

Notes on the historical phonology of Indo-Iranian loanwords in Northwestern Tibetan dialects

Recent research has shown that the vocabulary of certain Northwestern Tibetan dialects contains a significant number of Indo-Iranian loanwords. It is, however, still unclear if these loanwords have been borrowed from a single or from several sources, and whether their presence is the result of substratum or adstratum interference. Likewise, the exact genetic position of the donor-language(s) within the Indo-Iranian group so far remains undetermined. The study of all these issues should, no doubt, be based on facts of historical phonology. In this article I attempt to identify the most conspicuous historical-phonological features of the Indo-Iranian elements found in Northwestern Tibetan varieties. Furthermore, I make some preliminary conclusions concerning the linguistic geography of the region in the pre-Tibetan period, as well as the direction and relative chronology of Tibetan migrations.

Keywords: historical phonology; language contact; linguistic substratum; Indo-Iranian languages; Dardic languages; Tibetan dialects; Ladakhi language; Balti language; Purik language.

A recent etymological study of the vocabulary of certain Northwestern Tibetan varieties, namely Ladakhi, Balti and Purik, has revealed the existence of a significant Indo-Iranian lexical stratum in these dialects (Kogan 2019). This discovery is to some extent in line with the hypothesis put forth more than a century ago by the German scholar August Hermann Francke, and still popular among Tibetologists, according to which the pre-Tibetan population of Ladakh and adjoining areas was somehow related to the present-day speakers of Dardic languages (Francke 1907). It should, however, be kept in mind that the overall picture of ethnic and linguistic history of the region is still far from clear. The presence of Indo-Iranian loanwords in the Tibetan dialects of Ladakh and Baltistan raises more questions than it answers. We do not know if these loanwords have been borrowed from a single source or several sources, whether their existence is the result of substratum or adstratum interference,¹ and finally, which branch or branches of Indo-Iranian they represent.

It goes without saying that the answers to these questions must be based on facts of historical phonology. The only possible way to establish the exact genetic position of the donor language is to compare its historical-phonological peculiarities with those of all the four known branches of the Aryan subfamily, namely Indo-Aryan, Iranian, Dardic and Nuristani. The strongest indication of borrowing from several related lects is, no doubt, the presence of more than one type of phonological development in the same position of the word, inexplicable by secondary processes like analogical changes. The choice of substratum or adstratum alternative can hardly be made either without taking into account historical-phonological data. In our case, such a choice is actually a complex task. The present-day Indo-Iranian neighbors of Northwestern Tibetan dialects are two Dardic languages: Kashmiri and Shina.² There are

¹ Based on our present knowledge of language situation in the Tibetan Empire and later northwestern Tibetan kingdoms, the influence of an unknown Indo-Iranian superstrate on Tibetan dialects should be considered improbable.

² The Ladakhi and Purik dialects are also in contact with Brokskat, a Dardic language genetically close to Shina. However, as shown in the above-cited article (Kogan 2019), the speakers of this language must be comparatively recent migrants to their present habitat.

strong reasons to believe that both of them spread to their present-day speaking areas in the Middle Ages due to ethnic migrations. The population of Kashmir until the 13th century spoke an Indo-Aryan language, most likely closely related to West Pahari dialects (Kogan 2016), whereas in the now Shina-speaking valleys of Karakoram the language spoken one thousand-plus years ago was probably Burushaski (Jettmar 1975). It should be born in mind that Indo-Iranian influence on Tibetan varieties under study, whatever its nature, cannot be exactly dated, even though the Tibetan conquest of Ladakh and adjoining areas in the 8th century A.D. provides a *terminus post quem* for the process. In light of this fact, it cannot be ruled out that at least a part of loanwords were borrowed in the 2nd millennium A.D. directly from Shina and/or Kashmiri. Another possible source of lexical loans may have been East Iranian Saka dialects spoken before the 11th century A.D. in the western part of the present-day Xinjiang, immediately to the north-east of Ladakh and Baltistan.³ To sum up, borrowing from an adstrate language should be considered probable if a particular Indo-Iranian word shows Indo-Aryan, East Dardic (Shina or Kashmiri)⁴ or Iranian historical-phonological traits.

In the following sections an attempt will be made to identify the most conspicuous historical-phonological features of Indo-Iranian elements found in Northwestern Tibetan dialects.⁵ First, I shall address the development of vowels, syllabic sonorants and some sequences containing vowels and sonorants, then the development of single consonants and consonant clusters. Finally, some preliminary conclusions will be drawn.

Vowels and syllabic sonorants

Vowels (monophthongs and diphthongs)

There is no phonological vowel length in Northwestern Tibetan dialects. That is why PII **a* and **ā* are usually reflected in the same way:

1. Ladakhi *asur* ‘mustard seed as a spice used in Ladakhi pickle’ (Norman 2010: 1073).⁶ Cf. OIA *āsūrī* ‘*Sinapis ramosa*’, Sindhi *ahuri* ‘mustard seed’, Lahnda *ôhur*, *ahūr*, *āhur*, Punjabi *āhur*, Hindi-Urdu *āsūrī* ‘mustard’, Kashmiri *āsoru* ‘plant, *Sinapis ramosa*’.

2. Balti *basanda* ‘dandelion’ (Sprigg 2002: 27). Cf. OIA *vāsanta-* ‘vernal, pertaining to spring’, OIA *vasanta-*, Shina *bazōn*, Phalura *basānd*, Bashkarik *basan*, Torwali *basān*, Kalasha *bāsun*, Khowar *bosun*, Pashai *wahón(d)*, Gawar-Bati *wasand* ‘spring’.

3. Balti *bat* ‘boiled mixture of germinated grain flour and ordinary flour (made during Ramzan)’ (Sprigg 2002: 27). Cf. OIA *bhakta-* ‘food; boiled rice’, Khowar *bot* ‘evening meal’, Kalasha *batay* ‘flour taken to be eaten in high pastures’ (Trail 1999), Shina *bat*, Bashkarik *batt*, Torwali *bāt*, Kashmiri *batī*, Lahnda, Punjabi *bhatt*, Hindi-Urdu, Nepali, Gujarati, Marathi *bhāt* ‘boiled rice’, Wakhi *bat* ‘wheat flour gruel’, Ishkashimi *bat*, Shughni *bāt*, Munji *bātək* ‘ritual food made of wheat flour and butter, cooked in water or milk’.⁷

³ It is, however, worth noting that a tentative attempt to find Iranian phonological features in Ladakhi, Balti and Purik words of Indo-Iranian origin made in Kogan 2019 was not successful.

⁴ The existence of the East Dardic subbranch consisting of Kashmiri, Shina and Kohistani languages was first hypothesized by George Abraham Grierson (1906; 1919). For the latest views on this group and its historical-phonological features see Kogan 2015; Kogan 2016.

⁵ The material studied here as well as lexical comparisons are drawn chiefly from Kogan 2019, where the respective etymologies are also discussed. If a word not mentioned in this article is analyzed, its etymology will be discussed in the present text.

⁶ For ease of reference, etymologies are numbered consecutively.

⁷ The above-cited East Iranian words are most probably borrowed from some Dardic source.

4. Balti, Purik *chal* ‘overflow, spill over’, Ladakhi (Leh and Shamskat dialects) *chal-ces* ‘to splash, to spill over’ (Sprigg 2002: 41; Norman 2010: 297). Cf. OIA *kṣarati*, *kṣalati* ‘flows, trickles’, *kṣālayati* ‘washes’, Proto-Iranian **xšar-* ‘to flow’ (> Persian *šārīdan* ‘to trickle’, (*āb*)*šār* ‘waterfall’, Ossetic *āxsārdzān* ‘waterfall’ (< **xšar-čana-*), Middle Persian *Xšart* ‘the river Jaxartes’ (Livshits 2003)), Kashmiri *čhalun* ‘to wash’, *čhar* ‘a sprinkle of water etc. from the fingers’, Gawar-Bati *čhār* ‘rapids in a stream’, Phalura, Indus Kohistani *čhār*, Bashkarik, Kalasha *uḥār*, Shina *char* ‘waterfall’.

5. Ladakhi *darak*, Purik *deraq* ‘stiff, hard’ (Norman 2010: 438) < PII **dhāraka-*. Cf. Kashmiri *dor* ‘firm, hard, strong, compact, durable, solid’, *darun* ‘to become steady (of something in motion), to become firm, to stand steady’ (Grierson 1915–1932: 238, 246), Pashai *d(h)ar-*, Wotapuri *dar-*, Phalura *dhāara-* ‘to remain, stay’, Torwali *dērī* ‘they remained’, Khowar *dorik* ‘hold back, wait, keep’, Hindi, Punjabi *dharnā* ‘to keep’, Gujarati *dharvū* ‘to hold, catch’, OIA *dhārayati*, *dharati* ‘holds, keeps’, Av. *dāraiehi* ‘(you) hold’.

6. Balti *gzar* ‘to flow’ (Sprigg 2002: 72), Purik *zar* ‘id.’, Ladakhi *zar-ces*, *dzar-cas* ‘to drip, to run down, to trickle out’ (Norman 2010: 790). Cf. Av. *γžar-* ‘to flow’, Ossetic *ğzælyn* ‘to pour down, drip’, OIA *jhara-* ‘waterfall’, *jhārī* ‘river’, Prakrit *jharāi* ‘drips’, Hindi *jharnā* ‘to ooze, trickle away’ < PII **gʷjhar-* < PIE **dʰgʷher-* (Cheung 2007: 124) or **gʷgh’er-* (LIV: 213–214).

Etymological **u* and **ū* also merge:

7. Balti, Purik, Ladakhi *mulṭuk*, *multuk* ‘fist’ (Norman 2010: 705; Sprigg 2002: 118) < **mulṭak* < **muṣṭaka-*.⁸ Cf. OIA *muṣṭi-*, Av. *muṣti-*, Khotanese *muṣtu*, Sindhi *muṭhi*, Lahnda, Punjabi *muṭṭh*, Hindi-Urdu, Gujarati, Marathi *mūṭh*, Nepali *muṭhi*, Persian *mošt*, Shughni *mut*, Wakhi *məst*, Shina *muṭ(h)*, Kashmiri *mōṭh*, Phalura, Kalasha, Khowar *muṣṭi*, Gawar-Bati *muṣṭāk*, Burushaski (loanword) *muč* ‘fist’.

8. Balti *mulak*, *mulu*, Purik *mulaq* ‘turnip’ (Sprigg 2002: 118; Zemp 2018: 944). Cf. Burushaski *múlo*, Shina *muúlo* ‘id.’, OIA *mūla-* ‘root’, *mūlaka-* ‘radish’, Shina *mūlī*, Khowar *mūl* ‘root’, Pashai *mūluk*, Hindi-Urdu *mūlī* ‘radish’ < PII *mūla-* ‘root’.⁹

No reflexes of PII initial or medial **i* and **ī* have been attested in our material. There is, however, one instance of the drop of **ī* in the word-final position (see 1.).

As we can see, Balti shows double reflexes of PII word-final **a*. This vowel is sometimes preserved (cf. *basanda* ‘dandelion’) and sometimes dropped (cf., e.g. *bat* ‘boiled mixture of germinated grain flour and ordinary flour (made during Ramzan)’). It cannot be ruled out that the difference in reflexes is due to different borrowing sources.

In one example, we find an unusual vowel sequence *aa* in Balti, which may correspond to short *a* in Old Indo-Aryan as well as in Iranian and Dardic:

9. Balti *baan* ‘man or men who sing religious songs and foretell the future’ (Sprigg 2002: 24). Cf. Kashmiri *wan-*, Indus Kohistani *ban-* ‘to say’, Kalasha *bandek* ‘1. to teach; 2. To announce so as to inaugurate, to order or command’ (Trail 1999: 27), OIA *vandate* ‘praises, worships’, Av. *vaṇd-*, Khotanese *van-*, Parthian *wynd-* ‘to praise, honor, worship’.

It is, however, quite possible that the source of the Balti word reflects some lengthened-grade derivative of the above-cited root. Such derivatives are attested in several Dardic languages. Cf., e.g. Wotapuri *-bān* ‘*sprechend*, *Sprecher*’ in *alīk-bān* ‘*Lügner*’ (*alīk* ‘*Lüge*’) with the medial *ā* being the regular continuant of the etymological **ā* (Buddruss 1960: 20, 87, 92).

Before a final nasal the change **a* > *o* takes place:

10. Balti (Skardu dialect) *dom* ‘sadness, trouble, difficulty, adversity’ (Norman 2010: 493). Cf. Shina *dāmīzhār* ‘adversity, trouble’ (Bailey 1924), Burushaski *dam(i)jar* ‘trouble, inconven-

⁸ For more details on the change **ṣṭ* > *lṭ* in Northwestern Tibetan dialects see Kogan 2019.

⁹ Probably, related to Middle High German *mūl*, German *Maul* ‘muzzle’ (Mayrhofer 1996: 369).

ience, worry’ (Lorimer 1938),¹⁰ OIA *damayati* ‘tames, subdues’, Ossetic *domyn* ‘to tame; exhaust; demand’.

11. Balti *chon* ‘vain’ (Sprigg 2002: 43), Purik *chon* (“tʰon”) ‘useless, in vain, unfounded’ (Zemp 2018: 109), Ladakhi *chon* ‘spontaneously, for no reason, gratuitously; in vain, for nothing, to no avail, useless; free, for free, at no cost’ (Norman 2010: 310). Cf. Bashkarik *chan*, Phalura *cheṇiko* (*e* < *a* in i-umlaut position), Shina *chon*, Kashmiri *chon* (< **chanu* with *o* < *a* in u-umlaut position), Gawar-Bati *chēnika* (*e* < *a* in i-umlaut position), Pashai (Kurangali dialect) *chāni* (*ā* < *a*) ‘empty, void’, Burushaski *chan* (Hunza), *čan* (Yasin) ‘*leer* (Schüssel), *unterbeschäftigt, ohne Arbeit, frei*’ (Berger 1998: 106).¹¹

See also 48.

This change does not occur before clusters and the historical vowel **i*. The etymological *a* is preserved in these positions:

12. Balti *zan-zos*, *dzan-zos* (Turtuk dialect), Purik *zan-zos* ‘wife; family’ (Norman 2010: 842; Sprigg 2002: 180).¹² Cf. Phalura *jeeni* ‘female person’ (Liljegren, Haider 2011: 76), Kashmiri *zan*, Sindhi, Lahnda, Punjabi *janī* ‘woman’, Bashkarik *jin kar-* ‘to marry’, Indus Kohistani *zhāl* ‘marriage’ (Zoller 2005: 202), OIA *jani-*, Av. *ǰaini-* ‘woman, wife’.

See also 2.

The Proto-Indo-Iranian diphthong **au* is monophthongized (**au* > *u*):

13. Ladakhi, Purik *kulik* ‘lock; key; the joint of the jaw’ (Norman 2010: 12) < PII **kaula-*, **kaulikā-* ‘curved’ (= Proto-Iranian **kaura-* > Khotanese *kūra-* ‘crooked’ (Bailey 1979: 62)). Cf. Tirahi *kóolə*, Pashai *kōlā*, Shumashti *kolāṇṭa*, Khowar *koli*, Bashkarik *kōl*, Torwali *kōl*, Phalura *kūulo*, Shina *kōlū* ‘curved, crooked’¹³.

I.2. Syllabic sonorants

The Proto-Indo-Iranian syllabic sonorant **ɾ* is vocalized in interconsonantal position:

14. Ladakhi *kit-ces* ‘to catch, seize, capture’ (Norman 2010: 82). Cf. Bashkarik *gītə* ‘took, bought’, Woṭapuri *gat* ‘took’, Gawar-Bati *gūtīm* ‘I took’ < **gṛpta-* < **gṛb-ta-* (Vedic *gṛbdha-*, Av. *gəṛapta-* ‘taken’).

15. Ladakhi *shen-ces* ‘to squash’ (Norman 2010: 986). Cf. OIA *śṛṇāti* ‘crushes, breaks’, Av. *a-sarə-ta-* ‘not discouraged (lit. ‘not broken’)', Khowar *šeník* ‘to crush’ < PIE **k’erh₂-* ‘brechen, zerbrechen (intr.)’ (LIV: 327–328).

¹⁰ The Burushaski word is likely to have been borrowed from Shina.

¹¹ In theory, there is a possibility that the Balti and Purik words have been borrowed from Shina, whereas the Ladakhi word is a loan from Kashmiri. Should this be the case, the Ladakhi example cannot illustrate the aforesaid historical phonological process, because in Kashmiri the change **a* > *o* was caused by *u*-umlaut rather than the following nasal. Semantic data, however, show that separate borrowing of the lexeme under analysis is rather unlikely. In all the three Tibetan varieties, this lexeme has undergone the same semantic development (‘empty’ > ‘(in) vain’). Technically, it cannot be ruled out that such a development took place independently in all the dialects, but it seems more probable that it occurred either in the donor language before borrowing or after borrowing but before the divergence of Ladakhi, Balti and Purik. In either case, the source language could hardly have been Kashmiri. In the latter language the development of *u*-umlaut dates back not earlier than to the 17th century (Kogan 2016), i.e. to a period when linguistic Tibetanization of not only Ladakh but also Baltistan was, no doubt, an accomplished fact. This said, direct borrowing from Kashmiri into Balti is scarcely possible, because these two languages, being geographically separated by the Shina-speaking area, are not in close contact with each other.

¹² This word is most probably a compound. Its second component may be connected with Tibetan *tshos* ‘color’ (Kogan 2019).

¹³ For the semantic change ‘crooked’ > ‘key’ cf. Greek *κλήις*, Latin *clāvis*, Old Church Slavic *kl’učь* (κλυучь) ‘key’ < PIE **klēu-* ‘Haken, krummes Holz’ (Pokorny 1959: 604–605).

The vowel *e* in the latter example may have been a free allophone of *i* in the donor language. Cf. the free variation of *i* and *ě* after *š* in Kashmiri.

I.3. Vowel sequences and sequences with intervocalic sonorants

Secondary vowel sequences which appear due to deletion of intervocalic consonants contract to monophthongs.

**a+*u* > *o*:

16. Ladakhi *po-ze* ‘ram, full-grown male sheep’ (Norman 2010: 540) < **pau-* < PII **paśu-* ‘cattle’ (> OIA *paśu-*, Av. *pasu-* ‘id.’, Pashto *psə*, Ossetic *fis* ‘sheep’).¹⁴

**a+*i* > *e*:

17. Ladakhi *tshele* ‘thatched roof, shade made of branches and leaves, thatched-roofed house or hut’ (Norman 2010: 773), Balti *tshele* ‘hut’ (Sprigg 2002: 170) < **ščaila-* < PII **ścad-ila-*. Cf. OIA *chādayati* ‘covers’, *chadis-* ‘cover, roof’, Lahnda, Punjabi *chatt*, Hindi *chat* ‘roof’, Kashmiri *chěy* ‘a kind of thatching grass’ (Grierson 1915–1932: 1066), Proto-Iranian **sād-* ‘to cover’ (> Pashto *psolal* ‘to adorn’ from **upa-* or **pati-sād-*) < PIE **sk’ed-*.

Sequences containing intervocalic sonorants, both historical and secondary, yield monophthong vowels.

**aya* > *e*:

18. Ladakhi *shen* (Leh dialect), *šen* (Shamskat dialect) ‘(wooden) floor’ (Norman 2010: 986) < PII **śrayana-* ‘leaning, foothold’. Cf. OIA *śray-*, Av. *sray-* ‘to lean’.

**āwa* > *o*:

19. Ladakhi *phok* ‘incense or burning juniper leaves or other fragrant burning materials’ (Norman 2010: 580) < PII **pāwaka-* ‘purifier’. Cf. OIA *pāvayati* ‘purifies’, Persian *pāk* ‘pure’ (< **pāwaka-*).

**āiya* > *e*:

20. Balti *be-kar* ‘court singer and dancer who improvises poems and songs’ (Sprigg 2002: 28), Ladakhi *be-da* ‘member of the caste that used to be itinerant musicians’, *be-mo* f. (Norman 2010: 640–641) < **bāiya-* < PII **wād(i)ya-* ‘music, musical instrument’, cf. OIA *vādya-* ‘musical instrument, music’, *vādyakara-* ‘musician’, *vādayati* ‘plays a musical instrument’, Punjabi *vajjā* ‘musical instrument’, Hindi-Urdu *bājā* ‘music’, Kashmiri *waz-* ‘to sound (of bell, clock etc.)’, *wāy-* ‘to play a musical instrument’, Shina (Guresi) *baž-* ‘to strike (of a gong etc.)’, Ossetic *wadynz* ‘flute, panpipe’ < **vādāničī*.

I.4. Vowel assimilation

Vowel assimilation, both progressive and regressive, was noted:

21. Balti *men-ze* ‘lump of dough’ (Sprigg 2002: 113), Ladakhi *men-ze* ‘ball of dough ready to be shaped and cooked, flattened dough ball, or shaped bread ready to be cooked’ (Norman 2010: 710). Cf. Burushaski *máano* ‘grösserer Teigklumpen’ (Berger 1998: 272), Shina *míno* ‘id.’, Brokskat *manīli*, Dameli *man* ‘bread’, Bashkarik *man*, Phalura *māṇḍili* ‘very soft bread’ (Morgenstierne 1940; 1941), Phalura *māṇḍ-* ‘to knead’, Romany *ma(n)ro* ‘bread’, Sindhi *mānī* ‘bread, loaf, food’, Punjabi *maṇḍā* ‘a thick cake’, West Pahari (Bhalesi) *mānni* ‘a large cake’, Assamese *mar-* ‘to knead (dough)’, Maithili *mār-* ‘to knead’, OIA *mṛdnāti* ‘crushes, kneads, rubs’, Av. *mōrandən* ‘(they) destroy’.

¹⁴ As can be seen, the development of this sequence differs from that of the Proto-Indo-Iranian diphthong **au*. This fact implies that the two sound changes could hardly be simultaneous. The element *-ze* in the Ladakhi word is historically a suffix found also in Classical Tibetan (Kogan 2019).

22. Balti *monḥok* ‘clod of earth’ (Sprigg 2002: 118) < **maṭ-ṭhok*? Cf. OIA *mṛttikā*- ‘earth, clay’, Hindi-Urdu, Punjabi *miṭṭī*, Nepali *māṭo*, Dameli *maṭhi* ‘id.’, Phalura *mēṭhi*, Shina *māṭi* ‘clay’; Burushaski *ṭhóko* ‘*Klumpen, Ballen (von Zucker, Salz, Butter)*’ (Berger 1998: 452), Indus Kohistani *ṭhokh* ‘a clod (earth, salt)’ (Zoller 2005: 220).

See also 7.

Consonants

Single consonants

The devoicing of the historical initial voiced consonants, characteristic of Central and Upper Ladakhi dialects including Leh, affected several Indo-Iranian loanwords:

23. Ladakhi *pat*, *paddi*, *patsi* ‘totally, completely’ (Norman 2010: 533) < **bad*. Cf. OIA *bad-dha*- ‘bound’, Gujarati *bādhū* ‘whole, entire’, West Pahari (Kotgarhi) *baddho* ‘all, entire, (pl.) all together’, Kashmiri *bod* ‘handful’.

24. Ladakhi *tul* ‘powdered dung’ (Norman 2010: 448) < **dul*. Cf. OIA *dhūli*- ‘dust, powder’, *dhūlikā*- ‘pollen, fog, mist’; Prakrit *dhūlī*, Hindi *dhūl*, Punjabi, Lahnda, Gujarati *dhūr*, Sindhi *dhūri*, Kumauni *dhuli*, Bengali *dhul*, Marathi *dhūl* ‘dust’; Nepali *dhulo* ‘dust, powder’; Tirahi “*dūda*”, Kalasha *udhrū*, Shina *ūdū*, Phalura *dūri* ‘dust’; Pashai (Wegali dialect) *duṛi* ‘dust-storm’; Torwali *dur* ‘mist’.

See also 14.

The above examples give reason to believe that the process of borrowing may have begun before and finished after the devoicing of the initial mediae (Kogan 2019).

When not devocalized, the etymological initial voiced aspirated obstruents lose aspiration and merge with their plain voiced counterparts. Thus, **bh* > *b*, **dh* > *d*. See 3, 5.

The process of deaspiration, followed by dentalization and assibilation, seems to have affected the Proto-Indo-Iranian affricate **jh* (**jh* > **j* > (*d*)*z*): *gzar* ‘to flow’ < PII **gjarh*- (see 6).

25. Another probable instance of this sound change is Ladakhi *zanggi* ‘a species of tiny flying biting insect’ (Norman 2010: 825). This word is likely to reflect some derivative of PII **jhan*- ‘to strike, kill, injure’ (> OIA *han*-, Av. *jan*- ‘id.’). Cf. the semantic development of this root in some Iranian languages, e.g. Talysh *žan*- ‘to bite, sting (of insects and snakes)’ (Rastorguyeva, Edelman 2007: 136). If we assume the same development for the Indo-Iranian lect from which the Ladakhi noun has been borrowed, we may consider the word in question to be the reflex of PII **jhan-aka*- with the meaning ‘sting, stinger’ to which the adjectivizing suffix *-ī* (< PII **-in/-ī*) has been added. This suffix was widely used, e.g. in Old Indo-Aryan, to form adjectives with a possessive sense, often prone to nominalization: *keśin*-, Nom Sg *keśī* ‘long-haired’ (cf. *keśa*- ‘hair’), *pakṣin*- Nom Sg *pakṣī* ‘winged; bird’ (cf. *pakṣa*- ‘wing’), *hastin*- Nom Sg *hastī* ‘possessing a hand; elephant’ (cf. *hasta*- ‘hand’). The original meaning of the source form of the Ladakhi word should thus have been “possessing a sting”. The elision of the interconsonantal short *a* (**jhanaka*- > **zang*-) is not uncommon in some Dardic languages (cf., e.g. Pashai *šūṅg* ‘dog’ < **śunaka*-, *sāṅg* ‘earth’ < **samaka*-). This sound change must have resulted in the formation of the secondary cluster *ng*, which in turn must have precluded the change *a* > *o* in the initial syllable.

Other etymological palatal affricates are also dentalized word-initially (**č* > *c* (=ts),¹⁵ **j* > (*d*)*z*):

¹⁵ Balti, Purik and Ladakhi words are transcribed using the standard Tibetological transcription system, where *c* and *ch* stand for the voiceless palatal and voiceless palatal aspirated affricate respectively, while their den-

26. Ladakhi *tsapik* ‘a little, a bit, a little while’ (Norman 2010: 731) < PII *čap- ‘to catch, snatch, pick, pinch’, cf. Shina čap-, Indus Kohistani cap-, Gawar-Bati cep- ‘to bite’, Pashai čip- ‘to bite off’, Kashmiri cop ‘a bite’, Balochi čāmp- ‘to snatch’, Ossetic cæfsyn ‘to stick, glue’, Munji cəb- ‘to pluck’.

See also 12, 17.

The initial palatal affricates in turn usually correspond to retroflex affricates or their regular reflexes in the Dardic languages (see 4, 11).

The two above types of correspondences may have resulted from a kind of affricate shift involving the dentalization of the historical palatal affricates and the concomitant palatalization of the retroflex ones. A similar historical-phonological process has affected the affricate system of Kashmiri (Kogan 2016).

It is worth noting that the initial palatal affricates in Burushaski loanwords are sometimes preserved as such:

27. Ladakhi *cancel*, *chancil* ‘the green outer shell or fruit of walnut’ (Norman 2010: 267). Cf. Burushaski *čanjil* (Nagir, Hunza), *čanjil* (Yasin)¹⁶ ‘die grüne äussere Schale der Walnuss, Häutchen zwischen den Teilen des Walnusskerns’, Shina *cháćiil* ‘id.’ (Berger 1998: 96).

28. Balti *cangti* ‘drop’, Purik, Ladakhi (Shamskat and Leh dialects) *cangti* ‘leak in a roof’ (Norman 2010: 267). Cf. Burushaski *čan man-*, Shina *cháčan-* ‘(Wasser) sickern, herabfliessen’ (Berger 1998: 97).

29. Balti *cha* ‘millet’ (Sprigg 2002: 41), Ladakhi (Shamskat dialect) *cha* ‘a variety of millet, a cereal grain which was grown in Ladakh in the past, but very little now’ (Norman 2010: 290). Cf. Burushaski *ča* (Nagir, Hunza), *ča* (Yasin) ‘Hirse, Kolbenhirse, *Setaria italica*’ (Berger 1998: 95).

Etymological intervocalic stops, both voiceless and voiced, are dropped in a number of cases:

30. Balti (Skardu dialect) *spa* ‘taste, tasty (neutral), (sexual) enjoyment’, Purik, Ladakhi (Nubra dialect) *spa* ‘taste, flavour’ (Norman 2010: 542; Sprigg 2002: 159) < PII *swāda-. Cf. OIA *svāda-* ‘taste’, Iranian *χwāda- (> Persian *χvā* ‘good taste’, Balochi *wād* ‘salt’), Shina *ispāvu* ‘tasty’.

31. Ladakhi *perak* ‘Ladakhi women’s head-dress, covered with turquoise and coral’ (Norman 2010: 539) < PII *paridhāka-. Cf. OIA *paridhā-* ‘to put on (clothes)’, Sindhi *paharaṇu*, Nepali *pairanu*, Hindi *pahirnā* ‘to put on, wear’, Kashmiri *pārun* ‘to put on; to adorn, ornament’, Khovar *purduik* ‘to cover oneself, put on a cloak’.

32. Ladakhi *shanṭi* ‘a leafy vegetable’ (Norman 2010: 975) < PII *śāka- ‘green vegetable’. Cf. OIA *śāka-*, Shina, Indus Kohistani *šā*, Phalura *šō*, Bashkarik *ša*, Kashmiri *hākh*, Kalasha *šak*, Khovar *šax*, Waigali, Dameli *cā* ‘id.’, Indus Kohistani *šṛī* ‘a green vegetable with round leaves’ (Zoller 2005: 378) < PIE *k’ēko- (Mayrhofer 1996: 628).

See also 17, 20, 48.

Intervocalic voiceless stops, however, are sometimes preserved or sonorized. Such instances require a separate analysis. The retention of intervocalic *k* is observed exclusively in probable reflexes of the reconstructed formations with the suffix *-k-, i.e. masculine nouns in *-aka- or feminine nouns in *-ikā-:

33. Ladakhi *kuruk* ‘donkey foal; foal’ (Norman 2010: 12), Balti *bong-kúru* ‘donkey colt’ (Sprigg 2002: 30) < *kurtaka-. Cf. Kalasha *kúak* (< *kurak) ‘child (male or female); the offspring of a human or animal’ (Trail 1999), Shumashti *kuṛ*, Dameli *kuṛa* ‘child’, Ashkun *kūṛə* ‘child, foetus’, Kati *kru*, *kuṛuk* ‘young of animals’, Prasun *kyūru* ‘young of animals, child’ < *kurtaka-, Kurdish *kurr* ‘son’, Middle Persian *kurra*, Persian *kurra* ‘foal’ < *kurna(ka)-.

tal counterparts are conveyed by *ts* and *tsh*. In Dardic and Iranian examples, as well as in Proto-Indo-Iranian reconstructions, *č*, *ch* mark the palatal affricates, while *c*, *ch* are used for the dental ones.

¹⁶ In Berger’s notation *ć* stands for the voiceless palatal affricate.

See also 5, 7, 8, 13, 19, 31, 48.

All such cases can be easily explained, if we assume that the suffix **-k-* retained its productivity in the donor language for a long time and was perceived as a separate morpheme when intervocalic consonants were dropped. In this connection, an interesting fact is that this formant has probably been suffixed to some roots of Burushaski origin:

34. Balti *kulak* ‘meal (quickly made mixture of buttermilk and flour)’ (Sprigg 2002: 94), Purik *kholak* ‘a certain dish’, Ladakhi (Leh, Shamskat and Nubra dialects) *kholak*, *qholak* ‘ready-to-eat dough of roasted flour’ (Norman 2010: 117). Cf. Burushaski (*d*)-*q(h)ul-an-* (Hunza, Nagir), *d-χul-an-* (Yasin) ‘(*Teig*) *kneten*’ (Berger 1998: 357).

The sonorization of the historical intervocalic *t* can be observed after a non-etymological nasal, or perhaps nasalization:

35. Balti *rindi* ‘lead, bullet’ (Sprigg 2002: 139), Ladakhi *rindi* ‘bullet; lead (metal)’ (Norman 2010: 911) < PII **rīti-* ‘flowing, melting, (metal) casting’ Cf. Bashkarik *rīd*, Torwali *žit* (*ž* < *r*) ‘brass’, Shina *rīl* ‘brass, bronze, copper’, Gawar-Bati *rīt* ‘copper’, OIA *rīti-* ‘stream; yellow brass, bell-metal’.

The development of secondary nasals and nasalization is a well-attested phenomenon in a number of Dardic, Indo-Aryan and East Iranian languages.

In the language under study rhotacism, i.e. the change of PIE **l* to *r*, seems not to have been a regular process in intervocalic position. Like in Nuristani, Dardic and most Indo-Aryan dialects, the distribution of reflexes of the two sonorants is not always clear. Possible, albeit not always provable cases of retention of PIE **l* as well as lambdacism (the change of PIE **r* to *l*) have been attested:

36. Ladakhi *puli*, *polo* ‘Ladakhi biscuits of a particular type’ (Norman 2010: 541). Cf. OIA *pūra-* ‘cake’, *pauli-* ‘a cake of scorched grain and ghee’, Sindhi, Punjabi, Hindi, Kumauni *pūrī*, Gujarati, Marathi *purī* ‘fried cake’, Kashmiri *pūr* ‘a kind of cake fried in ghee’ < PIE **pūr-* ‘corn, wheat’ (Gamkrelidze, Ivanov 1984: 657).

See also 4,¹⁷ 8.

PII **s*, both prevocalic (word-initial) and intervocalic, is preserved:

37. Ladakhi *sale* ‘knitting needle’ (Norman 2010: 1005). Cf. Phalura *silēni* ‘needle’, OIA *sīvyati* ‘sews’, Khotanese *hīya* ‘sewn stuffs’, Ossetic *x^wiyjīn* ‘to sew’.¹⁸

See also 2.

PII initial **ś* (< Proto-Aryan¹⁹ **ć* < PIE **k*) is retained:

38. Balti *shang* ‘wisdom, sense’ (Sprigg 2002: 151)²⁰, Purik *šanj* ‘consciousness’ (Zemp 2018: 931), Ladakhi *shang* ‘alertness, awareness, caution, prudence’ (Norman 2010: 553). Cf. Burushaski (loanword) *šanj* ‘awake, aware; care, heed, attention’ (Lorimer 1938: 322), Shina *šori* ‘care, anxiety; awake, alert’ (Bailey 1924), Khowar *šanig* ‘fear, suspicion’ (Morgenstierne 1973), OIA *śāṅkā-* ‘fear, distrust’, *śāṅkate* ‘is afraid, distrusts’ < PIE **k’enk-* ‘in der Schwebe sein, hängen (intr.)’ (LIV: 325).

See also 15, 32.

¹⁷ Possibly, PII **kšar-* (>Balti, Purik, Ladakhi *chal* ‘overflow, spill over’, see 4) is related to PII **gīhar-*, reflecting another variant of the same Proto-Indo-European root (Rastorgueva, Edelman 2007).

¹⁸ Burushaski *sel* ‘Nadel, Stecknadel’ (Berger 1998b: 377) may have been borrowed from the same Indo-Iranian source as the Ladakhi word.

¹⁹ The terms “Aryan” and “Indo-Iranian” are not used as synonymous in the present work. Following D.I. Edelman (1992), we believe that the split of the Aryan unity began with the separation of the Nuristani branch. The remaining three branches, i.e. Indo-Aryan, Iranian and Dardic, form the Indo-Iranian unity, whose disintegration took place at a later date. Thus, Indo-Iranian is considered a lower-level taxon within the Aryan subfamily.

²⁰ In traditional Tibetological transcription *sh* stands for voiceless palatal sibilant.

As for intervocalic PII *ś, there are possible cases of its retention as well as elision:

39. Balti *lashi* ‘a resinous wood used as a candle or torch because it burns slowly’ (Sprigg 2002: 98). Cf. Kashmiri *lašhi* ‘a torch’ (Grierson 1915–1932: 533), Shina (Gilgit dialect) *lāi* ‘torch (unlit)’ (*i* regularly < *ś in intervocalic position).

See, by contrast 16.

Since the Indo-Iranian etymology of the above word for ‘torch’ is unclear, and no apparent cognates outside Dardic have been found for it so far, this example should be considered doubtful.

PII intervocalic *ś²¹ (=OIA ś) > *y*:

40. Balti *múyu*, *myo* ‘mouse’ (Sprigg 2002: 119). Cf. OIA *mūṣ-*, *mūṣā-*, *mūṣika-* ‘mouse, rat’, Persian *mūš*, Pashto *māžak*, Ossetic *myst*, Pashai *mūč*, Shumashti *múšo*, Gawar-Bati *muša*, Torwali *mūš*, Phalura *mūšo* ‘mouse’, Bashkarik *mūš* ‘mouse, rat’, Shina (Guresi dialect) *mūzu* ‘rat’, *mūžai* ‘mouse’, West Pahari (Jaunsari) *mūśā*, Romany *mušó* ‘mouse’, Kumauni, Nepali *muso* ‘mouse, rat’.

After a non-etymological nasal (or secondary nasalization) this sibilant is sonorized (*ś > ž):²²

41. Balti *munzhur* ‘small mole-like mouse’ (Sprigg 2002: 119).

PII initial *w > *b*. See 2, 9, 20.

One probable instance of consonant dissimilation across a morphemic boundary has been attested. See 22.

Consonant clusters

PII initial *dw > *b*:

42. Balti *bar-ban* ‘window (in a wall), glass-pane window’ (Sprigg 2002: 26), Purik *barban* ‘window’ (Zemp 2018: 945). Cf. Brokskat *barban* ‘id.’, Pashai *darī*, Gawar-Bati *derī*, Kalasha *durík*, Phalura *darúri*, Indus Kohistani *darī*, Shina *darií*, Kashmiri *dār* ‘window’, OIA *dvār-* ‘door, gate’, Av. *duuar-* ‘gate’.

PII *kš (> OIA *kṣ*, Proto-Iranian *xš, Dardic *čh) > *čh*. See 4.

This process may have gone through the intermediate stage of *čh*, the latter having changed to palatal *čh* in the wake of affricate shift.

Like in most Dardic and New Indo-Aryan languages, in etymological groups of two voiceless stops the first component is lost (*kt, *pt > t).²³ See 3, 14.

The etymological voiceless stops are sonorized after nasals (*nk > ng > ŋ, *nt > nd).²⁴ See 2, 38.

Proto-Indo-Iranian clusters of the type “nasal + voiced stop” develop in different ways. Word-finally PII *ndh > *n*:

43. Balti *ban* ‘fence’ (Sprigg 2002: 25) < PII *bandha-, cf. OIA *bandha-* ‘border, framework, damming’.

In the same position PII *ngh has apparently lost its nasal component and was simplified to *g* with subsequent devoicing (*ngh > g > k, q):

44. Ladakhi *tak-tak* (Shamskat dialect), *ṭak-ṭak*, (Leh dialect) ‘taut, stretched tight, tight (e.g. of curtain, clothing, greenhouse plastic)’, Purik *taqtaq* ‘tight’ (Norman 2010: 370), Balti

²¹ Developed from PIE *s by the RUKI-rule.

²² In Tibetological transcription the voiced palatal sibilant is conveyed by *zh*.

²³ It is, however, unclear, whether or not the development of these clusters involved regressive assimilation and the formation of geminates, as was the case, e.g. in the history of Indo-Aryan.

²⁴ Cf. the aforementioned sonorization of voiceless consonants after a non-etymological nasal or secondary nasalization.

ʔak-ʔak ‘hard’ (Sprigg 2002: 184) < *tag < PII *tangh- (> Proto-Iranian *ʔanj- ‘to pull, draw’ (> Av. ʔanj- ‘id.’, Persian *sanjidan* ‘to measure; reflect; compare, put in balance’, Ossetic *tinʒin* ‘to spread, stretch out; to crucify’), Kashmiri *ʔanz* ‘extreme and urgent desire’, Khowar *tonjeik* ‘to destroy, pull down’).

The Proto-Indo-Iranian cluster *rt seems to have undergone two different developments. In one Balti example the first component of this cluster moved from word-medial to word-initial position:

45. Balti *rkat* ‘to cut down with a sword’ (Sprigg 2002: 140). Cf. PII *kart- ‘to cut’ > OIA *kr̥ntati*, *kartati* ‘cuts’, Av. *kərəntaiti* ‘cuts’, *karəta-* ‘knife’, Hindi-Urdu *kattā* ‘curved knife’, *kattī* ‘sword, knife, dagger’, Tirahi *katāri* ‘knife’, Pashai *kāṭare* ‘spear’, Gawar-Bati *kaṭāro* ‘large knife’, Kalasha *katār*, Khowar *kuter* ‘knife, dagger’, Bashkarik *kāṭer* ‘knife’, Savi *kaṭārēi*, Phalura *kaṭōro* ‘dagger’, Shina *khāṭaru* ‘knife’.

On the other hand, there is an apparent instance of the development *rt > r (perhaps, through the retroflex ʀ as an intermediate stage): Ladakhi *kuruk* ‘donkey foal; foal’ (Norman 2010: 12), Balti *bong-kūru* ‘donkey colt’ (Sprigg 2002: 30) < PII *kurtaka- (see 33).

It should, however, be noted that, technically, PII *kurtaka- cannot so far be proven to be the only possible prototype for the immediate Indo-Iranian source of the Ladakhi and Balti words. Since the reflex of PII *rn in the language under study is not clear, the prototype *kurnaka-, similar to that which is reconstructed for the Iranian forms, cannot be excluded.

The PII cluster *rth seems to be reflected as the retroflex or dental unaspirated stop (*rth > Balti, Purik ʈ, Ladakhi ʈ):

46. Balti, Purik *gaṭ* ‘knot, joint of body’ (Sprigg 2002: 58; Zemp 2018: 64), Ladakhi (Leh dialect) *changgat* ‘knee-joint’ (Norman 2010: 313), *gat* ‘obstacle’ (Norman 2010: 241) < *garthi-, cf. Indus Kohistani *gār*, Brokskat *gaṭhi* ‘knot’, Pashai *gaṭanā* ‘joint’, *gaṭh-* ‘to tie’, Hindi, Bengali, Marathi, Gujarati *gāṭh*, Punjabi, Lahnda *gaṇḍh*, Sindhi *g’andhi* ‘knot’, Pashto *yaṛəl* ‘to twist, spin, plait’ (< *garṭh-), Burushaski (loanword) *gaṭ* ‘Knoten (auch im Stengel von Pflanzen), Knöchel (des Fingers)’ (Berger 1998: 150). The reconstructed prototype *garthi- is, no doubt, a derivative of PII *grath-/granth- ‘to tie, bind’ > OIA *granthi-* ‘knot’, *granthayati* ‘ties’, Middle Persian *grih*, Persian *giriḥ*, Khotanese *grratha*, Sogdian *yr’nš*, Ishkashimi *yureḥ*, Ossetic *ælxync* ‘knot’.

PII *śc (< PIE *sk’) > ch. See 17.

This change has probably passed through the intermediate stage of *čh.

PII *sw > sp (see 30).

This phonetic change is regular in a number of Dardic languages (e.g., Tirahi, Kalasha, Khowar). In Shina, however, it coexists with the more widespread change *sw > s: *sa* ‘sister’ < *swasar-, *sācu* ‘dream’ < *swāpra- (Turner 1966: 805), *so-* ‘to sleep’ < *swapa-. It seems likely that words with *sp* < *sw were borrowed from some other, perhaps now extinct, Dardic lect.

PII initial *śr > ʂ, ś. See 18.

The cerebral and palatal reflexes are found in the Shamskat and Leh dialects of Ladakhi respectively. Since no examples of secondary cerebralization of the palatal sibilant were noted in Shamskat, it is reasonable to assume that the retroflex ʂ (< PII *śr) of the Indo-Iranian donor language was preserved intact in this dialect and palatalized to ś (*sh*) in the Leh variety due to the influence of the following front vowel.

PII *št > ʂṭ > lṭ, lt:

Balti, Purik, Ladakhi *mulṭuk*, *multuk* ‘fist’ (Norman 2010: 705; Sprigg 2002: 118) < *mulṭak < *muṣṭaka-. See 7.

The change of a sibilant to *l* before a voiceless retroflex stop is apparently recent. As was shown in Kogan 2019, this process has even affected English loanwords. The presence of the retroflex ʈ in the above Tibetan words clearly indicates that the Proto-Indo-Iranian cluster *št

has been cerebralized in the donor language, i.e. the development $\acute{s}t > \acute{s}\acute{t}$ has taken place. The existence of the variant with dental (*multuk*) may be attributed to the fact that the sequence $l\acute{t}$ is rather rare in Northwestern Tibetan dialects.

The OIA cluster $\acute{s}th$, unlike its counterpart with unaspirated stop ($\acute{s}t$), seems to correspond to a single retroflex consonant in the language under study. This consonant is reflected as \acute{d} in Balti and \acute{t} in Ladakhi:

47. Balti *kaḍik* ‘small branches’ (Sprigg 2002: 82), Ladakhi *kaṭik* ‘branches and leaves of trees as fodder for animals in the spring’ (Norman 2010: 4). Cf. OIA *kāṣṭha-* ‘piece of wood’, *kāṣṭhikā-* ‘small piece of wood’, Lahnda, Punjabi, Hindi-Urdu, Kumauni, Nepali, Gujarati *kāṭh*, Sindhi *kāṭhī* ‘wood’, Kati *kāṭ* ‘branch’, Kalasha *kaṭ* ‘board’ (Trail 1999), Shina *kāṭ* ‘wood’, Kashmiri *kāṭh* ‘wood’, *kāṭh* ‘small stick’.

The Ladakhi form looks more archaic than the Balti one. In the latter the intervocalic sonorization of an earlier voiceless retroflex stop must have occurred. The origin of the lexeme under discussion is still somewhat unclear. The most plausible and reliable etymology seems to be the one advanced by Otakar Klima and supported by Thomas Burrow (Klima 1970; Burrow 1975). According to these scholars, the Indo-European protoform of OIA *kāṣṭha-* should be reconstructed as **kolstho-* and construed as a derivative of PIE **kelā-* ‘to hew’ with an *s*-extension.²⁵ The element **-tho-* in this protoform is most probably etymologically identical to OIA *-thā-*, Av. *-θā-* < PII **-thā-* (cf. OIA *artha-* ‘object, aim’, Av. *arəθa-* ‘object, matter’ < *ar-* ‘to move, reach’; OIA *gāthā-*, Av. *gāθā-* ‘singing, chant’ < *gā-* ‘to sing’). For more details on this suffix see Wackernagel, Debrunner 1954: 717–722.

Based on the last two groups of examples, the following development of cluster-initial sibilants may be hypothesized: a sibilant disappears before a historical (Indo-Iranian) voiceless aspirate, and is retained, sometimes with subsequent changes, before an unaspirated voiceless stop. The same phonological development is characteristic of many Dardic languages and may be tentatively reconstructed for Proto-Dardic.

PIE **tkʷ* (> OIA *kṣ*, Proto-Iranian **š*) > *ch*:²⁶

48. Balti *tshon* ‘injury’ (Sprigg 2002: 171), *tshak* ‘rheumatism’ (Sprigg 2002: 169), Purik *tshaq* ‘pain with difficulty of breathing’, *tshak yong* ‘to ache’, Ladakhi *tshak* ‘sprain, pulled muscle, sudden cramp, sudden sharp pain’ (Norman 2010: 759–760). Cf. OIA *kṣaṇoti* ‘injures, hurts’, *kṣata-* ‘wounded’, *kṣataka-* ‘wound’, *kṣaṇana-*, *kṣati-* ‘injury, damage’, Pali *khaṇati* ‘destroys’, Khowar *čay* ‘illness’ (= OIA *kṣati-*?), Khotanese *vaṣanaurau* ‘destructive’ (Bailey 1979: 379) < **vi-šana-bara-*, Manichaean Sogdian *pʷšyyn* 3 Sg ‘to trap’ < **apa-* or **upa-šan-*, Christian Sogdian *ptšng*, Buddhist Sogdian *ptšʹnkh* ‘cross, torture-instrument’ < **pati-šana-ka-* (Gershevitch 1954: 25, 96), Greek *κτείνω* ‘I kill’ < PIE **tkʷen-* (LIV: 645).

The dental affricate *ch* must have regularly evolved from the earlier palatal *čh*. The latter phoneme is the likely Proto-Dardic reflex of PIE **tkʷ*. As was shown in Kogan 2019, this historical-phonological feature of the aforesaid Indo-Iranian loanwords is a strong argument in favor of their Dardic origin.

In addition to the historical-phonological innovations discussed above, one interesting archaism seems to be noteworthy. An Indo-Iranian loanword in Balti shows the retention of the initial stop in the Proto-Indo-Iranian cluster **gjh* (> Indo-Aryan *jh*, Proto-Iranian **gž*): *gzar* ‘to flow’ < PII **gjhar-* (see 6).²⁷

²⁵ In *Lexikon der indogermanischen Verben* this root is reconstructed as **kelh₂-* ‘*schlagen*’ (LIV: 350). For PIE **kolstho-* Klima adduces probable reflexes in Slavic.

²⁶ In Indo-Iranian languages this Proto-Indo-European cluster has reflected in the same way as PIE **kʷ*’s. This suggests that the two groups may have merged in Proto-Indo-Iranian.

²⁷ On the Iranian reflexes of PII **gjhar-* see also (Cheung 2007; Rastorgueva, Edelman 2007).

Preliminary conclusions

The above analysis has revealed a number of facts that are directly relevant to the questions formulated at the beginning of the present work. First of all, four apparent cases of non-uniform phonological development in the same position have been detected. These are:

- 1) PII medial **a* or **ā* > *aa* in Balti *baan* ‘man or men who sing religious songs and foretell the future’ but > *a* elsewhere;
- 2) PII final **a* preserved in Balti *basanda* ‘dandelion’ but dropped elsewhere;
- 3) the initial *r* of PII **-rt-* has moved to the word-initial position in Balti *rkat* ‘to cut down with a sword’, whereas in Ladakhi *kuruk* ‘donkey foal; foal’, Balti *bong-kúru* ‘donkey colt’ the same intervocalic cluster evolves into *r* (perhaps, through retroflex *r*);
- 4) PII intervocalic **ś* is probably preserved in Balti *lashi* ‘a resinous wood used as a candle or torch because it burns slowly’ but dropped in Ladakhi *po-ze* ‘ram, full-grown male sheep’.

It is easy to see that all the above cases involve words peculiar to Balti. As shown in Kogan 2019, the majority of Indo-Iranian and Burushaski loanwords in Northwestern Tibetan fall into two groups: those characteristic of all the dialects, and those found only in Muslim varieties, i.e. in Balti and/or Purik. All the attested instances of seemingly unmotivated irregularities in historical phonology belong to the latter category. This fact seems to be explicable, if we bear in mind that the route of Tibetan migration to the present-day Muslim regions must have inevitably passed through Ladakh. The Indo-Iranian speakers of this or adjoining areas must have been the first people whom the Tibetans had encountered in the course of their north-westward movement. It was the contact with these people that had most probably resulted in the acquisition of loan vocabulary, now shared by Ladakhi, Purik and Balti. Later on, a certain part of Tibetan speakers migrated from Ladakh further northwest to Baltistan, where a number of lexical items from local or neighboring Indo-Iranian dialects could have been borrowed. In other words, Muslim dialects seem to possess at least two chronologically different Indo-Iranian lexical strata. A remarkable fact is that the earliest stratum, common with Ladakhi, does not show any historical-phonological irregularities. It means that there is so far no indication that loanwords belonging to this stratum have been borrowed from more than one source. This is why we consider it reasonable to accept the single-source hypothesis as a working one.

Although the lexical data analyzed above are rather scanty, they nevertheless do give us some hints as to the genetic position of the donor language. In my previous paper it was noted that no unquestionable examples of typical Indo-Aryan phonological developments had been so far attested in the material (Kogan 2019). Now it can be added that certain historical-phonological features, discovered during our study, make it virtually impossible to classify the source-language of at least some of the Indo-Iranian loans as Indo-Aryan. Here undoubtedly belong the deaspiration of PII **jh* (< PIE **gh* and **g^wh* in the palatalizing position) with the subsequent change *j* > *z*,²⁸ and the distinction, at least in word-initial position, between the reflexes of PII **kš* and PIE **tk*.²⁹ On the other hand, the retention of the prevocalic (word-initial) and intervocalic PII **s* suggests that the language in question was not Iranian, since, as is widely known, in Iranian the lenition **s* > *h* occurred in both positions. The Nuristani origin of the above loanwords is hardly probable either, since no instance of a regular Nuristani development of the PIE palatal **k* into the dental affricate *c* has been noted. Instead, as we have shown, the change **k* > *ś* has taken place. Apart from this, in Nuristani etymological voiceless

²⁸ See the note on Ladakhi *zanggi* ‘a species of tiny flying biting insect’.

²⁹ As noted above, in Indo-Iranian this cluster seems to have merged with PIE **k*’s. Its Proto-Indo-Iranian reflex may perhaps be reconstructed as **śs*.

aspirates have been deaspirated with the result that the opposition in aspiration ceased to exist for consonants, whereas our data contain quite a number of words with voiceless aspirated stops and affricates.

Unlike Indo-Aryan, Iranian or Nuristani, certain apparently Dardic historical-phonological isoglosses have been detected in our material. They are as follows:

- 1) the deaspiration of Proto-Indo-Iranian voiced aspirates, including the aspirated affricate **jh*;
- 2) loss of historical sibilants in Indo-Iranian clusters of the type “sibilant + voiceless aspirate”;³⁰
- 3) PIE **tkʷ* > **čh* with subsequent dentalization to *ch*.

Two of the three above isoglosses are found outside Dardic (the isogloss (1) in Iranian and Nuristani, and the isogloss (2) in Middle and New Indo-Aryan) but the bundle as a whole seems to be unique to the Dardic group. The isogloss (3) indicates that the donor language, being Dardic, could not, however, have belonged to the East Dardic subgroup, where word-initially the Proto-Dardic **čh* was cerebralized into **çh* and thus merged with the reflex of PII **kš* (Kogan 2016; 2019).

Importantly, none of the three above-mentioned sound changes is found in lexical items peculiar to Muslim dialects alone. It means that these historical-phonological features most probably characterize the earliest stratum of Indo-Iranian loanwords, common to all the Northwestern Tibetan varieties. This fact substantially contributes to clarifying the nature of the Tibetan-Indo-Iranian language contact. As already noted, there are reasons to believe that the earliest (“common Northwest Tibetan”) loanwords have been borrowed from a single source. Since this source-language could have been neither East Dardic nor Indo-Aryan or Iranian, and it is to these groups that all the known Indo-Iranian neighbors of Ladakhi, Balti and Purik belong or belonged, the substratum influence in our case seems to be much more likely than the adstratum one.

The tentative picture of linguistic and ethnic contact obtained from the present research may thus be represented in the following way. As a result of the northwestward expansion of the Tibetans in the early Middle Ages, a number of Tibetan dialects fell under the influence of a certain Indo-Iranian substratum lect. Geographically, the zone of this substrate contact was most probably located in present day Ladakh. The substrate language must have belonged to the Dardic group but, no doubt, was not an early form of Shina, Brokskat or Kashmiri, nowadays spoken in zones adjacent to Ladakhi and Balti. After a certain time period, during which considerable loan vocabulary was acquired, a new migration to the northwest took place. Its consequence was the spread of the Tibetan language in what is now Baltistan and Kargil. The majority of the local pre-Tibetan population seems to have spoken a certain variety of Burushaski (Kogan 2019)³¹ but the presence of some Indo-Iranian forms of speech is also probable. The latter lects must have influenced the early form of Balti, as is clear from the existence of several Indo-Iranian loanwords specific to the Balti dialects. The exact source of such loanwords, as well as the type of contact that has resulted in their adoption, remain unknown because of the extreme scarcity of the material. There is, however, a possibility that future field linguistic research will yield sufficient new lexical data, instrumental in the clarification of these both issues.

³⁰ See the note on the intervocalic retroflex in Balti *kaḍik* ‘small branches’ and Ladakhi *kaṭik* ‘branches and leaves of trees as fodder for animals in the spring’.

³¹ Burushaski seems to have exerted some influence on the Indo-Iranian language of pre-Tibetan Ladakh. This is indicated by the existence of Burushaski loanwords common for Ladakhi, Balti and Purik. It is remarkable that some of these loanwords are attested with the Indo-Iranian formant *-k* (see 34).

Abbreviations for language names

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

References

- Bailey, T. Grahame. 1924. *Grammar of the Shina Language* ("Royal Asiatic Society, Prize Publication Fund", vol. VIII). London: Royal Asiatic Society.
- Bailey, Harold W. 1979. *Dictionary of Khotan Saka*. Cambridge University Press.
- Berger, Hermann. 1998. *Die Burushaski-Sprache von Hunza und Nager. Teil III. Wörterbuch Burushaski-Deutsch, Deutsch-Burushaski. Unter Mitarbeit von Nasiruddin Hunzai*. Wiesbaden: Harrassowitz Verlag.
- Buddruss, Georg. 1960. *Die Sprache von Wotapur und Katarqala*. Bonn: Selbstverlag des orientalischen Seminars der Universität Bonn.
- Burrow, Thomas. 1975. A new look at Brugmann's law. *Bulletin of the School of Oriental and African Studies*, Vol. 38, No. 1: 55–80.
- Cheung, Johnny. 2007. *Etymological Dictionary of the Iranian Verb. Leiden Indo-European etymological dictionary series*. Leiden / Boston: Brill.
- Francke, August H. 1907. *A history of Western Tibet: one of the unknown empires*. London: S. W. Partridge & Co.
- Gamkrelidze, Tamaz V., Vyacheslav V. Ivanov. 1984. *Indoevropskij jazyk i indoevropejcy: Rekonstrukcija i istoriko-tipologičeskij analiz prajazyka i protokultury*. Tbilisi: Izdatel'stvo tbilisskogo universiteta.
- Gershevitch, Ilya. 1954. *A Grammar of Manichean Sogdian. (Publications of the philological society)*. Oxford: Basil Blackwell.
- Grierson, George A. 1906. *The Piśāca languages of North-Western India*. London: Royal Asiatic Society.
- Grierson, George A. 1915–32. *A Dictionary of the Kashmiri Language. Vol. I–IV* ("Bibliotheca Indica". Work No. 229). Calcutta: Royal Asiatic Society of Bengal.
- Grierson, George A. 1919. *Linguistic Survey of India. Vol. VIII, pt. 2. Specimens of the Dardic or Piśāca Languages (including Kāshmirī)*. Calcutta: Superintendent Government Printing, India.
- Jettmar, Karl. 1975. *Die Religionen des Hindukusch (= Die Religionen der Menschheit. Bd. 4, 1). Mit Beiträgen von Schuyler Jones und Max Klimburg*. Stuttgart: Kohlhammer Verlag.
- Klima, Otakar. 1970. Das altind. kāṣṭhām. *Archiv orientální* 38: 26–27.
- Kogan, Anton I. 2015. Nekotojje voprosy geneticheskij klassifikatsii dardskix jazykov po dannym istoričeskij fonetiki. *Journal of Language Relationship* 13/1: 1–21.
- Kogan, Anton I. 2016. *Problemy sravnitel'no-istoričeskogo izučeniija jazyka kashmiri*. Moskva: "Fond razvitija fundamental'nyx lingvističeskix issledovanij".
- Kogan, Anton I. 2019. On possible Dardic and Burushaski influence on some Northwestern Tibetan dialects. *Journal of Language Relationship* 17/4: 263–284.
- LIV = Kümmel, Martin and Helmut Rix, eds. 2001. *Lexikon der indogermanischen Verben: Die Wurzeln und ihre Primärstammbildungen. 2nd ed.* Wiesbaden: Reichert.
- Liljegren, Henrik, Naseem Haider. 2011. *Palula vocabulary*. Islamabad: Forum for Language Initiatives.
- Livshits, Vladimir A. 2003. Drevneje nazvanije Syrdar'ji. *Vestnik drevnej istorii* 1(244): 3–10.
- Lorimer, David L. R. 1939. *The Burushaski language. Vol. III. Vocabularies and index*. Oslo: H. Aschehoug & Co. (W. Nygaard).
- Mayrhofer, Manfred. 1996. *Etymologisches Wörterbuch des Altindoarischen. II Band*. Heidelberg: Universitätsverlag C. Winter.
- Morgenstierne, Georg. 1940. Notes on Bashkarik. *Acta Orientalia* XVIII(3/4): 206–257.
- Morgenstierne, Georg. 1941. *Notes on Phalura, an Unknown Dardic Language of Chitral*. Oslo: J. Dybwad.
- Morgenstierne, Georg. 1973. Sanskritic Words in Khowar. In: Georg Morgenstierne. *Irano-Dardica*: 256–272. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Norman, Rebecca. 2010. *A Dictionary of the Language Spoken by Ladakhis*. Ms. (unpublished draft).
- Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch*. Bern / München: Francke Verlag.
- Rastorgueva, Vera S., Dzhoy I. Edelman. 2007. *Etimologičeskij slovar' iranskix jazykov. T. 3 (f–h)*. Moskva: «Vostočnaja literatura» RAN.

Abbreviations for language names

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

References

- Bailey, T. Grahame. 1924. *Grammar of the Shina Language* ("Royal Asiatic Society, Prize Publication Fund", vol. VIII). London: Royal Asiatic Society.
- Bailey, Harold W. 1979. *Dictionary of Khotan Saka*. Cambridge University Press.
- Berger, Hermann. 1998. *Die Burushaski-Sprache von Hunza und Nager. Teil III. Wörterbuch Burushaski-Deutsch, Deutsch-Burushaski. Unter Mitarbeit von Nasiruddin Hunzai*. Wiesbaden: Harrassowitz Verlag.
- Buddruss, Georg. 1960. *Die Sprache von Wotapur und Katarqala*. Bonn: Selbstverlag des orientalischen Seminars der Universität Bonn.
- Burrow, Thomas. 1975. A new look at Brugmann's law. *Bulletin of the School of Oriental and African Studies*, Vol. 38, No. 1: 55–80.
- Cheung, Johnny. 2007. *Etymological Dictionary of the Iranian Verb. Leiden Indo-European etymological dictionary series*. Leiden / Boston: Brill.
- Francke, August H. 1907. *A history of Western Tibet: one of the unknown empires*. London: S. W. Partridge & Co.
- Gamkrelidze, Tamaz V., Vyacheslav V. Ivanov. 1984. *Indoevropskij jazyk i indoevropejcy: Rekonstrukcija i istoriko-tipologičeskij analiz prajazyka i protokultury*. Tbilisi: Izdatel'stvo tbilisskogo universiteta.
- Gershevitch, Ilya. 1954. *A Grammar of Manichean Sogdian. (Publications of the philological society)*. Oxford: Basil Blackwell.
- Grierson, George A. 1906. *The Piśāca languages of North-Western India*. London: Royal Asiatic Society.
- Grierson, George A. 1915–32. *A Dictionary of the Kashmiri Language. Vol. I–IV* ("Bibliotheca Indica". Work No. 229). Calcutta: Royal Asiatic Society of Bengal.
- Grierson, George A. 1919. *Linguistic Survey of India. Vol. VIII, pt. 2. Specimens of the Dardic or Piśāca Languages (including Kāshmirī)*. Calcutta: Superintendent Government Printing, India.
- Jettmar, Karl. 1975. *Die Religionen des Hindukusch (= Die Religionen der Menschheit. Bd. 4, 1). Mit Beiträgen von Schuyler Jones und Max Klimburg*. Stuttgart: Kohlhammer Verlag.
- Klima, Otakar. 1970. Das altind. kāṣṭhām. *Archiv orientální* 38: 26–27.
- Kogan, Anton I. 2015. Nekotojje voprosy geneticheskij klassifikatsii dardskix jazykov po dannym istoričeskij fonetiki. *Journal of Language Relationship* 13/1: 1–21.
- Kogan, Anton I. 2016. *Problemy sravnitel'no-istoričeskogo izučeniija jazyka kashmiri*. Moskva: "Fond razvitija fundamental'nyx lingvističeskix issledovanij".
- Kogan, Anton I. 2019. On possible Dardic and Burushaski influence on some Northwestern Tibetan dialects. *Journal of Language Relationship* 17/4: 263–284.
- LIV = Kümmel, Martin and Helmut Rix, eds. 2001. *Lexikon der indogermanischen Verben: Die Wurzeln und ihre Primärstammbildungen. 2nd ed.* Wiesbaden: Reichert.
- Liljegren, Henrik, Naseem Haider. 2011. *Palula vocabulary*. Islamabad: Forum for Language Initiatives.
- Livshits, Vladimir A. 2003. Drevneje nazvanije Syrdar'ji. *Vestnik drevnej istorii* 1(244): 3–10.
- Lorimer, David L. R. 1939. *The Burushaski language. Vol. III. Vocabularies and index*. Oslo: H. Aschehoug & Co. (W. Nygaard).
- Mayrhofer, Manfred. 1996. *Etymologisches Wörterbuch des Altindoarischen. II Band*. Heidelberg: Universitätsverlag C. Winter.
- Morgenstierne, Georg. 1940. Notes on Bashkarik. *Acta Orientalia* XVIII(3/4): 206–257.
- Morgenstierne, Georg. 1941. *Notes on Phalura, an Unknown Dardic Language of Chitral*. Oslo: J. Dybwad.
- Morgenstierne, Georg. 1973. Sanskritic Words in Khowar. In: Georg Morgenstierne. *Irano-Dardica*: 256–272. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Norman, Rebecca. 2010. *A Dictionary of the Language Spoken by Ladakhis*. Ms. (unpublished draft).
- Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch*. Bern / München: Francke Verlag.
- Rastorgueva, Vera S., Dzhoy I. Edelman. 2007. *Etimologičeskij slovar' iranskix jazykov. T. 3 (f–h)*. Moskva: «Vostočnaja literatura» RAN.

- Sprigg, Richard K. 2002. *Balti-English, English-Balti dictionary*. New York / London: Routledge.
- Trail, Ronald L., Gregory R. Cooper. 1999. *Kalasha Dictionary with English and Urdu compiled by Ronald L. Trail, Gregory R. Cooper*. Islamabad: National Institute of Pakistan Studies, Quaidi-Azam University, Summer Institute of Linguistics.
- Turner, Ralph L. 1966. *A Comparative Dictionary of the Indo-Aryan Languages*. Oxford University Press.
- Wackernagel, Jacob. 1954. *Altindische Grammatik. Band II, 2. Die Nominalsuffixe von Albert Debrunner*. Göttingen: Vandenhoeck & Ruprecht.
- Zemp, Marius. 2018. *A grammar of Purik Tibetan*. Leiden / Boston: Brill.
- Zoller, Claus P. 2005. *A Grammar and Dictionary of Indus Kohistani. Volume 1: Dictionary*. Berlin / New York: Mouton de Gruyter.

А. И. Коган. Об историко-фонетических особенностях индоиранских заимствований в северо-западных тибетских диалектах

Как показали недавние исследования, в лексике ряда северо-западных тибетских диалектов имеется значительное количество индоиранских заимствований. Остается, однако, неясным, были ли эти заимствования усвоены из одного или нескольких источников и является ли их наличие следствием воздействия субстрата или адстрата. Не установлено и точное генетическое положение языка-донора внутри индоиранской языковой общности. Изучение данного круга вопросов, несомненно, должно опираться на факты исторической фонетики. В статье представлена попытка выявить наиболее яркие историко-фонетические особенности индоиранских элементов в северо-западных диалектах тибетского языка. Кроме того, делается ряд предварительных выводов, касающихся лингвистической географии исследуемого региона в дотибетскую эпоху, а также направления и относительной хронологии миграций тибетцев.

Ключевые слова: историческая фонетика; языковые контакты; языковой субстрат; индоиранские языки; дардские языки; тибетские диалекты; язык ладакхи; язык балти; язык пурик.