow: Academia.
- Bhaskararao, Peri. 1980. *Konekor Gadaba. A Dravidian language*. Pune: Deccan College.
- Bray, Denys. 1909. *The Brahui Language, part I*. Calcutta: Superintendent Government Printing.
- Bray, Denys. 1934. *The Brahui Language, part II — The Brahui Problem. Etymological Vocabulary*. Delhi: Manager of Publications.
- Brigel, Rev. J. 1982. *A grammar of the Tulu language*. New Delhi: Asian Educational Services.
- Brown, Charles Philip. 1903. *A Telugu-English dictionary*. Madras: Promoting Christian Knowledge.
- Burrow, Thomas, Sudhibhushan Bhattacharya. 1953. *The Parji language: A Dravidian language of Bastar*. Hertford: Austin.
- Burrow, Thomas, Sudhibhushan Bhattacharya. 1963. Notes on Kuvi with a short vocabulary. *Indo-Iranian Journal* 6: 231–289.
- Burrow, Thomas, Sudhibhushan Bhattacharya. 1970. *The Pengo Language*. Oxford: Clarendon Press.
- Das, A. Sisir Kumar. 1973. *Structure of Malto*. Annamalainagar: Annamalai University.
- Dames, Mansel Longworth. 1881. *A sketch of the northern Balochi language, containing a grammar, vocabulary and specimens of the language*. Calcutta: Asiatic Society.
- DEDR = Burrow, Thomas, Murray B. Emeneau. 1984. *A Dravidian Etymological Dictionary. 2nd edition*. Oxford: Clarendon Press.
- Droese, Rev. Ernest. 1884. *Introduction to the Malto Language*. Agra: Secundra Orphanage Press.
- Dumont, Louis. 1953. The Dravidian Kinship Terminology as an Expression of Marriage. *Man* 53: 34–39.
- Dziebel, German. 2001. *Fenomen rodstva: Prolegomeny k ideneticheskoj teorii* [Phenomenon of kinship: Prolegomena to the idenetic theory]. *Algebra rodstva: Rodstvo. Sistemy rodstva. Sistemy terminov rodstva* (6). Sankt-Peterburg: MAE RAN.
- Emeneau, Murray B. 1953. Dravidian Kinship Terms. *Language* 29(3): 339–353.
- Emeneau, Murray B. 1955. *Kolami. A Dravidian language*. Berkeley / Los Angeles: University of California Press.
- Fabricius, Johann Philipp. 1972. *Tamil and English dictionary*. Tranquebar: Evangelical Lutheran Mission Pub. House.

- Grignard, André. 1924. *An Oraon (Kurux) — English Dictionary*. Calcutta: Catholic Orphan Press.
- Gundert, Hermann. 1872. *A Malayalam and English dictionary*. Mangalore / London: C. Stolz / Trübner & Co.
- Gurov, Nikita V. 2013. Dravidian languages [Dravidijskie jazyki]. In Gurov, Nikita, Alexander Dubyanskiy, Andrej Kibrik, Elena Markus (eds.). *Languages of the world: Dravidian languages*. Moscow: Academia. 17-47.
- Gurov, Nikita, Alexander Dubyanskiy, Andrej Kibrik, Elena Markus (eds.). 2013. *Languages of the world: Dravidian languages*. Moscow: Academia.
- Gwynn, John Peter Lucius. 1991. *A Telugu-English dictionary*. Delhi / New York: Oxford University Press.
- Hockings, Paul, Christiane Pilot-Raichoor. 1992. *A Badaga-English Dictionary*. Berlin / New York: Mouton de Gruyter.
- Israel, Motchakon. 1979. *A grammar of the Kuvi language (with texts and vocabulary)*. Dravidian Association Publication No. 27. Madurai: Sangnam Printers.
- Kobayashi, Masatom, Bablu Tirkey. 2017. *The Kurux Language: Grammar, Texts and Lexicon*. Brill's Studies in South and Southwest Asian Languages, xvii. Leiden / Boston: Brill.
- Krishnamurti, Bhadriraju. 2003. *The Dravidian Languages*. Cambridge: Cambridge University Press.
- Krishnamurti, Bhadriraju. 1969. *Konda or Kubi. A Dravidian language. Texts, Grammar and Vocabulary*. Hyderabad: Tribal Cultural Research & Training Institute.
- Mahapatra, B. P. 1979. *Malto — An ethnosemantic study*. CIIL Occasional Monograph Series, 15. Mysore: Central Institute of Indian Languages.
- Männer, Rev. A. 1888. *English-Tulu dictionary*. Basel Mission Press.
- Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.
- Metsäranta, Niklas, Veronika Milanova, Terhi Honkola. 2023. Borrowability of kinship terms in Uralic languages. In: *Finnisch-Ugrische Forschungen* 68: 141–216.
- Milanova, Veronika. 2020. *Kinship and Affinity in Indo-European: Universal, Inherited, Contact, and Area Specific Features of Indo-European Kinship Terms (with a Special Focus on the Iranian Branch)*. Doctoral Thesis. Vienna.
- Milanova, Veronika, Sampsa Holopainen, Jeremy Bradley. 2020. Contact phenomena in Indo-European kinship and social terms and beyond (a pilot study with special focus on the Iranian and Uralic languages in the context of Central Eurasia). In: Romain Garnier (ed.). *Loanwords and Substrata: Proceedings of the Colloquium held in Limoges (5th-7th June, 2018)*: 345–366. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.
- Molesworth, James Thomas. 1857. *A dictionary, Marathi and English*. 2nd ed., rev. and enl. Bombay: Bombay Education Society's Press.
- Monier-Williams, Monier (revised by E. Leumann, C. A. Cappeller). 1899. *Sanskrit-English dictionary etymologically and philologically arranged with special reference to cognate Indo-European languages*. Oxford University Press.
- Morgan, Lewis Henry. 1997 [1871]. *Systems of Consanguinity and Affinity of the Human Family*. Lincoln and London: University of Nebraska Press.
- Olderogge, Dmitry. 1960. Osnovnye cherty razvitiya sistem rodstva [Main features of the development of kinship systems]. *Sovetskaya etnografiya* 6: 24–30.
- Parkin, Robert J. 1985. Munda Kinship Terminologies. *Man* 20(4): 705–721.
- Periyalwar, R. 1979. *Phonology of Irula with Vocabulary*. Annamalainagar: Annamalai University.
- Popov, Vladimir. 2015. Sistemy terminov rodstva [Systems of kinship terms]. In: *Bol'shaja rossijskaja enciklopedija* 30: 303. Moskva: Izdatel'stvo Bol'shaja rossijskaja enciklopedija.
- Rafath, Mohd. Ali et al. (eds.). 2009. *Gondi. Multilingual Dictionary with Semantic Category Section and Indexes for Telugu, English and Hindi*. Adilabad: Rajiv Vidya Mission (SSA).
- Ramakrishna Reddy, B. 2009. *Manda-English Dictionary*. Mysore: Central Institute of Indian Languages.
- Smirnitckaya, Anna. 2013a. *Cejlonskij dialekt tamil'skogo jazyka*. [Ceylon dialect of Tamil]. In: Nikita Gurov, Alexander Dubyanskiy, Andrej Kibrik, Elena Markus (eds.). *Jazyki mira: Dravidijskije jazyki*: 151–163. Moskva: Academia.
- Smirnitckaya, Anna. 2013b. *Dialekty tamil'skogo jazyka* [Dialects of Tamil]. In: Nikita Gurov, Alexander Dubyanskiy, Andrej Kibrik, Elena Markus (eds.). *Jazyki mira: Dravidijskije jazyki*: 163–173. Moskva: Academia.
- Smirnitckaya, Anna. 2016. Semantika terminov rodstva v tamil'skom jazyke s tochki zreniya tipologii semanticheskix perexodov. [Semantics of kinship terms in Tamil from the point of view of the semantic shift typology]. *Vestnik RUDN University. Series "Linguistics"* 20(2): 112–127.
- Smirnitckaya, Anna. 2020. Sistema rodstva jazyka brahui s tochki zreniya tipologii semanticheskix perekhodov. [Brahui language kinship system in terms of typology of semantic shifts]. *Etnografiya* 4(10): 101–117.
- Smirnitckaya, Anna. 2022. *Dravidijskie terminy rodstva s tochki zreniya tipologii semanticheskix perexodov* [Dravidian kinship terms from the point of view of typology of semantic shifts]. PhD thesis. Moscow: Institute of Linguistics.

- Starostin, Sergei (ed.). 1998–2005. *The Tower of Babel. An etymological database project*. Available online at: [starlingdb.org](http://starlingdb.org) [Accessed 07.09.2023].
- StarLing 2023 = Starostin, George. 2023. *Database of Dravidian Etymology*. Electronic resource, available at [starlingdb.org](http://starlingdb.org). [Accessed: 07.09.2023.]
- Starostin, Georgiy. 2013. *Yazyki Afriki. Opyt postroyeniya leksikostatisticheskoy klassifikacii. Vol. 1. Metodologiya. Kojjanskije jazyki. [Languages of Africa. An attempt at a lexicostatistical classification. Vol. 1: Methodology. Khoisan languages.]* Moskva: Jazyki slav'anskoj kul'tury.
- Starostin, Georgiy. 2000. *Rekonstrukcija fonologicheskoy sistemy pradravidijskogo jazyka. [Reconstruction of the phonological system of the Proto-Dravidian language]*. PhD thesis. Moskva: RGGU.
- Subrahmanyam, P. S. 1968. *A descriptive grammar of Gondi*. Annamalainagar: Annamalai University.
- Swadesh, Morris. 1971. *The Origin and Diversification of Language*. Ed. post mortem by Joel Sherzer. Chicago: Aldine.
- Tadmor, Uri, Martin Haspelmath, Bradley Taylor. 2010. Borrowability and the notion of basic vocabulary. *Diachronica* 27(2): 226–246.
- TL = Vaiyapuri Pillai, S. et al. (eds.). 1924–1936. *Tamil lexicon*. University of Madras.
- Trautmann, Thomas R. 1995. *Dravidian Kinship*. New Delhi: Vistaar Publ.
- Trawick, Margaret J. 1992. *Notes on Love in a Tamil Family*. Berkeley: University of California Press.
- Turner, Ralph L. 1966–1985. *A Comparative Dictionary of the Indo-Aryan Languages, plus Indexes, plus Addenda and Corrigenda* (ed. by J. C. Wright). London / New York / Toronto: Oxford University Press; London: School of Oriental and African Studies.
- Učida, Norihiko, Bando Bhimaji Rajapurhit. 2013. *Kannada-English Etymological Dictionary*. Tokyo: University of Foreign Studies.
- Wilden, Eva. 2018. *Grammar of Old Tamil for students. École française d'Extrême-Orient, Collection Indologie 137*. Institut français de Pondichéry. Available online at: <https://shs.hal.science/halshs-01892342v2>.
- Winfield, W. W. 1929. *A vocabulary of the Kui Language*. Calcutta: Asiatic Society of Bengal.
- Winslow, Miron. 1862. *A comprehensive Tamil and English dictionary of high and low Tamil*. Madras: P. R. Hunt.
- Yakobson, Roman. 1985. *Zvukovyje zakony detskogo jazyka i ix mesto v obshchej fonologii [Sound laws of children's language and their place in general phonology]*. In: Roman Yakobson. *Izbrannye raboty*: 105–116. Moskva: Progress.
- Zalizniak, Anna A. 2018. The catalogue of semantic shifts: 20 years later. *Russian journal of linguistics* 22(4): 770–787.
- Zhivlov, Mikhail, Tatiana Mikhailova, Maria Orlova, Anna Smirnitskaya. 2023. *Oboznacheniya muzha: klassifikaciya modelej nominacii (na baze lingvosemanticheskogo polya Evrazii) [“Husband” Designations: A Classification of Nomination Models (In the Linguo-Semantic Field of Eurasia)]*. *Izvestija Rossijskoj akademii nauk. Serija literaturny i jazyka* 82(6): 5–15.
- Zalizniak, Anna A., Marina Bobrik, Maria Bibaeva, Maria Bulakh, Mikhail Voronov, Dmitry Ganenkov, Ilya Gruntov, Timur Maisak, Tatyana Mikhailova, Maria Orlova, Maksim Russo, Anna Smirnitskaia. 2023. *Datsemshift: database of semantic shifts in languages of the world*. Moscow: Institute of Linguistics, Russian Academy of Sciences. Available online at: <https://datsemshift.ru>.
- Zide, Norman, B. P. Das. 1963. *Gutob Verb Lexicon*. Ms., unpublished.
- Zvelebil, Kamil. 1973–1982. *The Irula Language*. Wiesbaden, Harrassowitz.

А. А. Смирнитская. Обозначения братьев и сестер в дравидийских языках: прадравидийская реконструкция и проблема заимствований.

Статья посвящена исследованию развития системы номинации сиблингов в дравидийских языках, с упором на анализ степени устойчивости тех или иных терминов к заимствованию. Делаются выводы о заимствовании в дравидийских языках в первую очередь терминов для обозначения старших сиблингов, начиная со старшего брата; о преимущественном заимствовании целых подсистем и последующем параллельном их сосуществовании наряду с исконной системой; о большей устойчивости к заимствованиям терминов, обозначающих младших сиблингов и сохранности в их форме наиболее архаичных языковых черт, в частности форм, выражающих отношения неотчуждаемой принадлежности. Относительно последних предлагается считать для прадравидийской реконструкции и проблема заимствований.

видийского уровня наиболее вероятной систему \*y- 1SG / \*n- 2SG / \*t- 3SG, с последующей ее реструктуризацией в северно-дравидийской подгруппе курух-малто на основании аналогического влияния продуктивных косвенных форм личных местоимений (*eng*- 1SG / *ning*- 2SG / *tang*- 3SG).

*Ключевые слова:* термины родства; обозначения сиблингов; дравидийские языки; лексические заимствования; устойчивость к заимствованиям; прадравидийский язык; неотчуждаемая принадлежность.

## О правилах ударения в именных основах пушту

Данная статья посвящена попытке уточнения правил ударения в пушту в историческом аспекте. На основании этимологического анализа ряда морфологических категорий предлагается следующее объяснение траектории развития акцентуации в наиболее распространенных именных основах: исторические *\*ā*-основы и основы с суффиксом *\*-a-ka* сохраняют праиранское ударение, *\*a*-основы закономерно утрачивают окситонезу, а основы, которые дают в пушту зваракай или конечный долгий *-ā*, обобщают конечное ударение. Приведенные данные демонстрируют, что в пушту происходила масштабная разнонаправленная морфологизация изначального подвижного ударения.

*Ключевые слова:* пушту язык; ударение; праиранский язык; древнеиндийский язык; праиндоевропейский язык.

Язык пушту характеризуется разноместным свободным ударением и в этом отношении выделяется на фоне большинства современных иранских языков. Тем не менее, до середины XX в. эта его особенность не привлекала должного внимания исследователей. Так, Г. Хирт в посвященном ударению томе своей знаменитой «Индоевропейской грамматики» (Hirt 1929) касается развития ударения в отдельных группах индоевропейских языков, в частности, иранских. Описывая ударение в новоиранских языках, он сначала упоминает работы А. Мейе и его ученика Р. Готьё, посвященные ударению в древнеперсидском и согдийском, в которых исследователи впервые предположили, что ударение в этих языках в сущности напоминало ударение классического латинского, когда ударение падало на второй слог от конца в том случае, если он был долгим, и на третий слог от конца, если он был кратким, а в дальнейшем конечные заударные слоги отпали (Meillet 1897; Meillet 1900; Gauthiot 1916). Далее Г. Хирт с осторожностью характеризует выводы, содержащиеся в упомянутых работах, попутно касаясь описания ударения в пушту в таком авторитетном источнике, как «Основы иранской филологии»: «Ich kann die Richtigkeit dieser Arbeiten nicht beurteilen. Da ich kein Iranist bin, kann ich mich auf die Benutzung neuerschienenener Arbeiten nicht anlassen. Es ist aber bemerkenswert, daß sich in dem Grundriß der iranischen Philologie kein Wort über die Betonung des Afghanischen, des Balutschis oder des Neupersischen findet. [Von letzterem weiß man, daß der Hauptton mit einigen Ausnahmen auf der letzten Silbe ruht, offenbar, weil dahinter alles fortgefallen ist]» — «Я не могу подтвердить правильность (выводов) этих работ. Поскольку я не иранист, я не могу полагаться на использование недавно вышедших работ. Но примечательно, что в Grundriß'e иранской филологии нет ни слова об ударении в афганском, белуджском или новоперсидском языках. [О последнем известно, что место основного ударения, за немногими исключениями, приходится на последний слог, очевидно, потому, что все заударные слоги выпали]» (Hirt 1929: 196). Утверждение об игнорировании описания ударения в пушту можно было бы распространить не только на Grundriß иранской филологии, но и на многие другие ранние работы, в том числе описательные и словарные: например, такой признанный специалист по языку пушту, как Г. Дж. Раверти, никак не обозначает ударение ни в своей грамматике (Raverty 1856), ни в словаре (Raverty 1867). Практически полное отсутствие сведений об ударении в пушту в посвященных этому

языку исследования прослеживается на протяжении всего XIX и первой четверти XX в.; подробный обзор истории вопроса содержится в работе (Ве́йка 1969: 8–14).

Последовательное обозначение и синхронное описание ударения в пушту впервые появляется в работах Г. Моргенштерне: в статье «Notes on the Pronunciation of Pashto» (Morgenstierne; Lloyd-James 1928) и в кратком этимологическом словаре словаря пушту (EVP 1927). На основании данных норвежского исследователя Г. Хирт предварительно отмечает основы, которые могут сохранять следы старого ударения: «So heißt es auch hier *at*<sup>1</sup> ‘acht’, ai. *aštāu*; *ōrə* Wolke = ai. *abhrám*. Aber anderseits vergleicht er *bar* ‘on, above’ mit ai. *upári*, wobei also der alte ton erhalten geblieben wäre». — «Так, здесь также приводятся *at* ‘восемь’, др.-инд. *aštāu*; *ōrə* ‘облако’ = др.-инд. *abhrám*. Но, с другой стороны, он (Г. Моргенштерне) сравнивает *bar* ‘верх; на, наверху’, где, соответственно, должно бы сохраняться старое место ударения» (Hirt 1929: 197).

Начало масштабного изучения связи ударения в пушту с праиранским и индоевропейским ударением положено Г. Моргенштерне в статье 1942 г. «Archaisms and innovations in Pashto morphology» (Morgenstierne 1942), посвященной в первую очередь именной морфологии в историческом освещении. Во второй части статьи (с. 95–98; § 9–13) последовательно обсуждаются формы, указывающие на сохранение в пушту следов праиранского разноместного ударения. Норвежский исследователь указывает на ряд основ с сохранением архаического ударения, которые относятся к различным морфологическим классам (*áspa* ‘лошадь’ ~ вед. *áśvā-*; *stóray* ‘звезда’ ~ вед. *táraka-*) и справедливо отмечает, что прилагательные и имена на *-ay* (< \*-aka) могут иметь различное ударение в зависимости от баритонезы или окситонезы в праязыке.

Наконец, по-настоящему убедительные и проработанные результаты в отношении истории акцентуации в пушту были получены в 70-е гг. XX в. В. А. Дыбо, который в публикации «Афганское ударение и его значение для индоевропейской и балто-славянской этимологии» достаточно полно описал именные основы пушту с точки зрения ударения (Дыбо 1974). Исследователь на основании обширного материала пришел к выводу, что в целом именные основы могут быть поделены на баритонированные и окситонированные и место ударения в них совпадает с местом ведийского и праиндоевропейского акцента в соответствующих когнатах. Это касается и производных основ с историческим суффиксом \*-a-ka, которые характеризуются ударением «по контрасту». По формулировке автора, «непроизводные с точки зрения современного языка афганские имена на *-ay* делятся на два акцентных типа: тип А с ударением на корне и тип В с ударением на основообразующем форманте. Сравнение показывает, что в тип А входят главным образом имена, образованные от индоиранских окситонированных основ, в тип В — имена, образованные от индоиранских баритон» (Дыбо 1974: 90). Как показывает исследователь, аналогичным образом были устроены и производные в ведийском языке (Дыбо 1974: 102–105). Кроме того, он считает, что именно ударение в языке пушту имеет большое значение для индоиранской реконструкции в том числе по той причине, что в ряде случаев место древнеиндийского словесного ударения различно в сравнении с праиндоиранским и праиндоевропейским (это касается в основном прилагательных и причастий), тогда как афганская акцентуация сохраняет следы более архаичной индоевропейской системы (Дыбо 1974: 75–77). Во второй части обсуждаемой работы, опубликованной на 15 лет позже, В. А. Дыбо показал, что распределение ударения в системе презенса в пушту также указывает на изначальное различие баритонированных и окситонированных основ.

<sup>1</sup> Здесь и далее формы языка пушту приводятся в авторской транслитерации.

Он описывает и реконструирует сложную и достаточно стройную систему акцентуации глагольных форм с энклитиками, в частности с частицей *na* ‘не’ (Дыбо 1989). Кроме того, в 1972 году В. А. Дыбо кратко указал на реликты разноместного ударения в языке йидга (Дыбо 1972: 43–44), а также в дардских языках дамели и шина (Дыбо 1972: 38–42). Более подробный разбор ударения в языке шина, а также дамели в сравнении с индоиранским и индоевропейским состоянием был позднее опубликован в статье Дыбо 2011.

Результаты работ В. А. Дыбо были в целом приняты советскими иранистами и послужили основой для акцентологических описаний пушту в работах Грюнберг 1987, Грюнберг & Эдельман 1987, а также широко цитируются при разборе соответствующей лексики в «Этимологическом словаре иранских языков» (ЭСИЯ). Также интересующая проблема в свете новых данных с привлечением дополнительного материала была вновь рассмотрена Г. Моргенштерне (Morgenstierne 1973; Morgenstierne 1983). Некоторые идеи, высказанные в последней статье, имеют особенное значение для настоящей работы и будут упомянуты при дальнейшем изложении.

Предпринимались попытки описать акцентуацию пушту и ее историю и с других позиций. Л. Г. Герценберг предложил рассматривать афганское ударение в рамках концепции, напоминающей те, что применялись для описания исторического развития акцентуации в древнеперсидском и особенно в согдийском: исследователь постулировал различие тяжелых и легких основ в восточноиранском и полагал, что в истории пушту в двусложных основах происходило перемещение ударения на конечный тяжелый слог в том случае, если первый был кратким, тогда как ударение оставалось на первом тяжелом слоге (Герценберг 1981; Дыбо 1990: 10–11). Но В. А. Дыбо отметил в критической статье, что «ритмическая концепция Герценберга» не выдерживает проверки фактами и оказывается избыточной «в связи с полной тождественностью системы афганского разноместного ударения ведической» (Дыбо 1990: 12). Также стоит еще раз отметить относительно позднее происхождение «ритмического закона» для согдийского и невозможность его проекции на праиранский или правосточноиранский уровень (Novák 2013: 75–81), что является дополнительным аргументом против подобной гипотезы.

Дж. Ченг в своей статье, специально посвященной ударению в пушту, частично отвергает объяснение, предложенное В. А. Дыбо (Cheung 2010). Исследователь также критически рассматривает концепцию Л. Г. Герценберга, отмечая, что деление на тяжелые и легкие слоги по образцу других иранских языков, предложенное последним, выглядит произвольным (Cheung 2010: 111), и после рассмотрения материала предлагает свою достаточно сложную систему описания двух-, трех- и четырехсложных форм в пушту.

Согласно Дж. Ченгу, для праиранских двусложных форм надо различать 4, а для трехсложных — 3 случая последовательности кратких и долгих слогов, что в конечном счете оказывает влияние на ударение словоформ.

1. Последовательность  $\check{\text{~}}$ : *orá* ‘облако’ < \**abrá-*, *nar* ‘муж, мужчина, самец’ < \**nára-*, *psá* ‘овца’ < \**paśú-*;

2. Последовательность  $\check{\text{~}}$ : *as* ‘конь’ < \**áswa-*, *gwaǵ* ‘ухо’ < \**gáuša-*; *nən* ‘сейчас’ < \**nūnám*, *skam* < \**skambá-*, прилагательные *mər* ‘мертвый’ (< \**múrdu* < \**mṛtú*) < \**mṛtám* (sic!), *pən* ‘полный’ (< \**púrnu* < \**pṛnú*) < \**pṛnám* (sic!), *kuṛ* ‘глухой’ < \**karná-*; *stər* ‘большой’ < \**sturá-*; *um*, *om* ‘сырой’ < \**āmá-*; *yaw*, *yow* ‘один’ < \**aiwá-*; *žər* ‘быстрый’ < \**jírá-*.

3. Последовательность  $\check{\text{~}}$ : *mlā* ‘поясница’ (< \**melá*) < \**máidyā-*; *niná* ‘поджаренное зерно’ (< \**lená*) < \**dānyá-*; *p(ə)lá* f., pl. *p(ə)lé* (< \**pělá*) ‘жила, нерв’ < \**páidyā-*.

4. Последовательность  $\check{\text{~}}$ : e.g. *áspa* ‘лошадь, кобыла’ < \**áswā-*, *lúma* < ‘силлок, петля’ < \**dāmā-* и т. д.





Статистические данные, полученные И. Бечкой на основании обработки одной пятой наиболее полного словаря М. Г. Асланова, однозначно указывают на то обстоятельство, что в пушту значительная доля слов может иметь только конечное ударение или имеет выраженную тенденцию к конечнударности, а именно 68,3 процента всех слов (Ве́йка 1969: 57–58). С другой стороны, именно те два морфологические класса слов, которые изначально были описаны Г. Моргенштерне и — более подробно — В. А. Дыбо как надежно сохраняющие следы старой акцентуации, демонстрируют отклонение от общего распределения. «В словаре Асланова 307 слов, заканчивающихся на ‘-a’ и начинающихся на буквы алиф, *t*, *c* и *d*; 139 из них имеют ударный, 168 — безударный ‘-a’. В их число не входят слова, заканчивающиеся суффиксом *-āna*, число которых очень велико». Далее, «слов, заканчивающихся на /ay-/ под буквой *b*, 189, из которых 99 имеют ударение на последнем слоге (/ay-/ является ударным) и 90 имеют ударение на слоге, отличном от последнего (/ay-/ является безударным)» (Ве́йка 1969: 95). Здесь же имеет смысл привести наблюдение Г. Моргенштерне относительно \*ā-основ: «Besonders deutlich tritt diese Unterscheidung zwischen Oxytonen und Barytonen in den Feminina von zwei Klassen von Adjektiven hervor:

Vgl. *wúč-a* f. ‘trocken’; *spín-a* f. ‘weiß’; *zyár-a* f. ‘gelb’, usw. : Ved. *śúṣkā*; *śvítā*; *hárítā*. Dagegen: *sra* (m. *sur*) ‘rot’; *paxá* (m. *pox*) ‘reif, gekocht’; *tərwá* (m. *trīw*) ‘sauer’; *saṛá* (m. *soṛ*) ‘kalt’ : Ved. *śukrá*; *pakvā*; *tīvrā*; Lit. *šaltá* (m. *šáltas*). — «Особенно отчетливо это различие между окситонами и баритонами выступает в женском роде двух классов прилагательных»: ср. *wúč-a* f. ‘сухой’; *spín-a* f. ‘белый’; *zyár-a* f. ‘желтый’, и т. д. : вед. *śúṣkā*; *śvítā*; *hárítā*. Напротив: *sra* (m. *sur*) ‘красный’; *paxá* (m. *pox*) ‘спелый, готовый’; *tərwá* (m. *trīw*) ‘кислый’; *saṛá* (m. *soṛ*) ‘холодный’: Ved. *śukrá*; *pakvā*; *tīvrā*; Lit. *šaltá* (m. *šáltas*)» (Morgenstierne 1983: 169).

Анализ, проведенный И. Бечкой, свидетельствует также не в пользу правил В. А. Дыбо и Дж. Ченга в том виде, в котором они были первоначально сформулированы. Очевидно, что чисто фонетические преобразования вряд ли могли обеспечить такой перевес конечнударным словоформам.

Соответственно, для уточнения исторических аспектов развития системы ударения в пушту, в целом наиболее подробно и адекватно описанной с синхронической точки зрения И. Бечкой и с диахронической — В. А. Дыбо — требуется ряд поправок, которые помогут разграничить наиболее архаичные случаи и определенные инновации, действие которых и предопределило безусловное преобладание многосложных конечнударных форм по сравнению с имеющими ударение на слогах, отличных от последнего. Достаточно перспективным представляется поиск причин данного распределения в морфологизации акцента в определенных классах основ, а также в появлении значительного количества ударных суффиксов, которые перечислены в работе (Ве́йка 1969: 84; 91–93).

### Праиранские \*a-основы

Последовательное рассмотрение форм прилагательных в мужском роде и других слов, продолжающих основы на \**-á* в пушту, приводит к выводу, что их закономерным рефлексом в пушту являются такие слова, как *məṛ* ‘мертвый’ < \**mṛtá-*, *sur* ‘красный’ < \**suxrā-* и т. д., с потерей конечного согласного, неотличимые в мужском роде по форме от рефлексов баритонированных основ праиранского. Именно этим обстоятельством объясняются некоторые мнимые отклонения от исторической окситонезы, которые

вступают в противоречие с правилами В. А. Дыбо и Дж. Ченга. Также подобные изменения затрагивают некоторые окситонированные основы на *\*i/ī*. В предлагаемом списке, начиная с 19 примера, будут приведены основные прилагательные, имеющие окситонезу в женском роде и при этом односложную форму в мужском без дополнительных подробных пояснений относительно их этимологии — более полные сведения об их происхождении могут быть обнаружены в (NEVP) и (Дыбо 1974: 81–82).

1. *bəl*, ванечи *bī*, *biya* f. ‘второй’ < *\*dwityá-* или *\*dwitá-*; ср. др.-инд. *dwitá-* ‘имя божества’, буквально ‘второй’ (NEVP: 14). Данный пример с точки зрения акцентуации осложняется тем, что форма женского рода в пушту имеет вид *bála*, а не # *balá*. В свете данных ванечи может быть предпочтительнее праформа *\*dwityá-*, именно она приводится в (ЭСИЯ-2: 486), хотя для пушту не исключается и возможность возведения к *\*dwitá-*.

2. *cer* m. ‘подобие’ < *\*čitrá-*, ср. др.-инд. *citrá-* ‘очевидный, заметный; ясный’ (NEVP: 18). Окситонеза доказывается наличием параллельного существительного женского рода *cerá* ‘внешность, наружность, облик, лицо; вид, форма, образ; рисунок, картина и т. д.’ (ПРС: 350; Дыбо 1974: 80). Г. Моргенштерне уверенно относит *cerá* f. ‘рисунок, картина’ к заимствованиям из персидского, где рефлекс классической формы *čihra* с базовым значением ‘лицо, облик’ может произноситься как *čēra* (EVP: 17), однако, на наш взгляд, можно в лучшем случае думать о контаминации унаследованной и заимствованной основ в разных значениях. Поэтому более обоснованным представляется подход составителей «Этимологического словаря иранских языков», которые колеблются между версией об образовании пшт. *cerá* из *\*čitra-* + *-ā* и предположением о раннем заимствовании из персидского (ЭСИЯ-2: 258). Что касается семантики, в древнеиндийском от рассматриваемого выше прилагательного *citrá-* засвидетельствованы производные существительные со значениями, близкими тем, что наблюдаются в персидском и пушту: др.-инд. *citrá-* ‘Glanz; Schmuck(stück); Gemälde, Bild; Wunder’ (Mylius 1975: 157).

3. *car* m. зап. ‘пастьба, выпас’, *cará* f. ‘пастбище, выгон’ < *\*čartá-* и *\*čartā-*, ср. др.-инд. *caritá-* ‘ходьба, хождение’. В данных словах содержится незакономерный рефлекс *a* — перед группой согласных праиранский *\*a* регулярно дает пшт. *ā* (Грюнберг-Эдельман 1987: 21; Cheung 2011: 171–172), но такое развитие может быть обусловлено влиянием заимствованной из персидского глагольной основы *car-* ‘пастись’ (NEVP: 18).

4. *čār* m., *čārā* f. ‘работа’ < *\*kāri-*, др.-инд. *kāri-* ‘действие’ (NEVP: 20). Возможно, в истории пушту произошел переход в *\*a*-основу. Здесь стоит отметить, что ударение *čārā́*, указанное Г. Моргенштерне, противоречит данным других источников: данное слово имеет облик *čā́ra* f. ‘работа; дело’ в русскоязычных словарях, а конечное ударение отмечено для *čā́rá* ‘средство, способ, выход; лекарство’ (КАРС: 185; ПРС: 308). Соответственно, данный случай сомнителен.

5. *γal* m. ‘вор’ при *γla* f. ‘воровка’ < *\*gadá-* и *\*gadā-*, ударение устанавливается по основе женского рода (NEVP: 30). В древнеиндийском засвидетельствован корень *gadh-* ‘схватывать, захватывать’ в позитивном и негативном планах, ясное производное которого с ударением на корне — *gádhyā-* ‘добыча’; дальнейшая этимология представляется менее надежно установленной. Если традиционно сравнивать с лит. *gōdas* ‘алчность, жадность’ и др.-ирл. *gataim* ‘уношу, вору’ < и.-е. *\*g<sup>h</sup>ed-*, *\*g<sup>h</sup>end<sup>h</sup>-* (ЭСИЯ-3: 100–101; IEW: 437–438), можно отметить, что В. А. Дыбо реконструирует окситонированное имя *\*ghādhós* на основании лит. *godūs* ‘жадный’ и относит глагол *\*ghādhāmi* > *\*ghādhāmi* > др.-ирл. *gataim* ‘краду, вору’ к разряду образованных от окситонированных имен (Дыбо 1961: 22). Со-

ответственно, праязыковая окситонеза может предполагаться и на основании внешних данных.

6. *γul* m. pl. ‘кал, экскременты’ < \**gūthá-*, ср. др.-инд. *gūthá-* ‘экскременты; грязь; нечистоты’ (NEVP: 30). В словаре Асланова представлен вариант данной лексемы *γвул, γвал* m. pl. ‘кал, испражнения’ (ПРС: 618); это же слово в записи *γул (γвал)* m. pl. приводится в качестве перевода русского слова *кал*, а *γул* m. pl. без фонетических вариантов дается как эквивалент рус. *экскременты* в словаре (РАСП: 257; 804). Форма, засвидетельствованная в словарях Г. Моргенштерне и К. А. Лебедева, принадлежит диалекту вазири согласно (ЭСИЯ-3: 294).

7. *γar* m. ‘гора’ < \**gará-* при более ранней праформе \**garí-* ‘гора’ (NEVP: 32). На переход из \**i*-основы в тематическую могут указывать также согдийские и ягнобские данные: ср. согд. *γr-* [*γar-*] ‘гора’ и ягн. *γar* ‘гора; перевал’ «при ранней утрате \**i* и отсутствии умлаута» (ЭСИЯ-3: 191).

8. *māt* ‘сломанный’ < \**maštá-*, от праиранского глагола \**maž-*, который не сохранился в пушту (NEVP: 53). Ввиду отсутствия надежных древнеиндийских параллелей окситонеза предположительная, но вероятная в причастии на \**-ta* (неясно, может ли пушту показывать следы праязыковой баритонезы в \**-to*-причастиях подобно европейским языкам).

9. *muṭ*, ванечи *mīt* m. ‘кулак; горсть’ < \**mušti-*, ср. авест. *mušti-*, др.-инд. *muṣṭi-* m. ‘кулак’ (NEVP: 53). Гласный в ванечи должен указывать на умлаут, соответственно, переход в \**a*-основу невероятен.

10. *nal* m. ‘тростник, трубка’ < \**nadá-*, ср. др.-инд. *naḍá-* m. (n.) ‘тростник’, *nadá-* m. ‘тростник’ (NEVP: 56). Древнеиндийское слово считается унаследованной основой с индоевропейской этимологией, ср. хетт. *nata-*, *nati-* ‘тростник, трубка для питья; стрела’, лув. *natatta* ‘изделие из тростника’, арм. *net* ‘стрела’ (EWAia II: 7–8; ЭСИЯ-5: 415–417, Yakubovich, Mouton 2023, I: 67). На окситонезу указывают также формы собственно пшт. *nalá-* f. ‘трубка’ и вазири *nalla* f. ‘large hollow reed; urethra’ (NEVP: 56), где в \**ā*-основе женского рода переход из окситонезы в баритонезу ожидаемо не происходит.

11. *nən* ‘сегодня’ < \**nīnám*, ср. др.-инд. *nīnám* ‘сейчас’ (NEVP: 57).

12. *pal* m. ‘ступня, след ноги’ < \**padá-*, ср. др.-инд. *padá-* n. ‘шаг, след ноги, отпечаток ноги; место’ (NEVP: 62). Как отклонение в отношении акцентуации рассматривается Дж. Ченгом — данное слово очевидно противоречит сформулированным им правилам (Cheung 2010: 113). На самом деле это закономерный рефлекс, как показывают остальные случаи.

13. *skām* m. ‘стойка, столб’ < \**skambá-*, ср. др.-инд. *skambhá-* m. ‘шест, столб; колонна’ (NEVP: 73).

14. *səl* ‘сто’ < \**śatá-*, ср. др.-инд. *śatá-* ‘сто’ (NEVP: 74). Как и в случае с лексемой *bəl* ‘второй’, форма не соответствует ожиданиям Дж. Ченга, поскольку в последовательности двух кратких слогов не должен происходить перенос ударения. Как и для *bəl*, он предлагает видеть промежуточные праформы \**sla*, \**bla* с метатезой начальных кластеров \**sl-* и \**bl-* (Cheung 2010: 115; Cheung 2011: 182), но утверждение о том, что для пушту начальный кластер *sl-* и подобные являются неприемлемыми, не соответствуют действительности: ср. *slet*, *sleṭ* m. ‘грифельная доска’ и *slex*, *slixt* m. ‘клей; вязкость, клейкость; прилипший’, *zrə* ‘сердце, душа’ и т. д.

15. *triw* m., *tarwá*, *trawá* f. ‘кислый (на вкус); солоноватый, соленый (о воде); пахта; вид соленого сыра’ < \**tīwrá-*, ср. др.-инд. *tīvrá-* ‘сильный, резкий, острый’ (NEVP: 83).
16. *wand* m. ‘плотина, ров’ < \**bandá-*, ср. др.-инд. *bandhá-* m. ‘цепь, связка’ (NEVP: 88; Cheung 2011: 174).
17. *wuz* ~ *wəz* m. ‘козел’; *wəzá*, *wza*, также *uzá*, *bza* и *bezá* f ‘коза’ < \**būzá-* и \**būzá-*; ударение устанавливается по форме женского рода (NEVP: 94).
18. *šəl*, вазири *šəl* m. ‘лестница из камней и земли’ < \**šritá-*, ср. др.-инд. *śritá-* ‘прислоненный’ (NEVP: 98). Сомнительный случай, потому что в словаре М. Г. Асланова засвидетельствовано *ḡal* m. ‘лестница’ (ПРС: 566), а также *шал* m. ‘лестница’ (ПРС: 551) с другим вокализмом, не выводимым из праир. \**i*.
19. *ím*, f. *umá* ‘сырой, недоваренный; неспелый, незрелый’ < \**āmá-* (NEVP: 8).
20. *sur*, f. *srá* ‘красный, алый; румяный’ < \**suxrá-*, авест. *suxra-* ‘красный’ (NEVP: 75).
21. *póx*, f. *raxá* ‘сваренный, вареный, спелый, зрелый’ < \**raxwá-* (NEVP: 67).
22. *sór*, f. *sará* ‘холодный, прохладный’ < \**śartá-*, авест. *sarta-* ‘холодный’ (NEVP: 76).
23. *tód*, f. *tawdá* ‘теплый, горячий, жаркий’ < \**taptá-*, авест. *tafta-* (NEVP: 81).
24. *már*, f. *mṛá* ‘мертвый, умерший’ < \**mṛtá-* (NEVP: 51).
25. *tríx*, f. *tərxá* ‘горький, соленый’ < \**taxra-* (NEVP: 83–84).
26. *kər*, f. *krá* ‘сделанный’ < \**kṛtá-* (NEVP: 40).
27. *pər*, f. *prá* ‘виновный, виноватый; побежденный’ < \**pṛtá-* (NEVP: 65).
28. *post*, f. *pastá* ‘мягкий; нежный’ < \**pastá-* < \**padtá-* (NEVP: 66).
29. *wər*, f. *wərá* ‘(при)несенный’ < \**bṛtá-* (NEVP: 92).
30. *wrost*, f. *wrastá* 1 ‘гнилой, испорченный; разложившийся, тухлый’ < \**fristá-* ? с несным вокализмом (NEVP: 91).
31. *wrost*, f. *wrastá* 2 ‘рваный, изношенный; истрепанный, пришедший в негодность (о денежном знаке)’ < \**wirastá-* < \**wi-rad-tá-* или \**wa-rastá-* < \**awa-rad-tá-*, ср. (ЭСИЯ-6: 328–331).
32. *prot*, f. *pratá* ‘лежащий; живущий, находящийся где-л.’ < \**parā-wastá-* (NEVP: 65).
33. *writ*, f. *writá* ‘жареный’ < \**brixtá-* (NEVP: 91).
34. *moṛ*, f. *maṛá* ‘сытый; удовлетворенный; состоятельный, богатый’ < \**hwartá-* (NEVP: 52, 59).
35. *zor*, f. *zará* ‘старый (о людях)’ < \**zār(a)tá-* (NEVP: 103).
36. *woṛ*, f. *wará* ‘маленький, мелкий; небольшой, невысокий; незначительный’ < \**wṛtá-* ? (NEVP: 92).

### Основы на зваракай

Абсолютно все слова, имеющие окончание *-ə*, имеют ударение на данной фонеме, для которой в афганистике используется традиционный термин зваракай (пшт. *zwarakáy* ~

*zwarakáy*). В этом смысле с синхронной точки зрения мнимы исключения, которые приводит В. А. Дыбо в разделе своей статьи, называемом «Примеры окситонезы вместо ожидаемой баритонезы»: *mēṛá* m. ‘мужчина, муж’ < \**márt(i)ya-* и *pinjá* ‘пять’ < \**pánca-* с определенной трансформацией основы (Дыбо 1974: 87–89), на странность рефлекса *-á* в слове ‘пять’ указывается в (Morgenstierne 1973a: 109). Также необходимо отметить, что Г. Моргенстьерне осознает данную проблему и высказывает определенные осторожные суждения относительно происхождения конечного ударного зварака: «Warum finden wir hier *-ə*, nicht *-a*, und warum ist die Endsilbe betont? Zur ersten Frage kann ich nur sagen, daß es mir in mehr als 50 Jahren nicht gelungen ist, die Verteilung von finalem *-a* und *-ə* in allen Einzelheiten zu erklären. Trotz aller Bemühungen geht die Patience für mich nicht völlig auf. — Gehen wir aber von *-ə* aus, so ist dieser Laut im Pscht. immer betont. Die einzige Möglichkeit, *sāṛá* und überhaupt das dreistufige Ablautsystem zu verstehen, scheint mir daher die Annahme zu sein, daß *-ə* sekundär den Akzent an sich gezogen hat, ohne daß ich diesen Vorgang erklären kann». — «Почему мы находим здесь *-ə*, а не *-a*, и почему конечный слог является ударным? На первый вопрос я могу только ответить, что мне не удалось за более чем 50 лет объяснить распределение конечных *-a* и *-ə* во всех деталях. Невзирая на все усилия, мое терпение в полной мере не помогает. Единственную возможность понять *sāṛá* и в целом трехчленную систему аблаута предоставляет, как мне кажется, предположение о том, что *-ə* вторично перетягивает на себя ударение, несмотря на то, что я не могу объяснить этот переход» (Morgenstierne 1983: 172–173). В рамках данной работы я также не могу провести полную реконструкцию исторического развития конечных гласных пушту и их точного происхождения, а также не могу объяснить наличие в основах прилагательных одновременно *ō*, *ā* и *a* в разных падежных формах, восходящих к \**a*. Данная проблема требует дальнейшего рассмотрения с опорой на работу (Cheung 2011), в которой сделаны многие ценные наблюдения и предложены конкретные удачные решения. С другой стороны, как кажется, можно сделать ряд предварительных замечаний о связи конечного зварака с определенными типами основ.

Любопытно отметить, что практически все имена на зваракай, отмеченные в (Дыбо 1974: 82–83), содержат суффиксы \**-(i)ya* или \**-(u)wa*, имеют дублеты в женском роде или являются pl. tantum, соответственно, во многих случаях их зваракай *ə* может продолжать \**-(i)ya*, \**-(u)wa* или \**-āh*. Впрочем, как уже было сказано, однозначная диахроническая интерпретация на данном этапе исследования морфологии и сравнительной фонетики пушту не представляется возможной.

#### Имена с суффиксом \**-(i)ya*

1. *lewá* m. ‘волк’ < \**daiw(i)ya-*, ср. др.-инд. *devyám* n. ‘божественное достоинство’ (NEVP: 45–46). В. А. Дыбо предполагает непосредственное возведение рассматриваемого слова к праир. \**daiwa-* ‘дэв’ и привлекает для сравнения др.-инд. *devá-* ‘бог, брахман, царь’ и прилагательное *devá-* ‘божественный’, указывая, что «поправка Г. Моргенстьерне (< *daēvya* (sic!) ‘daevic’), по-видимому, излишня» (Дыбо 1974: 82). Однако в свете изложенных выше соображений и данных, согласно которым \**a*-основы с исходной окситонезой теряют ее в пушту и выглядят так же, как исконно баритонированные имена, правильной должна быть именно реконструкция \**daiw(i)ya-*.

2. *ōwrá* m. sg. (и pl.) ‘облако’ < \**abr(i)ya-* (хотя отсутствуют следы палатализации гласного, NEVP: 10). Не стоит ли в таком случае предполагать, что форма *ōwrá* изначально множественного числа и *-á* < \**-āh*? Трактовка Дж. Ченга напрямую из \**abrám* сомнительна (Cheung 2011: 180).

3. *trə* ~ *tə́rə* m. ‘дядя по отцу’ < \**pitṛv(i)yá-*, ср. др.-инд. *pitṛvyá-* ‘дядя по отцу’.

4. *wrā́rə* m. ‘племянник (сын брата)’ < \**brāṭṛv(i)yá-*, ср. др.-инд. *brāṭṛvyá-* ‘племянник’.

5. *ū́má* m. ‘эфедра, хвойник’ < \**haum(i)yá-*, ср. др.-инд. *somyá-* ‘относящийся к коме, связанный с сомой’.

6. *čā́rə* f. ‘нож, короткий меч, бритва’ < \**kartí-* или \**kart(i)yá-*, ср. др.-инд. *kṛtí-* ‘род кинжала, нож мясника’;

7. *rā́šə* f. ‘груда зерна, хлеб (на току); продукты, продовольствие’ < \**rā́sí-* или \**rā́sí-(yá)-*, ср. др.-инд. \**rā́sí-* ‘куча, груда, множество, масса’;

8. *zṛə* m. ‘сердце, душа’ < \**ṛdyá-* < \**ṛdíya-*, ср. греч. *καρδία* ‘сердце’.

#### Имена с суффиксом \*-(u)wá

9. *psə* m. ‘баран; диал. мелкий рогатый скот’ < \**paśú-* или \**paś(u)wá-*, ср. др.-инд. *paśú-* ‘мелкий скот, жертвенное животное, скотина’.

10. *wā́zə* f. ‘сажень, мера длины, равная 2, 134 м.; шаг (коня)’ < \**bā́zú-* или \**baž(u)wá-*, ср. др.-инд. *bahú-* ‘рука, предплечье; мера длины’.

#### Имена, имеющие дублеты женского рода или pl. tantum

11. *tuṛyá*, *tə́ryə* m. ‘птица (часто хищная)’ < \**mṛgá-*, ср. др.-инд. *mṛgá-* ‘дикое животное, птица; дичь; газель, антилопа, олень и др.’. В диалекте вазири есть форма *maṛyá-* f. ‘небольшая птица’ < \**mṛgá-* или \**maṛgá-*. Более того, в пушту распространены производные с незакономерной окситонезой *maṛyá* ‘пташка’ и *maṛyá*, *maṛyá* ‘птичка, пташка’ < \**mṛgáka-* < *mṛgáká-* и \**mṛgákā-* < \**mṛgáká-* (NEVP: 51; ЭСИЯ-5: 381). Орм. лог. *morgá*, кан. *mirgá* ‘воробей, пташка’ могут быть рефлексами \**mṛgá-*, особенно если учесть, что в канигурамском диалекте это слово женского рода (Ефимов 1986: 127).

12. *bā́ṇə* m. sg. и pl. ‘ресница, ресницы’ < \**parná-*, ср. др.-инд. *parṇá-* ‘крыло, перо; лист; оперение стрелы’, наряду с *bəṇá* ‘перо’ и *bā́ṇə* ‘ресница’ (NEVP: 15); по всей видимости, данная лексема может восходить к pl. tantum \**parnáḥ*. Необходимо отметить наличие основы *bā́ṇú-* ‘ресница’ (ПРС: 107).

13. *γā́rmə* m. ‘жара, зной; жар солнца’, в отдельных диалектах ‘солнце’ < \**garmá-*, ср. др.-инд. *gharmá-* ‘солнечное тепло, жар; жаркое время года, день’, греч. *θερμός* ‘теплый, горячий, жаркий’. Наряду с основой мужского рода засвидетельствовано *γā́rmə* f. ‘полдень’ (NEVP: 32). Безусловно, три последних слова являются наиболее сложными для трактовки: приходится констатировать, что ударный зваракай продолжает окситонезу вопреки сформулированному выше правилу, возможно, под влиянием параллельных однокоренных форм женского рода на *-á* < \**á*. Важно отметить, что две из трех обсуждаемых праиранских основ, \**parná-* и \**garmá-*, имеют перед тематическим гласным консонантный кластер, а праформа \**mṛgá-* также может быть реализована на промежуточном этапе с медиальным кластером \**rg-*.

14. *mazgá*, *mā́gzə* m. pl. ‘костный мозг, мозг’ < \**mazgá(n)-*, ср. др.-инд. *majján-* m. ‘костный мозг’ и *majjá-* f. ‘костный мозг; середина растения’. Pl. tantum, ср. рус. *мозги*. Непонятно, является ли рассматриваемая форма старой, то есть *mazgá* < \**mazgáḥ* или к основе

на более позднем этапе присоединяется ударное окончание? В любом случае в пушту не \**n*-основа.

15. *šaudá, šōdád* m. pl. ‘молоко’ < \**xšwiptá-*, ср. др.-инд. *kṣiptá-* ‘брошенный, выгнанный’.

Для данной основы наблюдается значительная вариативность в оформлении, в том числе диалектная, но можно отметить преобладание форм множественного числа в словаре М. Г. Асланова: «*шодá* мн. м. и редко *шодá* ж. вост. ’1) молоко; 2) млечный сок растенный, латекс’; *шодé, шудé* мн. ж. диал. ‘молоко’» (ПРС: 555). Дублиеты женского рода *шудá* и *шудé<sup>2</sup>* отмечены в словаре (КАРС: 346; 348). В словаре Лебедева приводятся формы *шидé* и *шодé* f. pl. (РАСП: 336).

К рассмотренным 15 примерам, которые разбираются в статье (Дыбо 1974), можно добавить дополнительные случаи из этимологического словаря пушту Г. Моргенштерне:

16. *obá* f. pl. ‘вода’ (NEVP: 7). Несмотря на др.-инд. N. pl. *ápās*, ударение падает на конечный зваракай. Дж. Ченг, правилу которого противоречит данный случай, предполагает добавление окончания *-ə* в более позднюю эпоху к \**ob* (Cheung 2011: 178).

17. *olá* m. ‘сакман, сакмал (стадо ягнят)’ < \**wāθβiwa-*, ср. авест. *uiqθβa-* n., *uiqθβā-* f. ‘стадо’ (NEVP: 8). Весьма вероятна такая реконструкция, но нет внешних данных для установления характера первоначального ударения данной основы.

18. *ořá* m. pl. ‘мука’ < \**ā-rtāh* (NEVP: 10). Акцентуация праформы неясна.

19. *bejá* f. ‘stone slab to cover the niche of a sepulchre’ < \**ura-čayā-* ? (NEVP: 14).

20. *kālá* ‘дома’ < \**kātāh* ? (NEVP: 38). Форма определенно связана с *kor* m. ‘дом’ (NEVP: 38).

21. *lemá* m. pl. ‘глаза’ < \**daimā(h)-*, ср. авест. *daēman-* ‘глаз, глазное яблоко (дэвовского существа)’ (NEVP: 43). Подобно рассмотренной выше основе *mazgá, māgzá* m. pl. ‘костный мозг, мозг’ < \**mazgá(n)-* ~ др.-инд. *majjān-* m. ‘костный мозг’, непосредственная праформа для *lemá* выглядит как переход в \**a*-основу с потерей конечного \**-n*.

22. *mandá* m. ‘широкая лощина’, ванечи ‘широкая река в разливе’ < \**ham-ant(i)ya-* ? (NEVP: 50).

23. *māstá*, вазири *mostá* f. pl. ‘кислое молоко; варенец’ < \**mad-tāh* (NEVP: 52). Для слова с такой семантикой оформление в виде pl. tantum достаточно ожидаемо, ср. рус. *сливки*.

24. *niká* m. ‘дед’ — по всей видимости, связано с авест. *niīāka-* ‘дед’, но сохранение \**k* и конечный зваракай делает восстановление праформы сложным (NEVP: 56). Не так ясен и вокализм.

25. *nanjřá, nanzřá* m. pl. ‘деготь, смола’ с непрозрачной внутренней формой, но рассматриваемое слово, очевидно, соотносимо с *rang* ‘цвет’ (NEVP: 58).

26. *rāyá; rayá* f. ‘дол, луга; равнина (каменистая или песчаная)’ — связано с перс. *rāy* ‘то же’, согд. *r’γ*, хот.-сакск. *rrai* ‘равнина’, осет. *rāy* ‘спина; хребет; гребень горы’ (NEVP: 69).

27. *rindá* m. ‘нож шорника’ (NEVP: 70), также *рундá* f. (ПРС: 461). Необходимо отметить параллелизм с рассмотренными выше тремя случаями *тирγá, тэрγá* m. ‘птица (часто хищная)’ ~ *марγá-* f. ‘небольшая птица’; *bāñá* m. sg. и pl. ‘ресница, ресницы’ ~ *bāñá* ‘пе-

<sup>2</sup> Данная форма, очевидно, множественного числа, хотя это и эксплицитно не указано в источнике.



po' и *bānā* 'ресница'; *γārmá* m. 'жара, зной; жар солнца' ~ *γārmá* f. 'полдень'. Однако этимология и, соответственно, акцентуация, неясны.

28. *rānjá* m. pl. 'сурьма, антимоний' (NEVP: 70). Как и приведенные выше формы *nanjrá*, *nanzrá* m. pl. 'деготь, смола', данная основа может быть родственна *rang* 'цвет' — Г. Моргенштерне восстанавливает *\*rangiya-*, отмечая уникальность кластера *-nj-* для языка пушту.

29. *sābá* m. pl., obl. *sabó* 'овощи, зелень', также 'кормовая трава' (NEVP: 73).

30. *skāsá* m., *skāsá* f. 'призрак, привидение' < *\*us-kasya-* (NEVP: 74).

31. *šəwá* f. 'дерево шишам, *Dalbergia sissoo*' (NEVP: 80).

32. *tyāró* f. 'темнота' от *tor* 'черный, темный' (NEVP: 84).

33. *tiyāšá* f. вазири 'лемех' < *\*taš(i)ya-*, ср. перс. *tēša* 'мотыга, кирка' (NEVP: 84).

34. *wāxá* m. pl. 'трава'; ванечи *wušá* 'трава' < *\*wastr(iy)a-* (NEVP: 93).

35. *xwālá* m. 'раскрывающий секреты, тайны' < *\*xwaθya-* ? (NEVP: 97).

Можно заметить, что и в расширенном списке (примеры 16–35) в основном содержатся основы, в которых конечный зваракай с большой вероятностью восходит к *\*(i)ya*, *\*(u)wa* или *\*-āh*.

Стоит обратить внимание на тот факт, что в словаре (КАРС) содержится ряд форм с безударным зваракаем, но все они в подавляющем большинстве имеют стандартные дублиеты с окончанием *-a* в других источниках.

Немногочисленные отличия данных словарей Зудина от данных словарей Асланова и Лебедева сводятся к следующим:

КАРС = Зудин 1950	ПРС = Асланов 1985	РАСП = Лебедев 1973
<i>ácy</i> ( <i>ū</i> ) f. 'поиски, разыскивание' 16 Также <i>hačá</i> 'хлопоты, старания; иск, требование, притязание' 338	<i>áca</i> f. 'старание, стремление; попытка, проба; бросок; наступление, продвижение вперед' 29	<i>háca</i> f. 'стремление, старание' 698, 704
<i>axísty</i> ~ <i>axísty</i> f. 'выдержка, извлечение' 17–18	<i>axistá</i> f. 'прием; принятие, получение; взятие, схватывание; покупка, приобретение' 34	
<i>úrjy</i> ( <i>-ū</i> ) m. 'первое молоко после отёла, молозиво' 21	<i>uržá</i> m. 'молозиво; молоко новотельной коровы со специями' 87	<i>uržá</i> pl. f. 'молозиво' 336
<i>azāngy</i> ( <i>-ūna</i> ) f. 'эхo; гул' 25	<i>azāngá</i> , <i>azāngá</i> f. 'отголосок, отзвук, эхo' 45	<i>azānga</i> f. 'эхo' 809 <i>azāngá</i> f. 'отголосок, отзвук' 454, 456
<i>óžy</i> f. 'чеснок' 47	<i>úža</i> f. 'чеснок' 89	<i>húža</i> f. 'чеснок' 788
<i>ísty</i> f. 'устранение, отстранение, удаление; откладывание' 51	<i>istá</i> m. pl. 'выбрасывание, выкидывание; выгалкивание; исключение, удаление' 96	
<i>bádý</i> ( <i>-ū</i> ) f. 'цель, мишень' 62	<i>báda</i> f. 'пояс шаровар; мишень, цель; луна на ущербе; взятка' 114	
<i>nály</i> ( <i>-ū</i> ) f. 'лемех; осколок (напр. стекла)' 93	<i>nāl</i> m. 'лемех' 155	<i>nāl</i> m. 'лемех' 303

КАРС = Зудин 1950	ПРС = Асланов 1985	РАСП = Лебедев 1973
<i>nā́ngы (-ī)</i> f. 'капитал, состояние' 94 <i>nā́ngы (-ī)</i> f. 'то же' 118	<i>nā́nga</i> f. 'капитал; перен. достояние, ценность, богатство, ресурсы' (ПРС: 156).	<i>nā́nga</i> f. 'капитал' 259
<i>nā́ны (-ī)</i> f. 'клин для колки дров' 94	<i>nā́ná</i> f. 'клин (деревянный); колодки (для пыток)' 156	<i>nā́ná</i> f. 'клин' 267
<i>prá́ты (-ī)</i> f. '1. абсурд, глупость, бессмыслица, чепуха 2. нареч. кроме, без, исключая; отдельно' 101	<i>prá́ta</i> f. 'вздорный, бредовый, бессмысленный; сумбурный' и <i>prá́ta</i> 'без; кроме, помимо, сверх' 164	
<i>ná́rшы (-ī)</i> f. 'каменная глыба; плотный слежавшийся песок' 102	<i>ná́rxa</i> f. 'валун' 166	<i>ná́rxa</i> f. 'глыба (каменная)' 125
<i>trá́ды (-ī)</i> f. 'тесьма; лохмотья, рубище' 137	<i>trá́da</i> f. 'полоска; лента, тесьма' 222	
<i>týrxy (-ī)</i> f. 'полынь, перец' 139	<i>týrxa</i> f. 'полынь' 225	<i>týrxa</i> f. 'полынь' 531
<i>xýры (-ī)</i> f. 'рассвет, раннее утро' 232	<i>xé́ra</i> f. 'наносный ил; муть' 376 <i>nə́ šanə́ xé́ra</i> рано утром, на заре	
<i>sá́нды (-ī)</i> f. 'жалоба, плач, причитание'	<i>sá́nda</i> f. 'стенание, плач, оплакивание' 492	<i>sá́nda</i> f. 'причитание' 576
<i>sá́ngы (-ī)</i> f. 'ветка; пика, копьё' 308	<i>sá́nga</i> f. 'ветвь, ответвление; пика, копьё' 492	<i>śá́nga</i> f. 'ветвь' 68
<i>stá́вы (-ī)</i> f. 'штык, кинжал' 315	<i>stá́va</i> f. 'кинжал' 503	<i>stá́va</i> f. 'кинжал' 264
<i>špá́жы (-ī)</i> f. 'вошь' 338	<i>špá́ža</i> f. 'вошь' 542	<i>špá́ža</i> f. 'вошь' 90 <i>spá́ža</i> f. 'то же' 90
<i>má́лгы (-ī)</i> f. 'соль' 467	<i>má́лга</i> f. 'соль' 782	<i>má́лга</i> f. 'соль' 684

Необходимо указать на значительную вариативность в вокализме и консонантизме отмеченных основ; тем не менее, безударный *-ы (-ə)* в словаре Зудина в других словарях в одном случае соответствует стандартной односложной основе мужского рода ('лемех'), в трех случаях — ударному зваракаю и, наконец, ударному или безударному *-а* во всех остальных приведенных в других лексикографических источниках основах.

Таким образом, в именных основах безударный конечный зваракай в принципе исключен. Далее противопоставлены суффиксы *-ау < \*a-ka-*, который может быть ударным или безударным в зависимости от праиранского ударения, что основательно продемонстрировано в статье Дыбо 1974: 90–102, и его соответствие в женском роде, *-áу < \*ā-kā*, которое всегда ударно, ср. Ве́рка 1969: 85. Соответственно, можно предполагать, что и в данном случае происходила морфологизация ударения в сторону ударности окончания.

Для иллюстрации унаследованного характера различия *'-ау < \*'-a-ka- ~ áу < \*-a-ká-* можно привести ряд примеров, которые по большей части перечислены в статье Дыбо 1974.

#### Основы с безударным суффиксом *-ау*

1. Афг. *ózay* 'плечевая кость' < *\*bāzuka-* < *\*bāzú-*.

2. Афг. *zámay* 'зима' < *\*zímaka-* при производящем *\*zímá-*, ср. авест. *zə́maka-* 'зимняя буря', др.-инд. *himá-* 'холодный; зима'.

3. Афг. *máray* 'труп, мертвец' < *\*mṛtaka-* < *\*mṛtá-*, в древнеиндийском соответствиями являются формы *mṛtá-* 'мертвый' и *mṛtaka-* 'умерший, труп' с незасвидетельствованным ударением.

4. Афг. *tóžay*, ванечи *tárža* ‘томимый жаждой, жаждущий’ < \**tṛsuka-* < \**tṛsú-*, индоевропейские соответствия, указывающие на место ударения, — др.-инд. *tṛsú-* ‘жадный, жаждущий’; прагерм. \**purzu-* < \**pursú-* (др.-в.-нем. *durri*, др.-сакс. *thurri*, др.-англ. *pyrre* «с перестройкой *u*-основы в основу на *-ja-*») (Дыбо 1974: 90–91).

5. Афг. *žóway* ‘живой’ < \**jíwaka-* < \**jīwá-* (ср. др.-инд. *jīvá-* ‘живой, живущий; животворный’).

6. Афг. *stóray* ‘звезда’ < \**stáraka-*, слово имеет точную древнеиндийскую параллель *táraka-* ‘звезда’.

7. Афг. *wárgay* ‘почка’ < \**wṛtkaka-* с контрастной баритонезой, ср. др.-инд. *vṛkká-*, авест. *vṛəḍka-* ‘почка’, по всей видимости, от праиндоевропейского глагола \**wert-* ‘поворачивать’ (EWAia II: 571–572).

### Основы с ударным суффиксом *-áyu*

Здесь следует особо отметить, что в пушту при изначальной праязыковой окситонезе в трехсложных формах на \**-a-ká* происходит перенос ударения на предпоследний слог с дальнейшим усечением последнего слога \**-ka* до *-yu*. Похожие процессы наблюдаются также в парачи и ормури.

1. Афг. *mrayáy* ‘раб, невольник’ < \**maryaká-* < \**márya-*, ср. др.-инд. *márya-* ‘молодой мужчина; жених; молодой женатый мужчина’, авест. *mairiia* ‘мужчина (дэвовское слово); лжец; негодяй, подлец’ и др.-инд. *maryaká-* ‘человечек’.

2. Афг. *nwasáy* ‘внук’ < \**naptraká-*, но Д. И. Эдельман отмечает, что более вероятна реконструкция \**napā(t)saká-*, в ее формулировке — «ном. \**napāts* > \**napās* с наращением суффикса *-aka*» (ЭСИЯ 5: 473).

3. Афг. *wráy* ‘ягненок’ < \**waraká-*; окситонеза \**waraká-* образована по правилу контрастного ударения от вед. *úran-*, *úraṇa-* ‘ягненок’ и продолжает более раннее \**warṇká-*.

4. Афг. *axsáy* ‘шурин, зять’ < \*(*ā*)-*hwaśruká-* или \*(*ā*)-*hwaśuryaká-* при производящей основе \**hwaśura-*, ср. др.-инд. *śváśura-* ‘свекор’, герм. \**swēhuraz* ‘свекор’ (др.-в.-нем. *swehur*, др.-англ. *sweor*).

5. Афг. *zyaráy* ‘желтый, бледный’, *zēřáy* ‘желтуха’ < \**zaritaká-* < \**záríta-*.

Соответственно, на синхронном уровне конечные *-á* и *-áyu* всегда несут ударение, в отличие от *-a* и *-ay*. К сожалению, из-за слабой изученности сравнительно-исторической фонетики пушту и проблем с реконструкцией промежуточных ступеней фонетических переходов от праиранского состояния к современному невозможно оценить, насколько древним является морфологизация ударения на зваракае. Также непонятно, насколько она была вызвана фонетическими факторами.

Имеющиеся данные могут свидетельствовать в пользу того, что основы на \*(*i*)*ya* и \*(*u*)*wa*, а также словоформы, имевшие конечный формант \**-āh*, давали в результате *a*, а \**ā-kā(i)*-основы переходили в конце концов в *áyu*, с перетяжкой ударения. Соответственно, морфологизация \**a*-основ идет в противоположном направлении.

Также параллельная ситуация с ударным зваракаем и потерей окситонезы в именах с праиранской основой на \**-a* может служить свидетельством в пользу того,

что в истории пушту в *\*a*-основах на определенном этапе звукового развития мог появляться *\*u*, который становился редуцированным и терял ударение. Напротив, такие сочетания, как *\*(i)ya*, *\*(u)wa* или *\*-āh* могли давать долгий звук похожего качества, который удерживал и даже перетягивал на себя ударение в отдельных случаях, что в дальнейшем могло привести к окончательной морфологизации конечного ударения на сохранившийся зваракай. Ударный зваракай, соответствующий *\*-am*, сохраняется только в архаических основах за ‘я’ < *\*ahám* и *ōwá* ‘семь’ < *\*haptám*. Также реконструкция *\*abrám* предлагается Дж. Ченгом для *o(w)rá*, *wará* ‘облако’ < (Cheung 2011: 180). Не исключено, что *\*-as* и *\*-am* действительно различались по рефлексам, что влияет на сохранение *-ə* в указанных случаях при условии изначальной окситонезы. С другой стороны, такому правилу будут прямо противоречить разобранные выше примеры *nən* ‘сегодня’ < *\*nīnám*, ср. др.-инд. *nīnám* ‘сейчас’ и *pal* m. ‘ступня, след ноги’ < *\*padám*, что заставляет предпочитать отдельное объяснение для за ‘я’ < *\*ahám* и *ōwá* ‘семь’ < *\*haptám*.

Наконец, есть еще один малочисленный, но примечательный тип основ, а именно с окончанием на долгий *-ā*. Необходимо признать, что его происхождение до сих пор не является установленным, так как и Г. Моргенштерне, и Дж. Ченг возводят *mlā* f., pl. *mlāwi* и *mlāgāni* ‘поясница, чресла’ к праир. *\*mādyā* (NEVP: 49; Cheung 2010: 113), что не может быть верным, поскольку подобные же основы должны давать конечный *-a*, как показывают и приводимый Дж. Ченгом непосредственно рядом с *mlā* < *\*mēlá* < *\*mēlyā* < *\*māidyā* < *\*mādyā* f. пример *p(ə)la* ‘жила; сухожилие; связка’ < *\*pelá* < *\*pēlyā* < *\*pāidyā* < *\*pādyā* f., и другие многочисленные основы языка пушту на *-a*, в том числе разобранные выше. В этимологическом словаре Г. Моргенштерне содержатся следующие формы с конечным *-ā*:

*anā* ‘бабушка’ — может быть заимствованием, ср. тур. *ana*, или как-то соотноситься с осет. *ama* ‘мать’ или собственно пушту *nyā* ‘бабушка’ (NEVP: 8);

*byā*, ванечи также *byār*, *bā* ‘снова’ — возможно, индоарийское заимствование, ср. пенджаби и лахнда *biā* ‘again’ и т. д., выведение в EVP из иранского и сравнение с авест. *araiia* ‘позже, потом’ менее вероятно (NEVP: 16);

*dā* adj.; *day* m., *dā* f. (косв. *də* m., *de* f., pl. *duy*) pron. ‘этот’ — выведение < *\*ita*/*\*ayta* представляется наиболее вероятным, но появление долгого гласного в исходе не так ясно (NEVP: 22);

*γwā*, вазири *γwo* f. ‘корова’ < *\*gawā*- (NEVP: 33). Непосредственно такой вид основы представляется сомнительным ввиду того, что ожидалось бы сокращение конечного пшт. *-a* < *\*-ā*.

*hā* f., PQ *áya*, (*h*)*óya*, африди *woya*, вазири *yowya*, *yīya* ‘яйцо’ < Г. Моргенштерне реконструирует *\*āyā*- наряду с *\*āwyā*, ср. авест. *aīia*- ‘яйцо’ (NEVP: 35);

*kaṇyá* ‘несметный’ < *\*a-karana*- ?, ср. пехлеви *a-kanārag* ‘безграничный’. Окончание *-yā* при этом остается неясным (NEVP: 39)

*lā*, вазири и др. диалекты *lya* ‘еще, все еще, до сих пор’ — выведение из авест. *haða* ‘всегда’ не может фонетически объяснить форму *lya* (NEVP: 42);

*nyā*, вазири *nio* ‘бабушка’ — сопоставляется с авест. *nīiākā*-, но на это слово могло оказать влияние перс. *nīya* ‘дедушка’ (NEVP: 60);

*plā* f., pl. *-we* ‘путешествие, поездка за рубеж (в т. ч. с торговыми целями); раз, однажды; конец, поездка в одну сторону’ — данное слово сравнивается с авест. *paθā-* ‘дорога’, Г. Моргенштерне с сомнением приводится праформа *\*paθuyā-*? (NEVP: 69).

*šā* f. ‘спина’. Этимология неясна (NEVP: 79);

*tanā́, taṇā́* f. ‘гром’ — ср. перс. *tandar* ‘гром’, *ṇ* может быть вторичного происхождения (NEVP: 82);

*xwā* f.; вазири *xwo* ‘сторона, направление’. Связно с *pa-xwā́ ~ pə-xwā́* ‘до, прежде чем; прежде, раньше’, этимология неясна (NEVP: 97);

*xwərá* adv. ‘очень’ — ср. пехлеви *frāy* ‘много, больше’ и т. д. (NEVP: 97);

*zoḥā́* f. ‘вид сиропа (сделанного из ягод)’ — в этимологическом словаре пушту выводится из *\*zauštrā-*, ср. авест. *ziš-* ‘получать удовольствие, наслаждаться’ (NEVP: 104).

Все эти слова также имеют конечное ударение. Возможно, конечный *-ā* восходит к какой-то последовательности типа *\*āyā, \*āwā* и т. п., а в дальнейшем всегда получает конечное ударение вследствие морфологизации, как и в других случаях. Еще менее ясно происхождение ударного суффикса *-í* — для его трактовки просто нет достаточного количества этимологического материала.

Таким образом, при указании на архаичность ударения в пушту следует делать оговорку: архаическое праиранское ударение лучше всего прослеживается в словах, продолжающих *\*ā*-основы, на что обращал внимание Г. Моргенштерне (Morgenstierne 1983: 169), а также в словах на *\*-ay*, суффикс которых восходит к *\*-a-ka* — данный тип наиболее последовательно рассмотрен В. А. Дыбо, как было указано выше; это же касается ряда глагольных основ (Дыбо 1989). Многие другие основы, по всей видимости, морфологизировали свое ударение: *\*a*-основы утрачивали окситонезу и начинали выглядеть так же, как изначально баритонированные; напротив, слова пушту с конечным зваракаем и долгим *-ā* обобщают конечное ударение. Сюда же можно добавить многочисленные примеры конечноударных суффиксов. Развитие акцентуации в пушту походит в ряде отношений на эволюцию ее в парачи и ормури, где подобные процессы происходят с *\*a*-основами и рядом других, а старое ударение можно наблюдать только в ограниченном ряде имен и глаголов и их форм.

## Литература

- Герценберг, Л. Г. 1981. Об афганском ударении. *Иранское языкознание. Ежегодник 1980*. Москва: 48–56.
- Грюнберг, А. Л. 1987. *Очерк грамматики афганского языка (пашто)*. Ленинград: Наука.
- Грюнберг, А. Л., Д. И. Эдельман. 1987. Афганский язык. В кн.: В. И. Абаев, М. Н. Боголюбов, В. С. Расторгуева (ред.). *Основы иранского языкознания. Новоиранские языки: Восточная группа*: 6–154. Москва: Наука.
- Дыбо, В. А. 1961. Сокращение долгот в кельто-италийских языках и его значение для балто-славянской и индоевропейской акцентологии. *Вопросы славянского языкознания* 5: 9–34.
- Дыбо, В. А. 1972. О рефлексах индоевропейского ударения в индоиранских языках. В кн.: С. Б. Бернштейн (ред.). *Конференция по сравнительно-исторической грамматике индоевропейских языков (12–14 декабря): Предварительные материалы*: 38–44. Москва: Наука.
- Дыбо, В. А. 1974. Афганское ударение и его значение для индоевропейской и балтославянской акцентологии. I. Именная акцентуация. В кн.: Т. М. Судник (ред.). *Балто-славянские исследования*: 67–105. Москва: Наука.

- Дыбо, В. А. 1989. Афганское ударение и его значение для индоевропейской и балтославянской акцентологии. II. Глагольная акцентуация. В кн.: Н. И. Толстой (ред.). *Славянский и балканский фольклор. Просодия: Сборник статей*: 106–147. Москва: Наука.
- Дыбо, В. А. 1990. О ритмической концепции Л. Г. Герценберга в пуштунской акцентологии. В кн.: В. М. Гусалов (ред.). *Вопросы иранистики и алановедения (Научная конференция, посвященная 90-летию В. И. Абаева). Владикавказ, 18–20 октября 1990 г. Тезисы докладов*: 10–12. Владикавказ: Северо-Осетинский НИИ ИФЭ.
- Дыбо, В. А. 2011. Древнеиндийский акцент в дардском языке шина как проблема индоевропейской акцентологии. В кн.: Е. К. Молчанова (ред.). *Лексика, этимология, языковые контакты: к юбилею доктора филологических наук, профессора Д. И. Эдельман*: 92–167. Москва: Тезаурус.
- Ефимов, В. А. 1986. *Язык ормури в синхронном и историческом освещении*. Москва: Наука.
- КАРС = Зудин, П. Б. 1950. *Краткий афганско-русский словарь*. Москва: Государственное издательство иностранных и национальных словарей.
- ПРС = Асланов, М. Г. 1985. *Пушту-русский словарь*. 2-е изд. Москва: Русский язык.
- РАСП = Лебедев, К. А., Л. С. Яцевич, З. М. Калинина. 1973. *Русско-афганский словарь (пушту)*. Москва: Советская энциклопедия.
- ЭСИЯ = Рассторгуева В. С., Д. И. Эдельман. 2000-. *Этимологический словарь иранских языков*. Москва: Восточная литература РАН.

## References

- Bečka, Jiří. 1969. *A Study in Pashto Stress (Dissertationes Orientales, 12)*. Prague: Oriental institute in Academia.
- Cheung, Johnny. 2010. Selected Pashto problems. I: The accent in Pashto. *Persica* 23: 109–121.
- Cheung, Johnny. 2011. Selected Pashto problems II. Historical phonology 1: On vocalism and etyma. *Iran and Caucasus* 15: 169–205.
- Dybo, Vladimir A. 1961. Sokrashchenie dolgot v kel'to-italijskix jazykax i ego znachenie dl'a balto-slav'anskoj i indoevropskoj akcentologii. *Voprosy slav'anskogo jazykoznanija* 5: 9–34.
- Dybo, Vladimir A. 1974. Afganskoe udarenie i ego znachenie dlja indoevropskoj i baltoslavjanskoj akcentologii. I. Imennaja akcentuacija. In: T. M. Sudnik (ed.). *Baltoslavjanskije issledovanija*: 67–105. Moskva: Nauka.
- Dybo, Vladimir A. 1989. Afganskoe udarenie i ego znachenie dlja indoevropskoj i baltoslavjanskoj akcentologii. II. Glagol'naja akcentuacija. In: N. I. Tolstoj (ed.). *Slavjanskij i balkanskij fol'klor. Prosodija: Sbornik statej*: 106–147. Moskva: Nauka.
- Dybo, Vladimir A. 1990. O ritmicheskoj koncepcii L. G. Gercenberga v pushtunskoj akcentologii. In: V. M. Gusalov (ed.). *Voprosy iranistiki i alanovedenija (Nauchnaja konferencija, posv'ashchennaja 90-letiju V. I. Abaeva). Vladikavkaz, 18–20 okt'abr'a 1990 g. Tezisy dokladov*: 10–12. Vladikavkaz: Severo-Osetinskij NII IFÈ.
- Dybo, Vladimir A. 2011. Drevneindijskij akcent v dardskom jazyke shina kak problema indoevropskoj akcentologii. In: E. K. Molchanova (ed.). *Leksika, ètimologija, jazykovye kontakty: k jubileju doktora filologičeskix nauk, professora D. I. Èdel'man*: 92–167. Moskva: Thesaurus.
- Efimov, Valentin A. 1986. *Jazyk ormuri v sinkhronnom i istoričeskom osveschenii*. Moskva: Nauka.
- ESIJ = Rastorgueva, Vera S., Dzhoy I. Èdel'man. 2000-. *Ètimologičeskij slovar' iranskix jazykov*. Moskva: Vostochnaja literatura RAN.
- EVP = Morgenstierne, Georg. 1927. *An Etymological Vocabulary of Pashto*. Oslo: Jacob Dybwad.
- EWAia I–II = Mayrhofer, Manfred. *Etymologisches Wörterbuch des Altindoarischen. I–II Bde*. Heidelberg: Universitätsverlag Winter.
- Gauthiot, Robert. 1916. De l'accent d'intensité iranien. *Mémoires de la Société de linguistique de Paris* 20(1): 1–25.
- Grjunberg, Aleksandr L. 1987. *Očerki grammatiki afganskogo jazyka (pashto)*. Leningrad: Nauka.
- Grjunberg, Aleksandr L.; Edel'man, Dzhoy I. 1987. Afganskij jazyk. In: *Osnovy iranskogo jazykoznanija: Novoiranskije jazyki: Vostochnaja gruppy*. Moskva: 6–154.
- Herzenberg, Leonard G. 1981. Ob afganskom udarenii. *Iranskoe Jazykoznanie. Ezhegodnik* 1980: 48–56.
- Hirt, Hermann. 1929. *Indogermanische Grammatik. V. Der Akzent*. Heidelberg: Winter.
- IEW = Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch*. Bern: French & European Publications.
- KARS = Zudin, Petr B. 1950. *Kratkij afgansko-russkij slovar'*. Moskva: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarěj.

- Meillet, Antoine. 1897. *Recherches sur l'emploi du génitif-accusatif en vieux-slave*. Paris: Hachette.
- Meillet, Antoine. 1900. La déclinaison et l'accent d'intensité en perse. *Journal asiatique* 15: 254–277.
- Morgenstierne, Georg, Arthur Lloyd-James. 1928. Notes on the Pronunciation of Pashto (Dialect of the Hazara District). *Bulletin of the School of Oriental Studies* 5(1): 53–62.
- Morgenstierne, Georg. 1942. Archaisms and innovations in Pashto Morphology. *Norsk Tidsskrift for Sprogvidenskap* 12: 88–114.
- Morgenstierne, Georg. 1973. Traces of Indo-European Accentuation in Pashto? *Norwegian Journal of Linguistics (formerly Norsk Tidsskrift for Sprogvidenskap)* 27: 61–65.
- Morgenstierne, Georg. 1983. Bemerkungen zum Wort-Akzent in den Gathas und im Paschto. *Münchener Studien zur Sprachwissenschaft* 42: 167–175.
- Mylius, Klaus. 1975. *Wörterbuch Sanskrit-Deutsch*. Leipzig: Verlag Enzyklopädie, VEB.
- NEVP = Morgenstierne, Georg. 2003. *A new etymological vocabulary of Pashto / by Georg Morgenstierne; compiled and edited by J. Elfenbein, D. N. MacKenzie and Nicholas Sims-Williams*. Wiesbaden: Reichert.
- Novák, L'ubomír. 2013. *Problem of Archaism and Innovation in the Eastern Iranian Languages*. PhD. Diss. Charles University.
- PRS = Aslanov, Martiros G. 1985. *Afgansko-russkij slovar (pushtu)*. 2<sup>nd</sup> ed. Moskva: Russkij jazyk.
- RASP = Lebedev Konstantin A., L'udmila S. Jacevič, Zoja M. Kalinina. 1973. *Russko-afganskij slovar' (puštu)*. Moskva: Sovetskaja enciklopedija.
- Raverty, Henry G. 1856. *A Grammar of the Pukhto, Pushto or Language of the Afghans*. Calcutta: J. Thomas.
- Raverty, Henry G. 1860. *A Dictionary of the Pukhto, Pushto or Language of the Afghans*. London: Longman, Green & Co.
- Yakubovich, Ilya, Alice Mouton. 2023. *Hittite-Luwian Ritual Texts Attributed to Puriyanni, Kuwattalla and Šilalluḫi (CTH 758–763)*. Wiesbaden: Harrassowitz.

*Artem Trofimov*. On the rules of stress in Pashto nominal stems

The aim of this paper is to clarify the rules of accentuation in Pashto and its historical background. Based on etymological analysis of different morphological categories, I propose the following explanation of the development of Pashto accentuation patterns in the most common nominal stems: (a) words originating in Proto-Iranian *\*ā*-stems and stems with the suffix *\*-a-ka* retain the Proto-Iranian accent; (b) *\*a*-stems usually lose oxytonesis; (c) stems resulting in Pashto *ə* or final long *-ā* generalize the final stress. The data demonstrate that the original paradigms with mobile stress underwent morphologization in Pashto. As a result, the vast majority of modern Pashto forms have stress on the final syllable.

*Keywords*: Pashto language; Proto-Iranian language; accentuation; Old Indian language; Proto-Indo European.





**Issue 4 / Часть 4**

**Articles / Статьи**



## The uvularization of *g-/d-* in Tibetic languages<sup>1</sup>

This study reviews four competing explanations for the origins of uvular preinitials in Tibetic lects, making a specific case study of modern uvular preinitial reflexes from Old Tibetan *g-/d-*. The first explanation is from Huang (2012), who claims that uvular preinitials were phonologically present in Pre-Tibetan, and thus at least some of the uvular preinitials in modern Tibetic lects descend from this Pre-Tibetan strata. Her argument is predicated on the hypothesis that Tibetic lects broke up into different languages before Old Tibetan was reduced to writing in the 7th century AD. The second explanation is from Hill (2010), who argues that all uvular preinitials are not inherited from Pre-Tibetan but are the result of language contact with Qiangic and/or Mongolic languages. Differing from Huang's explanation, Hill's explanation rests on the theory that all modern Tibetic lects descend from Old Tibetan. The third explanation assumes that Hill is correct in the claim that there were no uvular preinitials in Old Tibetan, and claims that there is a regular sound change from *g-/d-* to velar/postvelar/uvular fricatives (except before velar initials, where the change is to *r-*) in Amdo lects and to uvular fricatives in Gyalrongic lects. For WAT lects, *g-* regularly changes to velar/postvelar/uvular fricatives, but *d-* changes to velar/postvelar/uvular or *ʃ-/r-*. The fourth explanation is that in Old Tibetan *ɣ* and *g-/d-* were in velar and uvular free variation, and thus uvular preinitials do come from Old Tibetan, but originate from phones and not phonemes. The first three explanations are scientific hypotheses; i.e., they can be tested through evidence and are falsifiable. The final explanation (appealing to free variation) is not a falsifiable. After examining the evidence on the timing of the breakup of the Tibetic lects, Huang's hypothesis is eliminated, leaving only Hill's explanation and 'Explanation 3' standing. However, Explanation 3 is the only explanation that proposes a set of regular sound changes to summarize the uvularization of *g-/d-*.

**Keywords:** Sa-skya Paṇḍita's Law; Pre-Tibetan language; Old Tibetan language; Gyalrongic languages; uvular reflexes of *g-/d-* in Tibetan.

### 1. Introduction

This study reviews four competing explanations for the origins of uvular preinitials in Tibetic lects, making a specific case study of modern uvular preinitial reflexes of Old Tibetan *g-/d-*. Section 2 discusses Sa-skya Paṇḍita's Law and the resulting complementary distribution of *g-* and *d-*, and Section 3 presents words with *g-/d-* reflexes in modern Tibetic languages and loanwords with *g-/d-* reflexes in modern Gyalrongic languages.

Section 4 discusses an explanation from Huang (2012), who postulates that uvular preinitials were phonologically present in Pre-Tibetan (inherited from Tibeto-Burman) and thus uvular *g-/d-* in modern Tibetic lects are reflexes, descending from an unwritten Old Tibetan

---

<sup>1</sup> This paper has received funding from: China Postdoctoral Science Foundation (73rd batch), Funding number: 2023M731815: "Anticausative marking in West Gyalrongic languages" and National Social Science Fund Key Project (21AYY024), People's Republic of China: "Research on Tibetan poetry, rhythm, and related prosodic phonology."

Glossing in the examples follow the Leipzig Glossing Rules. Old Tibetan and Written Tibetan transcriptions follow the Wylie transcription system (Wylie 1959), except in the case of  $\alpha$  where *ɣ* is used instead of an apostrophe, following Hill (2019b).

with phonological uvular preinitials. Her argument is predicated on the notion that there was a split in Old Tibetan that resulted in a non-uvular written Old Tibetan (based on an old U-Tsang lect) as one branch (with Written Tibetan and modern U-Tsang as descendants), and another branch of uvular Tibetic lects (e.g., Kham and Amdo)<sup>2</sup>.

Section 5 examines Hill's 2010 explanation, including his explanation for why *g-/d-* has uvular fricative reflexes in some modern Tibetic lects. Hill argues that uvular preinitials are not inherited from Pre-Tibetan but are the result of language contact with Qiangic and/or Mongolic languages. Differing from Huang's explanation, Hill's explanation rests on the theory that all modern Tibetic lects descend from Old Tibetan.

Section 6 discusses a third explanation. Explanation 3 assumes that Hill is correct in the claim that there were no uvular preinitials in Old Tibetan, and claims that there is a regular sound change from *g-/d-* to velar/postvelar/uvular fricatives (except before velar initials, where the change is to *r-*) in Amdo lects and to uvular fricatives Gyalrongic lects. For WAT lects, *g-* regularly changes to velar/postvelar/uvular fricatives, but *d-* changes to velar/postvelar/uvular or *ʃ-/r-*. Hill's notion of contact induced uvularization is acknowledged as a possible origin story, but in light of the regularity of the sound change it is not explored further.

Section 7 discusses the fourth explanation, which is that in Old Tibetan *h* and *g-/d-* were in velar and uvular free variation, and thus uvular preinitials do come from Old Tibetan, but originate from phones and not phonemes.

Section 8 evaluates each of these four explanations. The first three explanations are scientific hypotheses; i.e., they can be tested through evidence and are falsifiable. The final explanation (appealing to free variation) is not a testable hypothesis, but merely one possibility. After examining the evidence on the timing of the breakup of the Tibetic lects, Huang's hypothesis is eliminated, leaving only Hill's explanation and Explanation 3 standing. Explanation 3 is the only explanation that proposes a set of regular sound changes to summarize the uvularization of *g-/d-*, and thus this explanation is considered as the best choice thus far.

## 2. Sa-skya Paṇḍita's Law

In the synchronic phonology of both Written Tibetan and Old Tibetan, *g-* and *d-* are in complementary distribution<sup>3</sup>. Sa-skya Paṇḍita's Law claims that since Old Tibetan *g-* and *d-* are in complementary distribution, *g-* only coming before acutes and *d-* only coming before graves, then the following sound changes must have taken place: *\*g- > d-* before graves (labials and velars) and *\*d- > g-* before acutes (dentals and palatals) (Hill 2011: 443-444, 2019b: 24)<sup>4</sup>. This conditioned neutralisation happened before Old Tibetan (i.e., Pre-Tibetan), or may have even occurred much earlier in the split off of Proto-Bodish from Proto-Trans-Himalayan (Proto-Sino-Tibetan).

<sup>2</sup> Personally, I don't like the term 'Written Tibetan' as simply 'Tibetan' would suffice for this *état de langue*. I use 'Written Tibetan' to identify with other scholars who use this term and in order to avoid confusion with modern spoken Tibetan (or Tibetic) lects. In Tibetan literary studies the synonym 'Classical Tibetan' is used.

<sup>3</sup> Phonotactics play an important role in the allophony of *g-* and *d-*. The claim is only made for *g-* and *d-* as preinitials to be allophones, not for the initials *g* and *d* to be allophones. Thus, before the medial *-r-*, *g-* and *d-* are in contrast in both Old and Written Tibetan. See also Jacques 2014: 157.

<sup>4</sup> The terms 'grave' and 'acute', coined by Jakobson et al. (1951: 29–30), have fallen out of fashion in general linguistic literature. However, there are two main reasons for resurrecting these terms. Firstly, the term 'grave' is typologically quite satisfying, since labials and velars have a phonetic connection in many languages of the world. Secondly, these terms are helpful for engaging in related Sino-Tibetan historical comparison, as these terms are used for Old Chinese by Baxter (1992), Baxter & Sagart (2014), and Hill (2019b), among others.

It not impossible to imagine that *\*g-* and *\*d-* were prefixes in Proto-Trans-Himalayan. One hypothesis proposed by Jacques (2008: 53–57, 2014: 158–160), who provides comparative evidence with Tibetan and Gyalrongic, is that *\*g-* comes from an animal noun classification prefix (perhaps *\*gV-*) and *\*d-* is a lexicalized indefinite possessor prefix in inalienable (kinship and body part) nouns (perhaps originating as *\*dV-*)<sup>5</sup>. For our purposes at the moment it is suffice to be agnostic concerning their function, and even entertain the notion that *\*g-* and *\*d-* were never prefixes. Whether they were prefixes or not does not have a direct effect on our present study, but for now I will just assume that these were prefixes at some earlier stage.

Over time, these prefixes lost their meaning and became fossilized as non-productive pre-initials, merging with preinitials of the same phonological value but unrelated to the prior prefix; *\*g-* and *\*d-* at this stage remained separate phonemes. Next comes a purely phonological change; Sa-skya Paṇḍita's Law redistributes these fossilized prefixes through a new phonological constraint (*g-* only before acutes and never before graves, and *d-* only before graves and never before acutes), so that what was once in principle *\*dd*, *\*dts*, *\*ds*, *\*dz*, *\*dn*, *\*dsh*, *\*dj*, etc. is now *gd*, *gts*, *gs*, *gz*, *gn*, *gsh*, *gj*, etc., and what was once *\*gk*, *\*gg*, *\*gp*, *\*gb*, *\*gm*, etc. is now *dk*, *dg*, *dp*, *db*, *dm*, etc. Thus, *g-* and *d-* as preinitials are in complementary distribution, and become allophones of the same preinitial phoneme, while still remaining in contrast as initial and final position phonemes. If there was a written language for Pre-Proto-Bodish, then we would expect to find examples of at least some of the following forms: *\*dd*, *\*dts*, *\*ds*, *\*dz*, *\*dn*, *\*dsh*, *\*dj*, *\*gk*, *\*gg*, *\*gp*, *\*gb*, *\*gm*, etc. It is possible that part of the motivation for Sa-skya Paṇḍita's Law is the avoidance of double consonants, e.g., *dd* and *gg*, which then became *gd* and *dg*, respectively. Thus, Sa-skya Paṇḍita's Law is mostly a process of dissimilation.

If we combine Jacques' hypothesis (*g-* is from an animal prefix and *d-* is from a kinship/body part prefix) with Sa-skya Paṇḍita's Law we can assume that in some Tibetan words with a *g-* preinitial before an acute initial (e.g., in some animal terms) no change occurred (*g-* > *g-*), but when we have *g-* before an acute initial for a kinship or body part word (like *gzhang* 'anus') then the *d-* > *g-* change occurred; in some words with a *d-* preinitial before a grave initial (e.g., body part terms) no change occurred (*d-* > *d-*), but when we have a *d-* preinitial before a grave initial in animal words (*dgo ba* 'Mongolian gazelle', *dbyi* 'lynx') then *g-* > *d-*. The term *dbyi* 'lynx' also has an alternative spelling *g.yi*, which hints to this sound change and shows the complementary distribution (*g-* cannot occur before *b*), which led to the ousting of *b* and the retention of *g-* in *g.yi*. In Sertha, 'eye' is pronounced by some varieties as *ɣəniŋ* and other varieties as *ɣəniŋ*; turning the original preinitial into a uvular or velar fricative initial of a new syllable by adding an epenthetic schwa.

### 3. Reflexes of *g-/d-* in Tibetic words and Gyalrongic loanwords

Table 1 tabulates some examples of *g-* in Tibetic and Gyalrongic lects, and Table 2 gives examples of *d-*. Since Gyalrongic languages have been in contact with Tibetic languages for over a millennium, and have thus borrowed extensively from Tibetic, Gyalrongic languages provide

<sup>5</sup> This proposal has been called into question by Hill (2014: 629–630), because it is impossible to completely rule out chance operations (e.g., it would be preposterous to propose that the 'h' in 'head', 'hand', 'heart', 'heel', etc. indicate a *h-* prefix that is related to body parts in English). However, Jacques (2014: 158–160) is not proposing a body part prefix but an indefinite possessor prefix of inalienably possessed nouns, which is well attested in Gyalrong (Jacques 2021: 114). The *d-* preinitial in Tibetan is but a fossilized trace of that prefix. Since in Old Tibetan *g-/d-* are found covering a broad range of semantic domains, e.g., body parts, animals, colors, times of day, abstract notions, etc., it could be that there were several prefixes with velar or dental initials that merged in Proto-Tibetan (Bialek p.c.)

an excellent point of comparison. These tables are in no way exhaustive, but still provide ample evidence for this study.

The data in the columns ‘Balti’ and ‘WAT’ (Western Archaic Tibetan) comes from Bielmeier et al. (2018; Forthcoming). WAT examples give the lect name in parentheses after the WAT form. The data in the column ‘Amdo’ comes from Bielmeier et al. (2018; forthc.) and Hua (2001). Data from Bielmeier et al. (2018; forthc.) indicate the lect in parentheses after the WAT and Amdo form, but data from Hua do not indicate the lect (more explanation is given in the next paragraphs). In order to simplify the Amdo data from Hua, I only include examples from lects that have preinitials. If no preinitial was found I chose just one of the Amdo lects to represent them all, as lects that have lost the preinitial have no bearing on our study. Although Hua has chosen to use the voiceless glottal fricative *h-* as a preinitial, it is likely that the voiceless velar fricative *x-* would be a better choice for some if not all of these examples. In any case, there is no phonological contrast between [h-] and [x-] as a preinitial in these lects; and even if there is a reflex *h-*, this is still explainable by the hypotheses presented in this article.

The following are notes for the Amdo data from Hua (2001) in Tables 1 and 2, indicating which dialects the data reflects. If nothing is said about the Amdo forms below, then all the lects of Amdo from Hua have the same form. The Amdo for *gtsang ma* ‘clean’ is from Rebong (Tongren), Xunhua, Hualong, and Skakhog/Skachu (Hongyuan). According to Hua, in Xunhua *gser* ‘gold’ is *hse*, in Hualong *hsar*, and in Skakhog/Skachu (Hongyuan) and Themchen (Tianjun) it is *hser*; *gso* ‘nourish’ is from Bsangchu (Xiahe), Rebong (Tongren), Skakhog/Skachu (Hongyuan), Themchen (Tianjun). The data for *gzah hkhor* is from Skakhog/Skachu (Hongyuan); *hza* being the first part of *hzadawa* ‘Monday’. The word for *dpon* in Themchen (Tianjun) is *hwonbu*. Hua has recorded that *dgah* has become *hga* in Bsangchu (Xiahe), Rebong (Tongren), Xunhua, and Xualong, but in Skakhog/Skachu (Hongyuan) and Themchen (Tianjun) has become *rga*.

G.yukhog and Stau data come from my own fieldwork (Gates 2021). G.yukhog is traditionally considered to be an Amdo lect, but I display the data separately since it has close proximity to Stau and provides ample examples of uvular preinitials for a Tibetic lect. Geshiza data comes from Honkasalo (2019) and personal communication with Honkasalo (October 31 – November 1, 2021). Khroskyabs data comes from Lai (2017) and personal communication with Lai (October 31, 2021). Japhug data comes from Jacques (2015–2016).

In both Tables 1 and 2, blank places indicate that there is no relevant comparanda available; in the case of the Tibetic lects no cognates were found, in the case of Gyalrongic no loanwords were found.

Balti *g-/d-* reflexes are mostly *x-/ɣ-*, with the occasional *g-* reflex as *χ-* (e.g. *χsu* ‘feed’) and *ʈ-* (e.g. *ʈsanma* ‘clean’), and the occasional *d-* reflex as *ʂ-* (e.g. *ʂpuxma* ‘shoulder’). According to Bielmeier et al. (2018: 50, 52), in Skardu (Western Balti) and Khaplu (Eastern Balti) “*x* and *ɣ* are postvelar.” The most common *d-* reflex in WAT Purik lects is *ʂ-/r-*.

Amdo *g-/d-* reflexes tend to be velar fricatives, with the exception of Sertha and G.yukhog, which have uvular fricative reflexes. Sertha adds an epenthetic schwa after the *g-* uvular fricative reflex, as in *ʂəjak* ‘yak’.

In Gyalrongic languages, *g-/d-* loanwords have a uvular fricative preinitial *ʂ-/χ-* except for *lonbutɕʰi* ‘elephant’, in which the preinitial has been lost, and the forms for the etymon ‘like’, which all have the reflex *r-*<sup>6</sup>. The complementary distribution resulting from Sa-skya Paṇḍita’s

<sup>6</sup> There are many instances in Tibetic and Gyalrongic lects where *g-/d-* preinitials have been lost altogether, but this is a trivial sound change that needs no further discussion.

Written Tibetan	Gloss	Balti	Amdo	G.yukhog	Stau	Geshiza	Khroskyabs	Japhug
<i>gtsang ma</i>	clean	<i>ɦtsaɦma</i>	<i>ɦitsaɦma</i>	<i>ɦtsomæ</i>	<i>ɦtsomæ</i>	<i>atsɔɦ</i>		<i>ɦtso</i>
<i>gser</i>	gold	<i>xser</i> (title: Amacha)	<i>yser</i> (Arik)	<i>ksər</i>	<i>ksər</i>	<i>asər</i>	<i>ɦsər</i>	<i>ɦsər</i>
<i>gso</i>	feed	<i>ɦsu</i>	<i>ɦsu</i> (Arik)	<i>ɦsu</i> ‘eat (hon.)’	<i>ɦsu</i> ‘eat (hon.)’			<i>ɦsu</i>
<i>gsur</i>	offering for <i>yi dwaqs</i>			<i>ksər</i>	<i>ksər</i>			
<i>gzah ɦkhor</i>	week	<i>yza</i> ‘rainbow’	<i>ɦza</i>	<i>ɦzæɦk’or</i>	<i>ɦzæɦk’or</i>			
<i>gnad ɦgag</i>	problem			<i>ɦnæɦgag</i>	<i>ɦnæɦgag</i>			
<i>gnatj</i>	long ago		<i>ɦna</i> (Arik)	<i>ɦna</i>	<i>ɦna</i>		<i>ɦnâ</i>	
<i>glu</i>	song	<i>ɦlu</i>	<i>ɦlæ</i> (Sertha)	<i>ɦlæ</i>	<i>ɦlæ</i>	<i>alæ</i>	<i>ɦlê</i>	
<i>gla</i>	hire/salary	<i>ɦla</i> ‘tribute’	<i>ɦla</i> (Rkangtsha)	<i>ɦlæ</i>	<i>ɦlæ</i>	<i>ala</i>		
<i>gshal</i> ‘wash’	flow			<i>ɦsər</i>	<i>ɦsər</i>			
<i>gdam pa</i>	leader			<i>ɦdæɦbæ</i>	<i>ɦdæɦbæ</i>			
<i>gdan</i>	carpet	<i>ɦdan</i> ‘area’	<i>ɦdan</i> (Rkangtsha)	<i>ɦden</i>	<i>ɦden</i>			<i>ɦden</i> ‘cushion’
<i>g.ya</i>	rust	<i>ɦja</i>			<i>ɦja</i>			<i>ɦja</i>
<i>glang chen</i>	elephant	<i>ɦlangpotɦo</i>	<i>ɦlangɦ’en</i> (Rkangtsha)	<i>ɦlæɦotɦ’e</i>	<i>ɦlæɦotɦ’e</i>	<i>alæɦbutɦe</i>	<i>ɦlæɦbutɦ’e</i>	<i>læɦbutɦ’i</i>
<i>gcun gzan</i>	carnivorous animal	<i>ɦtɦanzan</i> ‘snow leopard’	<i>ɦtɦandzon</i> (Arik)	<i>ɦtɦendzen</i>	<i>ɦtɦendzen</i>	<i>ɦtɦæɦzæn</i>		<i>ɦtɦæɦzæn</i>
<i>g.yu</i>	turquoise	<i>ɦju</i>	<i>ɦjy</i> (Sertha)	<i>ɦjæ</i>	<i>ɦjæ</i>		<i>ɦjikæɦ</i>	<i>ɦju</i>
<i>g.yag</i>	yak	<i>ɦjaq</i> ‘crossbread’	<i>ɦjak</i> (Sertha)	<i>ɦjiɦgæ</i>	<i>ɦja</i>	<i>ajæ</i>		
<i>gzig</i>	leopard		<i>ɦzæk</i> (Arik)	<i>ɦzuɦq</i>				

Table 1: Words with g-

Tibetan	Gloss	WAT	Amdo	G.yukhog	Stau	Geshiza	Khroskyabs	Japhug
<i>dpe</i>	metaphor/example	<i>xpe</i> (Balti)	<i>χ'we</i> 'story' (Ndzo)	<i>κve</i>	<i>κpi</i>	<i>api</i>	<i>χpi</i>	<i>χpi</i>
<i>dpaŋ bo</i>	hero	<i>ξpao</i> (Chiktan Purik) <i>ξponbo</i> (Chiktan Purik)	<i>χ'wawo</i> (Rkangtsha) <i>χ'woupo</i> (Rkangtsha)	<i>κpawo/κawawo</i> <i>κwəni</i>	<i>κpawu</i> <i>κpən</i>		<i>χpawu</i> <i>χpən</i>	<i>χpawu</i> <i>χpən</i>
<i>dmaŋ</i>	low		<i>γma</i> (Arik)	<i>κmæ</i> 'low altitude'	<i>κmæ</i> 'low altitude'			<i>κma</i>
<i>dmag mi</i>	soldier	<i>ξmaqmi</i> (Kargil Purik)	<i>hmak</i>	<i>κmaqmiʒ</i>	<i>κmami</i>	<i>amawə</i>	<i>κmæʒ</i>	<i>κmami</i>
<i>dgaŋ</i>	like	<i>rgak'an</i> 'friend' (Chiktan Purik)	<i>hga/rga</i>	<i>rgæ</i>	<i>rgæ</i>	<i>rga</i>	<i>rga</i>	<i>rga</i>
<i>dpung pa/ma</i>	shoulder	<i>ξpuxma</i> (Balti)	<i>χ'wuyja</i> (Rkangtsha)	<i>κwō</i>	<i>κpō</i>			

Table 2: Words with *d-*



Law can be seen in Gyalrongic *g-/d-* loanwords. For example, if Tibetan loanwords in Stau have the preinitial *ɸ-* followed by a labial, then *ɸ-* corresponds to Tibetan *d-*, otherwise the preinitial *ɸ-* corresponds to Tibetan *g-* (compare also Lai 2017: 56). In Geshiza, Proto-West-Gyalrongic *\*ɸ-* has evolved into *ɣ-* (but *ɸ* becomes *ɣ* as an initial; e.g., Stau *ɸə* ‘head’, Geshiza *ɣuə* ‘head’), and Proto-West-Gyalrongic *\*ɣ-* has changed into *w-* (Honkasalo et al., to appear). The preinitial *x-* in the Geshiza word *xtɕænzæn* ‘carnivorous animal’ goes against this, demonstrating that some loanwords from Tibetan were borrowed into Geshiza following this sound change.

All Tibetic and Gyalrongic lects have the reflex *r-* for *dgah* ‘like’, though some Amdo lects use *h-* (Hua’s 2001 transcription).

#### 4. Huang’s hypothesis of uvular preinitials in Old Tibetan

Huang (2012) concludes that uvular preinitials were present in Old Tibetan based on the following evidence. Firstly, Huang demonstrates the possibility that uvular preinitials were inherited from Proto-Tibeto-Burman. She uses evidence from cognate sets that have uvular stops, nasals and fricatives in Ngwi, Qiangic, and Tibetan languages to bolster this claim (2012: 158–167). It is important to note at this stage that Huang is referring to both uvular initials and preinitials. She only gives examples of uvular initials in Ngwi and Qiangic and gives examples of uvular initials and preinitials in modern Tibetic lects<sup>7</sup>. Secondly, Huang shows that the majority of uvular fricative preinitials in Amdo (in particular the Skakhog/Skachu lect) correspond to Old Tibetan *g-/d-* (2012: 167–168). Thirdly, she demonstrates how Tibetan *d-* comes from *g-* (2012: 168–169). Fourthly, Huang proposes that Tibetan *g-* (which later splits into *g-* and *d-*) comes from the merger of *\*g-* and *\*ɣ-* (2012: 169–170). Her conclusion from all of this evidence is that at least some portion of uvular fricative preinitials in Tibeto-Burman come from a Proto-Tibeto-Burman voiced uvular stop preinitial *\*ɣ-* (2012: 170–172).

We can summarize Huang’s logic as follows, illustrated in Figure 1. Since many Tibeto-Burman languages have uvulars in cognate sets, and most of these languages are too widespread geographically to have borrowed uvulars from each other, Proto-Tibeto-Burman probably had a voiced uvular stop *\*ɣ*. This voiced uvular stop has voiced and voiceless stop and fricative reflexes in the Tibetan, Qiangic, and Ngwi subgroups. If the initial *\*ɣ* existed, and modern Tibetic and Gyalrongic lects have uvular reflexes for *g-/d-*, then we can postulate the preinitial *\*ɣ-* in Proto-Tibeto-Burman. Thus, for Tibetan *g-/d-*, Proto-Tibeto-Burman *\*ɣ-* and *\*g-* merged to become *g-/d-* in Old Tibetan, but became a uvular fricative preinitial (among other reflexes) in some modern Eastern Tibetan lects. By implication, Qiangic languages that have Tibetan loanwords with a uvular preinitial for *g-/d-* borrowed the uvular preinitial with the entire word.

Huang’s postulation of a uvular preinitial to Tibeto-Burman assumes the that history of uvular initials and preinitials are tied together; however, this may not be the case. The conflation of uvular initials and preinitials does not strengthen her argument, but it does not negate it either.

A central part of Huang’s proposal hinges on whether or not all modern Tibetan lects descend directly from Old Tibetan (the language that was reduced to writing in 650 AD), or if at least some of the lects (such as Amdo, Kham, and Balti) split off before the creation of Old Tibetan. Huang favors the later hypothesis, claiming that Central Tibetan (卫藏, or U-Tsang) was

<sup>7</sup> Thanks to an anonymous reviewer who pointed out that it is important to indicate when Huang is referring to uvular initials, preinitials, or both. There is an absence of uvular codas in her data.

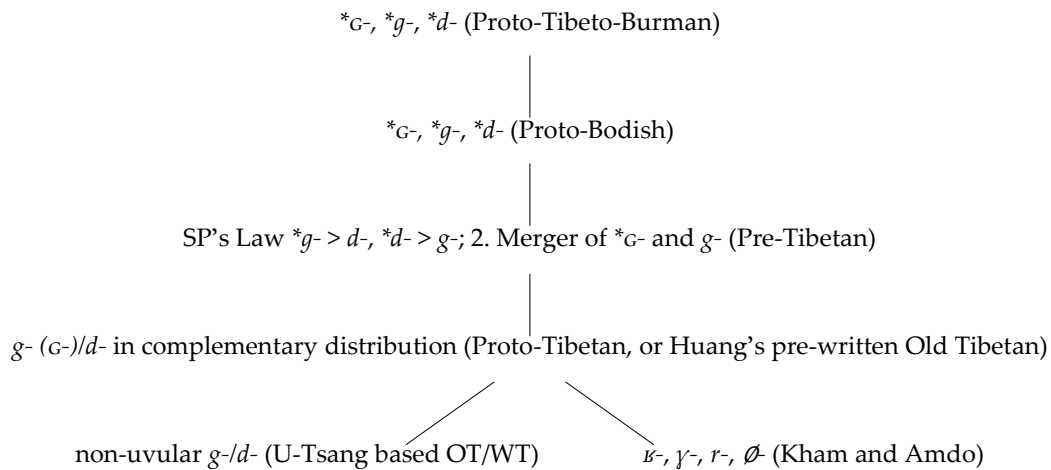


Figure 1. Huang's hypothesis

the variety used to create Old Tibetan (and thus also Written Tibetan), and that other lects split off before Old Tibetan was reduced to writing (2012: 166). Concerning uvulars in Tibetic languages, Huang states,

*Naturally, it can also be said that uvulars in [modern] Tibetan vernaculars are the reflexes of the consonants that existed in Tibetan during the pre-Tibetan period, because the literary language of Tibetan reflects the phonetic forms of Old Tibetan — the Tibetan written script has no letters that reflect uvulars. The Tibetan written script was created in the seventh century; at that time, the Tibetan language was already divided into three lects: U-Tsang, Kham, and Amdo. The Tibetan written script was created on the basis of the Central Tibetan lect, that is, the U-Tsang lect. The Tibetan written script reflects that the U-Tsang dialect did not have uvulars at that time, and some of the Amdo and Kang lects have preserved uvulars to this day, which come from the Old Tibetan language before the seventh century (Huang 2012: 166, translation mine)<sup>8</sup>.*

Of course, Huang is conflating Written Tibetan (or Classical Tibetan) to Old Tibetan (fortunately the two only differ slightly phonologically; Hill 2019b: 7), and it is not entirely clear if by 古藏语 'Old Tibetan' she actually means something like 'Pre-Tibetan'. Regardless, her claim places the split of Common Tibetan, the proto-language reconstructable from modern Tibetan lects by using the comparative method (Hill 2010: 112), before the writing of Old Tibetan, and that the written language of Tibetan descends from an old form of U-Tsang.

One way to avoid placing uvular preinitials as a distinctive uvular category in Old Tibetan is to use Huang's hypothesis, but modify it slightly. Pre-Tibetan would have uvular preinitials, but when certain lects like Amdo split from Pre-Tibetan they retained uvular preinitials, but Old Tibetan simply lost them. This modification of Huang's hypothesis is displayed in Figure 2.

In either scenario, Huang's hypothesis breaks up Tibetan (into Central, Amdo, and Kham) before Tibetan is reduced to writing. This hypothesis is an *a priori* hypothesis (unlike Hill's hypothesis discussed below in Section 5), because we know that the area for the Tibetan speaking population was small (and still not far from the Urheimat of the Yarlung Valley) and did not expand until after the writing system was created in 650 AD (Hill 2019b: 3; Bialek 2018).

<sup>8</sup> “自然也可以说，藏语土语的小舌音是藏语在前藏文时期就有的辅音的遗存，因为反映古藏语语音面貌的藏语书面语——文并无反映小舌音的字母，藏文创制于七世纪，其时藏语已分化为卫藏、康、安多三大方言，藏文是以中部藏语即卫藏方言为基础创制的，藏文反映那时卫藏方言已无小舌音，而来自七世纪前更古藏语的部分安多和康方言清小舌塞音还保留至今。” I am translating 藏文 as 'Tibetan written script' and 古藏语 as 'Old Tibetan'.

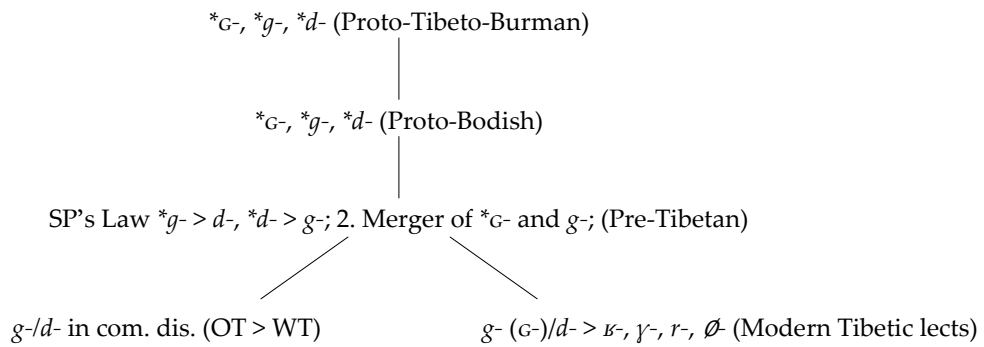


Figure 2. Huang's hypothesis modified

## 5. Hill's hypothesis

Hill claims that uvulars are not distinct in Old Tibetan (Hill 2010: 120). Based on Peiros and Starostin's Law, uvulars as a phonological category were lost through the merger of velars and uvulars in Proto-Bodish (or possibly Pre-Tibetan). At this point of in the discussion, I need to clarify that my focus is still on uvular preinitials. Peiros and Starostin's Law and Hill's claim deal with uvulars in general, that there was no uvular distinction in Proto-Bodish and Old Tibetan, respectively; but it stands to reason that if a uvular distinction did not exist in Old Tibetan, then it would not exist in either the initial or preinitial position.

Uvulars and velars were distinct in Trans-Himalayan (Hill 2019b: 32-34, 45). For our purposes, 'Trans-Himalayan' can be equated with Huang's 'Tibeto-Burman'. Hill's claim that there are no uvulars in Old Tibetan is supported by the fact that uvulars "were not orthographically distinguished from the velars" (Hill 2010: 120). The explanation that Hill gives for the presence of uvulars in modern Tibetic languages is that they are contact induced through Qiangic and/or Mongolic languages; languages in a region that he considers a "uvular-prone Sprachbund" (Hill 2010: 120).

Figure 3 is a summary of Hill's hypothesis regarding *g-/d-*<sup>9</sup>. Uvulars are lost long before Old Tibetan emerges (Peiros and Starostin's Law). Next Sa-skya Paṇḍita's Law, also occurring

<sup>9</sup> To be accurate, Hill's (2010) hypothesis is not just regarding *g-/d-*, but regarding both uvular initials and pre-initials to counter Sun (2003: 782), who reports that in the Tibetic lect Zhongu certain "instances of uvulars are of mysterious origin, as they occur in lexical items of unknown ancestry (e.g., /χeŋə/ 'hole', /q<sup>h</sup>etʂə/ 'neck'; /qeŋe/ 'hook', and in the suffixes /-qe/, /-qo/ attached to many nouns, as in /neqe/ <rna.ʔ> 'ear', /zeqo/ <zhwa.ʔ> 'hat')." Sun goes on to write, "It would be rash to attribute them indiscriminately to Qiangic substratal or areal influences, as very few of them are recognizable as Qiangic loanwords" (2003: 782-783). Sun also cites examples of uvular onsets in modern Zhongu words with clear origin in Old Tibetan. Sun suggests that this evidence (along with evidence of uvulars in other Tibetic lects) indicates the possibility that uvulars "predate standard written Tibetan." There are a couple of problems with this thinking. Firstly, how is the presence of modern uvulars more indicative of predating "standard written Tibetan" than coming after "standard written Tibetan" through natural sound change or through contact? Secondly, the presence of uvulars could still be the result of areal influences (such as from Qiangic languages) without the cited words being direct loanwords from Qiangic languages. For example, as long as some loanwords with uvulars were borrowed into Zhongu, these uvulars could effect change in other words, e.g., the ones Sun cites. Thirdly, there is the problem of the whether or not Old Tibetan had uvulars. While Sun can propose this possibility, most scholars agree that uvulars were not present in Old Tibetan. Finally, if Sun is proposing the same timing of the break up of Tibetan as Huang (i.e., before the Old Tibetan orthography was created), then once again there is the historical problem; all the evidence points to Tibetan speakers not leaving central Tibet (Yarlung, Ngaspo, and Rtsangbod) until after 650 AD (Bialek 2021: 342-351). Before 550 AD there is evidence that Pre-Tibetan speakers had already spread across the Dbus-Gtsang region (Bialek 2021: 342-351), but this is still not the eastward expansion that would come after 650 AD.

before Old Tibetan, puts  $g-/d-$  in complementary distribution in Old Tibetan, which continues on later in Written Tibetan. The uvular fricative reflexes of  $g-/d-$  found in Tibetic lects have been introduced through Qiangic (particularly Gyalrongic) or Mongolic languages. Hill places Peiros and Starostin's Law chronologically before Sa-skya Paṇḍita's Law, thus in Figure 3, Peiros and Starostin's Law is placed at the Proto-Bodish level and Sa-skya Paṇḍita's Law is placed at the Pre-Tibetan level. This is mainly for expository purposes; both laws could be placed at the same level at this stage of research, whether that be Proto-Bodish or Pre-Tibetan, as long as Peiros and Starostin's Law is placed before Sa-skya Paṇḍita's Law.

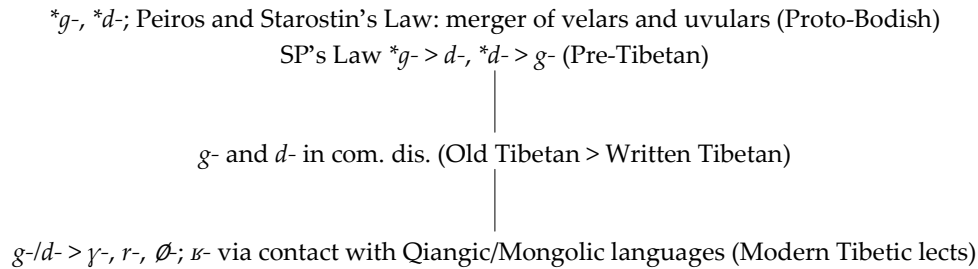


Figure 3. Hill's hypothesis

Hill's hypothesis is opposite to that of Huang's in that he claims that Common Tibetan descends from Old Tibetan (Hill 2010: 112). The Urheimat of Tibetan is the Yarlung Valley, and when the Old Tibetan orthography was created (650 AD) Tibetan speakers had not yet spread far from this region (see Bialek 2018 and 2021 for arguments in favor of this theory).

According to Hill's hypothesis, Gyalrongic languages borrowed words with  $g-/d-$  into their languages and uvularized  $g-/d-$  independently from Tibetic languages, but through contact with Gyalrongic languages Tibetic languages also uvularized  $g-/d-$ . In other words,  $g-/d-$  (perhaps as  $\gamma-/x-$ ) is substituted with  $\nu-/χ-$  in Gyalrongic languages, and then these Gyalrongic languages have influenced some Eastern Tibetic lects to also pronounce  $g-/d-$  reflexes as uvular fricatives. Although this is a reasonable hypothesis, Hill does not take into account the uvular fricative preinitials in Balti and WAT lects, which could not have developed into uvulars through contact with Mongolian and Qiangic languages as they are greatly removed geographically to the west of these languages. See Tables 1 and 2 for examples of uvular reflexes of  $g-/d-$  in Balti and WAT.

### 6. Explanation 3

'Explanation 3' builds on Hill's hypothesis to address the real issue at stake: is the uvularization of  $g-/d-$  regular or sporadic? Explanation 3 assumes that Hill is correct in the claim that there were no uvular preinitials in Old Tibetan, and claims that there is a regular sound change from  $g-/d-$  to velar/postvelar/uvular fricatives (except before velar initials, where the change is to  $r-$ ) in Amdo lects and to uvular fricatives Gyalrongic lects. For WAT lects,  $g-$  regularly changes to velar/postvelar/uvular fricatives, but  $d-$  changes to velar/postvelar/uvular or  $\xi-/r-$ . In the case of Amdo lects, the uvularization was triggered by contact with Qiangic and Mongolic, and in the case of WAT, the uvularization was triggered by contact with Indo-Aryan or Iranian languages.

However, the main emphasis of Explanation 3 is not on whether or not the change from  $g-/d-$  to uvular fricatives was contact induced, but rather that the changes were regular and not

sporadic. More research would be needed to reveal what motivated the regular change; whether contact induced or language internal reasons. In fact, since we know that the change has been regular, the issue of language contact induced change is not a highly important question, and may actually be unanswerable. If the change was completely sporadic there would be more motivation for trying to find an answer for the irregularities, e.g., language contact or inter-dialectal borrowing. The inability to point to the exact internal or external causes for a particular sound change is common in historical linguistics<sup>10</sup>.

Explanation 3 suggests that uvular preinitials in Balti and WAT could be the result of language contact with Indo-Aryan or Iranian languages. The main evidence for this explanation is found in the Balti data found in Table 1. As was mentioned above, Bielmeier et al. (2018: 50, 52) state that in Skardu (Western Balti) and Khaplu (Eastern Balti) “*x* and *ɣ* are postvelar.” The postvelar/uvular nature of these preinitials in Balti is confirmed by Caplow (2016: 208) with her transcriptions of *ʷɰzar.ba* ‘to drip’ from *ɦɰzar ba/gzar* or *ba/bzar ba*, *ʰχnot.pa* ‘to harm’ from *gnod pa*; and especially *χlat.pa* ‘brain’ from *klad.pa/glad pa*, which has a minimal contrast with *ʰlat.pa* ‘to be tired’ from *glad pa*. Note that *χlat.pa* ‘brain’ and *ʰlat.pa* ‘to be tired’ can still be considered in minimal contrast because the stress that she transcribes is deemed “marginally phonemic” (Caplow 2016: 191).

That being said, although uvulars are present in Indo-Aryan languages, for most of those languages they are not a prominent feature. For most of the the languages surrounding Balti, uvulars are only marginally phonemic, and usually they only have one unaspirated voiceless uvular stop (e.g., *q* in Hindustani, Khowar, Phasto, Shina), which does not occur in the preinitial position. However, Wakhi has *χ*, *ɸ*, and *q*, and thus it may be a candidate for inducing uvularization of preinitials in Balti. The isolate Burushaski could also be a source of uvulars, having *ɸ*, *qʰ*, and *q*<sup>11</sup>. Recent research is pointing to the heavy influence of languages like Burushaski on certain WAT lects (Kogan 2019). In the consideration of this hypothesis, it is important to bear in mind that for many of these Indo-Aryan languages, *q-* is only found in loanwords from Arabic<sup>12</sup>.

Explanation 3 proposes the following sound changes; note that the preinitial *r-* is also included, since in some situations there is a merger of *r-* with *g-/d-* in modern Tibetic lects:

1. Pre-Tibetan: *\*g-* > *ɣ-/x-*.
2. Pre-Tibetan: *\*d-* remains a dental.
3. Saska Paṇḍita’s Law: preinitials *ɣ-/x-* and *d-* become allophones of the same phoneme.
4. After the creation of Old Tibetan, but before the breakup of Tibetic languages, *d-* > [<sup>h</sup>*r-*] (allophone of *x-/ɣ-*).
5. Break-up of Tibetan (post-Old Tibetan): preinitial *g-* in Balti and Amdo became uvular from language contact or remain velar, and became uvular for Tibetan loanwords with *g-/d-* in Gyalrongic.
6. The preinitial *d-* > [<sup>h</sup>*r-*] (allophone of *x-/ɣ-*) > *ʂ*, *z*, *r* (before velars, and for WAT all environments), but became uvular or velar in other lects (except for before velars), due to language contact.
7. *r-* > *x-/ɣ-* > *χ-/ɸ-* (in many lects, also contact induced and a tendency towards merging all preinitials into velar/uvulars). *r-* > *r-* elsewhere.
8. A more recent change is *x* replacing *ʂ-*, e.g., Balti *xpe* ‘metaphor’.

<sup>10</sup> This discussion benefited from comments by an anonymous reviewer.

<sup>11</sup> I do not have enough information on the Gujarati and Domaaki languages to know whether or not they have uvulars.

<sup>12</sup> Thanks to Arnaud Fournet for reminding me of this constraint.

## 7. Explanation 4: *h* and *g-/d-* in velar and uvular free variation

If all Tibetic lects must descend from Old Tibetan, and if there are no uvular phonemes in Old Tibetan, and if uvulars in western Tibetic lects didn't develop through language contact, then perhaps there is another explanation to be explored for the uvular reflexes of *g-/d-* in Tibetic cognates and Gyalrongic loanwords. This explanation is that [ɣ-] and [ʁ-] were free variants of *g-/d-*. Voicing assimilation also occurs, producing [x-] and [χ-] before voiceless initials and [ɣ-] and [ʁ-] before voiced initials. This fourth explanation is illustrated in Figure 4.

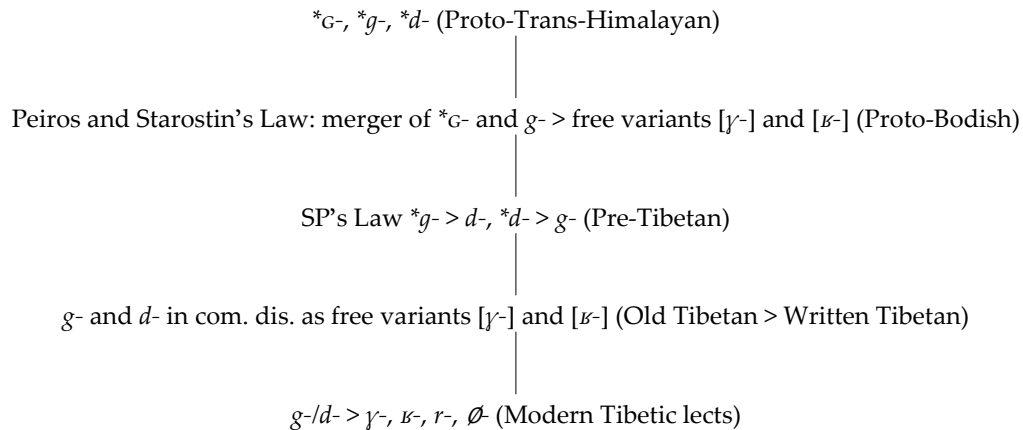


Figure 4. Explanation 4

As Figure 4 illustrates, the uvular reflexes of *g-/d-* in Tibetic descend directly from the Old Tibetan free variation of [ɣ-] and [ʁ-], which also explains why there are also velar reflexes for *g-/d-*. This explanation proposes that the reflexes of *g-/d-* found in Gyalrongic languages were borrowed in from Tibetic languages along with the entire word, albeit undergoing other changes specific to each language and lect.

This line of reasoning could also entail that *h* (ɣ) is in complementary distribution with C2 *g-* and *d-*, since *h* as [ɣ] never occurs in C2. *h-* does occur as N- in C2; the sound change from ɣ- > N- occurring before Old Tibetan (Hill 2005: 114–115, 127). Recently, Hill (2019a) has argued that *h* as a phoneme and *g-* as a preinitial/prefix have a common origin, and that *g-* has derived from *h*<sup>13</sup>. This would then imply, that while still not distinctive, there was free variation between [ɣ-] and [ʁ-] for *h*. Hill argues against *h* as a phonemic uvular consonant (Hill 2005, 2009). However, his claim still leaves open the possibility of *h* having the free variants of [ɣ-] and [ʁ-]; Hill admits that *h* “may have been articulated phonetically as a uvular” (2009: 124). While Explanation 4 is not contingent on *h* having uvular reflexes, it could be used to explain why some Tibetic languages also have uvular reflexes for *h* as a simple initial.

The problem with Explanation 4, while it is a possibility, is that it is not scientifically verifiable. There is just no way to falsify a proposal of free variation for a dead language with only written texts the sources of data<sup>14</sup>.

## 8. Conclusion

In summary, this article has explored four explanations for why there are uvular reflexes of *g-/d-* in some modern Tibetic words and Gyalrongic loanwords. Given Peiros and Starostin's

<sup>13</sup> For a counter argument see Bialek (2020: 317, fn. 134).

<sup>14</sup> I am grateful to an anonymous reviewer for this criticism.

Law, Sa-skya Paṇḍita's Law, and the data given in Section 3 (Tables 1 and 2), the most likely hypotheses are Hill's hypothesis in Section 5 (the result of contact with Mongolic and/or Qiangic languages), or Explanation 3 in Section 6, which proposes regular sound changes. Huang's hypothesis (explained in Section 4) rests too heavily on the hypothesis that Common Tibetan broke away from Old Tibetan before it was reduced to writing, a hypothesis that has been falsified. Explanation 4 (explained in Section 7), which states that *ḥ* and *g-/d-* were in velar and uvular free variation in Old Tibetan is not falsifiable, and thus cannot undergo scientific verification. All in all, Explanation 3 provides the most satisfying explanation to date. Explanation 3 provides a tentative set of sound changes and embraces the possible influence of Qiangic and Mongolic in the west and Indo-Aryan and Iranian in the east, but does not rely on contact with these non-Tibetic languages to support it.

### References

- Baxter, William H., Laurent Sagart. 2014. *Old Chinese: A new reconstruction*. Oxford University Press.
- Baxter, William H. III. 1992. *A Handbook of Old Chinese Phonology*. *Trends in Linguistics, Studies and Monographs* 64. Berlin: Mouton de Gruyter.
- Bialek, Joanna. 2018. The Proto-Tibetan clusters *sl-* and *sr-* and the periodisation of Old Tibetan. *Himalayan Linguistics* 17(2): 1–50.
- Bialek, Joanna. 2020. Old Tibetan verb morphology and semantics: An attempt at a reconstruction. *Himalayan Linguistics* 19(1): 263–346.
- Bialek, Joanna. 2021. Naming the empire: From Bod to Tibet. A philologico-historical study on the origin of the polity. *Revue d'Etudes Tibétaines* 61: 339–402.
- Bielmeier, Roland, Felix Haller, Katrin Häsler, Brigitte Huber, Marianne Volkart. 2018. *Comparative dictionary of Tibetan dialects (CDTD): Vol. 2: Verbs*. De Gruyter Mouton. doi: 10.1515/9783110554076.
- Bielmeier, Roland, Felix Haller, Katrin Häsler, Brigitte Huber, Marianne Volkart. Forthc. *Comparative dictionary of Tibetan dialects (CDTD): Vol. 1: Nouns*. De Gruyter Mouton.
- Caplow, Nancy J. 2016. Reconstructing stress in Proto-Tibetan: Evidence from Balti and Rebkong Amdo. *Linguistics of the Tibeto-Burman Area* 39(2): 180–221.
- Gates, Jesse P. 2021. *A grammar of Mazur Stau*. Paris: École des Hautes Études en Sciences Sociales dissertation.
- Hill, Nathan W. 2005. Once more on the letter ' '. *Linguistics of the Tibeto-Burman Area* 28(2): 111–141.
- Hill, Nathan W. 2009. Tibetan <*ḥ*> as a plain initial and its place in Old Tibetan phonology. *Linguistics of the Tibeto-Burman Area* 32(1): 115–140.
- Hill, Nathan W. 2010. An overview of Old Tibetan synchronic phonology. *Transactions of the Philological Society* 108(2): 110–125.
- Hill, Nathan W. 2011. An inventory of Tibetan sound laws. *Journal of the Royal Asiatic Society* 21(4): 441–457.
- Hill, Nathan W. 2014. Tibetan. In: Rochelle Lieber, Pavol Štekauer (eds.). *The Oxford Handbook of Derivational Morphology*: 620–630. Oxford University Press.
- Hill, Nathan W. 2019a. The derivation of the Tibetan present prefix *g-* from *ḥ*. *Acta Orientalia Academiae Scientiarum Hungaricae* 72(3): 325–332.
- Hill, Nathan W. 2019b. *The historical phonology of Tibetan, Burmese, and Chinese*. Cambridge University Press.
- Honkasalo, Sami. 2019. *A grammar of Eastern Geshiza: A culturally anchored description*. Helsinki: University of Helsinki dissertation.
- Honkasalo, Sami, Jesse P. Gates & Mathieu Beaudouin. To appear. *Preinitials and their development in Horpa and Tangut*. Ms.
- Hua, Kan. 2001. *Zangyu Anduo fanyan cihui* [藏语安多方言词汇 = *A vocabulary of Amdo Tibetan dialects*]. Lanzhou: Gansu minzu chubanshe (甘肃民族出版社).
- Huang, Bufan (黄布凡). 2012. Zàngmiǎnyǔ de xiǎoshéyīn [藏缅语的小舌音]. *Yǔyánxué Lùncóng* (语言学论丛) 45: 157–174.
- Jacques, Guillaume. 2008. *Jiāróngyǔ yánjiū* [嘉絨語研究 = *Study on the Rgyalrong language*]. Beijing: Minzu chubanshe.
- Jacques, Guillaume. 2014. On Coblin's law. *Studies in Chinese and Sino-Tibetan Linguistics* 2014: 155–65.

- Jacques, Guillaume. 2015–2016. *Dictionnaire Japhug-Chinois-Français, version 1.1*. Paris: Projet HimalCo. Online at: <http://himalco.huma-num.fr/>.
- Jacques, Guillaume. 2021. *A grammar of Japhug*. Berlin: Language Science Press.
- Jakobson, Roman, C. Gunnar M. Fant, Morris Halle. 1951. *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge: MIT Press.
- Kogan, Anton. 2019. On possible Dardic and Burushaski influence on some Northwestern Tibetan dialects. *Journal of Language Relationship* 17/3–4: 263–284.
- Lai, Yunfan. 2017. *Grammaire du khroskyabs de Wobzi*. Paris: Université Paris III dissertation.
- Sun, Jackson T.-S. 2003. Phonological profile of Zhongu: A new Tibetan dialect of Northern Sichuan. *Language and Linguistics* 4(4): 769–836.
- Wylie, Turrell. 1959. A standard system of Tibetan transcription. *Harvard journal of Asiatic studies* 22: 261–267.

*Джессу П. Гейтс. Увуляризация g-/d- в тибетических языках.*

В статье проанализированы четыре (относительно) взаимоисключающие гипотезы о возникновении увулярных пред-инициалей в языках тибетической группы, касающиеся в первую очередь современных увулярных отражений старотибетского g-/d-. Первая гипотеза принадлежит Хуан Буфань (Huang 2012), которая возводит увулярные пред-инициали к дотибетскому языковому состоянию, исходя из того, что распад тибетической ветви предшествует появлению письменного старотибетского языка в VII в. н.э. Второе объяснение принадлежит Н. Хиллу (2010), который утверждает, что увулярные пред-инициали не унаследованы от «дотибетского», а возникли благодаря контактам с цянскими и/или монгольскими языками. Третья гипотеза предполагает чисто фонетическое объяснение с постулированием регулярного развития g-/d- в велярные, поствелярные или увулярные фрикативные в диалектах Амдо (за исключением позиции перед велярными инициалами, где они развиваются в r-) и в увулярные фрикативные в гьярунгских диалектах. Наконец, четвертое объяснение заключается в том, что в старотибетском варианты *h* и g-/d- находились в состоянии свободного варьирования, т. е. современные увулярные пред-инициали восходят скорее к свободным аллофонам, чем к фонемам. Из этих гипотез первые три можно считать научными (поскольку они формально фальсифицируемы), четвертая (свободное варьирование) оказывается нефальсифицируемой. Анализ обстоятельств, связанных с распадом тибетического языкового единства, приводит к отказу от гипотезы Хуан, что оставляет в качестве единственно возможных объяснения Хилла и третье объяснение; при этом только последнее позволяет объяснить увуляризацию g-/d- в терминах регулярных фонетических изменений.

*Ключевые слова:* закон Сакья-пандиты; дотибетское языковое состояние; старотибетский язык; гьярунгские языки; увулярные рефлексy g-/d- в тибетских диалектах.



## \**Tubar(i)*:- a divine epithet reflected in Luwic onomastics

This article extends and discusses a study by M. Valério (2015), which deals with several groups of personal names from southern Anatolia. It is proposed here to reconstruct one onomastic stem, *\*tubar(i)*-, common to several names from different language corpora belonging to the Luwic sub-group (Luwian, Carian, Pisidian) and Greek epigraphic sources from southern Anatolia (from Caria to Cilicia). This prolific element is associated with various divine names or epithets in compounds, which suggests that it corresponded to a divine title. Its meaning can be reconstructed as “battle companion, comrade-in-arms”, originally qualifying various deities whose role was to guard their protégé on the battlefield.

**Keywords:** Luwic languages; Anatolian languages; Luwian language; Lycian A language; Lycian B language; Carian language; Pisidian language; Greek language; onomastics; theophoric names; Anatolian religion.

### 1. Introduction

For a long time, the Lycian place-name Tyberissos has been associated with the name of a Lycian hero, Τούβερις, indirectly referenced in a local myth documented by Stephanus of Byzantium.<sup>1</sup> In a complementary study to the present article, I sought to ascertain the validity of this connection — whether Τούβερις was genuinely linked to the toponym or if it was a derivative form originating from it.<sup>2</sup> During this investigation, I uncovered an onomastic stem *\*tubar(i)*- attested in numerous Luwic anthroponyms, which appear not only in Luwian-, Carian-, or Pisidian-language records, but also in Greek sources from southern Anatolia. The extensive scope of these forms precluded a comprehensive analysis within the previous article, primarily dedicated to the toponym Tyberissos. Hence, the primary goal of the current work is to delve into this deeper exploration.

The present contribution lists and analyses all the names containing the stem *\*tubar(i)*- across various corpora from southern Anatolia (§ 2). It further proposes a correlation between this onomastic stem, on the one hand, and Carian names in *°δubr-/°δýbr-* as well as *°νδυβερι-* / *°νδο/(υ)βα/ε/ηρα-* (in Greek transmission), on the other hand. While these names were dealt with separately in Valério 2015, it is argued that [d] in the latter group is an allophone of /t/ in the position after [n] or [r] (§ 3). After examining a few anthroponymic compounds where *\*tubar(i)*- appears as the first member (§ 4), the paper provides an overview of all personal names formed using this morpheme, categorizing them based on the accompanying element.

---

<sup>1</sup> This paper has been written under the auspices of the Alexander von Humboldt Fellowship, which enabled my extended academic stay at the University of Marburg. At various stages of its development, it greatly benefited from the comments of Ignasi-Xavier Adiego, Alcorac Alonso Déniz, H. Craig Melchert, Enrique Nieto Izquierdo, Diether Schürr, and Ilya Yakubovich, to all of whom I am grateful. Ilya Yakubovich deserves additional thanks for his stylistic advice and the Russian translation. However, I am solely responsible for any errors that may be found.

<sup>2</sup> For a discussion about the heronym and a full analysis of the toponym, I refer to Réveillhac forthc. a. The reader should be aware that sections 5, 6 and 7 of the present article are adaptations of passages already included in the article on Tyberissos. These paragraphs are, in fact, key points in the argumentation, which I feel justified in repeating for the sake of clarity.

It becomes apparent that, in the majority of instances, *\*tubar(i)-* is linked to a theonym or a divine epithet, possibly functioning itself as a substitute for a divine name (§ 5). The article endeavours to establish that this stem originally represented a divine title, later adopted as a divine hypostasis. This title might have initially meant “battle companion, comrade-in-arms”, referencing the protective roles of tutelary deities on the battlefield — a concept echoed in Hittite and Luwian texts concerning various deities (§ 6). The concluding section compares this interpretation with instances of Lyc. B *tuburiz*, mentioned three times on the Xanthos Stele, albeit within obscure contexts (§ 7).

## 2. The onomastic stem *tubar(i)-*

While the Lycian hero Τούβερις is only known from late Greek sources, the stem *\*tubar(i)-* is far better attested. Numerous personal names from southern Anatolia contain elements akin to the heronym conveyed by Stephanus of Byzantium. These potential candidates, found in various sources such as Hieroglyphic Luwian, Carian, Pisidian, and Greek inscriptions, are presented in Table 1. It is important to note that, at this stage, the comparison is purely formal and establishes an initial working hypothesis that needs further investigation.

Region (language of attestation)	Personal name	City	Period	Reference
Central Anatolia (Hieroglyphic Luwian)	/TuPar(i)-/ ( <i>tu-pari</i> , nom. sg.)	Hattusa (Böğazköy)	?	BoHa 19, no. 467
Egypt (Carian)	<i>D-týbr-</i>	Thebes	?6 <sup>th</sup> –5 <sup>th</sup> c. BCE	E.Th 2.1; Adiego 2007: 96
	<i>Dýbr-</i>	Thebes	?6 <sup>th</sup> –5 <sup>th</sup> c. BCE	E.Th 5.1; Adiego 2007: 97
	<i>Kśa-týbr-</i>	Thebes	?6 <sup>th</sup> –5 <sup>th</sup> c. BCE	E.Th 2.2; Adiego 2007: 96
	<i>Kud-tubr-</i>	Thebes	?6 <sup>th</sup> –5 <sup>th</sup> c. BCE	E.Th 9; Adiego 2007; 99
Caria (Greek)	Σαγγο-τβηρις	Arlisseis	?354 BCE	<i>I.Carie hautes terres</i> : 91.3
Lycia (Greek)	*Ερμα-τοβορις (gén. Ερματοβοριος)	Tlos	1 <sup>st</sup> c. BCE	TAM II: 550.21
Pisidia (Pisidian)	T[β]ερι-?	Solufar	2 <sup>nd</sup> –3 <sup>rd</sup> c. AD	N 15; Brixhe 2016: 84
Pisidia (Greek)	Περτα-τουβαρις	Termessos	Imp.	<i>I.Termessos Suppl.</i> IV: 145.1, 2-3; 146.1

Table 1. Personal names potentially formed with the stem *\*tubar(i)-*

The name /TuPar(i)-/ attested in Hattusa in hieroglyphic transmission might be a direct match of Τούβερις’s underlying name, but its isolated character hardly allows us to go any further. Another simple name, dating to the Roman period in Pisidian records, is T[β]ερι-, although unfortunately incomplete. Its restoration, based on Cilician compounds of Greek transmission in Tβερ° (see Section 4 below), gains additional support from the Carian name Σαγγο-τβηρις. To progress further, it is imperative to examine compound names, primarily in Greek transmission, which tend to be more understandable than those from Carian sources.

The vocalism of °τοβορις in Ερμα-τοβορις reflects a specific evolution found in Lycian, also observable in the toponym *Tuburehe(i)*- contrasted with its Greek equivalent Τυβερισσός. The latter is likely adapted from a more archaic dialectal form (as suggested by Réveilhac forthc. a).<sup>3</sup> While the first element of Ερμα-τοβορις is easily identifiable as the name of the Moon-god Arma (cf. *Erm̃me-neneli-*, Αρμα-πιας etc.), that of Περτα-τουβαρις is less apparent. However, the same base might be present in the Lycian place-name Παρταησσός<sup>4</sup> and the probable Lycian title *parttala-* (TL 35.14 and 16) of uncertain meaning, which does not significantly aid in its comprehension. Should one accept the equivalence between the Lycian name Παρτασις and the divine epithet CLuw. *Parattašš(i)*-, a title found twice for the Storm-god with a likely meaning of “of Impurity” (based on CLuw. *paratt(a)*- “impurity”),<sup>5</sup> the hypothetical Lycian stem-form \**partta-* might have been directly used to form compounds, similarly to Luw. /*piha-*/ “brightness, splendour” (*Piħa-muwa-*, *Piħa-nan(i)*-) compared to \**piħašša/i-* “of brightness, splendour” (*Piħašša-muwa-*, *Piħašša-walwi-*).

In the name Σαγγο-τβηρις, a syncope in the second member’s first syllable is plausible, as it is well attested in Carian for unaccented vowels (Adiego 2007: 241–242). The isolated first element, Σαγγο-, is also found in the simple personal name Σαγγως (KPN § 1369; Halicarnassus, 5<sup>th</sup> c. BCE). Although its nature and meaning require further clarification, a speculative hypothesis can be put forward: if the theonym <sup>d</sup>*Šalupiya-* (van Gessel 1998: 370), documented once in a Hittite cult inventory,<sup>6</sup> indeed corresponds to a compound (*Satzname?*), the first element being the name of the Hurro-Hittite goddess *Šalu-*, wife of Kumarbi (Laroche 1946: 57; van Gessel 1998: 370), then a similar analysis might apply to <sup>d</sup>*Šanhupiya-*, referring to a deity honoured in the northern regions of the Hittite Empire (van Gessel 1998: 371; Taracha 2009: 102). Thus, a hypothetical theonym <sup>d</sup>*Šanhū-* could be inferred from <sup>d</sup>*Šanhupiya-* and linked to Σαγγως and Σαγγο-τβηρις.

In the Carian corpus, the identification of a similar element is supported by *D-týbr-*, where the first element can be traced back to \**Ida-*. This connection is evident in other Carian and Lycian forms such as Car. *D-quq-* → *Iða-γυγος* or *D-wśol-/Id-uśol-* → *Ið-υσσωλλος*.<sup>7</sup> Similarly, in *Kśa-týbr-* one can isolate the stem *Kśa-*, which is potentially linked to the group of Luw. \*/*hassa-*/ “hearth” (inferred from CLuw. *ħaššanitt(i)*- “hearth”), Hitt. *ħašša-* “hearth” and Lyc. *xaha-* “id.”. An alternative hypothesis is to link it to the element found in the patronymic of Thales, Ἐξα-μύης,<sup>8</sup> and in Εξα-βοας/Αξα-βως.<sup>9</sup> The combination of the element Εξα- with -μυης (← *muwa-* “power, drive”)<sup>10</sup> hints at a possible association with a deity, although a de-

<sup>3</sup> This evolution proceeds from a progressive assimilation, favoured by the presence of the bilabial fricative rendered by (b). There is also a matching phenomenon of regressive assimilation *-e-u- > -u-u-*, as in the case of *urublije-* “monument” as opposed to *erublije-*, attested once and probably derived from the verb *eruwe-* “to build” (Martínez Rodríguez 2019: 221).

<sup>4</sup> This place name is documented in the *Life of Saint Nicholas of Sion* (Anrich 1913: § 57.16–17); the saint lived in Lycia under Justinian and his hagiography was written shortly after his death (Anrich 1917: 536). On the value of this Byzantine text for the study of local toponymy, see Robert 1955: 197–208.

<sup>5</sup> KUB 7.14(+) i 2-3 and KBo 29.30+ iii 6’ (CTH 694.1). See Laroche 1946: 69 (with an outdated semantic connection); van Gessel 1998: 350; Mouton and Yakubovich 2021: 32. Cf. CLuw. (≪)*paratta-* “impurity” (*vel sim.*), Melchert 1993: 167.

<sup>6</sup> About this text (CTH 510), see Cammarosano 2015: 221.

<sup>7</sup> The same element is found in Carian in forms such as *D-biks-* or *D-bkm-* (Adiego 2007: 334) and in Lycian in the simple name *Eiða* (KPN § 451-2) and in compounds such as *Ida-maxxa-* → *Iða-μαξας* or *Idazzala-* → *Eiða-σσαλας* (vs. *Zzala-* → *Σαλας* or *Ερμα-σαλας*); see Réveilhac 2018: 501-502.

<sup>8</sup> D.L. 1.22, 29: *Sud. s.v. Θαλιης*; St.Byz. *s.v. Μίλητος*.

<sup>9</sup> *Εξαβοας*: Balboursa, 2<sup>nd</sup> c. BCE; *Αξαβως*: Pisidia, 2<sup>nd</sup> c. BCE.

<sup>10</sup> On this lexeme, see now Valério 2023.

finitive etymology remains uncertain at this stage. Regarding *Kud-tubr-*, the element *Kud-* might mirror Luw. \*/huda-/ “haste, swiftness”, documented in Hittite as *Glossenkeilwort* in the form *hūda-* and as a derivational base in CLuw. *hūtarlānn(i)-* and *hutarli(ya)-*, and Lyc. *\*xuda-* “haste”, deduced from the agent noun *xddaza-* “slave” and the personal names *Xudali(je)-*, *Xudali(jē)-* and *Xudara-* (Rieken & Sasseville 2022). However, accepting this link requires assuming an evolution PLuw. *\*hu* > Car. *ku*, in addition to the established change into *qu* (e.g. *\*Tarhunt-* > *Trquδ-*, *\*huha-* “grandfather” > *quq-*), which demands further supporting evidence. If this hypothesis were correct, *Kud-* might represent an epithet that is close in meaning to Hittite *nuntar* in *nuntaraš* <sup>d</sup>LAMMA “Tutelary Deity of Swiftness” (see van Gessel 1998: 694 for the attestations).

Yet, it is crucial to examine the connection between the elements *týbr-* and *tubr-*, whose vocalism differs. To consider them as potential variants of the same morpheme, it is necessary to assume an effect of *i*-umlaut triggering the fronting of [u] > [y] in *týbr-*: PLuw. \*[tubari-] > PCar. \*[tyberi-] > Car. [tybr̥-]. Such a phenomenon is known in Carian (as well as Lycian) for other vowels, as seen in *ted-* “father” (cf. Lyc. *tedeli-*) vis-à-vis Luw. /tad(i)-/ “id.” or in *en-* “mother” (cf. Lyc. *ēneli-*) vis-à-vis Luw. /ann(i)-/ “id.” (Adiego 2007: 259). Adiego (2007: 257) suggested that the same pattern accounts for the ethnic suffix *-y/ýn-* (e.g. *kbdyn-* “Caunian”) derived from PLuw. *\*-wannali-* (cf. Luw. /-wanna/i-/ , Lyc. B *-wñneli-*, Lyc. A *-ñneli-*).<sup>11</sup> Melchert (2021: 108) gives arguments against the conditioned fronting of [u] in [y] that can be countered: 1) *ybt* need not correspond to the preterite “(he) offered” (< *\*uboto*, cf. Lyc. *ubete*), but could just as plausibly represent the present in *\*-ti* “(he) offers” (as already proposed by Adiego 2007: 259); 2) the analysis of *týn* as the infinitive of a form related to Luw. /tuwa-/ relies on the overall interpretation of the inscription C.Ha 1, which is not based on *communis opinio* (other hypotheses exist, see Adiego 2007: 283–284); 3) the final morpheme (better than “suffix”) *°eym-* in *Paraeym-* probably represents a participial stem in *\*-mmali-* subject to *i*-mutation, as suggested by the name *Οριδ-ηϋμος* (vis-à-vis *Αστ-ηϋμος*), which would justify the fronting of [u] → [y].<sup>12</sup>

Efforts have been taken to establish connections between the aforementioned names to other names from the Carian corpus and the Greek corpus in Asia Minor. The most comprehensive exploration on this subject, to the best of my knowledge, is found in Valério 2015, which builds upon earlier research concerning the Linear A *du-pu<sub>2</sub>-re* and the Hittite *t/labarna-* (Valério 2007). In this contribution, several pages are devoted to Carian names in *°DUbr-*,<sup>13</sup> Lycian names in *°vδυβερι-*, and Cilician names in *°vδο/(υ)βα/ε/ηρα-*, as well as Lycian *Tubure/i-*, *°τοβορι-* (Valério 2015: 333–339). Acknowledging the challenge posed by the interpretation of the personal names from the Carian sources, the paper eventually arrives at the conclusion that reducing all the concerned personal names to a singular and variable element *\*DUbr-* is unfeasible. Valério 2015 suggests instead the stem *\*(°)y/ýb(y)r-*, associated with the alleged element *\*(°)υβερι-* identified in Neumann 2007: 399. Following Neumann’s proposal, Valério cautiously attempts to link this virtual element with the group of Hitt. *warr-* “to help” and HLUw. /warri(ya)-/ “id.”, albeit acknowledging a major phonetic challenge: such a hypothesis would require admitting not a proto-form *\*warri-* but *\*\*uwarri-* > *\*uweri-* (Valério 2015: 335). This sug-

<sup>11</sup> The development must have been as follows: *\*-wanni* (common gender direct case) > (*\*-wni* >) *\*-uni-* > *-ýn-* (in the Egypto-Carian system)/*-yn-*. In addition to the rule [a] → [e] / (C) \_C[i], there would therefore also be [u] → [y] / (C) \_C[i] or, to subsume both, a more general phonological rule  $V_{[+back]} \rightarrow [-back] / (C) \_C[i]$ .

<sup>12</sup> It is tempting to see the element *°eym-* → *°ηϋμος* / *°ηϋμος* as a cognate of the stem found in the Lycian personal names *Uwēmeli-* and *Un-uwemeli-*: it would then represent the participle of a verb probably meaning “to look (favourably)” (CLuw. *awa-* “to see”; Lyc. A and B *uwe-* “to look favourably”; Lyd. *ow-* “id.”, according to Sasseville 2020: 289–290).

<sup>13</sup> Where D = voiced or voiceless dental, and U = [u] or [y].

gestion is also based on the assumption that an intervocalic \**w* can be strengthened into [β]. However, similar to the situation in Lycian, there is a lack of such a phonetic process in Carian.<sup>14</sup>

One argument put forward to advocate the Carian anthroponymic (and lexical) stem \*(°)y/ýb(y)r- is the existence of the simple personal name in the genitive *Ybrsś* (C.Hy 1, 7-8). Nevertheless, unlike the previous forms, it represents an *s*-stem. This anthroponym finds a plausible avatar in the name *Οβρασις* attested in Cilicia (LGPN 5b; Olba, 1st c. BCE), easily aligning with the larger group of personal names in *uppara-/o(v)πρα-/o(v)βρα-* (Houwink ten Cate 1961: 162–164; Melchert 2013: 44). It is conceivable that the [y] might have arisen due to the *i*-mutation affecting the suffix \*-*assa/i-* (as seen in *Οβρασις*). After excluding *Ybrsś* from the equation, a Carian stem \*(°)tUbr- (with U = [u] or [y]) emerges, showing similarity to *τοβορις/τουβαρις/τβηρις*, or even to the Luwian name /TuPar(i)-/, and the Pisidian name T[β]ερι-.

The case of Car. *Dýbr-* presents a more complex challenge. The initial ⟨d⟩ might have emerged from an interchange with ⟨t⟩, following the neutralization of the two underlying phonemes in the initial position, if one follows a plausible hypothesis that /d/ and /b/ may have undergone a devoicing in this position in Carian.<sup>15</sup> Another possibility is that it could be secondarily derived from the allomorph °*dybyr-* found after [n] or [r], as in *Ar-dybyr-* (see Section 3 below). A third solution could be regarding this name as a compound *D-ýbr-*, where the element \**Ida-*, previously identified in *D-týbr-*, is associated with the stem \**ybr-* found in *Ybrs-*, which I propose to connect to the *uppara-/o(v)πρα-/o(v)βρα-* group. To explain the initial [y], one might refer to rarer forms where this same stem seems to have a final *-i*, such as the names *Οβρι-μοτης/Οβρι-μωτης* (Pisidia). However, a caveat must be expressed: all the compounds in this group display the element *uppara-/o(v)πρα-/o(v)βρα-* as the first member, with the notable exception of *Κινδυ-οπρας/Κιδ-ουβρης*. While this doesn't necessarily refute this interpretation conclusively, additional unambiguous parallels would be required to reinforce it.

### 3. An allomorph after [n] and [r]

So far, I have intentionally omitted several names in \*(°)n-dUb(V)r(a/i)-,<sup>16</sup> due to their specific examination in Valério 2015. The names concerned are listed in Table 2.

This list does not include the Lycian name group *Xñtabure* → *Κενδεβορα/Κινδαβυρις/Κενεβορις*, of which *Καδοβορις* represents an additional variant with nasal reduction (see below, in this section) and vowel labialisation next to [b] or [β]. These forms mirror a compound *Xñtab-ure-* (cf. Lyc. B *xñtaba-* and Luw. /ura-/ “great”), similar to HLuw. MAGNUS.REX in structure and meaning (Yakubovich 2017: 46; Réveilhac 2018: 437).

<sup>14</sup> The outcome of PANat. \**w* in Carian is certainly not clear, but it is likely to produce [w] or [u] in certain sequences: e.g. *muo-* vis-à-vis Luw. /mawa-/ “four” (Adiego 2019a: 18–20, 25–26; followed by Simon in Bauer et al. 2022) and *kδουś*, if it reflects an earlier \**hantawa-*, as is reasonable to assume (see Simon 2022a for a detailed discussion). The data are more numerous for Lycian: PANat. \**arowā-* “freedom” > Lyc. *arawa-*, PANat. \**g<sup>w</sup>ow-* “ox” > Lyc. *wawa-*, PANat. \**duwV-* “to place” > Lyc. *tuwe-*. Pace Sasseville 2021: 177, it is methodologically problematic to rely on the Lycian B form *xñtaba-* to demonstrate the strengthening of intervocalic \**w* into [β] in Lycian B given that its exact meaning and formation are unknown. It is on the basis of this argument that I reject the connection sometimes made between the name of Tyberissos and HLuw. /tuwaris(a)-/ (Réveilhac forthc. a).

<sup>15</sup> Cautiously in Adiego (2007: 246). However, this seems to be confirmed by the group of names based on the root \**b<sup>h</sup>ēh<sub>2</sub>-* such as *Pikre-*, *Pik(a)rm-* etc.

<sup>16</sup> Here, U = rounded vowel going back to \**u*, and V = any vowel. The unorthodox notation of this sequence has a purely conventional and practical function, in order to group together different linguistic and graphic variants of the same stem.

Region (language of attestation)	Personal name	City	Period	Reference
Caria (Carian)	<i>Smδýbr-?</i> <sup>17</sup>	Halicarnassus	?	C.Ha 1; Adiego 2007: 114
	<i>Šodubr-</i>	Krya	?	C.Kr 1.2; Adiego 2007: 158
Caria (Greek)	Ζερωμεδυβερως	Caunos	?354 BCE	<i>I.Carie hautes terres</i> : 90.18 (Ξερωμεδυρος); 91.6
Lycia (Greek)	Ερωμανδοβερως (gen. Ερωμανδοβερως)	Phellos	2 <sup>nd</sup> -1 <sup>st</sup> c. BCE	SEG 53: 1696.4
	Ερωμανδυβερως	Limyra	Imp.	Wörrle 2012: 71. 4; 73.1; 3; 81.5; SEG 54: 1406.2
	Ξανδυβερως	Typallia	Imp.	Peterschen and von Luschan 1889: 149, no. 180.8; SEG 6: 624.4
	Περωπενδυβερως	Arykanda	Imp.	<i>I.Arykanda</i> : 143.2; 59; 128.5.; 136
Pisidia (Greek)	Μινδυβηρας (Μινδυβηρας)	Selge	1 <sup>st</sup> c. AD	Bean and Mitford 1970: 201a.4
Isauria	*Ξανδοβηρας (gen. Ξανδοβηρου)	Iotape	Imp.	CIG: 4413d.2
	Ταρκυνδβερωρας	Isaura Nea	Imp.	Sterrett 1888: 181.4
Cilicia	Ινδοβαρας	Kestros	1 <sup>st</sup> -2 <sup>nd</sup> c. AD	Bean and Mitford 1970: 167.2; 168.2; 182
	Ινδοβηρας	Lamos	Imp.	Bean and Mitford 1970: 193.2
	*Μανδουβιρας (gen. Μανδουβιρου) <sup>18</sup>	Diokais-areia/Olba	Byz.	<i>I.Westkilikien Rep</i> : 33
	*Μοτονδοβερας (gen. Μοτονδοβερου)	Olba	1 <sup>st</sup> c. BCE	<i>I.Cilicie</i> : 11 A II.39
	*Ρωνδοβερωρας (gén. Ρωνδοβερωρου)	Olba	1 <sup>st</sup> c. BCE	<i>I.Cilicie</i> : 11 A II.37
	Ρωνδβερωρας	Olba	2 <sup>nd</sup> c. BCE	Heberdey and Wilhelm 1896: 155 A II.34; 69; 84; 100

Table 2. Personal names in °(n)-dUb(V)r(a/i)-

<sup>17</sup> In the form *Smδýbrs*, which appears at the beginning of the inscription, the final -s can theoretically belong to the stem of the name (cf. *Ybrsś* in C.Hy 1a) or represent an inflectional suffix. The analysis depends closely on the interpretation of the inscription in which the form occurs and on the value of the s-case itself. On these thorny issues, see especially Adiego 2007: 283–284 and Melchert 2021: 111–112.

<sup>18</sup> Nominative *Μανδουβιος* in LGPN 5b. But in view of the other occurrences in the region, one would rather expect a final -ας.

After refuting the existence of a \*DUbr- stem in Carian,<sup>19</sup> Valério (2015: 335–338) offered an alternative account. He suggested that the alleged stem  $^{\circ}\nu\delta\upsilon\beta\epsilon\rho\iota\text{-}/^{\circ}\nu\delta\omicron\beta\eta\rho\alpha\text{-}$  and its variants should be interpreted as a reflection of Lyc. \**ñtuweri-*, which is inferred from the hapax *ñtuweriha*, found on one of the faces of the Xanthos Pillar (TL 44b.57): *se=dewē : zxxaza : se ñtuweriha : ade* “he made a gift to the fighters and *ñt*.” The Lycian term is analysed as an adjective derived from a substantive \**ñtuweri-*, which consists of the prefix *ñte-* “in, within” and \**weri-* < \**warri-* “help” (cf. Hitt. *warr-* “to help”, HLuw. /*warriya-*/), with *e* > *u* in contact with *w*. This etymology finds support in the Hittite expression *anda warišša-* “to come to one’s aid”. The concept of “assistance, help” (*vel sim.*) that emerges from this etymology fits well the context of the aforementioned Lycian passage and would not be unjustified in theophoric personal names such as  $\text{E}\rho\mu\alpha\nu\delta\upsilon\beta\epsilon\rho\iota\varsigma/\text{E}\rho\mu\alpha\nu\delta\omicron\beta\epsilon\rho\iota\varsigma$ , interpreted as possessive compounds “Who has the assistance of X”, where X represents a theonym like the Moon-god Arma.<sup>20</sup> Without disputing the meaning of the Lycian lexeme, the details of its formation are of major importance for the present purpose. *Ñtuweriha* is an action noun derived from the verb \**ñtuwerih-*, which is structurally comparable to Hitt. (*anda*) *warišša-* (Sasseville 2023; contra Neumann 2007: 253–254): on this basis, it would be necessary to reconstruct *ñtuweriha-* (Lyc. A), \**ñtuwerisa-* (Lyc. B), or even \**ñtuwarisa-* (PLyc.) stems, which can hardly be aligned to the onomastic stem. Additionally, a significant phonetic challenge arises: the glide [w] is never rendered by a ⟨β⟩ in Greek onomastics, but rather surfaces as ⟨ν⟩, as in *Hln̄midewe-* →  $\text{E}\lambda\mu\iota\delta\alpha\nu\alpha$  and *Tewinaza-* →  $\text{T}\epsilon\upsilon\iota\nu\alpha\sigma\alpha\varsigma$ , or as a hiatus after *u*, as in *Xuwata-* →  $\text{K}\omicron\alpha\tau\alpha$  (Réveilhac 2018: 360–361; Adiego 2020b: 51). Despite early fricativisation of Anatolian Greek [b] into [β] and then [v], potentially leading to the use of ⟨β⟩ to render the glide, it is implausible that Greek adaptations of *ñtuweriha-* would *exclusively* display a ⟨β⟩. As a result, the association between the alleged stem  $^{\circ}\nu\delta\upsilon\beta\epsilon\rho\iota\text{-}/^{\circ}\nu\delta\omicron\beta\eta\rho\alpha\text{-}$  etc. and Lyc. *ñtuweriha-* must be dismissed.

To elucidate these personal names, a more economical explanation might be proposed, suggesting their incorporation into the group of compounds having \**tubar(i)-* as their second member.<sup>21</sup> In this scenario, the initial dental may have undergone voicing assimilation following the final nasal of the first member, a phenomenon observed, e.g., in  $\text{T}\alpha\rho\kappa\upsilon\mu\text{-}\beta\iota\gamma\rho\epsilon\mu\iota\varsigma$  and  $\text{P}\omega(\mu)\text{-}\beta\iota\gamma\rho\epsilon\mu\iota\varsigma$  vs.  $\text{P}\iota\gamma\rho\alpha\mu\iota\varsigma$ . Let us begin by interpreting the simpler compounds:

- $\text{E}\rho\mu\alpha\nu\text{-}\delta\omicron\beta\epsilon\rho\iota\varsigma/\text{E}\rho\mu\alpha\nu\text{-}\delta\upsilon\beta\epsilon\rho\iota\varsigma$  ← /Arman(t)-/ (Moon-god),<sup>22</sup> thus a variant of  $\text{E}\rho\mu\alpha\text{-}\tau\omicron\beta\omicron\rho\iota\varsigma$ ;
- $\text{T}\alpha\rho\kappa\upsilon\nu\text{-}\delta\beta\epsilon\rho\rho\alpha\varsigma$  ← /Tarhun(t)-/ (Storm-god), with syncope of the initial vowel in the second member (as seen in  $\text{P}\omega\nu\delta\beta\epsilon\rho\rho\alpha\varsigma$  below);
- $\text{P}\omega\nu\text{-}\delta\omicron\beta\epsilon\rho\rho\alpha\varsigma/\text{P}\omega\nu\text{-}\delta\beta\epsilon\rho\rho\alpha\varsigma$  ← /Run(ta)-/ (Protective Stag-god).

The devil’s advocate could argue that the proposed segmentation in these theophoric compounds is unverifiable, given that all three theonymic stems end in *-nt-*. The case of  $\text{P}\epsilon\rho\pi\epsilon\nu\delta\upsilon\beta\epsilon\rho\iota\varsigma$  might be, however, less ambiguous, if the first member indeed reflects Lyc. (A and B) *perepñ* “further(more)” (perhaps to be understood as divine epithet), as in  $\text{P}\epsilon\rho\pi\epsilon\nu\eta\nu\iota\varsigma$  (*vis-à-vis*  $\text{N}\eta\nu\iota\varsigma$  ← *nēne/i-*) and  $\text{P}\epsilon\rho\pi\epsilon\nu\nu\nu\epsilon\mu\iota\varsigma$  (*vis-à-vis*  $\text{O}\nu\epsilon\mu\iota\varsigma$ ).

This hypothesis could also shed light on the formation of other anthroponymic compounds, although the identification of their first member is not always straightforward. As for

<sup>19</sup> See already the embarrassment of Schürr (2013: 28) to account for the formation of several Carian names.

<sup>20</sup> This was the analysis adopted in Réveilhac 2018: 490–491. Compare with the numerous Anatolian compounds of the type *X-muwa-* “Who has the might of X” (Houwink ten Cate 1961: 166–169; Laroche 1966: 322–324; Melchert 2013: 33–34).

<sup>21</sup> The Cilician data in *-ας* suggest that \**tubara-* replaced \**tubar(i)-* in Cilicia.

<sup>22</sup> Another possibility, suggested to me by Ilya Yakubovich, would be to see in the first member the cognate of Car. *armon* “interpreter”, then perhaps used as a divine epithet.

Μινδυβηρας and its likely variant Μανδουβιας, the first member can receive two interpretations. One interpretation would be to regard Μιν- and Μαν- as variants of Μην-, reflecting the name of the Moon-god Men, with one variant displaying iotacism — a common feature in the imperial period in Anatolia —<sup>23</sup> and the other retaining the older vocalism (Sittig 1911: 153–155; Masson 1990: 327–328). The worship of Men is well-documented in certain Anatolian regions, particularly in Roman Pisidia, where Μινδυβηρας<sup>24</sup> is attested. If this interpretation holds, Μινδυβηρας/Μανδουβιας could stem from a secondary adaptation of the group of Αρματοβορις/Αρμανδοβορις/Ερμανδυβερις, substituting Men’s name for that of Arma, who is also a Moon-god. An alternate analysis would be to see the first member as a \**mVnd-* element, supported by other anthroponyms: Lyc. *Mñtete(i)-*, Μενδις (LGPN 5b; Lycia, Olympus, 3rd c. AD)<sup>25</sup>, Μίνδης (see KPN § 920)<sup>26</sup>, Μενδεσις (LGPN 5b; Cilicia, Kestros, 1st-2nd c. AD), or Μινδριων (LGPN 5b; Milyas, Soklai, 4th-5th c. AD). The stem of these names may be tentatively related to Lyc. *miñte/i-* (cf. Gk. μινδις, μενδιτης), denoting a local authority sometimes linked to deities (cf. *mahāi miñtehi* “of the gods of the *miñte/i-*”; TL 58.5), or alternatively to Lyc. *mēte-* “damage, harm” (cf. Hitt.-Luw. *mantallali-* “poisonous, slanderous”, Lyd. *mētr(i)-/mētl(i)-* “harm”).

The first element of Ξανδυβερις and Ξανδοβηρας has been associated by Neumann (2007: 144–145) with Lyc. B *xzzāta-* (TL 44d, 44), potentially underlying the Greek name of the city of Ξάνθος (otherwise called *arñna-* in Lycian A), although this remains uncertain. The same stem appears in the Pamphylian anthroponym Ξανδαροιζας (LGPN 5c; Kibyra Mikra, imp.), which can be segmented as Ξανδα-ροιζας thanks to the comparison with Ποιζας (Cilicia)/Ποιζος (Lycia)/Πωιζις (Pisidia).<sup>27</sup> Based on this, it is tempting to place the morpheme boundary between Ξανδ-, on the one hand, and -υβερις/-οβηρας, on the other. However, factoring in Ερμαν-δοβερις/Ερμαν-δυβερις vis-à-vis Ερμανδας, Ταρκυν-δβερρας vis-à-vis Ταρκονδαιος, and Ρων-δοβερρας/Ρων-δβερρας vis-à-vis Ρωνδας, it is conceivable that the first element in Ξαν-δυβερις/Ξαν-δοβηρας vis-à-vis Ξανδα-ροιζας shows variation \*/*ksan(t)-*/ ~ \*/*ksanta-*/. This stem might tentatively correspond to HLuw. /*hasa-*/ “abundance”: it would reflect a derived possessive adjective \**has-ant(i)-* “having abundance”, to be linked to the Palaic adjective *hašāwant-*, used as a divine epithet of the Hattian deity Kamama (van Gessel 1998: 216),<sup>28</sup> who belongs to the group of the protective ʿKAL deities (Laroche 1973: 85-86). Also belonging to the ʿKAL group, the Luwian god Runtiya is also presumably hidden behind Ποιζας: in the compound Ξανδα-ροιζας, the Luwian protective deity too appears to be associated with abundance.

The names from Cilicia Ινδοβαρας and Ινδοβηρας contain a first element that can be linked to several simple names from the same region, namely Ινδας, Ινδης, or even Ινδους (see LGPN 5b: 213–214). These names have a root that resembles the Luwian adverb *inta* “below”.<sup>29</sup> This adverb is known in Cuneiform Luwian as an epithet of the Sun-god Tiwad, invoked in association with his celestial counterpart [*šarr*]i *Tiwat-*, meaning “Tiwad above”

<sup>23</sup> See also the Pamphylian onomastic element Μεινα-, Μινα- representing the name of Men (Brixhe 1976: 29 and 136).

<sup>24</sup> On the cult of Men in Pisidia, see especially Labarre 2010. Note that the divine name may have been converted into a personal name in Pisidian, as indicated by several attestations of Μηνι- in the Pisidian corpus.

<sup>25</sup> Despite the doubts expressed by Zgusta (KPN: 317, fn. 219), there is no need to resort to a Thracian interpretation of this name (as a variant of the well-known Βενδίς).

<sup>26</sup> Procopius of Caesarea, *De bellis*, 7.36.26, as the name of an Isaurian.

<sup>27</sup> Perhaps Ξανδα, attested in an Athenian inscription (IG II<sup>2</sup> 7926, 2<sup>nd</sup> c. BCE) as the name of a Galatian from Ancyra, should also be added, but this remains uncertain (KPN § 1060).

<sup>28</sup> See Sasseville 2022, for a discussion about the Palaic word, with references.

<sup>29</sup> I thank Ilya Yakubovich, who suggested this connection.



(KUB 32.10+ obv. 12'; Yakubovich and Mouton 2023, vol. 1: 209). This duality between a solar deity “above” and a solar deity “below” is likely explained in KBo 29.3+ ii 25–26, where the Sun-god (of Heaven) is associated with the living, while the Sun-goddess of the Earth is associated with the dead (Mouton 2023: 128). The root *Ivδ-* likely reflects the epithet “below”, which has replaced the divine onomastic sequence */inda Tiwad/*. Therefore, *Ivδοβαρας/ Ivδοβηρας* is based on a formula meaning “(Tiwad) Below *tubar(i)*”.

In the case of *Ζερωμεδυβερος*, one has to assume the disappearance of the preconsonantal nasal that triggered the voicing of [t] into [d]. Such an evolution is well documented, particularly in Greek adaptations of Anatolian names, such as *Κεδηβης* vis-à-vis *Κενδηβης/Κενδεβης* (cf. Lyc. B *xñtaba-*). The plausible first element of this compound might be traced in the name *Ζερωμουνδης* (Lycia),<sup>30</sup> which may in turn result from the anthroponymic conversion of *\*z(e)rm̃m̃nt(i)-* < *\*zarmant(i)-*. As per Neumann (1976: 139–141), such a form could be based on the lexeme *\*zarma-* < *zalma-* “protection”,<sup>31</sup> identified in the Cilician names *Τροκο-ζαρμας*, *Ρω-ζαρμας* and *Ια-ζαρμας* (cf. Luwian personal names of the type */Tarhu-zalma-/*).<sup>32</sup> *\*Z(e)rm̃m̃nt(i)-* might therefore represent a possessive adjective in *-nt(i)-* meaning “having protection”, unless it is the former participle of a verbal stem.<sup>33</sup> In the latter scenario, considering the close association between the element *\*zarma-* and divine names, *\*z(e)rm̃m̃nt(i)-* might correspond to a divine epithet meaning “protecting”.

Carian documents also provide two names of more ambiguous interpretation: *Šodubr-* and *Smđýbr-* (if this interpretation is favored over that of a stem *Smđýbrs-*),<sup>34</sup> where the grapheme ⟨δ⟩ indicates the biphonemic sequence */nd/*, like in the Storm-god’s name *Trqδ/Trqude* (cf. Luw. */Tarhunt-/*, Lyc. A *Trqqñt-*; see Adiego 2020a: 105). *Šodubr-* might potentially be a compound with the name of the War-god *Sanda*, widespread in southern Anatolia,<sup>35</sup> as first element. A specific relative chronology of sound changes is, however, required to justify the initial fricative [š] and the vocalism of the first member: *\*/Sanda tubar(i)-/* “*Sanda tubar(i)*” > *\*/Sñd(a) tubr̥-/* (syncope and apocope) > *\*/Šñ(d)tubr̥-/* (palatalization of the sibilant)<sup>36</sup> > *\*/Šñdubr̥-/* [Šñdubr̥-] > *\*/Šondubr̥-/* spelled *Šodubr-*. The final stage would correspond to the phonologisation of the “excrecent” vowel derived from the syllabic nasal, a process recently highlighted by Adiego (2020a: 115–118), elucidating variants such as *Pnuśoλ/Punwśoλś* (→ Πονυσωλλος).

Regarding *Smđýbr-*, it has previously been linked to the name *Σεμενδησις* found at *Istlada* (Lycia), rightly interpreted as derived from the divine name *Σομενδης/Σουμενδης* (Schürr 2013: 28). The latter is documented on various votive stelae from Central Lycia, indicating one

<sup>30</sup> Pace Valério (2015: 335), who regards the pair *Ζερωμεδυβερος/Ζερωμουνδης* as a hint that there was a stem *Ζερω-*, possibly extended into *Ζερωεδ-*, without however succeeding in offering a comprehensive explanation. Adiego (2019b: 147, fn. 2) prefers not to include the name *Ζερωμουνδης* in his study of names in *-zarma-*, *-sarma-*, *-ζαρμας*.

<sup>31</sup> The meaning of *zalma-* is inferred from the graphic alternation attested for the personal name *Iyarrazalma-/Iyarra-PAP-* (Laroche 1966: nos. 430 and 433), since the sumerogram PAP/PAB is used in Hittite texts as the equivalent of *paḥš-* “to protect”.

<sup>32</sup> Nevertheless, Adiego (2019b: 157–159) prefers to remain cautious about the origin of the stem *-ζαρμας*, which could also reflect the Luwian contracted variant *Sarma-* of the divine name *Sarruma-*, following the analysis of Houwink ten Cate 1961: 134–136).

<sup>33</sup> Indeed, it cannot be ruled out that the element *zalma-* is verbal, in which case all the names *X-zalma-* with *X* = divine name would be *Satznamen* meaning “*X*, protect (him)!”. Such an explanation would then shed light on forms such as HLuw. */Tarhu-zalma-/*, *Τροκο-ζαρμας*, or *Ρω-ζαρμας*, whose first element would correspond to the vocatives */Tarhu(n)/* and */Ru(n)/* (see Adiego 2022: 81–82).

<sup>34</sup> Cf. fn. 17 above.

<sup>35</sup> On the place of this divine name in southern Anatolian onomastics, see Houwink ten Cate 1961: 136–137.

<sup>36</sup> Compare with *šar-* < *\*sri-* (Adiego 2007: 261).

or several places of worship dedicated to this deity. It has been suggested elsewhere (Réveilhac forthc. c) that the divine name Σομενδης/Σουμενδης should be seen as an ancient epiclesis corresponding to CLuw. *zammant(i)-/zamnant(i)-*, also used as a *Glossenkeilwort* in Hittite contexts. Literally meaning “having the *zam(ma)n-*”, this adjective, mostly employed in ritual contexts, could convey the active meaning of “damaging” or the passive meaning of “damaged”, both possibly referring to the supernatural impact on the patient in purification rituals. The name Σεμενδησις likely denotes a possessive derivative in *\*-assa/i-*, implying “Belonging to Soumendis / the *zammant(i)-* deity”,<sup>37</sup> while *Smδύβρ-* might reflect the sequence *\*Zammant(i)-tubar(i)-*.

Nevertheless, it is essential to exercise caution in the last two cases. If a second member *°y/ybr-* is acknowledged in Car. *Dýbr-* (see Section 2 above), then it could also potentially be found in *Šodubr-* and *Smδύβρ-*. These two names, therefore, remain equivocal, although, as noted earlier, the prevalence of compounds with *uppara-/o(v)πρα-/o(v)βρα-* as the first member makes the interpretation with *\*tubar(i)-* more plausible in my perspective.

Moreover, the voicing of [t] into [d] extends beyond the position after [n(d)] and likely encompasses the position after [r], as observed in Lycian, evident in examples like *Natr-bbijēmeli-* (→ Ἀπολλό-δοτος) compared to *Mahane-pi[jēme/i]-* → Μανα-πιμ[ι]ς and *pije-* “to give” (Réveilhac 2021: 566). This pattern appears in variants of the same name, as depicted in Table 3 below.

Egypt (Carian)	<i>Ardybyr-</i>	Memphis	?6 <sup>th</sup> –5 <sup>th</sup> c. BCE	E.Me 52; Adiego 2007: 72
Caria (Greek)	Αρδυβερος	Halicarnassus	5 <sup>th</sup> –4 <sup>th</sup> c. BCE	SEG 43: 713C.2; 9; 36

Table 3. Two variants of a personal name in *°r-dybVrali-*

In the Carian name *Ardybyr-* → Αρδυβερος, the first member can also be found in *A[r]bikarm-* vis-à-vis *Pik(a)rm-*. This element is ambiguous and might correspond to: 1) an adverbial element meaning “high” (cf. Luw. /aru-);<sup>38</sup> 2) the Carian appellative *eri-* “companion” and Lycian *°eri-* “id. (?)” (cf. Hitt. *arā-*);<sup>39</sup> 3) a divine name, reflecting *°Ara/i-* found in the Hittite documentation (KBo 52.56 ii 14; cf. van Gessel 1998: 45).

#### 4. The stem *\*tubar(i)-* as first member in compounds

Houwink ten Cate (1961: 159) proposed an interpretation of Τβερασητας, Τβερημωσις, Ρωνδβεροας, and Ταρκυνδβεροας, presenting them as compound names containing an element akin to Luw. *\*tapara-* “ruler”. Nevertheless, Valério (2015: 342) rightfully dismissed this interpretation, based on the comparison with Luw. *\*uppara-* which became *O(v)πρα-* in Hellenistic forms, implying that one would expect *\*tupara-* > *\*\*TVβρα-*. Valério suggested viewing the first member of Τβερασητας and Τβερημωσις as a reflection of a Proto-Anatolian element

<sup>37</sup> Like in the Lyc. B personal name *Masasa/i-* “Belonging to the god(s)”.

<sup>38</sup> See also Houwink ten Cate (1961: 172–173), where the adverbial element *er(i)-* is identified in several names, although some of them actually display another adverbial element, *hri*, such as Αρμοας/Ερμοας, reflecting *Hrm̄muwe-* and having Σεριμυας as a Lycian B avatar (Réveilhac 2018: 486).

<sup>39</sup> For the Carian form, see Simon 2016–2017. The Lycian *°eri-*, which is not assured, could be the second member of a possible compound *tidereli-* “collacteus”, with in the first member *\*tide-* “teat” (see Melchert 2004: 66; and Neumann 2007: 360–361, with discussion and references). Adiego (2022: 83) suggests interpreting the Cilician names Ταρκυαρις and Τροκοαρις as *Satznamen* reflecting a clause with a vocative *\*Tarhu, aris* “Tarhunt, (be) his comrade!”, but it could also be *\*Tarhu, ari* “Tarhunt, raise (him)!”.

\**d(u)warā-* “long-lasting” (< PIE \**dwēh<sub>2</sub>-ró-* “long”) or a Luwian adjective \**tuwala-* (cf. Hitt. *tuwala-* “distant”) > \**t(u)wara-* (Valério 2015: 344–345).<sup>40</sup> While this proposal is ingenious, its foundation is tenuous due to the absence of additional lexical or anthroponymic data. Moreover, the semantic aspect is not as apparent as suggested by the author. The proposed translations “Man of long-lasting might” and “Long-lasting man” for Τβερημωσις and Τβηρασητας respectively are not conclusive, since compounds with a second member °*muwa-* and °*ασητας*<sup>41</sup> are usually theophoric, which suggests the need for a divine element (name or epithet) as the first element.

The association of two Carian forms with this small group of names from Cilicia, *Tbridbδ-* in Memphis (E.Me 42, 2; Adiego 2007: 66), and \*Τοβορορας (gen. Τοβοροροῶ)<sup>42</sup> from a Greek inscription in Halicarnassus (5<sup>th</sup>–4<sup>th</sup> c. BCE), suggests that the stem \**tubar(i)-* might serve as the first member of a compound. Although the integration of the first is provisional until the second virtual member is positively identified,<sup>43</sup> \*Τοβορορας can be interpreted as derived from \**Tubur-ura/i-* meaning “Great(est) *tubar(i)-*” *vel sim.* This formation finds parallels,<sup>44</sup> indicating the plausibility of the stem \**tubar(i)-* serving as a first member of a compound.

If one assumes a syncope of the vowel in the first syllable of the stem,<sup>45</sup> as observed in Σαγγο-τβηρις, Ρων-δβερρας (next to Ρων-δοβερρας), and Pis. Τ[β]ερι- (see Section 2 above), then the stem Τβαρα-/Τβερη-/Τβερα- documented in Cilicia (and possibly in Car. *Tbr*<sup>o</sup>) might be a variant of *tubar(i)-*. This phenomenon, quite widespread in names from Cilicia (as seen in Ρων-δβερρας), could either be a pan-Luwian or a regional (post-Luwian) feature. Consequently, the interpretation of these anthroponyms could be as follows:

- Τβαρα-μοτας < \*/Tubara-muwatta-/ “*Tubar(i)-* (of) Conquest” or “Having the Might of *tubar(i)-*” (cf. HLuw. /muwatta-/);
- Τβερη-μωσις < \*/Tubara-muwassa/i-/ “Mighty *tubar(i)-*” or “*Tubar(i)-* of Might” (cf. HLuw. /muwassa/i-/);
- Τβερα-ασητας < \*/Tubara-assatta-/ “*Tubar(i)-* (of) Peace” or “Having the peace of *Tubar(i)-*” (Luw. in Hittite transmission \**aššatta-*).<sup>46</sup>

## 5. Synthesis on the compounds with \**tubar(i)-*

So far, I have endeavoured to demonstrate the existence of the stem \**tubar(i)-* in southern Anatolian onomastics. This stem, which has sometimes undergone phonetic variations, clearly appears in compounds as a second member after a first member ending in a vowel (e.g. Ερμα-

<sup>40</sup> The first hypothesis is suggested with caution, however, since “no direct descendant of P.-Anat. \**dwarā-* is attested among the Anatolian languages”.

<sup>41</sup> Contrary to what has long been claimed, names in °*σητας* do not correspond to Luwian compounds in /°*zid(i)-*/, but are actually in °*ασητας*. See fn. 46 below.

<sup>42</sup> Better than \*Τοβορορος (SEG 43.713A; cf. LGPN 5b).

<sup>43</sup> The second member could be related to the Lycian place-name Idebessos (Ιδεβησσος, Ιδβησσος, Έδεβησσός; see KON § 361).

<sup>44</sup> Cf. Lyc. *Xñtabure-* → Κενδεβορα/Κινδαβυρις/Κενεβορις. See Section 3 above.

<sup>45</sup> For syncope in the languages of the Luwic subgroup, see Melchert 1994: 275–276 (Luwian, more limited than Lycian) and 318–321 (Lycian); Hajnal 1995: 175–188 (Lycian); Adiego 2007: 241–242 (Carian).

<sup>46</sup> *Contra* Houwink ten Cate (1961:171–172), the names in °*σητας* (Cilicia) or °*σατης* (Pisidia) from Greek sources cannot correspond to Luwian compounds with /*zid(i)-*/ “man” as second members for phonetic reasons (nature of the dental, vocalism of the stem). On this crucial issue, see now Réveilhac *forthc.* b, where a segmentation °*ασητας*/°*ασατης*/°*ασατας* is proposed, and this element is analysed as the abstract noun “peace” (Luw. \**aššatta-*, Lyc. B *asata-*, Lyc. A *ahata-*) used as a divine attribute.

τοβορις). I have argued that this same stem displays an allomorph with initial [d] when it comes after a first member ending in [n] (e.g. Ερμαν-δοβερις/Ερμαν-δυβερις) or [r] (e.g. *Ar-dybyr-/Aρ-δυβερος*). I have also suggested identifying *\*tubar(i)-* as the first member of (rarer) compounds, in full form in Τοβορο-ορας (Caria), but with syncope in a few Cilician names (e.g. Τβαρα-μοτας). If one accepts this argumentation, it is possible to make use combinatorial analysis in order to interpret the meaning of the stem under discussion. Below one finds a summary of the compounds with *\*tubar(i)-*, ordered according to the associated stem type.

Personal name	Divine name	Deity's function
<i>Ar-dybyr-</i> Aρ-δυβερος	Ara/i-? (or <i>*ar-</i> “super”)	?
<i>D-týbr-</i>	Ida-	?
Ερμα-τοβορις Ερμαν-δοβερις Ερμαν-δυβερις	Arma(nt)-	Moon-god
Ρων-δοβερορας Ρων-δβερορας	Runt(iy)a-	Protective Stag-god
Σαγγο-τρηρις	<i>*Sanhu-</i> (cf. <sup>d</sup> <i>Šanhupiya-</i> )?	?
? <i>Šodubr-</i>	Sanda-	War-god
Ταρκυν-δβερορας	Tarhu(nt)-	Storm-god

Table 4. Type 1, X-*tubar(i)-*, where X = divine name → “(Deity) *tubar(i)-*”

Personal name	Divine epithet	Epithet's class and meaning	Parallels
Ζερμε-δυβερος	<i>*zarmant(i)-?</i>	Participle/adjective: “protecting”?	Ζερμουνδης
Iv-δοβαρας Iv-δοβηρας	Luw. /inda/	Adverb: “below”	Ivδας Ivδης Ivδους
<i>Kśa-týbr-</i>	?	?	Εξα-μύης Εξα-βοας?
<i>Kud-tubr-</i>	Luw. <i>*/huda-/</i> , Lyc. <i>*xuda-?</i>	Substantive: “haste, swiftness”	Κουδεις?
Μαν-δουβιρας Μιν-δυβηρας	Lyc. <i>miñteli-</i> or <i>mēte-?</i>	Substantive: (official institution) ou “harm, damage”	Lyc. <i>Mñtete(i)-</i> Μενδης Μίνδης Μενδεις
Μοτον-δοβερας	?	?	Μουτνοιτης? (vs. Μο[κ]-οιτης)
Ξαν-δυβερις Ξαν-δοβηρας	<i>*hasant(i)-?</i>	Adjective: “having abundance”?	Ξανδα-ροιζας
Περπεν-δυβερις	Lyc. <i>perepñ</i>	Adverb: “further”	Περπενηνης
Περτα-τουβαρις	CLuw. <i>paratt(a)-</i>	Substantive: “impurity”	Παρτασις <i>Parattaššali-</i> Place-name Παρταησός
? <i>Smđýbr-</i>	CLuw. <i>zammant(i)-</i>	Adjective: “damaging; damaged”	Divine name Σομενδης/Σουμενδης Σεμενδησις

Table 5. Type 2, X-*tubar(i)-*, where X = divine epithet → “([Deity +] divine epithet) *tubar(i)-*”

Personal name	Divine epithet	Epithet's class and meaning	Parallels
Τβερ-ασητας	Luw. * <i>aššatta</i> -, Lyc. B <i>asata</i> -, Lyc. A <i>ahata</i> -	Substantive: “peace”	Μιρ-ασητας Οβρ-ασητας/ Οπρ-ασατης
<i>Tbr-idbd-</i>	?	?	Place-name Ιδεβησσος?
Τβαρα-μουτας	* <i>muwatta</i> -	Substantive: “conquest”	Ουρα-μουτας Αρσι-μουτας Οβρα-μουτας
Τβερη-μωσις	* <i>muwassali</i> -	Adjective: “mighty” or “of might”	Πιγρα-μουσις Ουπρα-μωσις
Τοβορ-ορας	* <i>urali</i> -	Adjective: “great”	<i>Tiwata-ura</i> -

Table 6. Type 3, *tubar(i)*-X, where X = divine epithet → “([Deity +] divine epithet) *tubar(i)*-” (same as Type 2) or “[Deity+] *tubar(i)*- (divine epithet)”

It quickly emerges that this stem is associated with various theonyms, but also with other lexemes belonging to different grammatical classes (nouns, adjectives or adverbs, when their identification is assured), used as likely divine epithets. On this basis, two hypotheses are conceivable:<sup>47</sup>

- 1) \**Tubar(i)*- is a theonym, as compounds of types 2 and 3 seem to suggest, and Τούβερις, indirectly attested as the name of a local hero or deity in Lycia, would reflect this theonym. In this case, type 1 compounds (theonym + \**tubar(i)*-) would be dvandvas combining two divine names, in line with what we already find in Anatolian onomastics:<sup>48</sup> in cuneiform sources, e.g. <sup>d</sup>SIN-<sup>d</sup>U (Laroche 1966: no. 138) “Moon-god (and) Storm-god”; in Hieroglyphic Luwian, e.g. /Sanda-tiwad(i)-/ (TÜNP 1.1; Hawkins 2000: 155) “War-god (and) Sun-god”; and in alphabetical documentation, e.g. Αρμα-ρωνζας (LGPN 5b), “Moon-god (and) Protective Stag-god”. However, apart from Stephanus of Byzantium’s late and indirect attestation, there is no other testimony of this word as a theonym in the various corpora, suggesting rather a secondary use as a divine name. Furthermore, the only avatar of this stem in the lexicon of a Luwic language, that is Lyc. B *tuburiz*, is in all likelihood an appellative used in the plural (see Section 7 below).
- 2) \**Tubar(i)*- is originally a divine epithet, associated with various deities in compounds of type 1. In type 2, therefore, the divine epithet used as first element would be a hypostasis, i.e., it would replace a divine onomastic sequence theonym + epithet. A similar process is frequently attested elsewhere, cf. *piḥassa/i*- “luminous; (subst.) lightning”, first used as a characteristic of the Storm-god (van Gessel 1998: 791) before constituting a hypostasis (e.g. *Piḥassa-muwa*- “Having the Might of [the Storm-god of] Lightning”) or Lyc. \**xttbile/i*- “destroyer” (cf. *xtta(i)*- “to harm, to damage”), associated with the Moon-god in anthroponymic compounds (e.g. Ερμε-κτιβιλις/Ερμα-κτιβιλος/Ερμα-κτυβελις), to finally replace the whole divine onomastic formula in the personal name Κτιβιλας “(Moon-god) The Destroyer” (Adiego 2022: 84-86). In type 3, finally, the elements are reversed in relation to type 2, since \**tubar(i)*- is used as first member, in association with another epithet: either the meaning of such a compound is equivalent

<sup>47</sup> The claim that the stem *tubar(i)*- refers to a deity finds independent support in the fact that it serves as the base for the place name Tyberissos, according to a mode of formation that finds parallels in Anatolia: see Réveillac forthc. a.

<sup>48</sup> On these personal names, see in particular Laroche 1966: 282; Zehnder 2010: 58–59; Melchert 2013: 36; Yakubovich 2013: 103–104.

to type 2 or, which seems more likely to me, *\*tubar(i)-* is to be interpreted here as a hypostasis and stands in for a formula theonym + *\*tubar(i)-*. This would then constitute the intermediate stage between its use as an epithet and its use as a divine name, which is attested later in Lycia in the form *Τούβερις*.

## 6. A warlike epithet?

The only attestation of the stem under discussion in the lexicon comes from Lycian B, in the form *tuburiz*, unfortunately mentioned in passages that are far from clear. As a result, it is necessary first to consider the semantics of the anthroponymic compounds listed in the previous section before attempting an etymological analysis of *\*tubar(i)-*.

The association of *\*tubar(i)-* with elements of such diverse nature suggests that it must not be a specific divine epithet: the underlying lexeme rather represents some kind of generic divine title or designates a quality compatible with different deities. In this respect, the previously listed names can be compared with the numerous Luwian and Lycian compounds having *nan(n)(i)-* “lord, leader” as second member,<sup>49</sup> particularly as several of these compounds have an identical first member: *Ερμα-τοβορις/Ερμαν-δοβερις/Ερμαν-δυβερις* vs. Hier. /*Arma-nan(n)(i)-*/, Lyc. *Erñmenēneli-* → *Ερμε-νηνις* etc.; *Ταρκυν-δβεροας* vs. Cun. <sup>md</sup>U-ŠEŠ, Hier. TONITRUS-FRATER<sub>2</sub> /*Tarhu(nta)-nan(n)(i)-*/; *Περπεν-δυβερις* vs. *Περπε-νηνις* (← Lyc. *\*Perepñ-nēneli-*).

In order to precisely define the meaning to *\*tubar(i)-*, it is beneficial to compare it with the matching lexical items and attempt a morphological analysis. The relevance of a Luwic lexical group containing the seme [STRIKE] becomes immediately apparent:

- HLuw. /*tub(a)-*/ “to strike down”; /*tub(a)i-*/ “to strike repeatedly”; /*tubas-*/ or /*tubassa-*/ “battle”;
- CLuw. *dūpi-/dūpai-* “to strike, touch, punish”;
- Lyc. A and B *tub(e)i-* “to strike down”; *tube-* “attack, battle (?)”.

These words, which are based on a Proto-Anatolian stem going back to the PIE root *\*(s)teup-* “to strike”, reflect two types of semantic development (see Sasseville et al. 2023):

- 1) [STRIKE] > [PUNISH; DESTROY]. This semantic value is attested in Luwian and Lycian in the context of curse formulae to designate the divine punishment to which any desecrator is exposed. Thus, in Hieroglyphic Luwian:

KULULU 5 (= Hawkins 2000, p. 485–486) :

§ 4 *za-ti-za-pa* DOMUS-*na-zá kwali-sá a+rali-na* CUM-*ni i-zi-i-ri+i*

[...]

§ 8 *wali-tu-ta za-zi* DEUS-*ni-zi wali-la* “PES”-*tu*

§ 9 *wali-na tu-pa-tu*

§ 10 *wali-tu-u [á-lali-ma-za ARHA]* DELERE-*nú-tu-u*

“(he) who will cause ARA(NA) to these houses, [...] for him may these gods come mortally! Let them strike him down, let them destroy his name...”

<sup>49</sup> On this noun, see the discussion in Oreshko 2014. Contra Réveillac (2018: 479–480), the translation by “lord, leader” on the basis of the connection with CLuw. *nana-* “to lead” (compare Latin *dux*) seems now plausible to me. It is also implicitly adopted in Adiego (2019b: 158). Nevertheless, the precise connection between this noun and the name of the “brother” poses a challenging issue that cannot be addressed within the scope of this present paper.

In this passage, the formula *wa/i-na tu-pa-tu* /w=an tubantu/ “let them strike him” is integrated into a network of divine threats. Similar formulae can also be found in Lycian epitaphs. Here are a few examples:

TL 72.2+3-4 (bilingual):

*m=ēne mahāi : tubeiti : nelez[i]*

“the gods of the agora will destroy him”

οἱ θεοὶ | ἀπολέσειαν : οἱ ἀγοραῖοι

“may the gods of the agora destroy (him)”

TL 80.3-4 :

*m=ene | [Trq]as tubidi se Maliya hrixuwama*

“Trqqñt and Maliya the Supervisor will destroy him”

TL 93.3 :

*m=ēne Trqas tubidi se muhāi huwedri :*

“Trqqñt and all the gods will destroy him”

2) [STRIKE] > [ATTACK; FIGHT]. This meaning is admittedly less certain than the previous one, since it is generally inferred from the warlike context for both Lyc. A *tube-* (TL 29.10; TL 44b.56) and HLuw. *tù-pa-sa<sub>6</sub>-ti* /tubas(s)adi/, which is a hapax:

TOPADA § 17-18 (= Hawkins 2000: 451-461) :

*wa/i-mu á-mi-sa<sub>4</sub> DOMINUS-ni-sa (DEUS)TONITRUS-zi/a-sa<sub>8</sub> (DEUS)SARMA-sa<sub>8</sub> (DEUS)L198-sa<sub>6</sub> (DEUS)BOS.L206.PANIS-sa<sub>8</sub>-ha PRAE-na L179-ia-ta<sub>x</sub> | wa/i-mi-ta<sub>x</sub> tù-pa-sa<sub>6</sub>-ti wa/i<sub>5</sub>-sú-ha*

“my lord Tarhunt, Sarruma, Ea and (god) X ran before me, and I succeeded by battle”

The form of the noun here is ambiguous, since it could theoretically correspond to a neutral *s*-stem /tubas-/ derived from the verb or to an adjective /tubassa/i-/ derived from of a PAnat. stem \**tub(a)-* “strike, attack”.

To this little dossier one can add Τουβασσις,<sup>50</sup> a Carian divine name of Greek transmission, which could originally be an epithet equivalent to the Luwian form. If the latter is indeed to be assigned to the stem /tubassa/i-/, as argued in Simon 2016, who believes that Τουβασσις can be interpreted as “(God) of Smiting” or “(God) of Fighting”. A similar semantic ambiguity exists for a related Luwian term known from Hittite sources, namely the abstract \**dupattar-* (gen. adj. acc. *dupattanaššin*), only used as an epithet of the Storm-god: does the epithet designate a god who strikes with lightning, one who fights, or even one who protects from smiting?<sup>51</sup> All are theoretically possible. Indeed, on the one hand, the *Annals of Muršili II* relate how the Storm-god defeated the king’s enemy by striking the land of Arzawa with his lightning;<sup>52</sup> on the other, it is common for gods to assist their protégé in battle. This applies not only to the Storm-god, but to all protective deities, as is shown clearly in the *Annals of Hattušili I*,

<sup>50</sup> Gen. Τουβασσις[ς]: Kızıllı et al. 2015: 379, 389.15.

<sup>51</sup> See the full discussion, with references, in Simon 2022b.

<sup>52</sup> CTH 61.I: *mahhan=ma iyahhat nu GIM-an INA UR.SAG Lawaša arhun nu=za dU NIR.GÁL EN-YA parā handandātar tekuššanut nu G<sup>IS</sup>kalmišanan šiyāit nu G<sup>IS</sup>kalmišanan ammel KARAS<sup>HA</sup>-YA uškit KUR URU Arzauwa=y(a)=an uškit nu G<sup>IS</sup>kalmišanaš pait nu KUR URU Arzauwa GUL-ahta ŠA m<sup>U</sup>hha-LÚ-ya URU Apāšan URU-an GUL-ahta (KBo 3.4.15–19; Goetze 1933: 46–47) “Wie ich aber marschierte, wie ich da zum Gebirge Lawaša gelangte, da zeigte der stolze Wettergott, mein Herr, seine göttliche Macht, und einen Donnerkeil schmetterte er hin. Und den Donnerkeil sah mein Heer. Auch das Land Arzawa sah ihn, und der Donnerkeil ging hin und traf das Land Arzawa. Auch des Uhha-LÚ-iš Stadt Apaša traf er.” (transl. A. Goetze).*

where the Sun-goddess of Arinna is reported to have “run before” (*p[é]ran hūwāiš*) the king on the battlefield, thus ensuring the surrender of the enemies.<sup>53</sup> The common motif of the deity “running before” on the battlefield is also found in the hieroglyphic Luwian inscription from TOPADA mentioned above.

In the case of *tubar(i)-*, the semantics of the stem *\*tub(i)-* is also ambiguous, but can possibly be clarified by its formation. There are two possibilities:

1) either an adjectival derivation in *-ra/i-* < *\*-ro-* on a reconstructed stem *\*tuba-* “strike” or “battle”, but this type of denominal formation seems unproductive in the synchrony of the Anatolian languages, although it does exist;<sup>54</sup>

2) or a compound *\*tub-ar(i)-* “companion of *tuba-*”, with the second member *\*ar(i)-*, virtual Luwian equivalent of Hitt. *arā-* c. “friend, companion”. Such a formation would then be comparable to Lyc. A *tidere/i-* “collacteus/-a, foster brother/sister”, with *\*tide-* “teat, breast” (cf. Hitt. *tēta(n)-*, CLuw. *tītan-* < PIE *\*d<sup>h</sup>eh<sub>1</sub>-j-* “to suckle”)<sup>55</sup> as first member. For such a compound, the meaning “battle, fight” is then much more relevant than “strike”, so *\*tub-ar(i)-* would designate the “battle companion, ally (σύμμαχος), comrade-in-arms”.

From a semantic perspective, this second hypothesis sheds light on the numerous uses of the epiclesis alongside various deities, each representing distinct entities tasked with accompanying a protégé in warfare. It has been observed that the intercession of tutelary deities is mentioned several times in Hittite and Luwian texts. This suggests that anthroponymic compounds containing *\*tubar(i)-* resonate with this divine function, encompassing both a martial and protective role.

A typological parallel is possible with the Greek texts, where warlike epicleses are not exclusive to Athena or Ares, as illustrated by some inscriptions. For example, the votive inscription from the Cyrenaean Paniskos *I.Pan du desert: 3* (Wadi Bir el-Aïn, Egypt, Ptolemaic period) displays a coherent network of warlike epicleses applied to several deities (Mother of Gods, Ares, Zeus, and Pan), who are asked by the dedicator to take part in the military expedition he is about to lead (see Bonnet & Galoppin 2021: 7–11).

It is not rare to find military or political titles used as divine epithets in Anatolian and Near Eastern languages. In some instances, these titles even replaced the original theonym.<sup>56</sup> The transformation of the epithet *tubar(i)-* into a divine name is suggested by anthroponymic compounds of type 3 (as discussed in section 5 above), where the stem is used as the first

<sup>53</sup> CTH 4.II: LUGAL.GAL *Tabarnaš* NARĀM dUTU URU *Arinna nu=mu=za=kan a[nda ginuwāš=š]a[š ḫal]iš[t]a<sup>?</sup> nu=mu kešša[rta ḫarta<sup>?</sup>] n=aš=mu MÈ-ya p[é]ran hūwāiš nu INA URU *Ninašša MÈ-ya pāun nu=mu mahḫan LÚ<sup>MEŠ</sup> URU *Ni-našša menahḫanda auer nu* (KÁ.GAL<sup>HL.A</sup>) EGIR-pa ḫeššir (KBo 10.2 i 27–32; Imparati 1965: 46–47; Melchert 1978: 8–9) “The Great King, the Tabarna, beloved of the Sun-goddess of Arinna — she placed him on her lap — she took me by the hand and ran before me in battle. In the city of Ninašša I went to battle, and when the men of Ninašša saw me facing them, they reopened the city gates.”**

<sup>54</sup> Note the possible existence of a *\*-ro-* derivation in the Hittite hapax *tupran* (c./n. acc.) “Herzklopfen (?)” (Oettinger 2001: 461, fn. 15), which could reflect a primary PIE derivation *\*(s)tup-ró-* “striking” (cf. lat. *stuprum* “disgrace, violation etc.”).

<sup>55</sup> A similar analysis is cautiously proposed by Billing 2023 for the Lyc. A hapax *tuveri* (TL 84.5), whose meaning is unclear. According to the author of the eDiAna entry — and quite convincingly, in my opinion — it could be a compound *\*tuw-eri-* litt. “companion of the offering” (cf. Lyc. B *tuwe/i-* “offering” and *°ere/i-* “companion”, perhaps in *tidere/i-*).

<sup>56</sup> The reader can find a survey of well-known examples (such as Bel “Lord” first associated with the Babylonian god Marduk, before he became the object of a specific cult until Late Antiquity) in Yakubovich 2021: 239–240, which also convincingly suggests that the theonym Ἐνυάλιος constitutes a borrowing from a Lydian epiclesis *ēnwa<sup>h</sup>ala-* to be reconciled with the Hittite (probably borrowed from Luwian) *annawali-/annauli-* “co-ranked, peer, colleague”.



member along with a divine epithet. Additionally, indirect evidence for this evolution comes from the mention of *Τούβερις* as a local hero or deity in Lycia by Stephanus of Byzantium.

### 7. Lyc. B *tuburiz*

In the Lycian B lexicon, a direct reflection of \**tubar(i)*- is evidenced by the term *tuburiz*, occurring three times in the poem from the Xanthos Pillar (TL 44c.53; c.57; d.31–32), engraved on the funerary monument of a ruler of Xanthos (TL 44c.32–d.71). This text, flanked by a lengthy Lycian A inscription (TL 44b.1–c.19) and a Greek epigram (TL 44c.20–31), remains the most cryptic one among the three inscriptions due to the scarce attestation of Lycian B: beyond it, only one other text, the poem of Antiphellos (TL 55), is known in this language. Despite multiple studies dedicated to TL 44c.32–d.71, the precise content of this text remains ambiguous, making it unrealistic to use it alone as definitive evidence for the proposed hypothesis. However, it is crucial to ensure that the context in which *tuburiz* appears is compatible with the proposed meaning.

Prior research has primarily associated *tuburiz* with the toponym *Tuburehe(i)*- *Τυβερισσός*, with limited advance beyond this link, as noted in Neumann 2007: 385.<sup>57</sup> Among recent contributions, there exist two conflicting interpretations. Schürr (2016: 174, 190–191), inspired by a suggestion by Zimmermann (2003: 293), draws a parallel between the poetic formula *tr̄m̄miliz* : *tbiplē* : *tr̄pplē* : *tuburiz* (TL 44c.51–52) and the relationship between the mythical figures *Τούβερις* and *Τέριμερις* indirectly reported by Stephanus of Byzantium. This connection leads Schürr to conclude that *tubar(i)*- refers to Tyberissos as an ethnonym. Nevertheless, the association between *tr̄m̄miliz* and *Τέριμερις*, forming the basis of this parallel, is approximate and, thus, fragile.

Offering an alternative perspective, Sasseville 2021 proposes to identify a lexical network in the poem associated with offerings made by the ruler to appease enraged deities. By considering context and etymological parallels, Sasseville suggests three offerings: *waxssa*- “bread” (cf. Hitt. <sup>NINDA</sup>*wageššar* “bread bites”), *mur(i)*- “bunch” (cf. Hitt. *muri*- “bunch, grape”) and *tubar(i)*- “vine” (cf. HLuw. /*tuwaris*-/ “vineyard”). While this proposal provides a comprehensive and coherent interpretation supported by etymology, it remains challenging in a context where circular arguments can arise. The analysis leans heavily on the assumption that *waxssa*- designates the “bread”, which is contentious and, in my opinion, highly uncertain. Additionally, the link between *tubar(i)*- and HLuw. /*tuwaris(a)*-/ is, in my view, phonologically unacceptable (see fn. 14 above). Overall, the interpretation remains open to alternative viewpoints.

Addressing all the issues raised by this text is unfeasible at the present stage. Instead, it is possible to offer a few hypotheses regarding the occurrences of *tuburiz*. As observed, the Xanthos poem is part of a coherent text group lauding the military deeds and piety of a ruler, likely Gergis-Xeriga. Notably, the Greek epigram connects the king’s achievements closely with the divine world, attributing his conquests to Athena “destroyer of cities”, acknowledging divine rewards, offering numerous tributes to Zeus, and culminating with the erection of a stele in dedication to the Twelve Gods. The Lycian A section also implies a votive inscription, mentioning several identifiable deities. One can make an educated guess that the Lycian B poem

<sup>57</sup> *Apud* Neumann, *ibid.*, V. Ševoroškin points out to the similarity of *tubar(i)*- with the supposed second member of anthroponyms of Greek transmission such as *Αρδυβερις*, *Ερματοβορις* or *Περπενδυβερις*, without, however, making the connection explicit.

addresses a comparable subject, which is supported by the use of theonyms (e.g. *Natri-*, *Trqqñt-*), heronyms (*Zrppedun-*), and military terms (e.g. *laxa-* “military campaign”) within the text.

Regarding *tubur(i)-*, it solely appears in the plural form *tuburiz* (nominative or accusative):

a) *trm̄miliz: tbiplē: trpplē: tuburiz: pduradi: xuzruwētiz* (TL 44c.52-53);

b) *murēnedi : tuburiz : uple<s>iz* (TL 44c.57);

c) *m̄qrimiz ñtuwiteni: uplesiz waxssadi: tuburiz* (TL 44d.29-31).

In each instance, *tuburiz* is associated with adjectives such as *trm̄mile/i-* “Lycian” and *xuzruwāt(i)-* “who has *xuzr-*”. Both in b) and c) *uplese/i-* appears: this possessive adjective has not yet found a satisfactory translation but its stem is well rooted in local onomastics, as illustrated by the names *Οπλης* (Lycia, Kibyrtis and particularly Pisidia), *Οπλεσις* (Cabalis, Pisidia),<sup>58</sup> *Οπλεσιανός* (Pisidia), *Οπλανιλις* (Kibyrtis), *Αρτιμ-οπλης* (Cabalis), and *Παναμυ-οπλης* (Kibyrtis, Pisidia).<sup>59</sup> The other terms remain equally obscure: the noun *waxssa-* in c), here in the ablative-instrumental (as *murēnedi* and perhaps *pduradi*), is also a well-known element used in the regional onomastics: Lyc. *Waxssepddimeli-*, *Waxssebe-*; Car. *Uksmu-/Waksmu-*, *Uksi-/Uqsi-*; *Ουαξαμουαξ/Ουαξαμωξ* (Isauria, Cilicia), *Ουαξαξ* (Isauria).

Comparing a) (extended to the following lines) to a passage in Lycian A is allowed by common place-names displayed in both sections:

Lyc. A (TL 44b.47-56)

*s=ed=de : ahata ha | [de] ēñē : q̄la=bi : eheteḥi : se mahāna : ehēt | [he] Arñna : Tuminehi : Kerθθi : Xākbi : epi=d | [e ñ]temlē (e)si tāma : s=e=we=nepe : astte : Trm̄m̄lis : s=ed=de=tuwetē : kumezija : tere tere | Trqqñti : pddātahi : qññākba : xrssēni : eh | bi : tabahaza : kumezija : padritahi : Arñ | na Tuminehija : kumezija : Xākbi : kume | zi[j]a : se tukedri : Kerθθi : ade : urublijē | Hātahe : tubehi : prñnezi : se lihbeze eh[b] | ije*

He has left peace<sup>60</sup> in the precinct of Peace and to the gods of Peace in Xanthos, in Tymnessos, in Kerththi, in Kandyba. (There) a temple is a building: Lycia had not (yet) made them. Everywhere he erected sanctuaries to the local *Trqqñt* (Storm-god) (and) to his twelve personal celestial gods, sanctuaries in Aphrodision (?) in Xanthos, Tymnesian sanctuaries, Kandybian sanctuaries and in Kerththi he made a monument to Handa (War-god) of Battle for his family and his *lihbeze*.

Lyc. B (TL 44c.51-54)

*se=de ker̄i trisu : qñnatbisu : prete | laxadi zrētēni : sebe=be qirzē : utakija Trm̄miliz : | tbiplē : trpplē : tuburiz : pduradi : xuzruwētiz :) | Xāzbi : Tuminesi : hñtawā : Kridesi :*

The mention of three cities, Tymnessos (Lyc. B *Tuminesi*; A *Tuminehi*), Kandyba (Lyc. B *Xāzbi*; A *Xākbi*) and Kertthi (Lyc. B *Kridesi*; A *Kerθθi*) in both passages justifies a comparison, although their contents are not identical.<sup>61</sup> Despite a lack of comprehension for several terms, I propose the provisional translation:

<sup>58</sup> This form probably reflects the anthroponymic conversion of the genitival adjective *uplese/i-*.

<sup>59</sup> KPN § 1097. The exact references of the anthroponyms are listed in LGPN 5b (Lycia) and 5c (Kibyrtis-Cabalis, Pisidia). Note also the strong presence of Greco-Anatolian *noms d'assonance* based on the similarity of the indigenous stem with Greek names: e.g. *᾽Οπλων* (73 individuals in southern Anatolia vs. 10 in the rest of the Greek world covered by the LGPN), *᾽Οπλωνιανός* (Pisidia), *᾽Οπλωνίς* (Pisidia), *᾽Οπλητιανός / ᾽Οπλητιανή* (Pisidia). On the question of Greco-Anatolian *noms d'assonance*, see most recently Réveilhac in press.

<sup>60</sup> The meaning “peace” for Lyc. A *ahata-* and Lyc. B *asata-* has recently been challenged by Melchert forthc., suggesting “good fortune” instead. However, the traditional meaning is defended in Réveilhac forthc. b.

<sup>61</sup> According to Craig Melchert (pers. comm.), the division of the sequence must be considered. Indeed, the symbol :) likely serves to separate prosodic units, making it unclear whether the three toponyms should be linked

Three times in the country he *pre-* twelve (divine) protectors during the expedition and he carries (?) as *utakija-* of the designated properties Lycian *tubur(i)-s xuzruwēt(i)-s* in double, in triple in Kandyba, in Tymnessos (and) a statue of the god Handa in Kerththi.

The formula *qñnatbisu* [...] *zrētēni* “twelve times... a divine protector” might be understood as a poetic expression akin to Lyc. A *qñnākba: xrssēni: ehbi* “to his twelve personal gods”, thus indicating the Twelve Gods protecting the king, hence a translation “twelve divine protectors”. As for the group *Trmīmiliz: [...] tuburiz* [...] *xuzruwētiz*, it might correspond to another poetic periphrase referring to the same deities, i.e. literally “the Lycian (divine) allies who have *xuzr-*”. Syntactically, this group would be on the same level as *hñtawā* (acc. sg.), likely derived from the Lycian theonym *Hāta-* (< *Sanda-*), which would therefore exceptionally retain the initial aspirate typical of Lycian A (in order not to distort the divine name), with the *-wa-* suffix (< *-wo-* + *\*-h<sub>2</sub>-*)<sup>62</sup> and therefore literally meaning “that which is like Handa”, a poetic turn corresponding to Lyc. A *tukedri* [...] *Hātahe* “statue of Handa”.

Overall, while subject to speculation, this analysis attempts to unify *tuburiz* within both martial and religious themes, showcasing its relevance in varying contexts.

## 8. Conclusions

The study yields several significant conclusions. Firstly, the existence of a Proto-Luwic stem *\*tubar(i)-* is affirmed through personal onomastics, indirectly evidenced in the place names *Tuburehe(li)-/Τυβερι(σ)ός* and the divine name or heronym *Τούβερις*. This stem manifests diversely in various languages, with variations such as */TuPar(i)-/* in Hieroglyphic Luwian, *°tubr-/°tjbr-* in Carian, *τ[β]ερι-* in Pisidian, *°τοβορις* in Greek from Lycia, *°τουβαρις* in Greek from Pisidia, and *°τβηρις* in Greek from Caria.

Moreover, a broader range of personal names reveals alternative forms of this onomastic stem: in Carian names as *°dubr-/°dybr-* and *°dybyr-* and in Greek-transmitted names as *°δυβερι-/°δο/(υ)βα/ε/ηρα* and *°δυβερος*. The initial dental of *\*tubar(i)-* undergoes voicing assimilation when in contact with the final [n] or [r] of the first member in these instances. This comprehensive explanation unifies and interprets multiple personal names that were previously explained through distinct and sometimes speculative processes.

Furthermore, the element *\*tubar(i)-* appears as the first member in some compounds, albeit less frequently and often in syncopated forms: *Tbr°* and *Toβορ°* in Caria, *Tβερ°*, *Tβαρα°* and *Tβερη°* in Cilicia.

An overview of these occurrences, detailed in Sections 2 to 4, reveals that this stem is consistently associated with a theonym or a divine epithet. Through the combinations in anthroponymy, it is suggested that *\*tubar(i)-* represents a divine title, which evolved into a divine hypostasis and then into a divine name or heronym, as implied by Stephanus of Byzantium's account about *Τούβερις*.

---

to the previous clause, particularly given the blank at the end of line 53. Yet, incorporating this sequence into the subsequent clause is challenging due to the numerous nouns it contains. As a result, the sequence at the start of line 54 might be an enjambment, possibly emphasized within the broader poetic structure.

<sup>62</sup> The *-wa-* suffix, attested in several Anatolian languages (Hittite, Luwian, Lycian, and Carian), was used to form words denoting a social category (Rieken & Sasseville 2014). However, it originally had an “equitative” function, which can still be seen in certain nouns such as *prīna-wa-*, lit. “that which is like a house (*prīna-*)”, hence “tomb-house, mausoleum” (Rieken & Sasseville 2014: 310).

Etymologically, *\*tubar(i)-* likely relates to the Luwian /*tub(a)-*/ “to strike” and Lycian *tube-* “battle”. It might be a compound /*tub-ar(i)-*/ meaning “battle companion, comrade-in-arms”. This concept echoes Hittite and Luwian texts describing deities engaging in battles alongside their protégés.

Finally, *tubar(i)-* finds a direct avatar in the Lycian B lexicon, with *tuburiz*. This word potentially holds the meaning of “(divine) allies” in the poem on the Xanthos Pillar, considering the martial and religious context, although this interpretation remains uncertain.

The prominence of *tubar(i)-* in Luwic onomastics suggests its use in defining and designating deities, evident in personal names across Luwian, Lycian, Carian, and Pisidian collections, Greek inscriptions from southern Anatolia, the heronym *Τούβερις* from indirect sources, and the place name Tyberissos (see Réveilhac forthc. a).

## References

- Adiego, Ignacio-Javier (Ignasi-Xavier). 2007. *The Carian Language*. Leiden / Boston: Brill.
- Adiego, Ignasi-Xavier. 2019a. A Kingdom for a Carian Letter. In: Ignasi-Xavier Adiego, José-Virgilio García Trabazo, Mariona Vernet, Bartomeu Obrador-Cursach, Elena Martínez Rodríguez (eds.). *Luwic Dialects and Anatolian Inheritance and Diffusion*: 11–50. Barcelona: Universitat de Barcelona.
- Adiego, Ignasi-Xavier. 2019b. The Survival of the God Name Šarruma in Cilician Names in the Greek Sources. *Alt-orientalische Forschungen* 46: 147–160.
- Adiego, Ignasi-Xavier. 2020a. Consonant Clusters, Defective Notation of Vowels and Syllable Structure in Caromemphite. In: Ronald I. Kim, Jana Mynářová, Peter Pavúk (eds.). *Hrozný and Hittite: The First Hundred Years. Proceedings of the International Conference Held at Charles University, Prague, 11-15 November 2015*: 95–119. Leiden / Boston: Brill.
- Adiego, Ignasi-Xavier. 2020b. Adaptación griega de nombres licios, adaptación licia de nombres griegos: aspectos fonológicos. In: C. Ruiz Darasse (ed.). *Comment s’écrit l’autre? Sources épigraphiques et papyrologiques dans le monde méditerranéen antique*: 43–60. Bordeaux: Ausonius éditions.
- Adiego, Ignasi-Xavier. 2022. Luwian *Tarhunaza-*, Cilician *Tqokovačac*, *Tqikovačac*. *Indogermanische Forschungen* 127: 75–90.
- Anrich, Gustav. 1913–1917. *Hagios Nikolaos: Der Heilige Nikolaos in der griechischen Kirche. Texte und Untersuchungen*, 2 volumes. Leipzig: B. G. Teubner.
- Bauer, Anna H., Elisabeth Rieken, Zsolt Simon, David Sasseville, Andreas Opfermann. 2022. Proto-Anatolian *\*m̥iēu-/m̥iu-* (eDiAna-ID 1440). In: *eDiAna*. Available online at: [www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=1440](http://www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=1440) [accessed on 19/10/2023].
- Bean, Georg. E., Terence B. Mitford. 1970. *Journeys in Rough Cilicia 1964–1968*. Wien / Graz: Böhlau Verlag.
- Billing, Oscar. 2023. Lycian A *tuwere/i-* (eDiAna-ID 1952). In: *eDiAna*. Available online at: [www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=1440](http://www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=1440) [accessed on 19/10/2023].
- BoHa 19 = Herbordt, Suzanne. 2005. *Die Prinzen- und Beamtensiegel der hethitischen Großreichszeit auf Tonbulln aus dem Nišantepe-Archiv in Hattusa (Boğazköy-Ḫattuša Ergebnisse der Ausgrabungen 19)*. Mainz am Rhein: von Zabern.
- Bonnet, Corinne, Thomas Galoppin. 2021. Divine Names on the Spot: Towards a Dynamic Approach of Divine Denominations in Greek and Semitic Contexts. In: Thomas Galoppin, Corinne Bonnet (eds.). *Divine Names on the Spot. Towards a Dynamic Approach of Divine Denominations in Greek and Semitic Contexts*: 1–17. Leuven / Paris / Bristol, CT: Peeters.
- Brixhe, Claude. 1976. *Le dialecte grec de Pamphylie: documents et grammaire*. Paris: Adrien Maisonneuve.
- Brixhe, Claude. 2016. *Stèles et langue de Pisidie*. Nancy: A.D.R.A.
- Cammarosano, Michele. 2015. Foreign Gods in Hatti. A New Edition of CTH 510. *KASKAL. Rivista di storia, ambiente e culture del Vicino Oriente Antico* 11: 199–244.
- CIG = 1860–1890. *Corpus inscriptionum Graecarum*. Berlin: Officina Academia/Reimer.
- CTH = Laroche, Emmanuel. 1971. *Catalogue des textes hittites* (2nd ed.). Paris: Klincksieck.
- eDiAna = Miller, Jared, Elisabeth Rieken, Olav Hackstein, Ilya Yakubovich (eds.). 2017-. *eDiAna. Digital Philological-Etymological Dictionary of the Minor Ancient Anatolian Corpus Languages*. Available online at: [www.ediana.gwi.uni-muenchen.de](http://www.ediana.gwi.uni-muenchen.de).

- van Gessel, Ben H. L. 1998–2001. *Onomasticon of the Hittite Pantheon*. 3 vols. Leiden / Boston / Köln: Brill.
- Goetze, Albrecht. 1933. *Die Annalen des Muršiliš*. Leipzig: Hinrichs.
- Hajnal, Ivo. 1995. *Der lykische Vokalismus. Methode und Erkenntnisse der vergleichenden anatolischen Sprachwissenschaft, angewandt auf das Vokalsystem einer Kleincorpussprache*. Graz: Leykam.
- Hawkins, J. David. 2000. *Corpus of Hieroglyphic Luwian Inscriptions. Vol. 1: Inscriptions of the Iron Age*. Berlin / New York: De Gruyter.
- Heberdey, Rudolf, Adolf Wilhelm. 1896. *Reisen in Kilikien, ausgeführt 1891 und 1892 im Auftrage der Kaiserlichen Akademie der Wissenschaften*. Wien: C. Gerold's Sohn.
- Houwink ten Cate, Philo H. J. 1961. *The Luwian Population Groups of Lycia and Cilicia Aspera during the Hellenistic Period*. Leiden: Brill.
- I.Arykanda = Şahin, Sencer. 1994. *Die Inschriften von Arykanda*. Bonn: Habelt.
- I.Cilicie = Dagron, Gilbert, Denis Feissel. 1987. *Inscriptions de Cilicie*. Paris: De Boccard.
- I.Carie hautes terres = Debord, Pierre, Ender Varinlioğlu. 2001. *Les hautes terres de Carie*. Bordeaux: Institut Ausonius.
- I.Pan du désert = Bernand, André. 1972. *Pan du désert*. Leiden: Brill.
- I.Termessos Suppl. = İplikçioğlu, Bülent, Güler Çelgin, A. Vedat Çelgin. 1991–2007. *Epigraphische Forschungen in Termessos und seinem Territorium*, 4 volumes. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- I.Westkilikien Rep. = Hagel, Stefan, Kurt Tomaschitz. 1998. *Repertorium der westkilikischen Inschriften. Nach den Scheden der kleinasiatischen Kommission der Österreichischen Akademie der Wissenschaften*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- IG = Deutsche Akademie der Wissenschaften zu Berlin. 1913–. *Inscriptiones Graecae*. Berlin: De Gruyter.
- Imparati, Fiorella. 1965. L'autobiografia di Hattušili I. *Studi Classici e Orientali* 14: 40–76.
- KBo = 1916–1921. *Keilschrifttexte aus Boghazköi*. Leipzig/Berlin: Gebr. Mann-Verlag.
- KON = Zgusta, Ladislav. 1984. *Kleinasiatische Ortsnamen*. Heidelberg: Winter.
- KPN = Zgusta, Ladislav. 1964. *Kleinasiatische Personennamen*. Prag: Verlag der Tschechoslowakischen Akademie der Wissenschaften.
- KUB = 1921–. *Keilschrifturkunden aus Boghazköi*. Berlin: Akademie-Verlag.
- Kızıl, Abuzer, Patrice Brun, Laurent Capdetrey, Raymond Descat, Pierre Fröhlich, Koray Konuk. 2015. Pidasa et Asandros: une nouvelle inscription (321/0). *Revue des Études Anciennes* 117: 371–409.
- Labarre, Guy. 2010. *Le dieu Mèn et son sanctuaire à Antioche de Pisidie*. Bruxelles / Fernelmont: E.M.E. & InterCommunications.
- Laroche, Emmanuel. 1946. Recherches sur les noms de dieux hittites. *Revue hittite et asianique* 7: 7–139.
- Laroche, Emmanuel. 1966. *Les noms des Hittites*. Paris: Klincksieck.
- Laroche, Emmanuel. 1973. Études de linguistique anatolienne. *Revue hittite et asianique* 31: 83–99.
- LGNP 5b = Balzat, Jean-Sébastien, Thomas Corsten, Richard W. V. Catling (eds.). 2013. *Lexicon of Greek Personal Names. Vol. 5b: Coastal Asia Minor: Caria to Cilicia*. Oxford: Clarendon Press.
- LGNP 5c = Balzat, Jean-Sébastien, Richard W. V. Catling, Édouard Chiricat, Thomas Corsten (eds.). 2018. *Lexicon of Greek Personal Names. Vol. 5c: Inland Asia Minor*. Oxford: Clarendon Press.
- Martínez Rodríguez, Elena. 2019. Phonotactics of the Lycian Labial Glide Clusters. *Indogermanische Forschungen* 124: 29–229.
- Masson, Olivier. 1990. *Onomastica Graeca selecta, I et II (introduction et index par C. Dobias et L. Dubois)*. Nanterre: Université de Paris X-Nanterre.
- Melchert, H. Craig. 1978. The Acts of Hattušili I. *Journal of Near Eastern Studies* 37/1: 1–22.
- Melchert, H. Craig. 1993. *Cuneiform Luwian Lexicon*. Chapel Hill: self-published.
- Melchert, H. Craig. 1994. *Anatolian Historical Phonology*. Amsterdam / Atlanta: Rodopi.
- Melchert, H. Craig. 2004. *A Dictionary of the Lycian Language*. Ann Arbor / New York: Beech Stave Press.
- Melchert, H. Craig. 2021. A Possible New Greco-Carian Contact Phenomenon. In: M. Bianconi (ed.). *Linguistic and Cultural Interactions between Greece and Anatolia. In Search of the Golden Fleece*: 108–115. Leiden/Boston: Brill.
- Melchert, H. Craig. Forthc. The Āssiya Deities. To appear in: *The proceedings of the 1<sup>st</sup> workshop The Gods of Anatolia and their Names held in Barcelona, May 17<sup>th</sup>–19<sup>th</sup> 2023*.
- Mouton, Alice, Ilya Yakubovich. 2021. Where Did One Speak Luwili? Geographic and Linguistic Diversity of Luwian Cuneiform Texts. *Journal of Language Relationship* 19/1: 25–53.
- Mouton, Alice. 2023. Solar Deities in the Kuwatalla Ritual Tradition: Epithets and Functions. In: Livio Warbinek, Federico Giusfredi (eds.). *Theonyms, Panthea and Syncretisms in Hittite Anatolia and Northern Syria. Proceedings of the TeAI Workshop Held in Verona, March 25–26, 2022*: 127–140. Firenze University Press.

- Neumann, Günter. 1976. Zu einigen hethitisch-luwischen Personennamen. *Zeitschrift für vergleichende Sprachforschung* 90/1: 139–144.
- Neumann, Günter. 2007. *Glossar des Lykischen* (überarbeitet und zum Druck gebracht von Johann Tischler). Wiesbaden: Harrassowitz.
- Oettinger, Norbert. 2001. Hethitisch *-ima-* oder: Wie ein Suffix affektiv werden kann. In: Gernot Wilhelm (ed.). *Akten des IV. Internationalen Kongresses für Hethitologie, Würzburg, 4.-8. Oktober 1999*: 456–477. Wiesbaden: Harrassowitz.
- Oreshko, Rostislav. 2014. The Strange Case of Dr. FRATER and Mr. DOMINUS: a Re-Consideration of the Evidence Concerning Luwian *nani-*. In: Piotr Taracha (ed.). *Proceedings of the Eighth International Congress of Hittitology, Warsaw, 5-9 September 2011*: 614–631. Warsaw: Agade.
- Réveillac, Florian. 2018. *Contact linguistique et emprunts onomastiques entre grec et lycien: apports à la phonétique et à la morphologie*. PhD dissertation. Sorbonne Université, Paris.
- Réveillac, Florian. 2021. Geminate Consonants in Lycian. A Twofold Interpretation. In: Annick Payne, Šárka Velhartická, Jorit Wintjes (eds.). *Beyond All Boundaries: Anatolia in the First Millennium BC*: 551–575. Leuven / Paris / Bristol, CT: Peeters.
- Réveillac, Florian. Forthc. As you like it: les noms d’assonance lyciens-grecs. Esquisse d’une méthodologie. In: Alcorac Alonso Déniz, Julian Méndez Dosuna, Enrique Nieto Izquierdo, Gilles van Heems (eds.). *Contact linguistique en Grèce ancienne: diachronie et synchronie (CoLiGA)*: 77–101. Lyon: MOM Éditions.
- Réveillac, Florian. Forthc. a. Tybérisos, un lieu divin : sur les traces d’une épiclèse en Anatolie méridionale. In: Alcorac Alonso Déniz, Enrique Nieto Izquierdo (eds.). *Noms de lieux en Méditerranée antique: approches linguistiques*. Nancy: A.D.R.A.
- Réveillac, Florian. Forthc. b. Anatolian Names in  $\text{ᾠ}\alpha\sigma\eta\tau\alpha\varsigma/\text{ᾠ}\alpha\sigma\alpha\tau\eta\varsigma/\text{ᾠ}\alpha\sigma\alpha\tau\alpha\varsigma$ , CLuw. \**aššatta-*, Lyc. B *asata-* and Lyc. A *ahata-*. To appear in: *News from the Lands of the Hittites* 8 (2024): 197–213.
- Réveillac, Florian. Forthc. c. Give a God a Bad Name and Beg Him: the Lycian Theonym  $\Sigma\omicron\upsilon\mu\epsilon\nu\delta\iota\varsigma/\Sigma\omicron\mu\epsilon\nu\delta\iota\varsigma$ . To appear in: *Proceedings of the 1<sup>st</sup> workshop The Gods of Anatolia and their Names held in Barcelona, May 17<sup>th</sup>-19<sup>th</sup> 2023*.
- Rieken, Elisabeth, David Sasseville. 2014. Social Status as a Semantic Category of Anatolian: The Case of PIE \*-*uo-*. In: H. Craig Melchert, Elisabeth Rieken, Thomas Steer (eds.). *Munus amicitiae: Norbert Oettinger a collegis et amicis dicatum*: 302–314. Ann Arbor / New York: Beech Stave Press.
- Rieken, Elisabeth, David Sasseville. 2022. Proto-Luwian \**Hudā-* (eDiAna-ID 2107). In: *eDiAna*. Available online at: [www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=2107](http://www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=2107) [accessed on 19/10/2023].
- Robert, Louis. 1955. *Hellenica. Recueil d’épigraphie, de numismatique et d’antiquités grecques, Vol. X*. Paris: Adrien Maisonneuve.
- Sasseville, David. 2021. Bread and Wine in Lycian Poetry. A Contextual Interpretation of TL 44c. *Kadmos* 60: 173–188.
- Sasseville, David. 2022. Palaic *ḥašāwant-* (eDiAna-ID 338). In: *eDiAna*. Available online at: [www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=2107](http://www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=2107) [accessed on 19/10/2023].
- Sasseville, David. 2023. Lycian A *ñtuweriha-* (eDiAna-ID 1806). In: *eDiAna*. Available online at: [www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=2107](http://www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=2107) [accessed on 19/10/2023].
- Sasseville, David, Caroline Schneider, Andreas Opfermann. 2023. Proto-Anatolian \*(s)*teup-* (eDiAna-ID 288). In: *eDiAna*. Available online at: [www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=2107](http://www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=2107) [accessed on 19/10/2023].
- Schürr, Diether. 2013. Kaunisch-Karisches in Krya: Revision der Grabinschrift und Vergleiche. In: O. Henry (ed.). *4<sup>th</sup> Century Karia. Defining a Karian Identity under the Hekatomnids*: 21–31. Istanbul: Institut Français d’Études Anatoliennes.
- Schürr, Diether. 2016. Zum Agora-Pfeiler in Xanthos V: Das Nordgedicht auf Cheriga (TL 44c, 32 ff.). *Kadmos* 55: 147–196.
- SEG = 1923-. *Supplementum epigraphicum Graecum*. Leiden: Brill.
- Simon, Zsolt. 2016. Toubassis. Zum Weiterleben anatolischer Gottheiten in der Eisenzeit. *Nouvelles Assyriologiques Brèves et Utilitaires* 2016/2: 82–84.
- Simon, Zsolt. 2016–2017. Zur Bedeutung von karisch *eri-*. *Münchener Studien zur Sprachwissenschaft* 70: 259–264.
- Simon, Zsolt. 2022a. Carian *kδous* (eDiAna-1071). In: *eDiAna*. Available online at: [www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=2107](http://www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=2107) [accessed on 19/10/2023].
- Simon, Zsolt. 2022b. Luwian \**dupattar/dupattn-* (eDiAna-ID 345). In: *eDiAna*. Available online at: [www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=2107](http://www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=2107) [accessed on 19/10/2023].
- Sittig, Ernst. 1911. *De Graecorum Nominibus Theophoris*. Halis Saxonum: Karras.

- Sterrett, John R. S. 1888. *The Wolfe Expedition to Asia Minor*. Boston: Damrell and Upham.
- TAM II = Kalinka, Ernst. 1920–1944. *Tituli Asiae Minoris. Vol. 2, Tituli Lyciae: linguis Graeca et Latina conscripti*. Vindobonae: R. M. Rohrer.
- Taracha, Piotr. 2009. *Religions of Second Millennium Anatolia*. Wiesbaden: Harrassowitz.
- TL = Kalinka, Ernst. 1901. *Tituli Asiae Minoris, Vol. 1, Tituli Lyciae: lingua Lycia conscripti*. Vindobonae: A. Hoelder.
- Valério, Miguel. 2007. “Diktaian Master”: a Minoan Predecessor of Diktaian Zeus in Linear A? *Kadmos* 46: 3–14.
- Valério, Miguel. 2015. Linear A *du-pu<sub>2</sub>-re*, Hittite *tabarna* and their Alleged Relatives Revisited. *Journal of Language Relationship* 13/4: 329–354.
- Valério, Miguel. 2023. Drive for power and the Indo-European source of Luwian *muwa-*. In: J.-V. García Trabazo et al. (eds.). *New Approaches on Anatolian Linguistics*: 265–291. Barcelona: Universitat de Barcelona.
- Wörrle, Michael. 2012. Die griechischen Sepulkralinschriften von Limyra. In: Jürgen Borchhardt, Anastasia Pekridou-Gorecki (eds.). *Limyra: Studien zu Kunst und Epigraphik in den Nekropolen der Antike*: 411–457. Wien: Phoibos-Verlag.
- Yakubovich, Ilya. 2013. Anatolian Names in *-wiya* and the Structure of Empire Luwian Onomastics. In: Alice Mouton, Ian Rutherford, Ilya Yakubovich (eds.). *Luwian Identities: Culture, Language and Religion between Anatolia and the Aegean*: 87–123. Leiden / Boston: Brill.
- Yakubovich, Ilya. 2017. The Luwian Title of the Great King. In: Alice Mouton (ed.). *Hittitology Today: Studies on Hittite and Neo-Hittite Anatolia in Honor of Emmanuel Laroche’s 100th Birthday*: 38–50. Istanbul: Institut Français d’Études Anatoliennes-Georges-Dumézil.
- Yakubovich, Ilya. 2021. The Anatolian Connections of the Greek God Enyalios. In: Michele Bianconi (ed.). *Linguistic and Cultural Interactions between Greece and Anatolia. In Search of the Golden Fleece*: 233–245. Leiden/Boston: Brill.
- Yakubovich, Ilya, Alice Mouton. 2023. *Luwili. Hittite-Luwian Ritual Texts Attributed to Puriyanni, Kuwatalla and Šilalluḫi (CTH 758-763), Vol. 1: Edition and Commentary, Vol. 2: Discussion and Glossary*. Wiesbaden: Harrassowitz.
- Zehnder, Thomas. 2010. *Die hethitischen Frauennamen. Katalog und Interpretation*. Wiesbaden: Harrassowitz.
- Zimmermann, Martin. 2003. Hafen und Hinterland. Wege der Akkulturation an der lykischen Küste. Vorbericht über die Feldforschungen in den zentrallykischen Orten Tyberissos und Timiussa in den Jahren 1999–2001, mit einem Beitrag von Oliver Hüllden. *Istanbuler Mitteilungen* 53: 265–312.

Флориан Ревеяк. Божественный эпитет \**tubar(i)*- в лувийской ономастике

Отправной точкой данной статьи является исследование М. Валерио (2015), посвященное нескольким группам личных имен из южной Анатолии. Здесь предлагается восстановить одну ономастическую базу, \**tubar(i)*-, общую для нескольких имен из различных языковых корпусов, принадлежащих к лувийской подгруппе (лувийский, карийский, писидийский) и греческим эпиграфическим источникам из южной Анатолии (от Карики до Киликии). Данный продуктивный элемент ассоциируется с многочисленными божественными именами или эпитетами в составе композитов, что указывает на то, что мы имеем дело с божественным титулом. Его значение можно восстановить как “боевой спутник, соратник”, изначально характеризовавший различных богов, чья роль заключалась в охране своих протезе на поле боя.

*Ключевые слова*: лувийские языки; анатолийские языки; лувийский язык; ликийский А язык; ликийский В язык; карийский язык; писидийский язык; греческий язык; ономастика; теофорные имена; анатолийская религия.

## Пермско-иранские языковые контакты по данным фонетики и морфологии удмуртского и коми языков

В статье рассматриваются некоторые результаты иранского языкового воздействия на языки пермской ветви финно-угорской группы уральской языковой семьи — удмуртского, коми-зырянского и коми-пермяцкого. В отличие от работ исследователей давнего и недавнего прошлого, упор в статье сделан не на лексическом фонде данных языков, а на их фонетических и морфологических системах.

*Ключевые слова:* иранские языки; скифский язык; осетинский язык; финно-угорские языки; пермские языки; удмуртский язык; коми-зырянский язык; коми-пермяцкий язык; фонетика; морфология; историческое языкознание; этническая история.

Среди древних внешних культурных и языковых контактов финно-угров Восточной Европы и Западной Сибири, которые в значительной мере определили ход их этнической истории в I тыс. до н.э. — I тыс. н.э., особое место занимают контакты со степным ираноязычным населением. Лексико-статистические данные свидетельствуют, что, наряду с правенгерским и, вероятно, предками поволжских финно-угорских (мордовских и марийского) языков, наиболее сильное иранское влияние в эту эпоху испытал пермский праязык — предок современных удмуртского и коми языков (см. напр. Напольских 2015: 53). Тема пермско-иранских языковых контактов уже давно и довольно обстоятельно изучается лингвистами. Достаточно упомянуть хотя бы соответствующие разделы уже ставших классическими трудов А. Йоки и К. Редеи, а также пространную статью В. И. Лыткина, специально посвящённую языковым контактам иранцев с древними пермянами (Joki 1973; Rédei 1986; Лыткин 1975).

Нельзя не заметить, что в названных работах, а также в прочих сочинениях, посвящённых в той или иной мере пермско-иранским языковым контактам, затрагивалась почти исключительно только одна сторона этого взаимодействия — лексические заимствования. По моему мнению, поиск результатов иранско-пермских контактов в других языковых сферах, в частности в области фонетики и морфологии, выглядит вполне перспективно, учитывая тот довольно значительный след, какой оставили эти контакты в словарном фонде пермских языков. Особенно перспективным мне представляется поиск подобных следов с привлечением данных и материалов восточноиранских языков сармато-аланского круга и их прямого потомка — осетинского языка, тем более что самые поздние лексические иранизмы пермских языков имеют специфически древнеосетинский облик (Напольских 2015: 139). Именно поэтому основное внимание в статье будет уделено взаимодействию пермских языков именно с этой частью иранских языков.

Следующий ниже текст представляет собой попытку выявить и предложить вниманию читателя возможные следы иранского (прежде всего, алано-осетинского) влияния в фонетических и морфологических системах удмуртского и коми языков.



## 1. Фонетика

1.1 Одной из наиболее заметных отличительных черт консонантизма пермских языков, которые выделяют их на фоне других финно-угорских языков, является озвончение части их согласных и появление ещё в прапермскую эпоху звонких взрывных, звонких сибиллянтов и звонких аффрикат (наряду с существовавшими и ранее глухими), коих не было в финно-угорском праязыке (ОФУЯ 1974: 118; ОФУЯ 1976: 136–137). Например, звонкие аффрикаты  $\text{ʒ}$  и  $\text{ʒ̣}$ , характерные для всех пермских языков и широко в них распространённые, нехарактерны практически для всех прочих финно-угорских языков.

Наиболее же ярким и показательным частным примером данного озвончения выглядит возникновение ещё в общепермскую эпоху (т. е. едва ли позднее рубежа I/II тыс. н. э.) в пермском праязыке анлаутных взрывных звонких фонем  $*b-$ ,  $*g-$  и  $*d-$  (в словах прафинно-угорского происхождения на месте старых  $*p-$ ,  $*k-$  и  $*t-$ , а также и в появившихся в пермских языках позже словах). Звонких взрывных согласных фонем в анлауте не только не существовало в финно-угорском праязыке, но и для подавляющего большинства финно-угорских языков такие согласные фонемы в начале слова нетипичны по сей день (ОФУЯ 1974: 119). Примеры:

— удм. *bur* ‘правый; добрый, хороший’, коми *bur* ‘добрый, хороший’ при мар. *poro* ‘добрый; добро’, эрз. *paro* ‘добро; добрый, хороший’, фин. *paras* ‘лучший, наилучший’ (КЭСК: 42; UEW: 724);

— удм. *giži*, коми *giž* ‘ноготь, коготь, копыто’ при мар. *küč*, эрз. *kenže*, фин. *kynsi* ‘ноготь, коготь’ (КЭСК: 84; UEW: 157);

— удм. *diñ* ‘комель (дерева); основание, фундамент’, кз. *din*, кп. *dijn* ‘комель, кочерыжка’ при мар. *tiiñ* ‘комель, основание, основа’, фин. *tyvi* ‘комель, основание’ (КЭСК: 94; UEW: 523) и т. д.

Даже в сравнительно поздних заимствованиях в подавляющем большинстве финно-угорских языков (за исключением пермских и, что важно (см. ниже в этом же разделе), венгерского) анлаутная звонкая взрывная согласная языка-источника, как правило, заменялась на соответствующую глухую, напр.: фин. *kenraali* ‘генерал’, *pankki* ‘банк’; мар. *paɣor* ‘багор’, *küsle* ‘тусли’ и т. п.

По мнению В. И. Лыткина и Т. И. Тепляшиной, озвончение части глухих согласных в пермских языках началось в прапермскую эпоху и завершилось ещё до отделения предков удмуртов от предков коми (ОФУЯ 1976: 138). С моей точки зрения, этот вывод о завершении данного процесса в пермских языках ещё до распада прапермской этноязыковой общности может быть поставлен под сомнение. Дело в том, что процесс озвончения, например, тех же анлаутных согласных в пермских языках и их диалектах прошёл неравномерно. Существуют пары общепермских лексем, где в коми языках в начале слова наличествует звонкая согласная, а в удмуртском — глухая, или наоборот. Более того, порой такие пары «глухая/звонкая согласная в начале одной лексемы» присутствуют параллельно даже в рамках одного отдельно взятого пермского языка. Примеры:

— кз. *bol*/*bolʔ* ‘пузырь’, кз., кп. *vabol* ‘пузырь, мозоль’ (где *va-* ‘вода, водяной’), при этом существуют кз. *polʔ*, кп. *pol*, кя. *pol* ‘пузырь, мозоль’ (см. также кз., кп. *pol’diñi* ‘набухать, вздуваться’); удм. *pułi* ‘тж.’; ср. эст. *pull* ‘пузырь’ с подобными параллелями в других финно-угорских языках (КЭСК: 40);

— удм. *biči* ‘почка, росток’; удм. *piči* ‘почка, верба’; кз. л. *poč* ‘почка’, печ. *poča* ‘сережка ивы’ (КЭСК: 226; ССКЗД: 295);

— кз., кп. *geb* ‘мошка’; удм. *kibi* ‘букашка, жучок’ (КЭСК: 75);

— кз., кп. *goz*, кя. *guz* ‘пара’; удм. *kuz* ‘пара’; при фин. *kansa* ‘народ’, эст. *kaasa* ‘супруг, супруга’ < праФП \**kansa* ‘Volk, Leute; Genosse, Freund’ (КЭСК: 77; UEW: 645);

— коми *gut* ‘муха’; удм. *kut* ‘тж.’ (КЭСК: 83);

— удм. *giž/kiž* ‘грязь (на теле), нечистоты’ (УРС: 117);

— удм. *gubiřes* ‘горбатый, сутулый’ / *kurıřes* ‘согнутый, горбатый’ (УРС: 114, 231);

— кз., кп. *žemđini*, удм. *čemtini* ‘запнуться, споткнуться’ (КЭСК: 88);

— кз. уд. *žied* ‘мелкий, незрелый, плохой (о ягодах, колосьях)’ (ССКЗД: 104), лет. *čijed*, др. кз. диал. *čijeb* (ССКЗД: 410) ‘маленький, хрупкий, худенький, слабенький’; удм. *čijed* ‘худощавый, тощий, хилый’ (КЭСК: 305);

— кз. *deṃni* ‘поставить заплату; запрудить (реку для ловли рыбы)’, *deṃas* ‘заплата’, *deṃed* ‘загородка, запруда (для ловли рыбы)’, кп. *deṃet* ‘тж.’; удм. *tiṃini* ‘запрудить’, *tiṃet* ‘пруд; запруда, плотина’ < праперм. \**tiṃ-* ‘stauen, abdämmen’; ср. также мар. *tumäš* ‘заплата’ и др. < праФУ \**toṃe-* ‘flicker’ (КЭСК: 96; UEW: 798);

— коми *žov/žol/žo/žu* ‘калина’; удм. *šu* ‘тж.’ < праФП \**šewz* ‘Maßholderbaum, Wasserholunder, Viburnum opulus’ (КЭСК: 102; КПРС: 142; UEW: 784);

— кз. уд. *kezan pu* ‘жимолость’; удм. *guzem pu* ‘тж.’ (КЭСК: 139; ССКЗД: 169);

— кз. скр. *čaręd*, вв. *žabed* ‘вешала, ряд кольев для просушки сетей’ (ССКЗД: 385) и др.

Как видим, процесс озвончения начальных (и не только) согласных корня слова выглядит незавершенным в том или ином пермском языке или его диалектах, он замер на разных стадиях в разных пермских диалектах и языках.

Помимо корневых основ, нечто очень похожее мы можем наблюдать и в некоторых аффиксах пермских языков, где также наблюдается чередование звонких и глухих согласных в разных пермских языках. Примеры:

— весьма продуктивному удмуртскому суффиксу прилагательных *-it* в коми-пермяцком языке соответствует *-it*, а в коми-зырянском — *-id*: удм. *šunıt* — кп. *šonıt* — кз. *šonid* ‘тёплый’, удм. *čurıt* — кп. *čorıt* — кз. *čorid* ‘твёрдый’, удм. *kurıt* — кп. *kurıt* — кз. *kurid* ‘горький’ и т. д. Отмечу, что верхнекамском (зюздинском) диалекте коми-пермяцкого языка, который во многих аспектах занимает промежуточное положение между коми-пермяцким и коми-зырянским языками, данный суффикс употребляется параллельно в обеих формах — с глухим и со звонким согласным — *-it/-id* (см. также ниже в данном разделе);

— удмуртскому отглагольному суффиксу имён *-et* в пермяцком соответствует суффикс *-et*, в зырянском — *-ed*: удм. *gožtet* — кп. *gižet* — кз. *gižed* ‘письмо’, удм. *pıket* — кп. *pıket* — кз. *pıked* ‘подпорка’, удм. *kerttet* — кп. *kerttet* — кз. *kerted* ‘завязка, привязь’ и т. п.;

— удмуртскому суффиксу лишительного падежа *-tek* в коми языках соответствует суффикс *-teg*: удм. *kitek* — кп., кз. *kiteg* ‘без руки’, удм. *ńańtek* — кп., кз. *ńańteg* ‘без хлеба’, удм. *purttek* — кп., кз. *purtteg* ‘без ножа’;

— соединительный (совместный) падеж коми-пермяцкого языка имеет суффикс *-ket*, которому в коми-зырянском соответствует *-ked*: кп. *oškęt* — кз. *oškęd* ‘с медведем’, кп. *kerkukęt* — кз. *kerkaked* ‘с домом’. Аналогично с упомянутым выше суффиксом *-it/-id* в верхнекамском диалекте коми-пермяцкого языка параллельно употребляются обе формы — и с глухим, и со звонким согласным — *-ket/-ked* (см. также ниже);

— интересная картина наблюдается в притяжательных суффиксах пермских языков. В удмуртском языке суффиксами, указывающими на принадлежность ко 2-му и 3-му лицу единственного числа существительных являются соответственно *-id/-ed* и *-iz/-ez* (в обоих случаях звонкие согласные), в коми-пермяцком — *-it* и *-is* (в обоих случаях глухие), в коми-зырянском — *-id* и *-is* (звонкий и глухой): удм. *kiid* — кп. *kiit* — кз. *kiid* ‘твоя рука’, удм. *kiiz* — кп., кз. *kiis* ‘его рука’, удм. *gurted* — кп. *gortit* — кз. *gortid* ‘твоя деревня’, удм.

*gurtez* — кп., кз. *gortis* ‘его деревня’. Следует заметить, что данные лично-притяжательные суффиксы имеют прафинноугорское происхождение и восходят к глухим праФУ \*-t (для 2-го лица ед. числа) и \*-s (для 3-го лица ед. числа) (ОФУЯ 1974: 273);

— такая ситуация повторяется при склонении пермских имён существительных с притяжательными суффиксами. В удмуртском языке суффиксами вступительного и местного падежей для существительных с принадлежностью 2-му и 3-му лицу единственного числа являются соответственно *-ad/-jad* и *-az/-jaz* (в обоих случаях звонкие), в пермяцком — *-at* и *-as* (в обоих случаях глухие), в зырянском — *-ad* и *-as* (звонкий и глухой): удм. *gurtad* — кп. *gortat* — кз. *gortad* ‘в твою деревню; в твоей деревне’, удм. *gurtaz* — кп., кз. *gortas* ‘в его деревню; в его деревне’;

— в верхнекамском (зюздинском) наречии коми-пермяцкого языка, занимающем промежуточное положение между коми-зырянским языком и остальными пермяцкими диалектами, вышеупомянутые именные суффиксы *-id*, *-ed*, *-ked*, *-ad* со звонким согласным *-d*, характерные для коми-зырянского языка, употребляются наряду с общепермяцкими формами с глухим согласным — *-it*, *-et*, *-ket*, *-at* (Баталова 1975: 222). И аналогичных примеров можно привести ещё немало.

Как видим, и здесь наблюдается разная степень озвончения согласных в пермских языках и их диалектах. Подобное положение вещей позволяет мне предполагать, что процесс формирования в пермских языках противопоставления звонких и глухих взрывных, сибиллянтов и аффрикат завершился не ранее в позднепрапермской эпохе (непосредственно накануне распада прапермской общности), шёл в разных прапермских диалектах неравномерно и закончился либо уже после её распада, либо в течение периода этого распада, который длился, по всей видимости, не одно столетие (Белых 2009). Таким образом, я бы датировал завершение данного фонетического процесса приблизительно концом I тыс. н.э. — первой половиной II тыс. н.э.

Здесь вполне логично возникает законный вопрос о том, как связать вышеизложенные лингвистические данные с историческими реалиями той эпохи.

Имея в виду всё вышесказанное, можно выстроить следующую цепочку рассуждений:

1. В прапермском языке имел место некий толчок к появлению в нём целой группы звонких согласных фонем, отсутствовавших в прафинно-угорском и прафинно-пермском языках и нетипичных для большинства современных финно-угорских языков.

2. Этот толчок мог быть вызван языковым влиянием извне.

3. Существует немалая вероятность, что это влияние напрямую связано с присутствием на территории пермской прародины или в непосредственной близости к ней в прапермскую эпоху ираноязычного населения, в языке которого были подобные звонкие согласные.

Почему именно иранские языки представляются мне наиболее вероятным источником данного воздействия на языки пермские? Прежде всего нужно сказать, что звонкие согласные имеют фонологический статус практически во всех ветвях индоевропейской языковой семьи, за исключением тохарской. Однако, среди этих языков только две их группы или ветви можно подозревать в роли такого рода источника — балто-славянские и иранские или, шире, индоиранские языки. (Пара)балто-славянский языковой компонент, проникший в Волго-Камье в середине I тыс. н.э. и оказавший определенное влияние на финно-угорские языки этого региона (Напольских 2015: 163–180), по моему мнению, вряд ли в полной мере подходит на данную роль.

Дело в том, что, по всей видимости, звонкие шумные стали проникать в язык предков пермян задолго до середины I тыс. н.э. через лексические заимствования из индоиранских языков (начиная с III тыс. до н.э. — см. напр. Напольских 2015: 133). Об этом

свидетельствует тот факт, что почти во всех разновременных индоиранских заимствованиях звонким согласным источника в пермских языках соответствуют также звонкие, а глухим — глухие (Лыткин 1975: 84–97). Полагаю, что именно под этим индоиранским, а позднее иранским (со II тыс. до н.э. — Напольских 2015: 133) воздействием в языке предков пермян постепенно звонкие шумные стали появляться на месте глухих и во многих словах исконного лексического фонда (общеуральских, общефинно-угорских, общефинно-пермских и т. д.). Ну а завершился этот процесс озвончения, как уже сказано, не ранее позднепрапермской эпохи, о чём подробно изложено выше в данном разделе.

Позволю себе предположить, что вышеупомянутый процесс в исторической фонетике пермских языков на завершающей своей стадии мог быть следствием достаточно интенсивных контактов пермян с носителями иранских языков сармато-аланского типа, в которых звонкие взрывные *b*, *g* и *d* (в том числе, и в начале слова), звонкие сибиллянты и звонкие аффрикаты, противопоставленные их глухим парам, вполне обычны, типичны и реконструируются для древних североиранских языков, как их именовал В. И. Абаев, не позднее скифской эпохи (ОИЯ: 323). Именно с носителями этих иранских языков контактировали предки пермян в позднепрапермскую эпоху.

Мне остаётся только добавить, что идея об иранском влиянии на развитие звонких согласных в пермских языках осторожно высказывалась В. И. Лыткиным и Т. И. Тепляшиной ещё в середине 1970-х гг. (ОФУЯ 1976: 138).

Замечу, что среди финно-угорских языков, помимо пермских, озвончение *\*p*- и *\*t*- в начале слова и превращение их соответственно в *b*- и *d*-, а также само появление анлаутных звонких взрывных (*b*-, *d*-, *g*-), развитие противопоставления глухих и звонких согласных в анлауте за счёт появления и широкого распространения звонких смычных, сибиллянтов и аффрикат имело место ещё и в истории венгерского языка (ОФУЯ 1974: 119; ОФУЯ 1976: 380; Хайду 1985: 267–268, 268–269), предки носителей которого, также как и древние пермяне, в течение своей истории активно контактировали с иранцами. Во всяком случае, можно с уверенностью говорить, что среди финно-угорских языков именно в пермских языках и в венгерском эти звонкие согласные фонемы употребляются наиболее активно, а противопоставление звонких и глухих согласных выглядит сформированным наиболее последовательно.

В финно-угорском языкознании интерпретация и датировка данного фонетического явления и некоторых других общих тенденций фонетического развития в венгерском и пермских языках (см. раздел 1.2 настоящей статьи) остаётся предметом дискуссий. Так, например, Э. Итконен и К. Редечи полагали, что эти общие тенденции фонетического развития сформировались в венгерском и пермских языках независимо друг от друга и не были следствием вторичных контактов носителей правенгерского и прапермского языков (Rédei 1964). Данная точка зрения представляется мне по меньшей мере сомнительной, т. к. выглядит совершенно невероятным, что многие не только фонетические, но и морфологические и лексические схождения и параллели сформировались отдельно и независимо в пермских и венгерском языках. Гораздо более обоснованным и убедительным выглядит мнение, что хотя бы часть из них является следствием достаточно интенсивных вторичных контактов предков пермян и венгров (см. Хелимский 1982: 19–20; Напольских, Смирнов 2022: 622–623; а также разделы 1.1, 1.2, 2.2, 2.5 настоящей статьи).

Выглядит не вполне убедительным, хотя и заслуживающим внимания, мнение Т. Э. Уотилы и К. Редечи о том, что озвончение анлаутных смычных, сибиллянтов и аффрикат в пермских языках связано с ассимилятивным воздействием на них последующих звонких согласных слова. В качестве примеров они приводят: удм., коми *bad'* 'ива' при фин. *raju* 'тж.'; кз. *bajdeg* 'куропатка' при венг. *fajd* 'Auerhahn'; удм. *diń* 'комель (дерева);

основание, фундамент', кз. *din*, кп. *din* 'комель, кочерыжка' при фин. *tyvi* 'комель, основание' и др. (Rédei 1964: 255, 256). Я считаю высказанное Т. Э. Утилой и К. Редееи суждение не вполне убедительным по следующим причинам:

Во-первых, существует немало примеров пермских лексем, где звонкая согласная середины или конца слова не оказала никакого ассимилятивного влияния на анлаутную глухую согласную: удм. *kiz*, кз. *koz*, кп. *kez* 'ель' < праур. \**kuse* ~ \**kose* 'Fichte, Tanne' (КЭСК: 127; UEW: 222); удм. *kur* 'луб, кора', кз. *kor* 'кожура, шелуха, луб' < праФУ \**kere* 'Rinde' (КЭСК: 133; UEW: 148–149); удм. *palez*, *palez*, коми *pelis*, диал. *peliz* 'рябина' < праФУ \**picla* 'Vogelbeere, Vogelbeerebaum' (КЭСК: 218; UEW: 376); удм., коми *pel* 'ухо' < праФУ \**peljä* 'Ohr' (КЭСК: 218; UEW: 370); удм. *todini*, коми *tedni* 'знать' < праФУ \**tumte-* 'fühlen, anrühren, tasten, betasten' (КЭСК: 283; UEW: 536) и др.

Во-вторых, с другой стороны, в пермских языках имеются примеры, когда анлаутная согласная озвончилась, несмотря на то, что за ней в составе слова следовала глухая: удм. *biškinj*, диал. *bičkinj*, коми *bičkinj* 'колоть, уколоть' < праФП < \**pičkz* '(durch)stechen, stoßen' (КЭСК: 45; UEW: 739); удм. *gor* 'яма, ямка, углубление; ложбина', коми *ger* 'углубление, ямка, лунка; лужа; маленькое озеро' (КЭСК: 80) (в UEW удивительным образом эти пермские слова никак не связываются с фин. *kuorra* 'яма, ямка, впадина' — пермских слов в соответствующей словарной статье просто нет — и соответственно для них не предлагается более древняя этимология); коми *gos* 'жир, сало', удм. диал. *gusjanj* 'накормить, подкормить; закусить, немного поест' (КЭСК: 79), удм. *žoskit* 'узкий, тесный', кз. *žeskid*, кп. *žeskj* 'тесный' (КЭСК: 90) и др.

В-третьих, мною уже были перечислены некоторые общепермские лексемы, в которых звонкая анлаутная предшествует последующей глухой согласной в одних пермских языках/диалектах, а в других в анлауте имеет место глухая согласная; или наоборот, в одних пермских языках/диалектах анлаутная глухая предшествует последующей звонкой согласной, а в других в анлауте в той же лексеме имеется звонкая согласная (см. выше в этом же разделе примеры с *bol'* ~ *pol'* ~ *pul'i*, *gut* ~ *kut*, *geb* ~ *kibi*, *goz* ~ *kuz* и др.).

В то же время, мне представляется, что данные факты являются важным подтверждением того, что озвончение анлаутных согласных в прапермском языке нашло своё завершение, как уже было сказано, сравнительно недавно, уже в позднепрапермскую эпоху, с чем, кстати говоря, соглашался и К. Редееи (Rédei 1964: 256).

Следует сказать, что в независимости от того, считаем мы мнение Т. Э. Утилы и К. Редееи верным или нет, это никак не мешает нам предположить, что сам факт появления звонких взрывных, сибилантов и аффрикат и их противопоставления своим глухим парам в прапермском языке вполне может быть вызван (индо)иранским языковым и культурным влиянием на пермян. В конце концов, тот же Т. Э. Утила считал, что звонкие смычные в начале слова существовали ещё в раннепрапермском языке, ссылаясь на такие арийские заимствования в пермских языках, как удм. *duri*, коми *dar* 'половник, разливательная ложка' (ср. санс. *dārvi-* 'ложка'), удм., коми *das* 'десять' (ср. осет. *dæs* 'тж.') (Rédei 1964: 255). Кстати говоря, по крайней мере последнее заимствование вполне могло попасть к пермянам уже в позднепрапермский период, о чём свидетельствует очень близкая фонетически осетинская параллель, но речь сейчас не об этом. Подобные примеры как раз и могут косвенно подтверждать, что именно проникновение в прапермский язык таких заимствований из древних (индо)иранских языков стимулировало возникновение в нём звонких смычных, сибилантов и аффрикат.

1.2. Ещё одной яркой отличительной особенностью фонетики пермских языков является деназализация, произошедшая ещё в общепермскую эпоху в прафинно-угорских

сочетаниях согласных \*-mp-, \*-nt-, \*-nk-, \*-ŋk-, \*-nč-, \*-nč-. В данных сочетаниях носовой согласный (*m*, *n* или *ŋ*) в пермских языках исчез, а второй компонент озвончился. Примеры:

— удм. *podini* ‘закрывать (напр. окна, двери); придавить, прижать, прищемить’, кп. *pednavni* ‘закрывать, затворять’, кз. нв., уд. *pednavni* ‘закрывать крышкой; закрыть (печную трубу)’ (ССКЗД: 41, 296), венг. *fedni* ‘закрывать’; при хант. *pent-* ‘закрывать, прикрывать’, манс. *pänt-*, *pant-* ‘закрывать’ и др. < праФУ \**pentz-* ‘zumachen, decken’ (ОФУЯ 1974: 138; UEW: 371);

— удм. *udini*, коми *udni* ‘напоить, дать пить’, венг. *adni* ‘дать’; при фин. *anta-* ‘дать’; мар. *omdaš* ‘скапливаться, скопиться (о молоке в вымени)’ и др. < праФУ \**amta-* ‘geben’ (ОФУЯ 1974: 139; UEW: 8);

— удм. *todini*, коми *tedni* ‘знать, узнать’, венг. *tudni* ‘знать; уметь; мочь’; при фин. *tunte-* ‘чувствовать, ощущать; знать, узнавать’, эст. *tunde-* ‘тж.’ и др. < праФУ \**tumte-* ‘fühlen, annrühren, tasten, betasten’ (ОФУЯ 1974: 139; UEW: 536);

— удм. *vižini*, диал. *vižini* ‘перейти, переправиться, переехать’, удм. *viž*, диал. *viž* ‘мост; пол’, коми *vižni* ‘перейти, переправиться’; при мар. *wončaiš* ‘перейти, переехать, переправиться’, манс. *onš-*, *wunš-*, *unš-* ‘тж.’ и др. < праФУ \**wanča-* ‘überschreiten’ (ОФУЯ 1976: 138; UEW: 557);

— удм. *ibini* ‘стрелять’; при фин. *amria* ‘стрелять’, эст. *ambi* ‘лук (оружие)’ и др. < праФУ \**amrз-* ‘werfen, schießen’ (ОФУЯ 1974: 139; UEW: 606) и т. п.

Важно заметить, что, как мы видим из первых трёх примеров, подобная деназализация развилась и в венгерском языке (см. ОФУЯ 1974: 137–139; ОФУЯ 1976: 380; Хелимский 1982: 19; Хайду 1985: 281–283), что может свидетельствовать о сепаратной пермско-венгерской фонетической инновации, а это, учитывая неблизкородственность пермских языков с венгерским, наводит на мысль о возможном одновременном влиянии на них извне, которое стало причиной этой деназализации. Что же могло быть источником такого влияния?

В этой связи обращает на себя внимание фонетическое явление регрессивной ассимиляции согласных в иронском диалекте осетинского языка, которое несколько напоминает пермско-венгерскую деназализацию, рассмотренную выше. В иронском диалекте в некоторых сочетаниях согласных имеет место исчезновение носовой согласной фонемы и уподобление её последующей согласной или (реже и только в исходе слова) просто выпадение носовой. В итоге произошла потеря носового, т.е. деназализация в следующих группах согласных иронского диалекта: \**nd* > *dd*, \**ns* > *ss/s*, \**nx* > *xx/x*, \**nğ* > *ğğ*, \**nv* > *vv* и некоторых других. Примеры:

- ирон. *æddæ* при диг. *ændæ* ‘снаружи, вне’;
- ирон. *xussar* при диг. *xonsar* ‘юг’;
- ирон. *fyssyn* при диг. *finsun* ‘писать’;
- ирон. *bos* при диг. *bons* ‘перевязь’;
- ирон. *xox* при диг. *xonx* ‘гора’;
- ирон. *æxxiys* при диг. *ænxis* ‘помощь’;
- ирон. *æğğæd* при диг. *ænğæd* ‘довольно’;
- ирон. *ævvaхs* при диг. *ænvaxs* ‘близко’ и т. п. (Абаев 1949: 380).

К сожалению, мне не удалось отыскать какие-либо сведения о датировке этого иронского фонетического сдвига. Однако, на принципиальную возможность достаточно древних контактов пермских языков с носителями алаано-осетинского языка именно иронского типа указывают такие лексические заимствования как удм. *badžim* ~ *bažžim* ‘большой, крупный’ < осет. ирон. *bæžžyn* ‘толстый, густой’, при дигор. *bæzgin* ‘тж.’, где в иронском варианте праосетинская звонкая смычная -g- закономерно превратилась перед гласной -y- в звонкую аффрикату -ž- (Абаев 1949: 376), что и отразилось в фонетике удмуртского слова.

Строго говоря, между пермско-венгерской деназализацией и осетинской (иронской) регрессивной ассимиляцией не так уж много общего. Собственно, в обоих случаях общим является только факт исчезновения первой (носовой) согласной фонемы в сочетании двух фонем. Однако, в пермских и венгерском языках вторая согласная сочетания озвончилась, а в иронском — осталась без изменений. Поэтому, настаивая на обязательном иранском влиянии на пермско-венгерскую деназализацию, я всё же считаю возможным обратить внимание на данное сходство и очень осторожно предполагаю, что происходившая в древнеосетинском (древнеиронском) ассимиляция согласных, могла оказать влияние на процесс деназализации в прапермском и правенгерском. Вкупе же с проходившим параллельно процессом озвончения глухих согласных (см. выше, раздел 1.1.) этот процесс мог привести к тому, что вторая согласная сочетания в пермских языках и в венгерском стала звонкой.

Замечу, что, например, К. Редеев и П. Хайду, отметив поразительное сходство озвончения согласных и деназализации их сочетаний в венгерском и пермских языках, всё же считали эти процессы произошедшими в разное время и независимо в правенгерском и прапермском (см. напр. ОФУЯ 1974: 137; Хайду 1985: 282–283). Впрочем, другие исследователи полагали возможным считать эти общие тенденции развития фонетики венгерского и пермских языков обусловленными тесными вторичными контактами предков венгров и пермян (см. напр. Хелимский 1982: 19–20). На данный момент я полагаю достаточным указать на имеющийся параллелизм, не углубляясь в рассмотрение проблемы и не отдавая предпочтения какой-либо из возможных интерпретаций.

## 2. Морфология

Вероятные следы иранского влияния на пермские языки также могут быть зафиксированы и в морфологических системах удмуртского и коми языков.

2.1. Удмуртский аффикс прилагательных и существительных *-eg/-ek*, а также родственный ему коми аффикс *-eg/-og*, в словах типа удм. *žazeg* ‘туть’ (при кз. *žožeg*, кп. *žožo* ‘туть’), лит. *žeg*, диал. *žizeg* ‘рожь’ (при коми *ružeg* ‘тж.’), *lapreg* ‘низкий, невысокий’ (ср. *lap* ‘тж.’), *puteg* ‘трещина, щель’ (ср. *put koškini*, *putini* ‘треснуть, расколоться’), *šuleg/šubeg* ‘узкий’ (развёрнутый список подобных слов в удмуртском языке см. Алатырев 1976: 124–129) и т.п. По всей видимости, сюда же следует отнести аналогичный аффикс, представленный в общепермских дохристианских личных именах, типа удм. *Šudeg* / коми *Šudeg* удм. *Kužeg*, *Mušeg/Mušek/Mošeg*, *Mijeg/Mojeg*, *Ožeg*, *Ožmeg*, *Poreg*, *Užeg*, *Urseg*, *Ćudeg* и т. п. (см. Атаманов 1989: 169–181).

Часть финноугроведов предполагала для этих аффиксов прауральское происхождение (Атаманов 1989: 170). Однако нельзя не заметить, что данная морфема внешне очень напоминает обильно представленные в дошедших до нас скифских именах суффиксы *-ag* (< *-ak(a)*) и аналогичные им осетинские суффиксы существительных и прилагательных *-ag/-æg*.

В скифском данные суффиксы могли присоединяться как к глагольной основе, образуя отглагольные имена, так и к именной. Аналогичная ситуация наблюдается и в современном осетинском языке (Камболов 2006: 109–110). Нечто похожее мы видим и в пермских языках. Сравните, например, осет. *kærdæg* ‘косящий, косарь’, ‘косимое’ > ‘трава’ (от *kærdyn/kærdun* ‘косить’), *xudæg/xodæg* ‘смеющийся’, ‘смешной’ (от *xudyn/xodun* ‘смеяться’), *žumæg/zumæg* ‘зима’, *næwæg* ‘новый’ и удм. *puteg* ‘трещина, щель’ (от *putini* ‘треснуть, расколоться’), *kureg* ‘курица’, *lapreg* ‘низкий, невысокий’, *kušæg* ‘мелкий, неглубокий’, кз. *bidmeg* ‘растение’ (от *bidmini* ‘расти’), *otseg* ‘помощь’ (от *otsavni* ‘помогать’), *kureg* ‘курица’.

Считается, что осетинские суффиксы *-ag/-æg* имеют праиранское происхождение и восходят соответственно к ир. *\*-āka/\*-aka* (Камболов 2006: 381). И в древнем скифском, и в современном осетинском языках вышеуказанные суффиксы выполняли также функцию образования личных имён: скиф. *Aspak < aspa* ‘лошадь’, *Sanag < sana* ‘вино’, осет. *Kwyzæg < kwyz* ‘собака’ и т. п. (Камболов 2006: 110; Абаев 1959: 110). Такую же роль играл аффикс *-eg/-ek/-eg* в пермских языках, служа суффиксом, образующим личные мужские имена (см. выше и ниже в данном разделе).

По мнению Т. Т. Камболова, скифский (*-ag/-ak*) и осетинский (*-ag*) суффиксы придают корневой основе имён следующие значения: ‘происходящий откуда-то’, ‘принадлежащий к чему-то’, ‘предназначенный стать чем-либо’, ‘пригодный к чему-либо’, ‘достойный чего-либо’ (Камболов 2006: 111).

По В. И. Абаеву, будучи присоединённым к глагольной основе, суффикс *-æg* образует чаще всего причастия активного действия, со значением действующего лица: осет. *fyssyn* ‘писать’ — *fyssæg* ‘пишущий, писатель’, *kusyn* ‘работать’ — *kusæg* ‘рабочий, работник’ или причастия настоящего времени со значением имени действия: *zarun* ‘петь’ — *zaræg* ‘песня’, *хиуфун* ‘кашлять’ — *хиуфæg* ‘кашель’ и т. д.

В подобных случаях суффикс *-ag* образует в осетинском языке отглагольные формы, означающие постоянное свойство или склонность к какому-либо действию: *wajun* ‘скакать, бегать’ — *wajag* ‘скакун; быстрый’, *nwazun* ‘пить’ — *nwazag* ‘пьяница’, *kusun* ‘работать’ — *kusag* ‘работающий, трудолюбивый’ и т. п.

Присоединяясь же к именным основам суффикс *-ag* образует формы, указывающие на происхождение, принадлежность к чему-либо, либо предназначение: *wyryssag* ‘русский’ (*wyrys* ‘русские’), *хохаг* ‘горец’ (*хох* ‘гора’), *тојаг* ‘будущий муж’ (*тој* ‘муж’), *араqqag* ‘сусло для водки’ (*араqq* ‘водка’). Суффикс *-æg* сравнительно редко присоединяется к именным основам и придаёт им довольно схожее значение: *тæгwyræg* ‘бедняга’ (*тæгwyr* ‘бедный’), *galiwæg* ‘левша’ (*galiw* ‘левый’), *æxsæræg* ‘белка’ (*æxsær* ‘орех’) (Абаев 1959: 63–64, 108–110).

Грамматики удмуртского и коми-пермяцкого языков никак не объясняют значение вышеуказанных пермских суффиксов, а в грамматике коми-зырянского языка о суффиксе *-eg* говорится, что его значение трудно установить (СКЯ: 157). По не очень чётко сформулированному суждению В. И. Алатырева, данный суффикс в удмуртском языке означает некий признак, «присущий предмету или явлению, выраженному производящей основой» (Алатырев 1976: 124). Однако, в свете данных осетинского языка многие пермские апеллятивы на *-eg/-eg*, а также общепермские дохристианские антропонимы получают вполне логичную и довольно ясную этимологию. Например:

— удм. *val’eg* ‘скользящий, гладкий’ < *val’-val’* выражение, передающее зрительно-осознательное ощущение очень гладкой и скользкой поверхности (Алатырев 1976: 124);

— удм. *kuśeg* ‘мелкий, неглубокий’ < *kuś* ‘тж.’, при коми *koś* ‘порог, перекат; отмель’ (Алатырев 1976: 125);

— удм. *Asleg* < праперм. *\*asil* ‘утро’, ср. удм. сев. *asil*, коми *asil/asiv* ‘тж.’ (Атаманов 1989: 171), т. е. данное имя буквально означало, вероятно, ‘рождённый утром, принадлежащий утру, происходящий от утра’;

— удм. *Ižeg* < *ižinj* ‘спать’ (Атаманов 1989: 173), т. е. ‘любящий поспать, соня’;

— удм. *Kužeg* < праперм. *\*kuž* ‘длинный, высокий’ (Атаманов 1989: 173);

— удм. *Poreg* < *por* ‘мариец, марийский’ (Атаманов 1989: 175);

— удм. *Šudeg*, коми *Šudeg* < праперм. *\*šud* ‘счастье’ (Атаманов 1989: 178), т. е. ‘счастливым, обладающий счастьем’, ‘предназначенный для счастья’ или ‘происходящий от счастья’ и т. д.



Следует полагать, что данные суффиксы проникли в прапермский язык в составе алаано-осетинских лексических заимствований, вроде др.-перм. *ideg* ‘ангел’ < осет. *idawæg* ‘божество’ (Joki 1973: 264), а впоследствии распространились и на прочую лексику пермских языков.

2.2. Общепермская форма инфинитива на *-ni* (удм. *miñin* [1], коми *miñin* [1] ‘идти’, удм. *užan* [1], коми *užavn* [1] ‘работать’ и т.д.). В некоторых удмуртских диалектах инфинитив образуется без конечной *-i* (*miñin*, *užan*) (Грамматика 1962: 255; Кельмаков 1998: 153–154). Вполне возможна связь с осетинским показателем инфинитива *-yn/-un* (ирон. *sæwun*, диг. *sæun* ‘идти, ехать’; ирон. *kusun*, диг. *kosun* ‘работать’). Осетинский показатель инфинитива имеет очень древнее, возможно общеарийское происхождение, и вероятно восходит к ар. *\*-una* (Камболов 2006: 382).

См. также явно сближающийся с пермским венгерский аффикс инфинитива *-ni* (*menni* ‘идти, ехать’, *hinni* ‘верить’ и т. д.). Здесь важно отметить, что данная форма инфинитива — это сепаратная пермско-венгерская параллель, которая не находит никаких соответствий в прочих финно-угорских языках. Если же учесть, что удмуртский и коми языки, с одной стороны, и венгерский, с другой — это весьма дальние «родственники», то сам факт наличия такой сепаратной параллели наводит на мысль, что мы, вероятно, имеем дело со сравнительно поздней пермско-венгерской инновацией, причиной возникновения которой вполне могло быть заимствование из иранского источника.

В финно-угорском языкознании общепермский (а также венгерский) суффикс инфинитива традиционно принято считать возникшим из сочетания прафинно-угорского (или даже прауральского) суффикса нефинитных форм глагола *\*-n* с суффиксом латива *\*-k* (впоследствии отпавшего): праФУ *\*-nik* > удм., коми *-ni*, венг. *-ni* (ОФУЯ 1974: 346; ОФУЯ 1976: 187; Хайду 1985: 333; Csúcs 2005: 278).

Не отрицая такую возможность, всё же замечу, что данное толкование никак не объясняет совершенно поразительное формальное и функциональное сходство пермского суффикса инфинитива с венгерским. На мой взгляд, достойного объяснения этот примечательный факт не получил в уралистике до сих пор. Например, я не могу согласиться с мнением Б. А. Серебренникова о том, что суффиксы инфинитива пермских языков связаны с аналогичным суффиксом венгерского языка генетически (Серебренников 1963: 314), т. е. восходят, как минимум, к прафинно-угорской эпохе. Такое предположение выглядит очень маловероятным, учитывая, что, во-первых, пермские языки связаны с венгерским весьма дальними родственными связями, а, во-вторых, ни в обско-угорских, ни в финно-волжских языках, которые являются гораздо более близкими «родственниками» соответственно венгерского и пермских языков, мы не видим никаких хотя бы отдалённо похожих форм инфинитива. Также вряд ли приемлемой мне представляется идея К. Редее о том, что удивительное сходство общепермской и венгерской форм есть не более чем чистая случайность (см. напр. Csúcs 2005: 279). С точкой зрения К. Редее тем более трудно согласиться, если вспомнить, что такими же случайными совпадениями этот учёный считал и чрезвычайно похожие процессы развития звонких согласных, а также денанализацию групп согласных в пермских и венгерском языках (см. выше, разделы 1.1 и 1.2 настоящей статьи). Не слишком ли много «чистых случайностей»?

Всё вышесказанное позволяет мне заключить, что, скорее всего, формальное и функциональное сходство пермского и венгерского инфинитивов является не следствием генетического родства этих языков, а результатом вторичных контактов носителей прапермского и правенгерского языков, которые, как принято считать, имели место в Прикамье в эпоху средневековья, начиная со второй половины I тыс. н.э. Поэтому, не отме-

тая сходу возможность прафинно-угорского (прауральского) происхождения данного суффикса инфинитива, я всё же взял бы на себя смелость предложить альтернативный его вариант — происхождение из иранского источника.

2.3. Удмуртский аффикс отглагольных образований *-on/-n* (типа *jaraton, užan*). При помощи этого суффикса от любого глагола образуются существительные, которые можно можно характеризовать как имя действия: *jaratini* ‘любить’ — *jaraton* ‘любовь’, *vijini* ‘убить’ — *vijon* ‘убийство’, *užani* ‘работать’ — *užan* ‘работа’, *šijini* ‘есть, кушать’ — *šijon* ‘еда’, *jujini* ‘пить’ — *juon* ‘питьё’ и т. д. (Грамматика 1962: 111–114). В некоторых случаях существительные, образованные от данного суффикса, имеют двойную семантику, означая одновременно не только определённое действие или процесс, но и предмет, связанный своим назначением с данным действием. Например: *pukon* ‘сидение (действие, состояние)’ и ‘стул (предмет, связанный с сидением)’ < *pukijini* ‘сидеть’, *šulan* ‘свист, насвистывание (действие)’ и ‘свисток (предмет, предназначенный для свиста)’ < *šulani* ‘свистеть’, *juon* ‘питьё, выпивание, выпивка’ и ‘питьё, напиток’ < *jujini* ‘пить’ и т. д. (Грамматика 1962: 113).

Следует также отметить, что в ряде случаев данная форма может быть переведена не как имя существительное, а как прилагательное или причастие, образованное от глагола и указывающее на определённое качество того или иного предмета или объекта. Например, слово *jaraton* (< *jaratini* «любить») может быть переведено не только как ‘любовь’, но и как ‘любимый’; *ulon* (< *ulijini* ‘жить’) — ‘жизнь’ и ‘жилой, жизненный, житейский’. Ср. *jun jaraton* ‘крепкая любовь’ (существительное) и *jaraton pije* ‘(мой) любимый сын’ (прилагательное); *šudo ulon* ‘счастливая жизнь’ (существительное) и *ulon inti* ‘жилое место’ (прилагательное).

В коми языках удмуртскому *-on/-n* соответствует аффикс *-an*, посредством которого от глаголов образуются существительные, обозначающие орудие или объект действия: *vartan* ‘цеп’ — *vartni* ‘молотить’, *sjnan* ‘гребень’ — *sjnavni* ‘чесать, расчёсывать’, *juan* ‘питьё’ — *junj* ‘пить’, *šojan* ‘еда’ — *šojni* ‘есть’ и т. п. (СКЯ: 154–155, КПЯ: 205–206).

Интересно, что, в коми-пермяцком языке, а также, в некоторых случаях, в лузско-летском и верхнесысольском диалектах коми-зырянского, аффикс отглагольных существительных *-an*, помимо значения орудия или объекта действия, может иметь и значение самого процесса действия или состояния. Например: кп. *olan* ‘жизнь, житьё’ < *olnj/ovni* ‘жить’, *vundan* ‘жатва’ < *vundjini* ‘жать’, лл., вс., *višan* ‘болезнь’ < *višni* ‘болеть’ (ССКЗД: 475).

Кроме того, согласно грамматике коми-пермяцкого языка, при помощи данного суффикса от глаголов образуются прилагательные или причастия (отглагольные прилагательные), обозначающие признак по действию (состоянию), производимому предметами или объектами: *bjđman pu* ‘растущее дерево’ (< *bjđmijni* ‘расти’), *vartan mašina* ‘молотилка, молотильная машина’ (*vartni* ‘молотить’) и т. д. (КПЯ: 222, 279–280).

В коми-зырянских грамматиках формы на *-an* трактуются как причастия (отглагольные прилагательные) совершаемого над предметом или объектом действия, напр.: *gižan pišme* ‘писомое письмо; письмо, которое пишется’ (< *gižni* ‘писать’); *veleđan mort* ‘обучаемый человек’ (< *veleđni* ‘учить, обучать’) (СКЯ: 242).

Разумеется, здесь следует иметь в виду, что грамматические системы русского и пермских языков серьёзно различаются, терминология, вполне подходящая и логичная для названий частей речи русского языка, вряд ли может в той же мере годиться для удмуртского и коми языков. Тем не менее, как бы мы не именовали вышеуказанные формы — существительными, прилагательными или причастиями — одно мы можем утверждать уверенно: все эти образования являются отглагольными, отображающими те или иные качества предметов или объектов, связанные с соответствующими глаголами.

В финно-угроведческой литературе пермские имена на *-n/-on/-an* принято считать по своему происхождению субстантивированными причастиями (Серебрянников 1963: 297; ОФУЯ 1976: 150; Csúcs 2005: 280).

В этой связи привлекает внимание осетинский суффикс отглагольных и отымённых образований *-on* и соответствующий ему древнескифский суффикс *-ān*, которые восходят к ир. *\*-āna* (Камболов 2006: 111, 382). В осетинских отглагольных образованиях при помощи суффикса *-on* образуются причастия настоящего времени, чаще пассивные, реже активные: осет. *baron* ‘прощающий’ < *bar-* ‘прощать’, *waržon/warzon* ‘любимый’ < *warž-/warz-* ‘любить’ и т. п. (Камболов 2006: 111). Любопытно, что в ряде случаев, подобно пермским языкам, в осетинском данные отглагольные образования могут переводиться как существительные, обозначающие соответствующее действие: *fændon* ‘желание’ < *fændyn* ‘хотеть’, *kænon* ‘поведение’ < *kænyn* ‘делать, творить’ (Абаев 1959: 112).

В. И. Абаев отмечал, что отглагольные образования на *-on* в осетинском языке довольно редки, чаще данный суффикс присоединяется к именным основам (Абаев 1959: 112). Однако, здесь следует упомянуть и о другом суффиксе осетинского языка — *-æn* — который тоже может присоединяться к глагольным (чаще) и именным (редко) основам и образует слова, довольно близкие по смыслу к словам на *-on* (Абаев 1959: 110–112). Форма на *-æn* также может иметь отношение к вышеуказанным грамматическим формам пермских языков. Данный осетинский суффикс, как и суффикс *-on*, имеет праиранское происхождение и восходит к ир. *\*-ana* (Камболов 2006: 382).

По В. И. Абаеву, в отглагольных образованиях суффикс *-æn* указывает, в частности, на:

а) орудие действия — *k'axæn* ‘кирка’ < *k'axyn* ‘копать’, *qazæn* ‘игрушка’ < *qazyn* ‘играть’, *wasæn* ‘свисток’ < *wasyn* ‘свистеть’ (Абаев 1959: 110). Ср. напр. удм. *šudon* ‘игрушка’ < *šudini* ‘играть’, *šulan* ‘свисток, свистулька’ < *šulani* ‘свистеть’, кз. *čipsan* ‘свисток; манок, пищик’ < *čipsini* ‘свистеть, пищать, щебетать’.

б) место действия — *badæn* ‘сиденье, место для сидения’ < *badyn* ‘сидеть’, *xuyssæn* ‘ложе’ < *xuyssyn* ‘спать’, *xinajæn* ‘купальня’ < *xi najyn* ‘купаться’ (Абаев 1959: 111). Ср. удм. *pukon* ‘стул’ < *pukini* ‘сидеть’, *pilaškon* ‘умывальник’ < *pilaškini* ‘мыться, умываться, купаться’.

в) само действие или его результат — *t'yssæntæ* ‘запасы’ (где *-tæ* — аффикс множественного числа) < *t'yssyn* ‘совать, вкладывать’, *axodæn* ‘завтрак’ < *axodyn* ‘есть, закусывать’ (Абаев 1959: 112). Ср. удм. *dirton* ‘спешка, торопливость; гонка’ < *dirtini* ‘спешить, торопиться’, *nullon* ‘ношение, носка’ < *nullini* ‘носить’, *nunažejan* ‘обед’ < *nunažejani* ‘обедать’.

Как видим, в удмуртском языке такие значения характерны для очень многих отглагольных образований на *-on/-n*, а значение самого действия — практически для всех них. Так, например, вышеупомянутое удм. *šudon* имеет значение не только ‘игрушка’, но и ‘игра’, *šulan* — не только ‘свисток’, но и ‘свист, насвистывание’, *pukon* — не только ‘стул’, но и ‘сидение (процесс)’, *pilaškon* — ‘умывальник’ и ‘мытьё, умывание, купание’ и т. п. В коми языках так же, хотя и в несколько меньшей степени, подобными значениямиобладают отглагольные формы на *-an* (см. выше в данном разделе).

В фундаментальных работах по общему финно-угроведению нет единого устоявшегося мнения о происхождении пермского суффикса *-n/-on/-an* (ОФУЯ 1976: 150, 153). Чаще всего его считают производным от суффикса инфинитива (Серебрянников 1963: 296; Csúcs 2005: 280) и при этом сопоставляют с марийским деепричастным суффиксом *-n/-en/-æn*, мордовским причастным суффиксом *-ń*, мансийским суффиксом причастия настоящего времени *-nə* (Серебрянников 1963: 295–298). Однако, с моей точки зрения, именно с осетинскими суффиксами пермские проявляют наибольшую фонетическую и функциональную близость. Это обстоятельство позволяет мне считать данную альтернативу по крайней мере не менее привлекательной и приемлемой.

2.4. В пермских языках существует деепричастие (отглагольное наречие), указывающее на действие, протекающее одновременно с основным действием. В коми-зырянском языке, в зависимости от конкретного диалекта, такие деепричастия образуются путём присоединения к глагольной основе суффикса *-igen/-igin/-igen/-ige/-ig*. В диалектах коми-пермяцкого языка данному суффиксу соответствует *-ige/-ike/-ik*. В удмуртском языке коми формам соответствует деепричастие на *-ku/-ki*. Более распространённый сегодня фонетический вариант *-ki* принято считать более поздним, возникшим от варианта *-ki* по ассоциации с удмуртским вопросительным местоимением *ku?* ‘когда?’ (Кельмаков 1998: 160).

Примеры: кз. *tunigen*, кп. *tunike*, удм. *tiñiku* ‘идучи, во время ходьбы, когда шёл’ (от коми *tunñi*, удм. *tiñiñi* ‘идти, ходить’); кз. *šorñitigen* ‘разговаривая, во время разговора’ < *šorñitñi* ‘разговаривать’, кп. *gerike* ‘распахивая, во время пахоты’ < *gerni* ‘пахать’, удм. *užaku* ‘работая, во время работы’ < *užani* ‘работать’ (см. ССКЗД: 486; СКЯ: 244; КПЯ: 281; Баталова 1975: 191; Грамматика 1962: 283–287; Кельмаков 1998: 159–160).

Мне представляется, что упомянутые пермские суффиксы (особенно их коми-зырянские варианты) фонетически и по своему значению имеют некоторое сходство с осетинским суффиксом прилагательных и наречий *-ykkon/-ygon/-igon*: ирон. *zponykkon* ‘вчерашний’ < *zpon* ‘вчера’, *bonuygon* ‘днём’ < *bon* ‘день’, *særdygon* ‘летом’ < *særd* ‘лето’, диг. *sæumigon* ‘утром’ < *sæumæ* ‘утро’, *iuxattigon* ‘давешний’ < *iu xatt* ‘однажды’ и т. п. (Абаев 1949: 425; Абаев 1959: 113). Основную проблему здесь составляет то важное различие, что в пермских языках вышеуказанные суффиксы могут присоединяться исключительно к глагольным основам, а осетинские суффиксы, по всей видимости, тяготеют к основам именным, а не к глагольным.

Тем не менее, существует по крайней мере два обстоятельства, которые позволяют мне с известной осторожностью всё же указать на данную возможную параллель. Во-первых, по мнению В. И. Абаева, малопродуктивный в современном осетинском языке суффикс *-ykkon/-ygon/-igon* представляет собой осложнённый вариант уже рассмотренного выше (см. раздел 2.3) суффикса *-on* (Абаев 1959: 113), а тот, как уже говорилось, может присоединяться как к именным, так и к глагольным основам. Поэтому можно предположить, что данная осложнённая форма в прошлом была более продуктивна и могла присоединяться также и к глагольным основам. Во-вторых, довольно близкими выглядят значения пермских и осетинских суффиксов, придаваемые ими тем корневым основам, к коим они присоединяются. Речь здесь практически всегда идёт о каком-то временном периоде, отрезке времени, когда происходило то или иное событие, действие и т. д.

Ср., например: кз. *Řednej zavod jilís dumajtigen šelemislj veli kokñi*. ‘Думая о родном заводе, сердцу было легко’, т. е. букв. ‘**Когда он(а) думал(а)** о родном заводе, его/её сердцу было легко’ (СКЯ: 244); осет. *Асу qædy læg bonuygon dær fæzæğæl wuzæp*. ‘В этом лесу человек и **днём** заблудится’ (Грамматика 1963: 215).

В финно-угроведении вопрос происхождения данных пермских суффиксов деепричастий считается сложной проблемой, т. к. в родственных финно-угорских языках очевидных параллелей им нет (Серебренников 1963: 302; Csúcs 2005: 284). Такое положение вещей заставляет обратить на осетинские параллели ещё большее внимание.

2.5. Довольно продуктивен в удмуртском языке аффикс *-et*, при помощи которого от глаголов образуются существительные, которые обозначают:

а) орудие действия или предмет, имеющий определённое назначение: *piket* ‘подпорка’ < *pikini* ‘подпирать’, *mertet* ‘мерка, орудие измерения’ < *mertani* ‘измерять, мерить, взвесить’, *šobiret* ‘покрывало, одеяло’ < *šobirjani* ‘покрывать, укрывать, накрывать’ и др.

б) орудие действия и результат действия: *kerttet* ‘вязанка’ и ‘завязка’ < *kerttini* ‘завязать, связать, обвязать’, *šjret* ‘нож для распарывания материи’ и ‘стежок шва; то, что подрезано’ < *šjriini* ‘распороть, подрезать’, *suret* ‘мешалка, мутовка’ и ‘смесь, мешанина’ < *surani* ‘мешать, смешивать’ и др.

в) действие и результат действия: *valektet* ‘объяснение, разъяснение’ и ‘то, что разъяснено’ < *valektini* ‘объяснить, разъяснить’, *jurttet* ‘помощь, поддержка’ (как процесс и как результат) < *jurttini* ‘помогать, помочь, поддержать’ и др.

г) результат и объект действия: *l'uket* ‘часть, отрезок’ < *l'ukini* ‘делить, расчленивать’, *gožtet* ‘письмо, записка’ < *gožtini* ‘писать, написать’, *tjmet* ‘пруд, запруда, плотина’ < *tjmini* ‘запрудить’ и др. (Грамматика 1962: 115–116).

В коми языках удмуртскому суффиксу *-et* соответствует суффикс *-et* (в коми-пермяцком) и *-ed* (в коми-зырянском). При помощи упомянутого аффикса образуются отглагольные существительные, обозначающие:

а) предмет, предназначенный для действия: кп. *kerttet* ‘повязка, привязь’ < *kertavni* ‘привязать, завязать’, *piket* ‘опора, подпорка’ < *pikni* ‘подпереть’, *tupket* ‘затычка’ < *tupkavni* ‘заткнуть, заделать’; кз. *kerted* ‘повязка, завязка’ < *kertavni* ‘завязать’, *piked* ‘подпорка’ < *pikni* ‘подпереть’;

б) результат производимого действия: кп. *otset* ‘помощь, поддержка’ < *otsavni* ‘помочь, помогать’, *gižet* ‘письмо’ < *gižni* ‘писать’, *šoget* ‘болезнь’ < *šogavni* ‘болеть’; кз. *čipeđ* ‘зарубка, выемка’ < *čipni* ‘сделать зарубку’, *gižed* ‘письмо’ < *gižni* ‘писать’; и т. п. (КПЯ: 207; СКЯ: 157).

Данные формы пермских языков напоминают осетинские причастия прошедшего времени, образованные при помощи похожего суффикса *-t/-d*, которые также могут переводиться и как отглагольные существительные, обозначающие действие и/или результат действия: *card* ‘жизнь’ < *caeryn* ‘жить’; *fyst* ‘написанный; писание; письмо’ < *fyssyn* ‘писать’; *bast* ‘связанный, привязанный; вязанка, пучок, связка’ < *battyn* ‘связывать’; *mard* ‘мёртвый, убитый; мертвец, покойник’ < *maryn* ‘убивать’; *syğd* ‘сгоревший, горелый, выжженный; пожар’ < *suzyn* ‘сгорать, гореть’; *kuyst* ‘работа’ < *kusyn* ‘работать’; *zond* ‘ум; знание’; *zynđ* ‘знаемый, известный’ < *zonyn* ‘знать’ и т. п. (Абаев 1949: 570; Абаев 1959: 63, 114). Осетинское причастие прошедшего времени восходит к древнеиранскому причастию на *-ta* (Абаев 1949: 570; Грамматика 1963: 271).

Здесь я не могу не указать на весьма сходные по звучанию и значению с пермскими и осетинскими венгерские формы отглагольных существительных с суффиксами *-at/-et/-t*, которые обозначают действие или, чаще, результат действия: *fordulat* ‘поворот, разворот, оборот’ < *fordulni* ‘оборачиваться, переворачиваться’, *fonat* ‘пряжа; коса’ < *fonni* ‘прясть; плести’, *menet* ‘ход, проезд, рейс, марш’ < *menni* ‘идти, ходить, ехать’, *cselekedet* ‘поступок, действие’ < *cselekedni* ‘действовать, поступать’, *hit* ‘вера’ < *hinni* ‘верить’ и т. п. (Балашша 1951: 154).

Некоторая часть финноугроведов склонна видеть в венгерском суффиксе развитие прафинно-угорского или даже прауральского отглагольного суффикса *\*-tt* (ОФУЯ 1974: 356; ОФУЯ 1976: 388). Однако, такая точка зрения поддерживается не всеми исследователями, единого мнения здесь нет (ОФУЯ 1974: 356–357). В то же время, что интересно, венгерский суффикс, кажется, никем не сопоставляется с вышеупомянутыми отглагольными суффиксами существительных пермских языков. С моей же точки зрения, общность происхождения пермских и венгерского суффиксов выглядит весьма вероятной. Кроме того, мы, возможно, вновь сталкиваемся с примером иранского влияния на пермские и венгерские языки.

2.6. Удмуртский суффикс сравнительной степени прилагательных и наречий *-gem* (употребляемый наряду с суффиксом *-ges*): *viĵ'gem/viĵ'ges* 'новее' (*viĵ'* 'новый'), *kužmogem/kužmoges* 'сильнее' (*kužmo* 'сильный'), *ićigem/ićiges* 'меньше' (*ići* 'мало') и т. д. Важно отметить, что в словах с этим суффиксом присутствует некий оттенок неполноты, умеренности качества. Дело в том, что сравнительная степень прилагательных и наречий в удмуртском языке может образовываться и без данного суффикса. Если сравнить два удмуртских выражения, с этим суффиксом и без него, например, 1) *ton miñeštĵm kužmo* и 2) *ton miñeštĵm kužmogem*, между ними есть определенная разница. В первом случае выражение означает просто «ты сильнее меня», а во втором — «ты немного, незначительно сильнее меня» (см. Грамматика 1962: 138).

Значение неполноты качества, придаваемое данным суффиксом проявляет себя не только в формулировании сравнительной степени, но и в выражениях типа *ćagĵrgem/ćagĵrges śinĵos* 'голубоватые глаза', где *ćagĵrgem/ćagĵrges* 'голубоватый' < *ćagĵr* 'голубой'. Помимо прилагательных и наречий этот суффикс может присоединяться и к глагольным основам, также придавая им значение неполноты, умеренности качества: *ton viśiškogem/viśiškoges* 'я немножко болею', тогда как *ton viśiško* означает просто 'я болею'.

В лузско-летском, верхнесысольском, вымском диалектах коми-зырянского и в верхнекамском диалекте коми-пермяцкого языков имеется суффикс неполноты качества прилагательных и наречий *-gem*; вс. *ul'gem* 'сыроватый' (*ul'* 'сырой'), лет. *leżgem* 'синеватый' (*leż* 'синий'), вк. *ranogem* 'рановато' (< рус. *рано*) (ССКЗД: 476, Баталова 1975: 172–173).

По поводу происхождения данного пермского суффикса финно-угроведы по сей день затрудняются сказать что-либо определённое. В ОФУЯ 1976: 155 говорится, что его происхождение неясно, Ш. Чуч, со ссылкой на А. Кёвешу, довольно туманно сообщает, что это, якобы, «...первоначально был суффикс прилагательного, который произошёл от древних языковых элементов» («...war ursprünglich ein Adjektivsuffix, das sich aus uralten sprachlichen Elementen entwickelt hatte») (Csúcs 2005: 207).

Ср., однако, осетинский (иронский) суффикс *-gotaw*, указывающий на умеренную степень какого-либо качества: *ædylygotaw* 'глуповатый' (от *ædyly* 'глупый'), *sabyrgotaw* 'тихонько' (*sabyr* 'тихий') (Абаев 1959: 118). В дигорском диалекте иронскому *-gotaw* соответствует *-gon*: *bor* 'жёлтый', *borgon* 'желтоватый' (Абаев 1949: 422). Переход конечного слога или слова *-t* > *-n* в дигорском — явление типичное для этого диалекта; ср. ирон. *not* — диг. *non* (наряду с *not*) 'имя', ирон. *salam* — диг. *salan* 'привет' и др. (Абаев 1949: 379). По всей видимости, иронский суффикс — двусоставной (*-got* + *-aw*). Об этом свидетельствует дигорская форма суффикса, а также то обстоятельство, что в осетинском языке существует суффикс существительных, прилагательных и наречий *-aw*, который В. И. Абаев считал идентичным аффиксу уподобительного падежа (Абаев 1949: 422; Абаев 1959: 120).

Внимательный читатель может справедливо заметить, что суффиксы и другие словообразовательные морфемы при языковых контактах заимствуются, как правило, вместе с включающей их в себя лексикой, и этому следовало бы уделить отдельное внимание. На это я хочу ответить следующее. Дело в том, что данная статья мною изначально задумывалась как первая часть более обширного исследования результатов ирано-пермского языкового взаимодействия. Следующая часть этого исследования будет специально посвящена результатам данного взаимодействия в области лексики и, возможно, фразеологии.

Подводя некоторые предварительные итоги, мне хотелось бы подчеркнуть, что тема пермско-иранского взаимодействия в области фонетики и морфологии выглядит доста-

точно перспективной, а посему исследования в этом направлении вполне могут быть продолжены, т. к., по моему мнению, способны привести к интересным результатам.

Так, например, в ходе данного исследования выявляются своего рода особые отношения между пермскими, венгерским и восточноиранскими языками алано-осетинского типа (см. выше, разделы 1.1, 1.2, 2.2, 2.5). А это, вкупе с уже довольно давно зафиксированными «особыми отношениями» между венгерским и пермскими языками (см. напр. Хелимский 1982: 19–20), позволяет с известной осторожностью предполагать существование во второй половине I тыс. н.э. в Прикамье и Предуралье языкового союза, включавшего в себя прапермский, правенгерский и алано-осетинский языковые компоненты. Этот вопрос безусловно заслуживает и пока ещё только ждёт отдельного обстоятельного и развёрнутого исследования. Хотя, справедливости ради, нужно отметить, что на данное тройственное ареальное взаимодействие уже обращали внимание исследователи этнической истории Волго-Уралья (Напольских, Смирнов 2022: 622–623).

### Список сокращений

ар. — арийский праязык; вв. — верхневычегодский диалект коми-зырянского языка; вк. — верхнекамское наречие/диалект коми-пермяцкого языка; вс. — верхнесысольский диалект коми-зырянского языка; венг. — венгерский язык; диал. — диалект, диалектное; диг. — дигорский диалект осетинского языка; др.-перм. — древнепермский (древнекоми-зырянский) язык; ир. — иранский праязык; ирон. — иронский диалект осетинского языка; кз. — коми-зырянский язык; кп. — коми-пермяцкий язык; кя — коми-язьвинский язык/диалект; л. — лузский говор лузско-летского диалекта коми-зырянского языка; лет. — летский говор лузско-летского диалекта коми-зырянского языка; лл. — лузско-летский диалект коми-зырянского языка; лит. — литературный язык, литературное; манс. — мансийский язык; мар. — марийский язык; напр. — например; нв. — нижневычегодский диалект коми-зырянского языка; осет. — осетинский язык; печ. — печорский диалект коми-зырянского языка; праперм. — прапермский язык, пермский праязык; праур. — пратуральский язык, уральский праязык; праФП — прафинно-пермский язык, финно-пермский праязык; праФУ — прафинно-угорский язык, финно-угорский праязык; санс. — санскрит; сев. — северное наречие удмуртского языка; скиф. — скифский язык; скр. — присыктывкарский диалект коми-зырянского языка; см. — смотрите; ср. — сравните; тж. — то же; уд. — удорский (вашско-мезенский) диалект коми-зырянского языка; удм. — удмуртский язык; фин. — финский язык; хант. — хантыйский язык; эрз. — эрзянский, эрзя-мордовский язык; эст. — эстонский язык.

### Литература

- Абаев, В. И. 1949. *Осетинский язык и фольклор*. Москва-Ленинград: Издательство Академии наук.
- Абаев, В. И. 1959. *Грамматический очерк осетинского языка*. Орджоникидзе: Северо-осетинское книжное издательство.
- Алатырев, В. И. 1976. *Именные деривационные аффиксы. Этимология некоторых слов*. Ижевск: Ротапринт республиканской типографии Управления по делам издательств, полиграфии и книжной торговли Совета Министров УАССР.
- Атаманов, М. Г. 1989. *Общепермский пласт удмуртских антропонимов. Вопросы финно-угорской ономастики: 169–181*. Ижевск: Удмуртский ИИЯЛ УрО АН СССР.
- Балашша, Й. 1951. *Венгерский язык*. Москва: Издательство иностранной литературы.
- Баталова, Р. М. 1975. *Коми-пермяцкая диалектология*. Москва: Наука.
- Белых, С. К. 2009. *Проблема распада прапермской этноязыковой общности*. Ижевск: Удмуртский госуниверситет.
- Грамматика 1962 = Перевошиков, П. Н. (ред.). *Грамматика современного удмуртского языка. Фонетика и морфология*. Ижевск: Удмуртское книжное издательство.
- Грамматика 1963 = Ахвледзиани, Г. С. (ред.). *Грамматика осетинского языка. Том 1. Фонетика и морфология*. Орджоникидзе: НИИ при Совете министров Северо-Осетинской АССР.

- Камболов, Т. Т. 2006. *Очерк истории осетинского языка*. Владикавказ: Ир.
- Кельмаков, В. К. 1998. *Краткий курс удмуртской диалектологии: Введение. Фонетика. Морфология. Диалектные тексты. Библиография*. Ижевск: Издательство Удмуртского университета.
- КПЯ 1962 = Лыткин, В. И. (ред.). 1962. *Коми-пермяцкий язык. Введение, фонетика, лексика и морфология*. Кудымкар: Коми-пермяцкое книжное издательство.
- КПРС 1985 = Баталова, Р. М., А. С. Кривошекова-Гантман. 1985. *Коми-пермяцко-русский словарь*. Москва: Русский язык.
- КЭСК = Лыткин, В. И., Е. С. Гуляев. 1970. *Краткий этимологический словарь коми языка*. Москва: Наука.
- Лыткин, В. И. 1975. Пермско-иранские языковые контакты. *Вопросы языкознания* 3: 84–97. Москва: Наука.
- Напольских, В. В. 2015. *Очерки по этнической истории*. Казань: Издательский дом «Казанская недвижимость».
- Напольских, В. В., О. В. Смирнов. 2022. Этнолингвистическая карта Поволжья в эпоху Великого переселения народов. *Археология Волго-Уралья. Том 4. Эпоха Великого переселения народов*: 610–622. Казань: Издательство АН РТ.
- ОИЯ = Расторгуева, В. С. (ред.). 1979. *Основы иранского языкознания. Древнеиранские языки*. Москва: Наука.
- ОФУЯ 1974 = Лыткин, В. И., К. Е. Майтинская, К. Редеи (ред.). *Основы финно-угорского языкознания. Вопросы происхождения и развития финно-угорских языков*. Москва: Наука.
- ОФУЯ 1976 = Лыткин, В. И., К. Е. Майтинская, К. Редеи (ред.). *Основы финно-угорского языкознания. Марийский, пермские и угорские языки*. Москва: Наука.
- Серебренников, Б. А. 1963. *Историческая морфология пермских языков*. Москва: Издательство АН СССР.
- СКЯ = Лыткин, В. И. (ред.). 1955. *Современный коми язык. Учебник для высших учебных заведений. Часть первая. Фонетика, лексика, морфология*. Сыктывкар: Коми книжное издательство.
- ССКЗД = Сорвачева, В. А. (ред.). 1961. *Сравнительный словарь коми-зырянских диалектов*. Сыктывкар: Коми книжное издательство.
- УРС = Вахрушев, В. М. (ред.). 1983. *Удмуртско-русский словарь*. Москва: Русский язык.
- Хайду, П. 1985. *Уральские языки и народы*. Москва: Прогресс.
- Хелимский, Е. А. 1982. *Древнейшие венгерско-самодийские языковые параллели (лингвистическая и этногенетическая интерпретация)*. Москва: Наука.

## References

- Abaev, Vasilij. 1949. *Osetinskij jazyk i fol'klor*. Moskva — Leningrad: Izdatel'stvo Akademii nauk.
- Abaev, Vasilij. 1959. *Grammatičeskij očerok osetinskogo jazyka*. Ordzhonikidze: Severo-osetinskoje knizhnoje izdatel'stvo.
- Alatyrev, Vasilij. 1976. *Imennyje derivatsionnyje affixy. Etimologija nekotoryx slov*. Izhevsk: Rotaprint respublikanskoj tipografii Upravlenija po delam izdatel'stv, poligrafii i knizhnoj trgovli pri Sovete Ministrov UASSR.
- Atamanov, Mikhail 1989. Obschčepermskij plast udmurtskix antroponomimov. *Voprosy finno-ugorskoj onomastiki*: 169–181. Izhevsk: Udmurtskij IJaL UrO RAN SSSR.
- Balassa, József 1951. *Vengerskij jazyk*. Moskva: Izdatel'stvo inostrannoj literatury.
- Batalova, Raisa 1975. *Komi-perm'atskaja dialektologija*. Moskva: Nauka.
- Belykh, Sergej. 2009. *Problema raspada prapermskoj etnojazykovoj obščnosti*. Izhevsk: Udmurtskij gosuniversitet.
- Csúcs, Sándor 2005. *Die Rekonstruktion der permischen Grundsprache*. Budapest: Akadémiai Kiadó.
- Grammatika 1962 = Perevoshchikov, P. N. (ed.). *Grammatika sovremennogo udmurtskogo jazyka. Fonetika i morfologija*. Izhevsk: Udmurtskoje knizhnoje izdatel'stvo.
- Grammatika 1963 = Akhvfefiani, G. S. (ed.). *Grammatika osetinskogo jazyka. Tom I. Fonetika i morfologija*. Ordzhonikidze: NII pri Sovete Ministrov Osetinskoj ASSR.
- Hajdu, P. 1985. *Ural'skije jazyki i narody*. Moskva: Progress.
- Helimskij, E. A. 1982. *Dreonejšije vengersko-samodijskije jazykovyje paralleli (lingvističeskaja i etnogenetičeskaja interpretatsija)*. Moskva: Nauka.
- Joki, Aulis J. 1973. Uralier und Indogermanen. *Mémoires de la Société Finno-Ougrienne. Vol. 151*. Helsinki: Suomalais-Ugrilainen Seura.
- Kambolov, Tamerlan. 2006. *Očerok istorii osetinskogo jazyka*. Vladikavkaz: Ir.
- Kel'makov, Valej. 1998. *Kratkij kurs udmurtskoj dialektologii: Vvedenie. Fonetika. Morfologija. Dialektnye teksty. Bibliografija*. Izhevsk: Izdatel'stvo Udmurtskogo universiteta.



- KpJa = Lytkin, Vasilij (ed.). 1962. *Komi-permiatskij jazyk. Vvedenie. Fonetika, leksika, morfologija*. Kudymkar: Komi-perm'atskoje knizhnoje izdatel'stvo.
- KPRS = Batalova, R. M., A. S. Krivoshechekova-Gantman. 1985. *Komi-perm'atsko-russkij slovar'*. Moskva: Russkij jazyk.
- KESK = Lytkin, V. I., E. S. Guliajev. 1970. *Kratkij etimologičeskij slovar' komi jazyka*. Moskva: Nauka.
- Lytkin, Vasilij. 1975. Permско-иранские языковые контакты. *Voprosy jazykoznanija* 3: 84–97. Moskva: Nauka.
- Napol'skikh, Vladimir. 2015. *Očerki po etničeskoj istorii. Kazan'*: Izdatel'skij dom "Kazanskaja nedvizhimost'".
- Napol'skikh, Vladimir, Oleg Smirnov. 2022. Etnolingvističeskaja karta Povolzhja v epokhu Velikogo pereselenija narodov. *Arxeologija Volgo-Uralja. Tom 4. Epoxa Velikogo pereselenija narodov*: 610–622. Kazan': Izdatel'stvo AN RT.
- OIJa = Rastorgujeva, V. S. (ed.). 1979. *Osnovy iranskogo jazykoznanija. Drevneiranskije jazyki*. Moskva: Nauka.
- OFUJa 1974 = Lytkin, V. I., K. E. Majtinskaja, K. Redei (eds.). *Osnovy finno-ugorskogo jazykoznanija. Voprosy proisxozhdenija i razvitija finno-ugorskix jazykov*. Moskva: Nauka.
- OFUJa 1976 = Lytkin, V. I., K. E. Majtinskaja, K. Redei (eds.) *Osnovy finno-ugorskogo jazykoznanija. Marijskij, perm'skije i ugorskije jazyki*. Moskva: Nauka.
- Rédei, Károly 1964. Vannak-e az előmagyar-permi érintkezésnek nyelvi nyomai? *Nyelvtudományi Közlemények* 66: 253–261. Budapest: Akadémiai Kiadó.
- Rédei, Károly 1986. Zu den indogermanisch-uralischen Sprachkontakten. *Österreichische Akademie der Wissenschaften. Philosophisch-historische Klassen. Sitzungsberichte*, 468. Wien: Verlag der Österreichischen Akademie der Wissenschaften,
- Serebrennikov, B. A. 1963. *Istoričeskaja morfologija permskix jazykov*. Moskva: Izdatel'stvo AN SSSR.
- SKJa = Lytkin, V. I. (ed.). 1955. *Sovremennyj komi jazyk. Učebnik dlja vysshix učebnyx zavedenij. Čast' pervaja. Fonetika, leksika, morfologija*. Syktyvkar: Komi knizhnoje izdatel'stvo.
- SSKZD = Sorvacheva, V. A. (ed.) 1961. *Sravnitel'nyj slovar' komi-zyrianskix dialektov*. Syktyvkar: Komi knizhnoje izdatel'stvo.
- UEW = Rédei, Karoi. 1988–1991. *Uralisches Etymologisches Wörterbuch. Band I–III*. Budapest/Wiesbaden: Verlag der Österreichischen Akademie der Wissenschaften.
- URS = Vakhrushev, V. M. (ed.). 1983. *Udmurtsko-russkij slovar'*. Moskva: Russkij jazyk.

*Sergey Belykh*. Permic-Iranian language contacts according to phonetics and morphology of the Udmurt and Komi languages

In this paper I discuss some results of the Iranian linguistic influence on the Permic branch languages of the Finno-Ugric group of the Uralic language family — Udmurt, Komi-Zyryan and Komi-Permyak. Unlike works by researchers of the long and recent past, the emphasis in the article is not on the lexical fund of these languages, but on their phonetic and morphological systems.

*Keywords*: Iranian languages; Scythian language; Ossetian language; Finno-Ugric languages; Permian languages; Udmurt language; Komi-Zyryan language; Komi-Permyak language; phonetics; morphology; historical linguistics; ethnic history.

## Из истории изучения дальнего родства: шесть фил в Северной Америке<sup>1</sup>

В статье впервые детально исследуются события 1913–1933 гг., связанные с появлением ключевых гипотез дальнего родства языковых семей в аборигенной Америке. Восстанавливается контекст состояния науки того времени и обращается внимание на непростые межличностные отношения Э. Сэпира, А. Л. Крёбера, П. Радина и др. участников описываемых дискуссий, а также на обстоятельства получения ими лингвистического материала, зачастую от последних носителей языков. Важным выводом является призыв рассматривать взгляды авторов «шести фил» в развитии, как некий компромисс между подходами, сложившимися в европейской и американской школах. Некоторые из выделенных Сэпиром фил (ацтекско-таноская), по-видимому, никогда не понимались им в духе классического сравнительно-исторического языкознания (как древо с единственным праязыком и проч.).

**Ключевые слова:** Э. Сэпир; А. Л. Крёбер; П. Радин; фила; суперсемья; генеалогическая классификация; компаративистика; скрещенные языки; американские индейские языки.

На сегодняшнем этапе для успешного продвижения в решении многих проблем, связанных с установлением дальнего родства языков, важную роль приобретает анализ, а порой и ревизия методологии предшествовавших поколений исследователей. Двадцатилетие 1913–1933 гг. вошло в историю науки как исключительно плодотворный период, когда было выдвинуто сразу несколько влиятельных гипотез «ламперов» по очередной перегруппировке языков коренного населения Америки. На протяжении длительного времени достижения американских ученых воспринимались как своего рода образец для многих регионов, получив, таким образом, общетеоретическую значимость. На основании различных источников, в том числе эпистолярных, в статье исследуется, как основные акторы этого периода (Э. Сэпир, А. Л. Крёбер, П. Радин и др.) в сложных взаимоотношениях друг с другом достигали своих результатов.

Еще Д. Бринтон обратил внимание на то, что в перечне семей американских индейских языков Дж. У. Пауэлла «не менее сорока», на самом деле 33–34, но во всех случаях больше половины, происходят из «узкой полоски земли между Скалистыми горами и Тихим океаном» (Brinton 1901: 57). Первая значительная ревизия авторитетной пауэлловской классификации (Powell 1891) в сторону укрупнения приходится именно на Калифорнию, где, начиная с 1900 г., работал А. Л. Крёбер. Во время полевой работы он лично записывал носителей языков, часто последних, многих калифорнийских «племен»: в 1900 г. — вийот и особенно юрок (вплоть до 1907 г.); с 1902 г. — йокатс (до 1904 г.) и мохэви (до 1911 г.); в 1907 г. — вашо; в 1908 г. — шошоноязычных серрано, кауийя, чемеуэйви, габриэлино и агуа-кальенте. В сотрудничестве с Роландом Диксоном, гарвардским специалистом по майду, Крёбер подготовил серию публикаций (Dixon, Kroeber 1903, 1913a,

---

<sup>1</sup> Исследование выполнено за счет гранта Российского научного фонда № 20-18-00159, <https://rscf.ru/project/20-18-00159/>, организация, осуществлявшая финансирование, — Институт языкознания Российской академии наук (ИЯз РАН).

1913b, 1919; Kroeber 1913), содержащих новый, редуccionистский взгляд на таксономию калифорнийских языков.

Ученики Франца Боаса, включая Крёбера, прошедшие в Колумбийском университете через семинар «Языки американских индейцев» их учителя, настроены были изучать языковое родство на основе анализа структуры языка в целом. В первой совместной публикации Крёбер и Диксон провели морфологическую классификацию, исходя из степени инкорпорации, выраженной в различных языках. В фокус попали самые разные признаки: местоименные аффиксы, падежи синтаксические («субъектив, объектив, посесив») и наречные, род, формы двойственного и множественного числа, редупликация, а кроме того, общий характер фонетики («трудная», «нетрудная», «легкая») и лексические сходства. Известные к тому времени языковые семьи, исключая яна и юмскую, оказались распределены по трем типам: юго-западному («чумаш»), северо-западному («юрорк») и центральному («майду») (Dixon, Kroeber 1903: 3, note 2).

Очень быстро типология, но столь аморфная, приобрела черты генетической классификации, и Крёбер и Диксон приступили к сведению 16 калифорнийских семей к более крупным, вначале четырем: шаста, чимарико, карок, помо, яна и юма были «консолидированы» ими в *хокскую* («хоканскую», Hokan); чумаш и салинская в — *искомскую* (Iskoman); вийот и юрок — в *ритвскую* («ритванскую», Ritwan); винтун, майду, йокатс, мивок и костано — в *пенутийскую* (Penutian, сегодня произносится rə'nu:ʃən, но исторически — rə'nu:tɪən). Двусмысленное положение сохранялось у эсселен — в статье, опубликованной в 'American Anthropologist', авторы причислили его к хокским (Dixon, Kroeber 1913b: 651 passim.), а в заметке в 'Science' не решились этого делать (Dixon, Kroeber 1913a: 225). За основу для новых искусственных названий, которые появляются ок. 1913 г., было взято числительное '2' в формах, свойственных языкам соответствующей группы, ср.: атсугейви *hoqi*, чумаш *iskom*, вийот *rit(w)-*; последнее название (пен+ути) представляло собой композит из прото-майду \*rə'ne и прото-костано \*uʔxi (Dixon, Kroeber 1913b: 653, 654; Shipley 1978: 82, 85; Campbell 1997: 67–68).

Методологически же столь быстрый сдвиг интересов вверх Крёбера в растерянность. В статье, опубликованной им в венском 'Anthropos', вроде бы перечислены компаративистские техники, нужные для надежного установления языкового родства, но видно, что сам автор вряд ли овладел всеми приемами в равной мере, кроме разве что лексических сравнений (Kroeber 1913: 394). И прочитавший статью индоевропеист Антуан Мейе подметил эту слабую сторону исследования Крёбера. Л. Кэмпбелл тоже считает, что крёберовские конструкции были «основаны в сущности на тех же методах, что и у Пауэлла» (Campbell, 1997: 68, 208), что вряд ли справедливо — стоит хотя бы просмотреть заключительную работу хокско-пенутийского цикла, содержащую разбор морфологии сравниваемых языков (Dixon, Kroeber 1919: 90–98).

Другое дело, что в поиске доказательной базы действительно ощущался некий тупик, и Крёбер прекрасно осознавал это («Сейчас меня удивляет, что мы упустили ключ к разгадке» — ALK/ES 03.01.1913; Golla 1984: 74). На протяжении нескольких лет перед появлением прославленных «шести фил» он буквально забрасывал Э. Сэпира лестными характеристиками в письмах, убеждая подключиться к более серьезному обоснованию только что выдвинутых радикальных предположений и, прежде всего, своих. Вот одно из таких посланий:

«Старик (так и есть, “Old Man” — И. К.), клянусь Юпитером, если бы у меня были твои знания и сила усвоения, и навыки обращения с чертовым материалом, я бы давно лишил тебя репутации лингвиста. По крайней мере, я не боюсь попробовать» (ALK/ES 24.07.1913; Darnell 1990: 112).

Но поначалу «старик», кажется, не спешил окунуться в калифорнийские дела, что объяснимо: в 1907–1908 гг. Сэпир уже ощутил непростой климат крёберовского департамента антропологии, попав туда вскоре после защиты докторской. Для смелых редуционистских шагов ему почти всегда хватало собственных сил, опыта и полевых данных. Еще до Калифорнии он изучал языки вишрам низовий р. Колумбия (чинукская семья) и такелма в шт. Орегон; у Крёбера занимался яна (вторично в 1915 г., когда будет обнаружен Иши — последний яхи). В год, проведенный в Пенсильванском университете, фокус интересов Сэпира переместился на южный пайют. Начиная с 1910 г., в свой оттавский, самый плодотворный для систематизации индейских языков период (хоть и скучный и не самый успешный с карьерной точки зрения), он исследовал нутка о-ва Ванкувер, сенека и мохок, несколько алгонкинских языков в окрестностях канадской столицы, цимшиан, а также сарси (Sargsee) и прочие атапаскоязычные группы. Атапаскские языки (часть будущих на-дене) останутся в фокусе его интересов и после возвращения в 1926 г. в Соединенные Штаты, в Чикаго.

Все же гонка постепенно захватила Сэпира. Возможно, первое свидетельство — его рецензия на исследование Р. Диксона о чимарико (1910), язык которых был записан фрагментарно, лишь со слов двух последних носителей — Полли Дайер, по определению Крёбера «беззубой старухи» (Kroeber 1925: 109), и индейца по прозвищу «Пятница» (Friday), частично чимарико, частично хупа и винтун (Janu 2009: 7–10). Рецензент заметил, что в фонетике автор «не всегда тщательно отделяет <...> сильные (fortes) от обычных глухих»; утверждает, что в чимарико отсутствует редупликация прилагательных со значением цвета, забывая об обратных примерах у себя же; более того, обосновывая объединение чимарико с шаста в одну семью, он упускает из внимания самые поразительные примеры их сходства: что и в тех, и в других языках местоименные элементы используются как префиксы и суффиксы. Словом, «доктор Диксон, похоже, не полностью утилизировал весь свой материал» (Sapir 1911: 142, 143).

**Компаративизм Сэпира.** Для общего понимания подхода Сэпира важна работа 1916 г. “Time Perspective in Aboriginal American Culture” (Sapir 1916), которую можно считать отправной точкой в развитии его взглядов на интересующую нас проблему. Прежде всего в ней заявляется, что «настоящая классификация генетически родственных языков всегда стремится принять вид генеалогического дерева» (Сэпир 1993: 558), т. е. *Stammbaum*, как в индоевропейском языкознании. В равной мере Сэпир допускает «возможность установления исторически равноценных языковых объединений, а также различия исторически первичного и относительно вторичного расхождения языков, даже когда последнее характеризуется сравнительно большей величиной» (Сэпир 1993: 561, сноска 49).

Признание существования сестринских языков подкрепляется у автора верой в реальность «сравнительно мало дифференцированного» праязыка (Сэпир 1993: 558). Все-таки дифференцированного, что может подразумевать не единый и монолитный *Ursprache* в духе А. Шлейхера, а диалектный континуум. Но такое допущение Сэпир никак не обсуждает, хотя буквально в то же самое время работает с материалами яна — очень похожим случаем, насколько будет показано дальше. В книге «Язык. Введение в изучение речи», написанной пятилетием позже, он прямо назовет состояние праязыка до распада «до-диалектным периодом» (Сэпир 1993: 157).

Про сэпировское понимание сути фонетических законов, постулирование которых превратилось чуть ли не в самоцель для тех, кто разделял методологию младограмматиков, в статье не говорится вовсе. Констатируется лишь, что знание подобных закономерностей чрезвычайно полезно при «стратификации заимствованной культурной лексики»

(Сепир 1993: 556, прим. 35). Но опять-таки в «Языке» этому посвящена целая глава, из которой видно, что, по мнению Сэпира, изменение в звуковой системе языка не происходит в отрыве от морфологии и синтаксиса и не сводится к действию известного принципа аналогии. Оно складывается из результатов действия, как минимум, трех «стихий», во-первых, общего дрейфа языка (*drift*), «о природе которого мы почти ничего не знаем», например, тенденции к усилению ударения, озвончения и проч.; во-вторых, выравнивания; и в-третьих, предохраняющей тенденции, работающей тогда, когда общий дрейф начинает угрожать «слишком серьезным морфологическим расстройством» (Сепир 1993: 169). При этом то, до какой степени невянятно он понимал дрейф языка, видно из его письма к Р. Лоуи, которое вообще посвящено растолкованию смысла 5-й и 6-й глав книги (ES/RL 23.05.1921; Lowie 1965: 49):

*«“Мистицизм” дрейфа происходит из факта, что непонятна точная природа (психологическая или иная) последовательного процесса, включающего ускоренное движение к типу. Мы способны чувствовать факты и процесс, даже если не можем разумно этот процесс определить. Мы получаем эти дрейфы и в истории искусства, религии, общественных сил. К детерминантам понятия «дрейф», несомненно, привлечена математическая и квазиэстетическая интуиция. «Эволюция» в обычном понимании, вероятно, представляет собой совершенно другой процесс. Может быть, «дрейф» и плохое слово, но у него есть то преимущество, что оно ни к чему не обязывает. Слишком невинно, чтобы причинять много боли».*

Относительно хронологии языковых разделений в вышеупомянутой статье предлагается учитывать фактор усложнения генеалогического дерева. Чем сильнее дифференцированы языки внутри семьи и чем шире ее география, тем больше времени должен занимать процесс дивергенции. Поэтому локализацию прародины («географического центра тяжести, в историческом смысле») резонно привязывать к размещению не всех ветвей семьи, а лишь основных (происходящих от сестринских языков), «независимо от того, насколько дробно они в свою очередь разделены» (Сепир 1993: 558, 561).

Вместе с тем, подчеркивается и другое — что современные носители языка, даже их большинство, могут не быть «потомками гомогенной группы людей, говоривших на гипотетическом праязыке» языковой семьи, к которой он относится (Сепир 1993: 559, прим. 43). Поэтому арии и семиты, о происхождении которых столько спорили в Старом Свете, не более, чем миф. Получается, что сопряженность лингвистических и этнических общностей относительно высока на уровне отдельных языков, но не семей в целом. Глубже во времени она ослабевает и теряется. «Грубо говоря, языковые элементы соответствуют элементам и комплексам элементов культуры, языковые группы — культурным ареалам», — в этой сэпировской формуле лучше всего выражены колебания и сомнения, насколько языковое родство надо понимать буквально, «генетически». И это при том, что в первой, посвященной культуре части своей статьи автор всячески убеждает, что культурные ареалы — не более, чем модель (Сепир 1993: 539).

Но в отличие от Боаса, судя по всему, отбросившего «догмы» индоевропейского языкознания, если не в ходе знаменитого семинара по языкам североамериканских индейцев, то уж точно во время подготовки “Handbook of American Indian Languages” (Boas 1911), у Сэпира, кажется, все проходило дольше и сложнее. В «Языке», в местах, где обсуждаются типы языковой структуры, он куда менее консервативен: признавая, например, что английский язык легче и лучше, чем немецкий усваивает заимствования, которые повлекли за собой даже отклонения от общего дрейфа. Последний конкретно в английском зависит от чего-то, что вполне можно было бы назвать последствием языкового

скращения: в языке уже имелось что-то такое, что «благоприятствовало усвоению новых слов», а заимствования в свою очередь «компенсировали какое-то внутреннее ослабление самого языка» (Сепир 1993: 156). Не так в более ранней «Time Perspective». В ней продолжают доминировать модели и схемы, имевшие хождение среди тогдашних европейских младограмматиков, возможно, именно из-за неуверенности ее автора, еще не настолько глубоко погрузившегося в лингвистику Нового Света.

Сэпир пробовал свести вместе горизонты культурной антропологии и лингвистики. А его мягкий компаративизм, местами даже синкретизм, несомненно проистекал из более общих ориентиров, заложенных в антропологии Исторической школы. Р. Дарнелл охарактеризовала его статью, как «самый крайний в боасовской антропологии пример лингвистики, выступающей служанкой этнологии», притом исходивший чуть ли не «от единственного боасовца, обладавшего подготовкой индоевропеиста» (Darnell 2021: 197).

И работа об определении временной перспективы в индейских языках и культурах, и уже упоминавшаяся книга, а помимо этого еще и одноименная с ней заметка, написанная для «Энциклопедии социальных наук» (Сепир 1993: 237–238), служат проявлением «мании классификаторства», по меткому определению Дарнелл (Darnell 1990: 110–117): то, что какие-то языки уже прочно ассоциируются с определенным перечнем языковых семей, вовсе не означает, что в ближайшем будущем остальные не займут место в них же или в каких-то еще до сих пор не выделенных группировках — «гораздо труднее доказать <...>, что некоторые языки, разделяющие немногие явно сходные черты, нельзя считать восходящими к общему источнику» (Сепир 1993: 558). Но, с другой стороны, необычайно высокая степень лингвистической пестроты в Новом Свете, должна свидетельствовать о том, что «языковая дифференциация» «только в незначительной части (на последних стадиях)» проходила здесь, а не в Азии еще до переселения (Сепир 1993: 539). Поэтому вскоре надо ожидать новых предположений и в том числе таких, которые выведут нас далеко за пределы американского континента.

**Алгонкино-ритвская фида.** В 1913 г. Сэпир изложил свою первую гипотезу, о родстве крёберовско-диксоновской ритвской семьи с алгонкинской на Востоке. Сравнению подверглись лексемы отдельных алгонкинских языков с вийот и юрок (чаще всего с каким-то одним, отрывочный материал не всегда позволял учитывать реалии обоих). Два последних языка — те нечастые случаи, когда он, не имея собственных полевых данных, вынужден был положиться на чужие наработки. Самому ему доведется добраться до этой части Калифорнии значительно позднее, уже в конце 1920-х.

Для Крёбера, напротив, юрок, поселенные в одноименную резервацию, еще на рубеже столетий стали «своими». Главный информант из их числа, Роберт Спотт, получил кое-какое образование, воевал на фронтах Первой мировой войны (пострадав от газовой атаки), председательствовал в племенном совете. До него Крёбер работал с его приемным отцом, Капитаном Споттом из Рекуой (Requa) в устье р. Кламат, как и с биологическим — Вейтчпекским Фрэнком (Weitchpec Frank) из дома Wogwu, выше по реке (Spott, Kroeber 1942: v; Kroeber 1954: 282). Значительно хуже в те времена были описаны вийот, подвергшиеся децимации, за катастрофическими последствиями которой антропологи просто не поспевали. «Кульминационным актом варварства и бесчеловечности со стороны... порочных белых», как написал крёберовский студент Ллуэллин Лауд, была резня, учиненная в 1860 г. старательским отребьем на о-ве Гантер (Индиан-Айленд) в заливе Гумбольдта (Buckley 1996: 285).

Сэпир представил список из 150 юрок-вийот-алгонкинских соответствий, не все из которых выглядели, однако, вполне бесспорными. Но он разбил их на 10 групп — суще-

ствительные, обозначающие индивидов, части тела, животных и растения, природные и искусственные объекты; числительные; местоимения; обозначения пространства; основы глаголов и прилагательных, — очевидно, имея в виду разную степень их важности как доказательств родства. Так он подчеркивал особую ценность совпадений в сравниваемых глаголах, перебрасывая мостик к морфологии: «*Очень важно отметить, что некоторые алгонкинские вторичные глагольные основы <...>, по-видимому, родственны первичным основам вийот*», при том, что были известны случаи, когда основы, первичные в одном алгонкинском диалекте, выступали как вторичные (заимствованные) в другом (Sapir 1913a: 629).

Среди свидетельств генетической общности чисто морфологического характера фигурировали посессивные местоименные префиксы, префиксы глагольных времен и залогов, показатели множественного числа и др. Сэпир описал также 8 закономерных звуковых преобразований, объясняющих, по его мнению, специфику ритвских форм в сравнении с алгонкинскими. В итоге он засомневался, надо ли по-прежнему настаивать на выделении ритвской семьи, хоть и родственной алгонкинской, но особой, или же вийот, юрок и алгонкинские языки («равнинно-атлантические») правильнее считать равноудаленными ветвями единого целого (Sapir 1913a: 646)?

Крёбер отреагировал на новость феерически (ALK/ES 30.07.1913; Golla 1984: 112–13): «*Дорогой д-р Сэпир, Ваш козырь выигрывает. Я уверен, что всегда думал об арапахо bā-, когда имел дело с вийот или юрок те-, we-, но кроме как о совпадении никогда ни о чем другом не мечтал*». По его словам, ему сразу захотелось все проверить и добавить свою толику к открытию, но материал по юрок находился у Уотермэна, а по арапахо — в офисе под ключом. Все же, прежде чем, как планировал, передать все свои записи по арапахо Майклсону, он решил еще раз взглянуть на них. Дальше следовал небольшой список из дополнительных соответствий: вийот *we's* ‘рука’ и арапахо *bäetcet* с тем же значением; вийот *weser* ‘женская грудь’ и арапахо *bäthän-i* ‘то же’; вийот *tan* и арапахо *-ot, n-ot* со значением ‘живот’... Против последних двух стояли вопросительные знаки. По-видимому, Крёбер сомневался в правильности своей транскрипции («*у меня такая плохая память*»). На скорую руку он попробовал даже вывести фонетический закон: «*вийот r < r̥ = алгонк. l > < n переход (?)*». Сэпир благодарил и уточнял (ES/ALK 05.08.1913; Golla 1984: 117):

«*Арапахо ba-θan-i “женская грудь” = В[ийот]. we-ser < \*-sen — чудесно, если процитировано правильно по памяти. До этого я не мог найти для ne-ser родственных алгонкинских слов, хотя смотрел внимательно. <...> Тебе будет интересно узнать, что у меня есть доказательство, показывающее что В.-Ю[рок]. L становится алгонкинским s̄ (или c̄)*».

В конце следующей декады крёберовские заготовки фонетических законов были и вовсе отвергнуты. Обмен двух «таксономистов» данными и идеями вокруг алгонкинско-ритвских сравнений продолжился, пока к декабрю не иссяк, причем Сэпиру, начиная с сентября отвлекавшемуся на другие темы, он видно наскучил еще раньше (ALK/ES 05.08, ES/ALK 05.08, 06.08, 08.08, 12.08, ALK/ES 12.08, ES/ALK 13.08, 14.08, ALK/ES 14.08, ES/ALK 18.08, 23.08, 26.08, 12.09, ALK/ES 23.09, ES/ALK 28.09, ALK/ES 30.12.1913; Golla 1984: 117–132).

**Юто-ацтекская фила.** Почти одновременно с алгонкино-ритвской гипотезой, но, судя по всему, все-таки на пару недель раньше, началась углубленная разработка другой, юто-ацтекской, которой повезло даже больше, в том смысле, что в итоге она оказалась лучше верифицирована. С 1913 г. Сэпир приступил к публикации результатов масштабного сравнительно-исторического исследования фонологии языков южный пайют и науатль (ацтекского) (Sapir 1913b, 1915b), с реконструкцией звукового состава общего

праязыка и постулированием ряда звуковых законов. К тому времени он уже посетил резервацию Юинта, где записывал ютоязычного Чарли Мэка, но главным информантом ему все же служил пайют Тони Тиллохаш из Карлайлской школы в Пенсильвании (Darnell 1990: 17). 30 мая 1913 г. Сэпир написал Крёберу (ES/ALK 30.05.1913; Golla 1984: 103–104):

«<...> [П]осылаю экземпляр (a carbon copy) первой части моей статьи о юто-ацтекской [семье], думаю, что Вам будет интересно увидеть материал перед тем, как он выйдет в виде публикации, возможно еще через какое-то время. Трактовка консонантов последует во второй части, в то время как третья посвящается рассмотрению пунктов морфологического сходства, многие из которых упомянуты, в действительности случайно, [уже] в настоящей части».

Все так. Парижское Общество американистов быстро начнет публиковать «Южный пайют и науатль». Вторая часть текста, из-за разгорающейся мировой войны, будет переложена в *'American Anthropologist'* и выйдет с двухгодичным запозданием, разбитая на два выпуска. Но морфологическая часть так никогда и не будет написана (Golla 1984: 104, note 1).

Чуть позже, но точно неясно, когда, Сэпир подхватит идею родства кайова с тано (к тому времени испанское *ñ* в названиях пауэлловских семей «американизировалось» в *n*), и тано с шошонскими, высказывавшуюся еще Дж. П. Харрингтоном, а до него И. Бушманном, А. Гэтшетом и др. (Harrington 1910: 119–123), расширяя таким образом юто-ацтекскую семью до ацтекско-таноской.

**На-дене́**, от хайда *na* 'населять; дом', либо тлинкитского *na* 'люди', и общеатапаскского *déne* с тем же значением — название, предложенное в 1915 г. Сэпиром (Sapir 1915a) еще для одной большой семьи, объединившей три упомянутых «ствола» из перечня Пауэлла. Из них тлинкит и хайда считались изолятами. По этим языкам у Сэпира были свои информанты: с тлинкитским помогал Луис Шотридж, с хайда — преподобный Питер Келли из Скидегейт, приехавший в Оттаву (Darnell 1990: 18).

В том же письме Крёберу, где объявлено было начало изучению юто-ацтекской фонологии, Сэпир упоминал, что в последнее время «занят атабаскскими, тлинкитским и хайда», и «собрал достаточно свидетельств, чтобы убедить себя, по крайней мере, в генетическом родстве этих трех» (ES/ALK 30.05.1913; Golla 1984: 104). В другом сэпировском письме (ES/ALK 05.08.1913; Golla 1984: 117), наряду с коррекцией звуковых переходов в алгонкинских и ритвских языках, говорилось: «Мною накоплены свидетельства, столько же хорошие, а может и лучше, по генетическому единству хайда, тлинкитского и атабаскских». Вырисовывается, что работа над на-дене, юто-ацтекской и алгонкино-ритвской семьями велась практически параллельно.

Но первую (и единственную) публикацию в *'American Anthropologist'* специально по на-дене Сэпир представил двумя годами позже, и лишь как предваряющую более основательное исследование, впрочем, так им и не осуществленное, что уже происходило и еще не раз повторится. В ней был предпринят комплексный анализ морфологии, лексики и фонологии языков сравниваемых групп и, в частности, сделан вывод, что общее своеобразие на-дене определяют 9 морфологических признаков, касающихся в основном глагола. Среди последних: корни существительных инкорпорируются в глагольный комплекс в качестве префиксов (инструментальных и др.); глагольной основе предшествуют местоименное подлежащее и дополнение, которое следует в самом начале; хайда и атапаскские («атабаскские») обладают аффиксами места, но в первом они выступают как



суффиксы, а во вторых — как префиксы; в атапаскских и тлинкитском имеется большое количество префиксальных «модальных» элементов со значением наречия; во всех языках на-дене для обозначения логического субъекта глагола используются субъектные/объектные местоименные элементы и др. Лексические сходства в статье содержали 98 примеров (меньше всего для хайда). Итоговым выводом было: «атапаскские, хайда и тлинкитский следует считать генетически родственными» (Sapir 1915a: 543–544, 551–554).

Еще Боасу в 1888 г. принадлежала мысль о возможной структурной близости тлинкитского и хайда, как и прозорливое наблюдение о том, что у них наиболее «азиатский» облик (Rohner 1969: 89, 98). В 1920–1921 гг. сходный вывод попытается обосновать Сэпир («языки на-дене, безусловно, самые «неамериканские» из всех языков, на которых говорят на севере континента» — ES/RL 15.02.1921; Lowie 1965: 44–46). Он задумается и о возможном родстве на-дене с сино-тибетскими языками, но это уже, так сказать, совсем другая история, и к ней еще предстоит вернуться. Ирония же заключается в том, что по прошествии времени Боас именно в пункте, касающемся на-дене (как и в ряде других), не поддержит начинания своего наиболее одаренного лингвистически последователя.

**Пенутийская фила.** Сэпир изначально поддержал пенутийскую, как и хокскую гипотезы Крёбера. «Мы оба верим» «в родство чимарико с шаста и костаноских с мивок», — подчеркивал, надо полагать, удовлетворенно Крёбер в своем письме, датированном самым началом 1913-го (ALK/ES 03.01.1913; Golla 1984: 73). Все предвоенные годы Сэпир следил из Оттавы за новостями, связанными с таксономией пенутийских языков. К апрелю 1915 г. он и сам предложил существенно пополнить их число, пока «негласно» (публикация об этом появится только через 6 лет — Sapir 1921b), включив в их состав «орегонские стволы (stocks)», превращенные таким образом в орегонские пенутийские (ES/ALK 15.04.1915; Golla 1984: 182–183):

«Дорогой д-р Крёбер,

Писал ли я тебе когда-нибудь, что у меня есть свидетельства, демонстрирующие, что такелма и кус родственны? Не очень близко, однако. Так[елма]. структурой куда более синтетический. Кус вполне аналитический по типу. Сейчас я начинаю подзревать, что кус (в итоге, может также сайусло, алсия [Alsea] и калапуйа?) и так[елма]. являются северными пенутийскими, отделенными от ю. пенутийских вторжением на север (шаста-ач[омави]., чим[арико]., карок, яна, помо) хокских языков, которые, кажется, тяготеют к югу <...>».

Далее перечислялись довольно убедительные примеры: кус *та* и йокатс *таi* ‘люди’; кус *yípsEn* ‘З’ при такелма *xibini* и йокатс *corin*, костано *кархан* с тем же значением; такелма *xi* и костано *ci* ‘вода’ (также кус *ci* ‘пить’). Судя по всему, хокское перемещение на север остановилось у горы Шаста. Из других калифорнийских групп шошонская также продвигалась на северо-запад, но атапаскская совершала противонаправленное движение на юг. Мигрировавшие атапаски разделили такелма с остальными северо-пенутийскими языками. Сэпир предположил, что синтетизм такелма мог быть более поздним явлением, развившимся на аналитической основе, характерной для кус, а в Калифорнии — для йокатс. «Конечно, существуют громадные морфологические различия между кус, так[елма]. и, скажем, йокатс или майду, но мы должны смотреть на эти вещи не негибкими дескриптивными глазами (sic!), а исторически». Одной из важнейших объединяющих черт, по его мнению, являлось наличие в сравниваемых языках (кус, такелма, йокатс) корней существительных специфического типа — двусложных, с повторяющейся гласной (CVC<sub>1</sub>V). Через неделю он уже добавлял список из 145 кусско-такелмско-пенутийских сравнений.

Вкупе с очередной порцией материала предлагались некоторые регулярные звуковые явления. В постскрипуме Сэпир попробовал было присвоить новой семье альтернативное название, но этот шаг не встретил отклика (ES/ALK 21.04.1915; Golla 1984: 183–184):

«Мне не нравится “Penutian”. Ввиду кост[ано]. *ama*, йокатс *mai*, майду *mai-*, кус *ti*, я бы предложил “Mai” (маи) в качестве названия семьи (stock name). Надело -an! Семья маи — было бы достаточно хорошо».

Крёбер отреагировал только через месяц и без энтузиазма: «За нехваткой времени вследствие подготовки к моему годовичному бегству, у меня не было возможности перечитать твоё доказательство новых родственных связей хокских и пенутийских» (ALK/ES 29.05.1915; Golla 1984: 191). Возможно, что он действительно был занят приготовлениями к европейскому вояжу, которому планировал посвятить свой *sabbatical*. Но нашлось время для другого — он просил Сэпира прислать ему оттиск работы «Южный пайют и науатль» (ALK/ES 18.05.1915; Golla 1984: 190). Не исключено, что Крёбер беспокоился о сохранении личного приоритета в классификации калифорнийских семей. Действительно, лучше всего, рассуждал он, если бы мы подготовили совместную статью о кус и такелма в связи с пенутийской проблемой, но это невозможно: «На протяжении шести месяцев у меня не будет времени, поэтому и не могу обещать». В любом случае «мы не представили ещё доказательства» (ALK/ES 29.05.1915; Golla 1984: 191)?!

Рукопись с «доказательством» Сэпира всплывет в архиве Ф. Боаса и только *post mortem* будет издана Моррисом Сводешем. Новый вариант содержит даже больше примеров орегонско- и калифорнийско-пенутийских сходств — 152. На вопрос, почему Сэпир не публиковал свою работу сам, Сводеш ответит, что включение в семью кус и такелма потянуло бы за собой необходимость обоснования ее дальнейшего расширения за счет сайусло, калапуя, чинукских и цимшиан, а для этого необходимы были уже новые свидетельства, которые только предстояло собрать (Sapir, Swadesh 1953: 133–134).

Годом позже Сэпир поделится с Крёбером (ES/ALK 14.07.1917; Golla 1984: 242–244), что обнаружил-таки несколько пар когнатов в ниска (один из языков цимшиан) и йокатс (один из калифорнийских пенутийских «стволов»). Так, йокатс *n-atet* ‘отец’ родственно, по его мнению, *häd-i<sup>’i</sup>* в речи женщин ниска (вокатив, где *-i<sup>’i</sup>* ‘мой’). Кроме того, у них же в речи мужчин имеется синонимичное *pä<sup>’p</sup>* (вокатив), коррелирующее с *n-orop* в йокатс, точнее в йоуламни (у Сэпира — йавелмани, Yauelmani). Соответственно, появление в йокатс дуплетов *n-orop* и *n-atet* объяснимо былым существованием и в этом языке гендерных различий, со временем стершихся. Похожим образом связаны йокатс (йоуламни) *no<sup>’om</sup>* ‘мать’ и ниска *nó<sup>’-i</sup>* ‘моя мать’ (форма вокатива — *nä<sup>’ä</sup>*), как и йоуламни *ts<sup>’utsa</sup>* ‘дитя дочери женщины’ (первоначально ‘мать матери’) и ниска *nts<sup>’é<sup>’ts</sup>-i</sup>* ‘мать моего отца (моей матери)’. При образовании от последнего формы вокатива *tsi<sup>’ts</sup>* исчезает глоттализиция (*ts<sup>’</sup>* > *ts*), что Сэпир считает признаком детского языка. Возможно, что того же, «детского» происхождения — *t<sup>’uta</sup>* (вместо *ts<sup>’utsa</sup>*) в йодэнчи, еще одном диалекте йокатс.

Установить все эти связи позволил один подмеченный ученым факт: в приведенных примерах, как правило, в вокативе исчезает *n-*, ср. йоуламни *n-orop* ‘отец’ при йодэнчи (вокатив) *oro-yo* ‘то же’: «Возможно мы имеем здесь след посессивного аффикса особого класса, используемого только в терминах родства». И по мнению Сэпира, ниска *nyé<sup>’e</sup>* ‘мой дед’ (*yé<sup>’e</sup>* – вокатив) изначально мотивировано как ‘я-дед’, указывая на принадлежность говорящему. Затем при превращении в невокативное ‘дед вообще’, после утраты начальным показателем первого лица ед. числа (ниска *n-* ‘я’) своего архаичного значения, появилась потребность в новом местоименном аффиксе *-e<sup>’i</sup>* ‘мой’. Если эта модель релевантна, то

могут заиграть и менее явные примеры пенутийско-цимшианских схождения, такие как: йокатс *bar* ‘мать отца’ и ниска *pəbéb-i* ‘брат моей матери’ (вокатив — *bi'p'*); йокатс *na-hamish* ‘тесть (свекор)’ и ниска *tātc* ‘тесть, теща (свекровь)’.

**Хокские языки.** Интересно, что, если инициатором расширения пенутийской семьи, похоже, выступал Сэпир, то хокской — определенно сам Крёбер, писавший 3 января все того же 13-го, богатого на ламперские начинания года (ALK/ES 03.01.1913; Golla 1984: 73–75):

*«В последнее время меня опять весьма заинтересовал яна. Помнишь, как два или три года назад Диксон и я собирали лексическую информацию по всем диалектам Калифорнии с намерением определить характер и степень заимствования между неродственными языками. <...> Помнишь, как ты достаточно любезно снабдил меня для (этой) цели перечнем из примерно 250 терминов яна, который, между прочим, оказался чрезвычайно удобен на ранних этапах нашего знакомства с Иши».*

«*Два или три года назад*» адресант, вместе с Р. Диксоном, приступил к исследованию, как раз и позволившему вскоре объявить об открытии новых языковых семей в аборигенной Калифорнии. Т. Уотермэн был первым из антропологов, кто две недели в сентябре 1911 г. пробовал установить контакт с Иши, «последним диким индейцем», в ку-тузке Оровилла (об обстоятельствах знакомства с ним и его трагической судьбе как информанта, см.: Kroeber, Th. 1961, 1970). В руках у Уотермэна находился тогда список слов яна, а ассистировал ему Сэм Батви. Насколько можно судить, кроме «250 терминов» Сэпира список содержал диксоновские материалы, собранные осенью-зимой 1900 г. по поручению Американского музея естественной истории. Диксон тоже работал с С. Батви, правда застал еще и вождя *Viī'yas'i*, известного белым как Джек с Раунд-маунтин.

Сэпир записывал яна в 1907 г.: в июле-августе у Чидаймии (Бетти Браун) нарративы на северном диалекте и в декабре у Батви — на центральном). Речь первой страдала чрезмерной лапидарностью. Из-за того, что единственной собеседницей Сэма уже долго являлась жена, он часто и не по делу переходил на женский язык, в шутку оправдываясь, что просто много мечтает и думает о женщинах. До Диксона и Сэпира лишь Дж. Кёртину удалось получить несколько текстов от Джека с Раунд-маунтин и еще от одного, давно умершего монолинга по кличке “The Governor”, прибегая к услугам все того же Батви (Sapir, Dixon 1910: 3–4, 6, 129). Как бы то ни было, южные диалекты оставались практически неизвестны науке вплоть до знакомства с Иши.

Хокскую принадлежность яна следует поставить под вопрос, — продолжал в письме Крёбер, готовый поначалу признавать скорее наличие древних контактов этого «ствола» с шаста и помо, нежели их общее языковое родство. Против такого предположения свидетельствовало, однако, еще меньшее количество заимствований, образовавшихся в результате контактов яна с исторически известными соседями — винтун и майду, рассматриваемыми как пенутиязычные (ALK/ES 03.01.1913; Golla 1984: 73–75). Сэпир же благодарил за присланную рукопись Диксона, но предупреждал, что в ней имеются недостатки и, в частности, путаница с отображением отдельных серий взрывных (ES/ALK 27.02.1913; Golla 1984: 89–90):

*«Как ты помнишь, я указывал в своей рецензии на «Чимарико» Диксона, что трудно понять, как расценивать лексические соответствия без определенного знания также и грамматических черт. <...> Извини, но фонетика материала, с которым ты имеешь дело, оставляет желать лучшего».*

Получается, что, сравнивая лингвистический материал яна и чимарико, Крёбер опирался на неверно зафиксированные глоссы. Например, в чимарико *qa'a* 'камень' следовало бы читать не велярный *q-*, а глоттализированный *k!*-. И скорректированное таким образом *k!a'a* оказывалось значительно ближе к яна *k!ai-* с тем же значением и т. д. и т. п.

В другом случае Крёбер предлагал Сэпиру сопоставить конечные элементы в яна *ha-na* 'вода' и эсселен *asa-nax* 'то же', соответственно *-nex* || *-nax* и *-na* (ALK/ES 12.04.1913; Golla 1984: 94). В третьем послании ученый из Калифорнии все еще продолжал гадать насчет возможного родства лутуами (*sic!*) с хокской семьей. Ему все больше и больше представлялось перспективным ее расширение, например, путем окончательного помещения в юмский дочерний «ствол» следующих языков — сери о-ва Тибурон (вопреки мнению Дж. Н. Б. Хьюитта) и текистлатекского Д. Бринтона в Южной Мексике: «[К]акими бы неопределенными ни были эти указания, вынужден буду удивиться, если наша хокская группа не докажет полностью, что является ядром очень большого ствола» (ALK/ES 08.05.1913; Golla 1984: 97).

Сэпир вернется в Калифорнию на три летних месяца 1915 г., чтобы поработать с Иши, пока того не сразит туберкулез, не оставив шансов на выздоровление. «Самая нервирующая» из всех работа будет проходить очень непросто. Она вообще бы не состоялась, если бы индеец был таким же «угрюмым, как некоторые более северные яна», например, Чидаймия, к тому времени уже покойная. Информанта, непривычного к общению с белыми, тяжело было «держать на привязи», чтобы записывать его под диктовку (Sapir 1923b: 264). Он так и не овладел в достаточной мере английским, изъясняясь со служащими музея и редкой вхожей профессурой на своего рода «пиджин-инглиш» (по определению В. Голлы): *hims no good* 'он — плохой', *sista* 'сестра (но и брат)', *hat-na* 'шляпа (с характерным для яна суффиксом абсолютива)' и т. п. (Pore 1920: 188; Golla 2003: 212–214). Батви и здесь помогал переводить (Sapir 1923b: 264).

Родным для матери Батви был южный диалект, но поскольку Сэм рано переселился в окрестности современного Миллвилла на Кау-крик, то почти полностью забыл его, перейдя на центральный — диалект своего отца (Sapir, Dixon 1910: 6, note 3a; Sapir, Spier 1943: 239). К 1907 г. южный яна уже полностью вымер, а Батви способен был выдать одни только разрозненные термины и фразы. Чидаймия тоже воспроизвела Сэпиру несколько фраз и слов на диалекте, некогда звучавшем, как она утверждала, на берегах Антелоп-крик (территория южных яна). После знакомства с Иши этот диалект представился ученому гораздо более близким к речи яхи, чем к северному или центральному яна, и он предложил различать внутри соответствующего «ствола»: северный яна, центральный яна и яхи, или южный яна, в свою очередь подразделявшийся на северный яхи (с Антелоп-крик) и южный яхи (речь Иши) (ES/ALK 10.01.1916; Golla 1984: 206–207). Позже, однако, Сэпир выступит с другой, четырехчленной классификацией, выделяя северный, центральный, южный яна и яхи, и видя в южном диалекте, толком так им и не задокументированном, «связующее звено между диалектами центральным и яхи, <...> с уклоном, скорее в яхи, чем в центральный яна» (Sapir 1923b: 263).

Но как бы ни было, яхи обнаруживал наименьшую близость по отношению к остальным диалектам яна. Именно по этой причине коммуникация Батви с Иши оказалась настолько затруднена. На севере же ареала, напротив, царила взаимопонятность, и между диалектами отсутствовали четкие границы. Так центральный яна у р. Кловер-крик, содержал уже определенные элементы северного диалекта и т. п. (Sapir, Spier 1943: 243). Все это свидетельствовало о том, что яна представляли собой не просто семью родственных языков, возникших путем дивергенции из некоего праязыка, а своеобразный, развивающийся в истории диалектный континуум. С другой стороны, яхи, особенно

в фонетическом отношении, показался Сэпиру «самым архаичным», близким к прото-яна и «заметно более резким для слуха» (Sapir 1923b: 263), а северный яна — напротив, наиболее ушедшим в развитии (Sapir, Spier 1943: 244).

Картину дополняло своеобразие политического устройства яна. Ни одна из диалектных групп, ни тем более яна в целом, никогда не представляли собой единства, распадаясь, как и в других частях аборигенной Калифорнии, на многие трайблеты, каждый с собственной отграниченной территорией и центральной деревней или ранчерией, где восседал наследственный вождь, собиравший дань, и располагалась общая церемониальная баня-потельня (sweat house). Иши обозначал термином *gari'si* центральную группу, не делая различий между ней и северными яна (Kroeber 1925: 345). Взгляд с севера на политическое деление был зеркально симметричным. Там похожим словом — *ga'ri'i* (*gari'ei* в сэпировской записи, т. е. на северном яна) — именовали как раз северных яна, а *gat'a'i* (*gat'a'ei* или *gat'a'ea* в северном произношении) — в совокупности центральных и южных. В основе всех этих терминов лежит глагольный корень *ga-* ‘говорить’, значение суффикса *t'a-* туманно, но *-ri(-)*, по-видимому, означает ‘низ, внизу’ (Sapir, Spier 1943: 243, note 19). Выходит, что яна в своей символической географии помещали верх на юге, а низ где-то севернее.

Существовали еще и особые женские и мужские речевые варианты, причем, если верить Сэпиру, во всех четырех диалектах. Женский яна отличался употреблением усеченных форм (или основ) почти всех слов. Исключение составляли лишь синтаксические частицы (артикли и проч.), два субстантивных глагола («быть», «есть») и пассивные формы с конечной долгой согласной, одинаково произносимые в обоих вариантах. На женском, теоретически используемом втрое чаще мужского, говорили как женщины, так и мужчины с женщинами. Но мужской тоже не составлял табу для противоположного пола — его могли употреблять женщины, например, цитируя речь двух мужских персонажей в мифах (Сепир 1993: 456, 460–461). По мнению Сэпира, сложные формализованные системы, какими являлись специфически женский и мужской языки яна, постепенно развились из «двух психологически отдельных источников» (two psychologically distinct sources). Так в речевом поведении женщины-яна длительное время подчеркивался ее второстепенный статус, обычаем ей предписывалось изъясняться не настолько свободно как мужчине, прибегая к редуцированным формам (Sapir 1949: 211), в чем, наверное, можно видеть отдаленный аналог античному лаконическому стилю.

Удивительная реальность языковой коммуникации яна научит Сэпира учитывать в своих этимологиях пол, возраст и родство говорящих. Вышеприведенные рассуждения о женских и мужских формах в ниска и, при правильной реконструкции еще в (прото-)йокатс, с высокой долей вероятности навеяны всплеском его интереса к яна и к персоне Иши. Теперь мы знаем, что гендерно маркированные социальные диалекты не такое уж редкое явление. Среди возможных «родственников» яна они, вероятно, имелись еще в салинском (Turner 1987: 115); за пределами Калифорнии — в языках пуэбло (Sims, Valiquette 1990), криков, по крайней мере, в коасати (Haas 1944, Bell 1990); в остальном мире — в чукотском, японском, шумерском и др.

Наиболее полно статус яна внутри хокского «ствола» будет обоснован в работе 1917 г. (Sapir 1917a: 1–34). В ней Сэпир приведет 192 материальных свидетельства в его пользу — корни, глагольные и именные суффиксы, местоимения и проч. Он также опишет несколько явлений фонологии, предположительно общехокских: наличие гласной в начале слова (как раз в яна она, как правило, утеряна); чередование гласных фонетической либо морфологической природы; а также их выпадение (особенно в шаста-ачомави) при определенных акцентологических условиях. Что касается «вывода» хокских языков за пределы Калифорнии и вообще поиска их дальних родственных уз, который закончится

постулированием хокско-сиусской филы, то, как уже отмечалось, первую скрипку здесь сыграет Крёбер.

Еще в 1915 г. он опубликовал таблицу из 35 возможных схождения, с одной стороны, языков «чонталь» и сери, а с другой — мохэви и остальных хокских (Kroeber 1915: 279, 282–283). Сэпир в письме поддержал этот давно ожидаемый шаг, но выразил скептицизм относительно корректности предложенных лексических параллелей (ES/ALK 05.03.1915; Golla 1984: 171):

«[М]ы вынуждены будем плутать в темноте до тех пор, пока не узнаем еще, какой точно размер у твоего нового хокского ствола, какие у него основные деления и подразделения, и что за фонетические законы действуют в отдельных диалектах».

Куда больше его волновала история хокской семьи в целом и, особенно, судьба корней специфического типа — VCV(C), будто бы лучше всего сохранившихся на южной окраине ареала у юма и на севере у чимарики. Что до яна, из всех хокских наиболее отделившегося, то «он отошел, скажем, от шаста и юмских, даже дальше, чем сери и чонталь». В том же послании Сэпир внушал своему корреспонденту мысль о возможности присоединения эсселен и искомских (чумаш и салинского) языков к «северным хокским» (шаста-ачомави, чимарики, яна, помо), единство которых казалось лучше обоснованным (ES/ALK 05.03.1915; Golla 1984: 172). До того Крёбер в «Новых языковых семьях в Калифорнии» осторожно привел девять возможных когнатов («вода», «дом», «небо» и др. — Dixon, Kroeber 1913b: 653), а Дж. П. Харрингтон на последней странице последнего номера того же тома *'American Anthropologist'* застолбил это направление поисков, не раскрыв, однако, никаких материалов (Harrington 1913: 716).

К концу марта-апрелю 1915 г. Сэпир постарался выделить первый фонетический закон (эсселен *c, s* : хокские *h, x*) и, как ему казалось, распознал в искомских некие «префиксальные элементы» *t-, n-, l-*, с учетом которых можно было бы проводить новые сопоставления с хокскими: мохэви *ime* 'нога' и чумаш *t-em, n-ime-l* 'то же'; мохэви *аттауа* 'небо' и чумаш *a-l-ара(уа)*, салинское *l-эм(о)* 'то же' и т. д. (ES/ALK late March 1915?; Golla 1984: 176). Очередная догадка содержалась в следующем письме — общее для всей семьи именное окончание *-l*: помо *-l*, чимарики *-l, -r*; дегэйньо *-ly*, в одном ряду с которыми оказались также яна *-на* и эсселен *-нах* (ES/ALK 07.04.1915; Golla 1984: 182) — два года назад Крёбер уже обращал внимание на эту пару соответствий (см. выше). Там же Сэпир, как кажется, впервые поколебал статус еще одного изолята: «Как насчет вашо? У меня есть смутное предчувствие, что он окажется сильно отошедшим хокским языком! Вашо *d-*, как и в чумаш и салинском, < хокского *\*da* "том" <...>. *d-апа-l* < *\*d-ата-l(a)* выглядит необычайно по-хокски. Я подметил и другие хокские точки схождения. Хочешь их?» (ES/ALK 07.04.1915; Golla 1984: 182).

На следующий год после первого своего упоминания вашо Сэпир подготовил небольшой текст, который еще через год увидит свет в *'American Anthropologist'* (Sapir 1917b). По прочтении машинописи (в отличие от Мерриама и Харрингтона боасовцы свободно обменивались дореченными версиями своих трудов) Крёбер несколько растерялся (ALK/ES 19.03.1916; Golla 1984: 209):

«Буду очень рад, если ты сможешь установить статус вашо. Ты знаешь, что я чувствую по этому поводу; но я слишком медлителен, чтобы в свободные моменты разбираться со звуковыми переходами (soundshifts), и, вероятно, никогда не пропитаюсь настолько, чтобы начать их угадывать. Кстати, фонетика вашо у меня особенно гнилая. Я никогда не слышал языка, кроме нескольких дней, и даже в то время в принципе не отличал один звук от другого».

Сэпировский «Статус вашо» вышел вдогонку новости с той, другой стороны — об «открытии» Харрингтоном родства вашо и чумаш (Harrington 1917: 154). Поскольку хокский характер чумаш уже «подозревался докторами Диксоном и Крёбером» (sic!), а о родстве с юмской, «типично хокской группой диалектов» заявлял Харрингтон, разумно поставить вопрос о связях вашо с хокскими языками. Сэпир пообещал, что в скором времени опубликует доказательства (как известно, этого не произошло), опять повторял, что яна грамматически ушли даже дальше вашо. На второй половинке одной-единственной страницы он умудрился проинформировать о текущем состоянии хокских исследований, в частности еще об одном кандидате на членство в разрабатываемой им и Крёбером филе, а именно о «недавно установленной д-ром Суантоном» группе языков западного берега Мексиканского залива (Gulf, галф). Действительно, последний объединил в 1915 г. коауильтеко, включая кокекрудо и котонаме, с каранкава, тонкава и как будто бы еще с ат(т)акапа (Swanton 1915: 17–40): «Очевидно, что предостережения, к которым призывали некоторые более консервативные исследователи, <...> не оказывают чрезмерно сдерживающего влияния» (Sapir 1917b: 450). Весьма наглядно выходит, указывал автор заметки, что территория, отделяющая коауильтеко-каранкава-тонкава от юма-чумаш-вашо, почти полностью занята (была занята) шошонскими и атапаскскими народами, переселившимися, очевидно, позднее, первые — с канадского севера, а вторые — из нынешних южной Аризоны и северной Мексики.

Дальше Крёбер объяснял, почему до сих пор в своих публикациях с Диксоном не включал вашо в состав хокского «ствола», но сейчас, в итоговой, меняет позицию: «Теперь мне кажется очень неразумным, чтобы и мы, и ты доказывали один и тот же момент независимо друг от друга и примерно одновременно, не говоря уже о Харрингтоне, возможно, превращающем [все] это в деловой треугольник». Вдруг он вспомнил, что у Диксона имелся некий полевой блокнот по вашо, «несомненно, больше, чем малость, которую я опубликовал». И теперь в интересах общего дела можно было действовать двояко. Либо Крёбер попросит своего соавтора передать все материалы Сэпиру для продолжения, либо, если последний предпочтет, чтобы его «освободили от работы доказывать, что вашо является хокским», они доделают все сами. «Все, чего я хочу в нашей статье, — это поместить язык туда, где он должен находиться» (ALK/ES 11.10.1917; Golla 1984: 251–252). Сэпир, ссылаясь на занятость, с легкостью, по крайней мере, внешней, «свалил все» на Крёбера, отослав ему заказным письмом свои карточки с лексическими сходствами в корневых элементах, префиксах и суффиксах (ES/ALK 17.10.1917; Golla 1984: 254):

*«Общие морфологические сходства, которые, возможно, даже более важны, я никогда не записывал: они в основном у меня в голове или могут быть легко проверены путем небольшого поиска исходного материала. Есть также ряд разработанных мной фонетических законов, которые имплицитно содержатся в посылаемом тебе материале. Полагаю, что уже писал тебе какое-то время назад относительно некоторых из них».*

Любопытно, что никому из них так и не пришло в голову подготовить что-либо в соавторстве, хотя к тому времени Крёбер должен был точно убедиться, что Сэпир, как лингвист и «таксономист», намного превосходил Диксона. Единственное, о чем просил Сэпир, так это позволить ему взглянуть на гранки в той части, которая будет посвящена вашо, чтобы, если понадобится, «исправить любые возможные недоразумения с вашей стороны при использовании моего материала или предложить дальнейшие рекомендации» (ES/ALK 17.10.1917; Golla 1984: 254). Все же после ознакомления с «Языковыми семьями Калифорнии» он написал (ES/ALK 10.11.1919; Golla 1984: 316–317): «Дорогой Крёбер, <...> Я не вижу, почему ты выражаешься так осторожно о вашо. Думаю, между нами, что мы на-

копили достаточно свидетельств, чтобы доказать хокский характер вашо при отсутствии обоснованного сомнения». Видимо, Крёбер так и не смог себя пересилить, не разглядев в свое время этого структурного сходства.

И снова: в сравнении с яна, а может быть и с помо, вашо по своей структуре предстает даже более хокским языком! Самыми же характерными для хокских надо считать чимарики, вероятно еще шастские; для пенутийских — йокатс. Следом шел поток предположений, расцвечивающих разными деталями историю калифорнийских языков. Например, пенутийский майду подвергся, особенно в морфологии, сильному влиянию хокских и в определенных инструментальных префиксах, возможно, еще шошонских. Наоборот, дрейф помо определяли контакты с пенутийскими. Во многих аспектах этот язык настолько близок шастским, чимарики и яна, что, вероятнее всего, географически когда-то соприкасался с ними, а затем покинул свою изначальную территорию, занятую юки и каким-то пенутийским, возможно, винтун. В свою очередь в юки Сэпир подозревал «изолированный осколок», весьма отклонившийся от других хокских, который подобно помо, но еще раньше, начал развиваться под пенутийским воздействием. Из параллелей к юки, предлагаемых Крёбером, его больше убедили хокские, нежели пенутийские. «Возможно, юки нужно сравнивать скорее с хокско-коауильтекскими, а не с хокскими в их узком значении» (ES/ALK 10.11.1919; Golla 1984: 316–317).

Калифорнийская мозаика, с упорядочения которой все началось, сыграла решающую роль в развитии сэпировской компаративистики. Судя по всему, в этой части Америки антропология столкнулась с ситуацией, не менее архаичной, чем лингвистическая непрерывность, известная в других концах света — в центре Индостана, на Новой Гвинее и др., — но которая в ряде аспектов была прямо обратной этим случаям. Здесь части языковых семейств, или стволов, территориально перемежались друг с другом, границы анклавов были более чем очевидны, но внутри них в редких случаях находились дочерние языки, отчетливо дифференцированные на диалекты, как например, помо и юмские, чаще же — осколки бывшего диалектного континуума, как яна. Хокская гипотеза, особенно во всех ее расширительных вариантах — хокско-искомской, хокско-коауильтекской, юки-хокской и проч. — воплотилась в сложных и компромиссных моделях, «трехмерных», в сравнении с «плоскими» филами алгонкино-ритвской и на-дене, которые больше походили на классическое генеалогическое древо. Хотя выявление лексических сходств и фонетических законов сохранялось в приоритете, о чем неустанно повторялось, все же непрекращающаяся подгонка методики приводила к тому, что от реконструкции праязыков на их основе пришлось отказаться на неопределенный срок. Сэпир настойчиво пытался отделить субстратные явления от адстратных, пробовал представить дрейф каждого языка, в общем-то пробивая брешь в лингвистику контактов, скрещенных языков и пресловутой «волновой» теории. Его все сильнее завораживал поиск суффиксов, окаменевших частиц в сравниваемых словах и всякого рода специфической лексики, которая могла бы увести в более глубокие хронологические пласты. И удивительно, на новом витке классификация калифорнийских «стволов» возвращалась к исходным типологическим ориентирам Крёбера и Диксона. Ветви схематичных «стволов» оживляли хаотично прораставшие то здесь, то там «ризомы». Все это не сильно напоминало ортодоксальное сравнительно-историческое языкознание.

**Америндская гипотеза Радина.** Примерно в 1915 г. к классифицированию североамериканских языков подключился Пол Радин, обменявшись подготовительными материалами с Сэпиром и Крёбером. Как и Сэпир, Радин сотрудничал в тот период с Геологоразведкой Канады, но приезжал и в Калифорнию. Что-то из его сопоставлений заин-



тересовало Сэпира, хотя методология явно раздражала. Но, как бы там ни было, без новых, привнесенных Радином идей вряд ли появились бы «шесть фил» в известном сегодня виде.

Регна Дарнелл представляет дело так, будто радинские начинания в этой области носили не более чем дилетантский характер. Характеризуя результаты его деятельности как «фиаско», исследовательница подводит к мысли, что они не имели никакого значения ни для Сэпира персонально, ни для развития соответствующего направления поисков в целом (Darnell 1990: 118–122). Ниже будет показано, что это не совсем так. Отечественная историография неоправданно скупа на дискуссию по вопросу о вкладе Радиана. В канонических «Народах Америки» разведены сэпировская таксономия и гипотеза Радиана, о которой упоминается в разделе «Современное состояние», а его имя идет в одном ряду с Тромбетти: *«Более осторожный американский исследователь Пол Радин (1919) пытался установить наличие некоторых общих черт в грамматических формах всех по крайней мере североамериканских языков»* (Токарев, Бломквист 1959: 41). Вопреки реальной цепочке событий создается впечатление, что Радин проводил свои сопоставления не параллельно с Сэпиром, а вслед за ним и на основе его достижений.

Между тем, еще в 1918 г. Сэпир *«очень внимательно просмотрел таблицы Радиана»* — рукопись будущей резонансной “Genetic Relationship of the North American Indian Languages” (ES/ALK 19.07.1918; Golla 1984: 278). Разумеется, вновь примкнувшему «таксономисту» не хватало профессионального тренинга. Едва ли он владел в должной мере техникой определения регулярных фонетических соответствий. Радин и сам как-то писал, отшучиваясь по этому поводу (PR/ES 06.08.1913; Darnell 1990: 119): *«Почему люди всегда смотрят на меня как на лингвиста, если я так старательно избегаю общения с ними?»*. И все же его заслуги склонен был признавать никто иной, как Сэпир, высокий профессионализм и лингвистическое чутье которого не подвергались сомнению, советовавший Крёберу прочитать манускрипт, содержащий свежие смелые выводы (ES/ALK 19.07.1918; Golla 1984: 278):

*«Я на самом деле думаю, что он [Радин — И. К.] добился настоящего успеха в своих сиуских сравнениях, и, как и сам я не раз подозревал, думаю также, что юки и сиуские окажутся генетически родственны хокским. <...> На твоём месте я бы внимательнее изучил его сиуский материал; выкажи ему похвалу, какой он реально желает, и попроси продолжить эту часть его проблемы тщательно и систематически».*

Юки — единственный изолят в Калифорнии, никем еще не аффилированный ни с какой из известных языковых группировок. Попытки своего коллеги установить родство этой маленькой диалектной группы с пенутийскими, атапаскскими, вийот-юрок и алгонкинским Сэпир отметал сходу. Но вот хокские и сиуские языки — совсем другое дело. Юки мог оказаться «главным членом» (a major member) хокской «увеличенной группы», но, возможно, существовало также особое родство юки с сиускими, и тогда все три ветви надо было признать сестринскими. Все же *«совершенно необходима, — напоминал он, — [п]редварительная работа по сравнительной сиуской фонологии»* (ES/ALK 19.07.1918; Golla 1984: 278–279).

Радин овладел «сиуским материалом» благодаря полевой работе у уиннебэйго (хочанк), блестяще проведенной им еще в годы докторантуры при Колумбийском университете. Теперь (весна и лето 1918 г.), получив позицию в Миллс-колледже в Окленде, шт. Калифорния, он отправился к ваппо — подразделению юки, члены которого проживали на территории графства Сонома, недалеко от городка Хилдсбург. Как и предшественникам, ему посчастливилось разыскать последних носителей языка (ваппо вымер

в 1990 г.), а среди них подходящих информантов — Джима Трайпо (Tippo) и Джо Мак-Клауда (Radin 1924: 2). Естественно, помимо привычной для боасовцев записи мифологических текстов, Радина интересовали лингвистические свидетельства внешней аффилиации ваппо, и получалось, что их речь обнаруживает *«довольно четкие и определенные связи с хокскими, с одной стороны, и с сиускими, с другой»*. Позже он распознал еще и ее *«замечательное сходство с атапаскскими и пенутийскими языками»* (Radin 1919a: 490), на что Крёбер из Беркли, где привыкли к большей строгости, сыронизировал (ALK/Gifford 18.06.1918; Darnell 1998: 241):

*«Все мы 17 лет возились с Калифорнией, и он за две недели объединяет с ваппо пол-континента. Ваппо может быть сиуским; но у вас не получится заставить сиуский [,] атапаскский и хокский прогуливаться в Хилдсберге».*

Последнее, что надо иметь в виду, это то, что Радину явно свойственно было гейткиперство. В апреле 1915 г. Сэпир отослал в Сан-Франциско выполненный под копирку экземпляр своей только что написанной работы о хокско-коауильтекском родстве (ES/ALK 03.04.1915; Golla 1984: 179). Примерно через два месяца, по тем временам достаточно оперативно, Крёбер извинялся, что университетское издательство не может ее опубликовать, поскольку автор никак не связан с университетом (ALK/ES 29.05.1915; Golla 1984: 192). Сэпир послал машинопись Радину, который, находясь тогда в Санте-Фе, выступил соорганизатором Юго-западной антропологической ассоциации и, как никто, мог бы содействовать публикации (ES/ALK 25.11.1915; Golla 1984: 198): *«Моя статья об этом предмете сейчас в руках Радина, для его нового Юго-западного общества»*. Но тот вместо ожидаемой помощи неизвестно сколько продержал текст: *«Радин еще не научился возвращать рукопись, порой по два или три года после ее получения, если вообще возвращал»* (ES/ALK 19.07.1918; Golla 1984: 278). Похоже, что именно из-за него «Хокские и коауильтекские языки» (Sapir 1920) — сэпировская работа, исключительно важная для верификации сразу нескольких предлагаемых языковых группировок, от хокского «ствола» до хокско-сиуской филы, — стала публично доступной слишком поздно, уже после обнаружения «шести фил».

В 1918-м была готова и в начале следующего года появилась в крёберовской серии публикаций по американской археологии и этнологии 14-страничная работа Радина о генетическом родстве североамериканских индейских языков. С самого начала автор предпочел оговориться, что *«не предпринималось никаких попыток последовательно проследить все морфологические формы и словари по всем языкам»*. А кроме того, в статье, якобы из-за ограничений по объему, пришлось отказаться от доказательства истинности выделенных им *«морфологических элементов, которые ранее таковыми не признавались»* (Radin 1919a: 489, note 1). В отличие от Сэпира, считавшего, что предки коренных американцев были изначально лингвистически гетерогенны, Радин предположил их единство. Поскольку даже с точки зрения здравого смысла Северо-восточная Азия, откуда шло переселение в Новый Свет, никак не могла выступать в роли Вавилонской башни языков, наблюдаемое сегодня языковое многообразие должно было сформироваться уже на новом месте! Забегая вперед, отметим, данные современной генетики свидетельствуют о наличии в прошлом, по крайней мере, трех отдельных переселенческих волн (Gibbons 2012: 144; Achilli et al. 2013: 14311), и не дают возможности для столь радикального объединения намеченных фил. Дарнелл упрекает Радина еще и за то, что в своей датировке процесса миграции в Америку — ок. 15 тыс. лет назад — он опирался на *«неадекватные по общему признанию»* археологические данные (Darnell 1990: 119). Но в том то и дело, что теория “Clovis First”, господствовавшая в годы, когда писалась ее собственная книга, в наши дни устарела, и сегодня исследователи называют даты все ближе и ближе к радинской.

В поиске общих для всех языков Северной Америки черт Радин сосредоточил внимание на восьми грамматических признаках, половина из которых будто бы была характерна для большинства языковых семей, выделяемых на континенте, а другая — поголовно для всех них: пассивный залог, иррегулярность форм множественного числа, редупликация, суффиксы или инфиксы глагольного вида. Он также попробовал сгруппировать известные семьи в три большие «подгруппы», очевидно, в очередной раз стирая грань между генеалогической таксономией и типологией (Radin 1919a: 491–492): I. Сэлишская, квакиутль, кутенэй, алгонкин; II. Пенутийская, лутуамийская, сахаптин, шошонская, таноская, юкийская, михе, сапотекская, кэддоская, ирокезская; III. Атапаскская, хокская, майя, сиуская, маскогская. При этом, продолжал он, атапаскские языки, вероятно, занимали промежуточное положение между, с одной стороны, хокскими, а с другой — цимшиан и сэлишскими; юки равноудален от атапаскских, хокских и сиуских, но находился ближе всего к пенутийским; шошонские, наиболее походившие на тано и сиу, в то же время одинаково далеки как от пенутийских, так и от юки; самые близкие алгонкинским ирокезские языки в равной мере удалились и от кэддо, и от маскогских. Эти его рассуждения иллюстрировало внушительное число «соответствий», требующих, однако объяснений, например:

‘Голова’ (№ 19): пенутийское *to-i*, *tco-i*; пайют *tco-*; атапаскское (бивер) *tsi'*; хокское (помо) *ci-na*, *xi-ya*; хокское (чимарики) *hi-ma*; квакиутль *saia* (‘лицо’); цимшиан *tsa-i* (‘лицо’); алгонкинское (оджибва) *cti-gwan*; лутуамийское (кламат) *teli-sh*;

салинское *тоо-с*; квакиутль *erne* (‘лицо’); сиуское (дакота) *pa*; маскогское (чокто) *антра(?)*; майя *po-i*; уабе (Huave) *ta-i*; вийот *ba-L* (‘волосы’); кэддоское *pu-ks*.

Всего им было приведено 75 морфологических элементов, 42 имени, 25 глаголов, 7 личных местоимений, 4 указательных местоимения и 2 числительных, чтобы подкрепить свои догадки (Radin 1919a: 493–502). Сэпир вспылит, обескураженный вопиющей методологической и просто технической сыростью радинского письма (ES/ALK 19.07.1918; Golla 1984: 278):

«Как я и ожидал, ужасающая сумма явной гнили, беспорядочно перемешанной (higgledy-piggledy) с некоторыми действительно хорошими вещами. <...> Подменить атапаскским на-дене, и затем дальше — [языком] хуна атапаскский, не принимая в расчет весьма специфическое диалектное развитие форм, которые сами по себе являются вторичным развитием других форм, это и есть [радинская] идея исторического лингвистического метода».

Намереваясь опубликовать разгромную рецензию в ‘*American Anthropologist*’, Сэпир, однако, вскоре передумает, к неудовольствию главного редактора П. Э. Годдарда, который, по язвительному замечанию Крёбера, наслаждался «разногласиями в нашем лагере». Персонально Крёбер отнесется к предложению Радина скорее спокойно и попробует даже отговаривать Сэпира от совершения резких шагов, могущих навредить их тройственной дружбе (ALK/ES 23.08.1919; Darnell 1990: 120). Тот же, после выяснения отношений с Полом в паре писем, сосредоточится на доведении хоть до какого-нибудь конца своих собственных исследований. Сэпира не меньше остальных заботил личный приоритет, поскольку со стороны могло показаться, что его приятель, а теперь оппонент шел с ним вровень. (При этом Радин в публикации аккуратно помечал, где материал Сэпира, а где его собственный.)

«Шесть больших групп». Оба придут к решению представить и обсудить новую классификацию языков Северной Америки на собрании Американской антропологиче-

ской ассоциации в конце года, но в реальности это произойдет годом позже и на форуме другой ассоциации. Пока же Сэпир продолжит делиться свои разработками с коллегами. Первым в их ряду будет, естественно, Ф. Боас, реакция которого оказалась на удивление мягкой, но при видимом его спокойствии чувствовалось внутреннее напряжение (FB/ES 18.09.1920; Darnell 1990: 122):

*«Я не думаю, что наши мнения на самом деле настолько различны, как это могло бы показаться постороннему <...>. Думаю, однако, что мы недостаточно знакомы с феноменами взаимного влияния языков в примитивной жизни, чтобы решить, имеем ли мы дело с постепенным развитием дивергенции, или же в целом языковые явления не следует рассматривать с одной и той же точки зрения, как и любые этнические явления <...>».*

Вторым будет Р. Лоуи. В адресованном ему письме (ES/RL 09.09.1920; Lowie 1965: 39–40) гипотеза излагалась уже почти в законченном виде, а аргументация — даже полнее, чем в резюмирующей все итоговой публикации. Интересно, что параллельно Сэпир писал пятую главу своего «Языка», посвященную грамматическим значениям (форме в языке), одну из важнейших в теоретическом отношении. *«Сейчас я заигрываю (I am flirting) с идеей предпринять небольшую работу по группировке всех американских языков (скажем, к северу от тарасского или уабе), морфологически и генетически».* Несмотря на всю амбициозность исследования, которое *«должно представлять большой интерес для этнологов»*, обеспечив его *«определенным историческим бэкграундом»*, оно вполне осуществимо, — *«если есть метод».* В отличие от Радина у Сэпира на этот момент он уже имелся.

*«Шесть больших групп»* это — эскимосско-алеутская, на-дене, алгонкино-вийот-юрок (с ними увязывались кутенэй и вакашско-сэлишско-чемакуаские), пенутийская (включая калифорнийскую и орегонскую группы, чинук, цимшиан), хокско-сиуская (т. е. хокские, юки, коауильтекский, керес, «группа читимаш» (Chitimach group), сиуско-маскоги-ючи и ирокезо-пони), юто-ацтекская (вместе с тейва-кайова?). — *«На такой основе, если доказать ее достоверность, можно почти что увидеть перемещения населения».* Правда, исчерпывающего доказательства опять не последовало; безусловно, основные силы и время отвлекла работа над книгой.

Глубокий контраст обнаруживался между пенутийской и хокско-сиуской большими группами: языки первой показывали тенденцию к флективности, второй — к агглютинации. И эскимосско-алеутская, и на-дене, и хокско-сиуская объединяли сильно полисинтетические языки (две последние поместил в одну «подгруппу» также и Радин), но, по мнению Сэпира, это не означало их генетического родства. Подобно тому, как полисинтетизм чинук — скорее всего результат более позднего развития на совершенно иной аналитической основе, указанное явление могло появиться и в изначально корнеизолирующих языках на-дене (*«мне кажется это практически определенным»*). Юто-ацтекские же могли оказаться плодом древнего смешения (old intermixture) пенутийских и хокско-сиуских, а алгонкино-вийот-юрок — специализированным боковым ответвлением (offshoot) пенутийских.

Теперь он мог подытожить все свои наблюдения конкретных примеров ареального взаимодействия языков в Северной Америке: юки, отнесенный к хокско-сиуским, определенно подвергся некоторому пенутийскому влиянию; точно также, вероятно, и яна; наоборот, пенутийские майду и такелма находились под несомненным воздействием хокских; и во всех случаях прослеживалось шошонское влияние.

Уже на данном этапе Сэпир готов был сделать вывод о том, что примерно те же группировки можно выделить на основе фонетических данных. Так, эскимосско-алеут-

ские языки обладали лишь одним взрывным согласным — *p* (*f* и *v* являлись у них заимствованием); на-дене и хокско-сиуские — изначально тремя (*p'*, *b* и *p'*, причем губные часто отсутствовали, как в на-дене, так и в ирокезских), пенутийские определенно (*p?*, *p*), а алгонкино-вийот-юрок с долей вероятности (*p?*, *\*p*), — только двумя; юто-ацтекские — также двумя, но различавшимися по иному принципу (*p*, *pp*). Для него было очевидно также, что эскимосско-алеутская группа стояла особняком; сдвиг в на-дене и пенутийских осуществлен параллельно, но на потрясающем удалении друг от друга; алгонкино-вийот-юрок и пенутийские обнаруживали соответствие; юто-ацтекские обладали собственными особенностями.

Разумеется, в фонетической сфере также происходили «вторичные нарушения»: изначальные системы чумащ, салинских и юмских, по-видимому, упростились (вполне возможно, что под юто-ацтекским влиянием); в то время как у такелма наличествовали целых три серии (наличие в нем *p'*, как и инструментальных префиксов, объяснялось влиянием со стороны шастских); три серии развил квакиутль, но в отличие от родственных нутка и сэлишских.

Далее, в послании Фрэнку Спеку Сэпир лишь намекнет о своих «далеко идущих идеях», которые «поставят наших друзей-консерваторов на уши». Каждый из шести больших «стволов» — они вот-вот обретут окончательные очертания, — «как мне кажется, представляет собой генетическое единство» (ES/FS 09.10.1920; Darnell 1990: 123).

Последним предварительную версию «шести фил» получит Крёбер (ES/ALK 04.10.1920; Golla 1984: 347–351). К тому времени «Язык» будет уже почти готов, и его автор снова, но, как оказалось, ненадолго, вернется в полной мере к вопросам таксономии: «[Ч]увствую, что хочу решить действительно большую проблему. Как только моя маленькая книжка будет закончена, я планирую составить действительно исчерпывающий вопросник по морфологическим и фонетическим особенностям языков Мексики и Северной Америки» (ES/ALK 04.10.1920; Golla 1984: 347). Прежде всего, Сэпиру хотелось проследить распространение следующих черт: употребление синтаксических падежей, противопоставление глаголов активных и стативных, образование диминутива при помощи *-tsi* или *-si* и т. д. Только затем он планировал применить «лексические тесты», чтобы выверить получающиеся группы. — «[У] меня лучшее чувство перспективы, более четкое, чем у большинства других, представление о пережитках старого vs. развитиях вторичных черт. Пол, например, не может сделать эту работу по-настоящему убедительной, потому что не знает, как оценивать; все является рыбой в его сети» (ES/ALK 04.10.1920; Golla 1984: 347).

Очевидно, что головоломка Радина заставила-таки Сэпира задуматься о возможности более широкого генетического единства американских индейских языков: «[П]ризнаю определенные обещающие «протоамериканские» черты (такие как отрицательные [частицы] *\*ka*, *\*ku*; диминутив на *\*-tsi*; мест. 1-го лица ед. числа N-; множественное и фреквентативное -l)» (ES/ALK 04.10.1920; Golla 1984: 348). Вместе с тем в письме снова повторены шесть фил и примерно в том же составе. Сэпир по-прежнему не знал, куда включить зуны, у которого еще в 1915 г. Крёбер начал искать сиуские и хокские связи («Это ни юто-ацтекский, ни атапаскский, ни алгонкинский. Я даже попробовал маскогские» — ES/ALK 28.11.1915; Golla 1984: 199). Против названия изолята керес — другого языка пуэбло, отнесенного к хокско-сиуским, Сэпир оставил пояснение — «решительно, судя по тому, что мне пишет Босас». Он все еще колебался относительно статуса модок-молале-сахаптин, которые представлялись ему в это время либо переходными между филами пенутийской и юто-ацтекской, либо даже «выбросом» (outlier) последней.

В сравнении с вариантом, изложенным в письме Лоуи, языки Мексики теперь были действительно задействованы активнее. — «[П]редставь, майя может принадлежать к F

(т. е. хокско-сиуским — И. К.)» (ES/ALK 04.10.1920; Golla 1984: 349). Сэпир всерьез рассматривал гипотезу Радина (Radin 1919b) о принадлежности уабе к со́ке-михе, считая также, что радинские материалы «доказывают абсолютно» сходство миштек-сапотек с отомí; а тарасский язык мог оказаться еще одним «выбросом» юто-ацтекской филы.

Как и в прошлый раз, Сэпир принялся фантазировать относительно более глубоких во временном отношении связей отдельных фил. Он снова предположил, что алгонкино-вакашские могли быть высокоспециализированным полисинтетическим ответвлением пенутийских; и что юто-ацтекские, возможно, происходили из «смешанного языка» (Mischsprache) — на основе пенутийского и хокско-сиуского. Но на-дене и в таком контексте стоял особняком (ES/ALK 04.10.1920; Golla 1984: 349).

Исследователя еще больше заворожило широкое распространение полисинтетизма в Северной Америке. Он продолжал настаивать, что само это явление возникало несколько раз и в различных местах — «генетически это не высокоценный критерий, но выражает скорее определенную крайнюю тенденцию к синтетическому выражению, как бы мы ее ни объясняли», и «оно интересно скорее для психологии, чем для истории». Лично ему для текущих целей куда ценнее казалась такая «темная особенность», как явное преобладание корней на начальную гласную в чимарико, салинских, керес, помо, шаста, ирокезских и маскогских. Или же противопоставление местоимений переходных и непереходных, активных и стативных — для пенутийских языков почему-то характерно первое, тогда как для хокско-сиуских — второе.

Было повторено, что наименее полисинтетическими в Северной Америке и ближайшими к «нашему флективному типу» являлись пенутийские. Как контраст, инкорпорирующий тип чинук, по Сэпиру, развился из сильно аналитической формы, которая была «сломана» (had broken down) влиянием языков типа такелма-кус. Подобным же образом в эскимосско-алеутских и алгонкино-вакашских полисинтетизм вырос вокруг старого флективного ядра, а в на-дене развился на «аналитической не-флексирующей (можно сказать, изолирующей) основе, да, в самом деле!» Ведь синтетическая форма в тлинкит, хайда и атапаскских при анализе легко разбиваема на односложные фрагменты, характеризующиеся значительной индивидуальной фонетической и функциональной самостоятельностью. — «[И] все же полисинтетизм на-дене — самых увлекательных из всех языков, когда-либо изобретенных, — в некотором роде выиграл собственную флективную систему!» В то же время в хокско-сиуских он — явно агглютинативного типа. «Психически» полисинтетический яна отличался, «как солнце от луны», от формально такого же по типу такелма. «Разве это тройственное развитие (threefold development) полисинтетизма не составило бы изящный кусочек в лингвистической теории?» (ES/ALK 04.10.1920; Golla 1984: 349–350). Интересно, что в примере с яна и такелма мелькнуло сэпировское понимание связи языка и мышления вполне в духе рассуждений следующего десятилетия о лингвистической относительности.

**На-дене и сино-тибетские языки.** Октябрьское письмо Крёберу знаменательно также в другом отношении, поскольку однозначно показывает жгучее стремление Сэпира (после письма к Лоуи прошло меньше месяца) вывести «шесть фил» за пределы Северной Америки. С первых строк очевидно, что внимание Сэпира привлекли австронезийская гипотеза В. Шмидта, статус мон-кхмерских языков и проч. И помимо новых мексиканских связей, рядом с эскимосско-алеутской семьей он поместил в скобках со знаком вопроса чукотско-корякскую, как, возможно, в нее входящую. Точно так же была осуществлена попытка перекинуть на Дальний Восток мостик от на-дене (ES/ALK 04.10.1920; Golla 1984: 350):

«Меня прямо сейчас заинтересовала еще одна большая лингвистическая возможность. Я с трепетом говорю о ней, хотя эту зародышевую идею носил с собой уже много лет. Я не считаю, что на-дене принадлежит к другим американским языкам. Я ощущаю ее, как крупную внедрившуюся общину (a great intrusive band), которая быть может разорвала старую эскимосско-вакашско-алгонкинскую непрерывность. И я решительно чувствую старую квази-изолирующую основу. Еще есть тон, который кажется старым (высокий и низкий) — я почти уверен, что тоном атабаскские и хайда подобны тлинкитскому. Короче, не считай меня ослом, коль я всерьез развлекаюсь идеей старого индо-китайского (сино-тибетского — И. К.) ответвления в Северо-западной Америке».

Сэпир отчитывался, что «уже просмотрел внимательно две тибетские грамматики <...> и нашел в тибетском языке почти того рода основу, из которой могла развиться генерализованная [семья] на-дене». Дальше перечислялись первые примеры сходств: тибетские постпозитивные *ta* 'в' и *di* 'к, у' используются для подчинения глагола, точно так же как в атапаскских и тлинкитском языках; как в тлинкитском, так и в тибетском переходный глагол как таковой явно пассивен; в тибетском каузативные или переходные глаголы отмечены префиксом *s-*, в языке тлинкит — *si-*; аблаут атапаскско-тлинкитского глагола практически повторяет соответствующую парадигму тибетского (наст. время — *byed* 'делаю', претерит — *byas*, будущее — *bya*, повел. накл. — *byos*):

«Я грежу? По крайней мере, я знаю, что дене гораздо ближе к тибетскому, чем к сиускому. Такие вещи, как инструментальные префиксы, с которыми так много возился Пол, никуда нас не приводят, поскольку префиксы на-дене этого типа — просто поздние композиции и даже не согласуются между собой (практически любое существительное могло бы стать инструментальной приставкой; что до этого, так полагаю, что китайский способен на любые вещи, как то, огонь [+ ] убивать, «убивать огнем»). Всем этим скорее озадачен».

Р. Дарнелл продвигает взгляд, согласно которому само обращение к возможным китайским связям на-дене произошло благодаря сотруднику Джесуповской экспедиции Бертольду Лауферу, четырежды побывавшему в Китае и являвшемуся, в сущности, единственным тогда в Соединенных Штатах профессиональным китаеведом и тибетологом (Darnell 1990: 127–131). Правда в ранней книге о Сэпире она называет письмо своего героя к Лоуи от 15 февраля 1921 г. «первым явным заявлением» (the first explicit statement) так называемой индо-китайской гипотезы. В подтверждение этой же мысли на последующих страницах она ссылается на сэпировские контакты с Лауфером после указанного срока. При этом, видимо, по недоразумению, несколько предложений из более раннего, разбираемого нами письма Сэпира к Крёберу дословно перенесено у Дарнелл в цитату из письма его же Лауферу от 1 октября 1921 г., ср.: “I do not feel that Na-dene belongs to the other American languages” (‘Я не считаю, что на-дене принадлежит к другим американским языкам’); “I feel it as a great intrusive band” (‘Я ощущаю ее, как великую внедрившуюся общину’); “<...> do not think me an ass” (‘<...> не считай меня ослом’) и проч. (Darnell 1990: 128).

Далее она изображает все так, будто Крёберу предположение о родстве на-дене с языками Азии покажется крайне неубедительным, он станет тянуть с ответом, откликнувшись лишь после деликатного напоминания своего приятеля и то больше соблюдая политес. Но это якобы снова произойдет только в 1921 г.: 24 ноября — повторное письмо Сэпира, затем 26 ноября — запоздалый отклик Крёбера (Darnell 1990: 129). В действи-

тельности же крёберовский ответ с реакцией по всем буквально пунктам, хоть и в присутствующей ему краткой, афористичной манере, последовал еще в декабре 1920 г. В нем, в частности, калифорнийский антрополог склонял Сэпира включить в алгонкино-вакашскую филу еще один возможный изолят — вымерший язык беотук о-ва Ньюфаундленд: «Предполагаю, что эти люди были ранней ветвью алгонкинов, которые изолировались на острове, и к тому времени, когда к ним присоединилась более поздняя волна (алгонкиноязычных микмак — И. К.), они значительно разошлись». При этом проблеме родства на-дене посвящался отдельный абзац, содержащий наряду с другими такие откровения и гиперболы (ALK/ES 27.12.1920; Golla 1984: 358):

*«У меня было бы меньше веры в твои открытия, если бы ты объединил их с чем-то американским. Что касается азиатского происхождения на-дене, я слишком невежественен, чтобы обладать мнением, хоть и сильно чувствую предельно моносилабический и, по существу, изолирующий характер атабаскских».*

Ясно, что «индо-китайская» гипотеза родилась до, а не после декабря 1920 г., когда предварительные результаты вынашиваемого семь лет проекта по группировке североамериканских языков были публично представлены на чикагском собрании Американской ассоциации содействия развитию науки. В более современных работах Дарнелл вроде бы устраняет допущенную оплошность и даже приводит очередное письмо — от 3 октября 1920 г. К. Уисслеру, неопровержимо доказывающее, что Сэпир обсуждал-таки уже тогда наиболее смелую из своих догадок. Как он сам заявлял, его вряд ли бы удивило, что пра-сино-тибетский, «изначально северный язык, на расстоянии броска камня от определенных индивидов, которые должны были совершить путешествие в Северную Америку», может обнаруживать столько близости с на-дене. — «Быть может не зря индейцы цимшиан говорят, что хайда звучит как китайский» (Darnell 2021: 167–168)!

**«Взгляд с высоты птичьего полета».** Итак, на собрании Американской ассоциации содействия развитию науки ничего не будет сказано по поводу китайско-тибетского родства на-дене. В последний момент антропологи, участвовавшие в форуме, перенесут свою встречу в Балтимор, Сэпир же поедет в Чикаго, как и планировалось изначально. Это и есть причина, почему в аудитории соберется немного подлинных знатоков вопроса. Чикаго был важен для него, подчеркивает Дарнелл, из-за желания встретиться с Лауфером и продолжить обсуждение связей языков Старого и Нового Света. Сэпир назовет свой доклад «Взглядом с птичьего полета на американские языки к северу от Мексики». Главные выводы выступления будут проиллюстрированы и суммированы в виде карты и таблицы новых семей (Darnell 1990: 123):

I. Эскимосско-алеутская			
II. Алгонкино-вакашская:			
1.	Алгонкино-ритвская:	2.	Кутенэй
	(1) алгонкинская	3.	Мосская («мосанская», Mosan, resp. Wakashan-Salish):
	(2) беотук (?)		(1) вакашская (квакиутль-нутка)
	(3) ритвская:		(2) чимакуаская
	(а) вийот		(3) сэлишская
	(б) юрок		

Таблица 1. Шесть языковых фил в Северной Америке, вкл. Мексику (по Э. Сэпиру)



III. На-дене:			
1.	Хайда	2.	Континентальные на-дене: (1) тлингит    тлингит (Tlingit) (2) атапаскская    атабаскская (Athabaskan)
IV. Пенутийская:			
1.	Калифорнийские пенутийские: (1) мивок-костаноская (2) йокатс (3) майду (4) винтун	3.	(3) калапуя
		4.	Чинук
		5.	Цимшиан
2.	Орегонские пенутийские: (1) такелма (2) береговые орегонские пенутийские: (а) кус (б) сайусло (в) яконские	5.	Пенутийские Плато (Plateau Penutian): (1) сахаптин (2) вайилатпу (модала-кайюс) (3) лутуами (кламат-модок)
		6.	Мексиканские пенутийские: (1) михе-соке (2) уаве, уабе (Huave)
V. Хокско-сиуская («хока-сиу», Hoka-Siouan):			
1.	Хокско-коауильтекская: А. Хокская: (1) северные хокские: (а) карок, чимарики, шаста-ачомави (б) яна (в) помо (2) вашо (3) эсселен-юмские: (а) эсселен (б) юма (4) салинско-сери: (а) салинские (б) чумаш (в) сери (5) текистлатекские (чонталь) Б. Субтиава-глаппанек (Subtiaba- Пар(р)анес): В. Коауильтекские: (1) тонкава (2) коауильтеко:	2.	(а) собственно коауильтеко (б) котонаме (в) комекрудо (3) каранкава Юки
		3.	Керес
		4.	Туникская: (1) туника-атакапа (2) читимаша
		5.	Ирокезо-кэддоская: (1) ирокезские (2) кэддоские
		6.	Восточная группа: (1) сиуско-ючи: (а) сиуские (б) ючи (2) натчез-маскогские: (а) натчез (б) маскогские (в) тимуква (?)
VI. Ацтекско-таноская («ацтекско-таноанская», Aztec-Tanoan):			
1.	Юто-ацтекская: (1) науатль (2) пима (3) шошонские	2.	Кайова-таноская: (1) таноский (2) кайова
		3.	Зуны (?)

Таблица 1 (продолжение). Шесть языковых фил в Северной Америке, вкл. Мексику (по Э. Сэпиру)

*Эскимосско-алеутские* — языки полисинтетические и флективные; последовательно суффицирующие. В них хорошо развиты формальные аспекты глагола (наклонение, лицо); фундаментальное значение имеет переходность/непереходность; наличествуют локатив и два синтаксических падежа. Существительные во множественном числе и местоименные элементы обладают формальным, а не только материальным значением. Отсутствуют: редупликация, внутренняя модификация корня и словосложение (именная инкорпорация).

*Алгонкино-вакашские* — языки полисинтетические и флективные, из них вакашские — преимущественно суффицирующие. В отчетливо флективных алгонкинских тоже имеются суффиксы, как глагольные (модальные и местоименные), так и именные (показатели рода, числа и обвиатива); причем они намного старше префиксов, которые по происхождению — очевидно, проклитики. Вакашские явно менее флективны, но допускают важные внутрикорневые модификации. И в тех, и в других глаголы классифицируются как субъектные и объектные; падежи едва ли представлены (обвиатив выражен синтаксическим падежом, и вроде бы есть еще один местный падеж). Хорошо развита редупликация, словосложение же в обычном смысле отсутствует (частично присуще алгонкинскому глаголу), инкорпорация — умеренная, в алгонкинских. Весьма богаты второстепенные элементы, «суффиксы» с конкретным значением: локативным, инструментальным, адвербиальным, собственно глагольным.

*На-дене* — языки слабо полисинтетические, изолирующие в своей основе, выработавшие, однако, квази-флективность; умеренно суффицирующие, если принимать в расчет бытование определенных частиц в роли «префиксов», в особенности в хайда. Речь об односложных элементах, расположенных в определенном порядке, которые сливаются скорее психологически, чем морфологически; «слово» здесь фактически находится на полпути между коротким предложением и подлинным словом». Не происходит никакой настоящей спайки корня и прикрепленных элементов, если только не образуются новые основы. Истинная основа — односложная, типа С + V (вероятно, также С + V + носовой). Вторичные фонетические процессы вызвали развитие глагольных форм, но слияние субъекта (местоимения) с модальными «префиксами» в действительности не зависит от изменений основы глагола. Важны «залог» и «аспект», время не столь принципиально. Глаголы делятся на активные и стивные (включая переходный объекта и пассивный субъекта). Хорошо развиты послелого, преимущественно именного происхождения. Композиция тщательно разработана. Нет никакого редуплицирования или формального развития рода, падежа, числа. Послелого участвуют в образовании «относительных» форм, а также в номинации глагольных форм. Наличествуют тоны, высокие и низкие.

*Пенутийские* — языки не полисинтетические, а флективные; в первую очередь используют суффиксы; префиксы, если даже и встречаются, имеют явно вторичное происхождение. Суффиксальные элементы несут почти исключительно формальное значение и тесно спаяны с корнем. Местоименные суффиксы, элементы, указывающие на падеж, множественное число, время и аспект, а также на залог (глагола) придают формальный вид словам. Исключительно важны внутрикорневые изменения, включая редупликацию (иногда концевую) и наращение гласной. Наиболее характерны корни типа С + V + С<sub>1</sub> + V. Инкорпорация обычно не развивается, в целом глагол не терпит никаких усложнений; композиция либо отсутствует, либо развита весьма умеренно. Глаголы классифицируются на (субъектно) непереходные и (субъектно) переходные и на объектно переходные — вероятно, изначально особую, третью категорию. В такелма, майду и винтун присутствуют тоны, по-видимому, восходящего-нисходящего типа; значение их еще не ясно. Чи-

нук представляет самостоятельное полисинтетическое развитие на основе разрушенного пенутийского аналитизма. Приверженность их группе очевидна из лексических данных и рудиментарных особенностей. Цимшиан находился под сильным влиянием алгонкино-вакашских языков.

*Хокско-сиуские* — языки полисинтетические и агглютинативные; в них практически отсутствует какая-либо тенденция к флективности, несмотря на формальную детализацию и случайные внутрикорневые модификации. Представлены как префиксы, так и суффиксы, но характерны первые в виде более формальных элементов, в особенности местоименных (в глаголе). — «Яна обладает вторичными чертами». Вообще же способы аффиксации чрезвычайно многообразны, особенно в хокских, где имеются инструментальные префиксы (как и в сиуских), локальные суффиксы и вторичные глагольные корни, основанные на древнем сложении. Глаголы бывают активными и стативными (чимарико, сиу, ирокезские). Обычны послелогии. В этой «группе» прекрасно развились инкорпорация и настоящее словосложение, однако, редупликация не настолько типична, как в алгонкино-вакашских и пенутийских языках, и в ряде случаев вовсе отсутствует. Внутрикорневые модификации вообще не обнаруживаются, за исключением яна. Наиболее частотны корни типа  $V + C + V_1 (+ C_1 + V_2)$ . Восходящие и нисходящие тоны отмечены в ачوماви, мохэви, вероятно, в помо, но они нуждаются в дальнейшем изучении.

*Ацтекско-таноские* — языки умеренно полисинтетические и в лучшем случае слабо флективные; суффиксированные. Префиксы — либо бывшие проклитики, либо старые сложные корни. Суффиксы — формального значения, как и в пенутийских языках. «Возможно, полисинтетичность [ацтекско-таноских] сформировалась на основе IV [пенутийской] с помощью процессов простого сложения, и может быть благодаря контакту с V [хокско-сиускими]. Глагол: субъектно-объектный как в II [алгонкино-вакашских], заметно отличающийся от I–IV, III–V». Частое явление представляет редупликация, как и инкорпорация и словосложение. Довольно обычны послелогии. Резко различаются имя и глагол. Падежи развиты, но слабо. — «Все во всем, скорее смешанного, а не специализированного типа». Характерным типом корня является  $C + V + C_1 + V_1$ . Тоны зафиксированы в языках кайова-таноской ветви, но их роль до сих пор непонятна.

Чувствуется, что характеристики перечисленных типов пропитаны идеями только что завершенной Сэпиром книги «Язык». Ключевая из них — дрейф языка. Очевидно, в голове он держал своего рода заготовки профилей исторического развития для каждой из выделенных семей, увы, так нигде детально и не раскрытые, тем более обоснованные привычными для компаративистики средствами, будь то «лексические тесты» или надежно восстанавливаемые фонетические закономерности. Так, он упрямо трактовал ацтекско-таноскую филу как смешанную по происхождению. И в этой связи интересна еще одна авторская ремарка (Sapir 1990–2008, vol. 5: 86): «Впечатляет меня (пра-ацтекско-таноский язык — И. К.), как старый пенутийский, на который плотно наложился хокский (тот же процесс, который имел место в майду, но бесконечно более древний)».

В своем ответе непосредственно на радинскую гипотезу, америндскую, как сейчас ее принято обозначать, Сэпир продвинулся не сильно, лишь четче сформулировав четыре «протоамериканские» черты, которые вчерне перечислял уже в письме Крёберу: «1. Постоянство  $n$ - 'я',  $m$ - 'ты'. 2. Отрицательные  $ka$ ,  $ku$ . 3. Длительно-множественно-итеративное  $-l$ . 4. Уменьшительное  $-si$ ,  $-tsi$ » (Sapir 1990–2008, vol. 5: 86). Он допускал, что три семьи из шести — эскимосско-алеутскую, алгонкино-вакашскую и пенутийскую — в будущем можно будет объединить в группировку более высокого уровня, но это не казалось на-дене. Видно также, что он по-прежнему сомневался в оправданности выделения хокско-сиуской (Sapir 1990–2008, vol. 5: 86): «Группа показывает скорее малую стабильность».

О некоторых колебаниях во взглядах Сэпира на вариант классификации, впервые представленный академическому сообществу, свидетельствуют коррективы, внесенные его рукой уже в готовый машинописный текст доклада, точнее в таблицу. Прежде всего, им была добавлена, в основном следуя Радину, новая категория «Расширения в Ц. Америке»: «пенутийские: михе-соке, <Ниано> (уабе)??, шинка?? <...> Хокско-сиуские: майя, отоми-миштек-сапотек». И наоборот, категория «Неразмещенные» подлежала ликвидации: беотук добавлен к алгонкино-ритвским языкам, зуньи — к ацтекско-таноским, сахаптин-вайилатпу-лутуами — к пенутийским, при этом кламат выведен из вайилатпу в отдельную группу — лутуами. Дальше исследователь изменил вийот-юрок на «ритвские» и соответственно «алгонкинские» — на «алгонкино-ритвские». В рамках пенутийской филы костаноская группа была объединена с мивок, а чинук (первоначально — в составе т. н. орегонских пенутийских) получил статус отдельной сестринской ветви. Но больше всего изменения касались хокско-сиуских языков: хокские и коауильтекские соединялись вместе; чумаш-юмские (в составе хокских) разъединялись (чумашские сближались с салинскими и сери, юмские — с эсселен); яна все-таки был добавлен в группу северных хокских. Точно так же атакапа с туника теперь образовывали более тесное, чем с читимаша, единство внутри туникской ветви. К ним Сэпир подтянул ирокезо-кэддо-ские, которые прежде стояли в самом конце списка, на противоположном краю от хокских, и таким образом место, наиболее удаленное от последних, досталось сиуским, ючи и маскогским, поименованным в совокупности «восточной группой». К маскогским он причислил также натчез и малоизвестный язык тимуква, вымерший еще при испанцах; а нутка, вашо и шошонские языки выделил, объединя соответствующие названия кружками, так никогда и не прояснив почему. Финальной фразой конспекта стало оптимистическое: «Движения населения должны быть выявлены с помощью лингвистического исследования» (Sapir 1990–2008, vol. 5: 91–92).

«Взгляд с высоты птичьего полета» не был бы презентован столь поспешно, если бы не вызов, брошенный Радином, нуждавшийся в какой-то реакции. Многие из изложенного в докладе предстояло годами кропотливо дорабатывать, и непосредственно вслед за чикагским мероприятием Сэпир ограничился лишь публикацией анонса в “*Science*” в 60 строк, из которых 20 занимала таблица (Sapir 1921a: 408), в точности, как раньше делал Харрингтон и он сам, когда возникла необходимость срочно застолбить идею о принадлежности вашо к хокским языкам и галфских к хокско-сиуским. Публикация отражала менее продвинутую версию, резюмируя текст доклада, которого еще не коснулись авторские правки, поэтому в ней сохранились отдельно хокская и коауильтекская ветви, не были распределены по семьям изоляты зуньи, беотук и сахаптин-вайилатпу-лутуами, как и отсутствовали названия «ритвская» и «восточная» — последнее применительно к объединенной сиуско-ючи-маскогской ветви.

Отчасти в ближайшее пятилетие Сэпир на самом деле займется обоснованием намеченных ранее генеалогических гипотез. Не считая истории с хокско-коауильтекским родством, другими такими сюжетами будут: фонология и фонетические законы, тоны, типы корней, суффиксация; исконная лексика языков, место которых в классификации все еще оставалось шатким: салинских, северного яна, ритвских в Калифорнии, хайда и чинук на Северо-западном побережье, субтиава в Никарагуа и др. (Sapir 1921c, 1922, 1923a, 1923b, 1925a, 1925b, 1926). Всего лишь через полтора месяца после Чикаго он спешит сообщить Лоуи о достигнутом прогрессе в исследовании возможных внешних родственных связей на-дене (ES/RL 15.02.1921; Lowie 1965: 44–46):

«<...> [С]читаю само собой разумеющимся, что Америка была заселена рядом исторически различающихся волн. В настоящее время я придерживаюсь мнения, что

*волна на-дене является самой поздней из всех по целому ряду причин, среди которых не менее интересным является тот факт, что она наиболее резко контрастирует с эскимосскими и южными языками. <...> Мои представления об этих языках постепенно проясняются, и я предвкушаю самое увлекательное задание по распутыванию истории группы».*

Вроде бы процедуре такого «распутывания истории» больше ничего не должно мешать. Сэпир будет убежден в том, что результаты получены им с соблюдением всех научных требований, и что именно его методике присуща особая глубина и широта видения в сравнении с инструментарием, которым пользовались его современники — европейские компаративисты, например, Карл Майнхоф: «[Н]е могу полностью следовать Майнхофу. <...> Я решительно не думаю, например, что он доказал, что фуль и готтентотский хамитские. <...> [М]оим собственным материалом на-дене гораздо более консервативен, чем материал Майнхофа» (ES/RL 15.02.1921; Lowie 1965: 44–46). А уже осенью он в азарте раскрывает Крёберу свои ближайшие планы на этот счет: «Давно хотел написать тебе о на-дене и индо-китайском, но мои сведения накапливаются так быстро, что тяжело сесть и выдать идею» (ES/ALK 01.10.1921; Golla 1984: 374). Сэпир намеревался реконструировать пратапаскский язык и подготовить сравнительное исследование на основе уже собранной им лексики, представляющей все ветви на-дене — «около 300 сопоставимых радикальных элементов, к которым я постоянно добавляю». Две другие части воистину титанического труда должны быть посвящены морфологии и фонологии, но прежде он собирается «опубликовать специальные статьи по избранным частям грамматики на-дене, например, некоторым архаичным послелогам; или указательным основам; или общим моментам синтаксиса» (ES/ALK 01.10.1921; Golla 1984: 376).

В это особенно длинное письмо, машинописная копия которого будет послана Лауферу, войдет не только изложение прожектов, несбыточных, как ясно сегодня, но также примеры когнатов, скажем, почти очевидных: тлинкит *k'a* ‘поверхность’, навахо *k'ā* ‘то же’, тибет. *k'a* ‘то же’; кит. *t'an* ‘уголь’ и хайда *s-t'an* ‘то же’; др.-кит. *ti* ‘фазан’ и атапаскское *di* ‘куропатка’ и т. д. и т. п. — правда числом всего 5, но «[э]то только капля в море» (ES/ALK 01.10.1921; Golla 1984: 376). Антрополог сосредоточил свою энергию на развенчании предубеждений («каменной преткновения»), более всего мешающих признанию его открытия, которых насчитывал тоже три — это «[н]еспособность осознания того весьма исключительного типа языка, к которому принадлежит на-дене»; и «что языки на-дене и на треть не настолько синтетические, как кажется»; как и предубеждение относительно «природы самого индо-китайского языка». Он повторял, не уставая: контраст между языками на-дене и эскимосскими, алгонкино-вакашскими и проч. просто огромен. И ему очень хотелось доказать изначальный корнеизолирующий характер на-дене на манер китайского, и что предшественники, например, Дж. Суантон, работавший с тлинкит, переусердствовали по части выделения в них аффиксов, которые в действительности являют собой независимые корни, даже «маленькие глаголы». Китайский же язык представляет «весьма вторичное развитие», поэтому более раннюю стадию куда лучше отражает тибетский, действительно похожий на на-дене. В нем, как и в хайда, помимо морфологических сходств, которые Сэпир уже отмечал годом ранее, фундаментальную функцию выполняет имя, а глагол — не более чем деноминативная структура (ES/ALK 01.10.1921; Golla 1984: 376):

*«[Ч]тение тибетского текста дает тебе то же ощущение, что и чтение текста на хайда. <...> [С]ходство в ощущениях между тибетским и на-дене, по крайней мере, как между латынью и английским, а может и ближе».*

Но были в послании и такие строки, которые свидетельствовали, что его автора, выражаясь словами Краусса, «занесло далеко за пределы любых объективно оправданных выводов» (Krauss 1973: 963; Sapir 1990–2008, vol. 6: 139). Так в постскрипуме Сэпир писал, можно представить, задыхаясь от волнения, как наткнулся на своеобразный «групповой параллелизм» (sort of group-parallelism) больших семантических гнезд корней \**ti*, \**li*, предположительно общих для языков обеих сверхсемей — «не могу устоять перед искушением дать как-то поживее идею о замечательном способе, которым лексические элементы переплетаются в на-дене и индо-китайских» (ES/ALK 01.10.1921; Golla 1984: 377).

В любом случае многое из этого выглядело весьма перспективно. Однако из запланированной серии публикаций, как частенько бывало у Сэпира, выйдет всего лишь сравнительно небольшая статья об образовании в атапаскских языках относительных форм прилагательных, в сущности, отглагольных, с помощью т. н. модальных префиксов, по крайней мере часть которых имела и в сино-тибетских (Sapir 1923c). Итогом же пятилетней работы станет опять-таки заметка в приложении к “*Science*” репортажного типа (Sapir 1925c: xiii):

*«Д-р Сэпир открыл не только то, что индейцы группы надине (the Nadine groups) говорят с тональным акцентом, повышающим или понижающим голос, <...> сходным с тоновыми особенностями в раннем китайском, но также, что значение определенных слов идентично. Далее, он обнаружил факт, что индейцы сохранили префиксы и суффиксы, которые давно исчезли из китайской речи, но явно различимы в ранних формах».*

Совсем скоро исследователь войдет в турбулентный период. Уже в постпостскрипуме письма про «индо-китайскую» гипотезу он делился, что его супруге, болевшей последние годы, «до сих пор еще не хорошо, и она может быть никогда не будет прежней». «Эта мысль бросает тень на все мои планы и надежды» (ES/ALK 01.10.1921; Golla 1984: 382). В 1924 г. Флоренс Сэпир не станет. Эдуард подружится с Рут Бенедикт и заведет скоротечный роман с Маргарет Мид, затем рассорится с обеими. В 1925 г. к большой радости он получит-таки место ассоциированного профессора в Чикагском университете. Там антрополог продержится очередное пятилетие, чтобы затем уехать в Йель — финальный пункт своей карьеры и жизни.

В последние 15 лет Сэпир превратится из теоретика преимущественно в эмпирика, интересующегося частными вопросами изучения американских индейских языков. Когда в 1929 г. ему предстоит написать статью о языках Северной и Центральной Америки для Британской энциклопедии (Sapir 1929), Сэпир не найдет ничего лучшего, чем просто изложить кратко свой доклад 1920 г. Никаких новых данных и идей. После этого он, насколько известно, никогда не вернется к языковой классификации, а по утверждению Мэри Хаас (Haas 1964; Voegelin, Voegelin 1977: 310), будет, по крайней мере, одну из своих шести фил, хокско-сиускую, вообще называть годной разве что для мусорной корзины (“wastepaper basket” group).

К проблематике систематизации индейских языков охладает также Крёбер, долгое время служивший для Сэпира сильнейшим внешним стимулом. Начиная с 1924 г. он вовлечется в месоамериканскую, а затем андскую археологию. Не успокоится только Радин, тоже поменявший несколько университетов, но к 1930 г., относительно прочно осевший в Беркли.

**«Решение проблемы языков американских индейцев».** В Американском философском обществе среди бумаг Чарльза Вогелина сохранился примечательный документ

под названием “Solution of the American Indian Language Problem”, (APS, Mss. Ms. Coll. 68). В нем на двух страничках расписаны возможные генетические и исторические связи ряда семей индейских языков:

«А. Первичный — хокский субстрат

В. Канадская группа языков (алгонкин, кутенэй, вакашские, сэлиш). Из них алгонкин наиболее ранний по времени.

В'. Вторичная дифференциация вышеперечисленных через перемешивание с хокским и некоторые новые развития, дающие группу

в' Цимшиан, чинукский, такелма калифорнийские пенутийские и эти последние путем перемешивания со вторичным хокским, дающие

в" Юто-ацтекский

С. — Некий тибето-китайский язык, взятый на  $\frac{3}{4}$  в С. Ш. А.

с' — Атапаскский, отмеченный перемешиванием с языком Бр. Колумбии, т. е. вакашско-сэлиш на хокском субстрате

с" — Хайда-тлинкит, как вышеперечисленный, но с меньшим хокским субстратом

Д. Прямое влияние или вторжение китайского ок. 700–500 гг. до н. э., дающее тлапанек, сапотек-миштек, масатек, чинантек, чьяпанек, субтиава [–] все они на различных хокских или хокско-пенутийских субстратах

Пол Радин»

Очевидно, принимая допущение Сэпира о вероятном смешанном происхождении некоторых языков и языковых семей, по крайней мере, ацтеко-таноской филы, этот неутомимый пионер в установлении дальнего лингвистического родства пойдет значительно дальше и предложит аналогичные скрещенные модели также для цимшиан, чинукских, такелма и др. Так, согласно новой радинской схеме, впрочем построенной, видимо, целиком на интуиции, как и первая, атапаскские — плод от скрещивания «некоего тибето-китайского» и вакашско-сэлишского. Почти везде он обнаружит хокский субстрат. Надо полагать, хокская семья казалась ему наидревнейшим реликтом в Новом Свете, а несколько мелких семей и изолятов в Мексике — наоборот следами самого позднего передвижения населения из Азии (Китая), практически уже в историческое время. «Китай», «китайский», как и «США» названы так для простоты, и, разумеется, под ними подразумевались реалии соответствующих, куда более древних эпох. Вслед за хокскими в Америке оказались алгонкино-мосские(-вакашские) языки, а в результате смешения тех и других появились пенутийские (калифорнийские, орегонские, цимшиан и чинук), которые, еще более смешавшись с хокскими, дали начало юто-ацтекским. Что касается на-дене, то исследователь противопоставлял скрещенные, по его мнению, атапаскские тлинкитскому и хайда «с меньшим хокским субстратом». Обновленная версия классификации Радина датируется 23 февраля 1933 г. За ним и будет последнее слово.

Поразительно, но теории «таксономистов» описанного периода представлены в эпистолярном жанре больше, чем в журнальных статьях! Многие важные идеи так и дошли до нас в неразвернутом виде, и при жизни ученых были лишь продекларированы ими. Среди неопубликованного наследия находится, например, дене-китайский словарь Сэпира, включающий более 100 лексических сравнений, который до сих пор хранится в библиотеке Американского философского общества (Ms. 497.3 B63c Na20a.3, vol. 2). Будущие исследования позволят определить научную значимость этих и других собранных в те годы данных и верифицировать многие из высказывавшихся предположений. Однако уже сейчас видно, что некоторые из концепций, по крайней мере, Сэпира и Радина, толковались искаженно именно в силу крайне тезисного характера соответствующих публикаций.

## Список сокращений

ALK/ES — Alfred Louis Kroeber to Edward Sapir; ALK/Gifford — Alfred Louis Kroeber to Edward Gifford; APS — American Philosophical Society; ES/ALK — Edward Sapir to Alfred Louis Kroeber; ES/FS — Edward Sapir to Frank Speck; ES/RL — Edward Sapir to Robert Lowie; FB/ES — Franz Boas to to Edward Sapir; PR/ES — Paul Radin to Edward Sapir.

## References

- Achilli, Alessandro et al. 2003. Reconciling migration models to the Americas with the variation of North American Native mitogenomes. *Proceedings of the National Academy of Sciences of the United States of America* 110(35): 14308–14313.
- Bell, Amelia R. 1990. Separate people: speaking of Creek men and women. *American Anthropologist* 92(2): 332–345.
- Boas, Franz (ed.). 1911. *Handbook of American Indian Languages. Part 1. (Bulletin 40, Bureau of American Ethnology)*. Washington: Government Printing Office.
- Brinton, Daniel G. 1901. *The American race: a linguistic classification and ethnographic description of the Native tribes of North and South America*. Philadelphia: David Mckay.
- Buckley, Thomas. 1996. “The Little history of pitiful events”. The epistemological and moral contexts of Kroeber’s Californian ethnology. In: George W. Stocking, Jr. (ed.). *Volksgeist as Method and Ethic: Essays on Boasian Ethnography and the German Anthropological Tradition (History of Anthropology 8)*: 257–297. Madison: The University of Wisconsin Press.
- Campbell, Lyle. 1997. *American Indian languages: The historical linguistics of North America*. New York / Oxford: Oxford University Press.
- Darnell, Regna. 1990. *Edward Sapir: linguist, anthropologist, humanist*. Berkeley / Los Angeles / London: University of California Press.
- Darnell, Regna. 1998. *And along came Boas: Continuity and revolution in Americanist anthropology*. Amsterdam / Philadelphia: John Benjamins Publishing Co.
- Darnell, Regna. 2021. *The history of anthropology. A critical window on the discipline in North America*. Lincoln: University of Nebraska Press.
- Dixon, Roland B., Alfred L. Kroeber. 1919. *Linguistic families of California. (University of California Publications in American Archaeology and Ethnology 16)*. Berkeley: University of California Press.
- Dixon, Roland B., Alfred L. Kroeber. 1903. The Native languages of California. *American Anthropologist* 5(1): 1–26.
- Dixon, Roland B., Alfred L. Kroeber. 1913a. Relationship of the Indian languages of California. *Science* 37(945): 225.
- Dixon, Roland B., Alfred L. Kroeber. 1913b. New Linguistic families in California. *American Anthropologist* 15(4): 647–655.
- Gibbons, Ann. 2012. Genes suggest three groups peopled the New World. *Science* 337(6091): 144.
- Golla, Victor. 2003. Ishi’s language. In: Karl Kroeber, Clifton Kroeber (eds.). *Ishi in three centuries*: 208–228. Lincoln / London: University of Nebraska Press.
- Golla Victor (ed.). 1984. *The Sapir-Kroeber correspondence. (Survey of California and Other Indian Languages 6)*. Berkeley: University of California Press.
- Haas, Mary R. 1944. Men’s and women’s speech in Koasati. *Language* 20: 142–149.
- Haas, Mary R. 1964. California Hokan. In: William Bright (ed.). *Studies in California linguistics (University of California publications in linguistics 34)*: 73–87. Berkeley: University of California Press.
- Harrington, John P. 1910. On phonetic and lexic resemblances between Kiowan and Tanoan. *American Anthropologist* 12(1): 119–123.
- Harrington, John P. 1913. [Note on Chumashan and Yuman relationship]. *American Anthropologist* 15(4): 716.
- Harrington, John P. 1917. Work of Mr. John P. Harrington [Note on Washo and Chumashan relationship]. *American Anthropologist* 19(1): 154.
- Jany, Carmen. 2009. *Chimariko grammar. Areal and typological perspective (University of California publications in linguistics 142)*. Berkeley / Los Angeles / London: University of California Press.
- Krauss, Michael E. 1973. Na-Dene. In: Thomas A. Sebeok (ed.). *Linguistics in North America (Current trends in linguistics 10)*: 903–979. The Hague: Mouton.



- Kroeber, Alfred L. 1913. The Determination of Linguistic Relationship. *Anthropos* 8(2): 389–401.
- Kroeber, Alfred L. 1915. Serian, Tequistlatecan, and Hokan. *University of California publications in American archaeology and ethnology* 11(4): 279–290.
- Kroeber, Alfred L. 1925. *Handbook of the Indians of California*. Washington, D. C.: Government Printing Office.
- Kroeber, Alfred L. 1954. Robert Spott 1888–1953. *American Anthropologist* 56(2): 282.
- Kroeber, Theodora. 1961. *Ishi in two worlds*. Berkeley: The University of California Press. [Russian Translation: Kroeber, Teodora. 1970. *Ishi v dvou mirax. Biografija poslednego predstavitel'a indejskogo plemeni jana*. Moscow.]
- Kroeber, Theodora. 1970. *Alfred Kroeber. A personal configuration*. Berkeley / Los Angeles / London: University of California Press.
- Lowie, Robert H. (ed.). 1965. *Letters from Edward Sapir to Robert H. Lowie*. Berkeley, California (mimeo).
- Pope, Saxton T. 1920. The medical history of Ishi. *University of California publications in American archaeology and ethnology* 13(5): 175–213.
- Powell, John W. 1891. Indian Linguistic Families of America North of Mexico. In: *Annual Report 7, Bureau of American ethnology*: 1–142. Washington, D. C.: Government Printing Office.
- Radin, Paul. 1919a. The genetic relationship of the North American Indian languages. *University of California publications in American archaeology and ethnology* 14(5): 489–502.
- Radin, Paul. 1919b. The genetic relationship of Huave and Mixe. *Journal de la Société des Américanistes* 11: 489–499.
- Radin, Paul. 1924. *Wappo texts, 1st series (University of California publications in American archaeology and ethnology 19)*. Berkeley: University of California Press.
- Rohner, Ronald (ed.). 1969. *The ethnography of Franz Boas: Letters and diaries of Franz Boas written on the Northwest Coast from 1886 to 1931. Compiled and edited by R. Rohner. With an introduction by Ronald Rohner and Evelyn Rohner. Translated by Hedy Parker*. Chicago: The University of Chicago Press.
- Sapir, Edward. 1911. Review of R. Dixon: “The Chimariko Indians and language”. *American Anthropologist* 13(1): 141–143.
- Sapir, Edward. 1913a. Wiyot and Yurok, Algonkin languages of California. *American Anthropologist* 15(4): 617–646.
- Sapir, Edward. 1913b. Southern Paiute and Nahuatl, a study in Uto-Aztekan, I. *Journal de la Société des Américanistes* 10(2): 379–425.
- Sapir, Edward. 1915a. The Na-dene languages, a preliminary report. *American Anthropologist* 17(3): 534–558.
- Sapir, Edward. 1915b. Southern Paiute and Nahuatl, a study in Uto-Aztekan, II. *American Anthropologist* 17(1): 98–120; 17(2): 306–328.
- Sapir, Edward. 1916. Time perspective in Aboriginal American culture, a study in method. *Memoir 70, Geological survey of Canada, Anthropological series* 13: 1–87.
- Sapir, Edward. 1917a. The position of Yana in the Hokan stock. *University of California publications in American archaeology and ethnology* 13(1): 1–34.
- Sapir, Edward. 1917b. The status of Washo. *American Anthropologist* 19(3): 449–450.
- Sapir, Edward. 1920. The Hokan and Coahuiltecan languages. *International Journal of American Linguistics* 1: 280–290.
- Sapir, Edward. 1921a. A bird’s-eye view of American languages North of Mexico. *Science* 54: 408.
- Sapir, Edward. 1921b. A characteristic Penutian form of stem. *International Journal of American Linguistics* 2: 58–67.
- Sapir, Edward. 1921c. A supplementary note on Salinan and Washo. *International Journal of American Linguistics* 2: 68–72.
- Sapir, Edward. 1922. The fundamental elements of Northern Yana. *University of California publications in American archaeology and ethnology* 13(6): 215–234.
- Sapir, Edward. 1923a. The Algonkin affinity of Yurok and Wiyot kinship terms. *Journal de la Société des américanistes de Paris* 15: 36–74.
- Sapir, Edward. 1923b. Text analyses of three Yana dialects. *University of California publications in American archaeology and ethnology* 20: 263–294.
- Sapir, Edward. 1923c. A type of Athabaskan relative. *International Journal of American Linguistics* 2: 136–142.
- Sapir, Edward. 1925a. The Hokan affinity of Subtiaba in Nicaragua. *American Anthropologist* 27(3): 402–435; 27(4): 491–527.
- Sapir, Edward. 1925b. Pitch accent in Sarcee, an Athabaskan language. *Journal de la Société des américanistes de Paris* 17: 185–205.
- Sapir, Edward. 1925c. The Similarity of Chinese and Indian Languages (Report of an interview). In: *Science* 62(1607): xii–xiii.

- Sapir, Edward. 1926. A Chinookan phonetic law. *International Journal of American Linguistics* 4: 105–110.
- Sapir, Edward. 1929. Central and North American languages. In: *Encyclopaedia Britannica, 14th ed., vol. 5*: 138–141. London / New York: Encyclopaedia Britannica Co.
- Sapir, Edward. 1949. Male and female forms of speech in Yana. In: David G. Mandelbaum (ed.). *Selected writings of Edward Sapir in language, culture and personality*: 206–212. Berkeley: University of California Press.
- Sapir, Edward. 1990–2008. *The collected works of Edward Sapir, 16 vols.* Berlin / New York: Mouton de Gruyter.
- Sapir, Edward, Roland B. Dixon. 1910. *Yana texts, together with Yana myths collected by R. Dixon (University of California publications in American archaeology and ethnology 9)*. Berkeley: University of California Press.
- Sapir, Edward, Leslie Spier. 1943. Notes on the culture of the Yana. *Anthropological Records* 3(3): 239–298.
- Sapir, Edward, Morris Swadesh. 1953. Coos-Takelma-Penutian comparisons. *International Journal of American Linguistics* 19(2): 132–137.
- Сепир, Эдвард. 1993. *Izbrannyje trudy po jazykoznaniju i kul'turologii*. Moskva: Progress.
- Shipley, William F. 1978. Native languages of California. In: William C. Sturtevant (ed.). *Handbook of North American Indians* 8: 80–90. California / Washington, D. C.: Smithsonian Institution.
- Sims, Christine P., Hilaire Valiquette. 1990. More on male and female speech in (Acoma and Laguna) Keresan. *International Journal of American Linguistics* 56(1): 162–166.
- Spott, Robert, Alfred L. Kroeber. 1942. Yurok narratives. *University of California publications in American archaeology and ethnology* 35(9): 143–256.
- Swanton, John R. 1915. Linguistic position of the tribes of Southern Texas and Northeastern Mexico. *American Anthropologist* 17(1): 17–40.
- Токарев, С. А., Е. Э. Бломквист. 1959. Языки и пис'mенност' koren'nogo naselenija Ameriki. In: *Narody Ameriki 1*: 24–52. Moskva: Izdatel'stvo AN SSSR.
- Turner, Katherine. 1987. *Aspects of Salinan grammar*. PhD dissertation. Berkeley: The University of California.
- Voegelin, Charles F., Florence M. Voegelin. 1977. *Classification and index of the world's languages*. New York / Oxford / Amsterdam: Elsevier.

Igor Kuznetsov. Toward a history of the study of long-distance relationship: the six phyla of North America

The present paper is a pioneering attempt to examine in detail the events of 1913–1933 associated with the emergence of key hypotheses of distant relationship of language families in aboriginal America. The context of the state of science at that time is restored, with attention drawn to the difficult interpersonal relationships of E. Sapir, A. L. Kroeber, P. Radin and other participants in the discussions described, as well as the circumstances of their accumulation of linguistic material, often from the last native speakers. An important conclusion is the call to consider the views of the authors of the “six phyla” in their development as a kind of compromise between the approaches of European and American schools. Some of the phyla identified by Sapir (such as Aztec-Tanoan) were apparently never understood by the scholar in the proper spirit of classical comparative-historical linguistics (i.e. as a tree with a single proto-language, etc.).

*Keywords:* Edward Sapir; Alfred L. Kroeber; Paul Radin; phylum; superfamily; genealogical classification of languages; comparative-historical linguistics; mixed languages; Native American languages.