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

№ 17/3

Moscow 2019

Российский государственный гуманитарный университет
Институт языкознания Российской Академии наук

Вопросы языкового родства

Международный научный журнал

№ 17/3

Москва 2019

Журнал основан К. В. Бабаевым

Вопросы языкового родства: Международный научный журнал / Рос. гос. гуманитар. ун-т; Рос. акад. наук. Ин-т языкознания; под ред. В. А. Дыбо. — М., 2019. — № 17/3. — x + 86 с.

Journal of Language Relationship: International Scientific Periodical / Russian State University for the Humanities; Russian Academy of Sciences. Institute of Linguistics; Ed. by V. A. Dybo. — Moscow, 2019. — No. 17/3. — x + 86 p.

ISSN 2219-3820

http://www.jolr.ru/
gstarst@rinet.ru

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

Подписано в печать 10.09.2019. Формат 60×90/8.
Усл. печ. л. 12,0. Уч.-изд. л. 12,6. Заказ № 778.
Тираж 1050 экз.

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

Table of Contents / Содержание

Articles / Статьи

Contributors

Albert Davletshin — candidate of sciences (History), senior researcher, International Centre of Anthropology, Higher School of Economics (Moscow), aldavletshin@mail.ru

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

Ilia Gruntov — candidate of sciences (Philology), senior researcher, Institute of Linguistics, Russian Academy of Sciences (Moscow), altaica@yandex.ru

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

Bartomeu Obrador-Cursach — PhD, lecturer, Institut del Pròxim Orient Antic, Universitat de Barcelona, obradorcursach@gmail.com

Сведения об авторах

Грунтов, Илья Александрович — канд. филол. наук, ст. науч. сотрудник отдела урало-алтайских языков Института языкознания РАН, altaica@yandex.ru

Давлетшин, Альберт Иршатович — канд. истор. наук, ст. научный сотрудник Международного центра антропологии факультета гуманитарных наук ВШЭ (Москва), aldavletshin@mail.ru

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

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

Обрадор-Курсак, Бартомеу — PhD, преподаватель Института древнего Ближнего Востока, Университет Барселоны, obradorcursach@gmail.com

Note for Contributors

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

G. Starostin
Institute for Oriental and Classical Studies
Russian State University for the Humanities
125267 Moscow, Russia
Miusskaya Square, 6
E-mail: gstarst@rinet.ru

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

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

Будущим авторам

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

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

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

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

Albert Davletshin
Russian State University for the Humanities / International Centre of Anthropology of the National Research University Higher School of Economics, Moscow; aldavletshin@mail.ru

Laryngealized vowels and laryngealized consonants in the history of the Totonacan languages of Mexico

The paper argues that Papantla Totonac has lost the original contrast between modal and laryngealized vowels after sonorants and thus developed a system of plain and glottalized stops, affricates and sibilants from the original system of modal and laryngealized vowels. Similarly, Misantla Totonac lost the contrast between modal and laryngealized vowels after sibilants, which resulted in a system of plain and glottalized stops, affricates and sonorants. These sound changes can be explained as being due to different mechanisms of glottalization regarding three classes of consonants — stops, sonorants and sibilants. Apapantilla Totonac and Zapotitlán de Méndez Totonac remain conservative in their treatment of the original Proto-Totonacan laryngealized vowels.

Keywords: Tepehua-Totonacan languages; Totonac language; laryngealized vowels; creaky voice; glottalized consonants; ejective consonants; laryngealized consonants; glottalized sibilants; glottalized sonorants.

The Tepehua-Totonacan linguistic family consists of a dozen of closely related languages spoken by approximately 290,000 people in the states of Veracruz, Puebla, and Hidalgo of Mexico. The exact number of the languages is unknown; some of them are spoken in one town and most of them show considerable dialectal variations (García Rojas 1978, MacKay 1994a, Levy & Beck 2012). These languages are divided in two branches – Tepehua and Totonacan, hence the name I use for the family. Until recently, these languages have been considered as dialects of Totonac and Tepehua, probably because the languages of the family are extremely conservative from the phonological point of view. According to lexicostatistical estimations, Tepehua and Totonac separated from each other around 3000 B.P. The Totonacan branch split off around 2000 B.P. when Misantla Totonac separated from the other Totonac languages (Figure 1).

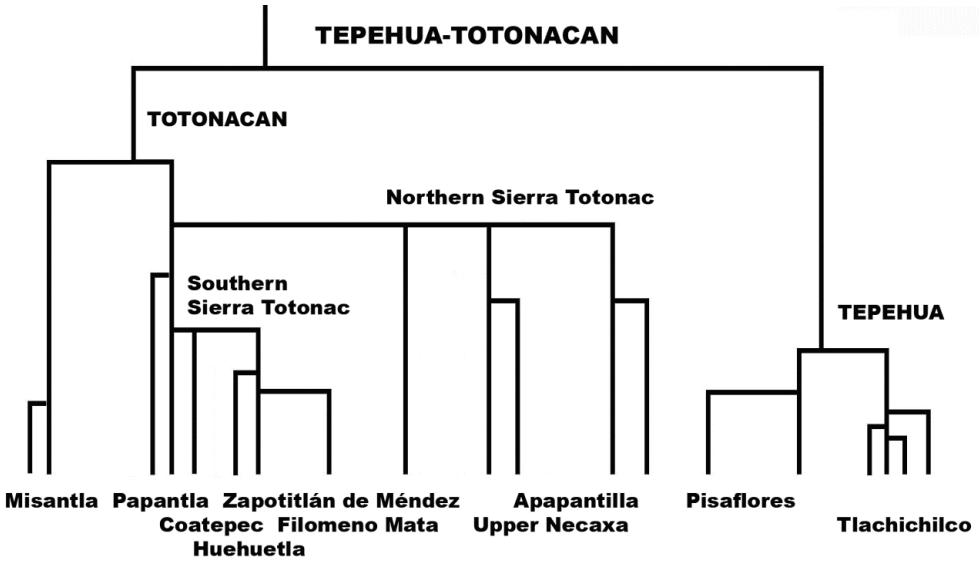


Figure 1. Classification of the Tepehua-Totonacan languages

Speakers of Tepehua-Totonacan languages quite often do not have names for their languages and use the names *totonaco* and *tepehua*. These are derived from Aztec ethnonyms *tepeːwa*ʔ 'highlanders' (literally 'owners of mountains') and *totoːnaka*ʔ 'people of the hot land' (literally 'ones of the place where it heats'). Specific designations for different languages come from the names of the principal towns where they are spoken, which is a common practice in Mesoamerican linguistics.

The basic phonological opposition in most of the Totonacan languages is the contrast between modal and laryngealized (creaky voice) vowels, while the Tepehua languages systematically contrast plain and glottalized stops and affricates. Importantly, glottalized consonants of the Tepehua languages correspond to laryngealized vowels of the Totonacan languages (Arana 1963: 124). Since there are relatively few languages in the world in which creaky voice has a phonemic status, the development of laryngealized vowels is interesting from a typological point of view. The family is also of particular interest for the laryngeal theory of Proto-Indo-European, because in at least one language of the family, Olintla Totonac, historically laryngealized vowels have developed new vocalic qualities (for data on Olintla Totonac, see Tino 2020). During the last decade, the reconstruction of Proto-Tepehua-Totonacan has become a subject for heated debates: no fewer than three different Proto-Tepehua-Totonacan inventories have been proposed – one consisting of modal and laryngealized vowels (Arana 1963, Brown, Beck *et al*. 2011, Watters 2018), one of glottalized stops and affricates (MacKay & Trechsel 2018a,b), and one more of glottalized stops, affricates and sonorants (Davletshin 2019).

I have an impression both from publications and informal talks that many fellow totonacanists tend to consider laryngealization as a highly unstable feature in the history of the Totonacan languages and easily assume irregular sound changes involving laryngealized vowels (e.g. MacKay & Trechsel 2018a,b). The author consents to the former statement but disagrees with the last one. In a recent paper, I have argued that the correspondences involving word-final laryngealized vowels in nominals in the Totonacan languages are numerous and complex but highly regular and phonologically motivated (Davletshin 2018). The aim of the present paper is to show that phonological systems of two Totonacan languages, which have been traditionally described as contrasting modal and laryngealized vowels, can be analysed as systems that contrast plain and glottalized consonants.

I will start with works in which the term "laryngealized vowel" was originally introduced by Hermann Aschmann and Kenneth Pike (in their phonological description of Zapotitlán de Méndez Totonac) and recapitulate their arguments. Then, I will show that this kind of analysis is not relevant in the case of Papantla Totonac, which should be rather described as a language with plain and glottalized stops, affricates and sibilants. I will proceed with Misantla Totonac and argue that the language can be alternatively described as one contrasting plain and glottalized stops, affricates and sonorants. Finally, I will try to show that both languages are innovative in respect to glottalized consonants and that Proto-Totonacan was a language that had both modal and laryngealized vowels.

Laryngealized vowels in Totonacan languages

Across the Totonacan language family, both the description and phonetic realization of the laryngealization feature significantly vary. Rafael Alarcón Montero (2008: 94–96) and Ester Herrera Zendejas (2009: 42–44) describe three forms of non-modal vowels in Papantla Totonac: 1) creaky voice throughout the entire duration of the vowel, characterized by irregular glottal pulses, 2) creaky voice for only a portion of the vowel, typically the beginning, and 3) stiff

voice, indicating a more subtle degree of laryngealization. Specific peculiarities of laryngealization depend on the context. In Papantla Totonac, laryngealized vowels often involve voicing of preceding stops (Herrera Zendejas 2009: 45). In many varieties, laryngealized vowels are followed by strong glottal stops in phrase-final position; in this position non-modal quality of the vowel is easily perceived (for example, in San Francisco Totonac, based on author's fieldwork data of 2007).

In many varieties, some speakers seem to laryngealize more clearly and reliably than others. Older speakers have hoarser voices because of well-known underlying physiological reasons and thus tend to overlaryngealize. On the contrary, children and teenagers tend to underlaryngealize, which might be due to the acquisition of laryngealization at a late age. Both sex and age variation in the realization of laryngealized vowels are interesting subjects for future studies. Many researchers, the author among them, have difficulties in identifying Totonacan laryngealized vowels by ear. I tend to hear clearly the laryngealized vowels following stops and affricates, possibly, because (phonetically) laryngealized vowels are restricted to these positions in the language of my primary fieldwork, which is Pisaflores Tepehua. Interestingly enough, even linguistically trained native speakers of Totonacan languages often have difficulties in identifying laryngealized vowels, for example, in writing.

Glottal stops are prominent in many Tepehua-Totonacan languages; these come from different sources and sometimes have different realizations. At the same time, the phonemic status of the glottal stop in Totonacan languages is not always clear; some researchers include it in the phonological inventory while others do not. Glottal stops are often accompanied by automatic laryngealization of the following vowel. The last but not the least important observation is that glottal stops, glottalized consonants and laryngealized vowels all involve articulatory movement of the glottis constriction. From this point of view, the widely established terms are misleading. The terms "glottalized consonants" and "glottalized vowels" might be more representative of the articulatory similarities between the corresponding sounds. One can say that the Tepehua and Totonac branches of the Totonacan family tree are differentiated in the temporal location of laryngealization in the syllable: Tepehua languages feature laryngealization in syllable onsets, while Totonac languages restrict it to the syllable nucleus (Arana 1963: 124).

However, the main subject of the present paper is not the phonetics of laryngealized vowels but their phonology — in other words, our main concern here are distributions and restrictions.

Zapotitlán de Méndez Totonac

Zapotitlán de Méndez Totonac was the first extensively documented language of the family; it boasts a republished dictionary and a number of papers on various grammatical topics (Herman Aschmann 1946, 1962, 1983, etc., Elizabeth Aschmann 1984, Aschmann & Wonderly 1952). Herman Aschmann calls the language Highland Totonac (in Spanish, *totonaco, dialecto de la Sierra*).

It is no wonder that laryngealized vowels were first noticed and described during the study of this language, although glottal stops in word-final position had already been found by Landis Christiansen in a variety of Sierra Totonac (Christiansen 1937: 152).[1]

The phonological system of Zapotitlán de Méndez Totonac can be presented in the following way in the International Phonetic Alphabet (IPA). Correspondingly, a subscripted tilde

[1] The exact location of Ladis Christansen's fieldwork is unknown. He indicates that there are three dialects of Totonaco –Totonaco of the Coast, Totonaco of Papantla and Totonaco of Sierra – and specifies that his study is limited to the Totonaco of Sierra (Christansen 1937: 52).

under the vowel (V̰) is used to indicate laryngealization. In works by Herman Aschmann and co-authors, laryngealized vowels are indicated by means of an apostrophe (V') or a superscript symbol of glottal stop above the vowel (V̓). The system is not described explicitly in the paper on the phonology of the language but can be easily deduced from it (Aschmann 1946).

Consonants						
p	t			k	q	
	t͡s	t͡ʃ	t͡ɬ			
	s	ʃ	ɬ			h
m	n					
			l			
w		y				

Vowels		
i, i: ḭ, ḭ:		u, u: ṵ, ṵ:
	a, a: a̰, a̰:	

We can see that the vocalic system consists of twelve phonemes where three vocalic qualities are combined with two degrees of vowel lengths and laryngealization or its absence. Five vocalic qualities are found on surface but mid vowels [e] and [o] are allophones of high vowels /i/ and /u/ in the context of uvular stops. Herman Aschmann (1946: 35–36) describes three different types of laryngealized vowels: 1) a modal vowel followed by a glottal stop, 2) a vowel followed by glottal stop but also affected by this glottal stop so as to be accompanied by a more or less rough glottal vibration, laryngealization or glottalization, and 3) vowel preceded and followed by a glottal stop with complete laryngealization of the vowel. These types depend on the nature of the preceding consonants. When the laryngealized vowel is preceded by a voiceless consonant, the first type occurs, or a mild form of the second type with only slight laryngealization at the end. When the laryngealized vowel is preceded by a sonorant, it may be completely laryngealized and the preceding sonorant on its off-glide is slightly assimilated to the laryngealized quality of the vowel. When the laryngealized vowel is preceded by another laryngealized vowel, the second one undergoes complete laryngealization as well. When two laryngealized vowels are separated by a sonorant, the second vowel is fully laryngealized and the medial sonorant also undergoes slight laryngealization. When the laryngealized vowel occurs in phrase-initial position, an initial glottal stop occurs, and the vowel is completely laryngealized. This word-initial glottal stop disappears in normal speech when a preceding consonant-final word is added. Importantly, modal vowels are banned in the initial position of lexical roots and prefixes.

Herman Aschmann (1946: 41) and Kenneth Pike (Pike 1947: 148) offer the following analysis. Both scholars deal with the same set of data and their descriptions seem to have resulted from shared conversations. They treat the phonetic glottal stop as part of a vowel nucleus, rather than a separate consonant because of the distribution of glottal stop (or laryngealization). The glottal stop can be treated as a separate consonant on a par with stops, affricates, nasals and so on, but one which differs widely from them in its distributional characteristics. Such an analysis would allow us to reduce the number of vocalic phonemes from twelve to six. However, distributional differences between the glottal stop and the other consonants are so striking that, according to the authors, the system of modal and laryngealized vowels seems to better reflect the structure of the language. The distributional evidence is as follows.

First, the glottal stop is used with much greater frequency than any single consonant. According to Herman Aschmann (1946: 41), in a sample text containing 1834 syllables, 732 contained a glottal stop.

Second, the glottal stop, as a consonant, would nearly double the number of consonant clusters, since it can occur before any word-medial or word-final single consonant or cluster. Furthermore, the distribution of /p/, /t/, /q/, /y/, /w/, /m/ and /n/ would seem odd, in that, word-finally, they would appear only after a vowel or as part of a cluster with a preceding [ʔ]; stating that in word-final position they occur after either laryngealized or non-laryngealized nuclei would be less controversial.

Third, laryngealized vowels are found in Spanish loans where they are substituted for non-laryngealized ones: *a̰ːnɨ̰maːɬ* 'beast' (Spanish *animal*), *wɨ̰ːlah* 'personal name' (Spanish, *Manuela*), *kwɨ̰ːrsah* 'by necessity' (Spanish, *fuerza*)[2].

Fourth, when there are alternate forms of a morpheme, and the particular usage is determined by whether or not the preceding morpheme ends in a consonant or a vowel, forms with final phonetic glottal stops are treated the same way as morphemes ending in modal vowels. For example, the suffix 'second person object' *-nɨ̰* loses its vowel after non-laryngealized (a) and laryngealized nuclei (b) but not after a regular consonant (c): *ʃkan* 'it bit you' (a), *ɬfɨ̰n* 'he tied you' (b) and *niknɨ̰* 'he hit you' (c). Aschmann (1946: 41) translates *ɬfɨ̰n* as 'he tied it', but this translation is mistaken (see also Pike 1947: 148). Similarly, morpheme-initial /y/ is retained after non-laryngealized and laryngealized vowels but not after regular consonants: *ʃkayaːw* 'we bite it', *ɬfɨ̰yaːw* 'we tied it', *nikaːw* 'we hit it'.

Fifth, in certain syntactic positions in the phrase syntactic juncture lengthens word-final vowels (a), but not vowels followed by a consonant (c); laryngealized vowels are lengthened, ignoring the presence of the phonetic stop (c).

snapapa ʃuːwa̰	*snapapaː ʃuːwa̰* (a)
'the white fur'	'the fur is white'
qama̰ liːwa̰t	*qama̰ː liːwa̰t* (c)
'tasty food'	'the food is tasty'
ɬkitit skuhnɨ̰	*ɬkitit skuhnɨ̰* (c)
'a lazy worker'	'the worker is lazy'

Sixth, in certain situations, a morphological process involves the substitution of a laryngealized vowel for a non-laryngealized one; treating the glottal stop as a regular consonant implies that the glottal stop here is technically an infix, which is not very convenient in view of other characteristics of the language as a whole; compare *liːsmaniyaːw* 'we (incl.) accustom ourselves to it' with *liːsma̰nɨ̰ya̰ːtit* 'you (pl.) accustom yourselves to it'. Importantly, the second person subject form involves laryngealization of each of the vowels which are preceded by a voiced continuant.

Seventh, laryngealized vowels in medial position are somewhat unstable. There are no cases of deletion of medial consonants which would parallel the assumed deletion of the phonetic glottal stop, represented on the surface by the phonemic substitution of non-laryngealized for laryngealized nuclei: [wanɨ̰ɬ] or [waniɬ] 'he said it' and [la̰wa̰] or [lawa̰] 'you did it'.

Herman Aschmann also notices significant dialectal differences in the realization of laryngealized vowels and considers this fact as an argument in favour of his analysis. According to him (Aschmann 1946: 42), in certain words where a phonetic glottal stop follows the vowel in Zapotitlán de Méndez that same word is pronounced in Coyutla with a glottal stop preceding

[2] Stress is not marked in works on Zapotitlán de Méndez Totonac, nor are stress assignment rules stated explicitly. There is a big deal of variation in stress assignment in Totonacan languages (see MacKay 2011), and I make no attempts to "reconstruct" stress in Zapotitlán de Méndez Totonac.

the vowel, and in Concepción it is realized as laryngealization of a vowel without any glottal stop present, for example, *ʃu̯ːn* 'bitter': [ʃuːʔn], in Zapotitlán de Méndez, [ʃʔuːn], Coyutla, and [ʃu̯ːn], Concepción).

We can see that for both Aschmann and Pike the decisive argument in favour of a system with modal and laryngealized vowels is the distribution of the phonetic glottal stop in relation to its adjacent vowels and consonants. We can also see that the offered analysis is inextricably linked to the Totonac variety under study: the distribution of laryngealization is analysed in relation to following consonants rather than preceding ones. We may assume that if the main subject of the study were Coyutla Totonac, the distribution of laryngealization in relation to the preceding consonants might have been likely considered as well, since this is where the phonetic glottal stop is heard in Coyutla Totonac.

It is worth mentioning that the seven distributional criteria presented by Aschmann and Pike do not contradict the idea of a system of phonologically significant glottalized consonants preceding phonetically laryngealized vowels. Such an idea requires us to assume (1) some morphophonetic alternations between glottalized and plain consonants and (2) a rule according to which glottalized consonants become deglottalized in syllable-final position. Both phenomena are quite common in languages with plain and glottalized consonants (Fallon 2002) and also found in Pisaflores Tepehua, for example (Davletshin 2016). It is important to emphasize here that to suggest an alternative description is not necessarily about being right or wrong. It is quite normal for the same language to be adaptable for different models of description, and importantly, with each of the suggested models having its advantages and disadvantages depending on different objectives and perspectives of research.

Interestingly enough, the distribution of laryngealized vowels in Spanish loans given by Aschmann depends on the preceding consonant — laryngealized vowels are found after sonorants and word-initially, while modal vowels are found after voiceless consonants: *a̯ːni̯ma̯ːɬ* 'beast' (Spanish *animal*), *wi̯ːlah* 'personal name' (Spanish, *Manuela*), *kwi̯ːrsah* 'by necessity' (Spanish, *fuerza*), *kapih* 'coffee' (Spanish, *kafé*) *ʃaːpatuh* 'Saturday' (Spanish, *sabado*), *pisiːntih* 'personal name' (Spanish, *Vicente*)[3]. The only exception to this rule is the second syllable of the name *wi̯ːlah*, where the vowel after the sonorant is followed by a glottal fricative. Modal vowels are banned word-initially in Zapotitlán de Méndez Totonac. I do not have a good explanation why either Spanish or Nahuatl sonorants should be borrowed as glottalized consonants into Totonacan languages. One theoretical explanation is that, since glottalized sonorants are rare in the language, they might be perceived as special markers of loanwords. However, I am not aware of other typological examples when loanwords are marked by phonemes absent in the donor language. A few Spanish loans show non-modal vowels following sonorants in the published dictionary: *kilaːntu* 'coriander (Spanish, *cilantro*)', *liːmuːnif* 'lime (fruit and tree, Spanish, *lemón*)', *ʃtiːlaːn* 'chicken (Spanish, *castilla*)', etc. They might represent a later stratum of borrowings or, alternatively, borrowings via Nahuatl, which played the role of *lingua franca* in the Early Colonial period.

Importantly, in Zapotitlán de Méndez Totonac laryngealized vowels are not attested following the glottal fricative /h/; when this phoneme is followed by a vowel in the same word, it is realized as a voiceless velar fricative [x] (Aschmann 1946: 36–37). This gap in the distribution of laryngealized vowels has been never commented upon in published literature. The only counterexample to the stated observation found in a published dictionary (Aschmann 1983) is the word *tihi̯* 'road' where the laryngealized vowel is found word-finally.

[3] According to Aschmann (1946: 41): "In adapting Spanish loans to the phonemic system of Totonaco, laryngealized vowels are sometimes (although nonpredictably) substituted for the non-laryngealized types of Spanish".

At least one second person subject form shows laryngealization after a glottal fricative: *ʧu̯hą̱* 'you (sg.) spit', cf. *ʧuh-* 'to spit' (Aschmann & Wonderly 1952: 139). Here laryngealization marks the second person subject. Laryngealized vowels in word- and phrase-final position are realized in a particular way, with a strong glottal stop, in many Totonacan varieties (see Coatepec Totonac in McQuown 1990 and Filomeno Mata Totonac in McFarland 2009). It is probable that word-final laryngealized vowels historically arise from a nominalizer **-ʔ* (Davletshin 2018). In other words, word-final laryngealized vowels can be phonologized as underlying modal vowels followed by a glottal stop. The ban on the glottal stop /ʔ/ after a glottal fricative /h/ is easy to explain because both consonants share the same place, but different manners of articulation. The ban on laryngealized vowels following glottal fricatives can be seen from the following table, which shows statistical correlations between laryngealized and non-laryngealized vowels in initial syllables of the lexical entries found in the published dictionary (Aschmann 1983).

	p_	t_	k_	q_	_
_V	189	201	81	87	!*
_V̰	46	38	61	39	182

	ʦ_	ʧ_	ʨ_
_V	33	28	13
_V̰	42	48	11

	s_	ʃ_	ɬ_	h_
_V	34	30	8	20
_V̰	10	14	3	!*

	m_	n_	l_	w_	y_
_V	175	12	222	11	3
_V̰	28	6	12	19	5

These statistics are somewhat misleading because some entries in the dictionary contain identical roots and, more importantly, the language employs a number of highly productive derivational prefixes. Nevertheless, the table clearly shows that there is a ban on laryngealized vowels following glottal fricatives and another one on word-initial modal vowels.

It is important to mention that laryngealized vowels carry low lexical load in Zapotitlán de Méndez Totonac, apart from examples where laryngealization distinguishes second person subject verbal forms from third person subject verbal forms. Here are all the minimal pairs found in published works (Aschmann 1983: 141, 143).

ɬkq̱aq̱ 'ashes' — *ɬkaka* 'spicy'
ʃku̯ta 'sour' — *ʃkuta* 's/he unties X'
ʃlakq̱an 'theirs' — *ʃlakan* 'his face'
ʃqq̱aɬnį̱ 'his blood' — *ʃqaɬnį̱* 'its opening'
stapu̯ 'biting midge (Spanish, *jején*)' — *stapu* 'beans'

Laryngealization marking second person subject forms results in many minimal pairs, since imperfective forms for second person singular and third person singular coincide in their surface realization if the verb stem ends in a non-sonorant. Here is an example from the paper where this phenomenon was first discussed (Nida 1949: 63, see also Aschmann & Wonderly 1952: 139):

paq̱ʃą̱ 'you (sing.) bathe' — *paʃa* 's/he bathes'

It is particularly instructive to compare Zapotitlán de Méndez Totonac with Pisaflores Tepehua. In Pisaflores Tepehua phonologically glottalized consonants may be superficially realized in different ways (author's fieldwork from 2011). Syllable-initially, glottalized bilabial and dental stops /p'/ and /t'/ are realized as implosives followed by a strongly laryngealized

vowel: [ɓV̰] and [ɗV̰]. Glottalized velar stops /k'/ are realized either as implosive or ejectives or plain stops followed by a laryngealized vowel: [ɠV̰ ~ k'V̰ ~ kV̰]. Glottalized affricates /ʦ/ and /ʧ'/ are realized either as ejectives or as plain consonants followed by a laryngealized vowel: [ʦ'V̰ ~ ʦV̰] and [ʧ'V̰ ~ ʧV̰]. In syllable-final position, underlying glottalized consonants are deglottalized. Glottal stops are realized with a strong laryngealization of the following vowel as well [ʔV̰]. These glottal stops can be of different origins: some of them come from historical plain and glottalized uvular stops *q and *q', others from glottalized glides *w' and *y', and still others from Proto-Tepehua-Totonacan glottal stops *ʔ. We can see that the main phonetic realization of glottalized consonants in Pisaflores Tepehua is laryngealization of the following vowels. However, the base system is not one of modal and laryngealized vowels, but rather one of plain and glottalized consonants, since laryngealization is not found after sonorants and sibilants, as opposed to Zapotitlán de Méndez Totonac.

Apapantilla Totonac

Apapantilla Totonac belongs to the Northern Totonacan subgroup (Reid & Bishop 1974, Reid 1991). The authors call the language Xicotepec de Juárez, but the town where the data were collected is Apapantilla (David Beck p.c., 2016). The published dictionary implies a phonological inventory similar to Zapotitlán de Méndez Totonac. The main difference is a five vowel system with mid vowels /e/ and /o/ developed from a combination of high and low vowels, also from high vowels in the context of uvular stops /q/ and velar fricatives /x/. The glottal fricative of Zapotitlán de Méndez Totonac /h/ corresponds to velar fricative /x/ in Apapantilla Totonac, which is (probably) also realized as uvular [χ].

Consonants:					
p	t			k	q
	ʦ	ʧ	tɬ		
	s	ʃ	ɬ	x	
m	n		l		
w		y			

Vowels:		
i, i: i̱, i̱:		u, u: u̱, u̱:
e, e: e̱, e̱:		o, o: o̱, o̱:
	a, a: a̱, a̱:	

Below is the statistical distribution of laryngealized and non-laryngealized vowels in initial syllables of the lexical entries found in the Apapantilla dictionary (Reid & Bishop 1974). Vowel-initial Spanish loans are written etymologically, without laryngealization of the first vowel indicated, and are excluded from statistics. As is explicitly stated in the grammatical sketch (Reid 1991: 2), modal vowels are banned word-initially.

	p	t	k	q	_
_V	262	514	129	126	!*
_V̰	52	116	85	73	492

	ʦ	ʧ	tɬ
_V	27	71	18
_V̰	48	86	6

	s	ʃ	ɬ	x
_V	43	98	18	33
_V̰	12	20	5	!*

	m	n	l	w	y
_V	473	15	479	22	6
_V̰	57	3	49	36	3

As in Zapotitlán de Méndez Totonac, Apapantilla Totonac bans modal vowels in word-initial position and laryngealized ones after the velar fricative /x/. Laryngealized vowels after sonorants are more frequent than after sibilants, but less frequent than after stops and affricates: the same pattern is observed in Zapotitlán de Méndez Totonac.

The word for 'road' (*ˈtex*) has lost the final vowel in Apapantilla Totonac as well as in the other Northern Totonacan languages; note that the derived form *kaːteˈxen* 'on the road' does not show a laryngealized vowel after the velar fricative /x/. I was able to locate fourteen items with a laryngealized vowel after the velar fricative /x/ in the dictionary; one of them is an ideophonic word, the others are second person subject forms. I give five examples: *naˈskuxa* 'you (sg.) work', *nakimaːtaːˈxiːya* 'you (sg.) are going to charge me', *taˈxuːpa ~ toˈxoːpa* 'you (sg.) stay in water', and *qeːloˈxoqo* 'loose (of clothes)' (Reid & Bishop 1974: 94, 306, 402–403). It is possible to phonologize second person subject forms in such a way that eliminates laryngealized vowels after velar fricatives in lexical representations. Alternatively, we can assume that the ban on laryngealized vowels after velar fricatives operated on a proto-level but is not active anymore.

I do not make attempts to incorporate into this study the data of another extensively documented language with laryngealized vowels, Upper Necaxa Totonac (Beck 2004 and 2011; Puderbaugh 2019). This Northern language is an innovative variety of Totonac which has developed glottalized sibilants from initial consonant clusters of the type "sibilant plus uvular stop" (Beck 2006).

Papantla Totonac

Papantla Totonac is another well documented language of the family, the only one which can boast two dictionaries (Aschmann 1973, García Ramos 2007) and two monographic studies dedicated to its phonology, one of them by a native speaker of the language (García Ramos 1980, Levy 1983). Papantla Totonac has the highest number of speakers in the family, although its dialectal variants have never been studied systematically. In his published dictionary, Herman Aschmann (1973) makes use of the apostrophe to indicate laryngealized vowels, but places it directly after consonants, that is to say, before vowels. Such an odd way of indicating laryngealized vowels is probably meant to show their peculiar phonetic realization in the language. Aschmann follows the same strategy in an unpublished dictionary of Coyutla Totonac (Aschmann 2000). Paulette Levy (1987, 1990) follows Aschmann's analysis and postulates a system of modal and laryngealized vowels. The following phonological system is explicitly stated in the work by Paulette Levy (1987: 9) and implicitly assumed in Aschmann's dictionary (1973).

Consonants:						
p	t			k	q	
	ʦ	ʧ	ʦˡ			
	s	ʃ	ɬ			h
m	n					
			l			
			(r)			
w		y				

Vowels:		
i, iː i̯, i̯ː		u, uː u̯, u̯ː
	a, aː a̯, a̯ː	

The /h/ is realized as velar fricative [x] syllable-initially and as glottal fricative [h] syllable-finally (Levy 1983: 44). The vibrant /r/ is found in loanwords and native ideophonic words, hence the use of parentheses (Davletshin 2014).

We can see that Aschmann proposed the same phonological system for Papantla Totonac and Zapotitlán de Méndez Totonac. However, distributional characteristics of laryngealized vowels are different in Papantla Totonac: phonetic laryngealized vowels are not attested following sonorants /m/, /n/, /l/, /w/ and /y/. Laryngealized vowels are also banned after glottal fricative /h/. Below is the statistical count of laryngealized and non-laryngealized vowels in initial syllables of the lexical entries from the dictionary of 1973. Laryngealized vowels are more frequent after stops and affricates are more frequent than after sibilants, in accordance with the pattern observed in Zapotitlán de Méndez Totonac and Apapantilla Totonac.

	p_	t_	k_	q_	_
_V	316	457	121	171	318
_V̰	77	93	96	58	!*

	ʦ_	ʧ_	tɬ_
_V	46	61	24
_V̰	77	86	18

	s_	ʃ_	ɬ_	h_
_V	58	66	16	38
_V̰	11	9	1	!*

	m_	n_	l_	w_	y_
_V	453	31	469	40	4
_V̰	!*	!*	!*	!*	!*

Again, these statistics are somewhat misleading because some entries in the dictionary contain the same root and, more importantly, the language has a number of highly productive derivational prefixes. All initial vowels are marked as non-laryngealized in the dictionary, but this is just a convention. Paulette Levy (1987: 61–62) states that in word-initial and morpheme-initial (in combinations with certain prefixes) positions vowels automatically acquire either an additional glottal stop, or a hiatus, or glottal stop accompanied by laryngealization, or simply laryngealization. The only lexical entry which starts with a lateral fricative and is followed by a vowel is *ɬḭt'hu* 's/he packs X with'. This is due to the rarity of *ɬV* syllables in Totonacan languages; some more examples of lateral fricatives before laryngealized vowels are found in the dictionary, e.g., *ˈpaɬa̰* 'hard, tough', *qa̰ɬa̰ːˈna* 'thief', *muːˈkɬṵn* 'strong unpleasant odour', *tuːˈɬṵkṵn* 'a kind of tree (Spanish, *hoja santa*)'.

The main reason for postulating a phonological opposition between modal and laryngealized vowels in Zapotitlán de Méndez Totonac is an even distribution of the laryngealized vowels in relation to their surrounding consonants. Thus, we can reanalyse the phonological system of Papantla Totonac as a system of modal vowels, plain and glottalized consonants: laryngealized vowels are eliminated, but glottalized stops, affricates and sibilants are introduced. It does not seem to be a coincidence that glottal stops are heard preceding vowels and directly following consonants in this language as opposed to Zapotitlán de Méndez Totonac where they are rather heard following vowels, since from a phonological point of view phonetic glottal stops belong to consonants in Papantla Totonac, but to vowels in Zapotitlán de Méndez Totonac.

Consonants:						
p p'	t, t'			k, k'	q, q'	
	ʦ, ʦ'	ʧ, ʧ'	tɬ, tɬ'			
	s, s'	ʃ, ʃ'	ɬ, ɬ'			h
m	n					
			l			
			(r)			
w		y				

Vowels:		
i, iː		u, uː
	a, aː	

Ejective fricatives are relatively rare sounds in the world's languages, but 18 of them are reported in the paper by Ryan Shosted and Sharon Rose (2011: 41–42). Importantly, at least one language, Lakota, has plain and glottalized stops, affricates and sibilants in its inventory, but the contrast is neutralized in sonorants, similarly to Papantla Totonac (Ingham 2003: 4).

Interestingly enough, a native speaker of the language Cresencio García Ramos (1980: 23, 2007: 50) postulates a system of plain and glottalized stops and affricates, in addition to glottal stop. I first met Cresencio García Ramos in 2007; his glottalized stops and affricates were very clearly audible, sounding nothing like plain consonants followed by laryngealized vowels. I suppose that this particular pronunciation is due to his dialect. The lack of glottalized sibilants in his works might be a dialectal trait too, or, alternatively, it might represent an unrecognized contrast. I wish I had taken Cresencio García Ramos's native-speaker intuition more seriously when I first met him. On the other hand, the recordings published by Paulette Levy (2012) do not show any phonetic glottalized consonants, but rather clear laryngealized vowels. In his phonetic study of Papantla Totonac, Rafael Alarcón Montero (2008: 96-98) mentions ejective stops as an alternative realization of laryngealized vowels and attributes it to dialectal variation.

The rules for the developments in Papantla Totonac can be presented as follows:

*RV̰ → RV; *T(S)V̰ → T(S)'V, *SV̰ → S'V

(R stands here for any sonorant, S for any sibilant and T for any stop).

Laryngealized vowels carry a low lexical load in Papantla Totonac (Aschmann 1973: vii). None of the minimal pairs involve laryngealized vowels following sonorants.

ḭʃla'ka̰n 'theirs'	*ḭʃlakan* 'his face'
'ɬka̰ka̰ 'ashes'	*'ɬkaka* 'spicy'
ʃa'pa̰ʃni 'dirty one'	*ʃa'paʃni* 'the washed one'
'ʃkṵta 'sour'	*'ʃkuta* 's/he unties X'
'sta̰pṵ 'biting midge'	*'stapu* 'beans'
'ʦḭʦḭ 'furuncle'	*'ʦḭʦi* 'warm, tepid (of water)'

The comparison of three languages — Papantla Totonac, Apapantilla Totonac and Zapotitlán de Méndez Totonac — suggests that Papantla Totonac has lost the opposition between modal and laryngealized vowels after sonorants. This loss resulted in the restructuring of the phonological system and in the development of glottalized stops, affricates and sibilants. The sound shift under discussion is easy to explain because glottalized stops, sibilants and sonorants are articulated differently — the sonorants, in particular, are produced with closer constriction of the vocal folds which interrupts or modifies normal voicing, but with neither a raising nor a lowering movement of the larynx: acoustically, they can be described as laryngealized or creaky voice sonorants (Ladefoged & Maddieson 1996: 78; Maddieson 2013). The lack of minimal pairs for modal and laryngealized vowels following sonorants and the ban on modal vowels after sonorants in word-final position in Proto-Totonacan might facilitate the loss of the contrast after sonorants.

	Papantla	Zapotitlán	Apapantilla	Misantla
's/he loosens X'	*ʃla'ha*	*ʃla̰hay*	*ʃla̰'xa*	—
'turkey hen'	*ʧa:wi'la*	*ʧuwila̰*	*ʧa:wi'la̰*	*ʧa:wa'la̰ʔ*
'wild tamarind'	*'lili:q*	*lḭlḭ:qa*	*la̰'la̰:k*	*la̰:'la̰:k*
'butterfly'	*ʃpḭpḭ'li:q*	*ʃpḭpḭlḭ:qa*	*ʃpḭpḭ'le:q*	*'ʃpḭpḭ*

	Papantla	Zapotitlán	Apapantilla	Misantla
‘heron’	*ˈlu̱:qu̱*	*lu̱:qu̱*	*ˈlo̱:qo*	—
‘clothes’	*ˈluʃu*	*lu̱ʃu̱*	*ˈlu̱ʃu̱*	—
‘long’	*ˈɬma:n*	*ɬma̱:n*	*ˈɬma̱:n*	*ˈɬma̱:n*
‘be lying (supine)’	*-ˈma:-*	*-ma̱:-*	*-ˈma̱:-*	*-ˈma̱:-*
‘bamboo (Spanish, *tarro*)’	*ˈmatɬu̱:k*	*ma̱tɬu̱:k*	*ˈma̱tɬu̱:k*	—
‘jug’	*ˈtɬa̱mank*	*tɬa̱ma̱nk*	*tɬa̱ˈma̱nk*	*ˈta̱mi̱:nk*
‘s/he pours on, waters X’	*muˈnu*	*mu̱nuy*	*mu̱ˈnu*	—
‘nine’	*-naˈhatsa*	*-na̱ha:tsa*	*-na̱ˈxa:s*	*-naˈha:tsa*
‘s/he winds or twists X’	*ˈsnata*	*sna̱ta*	*ˈsna̱ta*	—
‘vein, nerve’	*ˈʃnuhut*	*ʃnu̱hut*	*ʃno̱ˈxot*	—
‘s/he eats’	*wa:ˈyan*	*wa̱:yan*	*wa̱:ˈyan*	*wa̱:ˈyan*
‘mealing stone (Spanish, *metate*)’	*ˈʃwa:ti̱*	*ʃwa̱:ti̱*	*ˈʃwa̱:ti̱*	*ˈʃwa̱:t*
‘egg’	*ˈqa̱ɬwa:t*	*qa̱:ɬwa̱:t*	*qa̱ˈɬwa̱:t*	*qa̱:qˈɬuwa̱:t*
‘you (sg.)’	*ˈwiʃ*	*wi̱ʃ*	*ˈwi̱ʃ*	*ˈwi̱ʃ*
‘guava’	*a:ˈsiwi:t*	*a̱:siwi̱:t*	*a̱:siˈwi̱:t*	*a̱:ˈsiwi̱:t*
‘immature ear of maize (Spanish, *jilote*)’	*ˈʃiwi:t*	*ʃiwi̱:t*	*ʃiˈwi̱:t*	*ˈʃiwi̱:t*
‘cool’	*sqa̱ˈwiwi*	*sqa̱wi̱wi̱*	*sqe̱ˈwi̱wi̱*	—
‘hawk’	*wa:ˈya*	*wa̱:ya̱*	*wa̱:ˈya̱*	—
‘earth’	*ˈti̱yat*	*ti̱ya̱t*	*ˈti̱ya̱t*	*ˈti̱ya̱t*
‘squirrel’	*ˈstaya*	*stayi̱*	*ˈstayi̱*	—
‘mouse’	*ˈtsi̱:ya*	*tsi̱:ya̱*	*ˈtsi̱:ya̱*	*ˈtsi̱w*

Word-initial palatal glides are extremely rare in Tepehua-Totonacan languages and are probably of secondary origin. I was unable to find other examples of Proto-Totonacan **ya̱* and **yu̱* with cognates in Papantla Totonac.

Laryngealized vowels after sonorants in Apapantilla Totonac, Zapotitlán de Méndez Totonac and Misantla Totonac are lexically determined. Papantla Totonac forms a relatively shallow subgroup together with the languages of Southern Sierra, Zapotitlán de Méndez Totonac and Olintla are among them (see Figure 1). Thus, the loss of laryngealization after sonorants in Papantla Totonac is an innovation. In Proto-Totonacan final vowels in nominals after both sonorants and fricatives were obligatorily laryngealized (Davletshin 2018: 161). This situation is preserved in Apapantilla Totonac and Zapotitlán de Méndez Totonac; its reflexes are also found as final glottal stops in Coatepec Totonac, Huehuetla Totonac and Olintla Totonac, and as glottalization of final vowel-preceding consonants in Filomeno Mata (for the data see McQuown 1990, Troani 2004, McFarland 2009). In Papantla Totonac, final vowels after sibilants in nominals are laryngealized but sonorants are followed by modal vowels word-finally. Laryngealized nominal-final vowels after sonorants and fricatives are probably due to a single sound change. The ban on final laryngealized vowels after sonorants and the ban on final modal vowels after fricatives in Papantla Totonac imply that the rule of laryngealization of final vowels chronologically preceded the loss of laryngealization after sonorants; thus, they also imply that the system of glottalized stops, affricates and sibilants attested in Papantla Totonac is an innovation.

According to Aschmann (1946: 42), laryngealized vowels in Coyutla Totonac are realized with a glottal stop preceding the vowel, similarly to Papantla Totonac. If our suggestion is cor-

rect, the distribution of laryngealized vowels in Coyutla Totonac should be uneven, depending on preceding consonants. An unpublished dictionary of Coyutla Totonac was posted on-line (Aschmann 2000). Predictably, this dictionary of some 6000 lexical entries does not include any entries with laryngealized vowels following sonorants in initial syllables.

Misantla Totonac

Misantla Totonac is a member of the Totonacan subgroup but is separated from all the other Totonacan languages genetically, both from grammatical and lexical points of view. It is also isolated from the other languages of the family geographically. Today, this language is spoken in two towns of Yecuatla and San Marcos Atesquilapan, Veracruz, and considerable dialectal differences between the two towns have been attested. The language is highly endangered, with fewer than 133 speakers, most of whom are elderly (according to census of 2010). The language is intensively studied in works by Carolyn MacKay and Frank Trechsel, although its lexicon remains underdocumented (MacKay 1994b, 1999; MacKay & Trechsel 2005; Castro Guevara 2011). It should be mentioned that Carolyn MacKay's grammar of 1999 was a very important contribution to Totonacan studies. In a phonological sketch of Misantla Totonac, MacKay (1994b: 370) gives the following system.

Consonants:						
p	t			k	q	ʔ
	ʦ	ʧ				
	s	ʃ	ɬ			h
m	n					
			l			
w		y				

Vowels:		
i, i:		u, u:
i̱, i̱:		u̱, u̱:
	a, a:	
	a̱, a̱:	

Neither MacKay nor most other Totonacanists incorporate Spanish loans and native ideophonic words into the phonological system (MacKay 1994b: 386). Glottal stop has a peculiar distribution in the language: it only occurs word-finally after short vowels, and all syllable-initial glottal stops can be interpreted as epenthetic (MacKay 1994b: 382). Remarkably, the vowel after a glottal stop is always laryngealized, and the vowel before a glottal stop is always stressed. If /ʔ/ is not word-final (for example, when followed by a suffix), then it is deleted (MacKay 1994b: 400).

MacKay (1999: 384–385, also Trechsel & Faber 1992) emphasizes that the contrasts between long and short vowels and between plain and laryngealized vowels are established by numerous minimal pairs in Misantla. However, neither of their minimal pairs shows any contrast between modal and laryngealized vowels after sibilants or sonorants.

ˈsta̱:niɬ 's/he sold X for Y'
kiliːˈti̱ya̱t 'my land'
ˈʧi̱:n 'strong, very'
ʧu̱ˈʧu̱ 's/he sucks X'

ˈsta:niɬ 's/he looked at X'
kiliːˈtiyat 'my girlfriend'
ˈʧi:n 'pus'
ʧuˈʧu 's/he roasts X'

Laryngealized vowels of initial syllables show a rather intriguing distribution in a published word-list of some 700 lexical items (MacKay & Trechsel 2005).

	p_	t_	k_	q_	_
_V	53	78	27	23	1
_V̲	6	14	20	18	45

	ʦ_	ʧ_
_V	11	25
_V̲	21	11

	s_	ʃ_	ɬ_	h_
_V	19	19	7	10
_V̲	2	1	0	1

	m_	n_	l_	w_	y_
_V	104	9	65	5	3
_V̲	2	2	5	8	1

The only example of a syllable-initial modal vowel must be a typo (see a loan from Spanish *espeho* ‘mirror’) because the rule of automatic glottal stop insertion before a syllable-initial vowel and its subsequent laryngealization is explicitly stated in the phonological sketch. Laryngealized vowels are frequent after stops and affricates, but few words show laryngealized vowels after sonorants and sibilants in the word-initial syllable. However, laryngealization of vowels after sonorants is an important marker of the second person subject on verbs (MacKay 1999: 156–160) and the rule of the laryngealization of final vowels after sonorants in nominals is easily recognizable in the published wordlist.

Two words show a laryngealized vowel after a syllable-initial /h/: *ˈhi̲n* ‘smoke’ and *taˈhi̲ːn* ‘thunder’. This is unusual because the other languages of the family do not allow laryngealization in this position.

	Misantla	Zapotitlán	Apapantilla	Papantla
‘smoke’	*ˈhi̲n*	*hini̲*	*ˈhini̲*	*ˈhin*
‘thunder’	*taˈhi̲ːn*	—	*tą̲ːˈhiːn*	*tą̲ˈhiːn*

MacKay (1994b: 384) stresses that, although laryngealized vowels most frequently occur after obstruents, they also occur following sonorants. However, she does not mention that very few items show a laryngealized vowel following a sibilant: *miˈsi̲n* ‘jaguar’, *ˈsą̲k* ‘s/he brings X together’, *ˈsi̲ːn* ‘rain’ and *ˈʃu̲ːn* ‘bitter’. Five out of six words with a laryngealized vowel following a fricative show a dental nasal. This suggests that laryngealization in these words can be reinterpreted in a different way, for example, as surface realization of the underlying final glottalized nasal. We should take into consideration the fact that some of the items might be typos, for example, ‘s/he chooses’ is given in the word-list as *lakˈsak* but contains the same root as *ˈsą̲k* ‘s/he brings X together’. Alternatively, some of these words might have been elicited from an overlaryngealizing speaker, for example, an elderly person. Carlo Antonio Castro Guevara’s dictionary (2011) does not help much, because the author worked with the consultants from Misantla, who were likely to speak a particular dialect, different from Yecuatla and San Marcos Atesquilapan. He only heard laryngealized vowels after stops and affricates; no laryngealized vowels following sonorants and fricatives are found in his dictionary.

Interestingly enough, only one out of five words with a laryngealized vowel following a fricative, ‘bitter’, consistently shows a laryngealized vowel in the other languages of the family.

	Misantla	Zapotitlán	Apapantilla	Papantla
‘jaguar’	*miˈsi̲n*	*stą̲kuː misin* ‘ocelot’	*miˈsin*	*miˈsin ˈki̲wi* ‘a kind of tree’
‘to bring together’	*są̲k-*	*sak-*	*sak-*	*sak-*
‘downpour’	*ˈsi̲ːn*	*siːn*	*taːsą̲ˈi̲ːn*	*ˈsi̲ːn, ˈse̲ːn*
‘bitter’	*ˈʃu̲ːn*	*ʃu̲ːn*	*ˈʃu̲ːni̲*	*ˈʃu̲ːn*

I do not include here the word 'six', which is given as *-ɬaːˈʃa̰n* in the wordlist but as *-ɬaːˈʃan* elsewhere (MacKay & Trechsel 2005: 315, 237, 238). Its cognates do not show laryngealization either.

The transcription of the Misantla Totonac word *ˈsḭːn* 'downpour' is given as ['seːn], where the mid vowel is not accounted for by the postulated allophonic rules (MacKay & Trechsel 2005: 294). The set 'rain, downpour' deserves some comments. First, Zapotitlán de Méndez Totonac and Coahuitlan Totonac do not show laryngealization in this word (Moore 2017: 100). Second, the mid vowels are surface allophones of high vowels in the vicinity of uvulars in Papantla Totonac and Misantla Totonac, which means that the form ['seːn] should be phonologized as *sa̰ˈḭn* or *saʔin*, cf. Apapantilla Totonac *taːsa̰ˈḭːn*. Two alternative realizations in Filomeno Mata Totonac *saʔin ~ saˈyin* correspond to two alternative forms *ˈsḛːn ~ ˈsḭːn* in Papantla Totonac (McFarland 2009: 14). Third, the sequences /iy/, /yi/ and /uw/, /wu/ are banned in Totonacan languages, and all lexical and grammatical morphemes in proto-Tepehua-Totonacan had consonantal onsets, including those with a glottal stop which shifted to a palatal glide following a vowel (Davletshin 2019). Thus, a likely reconstruction of the word is **siʔan*, cf. also Pisaflores Tepehua *siːˈniː* 'the chief of Thunders', etymologically 'one who makes downpours to come' (author's fieldwork from 2011). The glottal stop **ʔ* disappears in many languages but results in laryngealization of the vowel in Apapantilla Totonac, Papantla Totonac and Misantla Totonac, because vowels after (phonetic) glottal stops are laryngealized in Tepehua-Totonacan languages. The high vowel /i/ and the low vowel /a/ coalesced, resulting in a long and mid vowel [eː] as the regular surface realization.

On the other hand, eleven words which consistently show laryngealized vowels after sibilants in the other languages of the family do not have laryngealization in Misantla Totonac, with the exception of *ˈʃṵːn* 'bitter'.

	Misantla	Zapotitlán	Apapantilla	Papantla
'flower'	*ˈʃana̰t*	*ʃa̰nat*	*ʃa̰ˈnat*	*ˈʃa̰nat*
'grand-'	*-ˈʃuːn*	*-ʃṵnaː-*	*-ʃṵyaː-*	*-ˈsṵːn*
'hole'	*ɬuˈkuk*	*ɬṵkuku*	*ɬṵˈkuku*	*ɬuˈkṵkṵ*
'sweat lodge'	*liːfaːˈqiːn*	*ʃa̰ːqa*	*ˈʃa̰ːq*	*ˈʃa̰ːqa*
'sweet'	*ˈsiksi*	*sa̰qsḭ*	*ˈsḛqsḭ*	*ˈsa̰qsḭ*
's/he blows X'	*suˈnṵʔ*	*sṵnuy*	*sṵˈnu*	*sṵˈnu*
'tongue'	*-siːmaˈqa̰ːt*	*-sḭːma̰qa̰ːt*	*-sḭːma̰ˈqa̰ːt*	*-sḭːmaˈqa̰ːt*
's/he has an itch'	*ˈʃin*	*ʃḭn*	—	*ˈʃḭn*
'nail'	*-maˈqasiːh*	*-sḭya̰n*	*-maqˈsḭːn*	*-sḭˈhan*
'squash'	*ˈnipʃi*	*nḭpʃḭ*	*ˈnipʃḭ*	*ˈnipʃḭ*
'bitter'	*ˈʃṵːn*	*ʃṵːn*	*ˈʃṵːnḭ*	*ˈʃṵːn*

For Papantla Totonac *ˈʃa̰ːqa* 'sweat lodge', see the wordlist published by Paulette Levy (1990: 164).

To sum up, Misantla Totonacan laryngealized vowels after sibilants correspond to modal vowels in other Totonacan languages, and laryngealized vowels after sibilants in other Totonacan languages correspond to modal vowels in Misantla Totonac. This implies that laryngealized vowels after sibilants have become neutralized with modal vowels in Misantla, since laryngealized vowels after sibilants are lexically determined in Totonacan languages. It is likely that Totonacan laryngealized vowels after sibilants come from two different sources. First, in Proto-Totonacan nominals, the final vowels after sibilants and sonorants were laryngealized (according to the rule of laryngealization for final vowels). Second, laryngealized vowels after sibilants in initial syllables most likely developed from combinations of glottal

stop-initial roots with the sibilant prefixes of size and intensity **s-*, **ʃ-* and **ɬ-* (Davletshin 2019). Misantla split first from the other Totonac languages, and so we can assert that laryngealized vowels after sibilants had never developed in that language: Misantla may have eliminated the glottal stop **ʔ* in the clusters **sʔ*, **ʃʔ* and **ɬʔ* in the same way as it happened in Tepehua languages. However, Misantla has lost laryngealization in the vowels after sibilants in word-final position as well, even though word-final vowels after sonorants are always laryngealized.[4] It should be mentioned that in Misantla Totonac word-final vowels were lost after fricatives and sonorants (Davletshin 2018: 166). However, in a few cases, when the vowel was preceded by a sequence of two consonants, it was preserved anyway.

	Misantla	Zapotitlán	Apapantilla	Papantla
'sweet'	*ˈsiksi*	*sa̰qsḭ*	*ˈsḛqsḭ*	*ˈsa̰qsḭ*
'squash'	*ˈnipʃi*	*nipʃḭ*	*ˈnipʃḭ*	*ˈnipʃḭ*

Again, laryngealized final vowels after sonorants and fricatives are probably due to a single sound change. The ban on final laryngealized vowels after sibilants and the ban on the final modal vowels after sonorants in Misantla Totonac imply that the rule of laryngealization for final vowels preceded the loss of laryngealization after sibilants. Importantly, it means that laryngealized vowels after sibilants are to be reconstructed for the Proto-Totonacan level. An additional argument is that two words seem to demonstrate migration of laryngealization from the vowel after a sibilant to the following syllable: *ˈʃana̰t* 'flower', cf. **ˈʃʔanat*, and *suˈnṵʔ* 's/he blows X', cf. **sʔunu-* 'to blow X'. It is also important to mention that laryngealization of the second person subject verbal forms does not involve vowels following sibilants, or other obstruents in Misantla (MacKay 1999: 156–160).

It is easy to figure out what could have been the motivation for the loss of laryngealized vowels after sibilants. Relatively few of the world's languages possess glottalized sibilants, probably because of an articulatory conflict: an ejective requires increased intraoral pressure and a fricative requires air to be continuously vented through a narrow constriction (Shosted & Rose 2011: 41-42).

The rules for the development of Misantla Totonac can be presented as follows.

*SV̰ → SV; *T(S)V̰ → T(S)'V, *RV̰ → R'V,

(R stands here for any sonorant, S for any sibilant and T for any stop.)

Accepting the loss of laryngealization after sibilants in Misantla Totonac, we automatically imply the existence of a system of modal vowels, plain consonants, and glottalized consonants at a certain period in history, when laryngealized vowels are eliminated, but glottalized stops, affricates, and sonorants are introduced in their place.

Consonants:						
p, p'	t, t'			k, k'	q, q'	ʔ
	ʦ, ʦ'	ʧ, ʧ'				
	s	ʃ	ɬ			h
m, m'	n, n'					
			l, l'			
w, w'		y, y'				

Vowels:		
i, i:		u, u:
	a, a:	

[4] The rule is not stated explicitly in the grammar but can be easily deduced from the published wordlist (MacKay & Trechsel 2005).

Interestingly, in the recordings of Misantla Totonac published by Carolyn MacKay (2012) laryngealization on vowels following stops and affricates is often heard close to the preceding consonant — a situation reminiscent of the one described by Herman Aschmann for Coyutla Totonac.

It is possible that after the loss of laryngealization in vowels after sibilants had been completed in Misantla Totonac, some new laryngealized vowels were introduced from other sources for the words *ˈhi̱n* 'smoke', *miˈsi̱n* 'jaguar', *taˈhi̱ːn* 'thunder' and *ˈʃu̱ːn* 'bitter'. I was unable to find any specific conditions which might be responsible for the new laryngealized vowels[5]. The comparisons suggest that we need a further study of Misantla Totonacan laryngealized vowels after fricatives, perhaps involving the use of instrumental phonetic methods. Such a study is made all the more urgent by the fact that Misantla Totonac is a highly endangered language.

Conclusions

Analysis of the distribution of laryngealized and plain vowels relative to preceding consonants in Totonacan languages results in several findings, some of them of considerable typological interest. First, in at least three Totonacan languages, laryngealized vowels are banned following the glottal fricative or its phonetic reflexes. Second, Papantla Totonac and Coyutla Totonac can be described as systems of plain and glottalized stops, affricates and sibilants. This system developed from an earlier contrast between modal and laryngealized vowels. Third, Misantla Totonac can be described as having a system of plain and glottalized stops, affricates and sonorants, at least when viewed in comparative and historical perspective. This system, too, has developed from an earlier system of modal and laryngealized vowels. It is possible to show that both Papantla Totonac and Misantla Totonac are innovative in this respect, and, correspondingly, Apapantilla Totonac and Zapotitlán de Méndez Totonac preserve the original system of modal and laryngealized vowels. Thus, these two languages are of special importance for the reconstruction of proto-Totonacan laryngealized vowels.

It is expected that laryngealized vowels tend to evolve differently depending on the preceding consonants, since glottalized stops, sibilants and sonorants involve different mechanisms of articulation. An unexpected finding is that laryngealized vowels can result in the development of glottalized consonants.

From the acoustic point of view, phonologically glottalized consonants in Papantla Totonac and Misantla Totonac (and in some Tepehua languages) are plain consonants followed by laryngealized vowels. Cross-linguistically, this is quite unusual. It is possible that laryngealized stops, affricates and sibilants might be a better term to describe them, allowing us to distinguish them from phonetically ejective consonants — hence the title given to the present paper.

Acknowledgements

My special thanks go to the speakers of Tepehua and Totonac languages for their knowledge, food and affection, in particular, my mentor Conrado García. I am very grateful to fellow Totonacanists for advice, heated debates and data that are so hard to get in the field: Carlos Gua-

[5] It is tempting to suggest that laryngealization in at least two examples come from the lost final vowel: *ˈhi̱n* 'smoke', cf. **hini*ʔ, and *ˈʃu̱ːn* 'bitter', cf. **ʃʔuːni*ʔ. However, some counterexamples do not allow me to postulate such a rule: *ˈɬiːn* 'pus', cf. **ɬiːna*ʔ. The word *miˈsi̱n* 'jaguar' would not be covered by this development.

dalupe Heiras Rodríguez, David Beck, Dorothy Herzog, Jim Watters, Jorge Tino, Michelle Gracía-Vega, Paulette Levy, Søren Wichmann and Susan Smythe Kung. Last but not least, Mikhail Zhivlov was always ready to discuss Totonacan reconstructions, while the Nostratic seminar in Moscow kindly hosted several of my talks on Tepehua-Totonacan languages.

References

Alarcón Montero, Rafael. 2008. Indicios acústicos de las vocales rechinadas del totonaco. In: Pedro Martín Butragueño & Esther Herrera Zendejas (coords.). *Fonología instrumental: patrones fónicos y variación*, 89–105. México: El Colegio de México.

Arana Osnaya, Evangelina. 1953. Reconstrucción del proto-totonacano. In: Ignacio Bernal & Eusebio Dávalos Hurtado (eds.). *Huastecos, totonacos y sus vecinos / Revista mexicana de estudios antropológicos* 13(2/3): 123–130.

Aschmann, Elizabeth D. 1984. The relative clause in Highland Totonac. *SIL Mexico Workpapers* 6: 1–27.

Aschmann, Herman P. 1946. Totonaco phonemes. *International Journal of American Linguistics* 12(1): 34–43.

Aschmann, Herman P. 1962. *Vocabulario totonaco de la sierra. Serie de vocabularios indígenas "Mariano Silva y Aceves"*, vol. 7. Instituto México: Instituto Lingüístico de Verano.

Aschmann, Herman P. 1973. *Diccionario totonaco de Papantla.* Vocabularios y Diccionarios Indígenas "Mariano Silva y Aceves" vol. 16. México: Instituto Lingüístico de Verano.

Aschmann, Herman P. 1983. *Vocabulario totonaco de la sierra. Serie de vocabularios indígenas "Mariano Silva y Aceves"*, vol. 7. Segunda edición. México: Instituto Lingüístico de Verano.

Aschmann, Herman P. 2000. *Coyutla Dictionary*. Bartholomew Collection of Unpublished Materials SIL International, Mexico Branch. Available at: www.sil.org/resources/language-culture-archives.

Aschmann, Herman P., William L. Wonderly. 1952. Affixes and implicit categories in Totonac verb inflection. *International Journal of American Linguistics* 18(3): 130-145.

Beck, David. 2004. *Upper Necaxa Totonac*. Munich: Lincom.

Beck, David. 2006. The emergence of ejective fricatives in Upper Necaxa Totonac. *University of Alberta Working Papers in Linguistics* 1. Available at: http://www.linguistics.ualberta.ca/UAWPL voll.cfm.

Beck, David. 2011. *Upper Necaxa Totonac dictionary*. Berlin: Mouton de Gruyter.

Brown, Cecil H., David Beck, Grzegorz Kondrak, James K. Watters, Søren Wichmann. 2011. Totozoquean. *International Journal of American Linguistics* 77(3): 323–372.

Castro Guevara, Carlo Antonio. 2011. *Léxico misanteca*. México: IIA-UNAMU.

Christiansen, Landis G. 1937. Totonaco. *Investigaciones lingüísticas* 4: 151–153.

Davletshin, Albert. 2014. *Los sustantivos, adverbios y adjetivos ideofónicos en el tepehua de Pisaflores, Veracruz, México*. Ponencia presentada en el Primer Congreso Internacional de Investigaciones sobre el Mundo Totonaco, Lipuntahuaca Huehuetla, Puebla, 27–29 de agosto, 2014.

Davletshin, Albert. 2016. *Glottalization of stops and affricates in San Pedro Tziltzacuapan Tepehua, Veracruz, Mexico*. Paper presented at the II Workshop on the Sound Systems of Mexico and Central America, Mexico City, March 30 – April 1.

Davletshin, Albert. 2018. Las vocales finales, los procesos fonéticos finales y mediales en el prototototonaco-tepehua: un primer acercamiento. In: Elsa Cristina Buenrostro, Lucero Meléndez Guadarrama, Marcela San Giacomo Trinidad (eds.). *Lingüística histórica de lenguas indomexicanas: hallazgos y discusiones recientes*: 139–186. IIA-UNAM: México.

Davletshin, Albert. 2019. *Some Topics in Tepehua-Totonac Historical Phonology* (*Некоторые вопросы исторической фонетики тепеуа-тотонакских языков*). M.A. Dissertation. Moscow, Russian State University for the Humanities.

Fallon, Paul D. 2002. *The Synchronic and Diachronic Phonology of Ejectives*. New York: Routledge.

García Ramos, Crescencio. 1979. Fonología del totonaco del Tajín, Veracruz. *Cuadernos Antropológicos (Veracruz)* 2: 133–176.

García Ramos, Crescencio. 2007. *Diccionario básico: totonaco-español y español-totonaco*. Xalapa: Academia Veracruzana de las Lenguas indígenas, Secretaría de Educación de Veracruz.

García Rojas, Blanca Rosa. 1978. Dialectología de la zona totonaco–tepehua. Tesis de licenciatura. México. Escuela Nacional de Antropología e Historia.

Herrera Zendejas, Esther. 2009. *Formas sonoras: mapa fónico de las lenguas mexicanas*. México: El Colegio de México.

Ingham, Bruce. 2003. *Lakota*. Languages of the World Materials 426. Munich: Lincom.

Ladefoged, Peter, & Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell Publishers.

Levy, Paulette. 1987. *Fonología del totonaco de Papantla, Veracruz*. México: UNAM.

Levy, Paulette. 1990. *Totonaco de Papantla, Veracruz*. Archivo de Lenguas Indígenas de México. México: El Colegio de México.

Levy, Paulette. 2012. Totonaco de Papantla: El Cerro del Carbón. In: Paulette Levy & David Beck (ed.). *Las lenguas totonacas y tepehuas: Textos y otros materiales para sus estudios*: 345–466. México: UNAM.

Levy, Paulette, David Beck (ed.). 2012. *Las lenguas totonacas y tepehuas: Textos y otros materiales para sus estudios*. México: UNAM.

MacKay, Carolyn J. 1994a. Prospects and proposals for Totonacan research. In: Leonardo Manrique, Yolanda Lastra, Doris Bartholomew (eds.). *Panorama de los estudios de las lenguas indígenas de México*. T. 1. Colección Biblioteca Abya-Yala, 16, 137–168. Quito, Ecuador: Ediciones Abya-Yala.

MacKay, Carolyn J. 1994b. A sketch of Misantla Totonac phonology. *International Journal of American Linguistics* 60(4): 199–248.

MacKay, Carolyn J. 1999. *A Grammar of Misantla Totonac*. Salt Lake City: University of Utah Press.

MacKay, Carolyn J. 2011. Una reconstrucción del acento primario en el proto-totonaco-tepehua. In: Ana Lidia Munguía Duarte (ed.). *Colección de Estudios Lingüísticos I. Fonética, morfología y tipología semántico–sintáctica*: 93–125. Hermosillo: Universidad de Sonora.

MacKay, Carolyn J. 2012. Totonaco de Misantla, Veracruz. In: Paulette Levy, David Beck (eds.). *Las lenguas totonacas y tepehuas: Textos y otros materiales para sus estudios*: 116–180. México: UNAM.

MacKay, Carolyn J., Frank R. Trechsel. 2005. *Totonaco de Misantla, Veracruz*. Archivo de Lenguas Indígenas de México. El Colegio de México: México.

MacKay, Carolyn J., Frank R. Trechsel. 2018a. An alternative reconstruction of Proto-Totonac-Tepehua. *International Journal of American Linguistics* 84 (1): 51–92.

MacKay, Carolyn J., Frank R. Trechsel. 2018b. Una reconstrucción alternativa del prototototonaco-tepehua. In: Elsa Cristina Buenrostro, Lucero Meléndez Guadarrama, Marcela San Giacomo Trinidad (eds.). *Lingüística histórica de lenguas indomexicanas: hallazgos y discusiones recientes*: 85-123. México: IIA-UNAM.

Maddieson, Ian. 2013. Glottalized Consonants. In: Matthew S. Dryer, Martin Haspelmath (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at http://wals.info/chapter/7.

McFarland, Teresa. 2009. The Phonology and Morphology of Filomeno Mata Totonac. Ph.D. Dissertation, University of California, Berkeley.

McQuown, Norman A. 1990. *Gramática de la lengua totonaca (Coatepec, Sierra Norte de Puebla)*. México: UNAM.

Moore, Devin. 2017. Subgrouping of Coahuitlán Totonac. *Canadian Journal of Linguistics / Revue canadienne de linguistique* 62(1): 84–117.

Nida, Eugene A. 1949. *Morphology*. 2nd ed. Ann Arbor: University of Michigan Press.

Pike, Kenneth L. 1947. *Phonemics. A Technique for Reducing Languages to Writing*. Ann Arbor: University of Michigan Press.

Puderbaugh, Rebekka. 2019. *Laryngealization in Upper Necaxa Totonac*. Ph.D. Dissertation, University of Alberta.

Reid, Aileen A. 1991. *Gramática totonaca de Xicotepec de Juárez, Puebla*. México: Instituto Lingüístico de Verano.

Reid, Aileen A., Ruth G. Bishop. 1974. *Diccionarío totonaco de Xicotepec de Juárez, Puebla*. Vocabularios y diccionarios indígenas "Mariano Silva y Aceves", vol. 17. México: Instituto Lingüístico de Verano.

Shosted, Ryan K., Sharon Rose. 2011. Affricating ejective fricatives: The case of Tigrinya. *Journal of the International Phonetic Association* 41(1): 41-65.

Tino, Jorge Antonio. 2020. Las vocales laringizadas y su manifestación fonética en el totonaco de Olintla. *Lingüística Mexicana, Nueva Época* 2(1): 7-30.

Trechsel, Frank R., Alice Faber. 1992. *Acoustic properties of plain and laryngealized vowels in the Misantla dialect of Totonac*. Ms. Universidad de las Américas, Puebla, México.

Troiani, Duna. 2004. *Aperçu grammatical du totonaque de Huehuetla, Puebla, Mexico*. Lincom: Munich.

Watters, James K. 2018. Sobre la relación histórica entre las vocales laringizadas y las eyectivas en totonaco-tepehua. In: Elsa Cristina Buenrostro, Lucero Meléndez Guadarrama, Marcela San Giacomo Trinidad (eds.). *Lingüística histórica de lenguas indomexicanas: hallazgos y discusiones recientes*: 125–138. IIA-UNAM: México.

А. И. Давлетшин. Ларингализованные гласные и ларингализованные согласные в истории тотонакских языков Мексики.

В статье доказывается, что язык тотонако-де-папантла нейтрализовал праязыковое противопоставление между модальными и ларингализованными гласными после сонорных, в результате чего возникла система, основанная на противопоставлении простых и глоттализованных смычных, аффрикат и сибилянтов. Сходными образом в родственном тотонако-де-мисантла было утрачено противопоставление между модальными и ларингализованными гласными после сибилянтов, вместо которого развилась система глоттализованных смычных, аффрикат и сонорных. Оба звуковых перехода понятны с той точки зрения, что три класса согласных (смычные и аффрикаты, сонорные и сибилянты) артикуляторно подразумевают разные механизмы глоттализации. Известно, что глоттализованные сонорные встречаются в языках мира реже, чем глоттализованные смычные и аффрикаты, а глоттализованные сибилянты реже, чем глоттализованные сонорные. В свете полученных данных, тотонако-де-апапантилья и тотонако-де-сапотитлан-де-мендес оказываются консервативными языками подгруппы и, таким образом, имеют особую значимость для реконструкции прототонакских ларингализованных гласных.

Ключевые слова: тепеуа-тотонакские языки; тотонакские языки; ларингализованные гласные; скрипучий голос; глоттализованные согласные; абруптивные согласные; ларингализованные согласные; глоттализованные сибилянты; глоттализованные сонорные.

Ilya Egorov
School for Advanced Studies in the Humanities, RANEPA, Moscow; i.m.jegorow@gmail.com

The origin and synchronic status of mid front vowels in Kazym Khanty

The present paper provides synchronic and diachronic analyses of two mid front vowels (*e* and *ɛ*) in the Kazym dialect of the Khanty language. Investigating the distribution of these vowels shows that they are phonemically contrasted in some positions and neutralized in others. The phonological status of both vowels is additionally confirmed by a perceptual experiment. The source for *e* and *ɛ* are two Proto-Khanty vowels, which are merged or distinguished depending on the consonantal context. Phonemic contrast is extended into some new positions through borrowings from Nenets, Komi-Zyrian, and other dialects of Khanty itself.

Keywords: vocalic systems; historical phonology; dialectology; Khanty language; Uralic languages.

1. Introduction

According to Wolfgang Steinitz's classification, Kazym dialect belongs to the northern group of Khanty dialects, while according to Nikolay Tereshkin's classification, it belongs to the western group. Existence of two *e*-type (mid front) vowels is a characteristic feature of this dialect. Kazym Khanty differs in this feature from the rest of the northern dialects, namely from Obdorsk (Nikolaeva 1999: 5), Sherkal (Steinitz 1950: 36), Shurishkar (Solovar, Nakhracheva & Shiyanova 2016: 22) and could also differ from Nizjam. According to Steinitz's description (1975: 5), only the Synja dialect distinguishes between *e* and *ɛ*, which are allophones of one phoneme. The aim of the present study is to clarify the nature of the relationship between two mid front vowels (free variation / allophonic alternation / phonemic contrast) and to conduct a diachronic analysis of these segments.

The remaining part of the paper is structured as follows. Section 2 gives a brief overview of the vowel system and specifies the issue. In section 3, I characterize the database on which my study is based. Section 4 provides the synchronic analysis of the distribution of *e* and *ɛ* as well as conclusions about their phonological status. In section 5, I give some additional arguments supporting the conclusions outlined in the previous section. Section 6 is devoted to the diachronic analysis. The last section establishes the main conclusions summarizing the results of the synchronic and diachronic analysis.

2. Vowel system of Kazym Khanty: a general overview

In Kazym Khanty, nine vocalic segments can be distinguished in the initial syllable. We describe them using three distinctive features: backness (front, central, back), height (close, mid, open) and tenseness (tense, lax). In non-initial syllables only five segments are possible.

Phonological status of all vowels, except for the pair *e*/*ɛ*, is certain, cf. the following minimal pairs: *taλ* 'empty' — *tăλ* 'winter' — *tɛλ* 'full', *χɯr* 'washtub' — *χur* 'photo' — *χər* 'glade,

vista' — *χor* 'bull, ox' — *χir* 'bag, sack', *peś* 'haunch' — *piś* 'cunning; opportunity', *vet* 'five' — *văt* 'so, well (particle)'. The vowel [ə] I consider to be an allophone of /ă/ in non-initial syllables.

Table 1. The vowel inventory of Kazym Khanty (without specification of the phonological status)

	Initial syllable							Non-initial syllable			
	front		central		back			front		central	
	tense	lax	lax	tense	lax	tense		tense	lax	tense	lax
close	i				ʉ	u		i			
mid	e	ɛ			ɵ	o		e	ɛ		
open			ă	a						ə	a

The formant space of the vowels can be visualized using the superposed bagplots diagram for F1 and F2 values of nine vocalic segments which are possible in the initial syllable. Each polygon is subdivided into internal and external areas. The internal area contains 50% of the values closest to the arithmetic meaning. This approach allows to measure the value of formants throughout the whole duration of the vowel, including accommodation areas on the border with consonants. Formant values on the border of the vowels are situated in the peripheral part of the polygon or considered outliers.

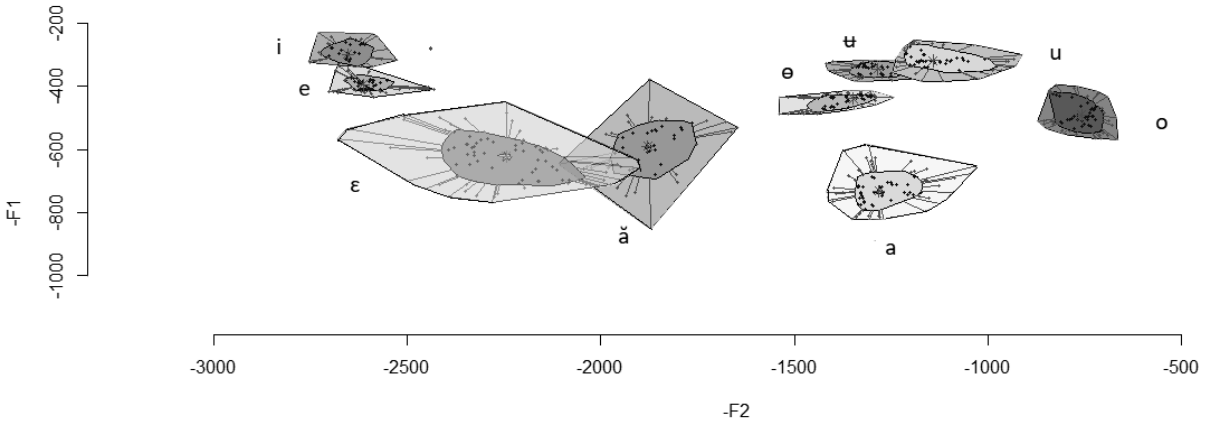


Figure 1. Formant space of the vowels in the initial syllable

Figure 1 shows that *e* and *ɛ* are completely different in their acoustic features. However, their status remains problematic, since no minimal pairs have been recorded for them. Strictly speaking, the existence of a minimal pair is not an absolute requirement for the confirmation of phonemic contrast. Thus, for example, it is widely known that there are no minimal pairs for /c/ and /č/ in standard Russian in the area of inherited and non-onomatopoeic lexicon. Providing a minimal pair is only a method for proving the existence of phonemic contrast, but not a requirement. To recognize two segments as different phonemes, two conditions must be fulfilled: 1) both segments must be possible in a similar phonological context; 2) segments must not show free variation, i.e. they must be lexically distributed.

As of now, there is a relatively small body of literature dealing with the Kazym Khanty vowel system. For one thing, it was the topic of Galina Kurkina's monograph (Kurkina 2000) which focuses on the acoustic features of the sounds. Based on an extremely small sample (83 items), Kurkina claims that *e* and *ɛ* are in complementary distribution depending on the following consonant (2000: 21–22), hence they should be considered allophones of one pho-

neme. Andrey Kaksin (2010: 28) considers *ɛ* as a variant of the phoneme /e/, however, he does not specify whether it is an allophone or a free variant. Neither are these segments characterized clearly enough in "Sketch of Khanty dialects" (Solovar et al. 2016: 16). Initially, both segments are defined as “phonemes”, although *e* is represented in slashes whereas *ɛ* is given in square brackets. Later on, the authors list the positions for the sound [e] (hence raising the question if it is really the main realization of the phoneme /e/), which appears 1) “before palatal consonants *ń*, *ś*, *λ´*, *j*”; 2) “after *k* and *χ*, palatals and bilabial *w*”. The second statement contradicts Kurkina’s analysis. Positional distribution of [ɛ] was not considered in Solovar et al. 2016. From this brief overview, it becomes clear that previous studies failed to determine the phonological status of *e* and *ɛ* reliably and to confirm the conclusions with an extensive corpus. However, in general, the scholars named above tend to describe these segments as variants of one phoneme.

3. Materials

The database on which the present study is based consists of two parts: synchronic and diachronic. The synchronic part was composed by Anton Kukhto on the basis of the wordlist from Valentina Solovar’s dictionary (Solovar 2014), which was corrected by me during a field session in Kazym in July 2019. The most important corrigenda are mentioned in Appendix 1. This part of the database consists of 720 lexemes. For each item, the left and right consonantal context, the vowel in the preceding and following syllable, the number of the syllable from the initial of the word have been determined. Additional data from Kaksin 2010 was used in order to investigate the distribution in non-initial syllables. The diachronic part includes data from Steinitz’s dictionary (DEWOS). A number of clarifications and corrigenda, based on Solovar 2014 and my field notes, are mentioned in Appendix 2. Kazym Khanty lexemes are cited with comparanda from other Khanty dialects, namely from Vakh, Vasyugan, Verkhne-Kalymsk, Vartovsky, Likrisovsky, Yugan, Maly Yugan, Tremyugan, Upper Demyanka, Konda, Nizyam, Sherkal, Synja and Obdorsk dialects. Each item has been analyzed in order to determine the origin of the vowel according to Mikhail Zhivlov’s (2007) reconstruction of the Proto-Khanty vowel system and its right consonant context.

4. Synchronic analysis of the distribution of e and ɛ

In word-initial syllables, perfect complementary distribution has been recorded in some positions, while several other positions demonstrate only a few exceptions deviating from the strict distribution, but in other positions phonemic contrast is certain. Free variation does not occur in the speech of the same native speaker, nor have we observed any interspeaker variation.

4.1. Perfect complementary distribution has been observed in the word onset: the lax vowel *ɛ* appears before *p*, *n*, *t*, *λ*, *k* and *ŋ*[1], whereas the tense vowel *e* is possible only before *w*. Word-internally ɛ appears before *m*, *p* and *l*, and *e* appears before palatals (*ś*, *j*, *ń*, *t´*) and *χ*. Positions where complementary distribution with no exceptions can be postulated are summarized in Table 2:

[1] No examples for other consonants attested in my corpus.

Table 2. Complementary distribution of *e* and *ɛ*

ɛ	e
_m, _p, _l #_p, #_n, #_t, #_*λ*, #_k, #_ŋ	_ś, _j, + single example for _ń, _t′, _*χ* #_w

4.2. There is no complementary distribution in other positions, therefore I can postulate phonemic contrast between tense /e/ and lax /ɛ/. The positions of phonemic contrast can be subdivided into two groups: 1) where the contrast is supported by the range of examples; 2) where only single «exceptions» contribute to the existence of the contrast. This dichotomy allows to make the following assumption: positions belonging to the first group reflect an old phonemic contrast, whereas in the second group it is an innovation. The analysis starts with the second group. Table 3 presents the main trends in distribution and all found exceptions:

Table 3. Distribution of ɛ and e: trends and exceptions

Context	Most cases	Exceptions
_w	e	*λɛwasa* ‘carelessly, negligently’
_r	ɛ	*keriti* ‘to fall’
_ŋ	ɛ	*leŋki* ‘poor fellow (Rus. бедняжка)’, *leŋkər* ‘harvest mouse’
_n	ɛ	*penśar* ‘tambourine’
_k	e	*šɛk* ‘log, stump’, *šɛk* ‘ankle’

Omitting numerous examples of *e* before *w*, I have to mention here the only example where the lax vowel appears in exactly the same position, i.e. between *λ* and *w*. It is the hydronym *λew* ‘Sosva river’, which is not widely used by speakers currently living in Kazym, situated far from this river. Nevertheless, this word has reliable Khanty cognates. The etymology of *λɛwasa*, on the other hand, is unclear. This word remains the only example of *ɛ* before *w*.

In most cases, the lax vowel appears before *r*. The only exception is *keriti*, cf. the examples with identical consonant context: *kɛr* ‘snow crust (Rus. наст)’, *kɛras* ‘high steep coast of a river’, *kɛrətti* ‘to go around’. The cited examples imply the presence of phonemic contrast before *r* or at least between *k* and *r*. Contradicting the assumption above, the contrast here goes back to Proto-Khanty; a detailed diachronic exploration of the situation will be presented in Section 6.1. However, the following cases must be interpreted as supporting the idea of an innovative nature of the contrast in this group.

In the position before *ŋ* the tense vowel appears only in two words: *leŋki* and *leŋkər*. The second word is not known to modern speakers, but the first one is commonly used also as a part of compounds with pejorative-diminutive meaning: *iki-leŋki* ‘lit. man-poor fellow (Rus. мужичонка)’, *puχ-leŋki* ‘lit. boy-poor fellow (Rus. мальчишка)’. Both *leŋki* and *leŋkər* have reliable Khanty comparanda (DEWOS: 844, 782). The second word is likely to be an interdialectal borrowing. In the Kazym dialect, the standard correspondence for Likr. *θ* and Irt., Ni. *t* is the lateral fricative *λ*. Initial *l-* points to a loan from Synja or Obdorsk dialect. However, Steinitz cites the form *λɛŋkər*, which demonstrates the predictably expected word-initial *λ* along with the lax vowel *ɛ* before *ŋ*.

Another example of a deviation from the general distribution in loans is *e* before *n* in *penśar* (< Nenets *pə́ēnt’śer’’*) ‘tambourine’. It is not excluded that the second consonant of the cluster actually affects the articulation of *e* in this particular case (before *ś* only *e* is possible). Neverthe-

less, in other cases the second consonant does not influence vowel quality. Steinitz cites the same word with *ń* (DEWOS: 1185); primary *ń* here could explain the tense vowel. Still, in modern Khanty this lexeme has been recorded with dental *n*, therefore I am forced to postulate phonemic contrast in this position as well.

One more example of loans breaking the distribution is a pair of homonymous lexemes: *šɛk* from Komi-Zyrian *žek* 'log, stump' and *šɛk* from Komi-Zyrian *šeg* 'ankle'. Inherited words demonstrate *e* before *k*: *wek* 'unlikely', *tekənti* 'to become full', *tekuptəti* 'to fill'. The same situation can be observed in Russian loans: *lekśitti* < Rus. 'to treat', *lekkər* < Rus. 'doctor, medical assistant', *wekša* < Rus. 'squirrel'.

The tense vowel *e* appears in word-final position in all the words, except for the lexeme *nɛ* 'woman, female', which also has the oblique stem *nɛŋ-*.

4.3. A certain phonemic contrast has been observed in word-internal positions before *t*, *ʎ*, *š*. The following quasi-minimal pairs confirm this:

pɛʎa 'to, in direction of (postposition)' vs *peʎi* 'affirmative particle'
kɛša 'for (postposition)' vs *keši* 'knife'

An obvious explanation involving assimilation by the second syllable vowel must be rejected in light of the following counterexamples: *ʎɛʎi* 'hungry', *pɛši* 'fawn, young deer', *keʎaŋ pɛləka* 'to the left', *keʎ-a* 'rope-DAT.SG'.

For the position before *t*, no minimal pair or even quasi-minimal pair has been found. All the examples from my database are cited below.

Position after *m*, *ʎ* and *χ* before *t*: *mɛt* 'the most, very', *mɛta* 'to enough', *mɛtšəti* 'to push, to shove (in); to put on', *mɛtšaśti* 'to push oneself away'; *ʎɛt-ut* 'food'; *ʎɛti* 'to eat; to burn (intr.)'; *noχər-ʎɛtnɛ* 'spotted nutcracker (*nucifraga caryocatactes*)', *χɛtškəti* 'to cry (of drake)'. The last lexeme is also given in Solovar 2014 in the form *χeškəti* 'to cry (of drake)'. This may be an onomatopoeic word.

Position after *w*, *j*, *k* before *t*: *wet* 'five', *wetmit* 'fifth', *wetχuśjaŋ* 'fifteen', *wetlow χăjəp* 'marsh sandpiper (*tringa stagnatilis*)', *wetsot* 'five hundred'; *jetšəti* 'to ripe, to be ready, to come to end', *jetšɛməti* 'to finish, to end'; *ketəmti* 'to touch'. I can postulate phonological contrast before *t* taking into account the fact that the left context does not determine vowel quality in this dialect except for the position after *j*, note on which can be found immediately below.

4.4. The only case in which left context synchronically affects a vowel is the position after *j* in the initial syllable. After *j* before *t*, *š*, *ʎ* phonemic contrast is neutralized, and only the tense vowel is possible here: *jetšəti*, *jetšɛməti*, *ješa* 'a few, a little', *ješawəʎ* 'soon', *ješək* 'dear', *jeʎ* 'far, into the distance', *jeʎʎi* 'forward', *jeʎpija* 'before (postposition)', *jeʎəmti* 'to put to shame', *jeʎməʎti* 'to be ashamed'. Only the tense vowel can be found between *j* and *s*, otherwise the lax vowel appears before *s*: *pɛsi* 'mourning', *pɛsəʎ* 'sedge', *rɛskəti* 'to hit strongly; to slam strongly', *sɛsi* 'trap'.

4.5. In initial syllables, phonological contrast of /e/ vs /ɛ/ is only found in limited positions, namely word-internally before *t*, *š*, *ʎ* (if no *j* precedes the vowel) and before *r*, *w*, *n*, *ŋ*, *k*. In the second group of positions it is supported by a few examples only (*keriti*, *ʎɛwasa*, *penśar*, *leŋki*, *šɛk*, *šɛk*). In all other contexts the contrast is neutralized. The archiphoneme /E/ is realized as a tense vowel in onset before *w* and before palatals (*t´*, *ś*, *ń*, *j*) and *χ* word-internally, as well as between *j* and *s*, *t*, *š*, *ʎ*; as a lax vowel before *p*, *n*, *t*, *ʎ*, *k*, *ŋ* in onset and before *p*, *m*, *l*, *s* (but [e] / j_s) word-internally.

In non-initial syllables, [e] and [ɛ] are in complementary distribution depending on the right consonantal context (Table 4). Hence, the phonemic contrast is neutralized in this position.

Table 4. Distribution of [e] and [ɛ] in the non-initial syllables

_p	_m	_w	_t	_s	_n	_l	_r	_λ	_k	_ŋ
ɛ	ɛ	e	e	ɛ	e	ɛ	ɛ	e	e	ɛ

5. Evidence from a perceptual experiment

For further confirmation of the phonological status of /e/ and /ɛ/, a perceptual experiment was conducted. Four informants were asked to evaluate six pairs of words in random order. In each pair, one word (given with a gloss in the table below) was correctly pronounced by Kazym Khanty native speaker, whereas in the incorrect counterpart (given without a gloss) the first vowel was substituted using a computer sound editor. The informant had to recognize the word, i.e. repeat it and translate it into Russian or to claim that the word is incorrect or absent in the Khanty language.

In Table 5, which shows the results of the perceptual experiment, I use the following symbols: ATL, ZAM, IMI, TRG — informants' initials; + — a word is correctly recognized; * — the informant claims there is no such word in Khanty; lexeme with a gloss — informant's interpretation.

Table 5. Results of perceptual experiment

	ATL	ZAM	IMI	TRG
[keši] 'knife'	+	+	+	+
[kɛši]	*	*kăši* 'to feel pain'	*kăši* 'to feel pain'	*
[kew] 'stone'	+	+	+	+
[kɛw]	similar to *kɛm* 'time'	*	*i kɛm* 'similar'	*
[kɛm] 'approximately'	+	*	+	+
[kem]	*	*	*	*
[kɛša] 'for (postposition)'	+	+	+	+
[keša]	*	*	*	*
[peś] 'haunch'	+	+	+	+
[pɛś]	*	*	*	*
[pɛši] 'олененок'	+	+	+	+
[peśi]	*piλa* 'together', *piλ* 'pair'	*piλ* 'pair'	*	*

It is important that when "rejecting" the word speakers considered it not as an incorrect pronunciation but as a totally non-existent form. This fact decisively confirms the phonological nature of /e/ and /ɛ/.

6. Diachronic analysis

On the basis of the following set of correspondence, Zhivlov (2007: 282) reconstructs two Proto-Khanty sources for the modern Kazym mid front vowels (*e* vs *ɛ* are not distinguished in Zhivlov 2007):

Table 6. Reflexes of the Proto-Khanty **ä* and **ē* in Khanty dialects

ProtoKh	V.	Vj.	Trj.	J	DN	Ko.	Ni.	Kaz.	O.
*ä	e/ö[1]	e/ö[1]	ă̯/ŏ̯[2]	ă̯/ŏ̯[2]	e	e	e	e	e/o[3]
*ē	i	i	i	i	e	e	je-, -e-	je-, -e-	i ~ e

Notes from the cited work: «1. *e* (*ö* adjacent to velars, if there is no *ä* in the second syllable). 2. *ă̯* (*ŏ̯*/*ŏ̯* after *k* not before velars, if there is no *ä* or *i* in the second syllable). 3. *e* (*o* ~ *e* adjacent to *ŋ*, *k*)».

6.1. My observations show that some positions of the synchronic phonemic contrast reflect Proto-Khanty opposition **ä* vs **ē*. In the position after *k* before *š* and *r*, *e* goes back to **ä* (1, 4) whereas *ɛ* goes back to **ē* (2–3, 5–6).

(1) Kaz. *keri-* 'to fall'; V., Vj. *kə̯rəγ*; VK *körəγ*; Vart., Likr., Mj., Trj. *kə̆rəγ*; J. *kə̆rəγ*; Irt. (DN, KoP., Kr.) *kerə*; Ni. , Š. *kerij-*; O. *kori-* (DEWOS: 676);
(2) Kaz. *kɛr* 'snow crust (Rus. наст)'; V., Vj., Likr., Mj., Trj., J., KoP., Kr. *kir*; Ni., Sy., O. *ker* (DEWOS: 661);
(3) Kaz. *kɛrət-* 'to go around'; V., Vj., Trj., J. *kirəγtə-*; Irt. (DN, KoP., Kr.), Ni., Sy., O. *kerət-* (DEWOS: 669);
(4) Kaz. *keši* 'knife'; V., Vj. *kə̯čəγ*; VK *köčəγ*; Vart., Likr., Mj., Trj. *kə̆čəγ*; J. *kə̆čəγ*; Irt. (DN, KoP., Kr.) *kečə*; Š. *kešə*; Sy. *keši*; O. *kesi* (DEWOS: 593);
(5) Kaz. *kɛš* 'fur stocking (Rus. чижи)'; V., Vj., Vart., Likr., Mj., Trj. *kiṇč*; J. *kinč*; KoP., Kr. *kenč*; Ni. *keša*; Sy. *keš*; O. *kis* (DEWOS: 646);
(6) Kaz. *kɛša* 'for (postposition)'; Trj. *kičə́*; DN, KoP., Kr. *kečə́*; Ni., Š., Sy. *keša*; O. *kossi, kosi* (DEWOS: 592).

If the left context is different, **ä* yields *ɛ* before *š* (7); before *r*, **ä* and **ē* are merged (8–12).

(7) Kaz. *pɛši* 'fawn, young deer (under one-year-old)'; V. *pečəγ*; KoP., Kr., Ni., Š. *pečə*; O. *pēsi* (DEWOS: 1095);
(8a) Kaz. *ńɛr* 'ire'; J. *ńir*; KoP., Ni., Š. *ńer* (DEWOS: 1069);
(8b) Kaz. *ńɛrəŋ* 'nervous, chippy'; V., Vj., Trj., J. *ńirəŋ*; KoP., O. *ńerəŋ* (DEWOS: 1069);
(9) Kaz. *nɛr-* 'to rub' V., Vj. Vart., Likr., Mj., Trj., J. *nir-*; DN, Kr., Ni., Š., O. *ner-* (DEWOS: 1012);
(10) Kaz. *sɛr, sɛri* 'deaf (of elk)'; V., Vj. *serɜ*; Vart., Likr., Mj., Trj. *sə̆rɜ*; J., KoP. *sə̆rə*; Ni. *serə*; Š., O. *ser* (DEWOS: 1368);
(11) Kaz. *wɛr* 'work, matter'; V., Vj. *wer*; Trj., J. *wə̆r*; Irt. (DN, KoP., Kr.), Ni., Š., Sy., O. *wer* (DEWOS: 1613);
(12) Kaz. *λɛr* 'root'; V. *ler*; Vj., VK *jer*; Likr. *θə̆r*; Mj., Trj., J. *λə̆r*; Irt. (DN, KoP., Kr.), Ni., Š., *ter*; Sy., O. *ler* (DEWOS: 797).

Separate reflexes of **ä* and **ē* have also been attested in the position before *t* and *λ*: ProtoKh. **ä* > Kaz. *e* (14–15; 18–21); ProtoKh. **ē* > Kaz. *ɛ* (13; 16–17).

(13) Kaz. *λɛt-ut* 'food'; V. *lit-o̯t*; Vj. *int-o̯t*; Trj., J. *λit-ŏt*; Irt. (DN, KoP., Kr.) *tet-ăt*; Š. *tet-ŏt*; Sy. *let-ŏt*; O. *lit-it* (DEWOS: 714);

(14) Kaz. *ket-əm-* ‘to touch’; cf. V., Vj. *kə̯t*; VK *kə̯t*; Vart., Likr., Mj., Trj. *kə̆t*; J. *kə̆t*; Irt. (DN, KoP., Kr.) *ket* (DEWOS: 698);
(15) Kaz. *wet* ‘five’; V., Vj. *wet*; Trj., J. *wă̆t*; Irt. (DN, KoP., Kr.), Ni., Š., Sy., O. *wet* (DEWOS: 1641);
(16) Kaz. *pɛλa* ‘to (postposition)’; V., Vj. *pil*; J. *piλ*; KoP. *pet-*; Ni., Š. *peta*; Sy. *pela*; O. *pelȧ* (DEWOS: 1145);
(17) Kaz. *λɛλt-* ‘to load (up), to ship; to seat’; V. *lilt-*; Vj. *ilt-*; Trj., J. *λiλt-*; DN, KoP., Kr., Ni., Š. *tettə-*; Sy., O. *lelt-* (DEWOS: 748);
(18) Kaz. *keλ* ‘rope’; V., Vj. *kə̯l*; Vart. *kə̆λ*; Likr. *kə̆θ*; Mj., Trj. *kə̆λ*; J. *kə̆λ*; Irt. (DN, KoP., Kr.), Ni., Š. *ket*; Sy., O. *kel* (DEWOS:);
(19) Kaz. *ńeləŋ* ‘greedy’; V., Vj. *ńeləŋ*; Trj., J. *ńă̆ləŋ*; Irt. (DN, KoP., Kr.), Ni., Š. *ńetəŋ*; Sy., O. *ńeləŋ* (DEWOS: 1042);
(20a) Kaz. *weλ-* ‘to hunt, to kill’; V., Vj. *wel-*; Likr. *wă̆θ-*; Mj., Trj., J. *wă̆λ-*; Irt. (DN, KoP., Kr.), Ni., Š. *wet-*; Sy., O. *wel-* (DEWOS: 1580);
(20b) Kaz. *weλpəs* ‘hunting, hunt, fowling’; Trj., J. *wă̆λpəs*; DN, KoP., Kr., Ni., Š. *wetpəs*; Sy., O. *welpəs* (DEWOS: 1581).
(21) Kaz. *weλəm* ‘marrow’; V., Vj. *weləm*; Vart. *wă̆λəm*; Likr. *wă̆θəm*; Mj., Trj., J. *wă̆λəm*; Irt. (DN, KoP., Kr.), Ni., Š. *wetəm*; O. *weləm* (DEWOS: 1584);

For an exact description of the development of Proto-Khanty **ä* and **ē*, two additional rules are needed. The vowel **ä* yields *ɛ* (22–23) after labial non-approximants (*m*, *p*). Labials *m* and *p* are put in the same class opposed to the labial approximant *w* when it comes to the realization of the synchronic archiphoneme /E/. After *t* and before *λ*, **ä* is reflected as *ɛ* (24–25). Upon first sight, such a position seems quite unnatural; nevertheless, one can find a typological parallel even within Finno-Ugric languages. Thus, Proto-Permic **ȯ* yields Proto-Udmurt **a* before *l* after dentals, whereas normally it yields **o* (Zhivlov 2010: 171, Lytkin: 1964: 128–129).

(22) Kaz. *mɛt-* ‘to get tired’; V. *met-*; Trj., J. *mă̆t-*; Ni., Sy., O. *met-* (DEWOS: 971);
(23) Kaz. *pɛλ-* ‘to prick, to prick oneself, to stick oneself into’; V., Vj. *pel-*; Trj., J. *pă̆λ-*; DN, KoP., Kr., Ni., Š. *pet-*; O. *pel-* (DEWOS: 1138);
(24) Kaz. *tɛλ* ‘clothes; vessel’; Vj. *tel*; KoP., Ni., Š. *tet*; O. *tel* (DEWOS: 1427);
(25) Kaz. *tɛλ* ‘full’; Vj., Vk *tel*; Likr. *tă̆θ*; Mj., Trj. *tă̆λ*; DN, KoP., Kr., Ni., Š. *tet*; Sy., O. *tel* (DEWOS: 1425);

The only example which cannot be explained by the phonological laws is (26). I assume here some kind of ablaut, which has undergone analogical levelling in Western Khanty but has been preserved in Eastern Khanty, cf. *λɛλt-* in example (17).

(26) Kaz. *λɛλ-* ‘to get on (sledges or some other transport)’; V. *lel-*; Vj. *jel-*; Trj., J. *λă̆λ-*; Irt. (DN, KoP., Kr.), Ni., Š. *tet-*; Sy., O. *lel-* (DEWOS: 747).

6.2 Another position of the contrast of Proto-Khanty **ä* and **ē*, which was already noted in (Zhivlov 2007), is a word onset. My observations allow to add some details to the whole picture of reflexes in onset.

Table 7. Reflexes of the Proto-Khanty **ä* and **ē* in the word-initial position

	#_C	#_n	#_w	#_ŋ
*ä	ɛ	ɛ	e	ɛ
*ē	je	jɛ	?	ɛ

Reflexes of two Proto-Khanty vowels are clearly distinguished before *λ* and *t*.

(27) Kaz. *ɛλ* ‘body’; V., Vj. *el*; Vart. *ĕλ*; Likr. *ĕθ*; Irt. (DN, KoP., Kr.), Ni., Š. *et*; Sy., O. *el* (DEWOS: 56);
(28) Kaz. *ɛλi* ‘capacious’; V., Vj. *elγi*; KoP., Ni. *etə*; O. *eli* (DEWOS: 73);
(29a) Kaz. *jeλ* ‘far, into the distance’; V., Vj. *il*; Trj. *iλ*; Ni. *jet*; Sy. *jel* (DEWOS: 58);
(29b) Kaz. *jeλn* ‘distantly’; V., Vj. *ilən*; Vart. *iλən*; Likr. *iθən*; Ni., Š. *jetn̮*; O. *jeln̮* (DEWOS: 60);
(29c) Kaz. *jeλta* ‘from afar’; Mj., Trj. *iλtə*; Ni. *jetta*; Sy. *jel’ta* (DEWOS: 59);
(30a) Kaz. *jeλɛm* ‘shame’; V., Vj. *ilim*; Vart. *iλim*; Likr. *iθəm*; Mj., Trj., J. *iλəm*; KoP *itim*; Ni., Š. *jetem*; O. *jelem* (DEWOS: 79);
(30b) Kaz. *jeλɛmt-* ‘to put to shame, to make ashamed’; Trj. *iλemtə-*; Ni. *jetəmt-* (DEWOS: 80);
(31) Kaz. *jeλpi* ‘foreside’; Trj., J. *iλpi*; KoP. *ippə*; Ni., Š. *jetpə*; Sy. *jelpi* (DEWOS: 60);
(32) Kaz. *ɛt-* ‘to grow’; V., Vj. *et-*; Trj., J. *ăt-*; Irt. (DN, KoP., Kr.), Ni., Š., Sy., O. *et-* (DEWOS: 201);
(33a) Kaz. *ɛtər* ‘clearly’; V., Vj. *etər*; Trj., J. *ătər*; Irt. (DN, KoP., Kr.), Ni., Š., O. *etər* (DEWOS: 218);
(33b) Kaz. *ɛtermə-* ‘to clear up’; Irt. (DN, KoP., Kr.), Ni., O. *etərmə-* (DEWOS: 218);
(34) Kaz. *ɛtmə-* ‘to go out, to appear’; V., Vj. *etəm-*; DN, KoP. *etəmə-*; Š., Sy., O. *etmə* (DEWOS: 202);
(35) Kaz. *ɛtəλtə-* ‘to introduce, to show in’; V. *etl̮tə-*; J. *ătl̮tə-*; Ni., Š. *etl̮tə-*; Sy., O. *etl̮’tə-* (DEWOS: 202);
(36) Kaz. *jetn* ‘evening’; V., Vj., Likr., Mj., Trj., J., *itən*; Irt. (DN, Kr.) *itən*, *itn̮*; Ni., Š., Sy. *jetn̮* (DEWOS: 217).

Collected examples demonstrate the same reflexes before *p*, *s* and *š*. However, no reliable conclusions can be drawn for these positions, since there is evidence for only one of two Proto-Khanty vowels.

(37) Kaz. *ɛpət-* ‘to overflow, to burst its banks’; Trj. *ăpət-*; Irt. (DN, KoP., Kr.), Ni., Sy., O. *epət-* (DEWOS: 153);
(38) Kaz. *ɛpəλ* ‘smell, taste’; V., Vj. *ewəl*; Vart. *ăwəλ*; Likr. *ăwəθ*; Mj., Trj., J. *ăpəλ*; Irt. (DN, KoP., Kr.), Ni., Sy., O. *epət* (DEWOS: 147);
(39) Kaz. *ɛsəm* ‘female breast’; Trj., J. *ăsəm-*; Irt. (DN, KoP., Kr.), Ni., Š., Sy., O. *esəm* (DEWOS: 194);
(40) Kaz. *ɛsλ-* ‘to hit’; V., Vj. *əsəl-*; Trj., J. *ăsλ-*; Irt. (DN, KoP., Kr.), Ni., Š. *estə-*; Sy. *esəl-*; O. *esl-* (DEWOS: 193);
(41) Kaz. *ješək* ‘expensive’; V., Vj., Vart., Likr. *ičəγ*; Mj., Trj., J. *ičək*, Irt. (DN, KoP., Kr.) *ječək*, Ni., Sy. *ješək*, O. *jesək* (DEWOS: 10).

The word-initial distinction of **ä* and **ē* before *n* has been preserved in a special way: **ä* > *ɛ* (42a-b), **ē* > *jɛ* (41). In connection to this, it is important to note the merge in word-internal position before *n*.

(42) Kaz. *jɛnλ* ‘edge’; V. *in’l*; Vj. *inl*; Trj. *in’λ*; J. *inλ-*; KoP., Kr. *jint*; Ni. *jentl*; Š. *jent*; Sy. *jetl*; O. *in’l̮* (DEWOS: 124);
(43a) Kaz. *ɛnəm-* ‘to rise, to grow’; V., Vj. *enəm-*; Likr., Mj., Trj., J. *ănəm-*; Irt. *enəm-*; Ni., Š., Sy. *enəm-* (DEWOS: 110);
(43b) Kaz. *ɛnməλt-* ‘to raise, to grow’; J. *ănməλtə-*; Ni. *enmət-*; O. *enməlt-* (DEWOS: 110).

Word-initially before *ŋ*, two Pronto-Khanty vowels are merged in *ε* (42–44).

(44) Kaz. *εŋəλ-* 'to groan, to moan'; V. *ö̭ŋəl-*; DN, KoP., Kr., Ni., Š. *eŋət-*; O. *eŋel-* (DEWOS: 141);

(45) Kaz. *εŋkər-* 'to curse'; J. *iŋkər-*; DN, Kr., Sy. *eŋkər*; O. *eŋχər-* (DEWOS: 139).

6.3 Remaining positions demonstrate merging of Proto-Khanty **ä* and **ē*. Before *p* (46–50), *m* (51–52), *n* (53–56), *r* (8–12) they yield *ε*. Before *w* (57–64) and *ś* (65–68), they are reflected as *e*. The data for other positions are insufficient to make any reliable claims. Before *l*, *ŋ*, *k*, only reflexes of **ä* are attested.

(46) Kaz. *jεpəs* 'skewer, spit'; V., *jiwəs*; Vj. *iwəs*; VK, Vart. *jiwəs*; Likr., Mj., Trj., *jipəs*; Irt (DN, KoP., Kr.), Ni., Sy., O. *jiwəs* (DEWOS: 397);

(47) Kaz. *kεpəλ* 'paws from the feet of hoofed animals (moose, deer)'; V., Vj. *kö̭wəl*; Vart. *kɔ̆pəλ*; J. *kŏpəλ*; Irt (DN, KoP., Kr.), Ni. *kepət*; Sy. *kepəl* (DEWOS: 655);

(48) Kaz. *rεp* 'mountain, hill'; Vj. *rew*; Vart. *răw*; Likr. *răp*; Trj., J. *răp*; Irt (DN, KoP., Kr.), Ni., Sy., O. *rep* (DEWOS: 1278);

(49) Kaz. *šεpəŋ* 'rotten'; Vj. *čewəŋ*; Mj., Trj., J. *čăpəŋ* (DEWOS: 292);

(50) Kaz. *λεp-* 'to go in'; Trj., J. *λăp-*; DN, KoP., Kr., Ni. *tep-*; Sy., O. *lep-* (DEWOS: 785);

(51) Kaz. *nεm* 'negative particle used before pronouns'; V., Vj. *nem*; Vart., Likr., Mj., Trj., J. *năm*; Irt (DN, KoP., Kr.), Ni., Sy., O. *nem* (DEWOS: 1000);

(52) Kaz. *sεm* 'eye'; V., Vj. *sem*; Likr. *săm*; Irt. (DN, KoP., Kr.), Ni., Sy., O. *sem* (DEWOS: 1338);

(53) Kaz. *mεna* 'bend, curve'; V., Vj., Trj. *miṇ* (DEWOS: 932);

(54) Kaz. *kεn* 'light'; Mj. *kɔ̆ṇəγ*; J. *kŏnəγ*; Irt. (DN, KoP., Kr.), Ni., Š. *kenə*; Sy. *keṇ*; O. *kon* (DEWOS: 648);

(55) Kaz. *pεnt-* 'to close, to lock'; V., Vj. *pent-*; Trj. *pănt-*; DN, KoP., Kr., Sy., O. *pent-* (DEWOS: 1179);

(56) Kaz. *sεn* 'nit'; V., Vj. *seṇəγ*; Likr., Mj., Trj., J. *sănəγ*; DN, KoP., Kr., Ni. *senə*; Sy. *seṇ*; O. *sen* (DEWOS: 1345);

(57) Kaz. *sew-* 'to plait, to weave'; V., Vj., VK *sö̭γ-*; Vart. *sɔ̆γ-*; Mj., Trj. *săγ̥-*; J. *săw-*; Irt. (DN, KoP., Kr.), Ni., Sy., O. *sew-* (DEWOS: 1308);

(58) Kaz. *sewi-* 'to bind'; V., Vj., Trj. *siγi-*; DN, Kr., *seγej-*; Ni., Š. *sewij-*; O. *siji-* (DEWOS: 1312);

(59) Kaz. *pew-* 'to freeze, to feel cold'; V. *pö̭γ-*; Trj. *păγ̥-*; DN, KoP. *peγ-*; Ni. *pew-*; Š. *peγ-*; Sy., O. *pew-* (DEWOS: 1115);

(60) Kaz. *sewər-* 'to hew'; Likr., Mj., Trj. *săγ̥ər-*; J. *săwər-*; Irt. (DN, KoP., Kr.), Ni., Š., Sy., O. *sewər-* (DEWOS: 1321);

(62) Kaz. *sewəs* 'stern of a boat'; V., Vj. *sö̭γəs*; VK *söγəs*; Vart. *sɔ̆γəs*; Likr., Mj., Trj. *săγ̥əs*; J. *săwəs*; Irt. (DN, KoP., Kr.), Ni., Š., Sy., O. *sewəs* (DEWOS: 1323);

(62) Kaz. *tewən* 'windless'; V.,Vj. *teγən*; Trj. *tăγ̥ən*; J. *tăwən*; DN *tewen*; KoP., Kr. *tewin*; Ni., Sy., O. *tewən* (DEWOS: 1415);

(63) Kaz. *wew* (commentary on the semantics see in Appendix 2); V., Vj. *wö̭γ*; VK *wöγ*; Vart. *wɔ̆γ*; Likr. *wăγ̥*, Mj. *wŏγ̥*; Trj. *wɔ̆γ̥*; J. *wŏw*; DN, KoP., Kr. *weγ*; Ni. *wew*; Š. *weγ*; Sy. *wew*; O. *wej* (DEWOS: 1571);

(64) Kaz. *λew* 'servant, labourer'; V. *lö̭γ*; Vj. *jö̭γ*; Likr. *θăγ̥*; Mj., Trj. *λăγ̥*; KoP. *tew*; Kr. *teγ*; Ni., Š. *tew*; O. *lew* (DEWOS: 733);

(65) Kaz. *peś* 'haunch'; Vj. *pit´-*; Irt. (DN, KoP., Kr.) *pet´-*; Ni., Š., Sy. *peś*; O. *piś* (DEWOS: 1251);

(66) Kaz. *śeś* (*seńś-*) ‘bast’; V., Vj., Likr., Mj., Trj., J. *sińt´*; DN, KoP., Kr. *seńt´*; Ni. *śeńś*; O. *śiś* (*śiṇś-*) (DEWOS: 1347);
(67a) Kaz. *weś* ‘beauty’; Vj., Trj. *wit´*; KoP., Kr. *wet´*; Š., O. *weś* (DEWOS: 1648);
(67b) Kaz. *weśəŋ* ‘beautiful (also of animals and clothes)’; V., Vj. *wit´əŋ*; Trj. *wit´əŋ*; DN, KoP., Kr. *wet´əŋ*; Š., O. *weśəŋ* (DEWOS: 1648);
(67c) Kaz. *weśəp* ‘beautiful’; Trj. *wit´əŋ*; KoP., Kr. *wet´əp*; Ni. *weśəp* (DEWOS: 1648);
(68) Kaz. *weś-* ‘to sew (on)’; V., Vj. *wet´-*; Trj. *wăt´-*; Ni. *weś-* (DEWOS: 1646).

6.4 Table 8 summarizes reflexes of Proto-Khanty **ä* and **ē* in word-internal position.

Table 8. Reflexes of the Proto-Khanty **ä* and **ē* in the word-internal position

	_p	_m	_w	_t	_n	_l	_r	_š	_λ	_ś	_k	_ŋ
*ä	ɛ	ɛ	e	e/ɛ[1]	ɛ	ɛ	ɛ/e[2]	ɛ/e[2]	e/ɛ[1,3]	e	e	ɛ
*ē	ɛ	ɛ	e	ɛ	ɛ	?	ɛ	?	ɛ	e	?	?

Notes: 1) **ä* > *ɛ* / *m, p_t, λ*; 2) **ä* > *e* / *k_r, k_š*; 3) **ä* > *ɛ* / *t_λ*.

7. Conclusions

In its current state, Kazym dialect demonstrates a phonemic contrast between the tense and lax front vowels only in limited positions. In positions of neutralization, vowel quality can be predicted by the following consonant and by the preceding *j*. The original source of the contrast are two Proto-Khanty vowels which yield different reflexes in one positions but are merged in others. The relative complexity of the rules which describe the distribution of reflexes confirms once again the Neo-Grammarian statement that the majority of imaginary “exceptions” can be explained by formulating additional sound laws. Being a result of analogical leveling, the only real exception in my database has no phonological explanation. Deviations from the postulated pattern are also found in borrowings, including interdialectal Khanty borrowings. So the diachronic study of Kazym mid front vowels is a perfect and concise example of the whole Neo-Grammarian methodology based on the principle of *Ausnahmslosigkeit der Lautgesetze*.

Acknowledgments

I am incredibly grateful to Mikhail Zhivlov for all the important notes he made and advice he gave after reading the draft of the present paper, and to Svetlana Toldova, Alexey Kozlov and especially to Ksenia Pecherskaya for the opportunity to take part in the expedition in Kazym.

The work is supported by the RFBR grant №19-012-00627 “Syntax and Semantics of Uralic and Altaic languages: converging functional typological and formal perspectives”.

Appendix 1. Corrigenda to Solovar 2014

Solovar 2014		Correction	Commentary
109	*jɛšawəʌ*	*ješawəʌ*	This variant is cited on p. 106, it is also confirmed by the field study.
146	*lɛkkər*	*lekkər*	This form has been recorded in my field notes. It is also more likely from the systemic point of view.
146	*lɛkśitti*	*lekśitti*	This variant is cited below on p. 146, it has also been confirmed by the field study.
230	*ńemәś karti*		The lexeme is not used currently; it is completely unknown to speakers in Kazym. The example has been excluded from consideration.
253	*pəsmɛkʌəti*	*pəsmekʌəti*	More likely *pəsmekʌəti*, the form is found in DEWOS. It is also more likely from the systemic point of view.
269	*pɛʌa (i)*	*pɛʌa*	Only the form *pɛʌa* has been confirmed by the field study. The form *pɛʌi* is recognized as one from a foreign dialect.
293	*seŋənwej*	*sewənwej*	The compound literally means 'knitted boots', i.e. 'socks'. The first part of the compound is related to the verb *sewti* 'to knit, to crochet' (Solovar 2014: 292).

Appendix 2. Corrigenda to DEWOS

In the transcription used in DEWOS for Kazym dialect, dental *n* and retroflex *ɳ* are distinguished. In modern descriptions (Kaksin 2010, Solovar 2014, Solovar et al. 2016) this distinction is not noted.

DEWOS		Correction	Commentary
782	*ʌɛŋkər*	*leŋkər*	The form *leŋkər* is cited in Solovar 2014: 146. Apparently, it is a loan from another dialect. Neither one nor the other form is used by speakers at the present moment.
828	*ļɛkəp*		The lexeme is not currently used; it is completely unknown to speakers in Kazym. The example has been excluded from consideration.
986	*ɳɛki*		The lexeme is not currently used; it is completely unknown to speakers in Kazym. The example has been excluded from consideration.
1069	*ńer, ńerəŋ*	*ńɛr, ńɛrəŋ*	According to modern data: *ńɛra-suχa* 'to spite somebody', *ńɛrəŋ-suχəŋ* 'obstinate' (Solovar 2014: 230), hence, *ńɛr* 'ire', *ńɛrəŋ* 'nervous' are more apparent. It is also more likely from the systemic point of view.
1427	*teʌ*	*tɛʌ*	According to Solovar 2014: 317, the form is *tɛʌ*; the same pronunciation has been confirmed by the field study.
1145	*pɛʌa, pɛʌi*	*pɛʌa*	Only the form *pɛʌa* has been confirmed by the field study. The form *pɛʌi* is recognized as one from a foreign dialect.
1406	*tɛkən-*	*tekən-*	Solovar (2014: 317) lists this lexeme with *e*. It has also been confirmed by field study.
1571	*wew*		Glossed in Solovar 2014 as 'weakness, tiredness; weak (of a man)', but as 'power, strength; heavy (work)' in DEWOS; the collocation *wewa jĭs* 'to become weak' is also cited in DEWOS. The semantic development remains unclear.

Abbreviations

Languages and dialects

DN — Upper Demyanka; Irt. — Irtysh; J. — Yugan; Ko. — Konda; KoP. — Konda on the basis of Paasonen's notes; Kr. — Krasnoyarskie on Konda; Likr. — Likrisovsky; Mj. — Maly Yugan; Ni. — Nizyam; O. — Obdorsk; ProtoKh. — Proto-Khanty; Rus. — Russian; Š. — Sherkal; Sy. — Synja; Trj. — Tremyugan; V. — Vakh; Vart. — Vartovsky; Vj. — Vasyugan; VK — Verkhne-Kalymsk

Sources

DEWOS = Steinitz 1966–1993.

References

Kaksin, A.D. 2010. *Kazymskij dialect xantyjskogo jazyka* [Kazym dialect of Khanty language]. 2nd edn. Khanty-Mansiysk: Yugorskij universitet.

Kurkina, G. G. 2010. *Vokalizm xantyjskogo jazyka* [Vowel system of Khanty language]. Novosibirsk: Sibirskij xronograf.

Lytkin, V. I. 1964. *Istoričeskij vokalizm permskix jazykov* [Historical study of vowel system in Permic languages]. Moscow: Nauka.

Nikolaeva, Irina. 1999. *Ostyak. Languages of the World. Materials 305*. München — Newcastle: LINCOM.

Solovar, V. N. 2014. *Xantyjsko-russkij slovar' (kazymskij dialekt)* [Khanty-Russian dictionary (Kazym dialect)]. Khanty-Mansiysk: Format.

Solovar, V. N., G. L. Naxračeva, A. A. Šijanova. 2016. *Dialekty xantyjskogo jazyka* [Dialects of Khanty language]. Khanty-Mansiysk.

Steinitz, W. 1950. *Ostjakische Grammatik und Chrestomathie mit Wörterverzeichnis* [Khanty grammar and chrestomathy with glossary]. 2nd ed. Leipzig: Harrassowitz.

Steinitz, W. 1966–1993. *Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache*. Berlin: Akademie Verlag.

Steinitz, W. 1975. Skizze der Phonetik der Synja-Mundart. In: W. Steinitz. *Ostjakische Volksdichtung und Erzählungen aus zwei Dialekten. Texte.* Janua Linguarum. Series Practica 254, vol. 1: 3–38. The Hague.

Zhivlov, Mikhail. 2007. K voprosu o rekonstrukcii obsko-ugorskogo vokalizma [Towards a Reconstruction of the Ob-Ugrian Vocalic System]. In: Ilya Smirnov (ed.). *Aspekty komparativistiki 2. Orientalia et Classica VI*: 281–309. Moscow: RSUH Publishers.

Zhivlov, Mikhail. 2010. Studies in Uralic vocalism I: a more economical solution for the reconstruction of the Proto-Permic vowel system. *Journal of Language Relationship* 4: 167–176.

И. М. Егоров. Происхождение и синхронный статус передних гласных среднего подъема в казымском диалекте хантыйского языка.

Настоящая статья посвящена синхронному и диахроническому анализу двух гласных переднего ряда среднего подъема (*e* и *ɛ*) в казымском диалекте хантыйского языка. Исследование дистрибуции этих гласных показало фонологический контраст в одних позициях и нейтрализацию в других. Фонологический статус обеих гласных дополнительно подтвержден перцептивным экспериментом. Источником e и ɛ являются две пра-хантыйские гласные, которые различаются или совпадают в зависимости от консонантного контекста. Фонологический контраст распространился в некоторых новых позициях благодаря ненецким, коми-зырянским и междиалектным хантыйским заимствованиям.

Ключевые слова: фонология, вокализм, историческая фонология, диалектология, хантыйский язык

Ilya Gruntov[†], *Olga Mazo*[‡]
[†] Institute of Linguistics of the Russian Academy of Sciences / Yandex, Moscow; altaica@yandex.ru
[‡] Russian State University for the Humanities / Higher School of Economics, Moscow; olga_mazo@list.ru

Mongghul, Mangghuer and beyond: estimating the proximity

The paper's chief goal is to evaluate the intensity of relations between various idioms of Mongolic languages within the Qinghai-Gansu Sprachbund, with special focus on Mangghuer and Mongghul dialects. On the basis of 58 grammatical and phonological shared innovations we attempt to deepen our understanding of the pathways and patterns of language change taking place in this area.

Keywords: Qinghai-Gansu Sprachbund; language contacts; Monguor language; Shirongolic languages; Mangghuer language; Mongghul language; Mongolic languages; shared innovations.

Mongghul and Mangghuer are two closely related endangered languages of the Mongolic family, spoken in Qinghai and Gansu provinces of China. For a long time they have been generally considered as a single Monguor language with two main dialects. Although the division of these idioms into at least two languages has gradually become accepted among specialists, new linguistic works continue to be published in which both languages are still referred to as Monguor dialects.

Our paper is an attempt to clarify and evaluate the intensity of relations between various idioms of QG Mongolic languages with special focus on Mangghuer and Mongghul, and to investigate various isoglosses within these languages. We do not discuss here the peculiarities and features shared by all QG Mongolic languages.

The languages of the region belong to the so called Qinghai-Gansu Sprachbund (Yellow River language union, Amdo Sprachbund, the Northwest China Sprachbund, etc., see Dwyer 1992, 1995; Slater 2003a; Janhunen 2007, 2012). This Sprachbund includes languages and dialects of four families: Sinitic (Northwestern Mandarin, Gangou, Tangwang, Wutun, Linxia /Hezhou/), Bodic (Amdo Tibetan), Turkic (Salar, Western Yugurs), and Mongolic (Mongghul, Mangghuer, Dongxiang, Baoan /Qinghai and Gansu varieties/, Kangjia, Shira Yughur).

Due to intensive interaction, languages of different types started to share phonetic, morphologic and syntactic features. Many of the speakers are proficient in two or more languages or dialects: Standard Chinese (Putonghua), Qinghai Mandarin, Amdo Tibetan. The choice of *lingua franca* in the region depends on religious identity: it is Amdo Tibetan for Buddhists, but Northwestern Mandarin for Muslims. Such interaction has caused serious changes in language structure and the appearance of idioms like Wutun, Gangou, Tangwang which are sometimes even labelled in linguistic works as "creole" or "hybrid".

The official Chinese classification of minorities refers to Mongghul and Mangghuer communities as Tu people (*Tuzu* 土族). Actually, *Tuzu* is one of the 55 national minorities, and it is a common name for people of different nationalities and speakers of different languages: Mongol-speaking Mongghul, Mangghuer, and Qinghai Baoan (Gansu Baoan are officially recognized as a separate Baoan nationality); speakers of Wutun (a Sinitic language); and Tibetan-speaking Shaowu Tuzu, who were considered Tibetans until 1986 (Janhunen et al. 2007).

According to Poston and Xiong (2014: 118), the number of the Monguor in 2010 was 289,565. The number of Monguor speakers in 2000 was about 100,000 (Zhang 2012: 37).

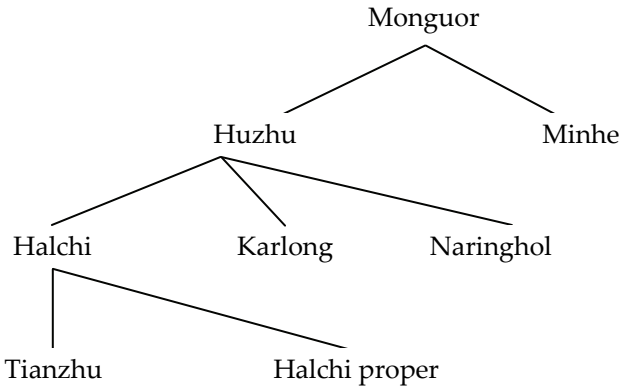
The Mangghuer (or Minhe Mangghuer, *минхэ*) speakers live in Minhe Hui and Tu Autonomous County (Qinghai); their number is approximately 25,000 (Slater 2003b: 307). Little is known about dialectal varieties of Mangghuer. Slater (2003b: 308) states that "Speakers report noticeable dialect differences, particularly in the areas of phonology and lexicon, but a systematic study has not yet been undertaken."

Mongghul (also named Huzhu Mongghul, *хуцзу*) speakers live in Huzhu Tu Autonomous County, Datong Hui and Tu Autonomous County, Ledu district, Menyuan Hui Autonomous County (Qinghai), Tianzhu Tibetan Autonomous County, Yongdeng county, Linxia county (Gansu) (Junast 1981: 1, Georg 2003: 286).

There are two main theories about the origins of the Monguor: the 'Mongol theory' and the 'Tuyuhun theory'. According to the first one, the Monguor are descendants of Mongol soldiers who arrived into the region in the 13th century. The second theory suggests that the Monguor are the descendants of the people of the Tuyuhun (吐谷浑) kingdom (3th–7th century) (for details see e.g. Cui Yonghong et al. 2015; Georg 2003: 287; Slater 2003a: 16–19; Schram 1954–1961; Lü Jianfu 2002; Li Keyu 1993, 2008, etc.).

Dan Xu and Shaoqing Wen (2017: 60–61) have studied Y-chromosomal data of the Huzhu Monguors. According to the results, the most frequent lineages were haplogroups R1a1a-M17 13,22% (predominant in some Turkic-speaking populations), D1-M15 10, 74% (associated with Tibeto-Burman populations), O3-M122, O3a2c1a-M117, O3a1c-002611 20.7% (predominant in Han and Hui populations, but rare or absent in Mongolic and Turkic populations). Haplogroups C3*-M217, C3c-M48 and C3d-M407, the common paternal lineage in Mongolic-speaking populations, appear with relatively low frequency — 9,09%.

De Smedt and Mostaert (1929: 145) and later Todaeva (1973: 11) list three Mongghul dialects: Naringhol, Halchighol, and Fulaan Nura. Janhunen (2006: 28; Janhunen et al. 2007: 179–180) considers Halchighol and Naringhol to be two separate languages. Georg (2003: 286) mentions a possible dialectal variety of Mongghul spoken in Menyuan Hui Autonomous County (Qinghai). Faehndrich (2007: 11–17) discusses Naringhol, Halchighol, Karlong (Fulaan Nura), Datong (possibly extinct), Tianzhu, Menyuan, Ping'an varieties. She proposes the following tentative Monguor family tree (Faehndrich 2007: 242):



The first records on the Monguor people (who were called *'Dada'*, *'Damin'*, *'Tuda'*, *'Tumin'*, and *'Turen'*) can be found in documents going back to the Ming dynasty (1368–1644; see Cui Yonghong et al. 2015: 18). The Catholic missionary Évariste-Régis Huc described Mangghuer spoken in Sanchuan as a Mongolic language with Chinese and Tibetan elements (Huc 1850: 36). N. M. Przheval'skii (1875: 199) in his materials gave a description of the local people of Qing-

hai named *Daldy*. G. N. Potanin mentions that the Chinese and the Dongxiang called them *Turen*, while other Mongols gave them the name of *Dalda* or *Doldo* (Potanin 1893: I, 342; II, 410). He included some Mangghuer words and data on other Shirongol languages in the materials of the expedition (1893: II, 410–425).

Monguor varieties are not evenly represented in scientific literature. Halchighol seems to be the most frequent idiom in linguistic descriptions. Dominik Schröder published a description of the religious life (1952–1953), two texts (1959–1970), and description of grammar (1964). "Mongorskii yazik" (Todaeva 1973) contains a detailed grammar of Halchighol Monguor with data on Naringhol, Fulaan Nara, Minhe, and different types of texts. Other Monggul materials based on Halchighol data are the grammar descriptions of Junast (1981) (with some comparison with Mangghuer), that of Chingeltei and Li Keyu (1988) (with comparison with other Mongolic languages), vocabularies of Hasbaatar (1985) and Li Keyu (1988), materials of Chingeltei (1986), Chuluu (1994), short descriptions of Georg (2003) and Sinor (1952), etc. Dpal-ldan-bkra-shis, K. Slater et al. (1996) published materials and a small dictionary of both Mangghuer and Mongghul. Different aspects of Mongghul phonetics, grammar, vocabulary and dialectology are discussed in Róna-Tas (1960, 1962, 1966), Chingeltei (1989), Kakudo (1987, 1997), etc.

Antoine Mostaert and Albrecht de Smedt wrote the first detailed description of Naringhol Mongghul including phonetics (1929–1931), grammar (1945), and a big Monguor-French dictionary (1933).

Faehndrich (2007) focused her dissertation on the Fulaan Nura or Karlong.

The Mangghuer grammar was published by Slater (2003a); sketches of Mangghuer grammar were written by Junast and Li Keyu (1982), Slater (2003b).

The phonology of the Mongol languages of Qinghai-Gansu is studied in detail by Nugteren (2011), who also compared some grammar and lexical features. The position of Monguor languages within the Mongolic family was discussed in many papers, such as Sanzheev (1952), Poppe (1955), Luwsanwandan (1959), Doerfer (1964), Binnick (1987), Nugteren (1997, 2011: 34–56), Rybatzki (2003: 386), Gruntov, Mazo (2015), etc.

Below we shall discuss the characteristics that differentiate Mangghuer from Mongghul, and Mongghul dialects from each other; the focus will be on what features they share with the other Mongolic languages of the region. In our paper we do not consider lexical borrowings (e.g. borrowing of numerals), but borrowing of grammatical elements is taken into account.

Phonetics

Mangghuer (Slater 2003b: 309)

Consonants

	labial	alveolar	retroflex	palatal	velar	uvular
stop	p b	t d			k g	q G
affricate		ts(c)[1] dz(z)	tʂ (ch) dʐ (zh)	tɕ (q) dʑ (j)		
fricative	f	s	ʂ (sh)	ɕ (x)	h	
nasal	m	n			ŋ	
liquid		l	r			
glide	w			j		

[1] Many authors use *pinyin*-based orthography; the corresponding transcriptions are given in brackets.

Fricative *f* generally occurs in Chinese borrowings.
Five short vowels: /a/, /e/, /i/, /o/, and /u/.

Mongghul (Georg 2003: 290)

Consonants

	labial	alveolar	retroflex	palatal	velar	uvular
stop	p b	t d			k g	G
affricate		ts(c) dz(z)	tʂ (ch) dʐ (zh)	tɕ (q) dʑ (j)		
fricative	f	s	ʂ (sh)	ɕ (x)	h	
nasal	m	n			ŋ	
liquid		l	r			
glide	w			j		

Five short vowels: /a/, /e/, /i/, /o/, /u/ and five long vowels: /ā/, /ē/, /ī/, /ō/, /ū/.
Affricates *ts*, *dz*, *tʂ*, *dʐ* occur mostly in Chinese and Tibetan loanwords.

In both languages stops and affricates are distinguished by aspiration; opposition of aspirated voiceless consonants vs. unaspirated voiceless consonants are often represented as opposition of strong vs weak consonants.

Mangghuer seems to be the only Mongolic language that has tones in native words. Dwyer (2008) discovered that several pairs of homophones differ in tones: one of the members of a pair is pronounced with a high tone, another with a low tone, e.g.

a. *wulang* [ʷu^{22}lã^{55}] 'drinking'
b. *wulang* [ʷu^{11}lã^{51}] 'many' (Dwyer 2008: 128).

There are also some indications that Baoan is currently in the process of developing a tonal system (Li:1986), but so far these tones have been found in loanwords only, although they do not match the tones in the donor language.

In the following cases the same features can be observed in both Mangghuer and Mongghul, but in Mongghul they appear more systematically or more often:

— development of strong obstruents. If both the word-initial obstruent and the word-internal obstruent are strong, the word-internal obstruent weakens, e.g. *huja-* 'to bark' < **kuca-* (Georg 2003: 291). The sequence of initial weak obstruent + internal strong obstruent changes to initial strong obstruent + internal weak obstruent, e.g. *pujig* 'book' < **bicig* (Georg 2003: 291). Weakening of the internal obstruent can also cause the appearance of secondary initial *h-*, e.g. Mangghuer *ʃɜu~sɜu* < **husun* < **usun* 'water' (Nugteren 2011: 38), Halchighol *xaldan* (Todaeva 1973: 372), Naringhol *xardam* (SM 1933: 160) < **altan* 'gold' (for details see Khelimsky 1984: 27; Georg 2003: 291–292; Nugteren 2011: 38, 252; Faehndrich 2007: 42–43). Such developments can be observed in all Mongolic languages of the region; in Monguor they are more frequent and systematical, and in Mongghul they appear more systematically than in Mangghuer;
— elision of vowels is also typical of Qinghai-Gansu languages, but it seems to appear more frequently in Mongghul, Baoan and SY. In many cases it leads to the appearance of new consonant clusters, including word-initial ones. Consonant clusters are found in all Mongolic languages of Qinghai-Gansu except for Dongxiang (but cf. Kim 2013: 351), whose syllabic structure has undergone strong Chinese influence. It is important to note that Mangghuer (for which we can compare historical records from the end of

the 19th century with modern data) has a tendency to eliminate secondary clusters via insertion of epenthetic vowels (Nugteren 2011: 91, 124–125, 127, 130–131, 133–134, 199). The precise number of such clusters differs: Georg lists 24 clusters in Halchighol (Georg 2003: 293). Faehndrich (2007: 73–74) notes that Karlong is in the process of developing new clusters (in addition to the 25 already attested clusters), which are not permitted in other Mongghul dialects. Junast and Li (1982: 478) state that Mongghul has twice as many clusters as Mangghuer, but they do not give the exact numbers. Anyway, Mongghul has many more initial clusters than Mangghuer.

The comparative data are given in Table 1.

There is a single phonetic shared innovation that separates Mangghuer and all Mongghul dialects from other Mongolic languages of Qinghai-Gansu: the reflex of **kï-/*ki-* (*ći* in Monguor and *ki*/*xi* in the other languages, see Todaeva 1973: 30–31; Georg 2003: 291; Nugteren 2011: 219–220).

Phonetic variability within Mongghul dialects may be illustrated as follows:

— **-l* is retained as *-l* in Halchighol and Karlong, but becomes *-r* in Mangghuer and Naringhol. The same innovation can also be found in Kangjia, where final **-l* has several reflexes: *-Ø /-r/-l/-lV/-n* (Sečenčogt 1999: 28; Nugteren 2011: 244);
— **-m* remains as *-m* in Naringhol and Halchighol, becomes *-n*/*-ŋ* in Mangghuer, and can be realized as *-m* or *-n* in Karlong. Mangghuer shares this innovation with Dongxiang and Kangjia. Sečenčogt (1999: 28) notes that in Kangjia **-m* can also be realized as *-r*.

The Monguor idioms have one innovation in common with SY, namely *ə* as a reflex of **U* in accented syllables (Nugteren 2011: 128).

The main phonetic differences between Mongolic languages of the region are presented below.

— CM **-b-* in intervocal position is retained as obstruent *-b-*, but turns to fricative *w*, *y* or *Ø* in the other languages (Nugteren 2011: 208–209);
— CM **-rb-* remains as *-rb-*, but usually develops into *-rw-* or *-r-* in the other languages (for details see Nugteren 2011: 209);
— CM **-bC-* is replaced by *-G-* (usually before affricates and sibilants). In Mongghul it is either retained as *-bC-* or changed into *s*, *ʂ*, *ś* +C; it is also retained in SY and Baoan, changes to fricative or affricate *+C / -r- +C* in Kangjia, and is dropped in Dongxiang (for details see Nugteren 2011: 213);
— final CM **-r* fuses with the preceding vowel into *ɔ˞* (Slater 2003a: 31) in Mangghuer, most probably under Chinese influence, but this change does not take place in Mongghul;
— CM **ǰ* and **č* develop into two sets (retroflexes and alveopalatals) in Mangghuer, but only into one set of affricates in other languages: alveopalatal in Mongghul and Baoan, palatal in SY and Kangjia and retroflexes in Dongxiang (Nugteren 2011: 218);
— velar **k(a)* is reflected as *q-* in Mangghuer and Dongxiang but as *h-* in Mongghul, Baoan and Kangjia; thus, Mongghul lacks aspirated uvulars. SY has both reflexes *q-* and *h-* (Nugteren 2011: 220);
— CM **h-* before originally rounded vowels develops into *f-* in Mongghul and Dongxiang in contrast with *h-* in Mangghuer and SY. In Baoan and Kangjia both reflexes exist (for details see Nugteren 2011: 250);
— the vowel *i* is realized as [ɿ] and [ʅ] following apicals and retroflexes in Mangghuer, Dongxiang and Kangjia under the influence of Chinese;

Table 1. Phonetic and phonological innovations in Qinghai-Gansu Mongolic.

	Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
tonal opposition	+	-	-	-	-	+-	-	-
*-l	-r	-r	-l	-l	-l	-l	-ŋ	-Ø /-r/-l/-n/-lV
*-m	-n/-ŋ	-m	-m	-m/-n	-m	-m	-n/ŋ	-n/-ŋ/-r
*U in accented syllables	usually ə				u			
*-r	V+r > əˑ	-r					Ø/-rV	
*k(a)-	q-	x-			q-/x-	x-	q-	x-
*-bV-	-b-	-v- / -Ø- / -j-			-β-/ Ø/-ɣ-	-β-/ Ø/-j-		v-/ Ø/
*-rb-	-rb-	-rv-/-r-			-rβ-	-br-/-wr-/-rw-	-r-/ -w-	-r-
*-bC[2]-	-G-	-bC- / s, ş, ś + C			-bC-		-C-	fricative +C / r+ C
*-d	Ø	-d/r			-d	Ø/-dV/r	-dV	-dV/r
*ǰ and *č	both retroflexes and alveopalatals	alveopalatal			palatal	alveopalatal	retroflexes	palatal
CM*hU-	h-	f-			h-	h-/f-	f-	h-/f-
distinction between short and long vowels	-	+				+ -	-	-
long vowel in the denominal verb suffix –rA / the intransitive marker -rA	-	+	+-			-		
[ɳ] and [ɭ] following apicals and retroflexes	+	-					+	

[2] C = sibilants or affricates.

— secondary distinction of long and short vowels is present in Mongghul, Shira Yughur, Ñantoq Baoan and Xiazhuang Baoan. Mangghuer, Dongxian, Dahejia Baoan and Kangjia lack it, although this loss might have been independent (and secondary as well); see details in Nugteren: 2011: 134–137. Chingeltei (1986) provides some examples of Mangghuer words with long vowels, but this phenomenon has not been confirmed in other grammatical manuals and descriptions;
— Nugteren (2011: 190–191) discusses the existence of vowel lengthening (or vowel length preservation) in Mongghul (and partly in SY) in the denominal verb suffix *-rA* and the intransitive marker *-rA*.

Word formation

Borrowed Chinese or Tibetan verbs are adopted by attaching the following sets of suffixes:

— Mangghuer *-la*, *-li* (Slater 2003a: 113–114), Mongghul *-la* (Todaeva 1973: 95–96), Karlong *-la* (Faehndrich 2007: 130), SY *-la* (Tenishev, Todaeva 1966: 63) and Dongxiang *-la*, *-lo*, *-lie* (Todaeva 1961: 40), e.g. Halchighol *gua* (Chinese 挂 *guà* 'to hang') — *guala* (Todaeva 1973: 96).
— Mangghuer *-ke*, *-ge* (Slater 2003a: 113), Naringhol *-ki*, *-gi*, Halchighol *-ki*, *-gi* (Todaeva 1973: 95), Karlong *-ki* (Faehndrich 2007: 130–131), SY *-ge* (Tenishev, Todaeva 1966: 63), Baoan *-ge* (Todaeva 1973: 95), Dongxiang *-gie* (Kim 2013: 352). Todaeva (1973: 75) notes that in Halchighol and Naringhol the variant of the suffix depends on the type of initial consonant: *laki* 'to lead', but *ta:gi* 'to step'.
— Mangghuer *-ra* and Dongxiang *-re ~ -ro* (Kim 2013: 352): Mangghuer *gaoxinra* ''be happy' < Chinese *gāoxìng* 高兴 (Sater 2013: 113).
— Mangghuer *tu*: *pintu* 'be equal' < Chinese *píng* 平 'equal' (Slater 2013a: 113).

In Karlong, the diminutive/singular suffix *-ŋge* < **nige* 'one' can be attached to some verbal forms to weaken the strength of the statement (Faehndrich 2007: 183–184) or to underline a quick change of successive actions (Todaeva 1973: 129).

Only in Mangghuer can one find the comparative marker *-her* (Slater 2003b: 312).

Karlong uses a specific adjective intensifier *-Gula* (Faehndrich 2007: 84, 138).

Halchighol *-mal* and Naringhol *-mar* have preserved the Common Mongolic resultative marker (Georg 2003: 294), while other idioms have lost it.

Table 2. Word formation innovations in Qinghai-Gansu Mongolic

	Mangghuer	Naringhol	Halchigol	Karlong	SY	Baoan	Dongxiang	Kangjia
adaptation of borrowed stems	la, li	la	la	la	la		la, lie, lo	
	ke, ge	gi	ki,gi	kɨ	ge	ge	gie	
	ra						re, ro	
	tu							
singulative suffix with verbs				ŋge				
intensifier				Gula				
comparative	her							
resultative		mar	mal					

Noun

Several Shirongolic languages (and Turkic Salar) have developed a postpositive indefinite article which goes back to Common Mongolic **nige-n* 'one', most probably due to Tibetan influence (Nugteren 2013: 227): Kangjia *-niγe/-nγe/-γe* (Sečenčogt 1999: 89), Bonan *-gə/-ngə* (Chen and Chingeltei 1986: 81), Mangghuer *-ge/-gi* (Todaeva 1973: 43) and Mongghul *-ŋgə* (Junast 1981a: 18), *-nge/-ge* (Todaeva 1973: 43; Faehndrich 2007: 84–86) (Table 3).

Cf. Kangjia:

enə kun niγe sɔ va
this person SG good be
'This person is good.' (Sečenčogt 1999: 90)

Table 3. Indefinite article in Qinghai-Gansu Mongolic.

Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
+	+	+	+	—	+	—	+

The dualis marker *-Gula/-Gurla* in Karlong surprisingly does not correspond to the paucalis marker *-Gula* in Baoan, since the Karlong form goes back to *Gu:r* 'two' + *la* 'Instrumentalis' (Faehndrich 2007: 89), while Chen and Chingeltei convincingly argue that Baoan *-Gula* goes back to grammaticalization of *Gulla* 'three persons' < *γurbaγula* (Chen & Chingeltei 1986: 85–86, 159).

The pluralis marker *-tang* mentioned in (Chuluu 1994: 5) for Halchighol is not confirmed by any other source and, moreover, Faehndrich underlines that all her language consultants refused to accept this form as normal; consequently, we did not include it into the table.

The Halchighol / Karlong affix *-maŋe/-maŋge* is used for associative plurality only (Georg 2003: 295, Faehndrich 2007: 88).

Data on dualis and plural markers are presented in Table 4.

Table 4. Dualis and plural markers in Qinghai-Gansu Mongolic

	Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
Dualis				-Gula/-Gurla				
Pluralis		-sGi	-sge	-sgï				
	-se			-sï	-s			-sʉn/-sɯ
			-ŋgula	-ŋgula				
			-ŋgu	-ŋgu				
				-la		-la	-la	-la
				-sïla			-si-la/-si-la/ -s-la	-la-sʉn
	-taŋ	-taŋ					-tan	
Associative plurality			-maŋe	-maŋge				

Case

An important feature of historical Mongolic morphonology is the reflexation of *-n*-stem declension. Unstable *-n* gets lost in NOM in Mangghuer and Mongghul, but is preserved in SY

and (not consistently) in Baoan, Dongxian and Kangjia (for details see Todaeva 1973: 30–31, 47; Georg 2003: 291, 295–296; Nugteren 2011: 45–47; Rybatzki 2003: 375). However, while in other Shirongolic idioms forms with originally unstable *-n* are lexicalized either in *-n* form or in *-Ø* form, in Mongghul *-n* is absent in NOM, but occurs in DAT.LOC in Karlong, and in DAT.LOC and ABL in Halchighol and Naringhol (Faehndrich 2007: 100).

Merging of GEN and ACC in one affix across the nominal paradigm (but not in Pronominal declension, see below) is a common feature for all QG languages, and thus we did not include it into the table. However, it is worth mentioning that Faehndrich (2007: 110) lists variants *-ni*, *-na* and *-nə* for Genitive in Tianzhu dialect, but only *-ni* and *-nə* for Accusative.

There is a peculiar reflex of DAT.LOC *DU > *di* (Rybatzki 2003: 377). According to Rybatzki, it occurs only in Shira Yughur and Halchighol; however, it is also present in Karlong (Faehndrich 2007: 94).

Mongghul locative *-rV* corresponds to Dongxiang Prolative *-rə* (Faehndrich 2007: 110). The Tianzhu dialect of Mongghul demonstrates a longer form of the locative *-tərа*, which possibly may imply that Todaeva (1973: 51–52) was right assuming that the locative *-rV* in Mongghul might be a grammaticalized form of the locative adverbs *dotu-ra* 'inside' and *de'e-re* 'over' etc. However, it might as well be an archaism preserving the same locative marker that can easily be found in the aforementioned locative Mongolic adverbs and in *dow-ra* 'below' (Gruntov, Mazo 2020 forthcoming), and in that case Tianzhu form might be an instance of a compound affix (DAT.LOC + PROL).

Monguor languages have developed a new DIR marker (Naringhol *-dzi*, *-dzidzi*, Mangguer *-dzi*, *-dzudzi*), which is a result of grammaticalization (Slater 2003: 171–172; Nugteren 2014).

Table 5. Case in Qinghai-Gansu Mongolic.

	Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
unstable *-n* in Nom	—	—	—	—	+	+-	+-	+-
DAT.LOC *DU > di			di	dɨ	di			
LOC –rV		-re	-ra/-re/-ri	-rɨ			-rə (Prolative)	
DIR	-dzi/-dzidzi	-dzi/-dzuǯi	—	—	—	—	—	—
INS | SOC	-la	-la/-ra	-la	-la	-lV	-ɢalə	-lə / -ɢala	-ɢala
COM	-tai	-dē	-dē/ -dī	—	-di	—	—	—

Personal and impersonal possession

Loss of 1st and 2nd-person possessive affixes in Monguor languages, Baoan and Kangjia seems to be an areal innovation (Rybatzki 2003: 380).

Reflexive (impersonal possessive particle) in QG languages goes back to the Common Mongolic reflexive particle **ben.*

Table 6. Possessive affixes in Qinghai-Gansu Mongolic

	Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
Reflexive particle	na / nang	nā	nā	na/ nā[3]	-an	—	ne	—
1st and 2nd Person possessive affixes	—	—	—	—	+	—	+	—

[3] Long vowel variant is attested in Faendrich 2007: 124.

Pronouns

The most striking feature of pronominal declension is a tendency to merge ACC and DAT.LOC forms. This merger for 1st and 2nd person sg. pronouns (and partly for plural pronouns) is fully completed in Baoan. In Mongghul 1st person sg. pronoun fuses DAT.LOC and ACC forms across all varieties of the language; however, the 2nd person sg. merges ACC and DAT.LOC only in Naringhol and Tianzhu Mongghul. Mangghuer (despite Rybatzki 2003: 376) has different variants of singular pronouns, except for the 2nd sg., where the DAT.LOC enclitic *-du* might be optional (Slater 2003a: 83).

Table 7. ACC and DAT.LOC forms of 1st and 2nd person pronouns in Mangghuer, Mongghul and Baoan (Todaeva 1973: 72; Slater 2003a: 83; Chen and Chingeltei 1986: 390, 176; Faehndrich 2007: 118)

	Mangghuer		Mongghul		Baoan	
	1Sg	2Sg	1Sg	2Sg	1Sg	2Sg
ACC	namei, nangda, nami, dami, damei, dangda	čimi (Todaeva), qimei (Slater)	ndā (Halchighol, Karlong, Naringhol), dā (Tianzhu)	čimu (Halchighol) qimu (Karlong) čimī (Naringhol) qimī (Tianzhu)	mənda, məndə, nada, nadə	qinda, qada, qədа, qədə
DAT.LOC	nameidu, namidu, damidu, nangda(du)	čimidu (Todaeva) qimei(du) (Slater)		qimɨ (Karlong) čimī (Halchighol, Naringhol), qimī (Tianzhu)		

Loc. in *-re* for the 1st personal pronoun (*munire*) was recorded only in Naringhol by de Smedt and Mostaert (Todaeva 1973: 70), but is absent in other Mongghul varieties and Mangghuer.

Verbs

Common Mongolic had markers for reciprocal voice *-(*V*)*ldV*- and for cooperative voice *-(*V*)*lča*- (cf. Janhunen 2003: 11). In modern QG languages their reflexes are often confused. In Mangghuer, Halchighol, Naringhol, Karlong, Kangjia, Dongxiang, and Shira Yughur reflexes of the reciprocal voice function as both voices. In Baoan, on the contrary, the reflex of the cooperative voice marker *-či* has acquired the functions of both voices (Todaeva 1964: 74). Descriptions of Tianzhu Mongghul lack the reciprocal or cooperative voice (Faehndrich 2007: 189).

Table 8. Merger of Reciprocal & Cooperative voices in Qinghai-Gansu Mongolic (Todaeva 1973: 102–103; Faendrich 2007: 189; Tenishev, Todaeva 1966: 65; Todaeva 1961: 42; Sečenčogt 1999: 135)

	Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
Merger of Reciprocal & Cooperative voices	-rde	-rdi/-di	-lde/-de	-ldɨ	-lda/-lde	-či	-ndu	-ndʉ/-dʉ

Imperative forms

Of all the forms of imperative in QG languages it is the desiderative marker which demonstrates the greatest diversity.

The origin of the formant *lax-*/*laš-* in Mangghuer and Mongghul (Faehndrich 2007: 156, 185; Todaeva 1973: 105) is unclear. Slater does not mention *-lašgi* in his description (Slater 2003a: 117), but Todaeva cites examples with both affixes. Thus, this element might be an innovation at the Proto-Monguoric level, since it is not attested in other QG languages.

For their equational constructions Mangghuer and Dongxian have borrowed the Chinese copula (是 *shi*) (Slater 2003a: 238; Todaeva 1961: 143) which does not have any inflectional variants (Slater 2003b: 318).

Table 9. Distribution of the desiderative marker and the copula *shi* in Qinghai-Gansu languages.

	Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
Desiderative 3Sg	-gi, -lašgi	-ragi	-ge, -laxge	-laxgi	-gene, -ɣane	-gə	-gə	-gʉ
Chinese copula *shi*	+						+	

Tense / aspect forms

Many languages of the Amdo Sprachbund, including most Mongolic languages of the region, have developed the category of perspective (subjective/objective distinction), i.e. marking of speaker involvement into the action, which strongly correlates with controlled vs. non-controlled actions. They have adopted this category from Tibeto-Burmanese languages (cf. detailed analysis in Slater 2003a: 194–218; Slater 2018; Brosig and Skribnik 2018: 575–579). Thus, the whole system of Common Mongolic tense/aspect affixes in most QG languages (except Dongxiang) was drastically restructured in order to fit into this new category. In some cases new affixes originated from converb/participle + copula constructions (e.g. Mangghuer *-ni* < modal converb *-n* + copula); in other cases old affixes acquired new functions (e.g. Past tense affix *-ba* in Mangghuer became a Subjective past affix).

In Shira Yughur the prospective category is clearly seen in Past tense. Nugteren (2003: 279–280) describes it as a 1st vs 2nd-3rd person opposition, but actually it is most probably the same phenomenon as perspective. It is worth noting that Todaeva also initially considered the opposition of subjective vs prospective in Monguor and Baoan as a sign of emerging personal conjugation and opposition of 1st vs 2nd & 3rd person (Todaeva 1973: 110–115; Todaeva 1964: 89), although she already proposed the possible influence of corresponding Tibetan copular construction on this phenomenon (Todaeva 1964: 89).

The present tense in Shira Yughur can also indicate similar processes. There are two present tenses: *-nAi*, which is described as referring to involuntary mental and physical actions, vs *-jla wai ~ -dla wai* referring to "specific concrete actions which are being consciously performed in the present" (Nugteren 2003: 279). It is clear that the second form goes back to the combination of imperfective converb *-ji* with the copula, a combination generally used in Mongolic for progressive aspect, hence the meaning "concrete actions". However, the opposition of voluntary and involuntary actions exactly matches the opposition of non-volitional, out of control objective forms vs. volitional controlled subjective forms in Mangghuer, described in detail by Slater in (2003a: 194–220).

Mangghuer Objective Future affix *-kun(i)ang* (Slater 2003b: 316) obviously goes back to a combination of future participle with the copula. Interestingly, Todaeva (1973: 119) explicitly

states that her Minhe Mangghuer materials do not contain any examples of such combinations.

Mongghul future affix *-m* in Karlong refers to both future and present situations (Faehndrich 2007: 156–157).

Mangghuer present forms might be the only reflex of CM confirmative tense **-laγa* in QG languages.

In Table 10 below we present the system of tense and aspect forms in QG languages. The data are drawn from the following sources: Slater 2003a, 2003b for Mangghuer; Faehndrich 2007 for Karlong and Naringhol; Todaeva 1973 for Halchighol and Naringhol; Junast 1981 and Georg 2003 for Halchighol; Tenishev, Todaeva 1966, Nugteren 2003, and Bulchulu and Jalsan 1990 for Shira Yughur; Todaeva 1964, Chen and Chingeltei 1986, and Wu Hugjiltu 2003 for Baoan; Todaeva 1961, Buhe et al. 1985, and Kim 2003 for Dongxiang; Sečenčogt 1999, 2002 for Kangjia.

Negation and prohibitive particles

General distribution of prepositive indicative negation particles demonstrates reflexes of Common Mongolic negation particle **ülü* for non-past verb forms and reflexes of **ese* for past forms. However, this distribution can vary. The particle *ī* in Naringhol and Karlong may be an irregular variant of **ülü* > *lie* > *ī*. In Halchighol *lī* is used with present and future, rarely with past. The particle *se* is used exclusively before the past form (Todaeva 1973: 136). In Naringhol *ī* occurs rarely (Todaeva 1973: 136), Georg transcribes this marker as *yii* (Georg 2003: 303); aside from that, descriptions of Naringhol also mention the particles *li* (SM 1933: 222) and *se* (SM 1933: 341). Karlong *lī* is used with finite verbs, *ī* with finite and non-finite verbs (Faehndrich 2007: 217–219). Reflexes of **ese* are not attested in Karlong. Mangghuer *lai* is typically placed before imperfective verbs, *sai* before perfective ones (Slater 2003a: 146–147). Baoan (Todaeva 1964: 107), Dongxiang (Kim 2003: 362), and Kangjia (Sečenčogt 1999: 202–203) preserve the basic distribution. SY has only *li ~ l'* (< **ülü*) (Nugteren 2003: 283).

For Mangghuer Todaeva (1973: 107) gives *bu* as a prohibitive particle, but Slater (2003a: 147–148) instead transcribes the prohibitive as *bao* and states that in folktales there are two rare alternative forms *bai* and *bu*, the latter probably a Chinese loanword. Mongghul data are taken from (Todaeva 1973:107; Faehndrich 2007: 218–219). Dongxiang *bu* (Todaeva 1961: 58) and Baoan sGo.dmar dialect *bə* (Chen and Chingeltei 1986: 302) go back to CM **bu*, but reflexes of CM prohibitive particle **bitügei* were preserved only in Baoan *təgə* (Chen and Chingeltei 1986: 302), SY *püti (*Nugteren 2003: 283) and Kangjia *bʉde* (Sečenčogt 1999: 139–140).

Table 11. Negation and prohibitive particles in Qinghai-Gansu Mongolic

	Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
Negation particles	lai/sai	ī (yī)?/ li/se	lī/se	lī/ī	li ~ l'	lə/sə	ulie/ese	ne/se
Prohibitive particle < CM *bü	bu, bao	bī	bī	bī		(bə)	bu	
Prohibitive particle < CM *bitügei					püti	təgə		bʉde

Table 10. System of tense and aspect forms in Qinghai-Gansu languages

		Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
Category of perspective		+	+	+	+	+ (?)	+	-	+
Finite forms and participles in predicative position									
Future	neutral		-m	-m			-m, -nə	-nə	
	SUBJ	-ni	-guī, -guni	-gi, -guni	-gi, -guni		-gi		-gi/ -gʉni
	OBJ	-kun(i)ang	-guna	-guna	-guna				-gʉa/ -gʉna
Present/ Non-Past	neutral				-m			-jiwə < Conv. -ji + Copula	
	SUBJ	-la bi	-ni	-ni	-ni	-jla wai ~ -dla wai (voluntary action)	-ji		-ni, -ne; -sini
	OBJ	-lang	-na	-na	-na	-nai,-nii, -nAn, -nAmnA (involuntary action)	-jo		-na; -sina
Past	neutral		-wa	-wa	-wa		-o/-wo	-wə	-wa/-pa
	SUBJ	-ba (-jia)	-ji	-ji	-ji	-wa			
	OBJ	-jiang	-ja	-ja	-ja	-j'wai			
Past perfect							-je		-ʧa, -ʃʤa, -ʃʤi, -ʧʧa
Perfective	neutral	–saŋ				-(G)sAn		-sən	
	SUBJ		-dzani	-sani	-sani		-saŋŋi		-sʉni
	OBJ		-dzana	-sana	-sana		-saŋŋa		-sʉni -va

Converbs

Imperfective converb

Mongghul (all dialects) *-ji*, Mangghuer *-ji* (Todaeva 1973: 123–124), Dongxiang *-jɨ* (Todaeva 1961: 49), Baoan *-ji* (Todaeva 1964: 95), SY *-ji*/*-ci* (Nugteren 2003: 276), Kangjia *-jɔ*/*-ʤi* (Sečenčogt 1999: 163). Since all the forms are similar, we do not include them into the table. However, it is interesting that at least in Dongxiang, Shira Yughur, and Kangjia this converb can be used as a reduplicative form (Todaeva 1961: 49; Sečenčogt 1999: 163; Nugteren 2003: 276). Cf., e.g., Dongxiang *hela xolujɨ xolujɨ.* 'They run and run...' (Todaeva 1961: 49).

Perfective converb

In QG languages reflexes of two different Common Mongolic forms **γa* and **γad* are used in the function of a perfective converb.

In Mongghul dialects there are different perfective converbs: Halchighol: *-ā (-ē, -ō)*, *-ānu* (*-ēnu*,*-ōnu*), *-wā*, *-wānu*, Naringhol: *-wā*, *-wānu*/*-jā*, *-jānu*, Karlong: *-ā*, *-ēd* (Todaeva 1973: 124–126), Tianzhu *-wanə* (Faehndrich 2007: 190). Todaeva considers the Mangghuer forms *-danaŋ*, *-daGanaŋ*, *-daGali* as converb markers, while Slater treats *danang* as a separate word (the conjunction 'after') and quotes the form *-da* with the same meaning (Slater 2003a: 263–266). Baoan has no such form (Todaeva 1973: 127), while Dongxiang forms *-də*, *-dənə* (Todaeva 1961:49) are cognates of Mangghuer ones. Kim (2003: 360) calls these Dongxiang forms quasiconverbs. SY forms *-ā*, *-ē*, *-ād*, *-ēd* (Tenishev, Todaeva 1966: 71), *-Gādu*, *-ādu* (Bulchulu and Jalsan 1990: 278–279; Nugteren 2003: 277) obviously correspond to those in Karlong. Kangjia has *-da*, *-ʤida* (Sečenčogt 1999: 154). Closely related might be such special forms as Karlong *-di*, Tianzhu *-di* (Faehndrich 2007: 180–181), which Faehndrich calls "completive".

Modal converb

Mangghuer *-n*, Mongghul *-n* (in Karlong one can also add the singular marker to this converb: *-neŋge)* (Todaeva 1973: 129), Dongxiang *-n* (Todaeva 1961:49), Baoan *-ŋ*, Tongren Baoan *–jaŋ* (Todaeva 1964: 95). In SY the modal converb *-(V)n* usually occurs in reduplicative sequences indicating a "repeated action performed during the action of the main verb, e.g. *ci muni zaghalidi* ***xalda-n xalda-n*** *üleyaan bar* "do your work **while looking** at my portrait!" (Nugteren 2003: 276). Rybatzki (2003: 382) states that Mangghuer has lost this participle, judging by the absence of this converb in Slater's description (Slater 2003b: 315); however, Todaeva (1973: 129) gives undoubtable examples of the existence of this converb in Mangghuer: *Te gerdu* ***oron*** *teŋgiji saGaba* '**Entering** (converbum modale) the house he asked in such a way...'.

Consecutive converb

Monghhul *-gulā*, Mangghuer *-kula* (Todaeva 1973: 130), Kangjia *-gʉ*, *-gʉma* и *-gʉda* (Sečenčogt 1999: 162), Baoan *-guʤi*, *-guma* (Chen and Chingeltei 1986: 222–223). Xiazhuang Baoan and Dadun Baoan also have *-təxaŋ* (Chen and Chingeltei 1986: 222–223; Todaeva 1964: 95). Dongxiang and Shira Yughur lack this converb. In Karlong there is a single example in Faehndrich's data (Faehndrich 2007: 168).

Conditional converb

The conditional converb is present everywhere in QG languages and its affixes differ only phonetically. Halchighol *-sa*, Karlong *-sa*, Naringhol *-dza*, Mangghuer *-sa* (Todaeva 1973: 131–132), Dongxiang *-sə* (Todaeva 1961: 60), Baoan *-sa* (Xiazhuang, Dadun *-sə*; Ganhetan *-sɿ*) (Chen and Chingeltei 1986: 223–224), SY *-sa*, *-se* (Tenishev, Todaeva 1966: 71), Kangjia *-sa* (Sečenčogt

1999: 157–159). We do not include it into the table, but list these forms here, since they help to understand how the concessive converb is constructed.

Concessive converb

In QG languages the concessive converb can be formed on the basis of the conditional converb *-*sA* with additional affixes. Mongghul *-sada* (Todaeva 1973: 132), Shira Yughur *-sada* (Nugteren 2003: 277), Kangjia *-sala* (Sečenčogt 1999: 159–161), Dongxiang *-seda* (Todaeva 1961: 51), but *-se-nu* in (Kim 2003: 360), and Baoan *-sede* (Todaeva 1964: 97). According to Todaeva, Dongxiang *-se-nu* is functionally synonymous with the perfective converb, while according to Kim, *-se-nu* basically serves as a concessive converb, but sometimes can also be used in the meaning 'after' (Todaeva 1961: 51; Kim 2003: 360). In Mangghuer concessive functions are performed by conditional converb *-sa* (Slater 2003a: 255–256).

Terminative converb

The terminative converb is also present in all QG languages, but the variation is interesting. For Mongghul Todaeva gives Halchighol *-delā*, Naringhol *-delā*/*-derā*, Fulaan nura (= Karlong) *-delā*, *-delāŋge* (*-delā* + *ŋge* — singular marker), and *-daŋge* (Todaeva 1973: 134). However, according to Faehndrich, Karlong *-delā* is not accepted by speakers (Faehndrich 2007: 190), and she gives the form *-tala* instead (Faehndrich 2007: 170–171). Mangghuer *-tula*/*-tala* (Slater 2003a: 254–255), Baoan *-tala*, *-tələ*, *-təla*, *-dala* (gNyan.thog Baoan *-tala* has variants *-sala*, *-la*, Xiazhuang Baoan has *-sələ*) (Chen and Chingeltei 1986: 226–227), Dongxiang *-tala* (Todaeva 1961: 51), SY *-tAlA* (Tenishev, Todaeva 1966: 71), Kangjia *-tala* (Sečenčogt 1999: 161). Actually, Baoan variants *-sala*, *-la*, *-sələ*, as well as Karlong *-daŋge* might indicate that the converb affix **tala* was originally a composite affix, and in descendant languages we witness the reflection of some old composite variants.

Abtemporal converb

Halchighol *-sār*, Mangghuer *-sar* (Todaeva 1973: 134), Mangghuer *-ser* (Slater 2003b: 315). According to Faehndrich, (2007: 190) *-sār* is not accepted by the speakers of Karlong, and *-n* with reduplication is used instead. SY has *-sār*/*-gsār* (Bulchulu and Jalsan 1990: 282–283). Todaeva shows that Qinghai Baoan *-ser* is used without any restrictions (Todaeva 1964: 95), but Chen and Chingeltei (1986: 230–231) report that *-sər* is used only in Gansu Baoan and Xiazhuang subdialect of Qinghai Baoan exclusively within the construction *-sər* + verb *sū-* 'live, sit' (Chen and Chingeltei 1986: 230–231). This form is absent in Dongxiang and Kangjia.

Zero-marked converb

Rybatzki mentions a zero-marked serial converb as an important innovation confined to Gansu-Qinghai "attested at least in Shira Yughur, Mangghuer, Bonan, and Santa" (Rybatzki 2003: 383). Todaeva (1973: 127–128) also cites Mongghul examples, and Sečenčogt gives examples from Kangjia (Sečenčogt 1999: 164). Thus, this isogloss is valid for all QG languages.

Numerals

There are several patterns for deriving ordinal numbers from numeral stems: prefixes (Baoan *aŋ-*, *aŋgə-* from Tibetan, see Chen and Chingeltei 1986: 157–158; Mangghuer *di-*, see Todaeva 1973: 91, Dongxiang *ǯi-*, see Todaeva 1961: 36–37, Kangjia *ǯi-*, *di-*, see Sečenčogt 2002: 69; all three from Chinese 第 *dì-*), or affixes (Mongghul *-dar*, see Todaeva 1973: 91, Dongxiang *-da*, see Todaeva 1961: 36–37 < Common Mongolic **-duγar*; SY *-cAAr* or *-rjAr*, see Nugteren 2003: 271–272 of unclear origin).

Table 12. Converbs in Qinghai-Gansu Mongolic

	Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
Reduplication of imperfective converb	—	—	—	—	+	—	+	+
Perfective converb	-danaŋ, -daGanaŋ, -daGali	-wā, -wānu / -jā, -jānu	-ā(-ē, -ō), -ānu (-ēnu, -ōnu), -wā, -wānu	-ā, -ēd; completive -di	-ā, -ē, -ād, -ēd / -Gādu, -ādu	—	-də, -dənə	-da, -ʤida
Modal converb	-n	-n	-n	-n	-n	-ŋ,-jaŋ	—	
Consecutive converb	-kula	-gula	-gulā	-gula	—	-guʤi, -guma	—	-gʉ, -gʉma, -gʉda
Consessive converb	—	-sada	-sada	-sada	-sada	-sede	-seda	-sala
Terminative converb	-tala/-tula	-delā/-derā	-delā	-delā, -delāŋge -daŋge (?), -tala	-tAlA	tala, -tələ, -təla, -dala, -sala, -la, -sələ	-tala	-tala
Abtemporal converb	-sar/-ser	—	sār	—	-sar/-gsar	-sər	—	—
Adding singulative marker to converbs	—	—	—	+	—	—	—	—

Several Qinghai-Gansu languages have borrowed Chinese numeral classifiers (Slater 2003a: 95–96; Faehndrich 2007: 135–136; Sečenčogt 2002: 70).

Table 13. Qinghai-Gansu innovations in numerals

	Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
Numeral prefixes (ordinal numbers)	di-	-dar	-dar	-dar	-cAAr/ -rjAr	aŋ-, aŋgə-	ǯi — -da	ǯi -
Numeral classifiers	+			+			+	+

Syntactic patterns

Causee marking in causative constructions varies: Halchighol uses ACC (for originally intransitive stems) and DAT.LOC (for originally transitive stems) (Todaeva 1973: 100–101); Karlong allows only Accusative in both cases (Faehndrich 2007: 182). In Naringhol the data are scarce, but there are at least examples of DAT.LOC marking for the causee (SM 1964: 20).

Mangghuer, according to Todaeva's records (Todaeva 1973: 100–101), only marks causee with ACC, and she underlines that this feature differentiates Huzhu Mongghul from Minhe Mangghuer, but Slater states that in Mangghuer causee receives accusative (for originally intransitive verbs) and dative (for originally transitive verbs) (Slater 2003a: 130). Kangjia (Sečenčogt 1999: 134), Dongxiang (Todaeva 1961: 41–42) and Baoan (Todaeva 1964: 72–73) also allow both dative and accusative marking. Shira Yughur demonstrates an alternative strategy with INS and DAT.LOC cases (Tenishev, Todaeva 1966: 64), which is more similar to Central Mongolic systems (see e.g. ACC, DAT.LOC and INS in Written Mongolian, Khalkha and Buriad, ACC and INS in Kalmyk).

Table 14. Causee marking in Qinghai-Gansu Mongolic

	Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
Causee marking	ACC/ DAT.LOC	ACC (?)/ DAT.LOC	ACC/ DAT.LOC	ACC	DAT.LOC/ INS	ACC/ DAT.LOC	ACC/ DAT.LOC	ACC/ DAT.LOC

Lexicon

A comparative analysis of the lexicon of QG languages, focusing on the percentage of borrowed and Common Mongolic etyma, is provided in Rybatzki (2003: 385–386). In this work V. Rybatzki claims that Mangghuer has only 39% of the native lexicon and thus "it is obvious that Mangghuer has suffered a massive loss of native vocabulary, making it, at least lexically, a good candidate for a 'mixed language'". Bonan (=Baoan) in his data demonstrates 50% of native vocabulary, Santa (i.e. Dongxian) has 56%, Monggul has 72%, and Shira Yughur has 77%.

Hattori (1959) and Kuzmenkov (1993) in their glottohronological studies on Mongolic lexicon considered Monguor to be a single idiom and did not use data from any other QG languages.

In our previous analysis (Gruntov, Mazo 2015) we scrutinized data for 110 basic lexical items from all Mongolic languages. It turned out that Mangghuer has 98% of common lexicon with Mongghul. Later we also analyzed the Leipzig-Jakarta list of 250 basic words for QG lan-

guages (unpublished) and found that 81% of the lexicon (76% if we eliminate all the borrowings) are still in common for Mangghuer and Mongghul. Thus, Mangghuer is lexically closer to Mongghul than to any other language, which is not surprising.

Shared innovations

We can summarize the previous data in the following table, showing the unique and shared innovations of Mangghuer, Halchighol, Naringhol and Karlong compared to other QG languages.

Table 15. Monguor innovations in Qinghai-Gansu languages

№	Innovation	Languages	Number of languages	Comment
1	*-kV* for the borrowed stems	Mgr, MN, MH, MK, SY, B, D	7	
2	reflex of Reciproc functions as both Reciproc and Cooperative voices	Mgr, MN, MH, MK, SY, D, K	7	
3	perspective	Mgr, MN, MH, MK, SY, B, K	7	
4	initial clusters allowed	Mgr, MN, MH, MK, SY, B, K	7	
5	**b > v*	MN, MH, MK, SY, B, D, K	7	
6	*-lV* for the borrowed stems	Mgr, MN, MH, MK, SY, D	6	
7	loss of resultative	Mgr, MK, SY, B, D, K	6	
8	indefinite article	Mgr, MN, MH, MK, B, K	6	
9	causee ACC/DAT.LOC	Mgr, MN, MH, B, D, K	6	
10	loss of 1st and 2nd poss.affixes	Mgr, MN, MH, MK, B, K	6	
11	**q > x*	MN, MH, MK, SY, B, K	6	
12	**hU > f*	MN, MH, MK, B, D, K	6	
13	concessive converb *-sVdV*	MN, MH, MK, SY, B, D	6	
14	**U > ɔ* in accented syllables	Mgr, MN, MH, MK, SY	5	
15	loss of prohibitive particle **bitügei*	Mgr, MN, MH, MK, D	5	
16	**-rb- > -r-*	MN, MH, MK, D, K	5	
17	**-d > r*	MN, MH, MK, B, K	5	
18	secondary long vowels	MN, MH, MK, SY, B	5	
19	**-m > -N*	Mgr, MK, D, K	4	
20	loss of unstable *-n* in Nom	Mgr, MN, MH, MK	4	
21	desiderative *-laCgi*	Mgr, MN, MH, MK	4	
22	adtemporal *-sVr*	Mgr, MH, SY, B	4	
23	numeral classifiers	Mgr, MK, D, K	4	
24	LOC *-rV*	MN, MH, MK, D	4	
25	merge of ACC/DAT 1Sg	MN, MH, MK, B	4	
26	perfective converb in *-VV*	MN, MH, MK, SY	4	
27	pluralis *-la*	MK, B, D, K	4	
28	**l > r*	Mgr, MN, K	3	
29	secondary vowel shortening	Mgr, D, K, B	3	Not all Baoan dialects demonstrate this secondary shortening

№	Innovation	Languages	Number of languages	Comment
30	**i> [ɿ] and [ɻ]*	Mgr, D, K	3	
31	merge of ACC/DAT 2Sg	Mgr, MN, B	3	Only partly in Mangghuer
32	future *-nV*	Mgr, B, D	3	
33	Chinese ordinal number prefix	Mgr, D, K	3	Baoan borrowed ordinal number prefix from Tibetan
34	pluralis *-sGV*	MN, MH, MK	3	
35	future *–m*	MN, MH, B	3	
36	DAT.LOC *-di*	MH, MK, SY	3	
37	**-d > 0*	Mgr, B	2	
38	*-rV* for the borrowed stems	Mgr, D	2	
39	DIR *-dʐi*	Mgr, MN	2	
40	copula *shi*	Mgr, D	2	
41	negative particle *ī*	MN, MK	2	
42	pluralis *-ŋgu/ -ŋgula*	MH, MK	2	
43	assotiative plurality *-maŋge*	MH, MK	2	
44	tonal opposition	Mgr	1	Baoan also has tonal opposition, but tones in Baoan and Mangghuer emerged independently, and it cannot be considered a shared innovation
45	**-r > ɚ˞*	Mgr	1	
46	**-bC- > G*	Mgr	1	
47	splitting of **ǰ /* č* reflexes	Mgr	1	
48	*-tu* for the borrowed stems	Mgr	1	
49	comparative *-her*	Mgr	1	
50	non-past *-la*-Copula	Mgr	1	
51	tendency to increase number of clusters	MK	1	
52	tendency to decrease number of clusters	Mgr	1	
53	pronominal LOC *-re*	MN	1	
54	intensifier	MK	1	
55	Dualis	MK	1	
56	completive *-di*	MK	1	
57	singulative marker with converbs	MK	1	
58	causee ACC	MK	1	

Conclusions

These results can be quantified as follows:

Table 16. Number of shared and unique innovations in Monguoric languages.

	Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
Mangghuer	34	15	13	14	8	12	14	13
Naringhol		29	24	23	11	15	10	11
Halchighol			28	25	13	15	10	11
Karlong				36	13	14	13	14

If we count grammatical and phonological innovations separately, we get the following distribution:

Table 17. Distribution of phonological innovations

	Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
Mangghuer	12	3	2	3	2	3	3	5
Naringhol		9	8	8	5	6	3	6
Halchighol			9	9	6	6	3	6
Karlong				11	6	6	4	7

Table 18. Distribution of grammatical innovations

	Mangghuer	Naringhol	Halchighol	Karlong	SY	Baoan	Dongxiang	Kangjia
Mangghuer	22	12	11	11	6	9	11	8
Naringhol		20	16	15	6	9	7	5
Halchighol			19	16	7	9	7	5
Karlong				25	7	8	9	7

We used 58 isoglosses (20 phonetical and 38 grammatical) to study the differences between Monguoric idioms. The data show that the most innovative idiom is Karlong with 36 innovations, followed by Mangghuer (34 innovations). In addition, Karlong is equally distant from other QG non-Mongghul idioms, Mangghuer included. Mangghuer and Karlong are the most innovative idioms in respect to both phonology and grammar. However, as far as phonology is concerned, Mangghuer has developed many unique distinct features whereas phonetical change in Karlong is generally within the areal Tibetan-oriented phonology trend: Karlong has much fewer shared innovations in phonology with Sinitic-oriented Dongxiang and Mangghuer than with other QG languages.

As expected, Mongghul dialects demonstrate a high level of shared innovations (23–25 innovations), and they can hardly be considered separate languages. Mangghuer has more shared innovations with Dongxiang and Kangjia than with SY. Most probably, this is the result of significant Chinese influence on these three languages. On the contrary, the number of innovations shared between Mongghul and Shira Yughur is noticeably larger than the respective number between Mangghuer and Shira Yughur. However, phonological innovations account for most of the difference. Grammatically, Mangghuer is almost at the same distance from SY as Mongghul idioms, but its phonology is much more siniticized. It might be interest-

ing that the number of Mangghuer — Mongghul shared innovations is not as high (compared to Mangghuer — Dongxiang — Kangjia shared innovations) as one might have expected. However, shared innovations in grammar confirm that Mangghuer is still the closest relative of Mongghul.

Abbreviations

ABL — Ablative; ACC — Accusative; B — Baoan; CM — Common Mongolic; D — Dongxiang; DAT.LOC — Dative-locative; DIR — Directive; GEN — Genitive; IMPERF — Imperfective; INSTR — Instrumental; K — Kangjia; LOC — Locative; Mgr — Mangghuer MK — Karlong; MH — Halchighol; MN — Naringhol; NOM — Nominative; OBJ — Objective; PROL — Prolative; QG — Qinghai-Gansu; SG — Singular; SM — de Smedt and Mostaert; SY — Shira Yughur; SUB — Subjective.

References

Binnick, Robert I. 1987. On the Classification of the Mongolian Languages. *Central Asiatic Journal* 31: 178–95.

Brosig, Benjamin, Elena Skribnik. 2018. Evidentiality in Mongolic. In: Alexandra Aikhenvald (ed.). *Oxford Handbook of Evidentiality*: 554–579. Oxford University Press.

Bùhé (布和). 1985. *Dōngxiāngyǔ hé Měnggǔyǔ*. Hohhot: Nèiměnggǔ rénmín chūbǎnshè (内蒙古人民出版社).

Bulchulu (保朝鲁**)** and Jalsan **(**贾拉森**)**. 1990. *Dōngbù Yùgùyǔ hé Měnggǔyǔ*. Hohhot: Nèiměnggǔ rénmín chūbǎnshè (内蒙古人民出版社).

Chén Nǎixióng (陈乃雄) and Chingeltei (清格尔泰). 1986. *Bǎo'ānyǔ hé Měnggǔyǔ*. Hohhot: Nèiměnggǔ rénmín chūbǎnshè (内蒙古人民出版社).

Chingeltei (清格尔泰). 1986. *Tǔzúyǔ Huàyǔ Cáiliào*. Hohhot: Nèiměnggǔ rénmín chūbǎnshè (内蒙古人民出版社).

Chingeltei (清格尔泰) and Lǐ Kèyù (李克郁). 1988. *Tǔzúyǔ hé Měnggǔyǔ*. Hohhot: Nèiměnggǔ rénmín chūbǎnshè (内蒙古人民出版社).

Chingeltei (清格尔泰). 1989. On Subjective and Objective Mood in the Monguor Language. In: Klaus Sagaster (ed.). *Gedanke und Wirkung: Festschrift für Nikolaus Poppe zum 90*: 67–75. Geburtstag, Wiesbaden: Otto Harrassowitz.

Chuluu, Üjiyediin. 1994. *Introduction, Grammar, and Sample Sentences for Monguor* [= Sino-Platonic Papers 57]. Philadelphia: University of Pennsylvania, Department of Asian and Middle Eastern Studies.

Cuī Yǒnghóng (崔永红), Zhāng Dézǔ (张得祖), Dū Chángshùn (杜常顺). 2015. The origin of Monguor. In: Gerald Roche and CK Stuart (eds). *Mapping the Monguor*. *Asian Highlands Perspectives 36*: 18–22.

Doerfer, Gerhard. 1964. Klassifikation und Verbreitung der mongolischen Sprachen. In: Bertold Spuler (ed.). *Mongolistik (Handbuch der Orientalistik 5.2)*: 35–50. Leiden: Brill.

Dpal-ldan-bkra-shis, Hu Jun, Hu Ping, Limusishiden, Keith Slater, Kevin Stuart, Wang Xianzheng, Zhu Yongzhong. 1996. *Language Materials of China's Monguor Minority: Huzhu Mongghul and Minhe Mangghuer* [= Sino-Platonic Papers 69]. Philadelphia: University of Pennsylvania.

Dwyer, Arienne M. 1992. Altaic Elements in the Línxià Dialect: Contact-induced Change on the Yellow River Plateau. *Journal of Chinese Linguistics* 20: 160–179.

Dwyer, Arienne M. 1995. From the Northwest China Sprachbund: Xúnhuà Chinese Dialect Data. In: *Yuen Ren Society Treasury of Chinese Dialect Data*, vol. I: 143–82.

Dwyer, Arienne M. 2008. Tonogenesis in Southeastern Monguor. In: Harrison, K. David, David Rood, Arienne Dwyer (eds.). *Lessons from documented endangered languages. Typological Studies in Language* 78: 111–128. Amsterdam: John Benjamins.

Faehndrich, Bergel Maria. 2007. *Sketch Grammar of the Karlong Variety of Mongghul, and Dialectal Survey of Mongghul.* PhD thesis. Honolulu: University of Hawai'i.

Georg, Stefan. 2003. Mongghul. In: Juha Janhunen (ed.).*The Mongolic languages*: 286–306. London / New York; Routledge.

Gruntov, Ilya, Olga Mazo. 2015. Klassifikacija mongolskix jazykov po lexicostatisticheskim dannym. *Journal of Language Relationship* 13/3: 205–255.

Gruntov, Ilya, Olga Mazo. 2020. Nominal Morphology: Case and Plurality. In: Martine Robbeets (ed.). *The Oxford Guide to the Transeurasian Languages*. Oxford University Press (forthc.).

Hasbaatar (哈斯巴特尔). 1985. *Tǔzúyǔ Cíhuì*. Hohhot: Nèiměnggǔ rénmín chūbǎnshè (内蒙古人民出版社).

Hattori, Shiro. 1959 [1993]. A Glottochronological Study of the Mongol Languages. Reprinted in: Selected Papers of Shiro Hattori, vol. 4: *Studies in Altaic Languages*: 87–8. Tokyo: Sanseido.

Huc, Évariste-Régis. 1850. *Souvenirs d'un voyage dans la Tartarie, le Thibet et la Chine pendant les années 1844, 1845 et 1846*, vols 1–2. Paris: Librairie d'Adrien le Clere et Cie.

Janhunen, Juha. 2003. Proto-Mongolic. In: Juha Janhunen (ed.).*The Mongolic languages*: 1–29. London / New York: Routledge.

Janhunen, Juha. 2006. The Monguor: The Emerging Diversity of a Vanishing People. In: Louis Schram (Kevin Stuart, ed.). *The Monguors of the Kansu Tibetan Frontier [1954, 1957, 1961] Part I: Their Origin, History and Social Organization; Part II: Their Religious Life; Part III: Records of the Monguor Clans: History of the Monguors in Huanchung and the Chronicles of the Lu Family*: 26–29. Xining City: Plateau Publications.

Janhunen Juha, Lionel Ha Mingzong, Joseph Tshe dPag rNam rGyal. 2007. On the Language of the Shaowa Tuzu in the Context of the Ethnic Taxonomy of Amdo Qinghai. *Central Asiatic Journal* 51(2): 177–195.

Janhunen, Juha. 2007. Typological Interaction in the Qinghai Linguistic Complex. *Studia Orientalia* 101: 85–103.

Janhunen, Juha. 2012. On the Hierarchy of Structural Convergence in the Amdo Sprachbund. In: Pirkko Suihkonen, Bernard Comrie, Valery Solovyev (eds.). *Argument Structure and Grammatical Relations*: 177–189. Studies in Language Companion Series 126. Amsterdam: John Benjamins.

Junast (照那斯图). 1981. *Tǔzúyǔ Jiǎnzhì.* Běijīng: Mínzú chūbǎnshè (民族出版社).

Junast (照那斯图), Lǐ Kèyù (李克郁). 1982. Tǔzúyǔ Mínhé fāngyán gàishù. In: *Mínzú Yǔwén Yánjiū Wénjí*: 458–487. Xīníng. Qīnghǎi rénmín chūbǎnshè (青海人民出版社).

Li, Charles N. 1986. The Rise and Fall of Tones Through Diffusion. In: *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*: 173–185.

Kakudo, Masayoshi. 1987. Tozokugo no Ka'i Hougen. *Oosaka Gaikokugo Daigaku Gakuhou* 75: 49–63.

Kakudo, Masayoshi. 1997. Tianzhu Tozokugo no Tokuchou: "Geseru Bunko" Daisanken no Shiryou ni Motozuite. *Oosaka Gaikokugo Daigaku Ronshuu* 17: 33–61.

Khelimsky, Evgenii A. 1984. A distinctive feature that became a phoneme: The case of Monguor. In: *5th International Phonology meeting. Abstracts*: 27. Wien.

Kim, Stephen S. 2003. Santa. In: Juha Janhunen (ed.). *The Mongolic languages*: 346–363. London / New York: Routledge.

Kuzmenkov, Evgenii A. 1993. Mongolskaja bazisnaja leksika v "Sokrovennom skazanii". In: *Mongolica*: 328–340. Moskva: Nauka.

Lǐ Kèyù (李克郁). 1993. *Tǔzú Měnggǔěr Yuánliú Kǎo*. Xīníng: Qīnghǎi rénmín chūbǎnshè (青海人民出版社).

Lǐ Kèyù (李克郁). 1988. *Tǔ Hàn Cídiǎn*. Xīníng: Qīnghǎi rénmín chūbǎnshè (青海人民出版社).

Lǐ Kèyù (李克郁). 2008. *Tǔzú Lìshǐ yǔ Yǔyán Wénzì Yánjiū Wénjí.* Běijīng: Mínzú chūbǎnshè (民族出版社).

Luwsanwandan, Shadawyn. 1959. Mongol xel ayalguunï ucir. *Studia Mongolica* 1: 1–16.

Lǚ Jiànfú (吕建福). 2002. *Tǔzúshǐ*. Běijīng: Zhōngguó shèhuì kēxué chūbǎnshè (中國社會科學出版社).

Nugteren, Hans. 1997. On the Classification of the "Peripheral Mongolic Languages". In: Árpád Berta (ed.). *Historical and Linguistic Interaction Between Inner-Asia and Europe* [= Studia Uralo-Altaica 39]: 207–16. Szeged.

Nugteren, Hans. 2003. Shira Yughur. In: Juha Janhunen (ed.). *The Mongolic languages*: 265–285. London / New York: Routledge.

Nugteren, Hans. 2011. *Mongolic Phonology and the Qinghai-Gansu Languages*. Utrecht: Lot 289.

Nugteren, Hans. 2013. Postposed indefinite articles in Mongolic and Turkic languages of the Qinghai-Gansu Sprachbund. In: Martine Robbeets, Hubert Cuyckens (eds.). *Shared Grammaticalization. With special focus on the Transeurasian languages*: 227–250. Amsterdam/Philadelphia: John Benjamins.

Nugteren. Hans. 2014. Amdo Altaic directives and comparatives based on the verb 'to see'. In: Martine Robbeets. Walter Bisang (eds). *Paradigm change in the Transeurasian languages and beyond*: 243–256. Amsterdam/Philadelphia: John Benjamins.

Poppe, Nicholas. 1955. *Introduction to Mongolian Comparative Studies* [= Mémoires de la Société Finno-Ougrienne 110]. Helsinki.

Poston, Dudley L., Qian Xiong. 2014. Are China's Minority Nationalities Still on the Margins? In: Isabelle Attané and Baochang Gu (eds.). *Analyzing China's Population: Social Change in a New Demographic Era*: 113–138. New York: Springer.

Potanin, Grigorii N. 1893. *Tangutsko-tibetskaya okraïna Kitaya i central'naya Mongoliya*, vols. 1–2. Sankt-Peterburg: Izdanie Imperatorskago Russkago Geograficheskago Obshchestva.

Przheval'skii, Nikolai. 1875. *Mongolija i strana tangutov: trekhletnee puteshestvie v vostochnoi nagornoi Aziï.* Sankt-Peterburg: Izdanie Imperatorskago Russkago Geograficheskago Obshhestva.

Róna-Tas, András. 1960. Remarks on the Phonology of the Mong[uo]r Language. *Acta Orientalia Hungarica* 10: 263–7.
Róna-Tas, András. 1962. On some Finals of the Mong[uo]r Language. *Acta Orientalia Hungarica* 14: 283–90.
Róna-Tas, András. 1966. *Tibeto-Mongolica: The Tibetan Loanwords of Monguor and the Development of the Archaic Tibetan Dialects* [= Indo-Iranian Monographs 7]. The Hague: Mouton.
Rybatzki, Volker. 2003. Intra-Mongolic Taxonomy. In: Juha Janhunen (ed.). *The Mongolic languages*: 364–390. London / New York: Routledge.
Sanzheev, Garma D. 1952. Mongol'skie jazyki i dialekty. *Uchenie zapiski Instituta Vostokovedeniya* 4: 30–125.
Schram, Louis M. J. 1954–61. *The Monguors of the Kansu-Tibetan Frontier: Their Origin, History, and Social Organization*, vols 1–3 [= Transactions of the American Philosophical Society, New Series 44/1, 47/1, 51/3]. Philadelphia.
Schröder, Dominik. 1952–3. Zur Religion der Tujen des Sininggebietes (Kukunor). *Anthropos* 47: 1–79, 620–58, 822–70; 48: 202–59.
Schröder, Dominik. 1959–70. *Aus der Volksdichtung der Monguor, 1. Teil: Das weisse Glücksschaf (Mythen, Märchen, Lieder), 2. Teil: In den Tagen der Urzeit (Ein Mythus vom Licht und vom Leben)* [= Asiatische Forschungen 6, 31]. Wiesbaden: Otto Harrassowitz.
Schröder, Dominik. 1964. Der Dialekt der Monguor. *Mongolistik* [= Handbuch der Orientalistik I: V, 2]: 143–58.
Sečenčogt (斯钦朝克图). 1999. *Kāngjiāyǔ yánjiū*. Shànghǎi: Yuǎndōng chūbǎnshè (远东出版社).
Sečenčogt (斯钦朝克图). 2002. Kāngjiāyǔ gàikuàng. *Mínzú yǔwén* 6: 66–77.
Sinor, Denis. 1952. Langues mongoles. In: Antoine Meillet, Marcel Cohen (eds.). *Les langues du monde* (2nd edition): 369–84. Paris: Champion.
Slater, Keith W. 2003a. *A Grammar of Mangghuer: A Mongolic Language of China's Qinghai-Gansu Sprachbund* [= Curzon Asian Linguistic Series 2]. London: Curzon.
Slater, Keith W. 2003b. Mangghuer. In: Juha Janhunen (ed.). *The Mongolic languages*: 307–324. London / New York: Routledge.
Slater, Keith W. 2018. Morphological innovations in Mangghuer and Shirongolic: Reconstructing the formal emergence of the subjective vs. objective distinction. In: Simeon Floyd, Elizabeth Norcliff, Lila San Roque (eds.). *Egophoricity* [Typological Studies in Language, 118]: 225–267. Amsterdam: John Benjamins.
de Smedt, Albrecht, Antoine Mostaert. 1929. Le dialecte monguor parlé par les Mongols du Kansou occidental, Ière partie: Phonétique. *Anthropos* 24: 145–66.
de Smedt, Albrecht , Antoine Mostaert. 1933. *Le dialecte monguor parlé par les Mongols du Kansou occidental, IIIe partie: Dictionnaire monguor-français.* Pei-p'ing: Imprimerie de l'Université Catholique.
de Smedt, Albrecht, Antoine Mostaert. 1945. *Le Dialecte Monguor. IIe partie: Grammaire*. Monumenta Serica 6, Journal of Oriental Studies. Peking: Catholic University.
Tenishev, Edgem R., Buljash X. Todaeva. 1966. *Yazyk zheltyx uigurov*. Moskva: Nauka (Institut narodov Azii AN SSSR).
Todaeva, Buljash X. 1961. *Dunsjanskij jazyk*. Moskva: Institut narodov Azii AN SSSR.
Todaeva, Buljash X. 1964. *Baoan'skij jazyk*. Moskva: Institut narodov Azii AN SSSR.
Todaeva, Buljash X. 1973. *Mongorskij jazyk: issledovanie, teksty, slovar'*. Moskva: Institut narodov Azii AN SSSR.
Wu, Hugjiltu. 2003. Bonan. In: Juha Janhunen (ed.). *The Mongolic languages*: 325–345. London / New York: Routledge.
Xu, Dan, Wen Shaoqing. 2017. The Silk Road: language and population admixture and replacement. In: Dan Xu, Hui Li (eds.). *Languages and Genes in Northwestern China and Adjacent Regions*: 55–78. Singapore: Springer Nature.
Zhang, Zhenxing (ed.). 2012. *Language Atlas of China. Zhōngguó Yǔyán Dìtújí. Shǎoshù Mínzú Yǔyán Juǎn*. 2 ed. Shànghǎi: Shāngwù yìnshūguǎn (商务印书馆).

И. А. Грунтов, О. М. Мазо. Хуцзу, минхэ и их соседи: оценка степени языковой близости

Целью настоящей работы является оценка степени языковой близости различных монгольских идиомов в рамках языкового союза Цинхай-Ганьсу, при этом основное внимание уделяется языку минхэ и диалектам хуцзу. На основании 58 грамматических и фонологических совместных инноваций в области фонологии и грамматики авторы пытаются проанализировать основные направления и модели языковых изменений на территории региона.

Ключевые слова: языковой союз Циньхай-Ганьсу; языковые контакты; монгорские языки; минхэ; хуцзу; монгольские языки; совместные инновации; широнгольские языки.

Bartomeu Obrador-Cursach
Institut del Pròxim Orient Antic, Universitat de Barcelona; obradorcursach@gmail.com

On the place of Phrygian among the Indo-European languages

The aim of this paper is to gather together certain relevant features of Phrygian based on our current knowledge of the language in order to determine its dialectal position inside the Indo-European family. The relatively large number of features shared with Greek is consistent with prior views about the close relation between the two languages, which may have formed a common proto-language. The relations proposed with certain other languages, such as Armenian, are not so strong despite sharing some features.

Keywords: Phrygian language; Indo-European dialectology; linguistic subgrouping; isoglosses; Proto-Greek language.

1. Introduction[1]

Over the last three decades our knowledge of the Phrygian language has increased immensely, especially in regard to historical linguistics. In the light of this new information, it is worth reconsidering previous theories about its genetical position within the Indo-European family. Because of the fragmentary nature of the vestiges of the language[2] the data available are very limited, and it must be added that many inscriptions (interestingly, some of the largest)[3] remain mainly undeciphered. Therefore, any conclusions presented in this paper must be considered provisional and, of course, subject to new findings and interpretations.

From a methodological point of view, direct sources (inscriptions) are the base of this work,[4] although indirect sources (glosses and toponyms) are occasionally considered. I exclude isoglosses that are extended throughout the Indo-European family, because here I will focus on the most restrictive isoglosses, highlighting the shared innovations. Needless to say, I operate with the clearest forms. As is the case with fragmentary languages, most of the current interpretations of the texts are mere possibilities that need to be confirmed by future findings.

---

[1] This is an extended version of a chapter devoted to the position of the Phrygian language within the Indo-European by Obrador-Cursach (2020: 121–127). A first draft of this paper was presented to the 6th "'Luwic' Dialects: Inheritance and diffusion" Workshop, held in Barcelona (March 28, 2019). I must thank the workshop participants for their useful comments. I am also grateful to Hrach K. Martirosyan for his remarks mainly on Armenian. Needless to say, any errors are my own responsibility.

[2] Direct sources for Phrygian are inscriptions found in central Anatolia. These inscriptions are gathered in two main corpora. Old Phrygian (OPhr., also known as Paleo-Phrygian) is the label for the inscriptions dated to between 800 and 330 BC and written in an epichoric alphabet closely related to the Greek one. The inscriptions here are quoted according to *CIPPh* and its supplements, i.e., an initial letter derived from the name of the area where the inscription was found with a number, and transcribed in Latin characters. The corpus called New Phrygian (NPhr., also Neo-Phrygian) comprises 118 inscriptions dated to the Imperial Roman period, which are written in the Greek alphabet and mainly contain cursing formulas against desecrators. They are quoted with the enumeration given by Obrador-Cursach 2020 followed by the traditional *numerus currens*. From Hellenistic times there is an important inscription, W-11 (also written in the Greek script), which scholars often consider a third stage of the language, called Middle Phrygian (MPhr.).

[3] Old Phrygian B-01, B-04 and B-08, Middle Phrygian W-11 New Phrygian 11.2 (18), 16.1 (116) and 43.1 (69).

[4] The treatment of each form is detailed in Obrador-Cursach 2020.

Consequently, arbitrary speculations are left aside. In addition, I should stress that I operate with the devoicing and deaspiration defended by Lubotsky (2004), known as Phrygian "Lautverschiebung".[5] Although some scholars have tried to nuance it by considering contextual shifts (e.g., Woodhouse 2006 and de Lamberterie 2013), it has the advantage of being the simplest explanation and, even more significantly, many Phrygian words have been identified and explained since the publication of Lubotsky's analysis. Finally, despite the difference with regard to the amount of data available, Martirosyan's paper on the place of Armenian (2013) was an excellent model for my purposes.

2. Compilation of isoglosses

2.1. Phonological isoglosses

2.1.1. In the attested positions, Phrygian shows the same reflexes as the old PIE laryngeals found in Greek, including the so-called "prothetic vowel": OPhr. *ev-* 'well' < PIE **h₁su-* (cf. Gk. εὐ-); NPhr. αναϱ 'husband' < PIE **h₂nēr* (cf. ἀνήϱ); OPhr. *onoman* 'name' < PIE **h₃nh₃-mn* (Ligorio and Lubotsky 2018: 1823). This feature is also shared with Armenian (Martirosyan 2010: 714–716, also 2013: 89).

2.1.2. Phrygian γλουϱεος 'golden' (MPhr-01 = W-11) shows that the Greek *CRh₂C > *CRōC (cf. χλωϱός 'id.') affected Phrygian as well.

2.1.3. Phrygian, as well as Greek, Armenian and Albanian (in some positions, see Demiraj 1997: 56), has lost the sibilant /s/ at the beginning of the word and in intervocalic position: OPhr. *egeseti*, NPhr. εγεσιτ, εγεδου 'hold, experience' < PIE **segʰ-*; NPhr. δεως 'god' < PIE **dʰh₁so-*. Note, however, that this feature is also shared, e.g., with Lycian (Melchert 1994: 288), Brythonic Celtic and Old Persian (Clackson 1994: 53–54), so this phenomenon is too common to be significant for our purposes.[6]

2.1.4. Phrygian belongs to the *centum* group of IE languages (Ligorio and Lubotsky 2018: 1824). Together with Greek, Celtic, Italic, Germanic, Hittite and Tocharian, Phrygian merged the old palatovelars with plain velars in a first step: NPhr. (τιτ-)τετικμενος 'condemned' < PIE **deiḱ-*; NPhr. γεγαϱιτμενος 'devoted, at the mercy of' < PIE **ǵʰr̥Hit-*; NPhr. γλουϱεος 'golden' < PIE **ǵʰl̥h₃-ro-*. However, two shifts affected this language. Phrygian merged the old labiovelar with the plain velar (the etymological and the resulting ones): OPhr. *ke(y)*, NPhr. κε (passim) 'and' < PIE **kʷe*; OPhr. *knais* (B-07), NPhr. ⱪναικαν 'wife' (16.1 = 116) < **gʷneh₂i-*. Secondly, in contact with palatal vowels (/e/ and /i/, see de Lamberterie 2013: 25–26), and especially in initial position, some consonants became palatalised: PIE **ǵʰes-r-* 'hand' > OPhr. ↑*iray* (B-05),[7] NPhr. ζειϱα (40.1 = 12) 'id.' (Hämmig 2013: 150–151). It also occurs in glosses: **ǵʰeu̯-mn̥* > ζευμαν 'fount, source' (Hesychius ζ 128). These two secondary processes, as happened in Tocharian and the Romance languages, lend Phrygian the guise of a *satəm* language.

[5] I avoid this term because Phrygian does not show a complete consonantal shift; PIE voiceless stops appears as such in Phrygian. On the history of the Phrygian studies I strongly recommend Morante Mediavilla's PhD dissertation (2007).

[6] Gusmani (1959: 13–14) considered that Phrygian and Lycian were in contact when this sound disappeared. However, this loss takes place in many unrelated languages.

[7] On the possible representation of "a voiced affricate or /z/" by the OPhr. arrow-letter, see Ligorio and Lubotsky 2018: 1820.

2.1.5. The ancient **ḱi-* became *s-* in Phrygian, as shown by the demonstrative pronoun (Ligorio and Lubotsky 2018: 1824, 1826): *sin* (B-05, masc. acc. sing.), *si* (B-01, M-01b, neuter nom.-acc. sg.), σεμουν (masc. and neuter dat. sg.), σαι (feminine sg. dat. 43.1 = 69), etc. The same treatment is found in Greek: σά < **ḱi-h₂* (Megarian, Ar.Ach. 757, 784, pl. nom.-acc. neut.).

2.1.6. The PIE suffix **-ih₂* yields *-iya* in OPhr. *niptiya* (B-05), a development also found in some conditions in Greek (Myc. *-iya*; alphabetic Greek -ια), also in Tocharian (PIE **n̥-ǵneh₃-tih₂* 'ignorant' > **n̥-ǵnō-tyH₂* > PTo. **āknātsā* > ToB *aknātsa*). In other languages, such as Armenian, Albanian (Demiraj 1997: 61) and Indo-Iranian it appears as *-ī*: compare πότνια 'mistress' (Myc. *poti-ni-ja*) < **potnih₂* with Ved. *pátnī* 'mistress, goddess', Y.Av. *paθnī-* and Old Lith. *vieš-patni* 'housewife'.

2.1.7. The shift **-m* > *-n* occurs in Phrygian (**dh₁-tó-m* > δετο(υ)ν); it is shared with Greek and Armenian (Kim 2018, 256) but it also occurs in Anatolian ("parallel but independent in Luvian, Palaic and Hittite", Melchert 1994, 181), Insular Celtic, Germanic, Balto-Slavic and Tocharian. Like the loss of sibilant /s/, the frequency of this shift makes it few relevant.

2.1.8. At some point in Phrygian pre-history, the inherited voiced stops became devoiced: NPhr. Τιαν (acc. sg.), Τιος (gen. sg.), Τι(ε) (dat.sg.) 'Zeus' < PIE **diēm*, **diu̯os*, **diu̯ei*; OPhr. *torv-* (B-05) 'wood' < PIE **doru̯-*/**dru̯-* (see Ligorio and Lubotsky 2018: 1823). Phrygian devoicing is the strongest difference with regard to Greek; according to Kortlandt (2016: 249–250) it happened together with Thracian and Armenian.[8]

2.2. Morphological isoglosses

2.2.1. The prohibitive particle **meh₁* is found in Phrygian as *me*/με (Obrador-Cursach 2020: 296–297, with references), as well as in Armenian *mi*, Skt. *mā́*, Avestan *mā*, Greek μή, Albanian *mo* and Tocharian A and B *mā* (also used for simple negation 'no' in To.).

2.2.2. Phrygian particle *ni(y)*/νι is also found in Pamph. καί νι + imperv. (Pamph. 3, see Brixhe 1976: 131–132), although its use is different in the two languages.

2.2.3. The use of the Phrygian conditional conjunction *ai, ay*/αι (PIE **éh₂i*, **áHi* according to *LIPP* II 348) is shared with Dorian and Aeolian Greek (αἰ instead of the most common εἰ, in Homer also αἴθε for εἴθε) and Messapian *ai*.

2.2.4. In the light of υψοδαν 'above, on the top' (Lubotsky 1993), Phrygian also has a local particle **-dʰn̥* attested in some Greek dialects as -θα (e.g. Aeol. πρόσ-θα 'before').

2.2.5. The demonstrative particle *e-* (considered the same of the verbal augment in *LIPP* II, 179) occurs attached to demonstratives: *e-sai≠t* (W-01b) and, possibly, ε-σ?αν (40.3 = 31; suggested by Neumann 1986a). This particle can be equated to the one found in Greek ἐκεῖνος and ἐχθές (as in Tzitzilis 2013).

2.2.6. The compound pron. **h₂eu̯-to-* is only found in Greek αὐτός and Phr. αυτος. Albanian has an adverb derived from the zero-grade of this formation: **h₂u-tó-* > *hut* 'in vain' (compare with Greek αὖτως 'in this very manner, even so, in vain', Demiraj 1997: 59). A similar formation is Alb. *aý* 'this, he' < **h₂eu̯-so-* (Demiraj 1997: 70 and Matzinger 2005: 384 num. 9).

[8] After the devoicing and before the first Phrygian texts, a deaspiration affected Phrygian inherited aspirated stops: NPhr. δεως 'god' < PIE **dʰh₁so-*; NPhr. γεγρειμεναν 'written' < PIE **gʰreiH-*; NPhr. βρατερε 'brother' < PIE **bʰreh₂ter-* (Ligorio and Lubotsky 2018: 1823).

2.2.7. The IE noun **dʰeh₁s-/*dʰh₁s-* is found elsewhere in the Indo-European languages (see *NIL* 102), but only in Greek, Phrygian and Armenian does it take the new meaning 'god'. Note, however, that, unlike Armenian *dik*' 'god' (*plurale tantum*), Phrygian *devos* (B-07; NPhr. dat. pl δεως) and Greek θεός (Mycenaean *te-o*) go back to the thematised form **dʰh₁s-ó-*.[9]

2.2.8. The singular nominative of the ancient masculine nouns and adjectives in **-eh₂* takes an *-s* in Phrygian, as shown in the ethnics *Tias*, Πουντας and **Gordiyas* (see Obrador-Cursach 2019). The same innovation is found in Greek: compare νεανίας, μαθητής with Latin *agricola*. This also happened to the adjective for 'big'. To the nominative singular of the PIE **mégh₂* 'big', Greek and Phrygian added the ending *-s*: compare OPhr. *mekas* (M-05, G-111, P-03, etc.; Obrador-Cursach 2016) and Greek μέγας with Sanskrit *máhi* and Armenian *mec*. A common reformulation of the *a*-stems in Phrygian and Greek was considered by Brixhe (2006: 42–43).

2.2.9. Phrygian seems to share with Greek the existence of a suffix **-eu̯-/-ēu̯-*: Gr. -εύς/-έως (Epic -ῆος), Phr. *-avos* (thematised). As happens in Greek, the Phrygian forms with this suffix are titles whose origin is not always clear: *akenan-ogavos* ('the holder of the *akenan*' (M-01a, W-01a); *ạkinanogavaṇ* (M-04, sg.acc.fem.?); cf. Greek ὀχεύς, -έως, Epic -ῆος) and *proitavos* (M-01b, M-02).

2.2.10. In the inflection of *knais*, *knays* 'woman' (B-07, HP-114), forms other than the sg.nom. takes the suffix *-ik-*, just as in Greek and (slightly differently) in Armenian.[10] Compare sg.acc. ḳναικαν (16.1 = 116) with Gr. γυναῖκα and sg.gen. κναικο<ς> with Gr. γυναικός.

2.2.11. The e-grade in the root of the adjective **gu̯ʰer-mo-* seems to be found in some toponyms: Γέρμα (*KON* § 204–2 from Galatia), Γέρμη (*KON* § 204–2 from Mysia), *Germe* (*KON* § 204–3 from Mysia). Traditionally, they have been linked to the Phrygian language[11] and, indeed, this is the best option to explain them. The e-grade of this ancient adjective (secondary substantive) is shared with Greek θερμός 'warm' (also θέρμη 'heat and, in plural, 'hot springs'), Albanian *zjarm, zjarr* 'heat' and Armenian *ǰerm* 'warm(th)', an appellative also found in place names (e.g. *Jerm*, gen. *-ay*, see Martirosyan 2010: 681–682). Other languages show an o-grade in the root: **gu̯ʰor-mo-*. See, e.g., Latin *formus* 'warm' and Sanskrit *gharmá-* 'heat'.

2.2.12. The Phrygian word for 'name', *onoman* (W-01b), derives from a form **h₃nh₃-mn*, as well as Greek ὄνομα and Armenian *anun* 'id.' (Martirosyan 2013: 96–97). This proto-form differs from the other languages, which "have generalised one of the two ablaut grades" (*EDG* 1085).

2.2.13. Three words show that some *i*-stem anthroponyms and theonyms added a *t*-enlargement to their genitive forms (Brixhe 2006: 40): *artimitos* (B-05, the Greek goddess Ἄρτεμις, genitive Ἀρτέμιδος Ἀρτέμιτος Alcm.54, Ἀρτάμιτος *CID* 1.10.8, 12), *manitos* (B-07, genitive of *manes*), *urgitos* (K-01, perhaps borrowed from Hurrian *Urḫi-*, as in *Urḫi-Tešub*, heteronym of Muršili III). This feature also occurs in some Greek words (see, e.g., Greek ὄρνις, -ῑθος 'bird').

2.2.14. Phrygian shares the *e*-augment with Greek, Armenian and Indo-Iranian: compare 3sg.aor. PIE **(h₁)é-dʰeh₁-t* '(s)he put', OPhr. *e-da-es*, Greek ἔ-θη-κα, Armenian *e-d* and Skt. *á-dhā-t*. This is considered a very significant morphological isogloss by Martirosyan (2013: 90).

2.2.15. Phrygian perfect middle participles are built through the PIE suffix **-mh₁no-* (NPhr. γεγαριτμενος 'cursed, devoted', γεγριμενον 'written', τιττετικμενος 'accursed', etc.), as well

[9] A parallel formation, **dʰh₁-s-ó-*, is found in some Anatolian languages: Hieroglyphic Luwian *tasa-* 'stele', Lycian *θθe-* 'sanctuary' and Lydian *tasẽv* 'statue (?)' (see Payne and Sasseville 2016: 76).

[10] On the PIE root of this noun and derivates, see the recent work by Opfermann (2017).

[11] See, e.g., Matzinger 2005: 384, num. 8 and, more cautiously, de Lamberterie 2013: 20.

as Greek (-μενος-, see de Lamberterie 2013: 43–44). The exact form is also found in ToA *-māṃ*, ToB *-mane* and Prakrit *-mina-* (slightly different Skt. *-a-māna-*, Av. *-a-mna-*) and frozen in some Lat. words such as *alumnus* (**h₂al-o-mh₁no-*) or *fēmina* (< **dʰeh₁-mh₁neh₂-*).

2.2.16. Phrygian has a 3sg. middle ending *-tor*: αββερετορ 14.1 = 73, 52.1 = 75, 59.3 = 79 and αδδακετορ 20.1 = 63, 24.1 = 40. According to Kortlandt (2016: 252), the forms with *-tor* (considered imperfect by him) "must be compared with the Armenian imperfect *berēr* < **-etor*, *beriwr* < **-etro*" (but see below § 3.2.) and are shared with Latin *-tur* and Old Irish *-thar*. This ending was lost in Greek after the generalisation of the ending -τοι, also found in Phrygian: OPhr. *edạtoy* (B-05), *tedatoy* (W-01a), *odeketoy* (B-07), NPhr. αββερετοι (10.2 = 113, 17.5 = 91, 19.2 = 129).[12] The presence of both *-tor* and *-toi* in Phrygian has been used to show the language's associations with Armenian (Kortlandt 2016: 252) and with Greek (Ligorio and Lubotsky 2013: 1816–1817) respectively. However, the presence of *-tor* in Phrygian is inherited from the proto-language and, as a residual feature, it is of little value for elucidating any dialectal position. Similarly, the sole presence of *-toi* does not show a special relation between two branches or languages, since it is "a simple case of analogical extension" also attested in Indo-Iranian and Germanic (Jasanoff 2003: 46).

2.2.17. In the Phr. corpus two denominative verbs are found: *kakoioi* (G-02c, also *kakuioi* P-04b; see Lejeune 1969: 299) and *kuryaneyon* (W-01c). The first is a verb in *-o-yo-* derived from an *o*-stem noun, *kako-*, while the second is in *-e-yo-*, derived from another *o*-stem noun, although only attested in Greek κοίρανος, -ου. In both cases one cannot rule out an early Greek borrowing dated to Myc. times. However, there is no phonetic reason for ruling them out as Phrygian cognates of the Greek forms.

2.3. Lexical isoglosses

2.3.1. A derived noun **bʰoh₂-t-*/**bʰeh₂-t-* from the PIE verbal root **bʰeh₂-* 'to shine' (*LIV*² 68–69) has been widely considered to yield Greek φῶς, φωτός 'man, mortal, hero' (Peters 1993: 101–108 and Vijūnas 2009: 72–76, following a previous suggestion by Brugmann & Delbrück 1906: 536; *NIL* 7). Recently, a formation of this kind has been claimed for the Phrygian DN βας (acc. βαταν, see Obrador-Cursach 2017).

2.3.2. The formation **(h₁)en-mén-* (*LIPP* II, 234) derived from **men* 'stand fast, remain' (*LIV*² 437) appears in the Greek adverb ἐμμενές 'unceasing' and Phrygian *iman* (sg.dat. *inmeney*, B-05) according to Vine (2010). It is also found in OIr. *ainmne* 'patience'.

2.3.3. Although the root **ǵʰl̥h₃-* 'green, yellow' is found in Greek χλωρός 'greenish-yellow, pale green', in Lat. *helus*, (*h*)*olus* 'green plants, vegetables, cabbage', in Skt. *hari-* 'fallow, greenish' (used in compounds), etc., the formation **ǵʰl̥h₃-ró-* is only found in Greek χλωρός and Phrygian γλουρεος 'golden' (MPhr-01 l. 2 identified by Brixhe 2004a, 17 through the Hesychius gloss γλούρεα, γ 659).

2.3.4. The stem *kako-* 'bad' is found both in Greek κακός 'bad, worthless' and, nominalised, in Phrygian κακουν 'harm' (*passim*). Thought of a Greek loan since Solmsen (1897: 52 fn. 4), there

[12] There are some other marginal forms. On the divergence in the endings, Kortlandt (2016: 252, reconsidering Lubotsky 1997: 127) considers that the Phrygian forms in *-tor* "were middle forms of the imperfect". However, things are not this simple: The forms on *-tor* occur in the same imprecative apodosis where a present is needed and the forms αδδακετ and αββερετ mainly occur.

is no reason not to consider it as a Phrygian inherited word shared with Greek. So, its status remains unclear. If a Greek borrowing, it is one of the oldest, since the stem is found in the verb *kakoioi* in G-02, an OPhr. inscription dated earlier than the 6th century BC.

2.3.5. Obrador-Cursach (2020: 269–270) suggests that OPhr. *keneman* 'cavity, niche' (M-01b) derives from **ḱen-e/o-* 'empty', a form found in Proto-Greek **ken(e)-u̯ó-* 'empty, idle' (Att. κενός, Ion. κεινός but Epic and Dor. κενεός) and Armenian *sin*, *sn-oy* 'empty'.[13] Previously, *keneman* was considered to derive from **kenh₁-*, attested in Sanskrit *khánati* 'to dig' (Lubotsky 1988: 15), but this root is not found outside Indo-Iranian.

2.3.6. The stem *koru̯-* (perhaps derived from **ḱerh₃-* 'grow' *LIV*² 329) is a Greek peculiarity attested in κόρη 'young girl, daughter', κοῦρος, κόρος 'boy' (*EDG* 752–753) but recently it has also been identified in MPhr. κορο- (MPhr-01, Lubotsky 2017), which follows the original *u̯*-stem inferred for the Proto-Greek forms (see Obrador-Cursach 2020: 177–178 and 277 and fthc.).

2.3.7. The adjective **mōro-* 'stupid' is found only in Greek μωρός and, nominalised, in Phrygian μουρου[ν] (61.1 = 100; plural μμυρα [sic] 35.1 = 5). As in the case of κακουν, it can be considered a Greek borrowing, although the word does not show any feature that rule it out as an inherited word (its etymology is still lacking).

2.3.8. The PIE verb **sleh₂gʷ-* 'take hold of, grasp' (*LIV*² 566) was only attested in Greek λαβεῖν 'id.' until Lubotsky (2004: 234–235) equated to Phrygian *lakedo* (W-01b).

3. Evaluation of the data

3.1. Phrygian as the closest language to Greek

To the best of our current knowledge, Phrygian was closely related to Greek. This affirmation is consistent with the vision offered by Neumann (1988: 23), Brixhe (2006) and Ligorio and Lubotsky (2018: 1816) and with many observations given by ancient authors.[14] Both languages share 34 of the 36 features considered in this paper, some of them of great significance: the "triple representation" of the vocalisation of the PIE laryngeals, the **-eu̯-*/**-ēu̯-* suffix or the addition or the ending *-s* to the nominatives of the ancient *h₂*-masculines. Other features are of little value if taken one by one, because they are shared with other unrelated Indo-European branches: the treatment **-ih₂ > -iya*, for example, is also found in Tocharian. The importance of features of this kind is their accumulation. The available data suggest that Phrygian and Greek coexisted broadly from pre-historic to historic times, and both belong to a common linguistic area (Brixhe 2006: 39–44). One of the features that emerged as a result is the shift **-ans > -ais* and **-ons > -ois* found in Phrygian and Aeolic, a common characteristic in the shared linguistic area (Brixhe 1990, 65–67 and 2004, 41–42). However, in such close languages, it is not always easy to differentiate the genetic and the areal features.

[13] *EDG* 672; Clackson 1994: 138–139 and Martirosyan 2010: 575 and 2013: 116.

[14] See, especially, Plato, *Cratilus*. 410a (Socrates is speaking): Ὅρα τοίνυν καὶ τοῦτο τὸ ὄνομα τὸ "πῦρ" μή τι βαρβαρικὸν ᾖ. Τοῦτο γὰρ οὔτε ῥᾴδιον προσάψαι ἐστὶν Ἑλληνικῇ φωνῇ, φανεροί τ᾽εἰσὶν οὕτως αὐτὸ καλοῦντες Φρύγες σμικρόν τι παρακλίνοντες· καὶ τό γε "ὕδωρ" καὶ τὰς "κύνας" καὶ ἄλλα πολλά. "Well, this word πῦρ is probably foreign; for it is difficult to connect it with the Greek language, and besides, the Phrygians have the same word, only slightly altered. The same is the case with ὕδωρ (water), κύων (dog), and many other words" (translation by Harold N. Fowler).

Table 1. Comparison of Phrygian features with Greek, Armenian, Albanian and Indo-Iranian (grey files indicate that borrowing cannot be totally ruled out)

	Phrygian feature		Greek	Armenian	Albanian	Indo-Iranian
Phonetic	2.1.1.	"prothetic vowels"	+	+	+	–
	2.1.2.	*CRh$_3$C > *CRōC	+	–	–	–
	2.1.3.	Loss of /s/	+	+	+	–
	2.1.4.	*Centum* treatment	+	–	–	–
	2.1.5.	*-ih$_2$ > -iya*	+	–	+	–
	2.1.6.	*ḱi- > s-	+	–	–	–
	2.1.7.	*-m > -n*	+	+	?	–
	2.1.8.	*M > T*	–	+	–	–
Morphological	2.2.1.	*meh$_1$	+	+	+	+
	2.2.2.	*ni(y)*/νι	+	–	–	–
	2.2.3.	Conditional *ai*	+	–	–	–
	2.2.4.	*-d^{h}n̥	+	–	–	–
	2.2.5.	*e*-demonstrative	+	–	–	–
	2.2.6.	*h$_2$eu̯-to-	+	–	+	–
	2.2.7.	*d^{h}h$_1$s-ó-	+	–	–	–
	2.2.8.	*-eh$_2$-s* masc.	+	–	–	–
	2.2.8.	*mégh$_2$-s	+	–	–	–
	2.2.9.	*-eu̯-/*-ēu̯-	+	–	–	–
	2.2.10.	*g^{u}neh$_2$-ik-	+	+	–	–
	2.2.11.	*g^{uh}er-mo-	+	+	+	–
	2.2.12.	*h$_3$nh$_3$-mn-	+	+	–	–
	2.2.13.	*t*-enlargement	+	–	–	–
	2.2.14.	*e*-augment	+	+	+	+
	2.2.15.	*-mh$_1$no-	+	–	–	–
	2.2.16.	*-toy*/-τοι	+	–	–	+
	2.2.16.	*-(t)or	–	?	–	–
	2.2.17.	verbs in *-o-yo-*	+	–	–	–
	2.2.17.	verbs in *-e-yo-*	+	–	–	–
Lexical	2.3.1.	*b^{h}oh$_2$-t-/*b^{h}eh$_2$-t-	+	–	–	–
	2.3.2.	*(h$_1$)en-mén-	+	–	–	–
	2.3.3.	*ǵ^{h}l̥h$_3$-ró-	+	–	–	–
	2.3.4.	*kako-*	+	–	–	–
	2.3.5.	*ken-*	+	+	–	–
	2.3.6.	*koru̯-	+	–	–	–
	2.3.7.	*mōro-*	+	–	–	–
	2.3.8.	*sleh$_2$g^u-	+	–	–	–

Also problematic are the lexical isoglosses. Because of the many similarities between Phrygian and Greek, it is often difficult to discern when words are loans and when they are shared or inherited. When a word lacks the expected inner evolution and can be explained as a mere borrowing, the latter explanation is preferred. Because the devoicing did not work in OPhr. *lavagtaei* (W-01a), it is preferable to consider it a loan from a (pre-alphabetic) Greek dialect (cf. Mycenaeans *ra-wa-ke-ta*; Doric λαγέτας 'leader of the people'). For its part, γλουρεα (MPhr-01 = W-11) must be a genuine Phrygian word, since it shows the result of the deaspiration in the first stop. Nevertheless, we do not have data of this kind in some lexemes. Both *kako-* (§2.3.6.) and **mōro-* (§2.3.7.) can be considered borrowings or inherited words as well, at least until a satisfactory etymology is found. A very specific problem concerns the shared substratum. If a word has no special Phrygian feature, the decision is not easy. The clearest example of this problem is Phrygian *vanakt-* (sg. nom. in M-04 *modro-vanak*, sg. acc. 6.1 = 88 ουανακταν, sg. dat. M-01 *vanaktei*) and Greek (Ϝ)άναξ (Mycenaean *wa-na-ka*). Was it borrowed from a substratum language by both Phrygian and Greek, was it borrowed only from that assumed language by Greek, and secondly borrowed from Greek by Phrygian? Honestly, we do not know;[15] for this reason, I exclude this word from this discussion. Because of the fragmentary nature of the Phrygian corpus, a large set of words which might shed light on this point is missing.

The affinity between the two languages cannot be denied. Indeed, they are so close that some scholars have suggested a Proto-Greco-Phrygian language, a common pre-historic stage (Neumann 1988). In the light of the available data, this is a very attractive idea. However, Greek has some features that distance it from Phrygian, and in fact they can be used in order to discern when this suggested common language split. A nice source of information is the treatment of the aspirate in the languages. It now seems clear that Grassmann's law did not affect Phrygian: compare OPhr. *bevdos* 'statue' (B-01, identified by Lubotsky 2008) and Greek πεύθομαι 'learn', both derived from the PIE root **b^{h}eu̯d^{h}-* (*LIV*^{2} 82–83). It remains unclear, however, when the dissimilation of the aspirate affected Greek, but if the idea of Proto-Greco-Phrygian is accepted, it must have occurred between the split of the two languages and post-Mycenean times (see Risch and Hajnal 2006: 303–304). Also, in relation with the aspirated stops, it would be interesting to know whether their devoicing in Greek also affected Phrygian.

3.2. Phrygian and Armenian, an insistent comparison

From the beginnings of the Indo-European theory to the works by Lejeune, who spoke against this idea (see especially 1979), Phrygian was considered to be extremely close to Armenian. Indeed, researchers such as Paul Anton de Lagarde, Christian Lassen, Andreas D. Mordtmann, Richard Gosche, Friedrich C. A. Fick and Otto Hass considered that Phrygian might even been an earlier stage of Armenian. Working basically on glosses, they inferred from them that Phrygian was a *satəm* language (the first ones *avant la lettre*) just like Armenian and concluded that the stops of the two languages shared their *Lautverschiebung*. These researchers also paid excessive attention to the three ancient passages mentioning the relation between the Armenians and the Phrygians: Herodotus' description of Xerxes' army against Greece, in which the Armenians are said to be settlers of the Phrygians,[16] the Stephanus Byzantinus' re-

[15] If we consider that another aristocratic title, *lavagtaei* (W-01a, dat. sg.), is a Greek borrowing, it is not difficult to assume that *vanakt-*, which shares the same semantic field, was also borrowed.

[16] Hdt. 7.73: Φρύγες δὲ ἀγχοτάτω τῆς Παφλαγονικῆς σκευὴν εἶχον, ὀλίγον παραλλάσσοντες. οἱ δὲ Φρύγες, ὡς Μακεδόνες λέγουσι, ἐκαλέοντο Βρίγες χρόνον ὅσον Εὐρωπήιοι ἐόντες σύνοικοι ἦσαν Μακεδόσι, μεταβάντες δὲ ἐς τὴν Ἀσίην ἅμα τῇ χώρῃ καὶ τὸ οὔνομα μετέβαλον ἐς Φρύγας· Ἀρμένιοι δὲ κατά περ Φρύγες

port of Eudoxos' opinion of the similarity between Armenian and Phrygian,[17] and the same observations reported by Eustathius of Thessalonica.[18]

Phrygian has constantly been compared to Armenian, but never with any great success. As we can see in Table 1, Phrygian shares fewer isoglosses with Armenian than with Greek, and only two of them are exclusive. Today, we know that Phrygian is a *centum* language and, despite Lejeune's (1979) attempts to deny it (followed by Brixhe 2008: 74, Matzinger 2005: 378 and de Lamberterie 2013: 27–28), Lubotsky (2004) showed that Phrygian stops were affected by devoicing and deaspiration, an incomplete *Lautverschiebung*. However, the phenomenon is not exactly the same as the one found in Armenian; in fact, the two languages only share the devoicing. Recently, Kortlandt (2016: 250) considered this shift to be an innovation which affected Armenian, Thracian (see below) and Phrygian, and this is one of the clearest differences between Phrygian and Greek. However, more than genetic, it should be considered a real feature of the proto-languages in their pre-historic homeland. Leaving Thracian aside (considered "an early dialect of Proto-Armenian" by Kortlandt 2016: 249), one must wonder if this is the only possible scenario for the Phrygian devoicing. Indeed, we do not know when it occurred, although it must have been after the split between Phrygian and Greek (if we assume so) and before the Phrygian deaspiration. It is possible, then, that the Phrygian deaspiration also occurred in the historical territory of Phrygian in central Anatolia, where the languages spoken lack voicing contrast. Kortlandt (2016: 252) also considered a second isogloss shared by Phrygian and Armenian and not by Greek, the 3sg. middle ending *-tor*. However, since both languages retained a common feature also found in other branches, this is irrelevant to the dialectal position of Phrygian. Even worse for Kortlandt's proposal, the Armenian ending *-ēr* may derive from **-ter* rather than from **-tor* and, consequently, the ending *-iwr*, considered a mere dialectal variant, is problematic.[19]

Note also that sometimes Greek σῦριγξ, ιγγος 'pipe, panpipe' and Armenian *sring* 'pipe, fife' have been considered as Phrygian in origin (if not elusive Mediterranean) and a proof of contact of the latter with the former two languages (see, e.g. Greppin 1990: 35 and Martirosyan 2010: 585 and 2013: 123). However, even if Armenian *sring* is a mere borrowing from Greek, any relation between this noun and Phrygian is ungrounded. Although we ignore the Phrygian equivalent, ancient sources are not explicit in this relation and its morphology excludes this proposal: an initial **s-* followed by non-front vowel is unexpected for a Phrygian inherited

---

ἐσεσάχατο, ἐόντες Φρυγῶν ἄποικοι. τούτων συναμφοτέρων ἦρχε Ἀρτόχμης, ἔχων Δαρείου θυγατέρα. "The Phrygian equipment was very similar to the Paphlagonian, with only a small difference. As the Macedonians say, these Phrygians were called Briges as long as they dwelt in Europe, where they were neighbors of the Macedonians; but when they changed their home to Asia, they changed their name also and were called Phrygians. The Armenians, who are settlers from Phrygia, were armed like the Phrygians. Both these together had as their commander Artochmes, who had married a daughter of Darius." (translation by A. D. Godley).

[17] St.Byz., s.v.: Ἀρμενία· χώρα πλησίον τῶν Περσῶν, ἀπὸ Ἀρμένου Ῥοδίου, ὡς Ἀντίπατρος ἐν γ Περὶ Ῥόδου. οἱ οἰκήτορες Ἀρμένιοι. ὡς Εὔδοξος α Γῆς περιόδου, Ἀρμένιοι δὲ τὸ μὲν γένος ἐκ Φρυγίας καὶ τῇ φωνῇ πολλὰ φρυγίζουσι. 'Armenia: a region near the Persians called after Armenus Rhodius according to Antipater's *On Rhodes* book III. The dwellers are called Armenians but, according to Eudoxus' *Chart of the earth* book I, the Armenians are native of Phrygia and their tongue resembles very much the Phrygian one.'

[18] Eust., *ad D. P.* 694: Ἰστέον δὲ ὅτι Ἡρόδοτος τοὺς Ἀρμενίους Φρυγῶν ἀποίκους φησὶ, λέγων καὶ ὅτι πολυπρόβατοί εἰσι. Καὶ Εὔδοξος δὲ ἐν Γῆς Περιόδῳ φησίν· "Ἀρμένιοι τὸ γένος ἐκ Φρυγίας, καὶ τῇ φωνῇ πολλὰ φρυγίζουσι." 'One must know that Herodotus says that the Armenians were settlers of the Phrygians and rich in cattle. Also, Eudoxus in *Chart of the earth* says: "the Armenians are native of Phrygia and their tongue resembles very much the Phrygian one."'

[19] I am in debt with Hrach Martirosyan for generously sharing with me the corresponding discussion for his unpublished work *Historical phonology and morphology of Classical Armenian* (§ M 504).

word and the suffix -γγ- is never attested in this language. Similarly, Arm. *sayl* 'wagon; Ursa Major and Minor, Arcturus', related somehow to Gr. σατίνη 'chariot' and the Hesychian gloss σάτιλλα· πλειὰς τὸ ἄστρον '*sátilla*: the cart constellation' (see Martirosyan 2010: 566–567 and 769, with references), defy any derivation from Phrygian. Precisely, its Phrygian equivalent can be found in another Hesychian gloss: κίκλην· τὴν ἄρκτον τὸ ἄστρον. Φρύγες '*kíklēn*: the constellation Ursa Major'.

Leaving aside the devoicing and, if still accepted, the retention of the ending *-*tor*, no other features are exclusively shared by Phrygian and Armenian; all the others are shared with Greek. With the available Phrygian data, the relation between this language and Armenian does not differ greatly from that between Greek and Armenian. It seems that rather than a close genetic affinity, Armenian and Phrygian may have shared the same linguistic area, as suggested recently by Kim (2018) for Greek and Armenian (see also Martirosyan 2013: 126). It is even possible that such an area (including its substrate language(s)) was shared by Proto-Greco-Phrygian and Proto-Armenian. However, if Greco-Armenian is still preferred, there is no strong evidence for a Proto-Phrygio-Armenian split from Proto-Greco-Phrygio-Armenian. In both scenarios, the relation of Phrygian and Armenian seems to be subordinate to the relation between Phrygian and Greek and the relation between Proto-(Phrygio-)Greek and Proto-Armenian.

3.3. The relation between Phrygian and Thracian: a minefield

Thraco-Phrygian is a label that has had more success among archaeologists[20] and linguists, despite being a creation of von Bradke (1890: 66). Although some affinities between the Thracians and the Phrygians have been considered in view of the remaining material culture, and since some ancient sources considered that the Phrygians were a tribe of the Thracians (e.g., Str. 7a.1.25), the comparison between the two languages is not easy, since the vestiges of Thracian are even more fragmentary than Phrygian and our knowledge of this language, mainly based on glosses, is very poor.

It remains unclear whether Thracian suffered a complete *Lautverschiebung*, as traditionally considered (see, e.g., Brixhe 2018: 1852), although at least deaspiration occurred in this language: see the toponyms in -διζα, -διζη, -διζον 'fortress' < *$d^{h}e\hat{i}g^{h}$- 'to knead clay, to build' (cf. Greek τεῖχος 'wall'). However, since the outcomes of the PIE laryngeals in this language seem to be different from Phrygian and Greek, it remains possible that Thracian was a *satəm* language. Note, however, that Brixhe (2018: 1853) considers that the ancient stops represented by ⟨ζ⟩ "may represent simply a palatalization (at least of *d/g) before front vowel", as considered for some Phrygian results. As can be seen, there are possible affinities in the phonology of Phrygian and Thracian, but they seem to point to ancient contacts rather than to genetic affinity.

3.4. The classical comparison between Phrygian and Ancient Macedonian

Since the ancient aspirates *b^{h}, *d^{h}, *g^{h} were represented as ⟨β, δ, γ⟩ in the available material related to Ancient Macedonian (see, e.g., the famous PN Βερενίκα = Attic Φερενίκη < *$b^{h}er$- and the gloss βρ(ο)ῦτος 'beer' < PIE *$b^{h}reu\underset{\frown}{\mu}$-) and ancient sources state that in the original homeland of the Phrygians in the Balkans they were neighbours of the Macedonians (Hdt. 8.138), the two languages were traditionally considered to be closely related. However, inferring from glosses and the few texts possibly written in Ancient Macedonian, there are strong arguments for considering that this language was a Greek dialect (see Méndez Dosuna 2012 for a convincing

[20] See an updated account by Tsetskhladze 2007, with references.

survey of the advantages and disadvantages). Although there is no universal consensus, internal Macedonian data dispel the idea of a special relation with Phrygian. Indeed, the use of ⟨β, δ, γ⟩ for ancient aspirates in Macedonian is better explained as the result of spirantization, *b^h, *d^h, *g^h (> *p^h, *t^h, *k^h) > /f θ x/ ~ [v ð γ], rather than deaspiration (Méndez Dosuna 2012: 134–137). Moreover, there is evidence that Macedonian was affected by the devoicing of the aspirates and Grassmann's law, which is incompatible with Phrygian: see the PN Κεβαλῖνος (= Att. Κεφαλῖνος) < *g^heb^h-. Therefore, whether or not Ancient Macedonian is accepted as a Greek dialect with "an inner-Greek *Lautverschiebung*", it does not show any special relation with Phrygian.

3.5. The limited relation between Albanian and Phrygian

Although the first testimonies of Albanian are very late, it shares certain features with the ancient Balkan languages. For this reason, it is sometimes considered in the discussion of Phrygian. As in the case of Armenian, the comparison is not easy because of the internal evolution of this language revealed in its different historical stages. However, the comparison between languages does not yield enough isoglosses to be able to conclude that Albanian and Phrygian were more closely related in pre-history than Greek and Albanian and there is no one isogloss that occurs only in these two languages. It is true that the PIE word *$ǵ^hes$-*r*- 'hand' (*NIL* 170–172, Hittite *keššar* 'id.', Greek χείρ, χειρός 'id.', Armenian *jern* 'id.', etc.) becomes an *a*-stem in Phrygian (OPhr. ↑*iraẏ* B-05 and NPhr. ζειραι 59.4 = 106, both sg. dat.) and in Albanian (*dorë*) but this is a general phenomenon in this latter language (Demiraj 1997: 140 with references).

4. Conclusion

With the current state of our knowledge, we can affirm that Phrygian is closely related to Greek. This is not a surprising conclusion: ancient sources and modern scholars agree that Phrygians did not live far from Greece in pre-historic times. Moreover, the last half century of scientific study of Phrygian has approached both languages and developed the hypothesis of a Proto-Greco-Phrygian language, to the detriment to other theories like Phrygio-Armenian or Thraco-Phrygian. The result of this study, as fragmentary as our knowledge of Phrygian, is consistent with this view. Future findings and new identifications may provide more information for scholars to work with.

Abbreviatures

CID = Rougemont 1977; *CIPPh* = Brixhe & Lejeune 1984; *EDG* = Beekes 2010; *KON* = Zgusta 1984; *LIPP* = Dunkel 2014; *LIV*^{2} = Rix 2001; *NIL* = Wodtko, Irslinger, Schneider 2008.

References

Beekes, Robert. 2010. *Etymological Dictionary of Greek*. Leiden/Boston: Brill.

von Bradke, Peter. 1890. *Über Methode und Ergebnisse der arische (indogermanischen) Alterthumswissenschaft*. Giessen: J. Ricker'sche Buchhandlung.

Brixhe, Claude. 1976. *Le dialecte grec de Pamphylie. Documents et grammaire*. Paris: Lib. d'Amérique et d'Orient Adrien Maisonneuve.

Brixhe, Claude. 1990. Comparaison et langues faiblement documentées: l'exemple du phrygien et de ses voyelles longues. In J. Dor, Jean Kellens (eds.). *La reconstruction des laryngales*: 59–99. Liège: Université de Liège.

Brixhe, Claude. 2004. Corpus des inscriptions paléo-phrygiennes: Supplément II. *Kadmos* 43: 1–130.

Brixhe, Claude. 2006. Préhistoire et début de l'histoire des dialectes grecs. *Incontri Linguistici* 29: 39–59.

Brixhe, Claude. 2008. Phrygian. In: R. D. Woodard, R. D. (ed.). *The Ancient Languages of Asia Minor*: 69–80. Cambridge: Cambridge University Press.

Brixhe, Claude. 2018. Thracian. In: J. Klein, J. Brian, M. Fritz (eds.). *Handbook of Comparative and Historical Indo-European Linguistics*. Vol. 3: 1850–1854. Berlin/Boston: Walter de Gruyter.

Brixhe, Claude, Michel Lejeune. 1984. *Corpus des inscriptions paléo-phrygiennes*. 2 vols. Paris: Institut français d'études anatoliennes.

Brugmann, Karl, Berthold Delbrück. 1906. *Grundriss Vergleichenden Grammatik der Indogermanischen Sprachen* 2.1. Berlin: Walter de Gruyter.

Clackson, James. 1994. *The linguistic relationship between Armenian and Greek*. Oxford/Cambridge: Blackwell.

Demiraj, Bardhyl. 1997. *Albanische Etymologien. Untersuchugen zum albanischen Erbwortschatz*. Amsterdam, Atlanta: Rodopi.

Dunkel, George E. 2014. *Lexikon der indogermanischen Partikeln und Pronominalstämme*. Heidelberg: Winter.

Greppin, John A. C. 1990. The Etymology of Greek σῦριγξ 'Pipe'. *Historische Sprachforschung* 103(1): 35–37.

Gusmani, Roberto. 1959. Relazioni linguistiche tra Frigia e Licia. *Archivio glottologico italiano* 49: 9–16.

Haas, Otto. 1966. *Phrygische Sprachdenkmäler*. Sophia: Die Phrygischen Sprachdenkmäler.

Hämmig, Anna Elisabeth. 2013. *Nevotan niptiyan*, die Fluchformel der Stele von Vezirhan. *Indogermanische Forschungen* 118: 125–54.

Jasanoff, Jay H. 2003. *Hittite and the Indo-European Verb*. Oxford: Oxford University Press.

Kim, Ronald I. 2018. Greco-Armenian. The Persistence of a myth. *Indogermanische Forschungen* 123: 427–272.

Kortlandt, Frederik. 2016. Phrygian between Greek and Armenian. *Linguistique Balkanique* 55: 2–3.

de Lamberterie, Charles. 2013. Grec, phrygien, arménien: des anciens aux modernes. *Journal des Savants* 2013, 1: 3–69.

Lejeune, Michel 1969. Notes paléo-phrygiennes. *Revue des études anciennes* 71: 287–300.

Lejeune, Michel. 1979. Regards sur les sonores I.E. en vieux phrygien. In: *Florilegium Anatolicum: Mélanges offerts à Emmanuel Laroche*: 219–224. Paris: De Boccard.

Ligorio, Orsat, Alexander Lubotsky. 2018. Phrygian. In: J. Klein, J. Brian, M. Fritz (eds.). *Handbook of Comparative and Historical Indo-European Linguistics*. Volume 3: 1826–1831. Berlin/Boston: de Gruyter.

Lubotsky, Alexander. 1988. The Old Phrygian Areyastis-Inscription. *Kadmos* 27(1): 9–26.

Lubotsky, Alexander. 1993. New Phrygian υψοδαν. *Kadmos* 32: 127–34.

Lubotsky, Alexander. 1997. New Phrygian Inscription No. 48: Paleographic and Linguistic Comments. In: R. Gusmani, M. Salvini, P. Vannicelli (eds.). *Frigi e frigio. Atti del 1º Simposio Internazionale, Roma, 16–17 ottobre 1995*: 115–130. Roma: Consiglio Nazionale delle Ricerche.

Lubotsky, Alexander. 2004. The Phrygian Zeus and the Problem of the 'Lautverschiebung'. *Historische Sprachforschung* 117(2): 229–37.

Lubotsky, Alexander. 2008. Old Phrygian *bevdos* 'Statue, Image', Greek βεῦδος 'Woman's Dress'. *The Journal of Indo-European Studies* 36: 96–97.

Lubotsky, Alexander. 2017. The Phrygian Inscription from Dokimeion and Its Meter. In: I. Hajnal, D. Kölligan, K. Zipser (eds.). *Miscellanea Indogermanica. Festschrift für José Luis García Ramón zum 65. Geburtstag*: 427–31. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Martirosyan, Hrach. 2010. *Etymological Dictionary of the Armenian Inherited Lexicon*. Leiden: Brill.

Martirosyan, Hrach. 2013. The place of Armenian in the Indo-European Language family: The Relationship with Greek and Indo-Iranian. *Journal of Language Relationship* 10: 85—137.

Matzinger, Joachim. 2005. Phrygisch und Armenisch. In: G. Meiser, O. Hackstein (eds.). *Sprachkontakt und Sprachwandel: Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17.–23. September 2000, Halle an der Saale*: 375–94. Wiesbaden: Dr. Ludwig, Reichert.

Melchert, H. Craig. 1994. *Anatolian Historical Phonology*. Amsterdam/Atlanta: Rodopi.

Méndez Dosuna, J. 2012. Ancient Macedonian as a Greek Dialect: A Critical Survey on Recent Work. In: G. Giannakis (ed.). *Ancient Macedonia: Language, History, Culture*: 133–45. Thessaloniki: Centre for the Greek Language.

Morante Mediavilla, Bibiana. 2007. *Studia Phrygia : la incidencia del testimonio de la literatura griega y romana en la descripción y las tentivas de determinar la posición lingüística de la lengua frigia desde la Antigüedad hasta el auge de la Lingüística Indoeuropea*. Barcelona: PhD dissertation.

Neumann, Günter. 1986. Zur Syntax der neuphrygischen Inschrift Nr. 31. *Kadmos* 241: 79–84.
Neumann, Günter. 1988. *Phrygisch und Griechisch*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
Obrador-Cursach, Bartomeu. 2016. Phrygian *mekas* and the recently discovered New Phrygian inscription from Nacoleia. *Indogermanische Forschungen* 121: 177–86.
Obrador-Cursach, Bartomeu. 2017. The Phrygian God Bas. *Journal of Near Eastern Studies* 76(2): 307–17.
Obrador-Cursach, Bartomeu. 2019. Γορδίας and Γόρδιον: a remark on Phrygian onomastics and ethnics. *The Journal of Indo-European Studies*, 47, 3–4, .
Obrador-Cursach, Bartomeu. 2020. *The Phrygian Language*. Leiden: Brill.
Obrador-Cursach, Bartomeu. Forthcoming. The last verse of the Middle Phrygian epigram from Dokimeion. *Indogermanische Forschungen*.
Opfermann, Andreas. 2017. Griechische, Armenische und Albanische ‚Frauen' und die Wurzel **Gu̯en-*. In: H. Bichlmeier, A. Opfermann (eds.). *Das Menschenbild bei den Indogermanen*: 55–80. Hamburg: Baar-Verlag.
Payne, Annick, David Sasseville. 2016. Die lydische Athene: Eine neue Edition von LW 40. *Historische Sprachforschung* 129: 66–82.
Peters, Martin. 1993. Beiträge zur griechischen Etymologie. In: L. Isebaert (ed.). *Miscellanea Linguistica Graeco-Latina*: 101–8. Namur: Société des études classiques.
Risch, Ernst, Ivo Hajnal. 2006. *Grammatik des mykenischen Griechisch. 1. Bd. Einleitung und Phonologie*. Avaiable: <http://www. uibk.ac.at/sprachen-literaturen/sprawi/mykgr.html> [Accessed 09.20.2019].
Rix, Helmut, 2001. *Lexikon der indogermanischen Verben*. Wiesbaden: Ludwig Reichert.
Rougemont, Georges. 1977. *Lois sacrées et règlements religieux*. Paris: de Boccard.
Tsetskhladze, Gocha R. 2007. Thracians versus Phrygians: about the origin of the Phrygians once again. In: A. Çilingiroglu, A. Sagona (eds.). *Anatolian Iron Ages* 6: 283–310. Leuven: Peeters.
Tzitzilis, Christos. 2013. Greek and Phrygian. In: *Encyclopedia of Ancient Greek Language and Linguistics*. Available: http://dx.doi.org/10.1163/2214–448X_eagll_COM_00000160 [Accessed 07.20.2019].
Vijūnas, Aurelijus. 2009. *The Indo-European Primary t-stems*. Innsbruck: Institut für Sprachen und Literaturen der Universität Innsbruck Bereich Sprachwissenschaft.
Vine, Brent. 2010. Old Phrygian iman. In: R. Kim, N. Oettinger, E. Rieken, M. Weiss (eds.). *Ex Anatolia Lux. Anatolian and Indo-European Studies in Honor of H. Craig Melchert on the Occasion on His Sixty-Fifth birthday*: 343–355. Ann Arbor: Beech Stave Press.
Wodtko, Dagmar S., Britta Irslinger, Carolin Schneider. 2008. *Nomina im indogermanischen Lexicon*. Heidelberg: Winter.
Woodhouse, Robert. 2006. Conditioned devoicing of mediae in Phrygian. *Studia Etymologica Cracoviensia* 11: 157–191.
Zgusta, Ladislav. 1984. *Kleinasiatische Orstnamen*. Heidelberg: Carl Winter.

Бартомеу Обрадор-Курсак. К вопросу о месте фригийского языка в индоевропейской семье

Основная цель статьи — исходя из наших текущих представлений о фригийском языке, собрать вместе те его особенности, которые могут быть релевантны для определения его генеалогической позиции в составе индоевропейской семьи. Оказывается, что большая часть этих особенностей объединяет его с греческим языком, что в целом согласуется с традиционными мнениями о близких связях между этими языками и об их возможном общем предке. Напротив, особые связи с такими другими индоевропейскими языками как армянский представляются не столь тесными, несмотря на наличие отдельных совместных изоглосс.

Ключевые слова: фригийский язык; индоевропейская диалектология; генетическая классификация языков; изоглоссы; протогреческий язык.

Александр Милитарев
Институт восточных культур и античности РГГУ; amilitarev@gmail.com

Этимология и лексическая реконструкция для древней и древнейшей истории ближневосточно-северноафриканско-средиземноморского ареала

В статье приводятся новые или обновленные афразийские этимологии и реконструированные праформы, которые могут способствовать реконструкции некоторых малоизученных и неизвестных аспектов, фрагментов и эпизодов древней и древнейшей истории Передней Азии, Северной Африки и Средиземноморья — таких, как ранее не выявленные лексические заимствования и их источники, указывающие на древние этнокультурные контакты; происхождение ливийского письма; представления, из которых предположительно развились такие разные концепты и образы, как египетское *Ка*, библейские *тоху ва-боху* и *шеол*; названия монстров в древних семитских культурах и реконструированные праафразийские термины, проливающие свет на социальный и эмоционально-этический аспекты жизни неолитического человека.

Ключевые слова: этимология; лингвистическая реконструкция; древняя история; афразийские языки; семитология; языковые заимствования; Библия.

Этимология и лексическая реконструкция могут пролить свет на некоторые аспекты древней и дописьменной этнокультурной истории, скорректировать устоявшиеся представления, стимулировать интерес историков и археологов к поиску новых или пересмотру имеющихся данных и интерпретаций, в частности, касающихся неисследованных этнокультурных контактов, на которые указывают выявленные лексические заимствования, иногда совершенно неожиданные. Цель настоящей статьи — не вдаваясь в теоретические и методологические основания такого общего утверждения, продемонстрировать конкретные возможности этимологического исследования в его разных проявлениях, в основном, на материале афразийских языков[1].

1. Берберское ‘железо' и проблема происхождения ливийского письма.

Одна из загадок берберской этимологии — происхождение ливийско-берберского[2] термина для железа: Old Lybian *zl*, Shilḥ *uzzal*, Wargla, Mzab *uzzəl*, Qabyle *uzzal*, Ghadames *wazal*, East Tawllemmet *uzzel*, Ahaggar, Ghat *t-az/ẓuli*, и т. д. < **ʔuzzāl*[3].

[1] Все примеры в статье, приводимые в принятой в семитологии и афразистике системе символов, содержатся в афразийской этимологической базе данных на сайте «Вавилонская башня» (starling.rinet.ru), составленной совместно с О. В. Столбовой (автором, в первую очередь, чадской части базы). В последние годы автор настоящей статьи продолжает дополнять и исправлять эту базу самостоятельно. Переводы примеров обычно цитируются на языках соответствующих двуязычных словарей и других источников.

[2] Об этой ветви афразийской макросемьи языков см. Айхенвальд & Милитарев 1991.

[3] В анлауте здесь мог бы быть и **ʕ*, но, во-первых, **ʔ* встречается в анлауте перед гласным во много раз чаще, а, во-вторых, в предполагаемом слове-источнике в анлауте *ʔ-*; по этой же причине, **ʔu-* здесь вероят-

У этого термина не находится ни внутренней берберской этимологии (например, переноса значения по цвету или еще какому-то признаку), ни параллелей в других афразийских языках, кроме западночадского хауса *ažiiloo* ‘iron band on spear shaft’, по-видимому, заимствованного из туарегских, и малонадежного соответствия позднеегипетского языка эллинистического периода *zn* ‘ploughshare' (если *n* < **l*), употребляющегося с детерминативом ‘металл’. Во втором примере, если связь все-таки есть, то египетский термин должен был быть заимствован из ливийско-берберских, но никак не наоборот, по соображениям и семантическим (‘железо' стало ‘железкой’), и хронологическим (термин явно реконструируется как праливийско-берберский, возникший до эллинистического периода египетского языка).

Традиционное объяснение заимствованием из семитского (в качестве источника заимствования приводится финик. *brzl* и соответствующее ему евр. *barzäl* בַּרְזֶל)[4] наталкивается на рутинную проблему ненадежных и сомнительных этимологий: невозможность объяснить фонетические изменения при заимствовании. В данном случае предполагаемое выпадение начального слога *bar* слова-источника или превращение первого семитского корневого согласного *b-* в берберский *w/u-* с полной ассимиляцией второго семитского корневого *-r* аналогичными процессами в берберском не подтверждается.

Вместе с тем никем не замеченным, насколько мне известно, остается объяснение ливийско-берберского термина одним пассажем из Иезекииля (27:19): *wə-dān wə-yāwān mə-ʔūzzāl... nātānnū barzäl ʕāšōt*...[5] (‘И Дан, и Йаван из Уззаля… платили (тебе)… выделанным железом…’). В словаре HALOT: 21 к топониму *ʔūzzāl*[6] дается следующий комментарий: “...in Arab. trad. the pre-Islamic name of *ṣanʕa*, capital of Yemen” [7].

The Anchor Bible dictionary (Freedman 1992) приводит следующую интерпретацию этого имени:

> UZAL (PERSON) [Heb *ʾûzāl* (אוּזָל)]. A son of Joktan and hence the name of a South Arabian tribe or place (Gen 10:27; 1 Chr 1:21). Since the time of Bochartus (1674: 130–34) this biblical name has usually been combined with ʾAzāl which, according to South Arabian tradition, is said to have been a Pre-Islamic name for Ṣanʿāʾ, the capital of Yemen. In fact, ʾAzāl is equated with Ṣanʿāʾ (al-Hamdān 1979: 193, 8), and the name occurs also frequently in poetry, e.g., in a verse of the Pre-Islamic poet ʾAlqama b. Dh Jadan (al-Hamdān 1979: 72, 8) or in a poem which is attributed to the Himyaritic king Asʿad Tubbaʿ (al-Hamdān 1979: 56, 5). According to a tradition not Shem, the son of Noah, but ʾAzāl, the son of Joktan, is said to have built Ṣanʿāʾ and her castle Ghumdān and to have called the town after his name (ar-Rāz 1974: 14–15). In the Sabean inscriptions, however, the later capital of Yemen is always referred to under the name ṣnʿw, Ṣanʿāw, and never under **ʾzl*, ʾAzāl, although the triliteral root *ʾzl* is attested in proper names, e.g., as epithet of the ancient town Waʾlān in the SE part of Yemen, *wʾln tʾzl*, Waʾlān Taʾzil (YMN 4, 2). According to one of the meanings of the Arabic verb *ʾazala*, “to restrain” the cognomen possibly expressed the wish that the fortified place should

нее, чем **w-* (*w* и *u* в берберских часто чередуются). Предложенный в праформе вокализм не только преобладает в производных берберских формах, но и полностью повторяет вокальную структуру слова-источника. Более убедительных аргументов привести не получается: афразийский (в т. ч. берберский) вокализм, в отличие от консонантизма, пока не поддается сколько-нибудь убедительной и универсальной систематизации.

[4] Об этом семитском корне см. Valério & Yakubovich 2010.

[5] Форма *ʕāšōt*, переводимая как ‘выделанное, обработанное’, дается в HALOT как слово с неопределенной этимологией. Возможно, что в масоретском тексте ошибка: вместо שׁ (shin) должно было стоять שׂ (ŝin), в этом случае слово является производным от глагола *ʕāŝā* ‘делать, выделывать’ и т.п.

[6] В HALOT приводятся оба варианта — с дагешем и без него; Biblia Hebraica Stuttgartensia в Ezekiel 27:19 приводит этот топоним с дагешем. Если наша интерпретация берберского **ʔuzzāl* верна, она подтверждает вариант с удвоенным *-z-*.

[7] Имеются и другие интерпретации и переводы этого стиха в целом и слова *ʔūzzāl* в частности, но толкование его как топонима представляется наиболее правильным.

prevent enemies to enter it; ʾAzāl as placename could be explained in the same way. Already Glaser (1890: 310) had pointed out that the hitherto accepted view of the identification of Uzal with Ṣanʿāʾ has to be given up. The Jews of Ṣanʿāʾ were probably the first who connected this town with a biblical name and called it ʾAzāl (Glaser 1890: 427). His proposal, however, to look for Uzal in the region of Yathrib/Mad na in the E Ḥijāz (Glaser 1890: 434), does not conform with our present historical knowledge of that region of Arabia.

On the other hand, no notice has yet been taken of the fact that aside from the use of ʾAzāl for Ṣanʿāʾ, the name ʾAzāl occurs two further times as a designation of places in Yemen (see al-Maqḥaf 1988: 26). The one ʾAzāl is a region in the district of ar-Raḍama, some 30 km E of Yar m; the other ʾAzāl is a region of the Banū ʿAmmār in the district of an-Nādira NE of Ibb on the upper course of the Wadi Banā. Both places are so far unattested in Sabean inscriptions since the areas referred to have not yet been investigated by archaeologists. In the last mentioned region there is the castle al-ʾAzāl which is currently in ruins but which was formerly among the famous castles of Yemen. Therefore it cannot be entirely excluded that this ʾAzāl might be the place in which one has to look for Uzal in the table of nations.

In Ezekiel 27, where the partners in the trade with Tyre and the merchandise imported into this commercial town are listed, MT *m ʾûzzāl* is attested (v 19). This might be interpreted as "spun yarn" if one interprets it as a Puʿal participle of the verbal root *ʾzl* (as a variant of **ʿzl*, "to spin"). Several Heb mss, however, offer the reading *m ʾûzāl* (LXX ex As l), which points to a place-name, "from ʾÛzāl". Making a conjecture at the beginning of the verse, which is presupposed by the LXX, the translation of the first half of the verse could read as follows: "and wine from Uzal they exchanged for your wares". Uzal has been identified with Izalla mentioned in Akkadian documents, a district NE of Mardin and Nisibin, especially since wine formed a part of the tribute paid to Ashurnasirpal by Izalla. By alteration of three waws into yods, the first three words of Ezek 27:19 have been interpreted and translated "and casks of wine from Izalla" (*m ʾ zāl*; see Millard 1962: 201–3). Since iron is mentioned in the second half in the verse, another (though less probable) proposal is that Uzal may be the Heb transliteration of the name Ušawalaš, known from Hittite documents as an Anatolian town in the region of modern Konya, where rich deposits of iron-ore exist; the translation of the unaltered MT would then be: "and Dān and Yāwān exchanged for your wares from Uzal iron …" (see Elat 1983: 323–30). In any case **ʿûzāl* in Ezek 27:19 is to be separated from *ʾûzāl*, the son of Joktan in the Table of Nations in the book of Genesis.

Совпадение берберского **ʔuzzāl* с библейским топонимом *ʔūzzāl* вкупе с упоминанием о торговле этого города железом дает возможность предположить происхождение ливийско-берберского термина для железа от названия города: перенос на название артефакта или любого товара названия места, откуда его доставляют или где его производят, — нередкое явление, ср., например, происхождение латинского обозначения меди *cuprum* от названия острова Кипр (Beekes 2009: 805). Что касается путей проникновения этого термина и обозначаемого им предмета из Южной Аравии в бербероязычный регион, то наиболее вероятных два — либо через Баб-эль-Мандебский пролив и далее через Нубию на север и северо-запад, либо через финикийские портовые города на средиземноморское побережье Африки. С одной стороны, ранние ливийско-нубийские контакты подтверждаются берберскими лексическими заимствованиями в нубийских языках, с другой стороны, 26 и 27 главы Иезекииля известны как «плач по Тиру», что указывает на финикийское посредничество.

Не вдаваясь в тонкости датировки железного века в Северной Африке, укажу на то, что широкое распространение железа на всем Ближнем Востоке вне Анатолии специалисты относят к XII–XI вв. до н.э. Если эта датировка обоснована, то она совпадает с полученной автором[8] датировкой дивергенции праберберского языка рубежом II–I тыс. до н.э.

[8] Данная датировка обоснована в ряде публикаций (напр., Милитарев 1991); в последней, обновленной версии — в подготовленной к печати статье *Once More on Tamâhaq Tuaregs in the Canary Islands in the Context of Ethno-Linguistic History (Linguistic and Inscriptional Evidence*) (предварительный вариант см. в Militarev 2018). Там же приведены аргументы В. Блажека в пользу более поздней праберберской дивергенции и мои контраргументы вместе с лексико-статистическими диагностическими списками по 13 берберским языкам, представляющим все основные группы и подгруппы этой языковой семьи.

Помимо весьма вероятного этимологического решения и вытекающего из него указания на торговые связи древних ливиоберберов с Южной Аравией, возникает еще одна интересная импликация, касающаяся совсем другой проблемы — происхождения ливийского письма.

В ряде работ, в частности, в статье, когда-то опубликованной в ВДИ (Милитарев 1991), я приводил аргументы — в виде сопоставительных таблиц знаков — в пользу того, что ливийское письмо отнюдь не входит в группу письменностей, в которую его включал И. Гельб и в которой «формы знаков изобретены произвольно» (Гельб 1982: 141), а в основе своей заимствовано из одной из разновидностей семитского «квазиалфавитного» письма (какие-то знаки могли, конечно, быть изобретены и добавлены «произвольно»).

Учитывая обширные финикийско-ливийские связи и тот факт, что название одного из вариантов ливийского письма — сохранившийся у туарегов *tifinaɣ* (с консонантным корнем *fnḳ*) — ассоциируется рядом исследователей с *Phoínikes*[9], естественно, считается, что этим видом семитского письма было финикийско-пуническое. Однако из приведенных таблиц в статье (Милитарев 1991) видно, что число знаков ливийского письма, сходных со знаками финикийско-пунического алфавита, меньше, чем число совпадений со знаками разных подвидов южносемитского письма[10]: с новопуническими символами совпадают три знака, с южносемитскими — семь, причем наибольшее сходство обнаруживается с северноарабским письмом[11]. Остальные ливийские символы (их девять) схожи как по форме, так и по фонетическому значению с символами представленными и в северной, и в южной разновидностях семитских алфавитов.

Было, однако, непонятно, какими путями южносемитская письменность могла стать известной ливийцам Северной Африки. Этимология берберского термина для железа позволяет предположить наличие торговых контактов между этими двумя регионами. А что касается эклектичного характера ливийского письма, возможно, оно заимствовало знаки из южной и северной семитской письменности в разные периоды.

2. Письмо из Египта — дописьменным соседям

На обширной территории северной Африки — от Канарских островов до Нубии — встречаются похожие слова, передающие идею письма. Удивительно, что эти слова засвидетельствованы в том числе и в бесписьменных и новописьменных языках. Интересно проследить происхождение приведенных ниже терминов и понять природу этого феномена.

---

[9] На то, что это вряд ли случайное совпадение, указывает и отсутствие альтернативной этимологии для *tifinaɣ*; однако единственным объяснением, каким образом туарегское письмо было названо «финикийским» по греческому этнониму (финикийцы так себя не называли), представляется то, что предки современных туарегов, которыми с высокой степенью вероятности были гараманты (см. Милитарев 1991), услышали о финикийском происхождении их письма от греков.

[10] Впервые предположение о происхождении ливийского письма из южносемитского, насколько мне известно, было высказано Энно Литтманом (Littmann 1904).

[11] А именно, самудским (Thamudic) и лихьянским (Lihyanite), которое сейчас принято называть даданитским (Dadanitic). Надписи самудским письмом, начиная с VIII тыс. до н.э., встречаются на огромной территории — от Йемена до Египта. См. The Online Corpus of the Inscriptions of Ancient North Arabia (http://krc.orient.ox.ac.uk/ociana).

В ливийско-берберских языках: эпиграфический восточно-нумидийский *tt-rb* и *t-rb-t-hn* — именные формы от глагола **rb* ‘писать’, ахаггар *t-êrəw-t,* pl. *t-êra* ‘lettre (missive); amulette consistant en un écrit; dessin d’ornement (consistant en lignes, points, figures géometriques)’, айр *t-irăw-t,* pl. *t-ira* ‘letter (message), amulet with an inscription on it’, гхадамес *ūrəḇ,* зенага *arha,* семлаль *ara,* кабильский *aru* ‘to write’, *t-ira* ‘writing’ и т. д. (см. берберские формы в Kossmann 1999: 88). Берберская праформа может быть реконструирована или как **Harah*^{w}, или как **Haraḇ* ‘писать’ (отглагольное имя **Hiraḇ*) — последний вариант подтверждается рефлексом *-b* в восточно-нумидийском.

В канарских: гран канариа *tara, tarja* ‘rayas en tablas, pared o piedras’, ‘señal para recuerdos’, *tarha* ‘marque pour les souvenirs’, ‘señal para recuerdos’[12].

В западночадских: хауса *rubùtā* ‘to write’ (сравнение приведено в Wölfel 1965: 511), в центральночадских: бура *rubwa, rubuta*, будума *rebōde,* афаде *ohárbotú,* etc. ‘to write’[13].

В канури: *revo-nin* ‘write’.

В древненубийском: *uran* ‘Schreiber’.

На первый взгляд, многие из этих форм не имеют между собой ничего общего, кроме согласного *r* и значений. Тем не менее, представляется возможным предположить их общий источник — египетский глагол *ḫ3b* (< **hrb*) ‘to send (a letter or message *inter alia*), to write a letter’, встречающийся в текстах Пирамид и давший существительное того же корня ‘letter, message’ в текстах Среднего и Нового Царств[14]. Из него, вероятно, заимствовано праберберское **Hiraḇ* ‘письмо, послание’ и **Haraḇ* ‘писать’[15]. Чадские формы представляют бóльшую сложность. Одна из них, афаде *ohárbo-tú*, являет удивительное свидетельство фонетической сохранности слова. Судя по форме, она была заимствовано из праберберского, как и хауса глагол *rubù-tā*, хотя оба слова теоретически могли заимствоваться и непосредственно из египетского (по глоттохронологии время выделения афаде — рубеж III–II тыс. до н.э., а хауса — конец IV тыс.). Формы в бура (*rubwa, rubuta*), будума (*rebōde*) и канури (*revo-nin*), скорее всего, заимствованы уже из хауса. Возможно, к берберским формам восходят, в конечном счете чадский глагол **rVn-* ‘to draw lines, to write’ (Stolbova 2005: 219) и древненубийское *uran* (одно из серии ранних заимствований в нубийские из ливийско-берберских[16]?), но *-n* в ауслауте объяснению пока не поддается.

Подводя итоги: египетское *ḫ3b* было заимствовано в праберберо-ливийский, а уже из берберских, скорее всего, распространилось по всей Северной Африке[17].

---

[12] Приведены в монументальном труде *Д.* Й. Вёльфеля (Wölfel 1965: 461-2), который, хотя и не нашел египетского источника, сопоставил канарские формы с большинством приведенных параллелей из берберских и других языков, из которых следует исключить баскский пример Вёльфеля: *ira-atsi* ‘añadir, escribir’, *ira-korri* ‘leer’, отвергнутый баскологами Р. *Л.* Траском и *Дж.* Бенгтсоном, поскольку *-ra-* в данных формах является показателем каузатива (я благодарен *Дж.* Бенгтсону за это уточнение). Есть еще подозрительно похожее на берберские формы древнегреческое τέρας (*téras*) ‘Götterzeichen’, ‘sign, emblem, wonder, monster’ — по мнению Фриска (Frisk 1960: 878), архаический термин, не имеющий надежных соответствий; я признателен индоевропеисту д-ру Коринне *Л*ешбер (Corinna Leschber) за консультацию по этой лексеме.

[13] Militarev 2018, #3.4.3.

[14] См. Erman & Grapow 1926–1961: II, 479–480.

[15] См. Милитарев 1984: 22–23. Гран канариа *tara, tarha* (< *ta-rha*, с префиксом имен *ta-*) скорее всего заимствовано из (прото)ахаггара (см. этимологию №8 в настоящей статье; развернутая аргументация приводится в Militarev 2018).

[16] См. Militarev 2018.

[17] Еще одно аналогичное и несомненное заимствование: праберберское **ta-matar-t* ‘sign of acknowledgement, recognition, testimony, etc.’ происходит из египетского термина Среднего Царства *mtr* ‘testimony, witness, instruction, etc.’ (см. Милитарев 1984: 22; Takács 2008: 733).

3. *Лосось*

Одно из самых распространенных в языках мира названий рыб восходит к латинскому слову *salmo,* gen. *salmōnis* (с кельтскими параллелями) без надежной индоевропейской этимологии, см., например, Andrews 1955: 310: “Salmo is probably either of Celtic origin or came from the pre-Indo-European Iberian of Aquitania. It is doubtful that the word has any connection with Latin *salire* ‘jump’, or with *saliva* ‘spit’, with a suggestion of sliminess and slipperiness”.

Мне нигде не попадалась ссылка на другую возможную этимологию латинского слова — из берберского **salm(-ay)*, мн.ч. **(i/a-)salm-an* ‘рыба’ (видовое название)[18]: ахаггар *a-sûlməy*, мн.ч. *i-sûlmäy-än*, зенага *siǯm-an* (< **silm-an*), снус, семляль, издег, кабильский *a-sləm*, риф *a-srʔm* (*r* < **l*), шенва *a-səlm*; этот термин засвидетельствован и на разных островах Канарского архипелага в виде *salema*.

Очевидно, что если латинское слово заимствовано из ливийско-берберских, то именно из формы **salm-an*[19]. Однако у этой формы есть варианты с метатезой корневых согласных: гхат *a-ləmšay* и гхадамес *u-lisma*. Из первого, по всей видимости, заимствовано хауса *lámsà* ‘вид рыб’, а второй представляет большой интерес тем, что указывает на более «глубокую» этимологию — связь с названием рыбы *nšmw.t* в египетском языке периода Нового Царства (Erman & Grapow 1926–1961: II, 340). Последняя форма с высокой вероятностью восходит к **lšmw.t* (иероглиф для *n* в египетском может передавать *n* и *l*): возможность случайных совпадений в афразийских корнях с тремя «твердыми» согласными и общим значением близка к нулю. Но и значение ‘рыба’ в *nšmw.t* оказывается вторичным — в более ранних текстах Среднего Царства это слово значит ‘чешуя’ (‘scales of fish’ — Faulkner 1962: 140), т. е. в более поздний период египетского оно, по-видимому, стало означать ‘чешуйчатую рыбу’. «Вторичность» египетской ‘рыбы’ указывает не на происхождение из общего праегипетско-берберо-чадского состояния, а на заимствование этого термина где-то во второй половине II тыс. до н.э. из новоегипетского в праливийско-берберский, после распада которого на ранние диалекты последовательность корневых согласных египетского слова-источника в большинстве этих диалектов изменилась, сохранившись только в гхадамесе (возможно, и в каких-то других восточноберберских)[20].

Таким образом, выстраивается следующая цепочка переходов значения и заимствований: египетское ‘чешуя’ → ‘чешуйчатая рыба’ → берберо-ливийская ‘рыба’ → латинское ‘лосось’.

---

[18] Впервые это предположение было высказано в Милитарев 1984: 23.

[19] Термин с такой последовательностью согласных заимствован (скорее всего, из какого-то туарегского языка) в центральночадские языки бура (*šalmwi* ’kind of fish’), мада (*ŝelémke* ’sorte de poisson sec plat’) и ульдем (*ŝēlémkē* ’poisson sec, ouvert’) — см. Stolbova 2007: 173.

[20] По-видимому, метатезы вызваны тенденцией к несовместимости в большинстве берберских языков между латеральными согласными *l* и *s* (< **ŝ*); последний (по установленному автором соответствию: берберское **s* ~ египетское *š* из праафразийского латерального сибилянта **ŝ*) еще в период заимствования произносился не как простой глухой, а именно как глухой *латеральный* сибилянт. Неисконная последовательность согласных в канарском названии рыбы указывает, скорее всего, на то, что оно не восходит к прaберберо-канарскому состоянию, а является поздним заимствованием из туарегского (см. Militarev 2018).

4. Египетская душа *Ка — Кур*?

Знаменитое египетское *Ka* (*kꜣ*), переводимое как ‘душа’, ‘двойник’, ‘дух’, ‘жизненная сила’ и т. п., не имеет надежной этимологии. Попытки найти сколько-нибудь убедительную внутреннюю египетскую этимологию результатов не дали. Версия, предлагаемая ниже, конечно, довольно гипотетична, но, тем не менее, заслуживает внимания.

Существуют серьезные основания[21] постулировать вокализацию *kꜣ* в виде *kuꜣ*. Однако фонетическая природа графемы *ꜣ*, которая может передавать три разных согласных — гортанную смычку *ʔ*, *r* и *l* (и, соответственно, восходить к трем разным общеафразийским согласным — **ʔ*, **r* и **l*), в этом слове неясна. Прояснить ее может только внешняя афразийская этимология. Никаких параллелей с консонантными корнями **kʔ* и **kl* и сколько-нибудь подходящим значением в других афразийских языках найти не удается. Зато находятся возможные и, на мой взгляд, достаточно интересные в плане семантики параллели с **kr*. Одна из них, впервые замеченная, по-видимому, египтологом и афразистом Карлтоном Ходжем (Carleton Hodge), заслуживает более детального обсуждения.

Хауса *kùrwā* ‘soul; ghost’ (Abraham 1962) приводится в более полном словаре Bargery 1934 в значении ‘the soul or personality which is supposed to leave a sleeper, returning when he awakes’[22] (корень существительного хауса *kur*, суффикс *-wa*, ср. *ḳane* ‘younger brother, sister’ vs *ḳán(u)wa* ‘young woman’). Если «души» в египетском и хауса, действительно, родственны, т. е. *kꜣ* соответствует **kur*, то вопрос о том, восходят ли оба термина к общеафразийскому корню или хауса заимствован из египетского, не совсем ясен. В пользу второго решения говорят очень специфическая семантика слова в хауса, почти полное[23] отсутствие его когнатов в чадских языках и, наконец, тот факт, что за последнее время выявлен[24] целый ряд египетско-чадских культурных терминов, не представленных в других афразийских языках и явно указывающих на древние этнокультурные контакты — скорее всего, египетские лексические заимствования в чадских.

Вполне вероятны, хотя и немногочисленны, и другие афразийские параллели: в берберских туарегских — ахаггар *tă-karaww-at*, айр *i-kerker-ăn* (редуплицированная форма мн. ч.) ‘épilepsie’, восточный тауллеммет *a-t-kər* (пассив глагола) ‘possessed (by an evil spirit)’, в семитских — араб. *kariya* (консонантный корень *kry*) ‘sommeiller’, тигре *kərri belä* ‘talk nonsense’ и *säb kəyar-ät* ‘soothsayers’[25].

Если все эти слова в конечном счете связаны и восходят к единому общесеверноафразийскому корню **kwr*/**kyr*/**kry*, его достаточно сложное значение следует, по-видимому, реконструировать как особое экстатическое состояние, «выход из себя», что может относиться к трансу исполнителя ритуала. В этом случае доисторические[26] истоки египетского *Ka* представляются весьма любопытными.

---

[21] См. Vycichl 1983: 74.

[22] На одной из конференций в 1980-е годы в московском Институте востоковедения АН я сделал сообщение с шутливым названием «Душа как *kurwa*: хаусанский двойник египетского *kꜣ*». Возможность читать *kꜣ* как [kur] была в ходе дискуссии подтверждена и, насколько я помню, даже поддержана петербуржским египтологом А. О. Большаковым, специально занимавшимся темой человека и его двойника в древнем Египте.

[23] Пока нашлись только две возможные параллели слову хауса в центральночадских языках бура (*kir* ‘the self; the will’) и гидар (*ma kro* ‘heart’). Другие возможные параллели: западночадские тангале *kɔrɔm* ‘divination’ и дотт *kártikáli* ‘tell fortune’ (в Stolbova 2011: 87 анализируются как дериваты **kVr-* ‘to see’).

[24] В основном, усилиями ведущего чадолога-компаративиста из Москвы О. В. Столбовой.

[25] Не исключено, что сюда же относится изолированное восточнокушитское слово в языке оромо: *ekeera* ‘ghost’.

[26] По моим глоттохронологическим подсчетам, северноафразийский праязык распался в 10 тыс. до н.э.

5. ʽМатеринская утробаʼ → ʽжалость, милосердиеʼ.

Одна из задач этимологического «детектива» — установить изменение значения от более раннего (пра)языкового состояния к более позднему; это особенно интересно, когда речь идет о дописьменной истории, да еще о категориях, связанных с человеческими эмоциями, для которых лексическая реконструкция может оказаться единственным источником информации. К таким категориям относится обсуждаемое ниже значение ‘милосердие’ в семитских языках.

Общесемитский глагол **rḥm* ‘жалеть, сострадать, быть добрым, милосердным к кому-либо’ и соответствующее ему существительное, восходящие к прасемитскому уровню (V тыс. до н. э. по глоттохронологии), происходят от общесемитского же анатомического термина **raḥ(i)m-/*riḥm-* ‘материнская утроба, матка’, который, в свою очередь, восходит к общеафразийскому **ra/im-*[27] ‘материнская утроба, беременность’, отраженному в чадских (как ‘ребенок’ — Stolbova 2016: 290; сюда же центральночадское: будума *olē-ram* ‘материнская утроба’, где *olē* ‘ребенок’) и в восточнокушитских: сомали *rim-ay* ‘матка’, *rim-an*, оромо *riimaa* ‘беременная (о самке животного)’, рендилле *rim-*, бурджи *rim-ees* ‘быть беременной’. Этот семантический переход известен: впервые, насколько я знаю, он был постулирован С. С. Майзелем как часть изосемантического ряда «матка (утроба, утробный плод, беременность) — жалость, сострадание, милосердие» в семитских (Майзель 1983: 221)[28]. Однако вероятность того, что любое постулируемое нетривиальное развитие значения, действительно, имело место, а не представляет собой случайное совпадение (приписанное ассоциативному мышлению древних людей лингвистами с современным, да еще и «европоцентристским» сознанием), существенно повышается[29], если удается продемонстрировать, что у реконструированного корня с предположительным исходным значением есть более широкая надежная этимология (афразийская)[30], т. е. он «древнее» (как в данном случае), а у корня с предположительным производным альтернативной этимологии нет, т. е. он «моложе».

6. «Вторые жены» эпохи раннего неолита

Одним из видов этимологической работы можно считать праязыковую реконструкцию, которая, как и в предыдущем примере, способна пролить свет на стороны жизни доисторического общества, трудно поддающиеся выявлению другими методами — такие, как, например, родственные и социальные связи, на которые, в качестве основного источника информации, экстраполируются сравнительно недавние этнографические данные; правомерность прямых параллелей между культурами, существенно различающи-

[27] Фарингальный *ḥ* в инлауте в семитских, не отраженный в сомали и рендилле, где он обычно сохраняется, судя по всему, вторичен (см. Takács 1997: 241–273).

[28] Я бы добавил к этому изосемантическому ряду еще один пример: семитское **ḫnn* ‘быть милосердным, жалеть’ из **ḫann-* ‘(feminine) genital organ, uterus’ и далее из афразийского **ḫanVn-* ‘genitalia’ (ibid.).

[29] Естественно, эта вероятность еще и тем выше, чем большее число синонимичных слов/корней («семантических рядов», по С. С. Майзелю) составляет данный «изосемантический ряд».

[30] Если гипотетически представить себе, что для семитского **raḥ(i)m-/*riḥm-* ‘материнская утроба, матка’ находится единственная афразийская этимология со значением ‘внутренности, кишки’, а для семитского **rḥm* ‘жалеть, быть добрым’ — единственная афразийская этимология со значением ‘хороший’ (или, скажем, ‘любить’), то такая ситуация не то, чтобы полностью исключила реальность нашего семантического перехода в прасемитском, но существенно ослабила бы его достоверность.

мися как хронологически — на многие тысячелетия, так и по уровню развития[31], не очевидна и требует подтверждения независимыми методами. Вот один пример.

Общесемитское **ki/ann-at-* (с суф. ж. р.) ‘невестка, невеста, прислуга, товарка’: аккад. (старовавилонский) *kinātu/kinattu* ‘menial, person of servile status attached to a household, doing agricultural and other work under supervision; person of equal social status, comrade, colleague’ (а также поздневавилонское и не вполне надежное *kinītu* ‘eine Nebenfrau’), арам. (считающееся аккадизмом) **kən-āt-* ‘companion’, араб. *kann-at-* ‘belle-fille, femme du fils; belle-soeur, femme du frère’;

общеберберское **ta-knaw* ‘co-wife, concubine’ (**i-knaw* ‘быть близнецами’, а в восточноберберском языке гхадамес *eknəw* — ‘ревновать’);

общечадское **k(ʷ)in-* ‘co-wife; virgin, maid; sister’, **kVn-*[32] ‘to marry’ (ср. Stolbova 2016: 208), ср. также в центральночадском языке гуду *kə̀-kə́nə* ‘grandmother or her siblings’, а в восточночадском языке тумак — *kùnáy* ‘orphan’;

центральнокушитское (агавское) **kiyan-* (и *keɣān* в северноагавском языке билин) ‘marriage’ и **(ʔə-)kʷin-* ‘woman’.

На афразийском уровне общую праформу, условно включая вокализм, можно реконструировать как **kʷayn-* с довольно сложным и несколько неожиданным для мезолита — раннего неолита значением, которое приблизительно описывается как ‘низкая социальная категория (молодых) женщин, которых берут в дом в качестве вторых жен, наложниц и/или работниц’[33].

Реконструкция сложных, не очевидных или плохо согласующихся с устоявшимися представлениями праязыковых значений для древнейшего и протописьменного периода в той или иной степени гипотетична. Приведенные семитские термины можно было бы разделить на две группы — ‘невестка, невеста’ (связь этих двух категорий подтверждается на материале многих языков — на русском, в частности) и ‘компаньон(ша), товарищ/ товарка, со-работник’ (и ‘прислуга’? или последнее значение скорее к ‘невестке’ — по хорошо известной народной традиции?) — и считать их восходящими к генетически разным корням, а берберские глаголы ‘быть близнецами’ и ‘ревновать’ или чадские существительные ‘бабка и ее сиблинги’ и ‘сирота’ вообще считать сюда никак не относящимися и не связанными между собой. Однако, явная принадлежность большинства рассмотренных терминов к женскому роду (или полу) и видимое отсутствие альтернативных этимологий для каждого из них препятствуют тому, чтобы «множить сущности», в данном случае — омонимы, и делают решение объединить их всех в один праафразийский корень с комплексным значением более обоснованным. Дальнейшее решение остается уже за историками, археологами, антропологами, этнографами и т. д.

---

[31] Например, вряд ли продуктивно прямо переносить черты архаичных культур охотников и собирателей, засвидетельствованные в новое или новейшее время, на неолитическую культуру прасемитов с ее развитыми земледелием, скотоводством, строительством жилищ и другими инновациями; даже подобная экстраполяция на мезолитическую-ранненеолитическую праафразийскую культуру XI–X тыс. до н. э. нуждается в обосновании.

[32] Знак *V* в праформе указывает на любой возможный гласный, означая, что конкретную артикуляцию (**a*, **i* или **u*) восстановить пока не удается даже с минимальной вероятностью.

[33] У этого реконструированного термина обнаруживаются обширные внешние параллели — скорее свидетельства очень глубокого родства, хотя и не доказуемого на сегодняшнем уровне развития компаративистики, чем случайные совпадения: ср. такие формы, как индоевропейское **gʷen-* ‘woman, wife’, алтайское **kune* ‘one of several wives’, синокавказское **qwänV*, австронезийское **kan*/**kɨn*, америндское **kuna*, банту **-kéntù* ‘woman’ в «Глобальной базе данных», составленной С. А. Старостиным (доступна в онлайн режиме на сайте «Вавилонская башня», http://starling.rinet.ru).

7. Одна этимология как вероятный индикатор миграции

Островные языки Канарского архипелага, судя по целому ряду морфологических и лексических изоглосс с берберскими языками, составляют с последними общую, берберо-канарскую семью, входящую в афразийскую макросемью; ряд отличий указывает на противопоставление берберской и канарской ветвей в пределах этой семьи. Между двумя ветвями устанавливается система регулярных консонантных соответствий[34], одним из которых является соответствие канарского *z (чаще всего передаваемого хронистами через *z*) общеберберскому *z, отражающемуся как [z] во всех живых берберских, кроме туарегских, где оно по разным языкам дает варианты *z*/s/š/ž/ǯ, а в ахаггарском регулярно отражается как *h*, например: канарское (о-ва Пальма) *zeloy* ‘soleil’, ‘el sol’ ~ берберск. **ā-zayl:* сиуа *a-zəl* ‘daytime’, семляль *a-zal* ‘sunlight’, изайан *a-zil* ‘heat of the day’, кабильск. *a-zal* ‘daylight’, айр *ə-žil* ~ ахаггар *a-həl* ‘sunlight’.

В ахаггар это «старое» *z* сохраняется при удвоении в определенных глагольных формах, ср. имперфектив *ighal* (<**əgzal*), но перфектив *gəzzul* ‘to be short’.

Наиболее правдоподобным представляется такое объяснение: на определенном этапе развития ахаггар *z* в речи стало «ослабевать» и постепенно переходить в *h*, но какой-то период сохранялись оба произносительных варианта слов — и с *z*, и с *h*; впоследствии за этими вариантами закрепилась функция морфонологического противопоставления двух глагольных форм.

В ахаггар с *h* из *z совпал *h* другого происхождения (<**h*)[35], причем следы разного происхождения сохраняются в позиции после *n*: перед *h* < *z *n* палатализируется и обозначается в транскрипции как *ñ*, а перед *h* < **h* не палатализируется. Так, ахаггарскому *te-ñhăr-t* (*h* <*z) ‘narine / nez’ соответствуют гхат *ta-nzər-t*, айр *čə-nžar-t*, указывая на праформу **ta-nzar-t*; а ахаггарскому *a-nhel* (*h* < **h*) ‘autruche’ соответствуют *a-nhil*, *a-năhil*, и т. д. в других туарегских, указывая на праформу **a-n(a)hil*.

В свете описанного фонетического явления в ахаггаре исключительно интересным представляется приведенное Вёльфелем слово с островов Тенерифе и Гран Канариа вместе с переводами на различные языки записавших его хронистов[36]: *añepa* ‘la asta que el rey llevaba del ante de sí’, ‘unas varas tostadas de tea y sabinas muy agudas’, ‘lanza de tea, que precedía al rey’, ‘gran lanza de tea fina, con una banderilla de juncos al extreme’, ‘lanza ó guión real’.

Вёльфель приводит также варианты этого слова в тенерифе: *anepa* ‘la lanza que el rey llevaba delante’, ‘a Scepter or Spear’, ‘étendard du roi’ и *anzpa*[37] ‘... una pertica sottile... ben lavorata, la quale era il segno reale’, находя к *añepa*/*anepa* нему ахаггарскую параллель: *ă-ñhəf* ‘batôn gros et long’[38].

---

[34] Установление этих соответствий осложняется тем, что канарские слова (их сохранилось около тысячи) воспринимались на слух, записывались, переводились и переписывались разными хронистами на разных европейских языках, да еще и в течение нескольких веков (а, например, в испанском в XVII в. произошли существенные изменения в консонантизме) — и многие из них дошли до нас в разных вариантах; с другой стороны, сопоставление этих вариантов помогает реконструировать фонетический облик канарских слов.

[35] В действительности, ситуация несколько сложнее (см. Militarev 2018), но в данном контексте это несущественно.

[36] Wölfel 1965: 477–478.

[37] Исключительно важный вариант, в котором Вёльфель напрасно заподозрил ошибку или опечатку.

[38] То, что в ахаггарском словаре Фуко (Foucauld 1952: 1348) приводятся размеры этой «палки» — от 0,3 до 0,5 см. в диаметре и от 1,5 до 1,8 м в длину — говорит о ее особой, по-видимому, некогда ритуальной

В работе Вёльфеля много отличных находок при сопоставлении канарских слов со словами из берберских, а иногда и других языков (хотя немало и фантастических сопоставлений). Однако будучи этнологом, а не лингвистом-этимологом, он не делает в данном случае решающего шага, заключающегося в том, что *ñ* в канарском *añera* происходит из палатализованного *n* с *h*, происходящим из *z, о котором говорилось выше. Эта гипотеза подтверждается родственным словом в другом севернотуарегском языке — гхат: *a-nžəf* ‘tison’[39]. Это означает, что канарское *añera* не родственно ахаггарскому *ă-ñhəf*, а просто то же самое слово, т. е. заимствование в канарский из ахаггарского, или какого-то ближайшего к нему диалекта, или их общего праязыка («протоахаггарского»), в котором имели место те же фонетические изменения (*z > *h*; **nz* > *ñh*). Вероятность того, что вся последовательность таких нетривиальных изменений могла произойти в неблизкородственных языках независимо друг от друга, практически исключена. И следующий логический вывод: это заимствование не объяснить ничем, кроме как нигде не засвидетельствованной миграцией ахаггароязычных туарегов на Канары.

Можно установить и относительные временны́е рамки этой миграции — время разделения ахаггара и того ближайшего к нему родственного языка, в котором *z переходит в *ž*, а не в *h*. Таким языком является входящий в севернотуарегскую группу гхат. Возможно сузить эти рамки до того начального периода, когда в ахаггарском еще свободно варьировались оба рефлекса *z > *z*/*h*, о чем свидетельствует почти чудом сохранившийся в тенерифе вариант *anzpa*. По глоттохронологии отделение гхата от языка-предка двух близкородственных *h*-диалектов, ахаггар и кель-уи, приходится на период VI–VII вв. н.э., что, видимо, и следует считать *terminus ante quem non* для протоахаггарской миграции на Канары.

Цель данного раздела статьи — продемонстрировать возможности даже одной этимологии для реконструкции неизвестного эпизода этнокультурной истории. Об остальных аргументах в пользу этой реконструкции и ее импликациях — обнаружении целой серии других «ахаггаризмов» в канарских диалектах, прочтении нескольких надписей ливийским письмом на о-ве Ферро (Йерро) с помощью ахаггарского словаря и др. — см. Militarev 2018.

8. *Tōhū-wā-ḇōhū*

Бывают случаи, когда этимология может пролить свет на многовековые богословско-философские дискуссии по поводу неясных библейских пассажей. Речь идет о знаменитом תֹּהוּ וָבֹהוּ (*tōhū-wā-ḇōhū*) из второго стиха первой главы Книги Бытия, переводимого как ‘[Земля же была] безвидна и пуста’ (Российское библейское общество и др. изводы), а в разных английских изданиях Библии как ‘without form, and void’ (New King James Version), ‘waste and void’ (American Standard Version), ‘unformed and void’ (Jewish Publication Society) и даже ‘[Earth was] a soup of nothingness, a bottomless emptiness, [an inky blackness]’ (The Message Bible).

Что дает этимология этих терминов?

функции. Общая праформа **a-nZef,* возможно, из **?a-n-Zef*; в этом случае она может быть родственна форме *šáafi* ‘размахивание поднятым копьем (в знак приветствия вышестоящего лица)’ в западночадском языке хауса и *šif* ‘палка-копалка’, *šif kingir* ‘(железный) жезл’ в центральночадском языке суккун (Stolbova 2009: 226; заметим «ритуальный оттенок» в приведенных значениях).

[39] Nehlil 1909; ср. также Prasse 1969: #541.

Еврейское *tōhū* переводится в словаре HALOT как ‘wilderness, wasteland, emptiness’, и в качестве параллелей закономерно приводятся угаритское *thw* ‘степь, пустыня’ и арабское *tīh-* ‘wilderness’, а также не имеющий к этому отношения постбиблейский еврейский глагол *tāhā* ‘to be rigid’, прокомментированный следующим образом: “A verbal form, together with a verbal form relevant to the derivation of *bōhū*, is found in Egyptian by Görg…: Egyptian *thʔ* ‘to deviate, miss’, Eg. *bhʔ* ‘to flee in panic’; Kilian…: through Egyptian there is a connection with Chaos, or rather interminableness, that is characteristic of the primaeval deities”. На самом деле егип. *bhʔ* ‘убегать в страхе’ не имеет никакого отношения к евр. *bōhū* (о его значении см. ниже), и тем самым теологема Килиана о связи библейского образа “через египетский с Хаосом — а пуще с *Бесконечностным*, что характерно для первозданных божеств” претенциозна и бессмысленна. Египетское *thỉ* (в другой транслитерации *thy*, но не *thʔ*) переводится как ‘to go astray, transgress, err, overstep’ и соответствует общесемитскому глаголу **tw/yh* ‘to (be afraid to) get lost, perish’ (вполне возможно, родственному вышеприведенным существительным со значением ‘пустыня’ — место, где страшно и можно заблудиться): сирийское *twh* ‘be alarmed, startled’, арабское *twh* ‘get lost (crossing the desert), wander about; be perplexed’, *tyh* ‘wander about, get lost’, геэз *tayyāhi* ‘fearful, terrified’, тигре *täwhä* ‘wander about’ (возможно, арабизм)[40].

Что касается еврейского *bōhū*, то его адекватный перевод в HALOT 111 как ‘emptiness, wasteness’ выводится почему-то (хотя и со знаком вопроса) из глагола *bhy* ‘удивляться’, представленного в постбиблейском иврите, который, в свою очередь, сравнивается с араб. *bhy* ‘быть пустым’. Конечно, не еврейский глагол ‘удивляться’, а библейское *bōhū* можно сравнивать с араб. глаголом *bhy* ‘to be empty (said of a furniture-free house)’ и именами *bāh-in* ‘empty’, *bahw-* ‘valley’ и ‘chest cavity’, а также с иудейско-арамейским глаголом в каузативе *h-bhy* ‘to clear (the field, thickets)’. Этимологическое значение еврейского *bōhū* — ‘пустое пространство’. Выводимое из приведенных форм довольно плохо сохранившееся центральносемитское **bahw-* с тем же значением подтверждается родственным корнем **bowh- ~ *bohw-* в кушитских: восточных (оромо *bowwā* ‘cliff, abyss, canyon, deep natural rift, gulley’ и дараса *bōwoʔà* ‘precipice, chasm’) и южных (**boohoo-nta* ‘pit, pitfall, hole in the ground’ и **boohoo-ngw* ‘valley, hole in the ground’).

Таким образом, исходя из этимологических значений и учитывая параллелизм *tōhū-wā-bōhū*, можно предложить следующий перевод Быт 1, 2: «Земля же была пустынна и пуста (или “пустыня и пустота”), и тьма над водной бездной/океаном» или же, более или менее следуя ритмике оригинала:

- - / - - / / - - / - - / - - / - /

«А земля была пуста и пустынна, над водной же бездной тьма».

Переводы *tōhū* вслед за ἀόρατος Септуагинты как ‘безвидный’, ‘unsightly’, ‘without form’, ‘unformed’, тем более ‘formlessness, confusion, unreality’ (не говоря уже о теологических шедеврах типа ‘soup of nothingness’ — буквально «суп из ничего»), не соответствуют значению этого слова, выводимому из его этимологии, и, по-видимому, нуждаются в пересмотре.

9. О локализации преисподней

В разных библейских текстах неоднократно упоминается шеол (*šəʔōl*) ˙преисподняя˙ как место, куда уходят умершие, однако сколько-нибудь внятное описание его отсутствует:

[40] У этого глагола есть когнаты и в других афразийских ветвях, все они имеют значение ‘потеряться, заблудиться, погибнуть’ (см. Militarev 2007).

складывается впечатление, что за ним стоит не фундаментальное представление о загробной жизни, а скорее мрачный поэтический образ, расхожая, но стертая метафора. Никакой убедительной этимологии в HALOT не предлагается, однако у этого термина есть правдоподобные семитские параллели с регулярными консонантными соответствиями. Наиболее вероятное этимологическое значение *šəʔōl*, не сохранившееся в самом библейском иврите, — ‘пропасть, котловина, русло потока’, ср. арабское *sāll-*, мн.ч. *sallān-* и *sawāll-* ‘gulch with steep slopes, the valley bottom, a current-bed’, *sayl-* ‘a current’, *syl* ‘to flow (of water), carry off (of a current); to be in trouble’, современное южноэфиопское (группа гураге): эндегень *səwel*, соддо *siyol* ‘cliff, abyss, precipice’ и современное южноаравийское: джиббали *sél* ‘to drain, run off; to rain’ (ср. еще аккадское *šalû/salû* ‘to submerge oneself (especially referring to the river ordeal)’).

У семитской праформы **šay/w/ʔal-* ‘gulch, precipice, current-bed’ есть параллели в западночадских (мангас *salà*, гурунтум *sál* ‘river’) и восточночадском языках (мокилко *sélò* ‘water reservoir, basin’).

Эта этимология косвенно подтверждается другим еврейским словом *bōr* — библейским параллелизмом к шеолу: «Но в ад (*šəʔōl*) ты низвергнут, в глубины преисподней (*bōr*)» (Ис 14:15); словосочетание *bōr šāʔōn* переводится в HALOT как ‘a pit of wasteland’ и означает преисподнюю[41]. Основные значения *bōr*, приводимые в HALOT, — ‘cistern’ (в т.ч. в значении ‘the entrance to *šəʔōl*’), ‘pitfall (often — deep hollow in rocky ground, used to store the water from the winter rains)’ и ‘grave’. В других семитских родственные слова значат ‘pit, cistern, well’ и ‘grave’, в других афразийских — берберских, чадских и восточнокушитских — слова, производные от корня **biʔVr- ~ *bVr-*, значат по языкам ‘dig; ditch, pit, well; grave’[42].

10. Названия монстров в семитских языках: сеанс мифологии с разоблачением

Ученому, борющемуся с мифологическими реликтами, в том числе в своем сознании, интересно понять, из чего рождаются фантастические образы и представления в мифологии и фольклоре. Не будучи специалистом ни в том, ни в другом, я, тем не менее, попытался эти образы и представления расклассифицировать по возможным источникам их возникновения (что, наверное, уже профессиональнее сделано другими авторами). Источниками фантастических и причудливых образов и представлений могут быть:

(1) индивидуальное воображение их создателей;
(2) сновидения;
(3) сознательное авторское творчество;
(4) трансформация вполне реальных объектов и феноменов
  (4.1) через мимолетные впечатления от непривычных и новых объектов;
  (4.2) через коллективную память, со временем обрастающую неправдоподобными характеристиками;
(5) народная этимология.

Если первые три источника образов могут быть объектами исследования различных областей психологии и когнитивных наук, то четвертый попадает в ведение исторических

[41] Интересно, что ту же идею передает арабское *dār-ul-bawār* ‘hell’ (дословно ‘house of pit’).

[42] Попутно замечу, что встречающееся в разных синонимичных афразийских корнях совмещение в одном корне значений ‘копать’, ‘яма’ и ‘могила’ ясно указывает на способ захоронения у праафразийцев — на всех праязыковых уровнях и, по-видимому, во всех языках-потомках.

дисциплин (от истории миграций, палеогеографии, этнозоологии и т. п. до мифологии и исторической лингвистики), а пятый целиком принадлежит этимологической науке.

Возможные примеры источника (4.2):

10.1. Семит. **ʔaṯ(h)al-* 'мифическая рептилия, дракон' ('a mythical reptile, dragon'). ◊ Арамейские: сирийский *ʔātalyā*, *ʔātehləyā* ʽdragon; eclipseʼ, мандейский *talia* ‘fictive dragon causing eclipse’[43]. Эфиосемитские: тигре *ʔashalät* ʽdragonʼ, тигринья *ʔasälät*, *ʔashalät* ʽanimale favoloso, di smisurata grandezza e della specie del coccodrillaʼ[44].

Афразийская реконструкция — **ʔač(h)ayl-/*haylač-* ʽкрупная рептилияʼ (ʽa large reptileʼ). ◊ Берберские: **Hassil-* ʽsnake, viperʼ: гхат *ašil* ʽserpentʼ, ахаггар *âššel* id., айр *aššel* ʽcoulevre, gros serpentʼ, мзаб *t-iššel-t* ʽvipèreʼ. Омотские: **haylaš-* ʽcrocodileʼ: северноомотские: волайта *aylaaš-*, гамо *haylašo*, гофа *haaylašo*, зала *aylašuwa*, дауро *allašo*; южноомотские: ари *haylɛša*[45].

Ср. Militarev & Kogan 2005, # 20.

10.2. Семит. **kVš-* ‘(мифическая водная) рептилия’. ◊ Аккадский (SB) *kušû* (Reiner & Roth 1961–2010: VIII, 602) ‘an aquatic animal’[46]. Эфиосемитские: геэз *kaysi* ʽserpent, dragonʼ, тигре *käyəs* id., тигринья *käyəs* ʽserpentʼ, амхарский *käysi* id.

Афразийская реконструкция — **kaw/ys-* ‘крупная опасная рептилия’. ◊ Чадские (западные): хауса *kwáasáa* ʽpuff-adderʼ (Stolbova 2011: 54), (восточные): бидийа *kóosò* ʽviperʼ (ibid.), кера *késé* ʽvaranʼ, леле *kōsō* ʽvaran de Nilʼ (ibid. #127), джегу *kesser* ‘kind of crocodile (small and brown)’, дангла *kɛ̀sírì* ‘crocodile’ (вероятно из **kVss-Vr-* с суффиксом *-r*).

10.3. Аккадское (начиная со староакк.) *laḫmu* (*laḫamu*) ‘a monster’, ‘ein mythisches Meerungeheuer’ (Militarev & Kogan 2005, #145).

Общесемитская реконструкция — **lVḫ(V)m-* ʽкрупное морское животноеʼ. ◊ Араб. *luḫm-* ʽkind of sea-fishʼ, диал. дасина *laḫam*, йеменский диал. *luḫam*, *laḫam* ʽsharkʼ. Современные южноаравийские: харсуси *léḫem*, джиббали *lḫum* ʽsharkʼ, сокотри *léḥem* ʽgrand poisson, requinʼ[47].

Афразийская реконструкция — **laḫ(m)-* ‘крупное водяное животное’. ◊ (?) Егип. *ḫrm*(*w*) (если < **ḫVlVm-*) ‘crocodile (?)’. Чадск. **lVḫV* 'hippo' (Stolbova 2005: 59): западночадские: дера *láhyò*; центральночадские: чибак *laḫa*, бура *laha*, гуде *láh*.

Возможные примеры источника (5):

10.4. ʽКозлорыбʼ, ставший ʽКозерогомʼ: шумерский ^{mul}*suḫur.máš*^{ku6}, дословно ʽрыбокозелʼ или ʽкозлорыбʼ (‘the Goat-Fish’) отождествляется с Козерогом (Capricornus).

Шумер. *suḫur.ku*$_6$ ʽвид карпаʼ (= аккад. *purādu* ʽcarpʼ, Reiner & Roth 1961–2010: XII, 516), ʽein Grosskarpfenʼ (von Soden 1965–81: 880); *suḫur.máš.ku*$_6$ ʽa kind of valued (carp)fishʼ

[43] Несмотря на явное сходство, по-видимому, не заимствовано из аккадского *attalû* ʽsolar or lunar eclipseʼ, как это принято считать, так как аккадский термин не связан с драконом, а сирийский часто обозначает дракона вне всякой связи с затмением (хотя значение ʽзатмениеʼ, скорее всего, возникло в арамейском под влиянием аккадского термина).

[44] Ср. возможно родственные геэз *nestāli* (*nesātāli*) 'serpent-idol of bronze; field snake' (LGz. 403), тигре *näsälät* 'serpent boa' (LH 235).

[45] Возможно, родственно также дахало (южнокушитский) *ṯáʔala* ʽpuff-adderʼ (если восходит к **čaʔal-*).

[46] В нескольких текстах — мифическое животное; по одной из интерпретаций, крупная хищная черепаха, по другой — акула (Militarev & Kogan 2005, #120). Ввиду эфиопских и афразийских параллелей аккадский термин вряд ли заимствован из шумерского KUŠÚ — скорее, наоборот.

[47] Заимствование в йеменские диалекты из южноаравийских (представленных и континентальными языками, и сокотри) более вероятно, чем в последние из арабского, учитывая конкретное значение в них (‘акула’) против более общего (‘какая-то морская рыба’) в классическом арабском.

(= аккад. *bitrû* 'outstanding, superb', Reiner & Roth 1961–2010: II, 279). В мифологии — персонаж из окружения Энки-Эа (Reiner & Roth 1961–2010: XVII, 351), изображаемый на камнях *kudurru* и печатях в виде козы с рыбьим хвостом; тот же образ выступает как символ созвездия. Возник он, скорее всего, не из воображения древних, а из их этимологических фантазий[48]. Именно второе значение *máš* = *uriṣu*, 'male goat' породило «козлиный» компонент образа, превратив его, в конечном счете, в «козлорыба». Аккадское *suḫurmāšu* 'The Goat-Fish' заимствовано из шумерского. // Ср. Kurtik & Militarev 2005, #25.

10.5. 'Поле', ставшее 'конем': шумерское ^{mul}*AŠ.GAN$_2$*, ^{mul}*AŠ.IKU*, ^{mul}*iku* 'a unit of measurement (of field)' обычно переводится просто как 'field'. В качестве созвездия отождествляется с областью современных созвездий Пегас и Андромеда.

Предположительно, греки интерпретировали *(aš).iku* как *ikû* по-гречески (вероятно, еще в микенский период), т. е. как 'конь', ср. линейное B *i-qo* 'horse' из индоевропейского **ek'wo-*; последняя форма должна была бы отражаться примерно как *iku ~ eku ~ iko ~ eko* в любом другом кентумном индо-европейском языке, который тоже мог оказаться посредником в формировании этой народной этимологии.

Естественно, гарантировать, что это пример именно народной этимологии, а не случайное совпадение, невозможно: данная интерпретация строится только на языковых аргументах, тогда как подобные замены в названиях созвездий принято объяснять через визуальные ассоциации, вызываемые конфигурациями звезд, или через связь ассоциированных с этими созвездиями божеств. Кроме того, необходимо отметить, что в картине месопотамского звездного неба имелось и собственное созвездие Коня (^{mul}*ANŠE.KUR.RA*) с очень спорной идентификацией: так, Пингри и другие авторы помещают его в пределах созвездия Кассиопеи, а Кох локализует его внутри Пегаса (если последнее верно, греческое название — перевод месопотамского); как бы то ни было, надежных данных по локализации или идентификации ^{mul}*ANŠE.KUR.RA* до сих пор не представлено. // Ср. Kurtik & Militarev 2005, #25.

10.6. 'Летучая рыба', ставшая 'Пегасом'.

Происхождение этого образа неясно. Традиционная этимология имени — от *pēgé* (πηγή) 'spring, fountain, fountain fed by a spring' (Chantraine 1968–1980: III, 894) — во-первых, скорее, народная и, во-вторых, ассоциируется только с одним эпизодом из мифов о Пегасе и его «крылатости» не объясняет. Предложенная связь Пегаса с лувийским богом-громовержцем *Pihaššašši* в качестве «аватары» последнего (Beekes 2009: 1183) тоже не объясняет, откуда взялся образ коня с крыльями. Ответ, похоже, находится между небесной сферой и пытливым ассоциативным мифологическим сознанием.

Шумерское созвездие ^{mul}*sim.maḫ* 'The Big Swallow' локализуется в западной (или юго-западной) области современного созвездия Рыб, включая западную область созвездия Пегас. По-аккадски это созвездие — *sinuntu* (и *šinūntu*) 'ласточка' (из общесемит. **su/inūn-at* id.), вторичное значение 'летучая рыба', ср. англ. 'swallow-fish'. Связь с водой подтверждается идентификацией в списке названий звезд: шумерское *mul.id$_2$.buranum* = аккад. *sinun-tum*, а *id$_2$.buranum* по-шумерски — Евфрат. Греки, по всей видимости, заимствовавшие из Вавилона как представление об этом созвездии, так и его название, перевели его как *khelidonías ikhthus* 'the swallow-fish'. Эта греческая 'ласточка-рыба' и ассоциация знакомого грекам шумеро-аккадского созвездия Ласточки с греческим созвездием Пегаса, возможно, проливают свет на происхождение образа крылатого коня. // Ср. Kurtik & Militarev 2005, #23.

[48] Подробнее об этом интереснейшем и малоизученном феномене, который я называю этимопоэтикой, см. Militarev 2010, Appendix I.

В заключение хотелось бы отметить следующее. Опыт показывает, что у «массовых» (разумеется, не в смысле «mass comparison») компаративистов и этимологов, т. е. специалистов, которые на хорошем профессиональном уровне реконструируют праязыки, составляют этимологические словари и большие базы данных, разбираются с генетической классификацией целых языковых семей и т. п., как правило, уже не остается ни достаточного времени, ни сил на приложение своих реконструкций и этимологий к историческому и, тем более, доисторическому материалу, включая сравнение лингвистических данных с данными археологии, популяционной генетики, физической антропологии, мифологии и других родственных дисциплин. Тем самым совершенно недостаточно используется мощнейший, а иногда и единственный ресурс расширения и верификации наших знаний о человеческом прошлом. Хотелось бы надеяться, что данный скромный опус послужит стимулом для подобного рода исследований (как в афразистике, так и в других областях исторического языкознания), которые в ближайшем будущем должны получить более широкое распространение[49].

Литература

Айхенвальд, А. Ю., А. Ю. Милитарев. 1991. *Ливийско-гуанчские языки*. В: И. М. Дьяконов, Г. Ш. Шарбатов (ред.). *Языки Азии и Африки. Афразийские языки. Т. IV, кн. 2*: 148–267. Москва: Наука. 1991.

Гельб, И. Е. 1982. *Опыт изучения письма*. Москва: Радуга.

Майзель, С. С. 1983. *Пути развития корневого фонда семитских языков*. Москва: Наука.

Милитарев, А. Ю. 1984. Современное сравнительно-историческое афразийское языкознание: что оно может дать исторической науке? В: И. Ф. Вардуль (ред.). *Лингвистическая реконструкция и древнейшая история Востока. Тезисы и доклады конференции. Часть 3*: 3–26. Москва: Главная редакция восточной литературы.

Милитарев, А. Ю. 1991. Глазами лингвиста: Гарамантида в контексте североафриканской истории (вместо послесловия). *Вестник древней истории* 3(1991): 130–158.

Милитарев, А. Ю. 2007. Значение этимологии для интерпретации древнеписьменных текстов (на примере еврейской Библии и Нового Завета). *Труды отделения историко-филологических наук РАН 2006*: 284–327.

References

Abraham, Roy Clive. 1962. *Dictionary of the Hausa Language*. University of London Press.

Aikhenvald, Alexandra, Alexander Militarev. 1991. Livijsko-guanchskije jazyki. In: Igor Diakonov, Grigory Sharbatov (eds.). *Jazyki Azii i Afriki. Afrazijskije jazyki. Vol. IV/2:* 148–267. Moskva: Nauka. 1991.

Andrews, Alfred C. 1955. Greek and Latin Terms for Salmon and Trout. *Transactions and Proceedings of the American Philological Association* 86: 308-318.

Bargery, George P. 1934. *A Hausa-English Dictionary and English-Hausa Vocabulary*. London: Humphrey Milford.

Beekes, Robert S. P. 2009. *Etymological Dictionary of Greek*. Leiden: Brill.

Chantraine, Pierre. 1968–1980. *Dictionnaire étymologique de la langue grecque: histoire de mots*. Paris: Klincksieck.

Erman, Adolf, Hermann Grapow. 1926–1961. *Wörterbuch der ägyptischen Sprache*. Berlin.

Faulkner, Raymond O. 1962. *A Concise Dictionary of Middle Egyptian*. Oxford: Griffith Institute.

de Foucauld, Charles. 1952. *Dictionnaire touareg-français*. Paris: Imprimerie nationale.

Freedman, David Noel. 1992. *The Anchor Bible dictionary*. *Vol. 6*. New York: Doubleday.

Frisk, Hjalmar. 1960. *Griechisches etymologisches Wörterbuch. Band I*. Heidelberg: Winter.

Gelb, I. E. 1982. *Opyt izuchenija pis'ma*. Moskva: Raduga.

HALOT = Köhler, Ludwig, Walter Baumgartner. 1994–1996, 1999–2000: *The Hebrew and Aramaic Lexicon of the Old Testament*. *Vols. I–V*. Leiden / New York / Boston: Brill Publishers.

[49] Автор выражает искреннюю благодарность В. Л. Цукановой за ряд содержательных и стилистических замечаний, позволивших улучшить итоговый текст статьи.

Kossmann, Maarten. 1999. *Essai sur la phonologie du proto-berbère*. Köln: Rüdiger Köppe Verlag.
Kurtik, Gennadij, Alexander Militarev. 2005. Once more on the origin of Semitic and Greek star names: an astronomic-etymological approach updated. *Culture and Cosmos* 9(1): 3–43.
Littmann, Enno. 1904. L'origine de l'alphabet libyen. *Journal Asiatique* 2: 423–40.
Maizel, Solomon. 1983. *Puti razvitija kornevogo fonda semitskix jazykov*. Moskva: Nauka.
Militarev, Alexander. 1984. Sovremennoje sravnitel'no-istoricheskoje afrazijskoe jazykoznaije: chto ono mozhet dat' istoricheskoj nauke? In: I. F. Vardul (ed.). *Lingvisticheskaja rekonstrukcija i drevnejshaja istorija Vostoka. Tezisy i doklady konferencii. Chast' 3*: 3–26. Moskva: Glavnaja redakcija vostochnoj literatury.
Militarev, Alexander. 1991. Glazami lingvista: Garamantida v kontekste severoafrikanskoj istorii (vmesto posleslovija). *Vestnik drevnej istorii* 3(1991): 130–158.
Militarev, Alexander. 2007. Znachenije etimologii dl'a interpretacii drevnepis'mennykh tekstov (na primere jevropejskoj Biblii i Novogo Zaveta). *Trudy otdelenija istoriko-filologicheskix nauk RAN 2006*: 284–327.
Militarev, Alexander. 2010. *The Jewish Conundrum in World History*. Boston: Academic Studies Press.
Militarev, Alexander. 2018. *Berber — Tuaregs — Canarians*. Ms., available online at: www.academia.edu/37897447.
Militarev, Alexander, Leonid Kogan. 2005. *Semitic Etymological Dictionary. Vol. II*: *Animal Names*. Münster: Ugarit-Verlag.
Nehlil, Mohammad. 1909. *Etude sur le dialecte de Ghat*. Paris: Editions Ernest Leroux.
Prasse, Karl G. 1969. *A propos de l'origine de H touareg (tahaggart).* Copenhagen: Royal Danish Academy of Science and Letters.
Reiner, Erica, Martha T. Roth (eds.). 1961–2010. *The Assyrian Dictionary of the Oriental Institute of the University of Chicago*. Chicago: Oriental Institute.
von Soden, Wolfgang. 1965–1981. *Akkadisches Handwörterbuch*: *unter Benutzung des lexikalischen Nachlasses von Bruno Meissner*. Wiesbaden: Harrassowitz.
Stolbova, Olga V. 2005. *Chadic Lexical Database. Issue I: L, N, NY, R*. Kaluga: Poligrafiya.
Stolbova, Olga V. 2009. *Chadic Lexical Database. Issue III: Sibilants and sibilant affricates*. Moscow: Institute of Oriental Studies.
Stolbova, Olga V. 2011. *Chadic Lexical Database. Issue IV: Velars*. Moscow: Institute of Oriental Studies.
Stolbova, Olga V. 2016. *Chadic Etymological Dictionary*. Moscow: Institute of Oriental Studies.
Takács, Gábor. 1997. The Common Afrasian Nominal Class Marker *ḥ. *Studia Etymologica Cracoviensia* 2: 241–273.
Takács, Gábor. 2008. *Etymological Dictionary of Egyptian. Vol. 3*. Leiden / Boston: Brill.
Valério, Miguel, Ilya Yakubovich. 2010. Semitic word for 'iron' as an Anatolian loanword. In: T. M. Nikolaeva (ed.). *Issledovanija po lingvistike i semiotike: sbornik statej k jubileju Vyach. Vs. Ivanova*: 108–116. Moscow: Languages of Slavonic Culture.
Vycichl, Werner. 1983. *Dictionnaire étymologique de la langue copte*. Leuven: Peeters.
Waldbaum, Jane C. 1978. *From Bronze to Iron: The Transition from the Bronze Age to the Iron Age in the Eastern Mediterranean*. Göteborg: Paul Astroms Forlag.
Wölfel, Dominik J. 1965. *Monumenta linguae canariae*. Graz: Akademische Verlagsanstalt.

Alexander Militarev. Some implications of etymology and lexical reconstruction for the history and pre-history of the Near Eastern / North African / Mediterranean areal.

In this paper, we propose several new or updated Afro-Asiatic etymologies and reconstructed protoforms which may assists in further reconstruction of certain unclear aspects and features of the history and prehistory of the Near East, North Africa, and the Mediterranean area — such as ancient ethnocultural contacts suggested by earlier undisclosed lexical borrowings; the possible origins of the Libyan script; ideas that may have given rise to such diverse concepts as Egyptian *Ka* or Biblical *tohu wa-bohu* and *sheol*; names for various monsters in ancient Semitic cultures; and certain social, emotional, and ethical aspects of Neolithic life hinted at by the respective reconstructed Afro-Asiatic protoforms.

Keywords: etymology; linguistic reconstruction; ancient history; Afro-Asiatic languages; Semitology; the Bible.