Tocharian B etswe ‘mule’ and Eastern East Iranian*

MICHAEŁ PEYROT

1 Tocharian B etswe

Tocharian contains an archaic stratum of Iranian loanwords, as illustrated by Toch. B tsain, pl. tsainwa ‘weapon’ < PToch. *tsainu ← OIr. *dʰainu- (cf. Av. zaēmū-1 ‘baldric’). The shape of the Tocharian word is archaic because 1) it preserves the OIr. diphthong *ai; 2) the plural suffix -wa reflects the original u-stem, which is likewise reflected in Arm. zēn ‘weapon’, gen. pl. zinwe’; and 3) it shows, as ts, the intermediate stage *dʰ of the development of PIIr. *dʰ > Av. z.

To this stratum of loanwords we can add the recently identified Toch. B word etswe ‘mule’ (Peyrot 2015:222–3). This word is attested in a Tocharian B–Old Uyghur bilingual of which the Old Uyghur part has been published by Maue (2015:499–507). The relevant sets are ‘Toch. B etswentsake, rendered with Old Uyghur katırlarka ‘mules’ (dat. pl.), and the following karsuwa, rendered with Old Uyghur yiidürmi[f] ‘loaded’ (Maue 2015:506). As I argued (l.c.), karsuwa must be a preterite participle, as also shown by the Old Uyghur equivalent, but it is not formed correctly. Therefore, the Toch. B words have to be resegmented as etswentsa kekarsuwa. The preterite participle is now correctly formed, and etswentsa has become a regular Toch. B perlative plural of a new noun etswe, which means ‘mule’ on the evidence of the Old Uyghur rendering with katatr. But it is not clear to which root kekarsuwa belongs. If it stands for kekar{k}uwa, then it would mean ‘bound’, and the correspondence would be etswentsa kekarkuwa ‘bound onto mules’ : katırlarka yiidürmi[f] ‘loaded onto mules’.

---

*This research was supported by the Netherlands Organisation for Scientific Research (NWO, project number 276-70-028). I am grateful to my fellow editors as well as to Federico Dragoni (Leiden) for valuable comments on an earlier draft. For important advice and information I thank Ablet Semet (Berlin) and William Baxter (Ann Arbor).

1This *u*-stem goes back to an earlier nu-stem, see de Vaan 2000:531.

2A full edition of the bilingual by Georges-Jean Pinault, Jens Wilkens, and myself is in progress. Our collaborative work has shown that a number of my readings and remarks (Peyrot 2015) have to be revised, but the interpretation of the correspondences under discussion here is not affected.

3The Old Uyghur word katatr [qatir] is borrowed from Iranian, cf. Sogd. xtrty /xartarē ‘mule’ (Gharib 1995:170a, 432a) < *xaratara- and Khot. khudara- ‘mule’ < *xaratara- (Bailey 1979:70b–71a). The function of the suffix -tara here is ‘a kind of’, i.e. ‘a kind of donkey’, cf. Skt. aṣṭara- ‘mule’ (i.e. ‘a kind of horse’; EW 1.1.140). Turkish katatr was borrowed into Mod.Pers. as qatir (Doerfer 1965–75:3.391–3).
Toch. B etswe is obviously borrowed from an Old Iranian *atswa- ‘horse’, further seen in e.g. Av. aspa- and Khot. asśa-. At this point, I cannot say why this Iranian word was borrowed despite the fact that Tocharian also inherited the word for ‘horse’ directly from Proto-Indo-European as Á yak, B yakwe ‘horse’. Probably the word was borrowed together with a kind of horse or horse-like animal that was sufficiently different from the “normal” Tocharian horse. If the Old Uyghur rendering katuv ‘mule’ is correct, this clearly warrants a borrowing next to the inherited word for ‘horse’, but it is difficult to understand how a mule could have come to be called *atswa- ‘horse’ in the Iranian source dialect. The question has to be left open. Since in the bilingual the “etswe” is used as a beast of burden, one might think of a mule, perhaps an onager, or simply a special kind of horse that was especially suited as a pack animal.

In the following, I will use the evidence of etswe to argue that the archaic stratum of Iranian loanwords in Tocharian is not from a dialect ancestral to Khotanese, but from another Iranian dialect possibly spoken to the north or to the east of Tocharian.

2 The Proto-Indo-Iranian cluster *éz

The newly identified etswe ‘mule’ conforms to the common characteristics of the Old Iranian loanwords so far identified in Tocharian and thus belongs to the same stratum as Toch. B tsain ‘weapon’ and other previously identified items (see Schmidt 1985; Tremblay 2000b; Isebaert 1980:82–123). The relevance of etswe lies in the preservation of the cluster *ts, which reflects Proto-Indo-Iranian *éy. The exact Proto-Iranian outcome of this Proto-Indo-Iranian cluster is debated. The cluster *ts, which is needed for the Old Iranian source dialect in order to explain the Tocharian form, is sometimes posited for Proto-Iranian as well (e.g. Schmitt 1989:27), even though it is not directly attested in a single Iranian language. In Avestan, the reflex of Proto-Indo-Iranian *éy is sp, e.g. aspa- ‘horse’, as in most other languages; in Old Persian it is š, e.g. asa- ‘horse’; and in Ossetic it is ñ, e.g. Dig. øfs æ ‘mare’ (<PIIr. *Haçya), probably from earlier sp. However, in Khotanese, Tumšuqese, and Wakhši the reflexes of the cluster are the palatals š, š, and ž, respectively, e.g. Khot. asša- ‘horse’, Tumš. biša- ‘all’ (~ Khot. bissä- < PIIr. *nučya), and Wakhši yaš ‘horse’ (Windfuhr 2009:18).

Usually the problem of the palatal reflex of *éy and its voiced counterpart *yu in Khotanese, Tumšuqese, and Wakhši is not addressed directly. Scholars focusing on Khotanese reconstruct Proto-Iranian or Proto-Indo-Iranian *éy (e.g. Emmerick 1989:216; Windfuhr 2009:18), while others postulate *ts, putting the conflicting evidence of Khotanese in brackets (Schmitt 1989a:27; Skjærvø 1989:375). A clear statement is that of Sims-Williams (1998:136), with which I fully agree: “Since the palatals ž, š can hardly be derived from *ts and *dz, it is simplest to assume Common Iranian *žw and *ju.”

4 For the sake of clarity I write the nonsyllabic counterpart of *u as *y. In reality, *u and *y were still allophones in Proto-Indo-Iranian and Common Iranian.

5 Kellens (1989:43) writes ţ, ż, but this reconstruction is not very widely accepted.

6 He further argues, “The palatal nature of II *č < IE *k seems also to have been preserved up to the Common Iranian stage in the case of the cluster *čr, cf. Khot. šram- [ʃra-] ‘good’ (= Avest. šrana-, OInd. śrīna- ‘beautiful . . . ’) (= Avest. šram- [ʃra-] ‘good’ (= Avest. šrana-, OInd. śrīna- ‘beautiful . . . ’)” (I.e.; similarly Emmerick and Skjærvø 1982:117). I doubt that this is correct. The normal development of PIIr. *čr in Khotanese is g (Emmerick 1989:215) as in guini- ‘hip’ < *ćrauni-, guiva- ‘fame’ ~ *ćraua- (Av. srauab-), gisga- ‘held’
The importance of this line of argument can hardly be underestimated. Since Khotanese š cannot be a secondary development of an intermediary *tsw < *éty, the only conclusion can be that Toch. B etsw was not borrowed from a dialect ancestral to Khotanese. Tocharian tsw points to *tsw or *dzw in the source, and nothing else: a palatal sibilant *š or a palatal affricate *č would have been represented by Toch. š or c [č] (pace Tremblay 2003a:678). Therefore, the Old Iranian stratum cannot be identified as “Old Sakan,” a term used by Tremblay for the ancestor of Khotanese, Tumšuqese, and Wakhī (Tremblay 2003b:422). Instead, it points to another Old Iranian dialect in which, as in most Iranian dialects, depalatalization of *éty and *fy to *tsw and *dzw did take place.

In theory, several scenarios are possible. For instance, one could imagine that Khotanese and Tumšuqese are relative newcomers in the Tarim Basin, and that they recently moved east from the Wakhan corridor, where the apparently more closely related Wakhī is spoken until today. Before Khotanese and Tumšuqese arrived, another dialect that better fits the archaic stratum in Tocharian may have been spoken in the Tarim Basin. It is also conceivable that the Tocharians had been in contact with an archaic Iranian dialect before they moved into the Tarim Basin. Finally, when the Tocharians had already settled in the Tarim Basin, there may have been Iranian speaking groups to their north, perhaps also to their east.

Obviously, the three options just mentioned are not exhaustive, and other variants could also be considered. In order to narrow down the number of possibilities, I will now first turn to the prehistory of Khotanese.

3 The prehistory of Khotanese: Linguistics

Little is known about the prehistory of Khotanese. Bailey (1970:68) writes, “it is likely that the two languages of Khotan and Tumshuq were spoken by two tribes of the Saka who about 200 BC or earlier settled with a monarchical or oecarchical system in this region of the Taklamakan. From the second century BC there is Chinese information on Khotan: no major invasion is recorded.” This gives us a date ante quem for the arrival of the Khotanese, and it is in accordance with the four legendary “accounts of the foundation of Khotan, all of which associate it with the son and ministers of the emperor Aśoka. This would place the foundation of Khotan firmly in the third century BC” (Emmerick 1992:1). As Emmerick points out, the evidence of these legendary accounts is weakened by the fact that they are only from the 7th century CE and later, thus leaving a gap of a thousand years between the foundation and its earliest account. He is cautious about the date of the arrival of Iranians in Khotan: “Nomadic tribes speaking languages of Iranian origin must have been wandering about Central Asia from a very early period, probably from the first half of the second

< *śrīša- (Av. snaē-). On the other hand, PIIr. *śi develops into Khot. ši, cf. e.g. šiśa- ‘copper, copper-colored’ < *śīsā- (Av. sīśā-). It is more likely that in *śīra- dissimilation to *śīra- took place (cf. Skt. śīra- ~ śīra-); with shortening of ō, *śīra- would regularly yield the attested Khot. śīna-. Dissimilation of the *-r- is also assumed by Emmerick (Emmerick and Skjærvø 1982:117) and Skjærvø (2004:11, 346b), both different in detail.

*In view of the development PIIr. *jī > Khot. š (see preceding note), I suppose that an intermediate stage was *ji, i.e. *éty > *jī > š. If the development of this cluster is interpreted in this way, it more clearly is an innovation, not an archaism, and therefore a strong argument for a Khotanese-Tumšuqese-Wakhī branch.
millennium BC, but we do not know when they first began to settle in permanent villages” (1992:2).  

An important aspect of the prehistory of Khotanese obviously is its exact relationship to Tumšuqese. It is generally accepted that the two languages are closely related, and indeed, a relatively long list of shared features in phonology, morphology, and lexicon can be made in spite of the poor attestation of Tumšuqese (cf. e.g. Bailey 1958:148–50). In most cases, the languages are very close, as for instance with the examples cited by Emmerick (1968:3), which are, however, especially selected to show their close relationship. Any differences in Emmerick’s examples are due only to different spelling conventions, e.g. Tumš. ´sazda- ‘snake’ ∼ Khot. šaysida- /śazda-./.

Bailey (1970:68; cf. also 1958:134–5), who calls the two varieties “dialects” that are “clearly of one language,” gives three examples to show that there are “striking differences”: Tumš. rorda- ‘given’ ∼ Khot. būda-; Tumš. šowarsana ‘it’ ∼ Khot. šūdasu; Tumš. patsasu ‘so’ ∼ Khot. pamjāsā. If we reconstruct these items for Proto-Tumšuqese-Khotanese, we do not seem to get a totally different language, but rather one that is relatively close to both daughter languages: *brorda- < *frī-bṛtṛ-; *śŏ-dasu vs. *śŏ-baru(h)-dasana < *parah-dasa- (Konow 1935:783; Emmerick 2009:390); *patsāsV < *pānčāsAt. For the differently formed numeral ‘1’, compare the higher numerals of Khotanese, e.g. šiūvare-bistā ‘1’ < *śŏ-baru(h)-wist-; Tumšuqese ‘1’ may easily be analogical after the higher numerals.

Other correspondences can be adduced to argue that the difference between the two languages is relatively large. In Tumš. mrida- ‘died’ ∼ Khot. mūda- and Tumš. zriđa- ‘old’ ∼ Khot. yādā- syllabic *r seems to have different reflexes, so that one would have to reconstruct syllabic *r for their common ancestor, i.e. *mīḍa- and *zrīda-. However, in both cases the Tumšuqese ri could probably be secondary. This is strongly suggested for zriđa- because Khot. yādā- < *zarta- seems to correspond perfectly to Av. zārta- < *ṛṛH-ta- (on Ved. jārīna- and jīra-, see Lubotsky 2007:232–3). Another case in point is the verb ‘do’. In Khotanese, all forms can be derived from a root yan-, ultimately < *κṛnu-, *κρναu- except for the preterite participle yādā-, yūdā- < *kṛtṛA- (Emmerick 1968:110–1). In Tumšuqese, the verb is ar-, ultimately also < *κṛnu-, *κρναu- (Maue 2016:126). Here the difference between the two languages is due to a special development in Khotanese: “No certain example of *rn- > -an- is found in Kh., where the normal treatment would be > -arr-.” (Emmerick 1968:112). Thus, the present of ‘do’ can be reconstructed as *(y)aṟn-. Nevertheless, the correspondence forces us to reconstruct *rn, which would otherwise not have been obvious from the regular reflex Tumš. rr, Khot. rr < *rn (Konow 1935:789). Finally, in Tumš. uzanvara-, usanāvara- ‘being’ ∼ Khot. uṣnora- < *uṣana-bara- (Bailey 1979:39b) or *uṣāna-bara-, the different contractions of *uṣa require a Proto-Tumšuqese-Khotanese *uṣanašbara-: a reconstruction *uṣanāora- would give us the Khotanese form, but not the Tumšuqese. In light of

---

8 He later suggested that the speakers of Proto-Tumšuqese-Khotanese settled “in oases around the Tarim basin some time during the second half of the first millennium BCE” (Emmerick 2009:377), but adduced no arguments in support of this idea.

9 There is no need to reconstruct *frī- for Proto-Tumšuqese-Khotanese since both languages agree in their special treatment of *frī- in this preverb instead of the otherwise regular development *frī- > *br- as in Khot. bṛiṣa- ‘dear’, Tumš. bṛi from the OIr. root *frī- (Konow 1935:787).

10 As Federico Dragoni points out to me (p.c.), an intermediate stage of the weakening of k- to y- seems to be preserved by the Old Khotanese spelling g- (e.g. Emmerick 1970:128a).
this correspondence, it seems that the monophthongization of *au < *au, *a3 to Tumš. o, Khot. ú was an independent process. Consequently, Tumš. rorda- ‘given’ ~ Khot. hūja-, reconstructed above as *brorda-, is probably to be reconstructed instead as *braurda- or *brauurdzda- < *brażda-. All in all, Konow (1935:797) seems to be quite right with his statement, “Die Zeit der Trennung kann nicht allzu kurz bemessen werden. Wir wissen aber nicht, wie lang sie gewesen ist.”

4 The prehistory of Khotanese: Historical sources

Konow further speculates (1935:801) that the speakers of the language ancestral to Khotanese and Tumšuqese may be identified with the Sāi塞 of the Chinese sources, who in the second century BCE moved south from Ili, the steppe area north of the Tian Shan mountains and the Taklamakan desert. According to him, this would account for the fact that the more archaic Tumšuqese is found in the north, while the speakers of Khotanese went further south. The Sāi were certainly seen as a Saka tribe by the Chinese historiographers, and their name derives from the word saka: sāi 塞 is from Early Middle Chinese sok (Pulleyblank 1991:273), Middle Chinese sok (Baxter and Sagart 2014:230; ɔ stands here for ə or ʌ, see o.c. 13). In view of the generally accepted identification of Khotanese and Tumšuqese as “Saka” languages (on which see further below), Konow’s suggestion is an obvious option that should be discussed.

The migration of the Sāi that Konow refers to is found in the Hanshu (the History of the Former Han Dynasty): “At the time the Yuè-chih [Yuèzhī] had already been defeated by the Hsiung-nu [Xiōngmù]; making for the west they attacked the king of the Sai. The king of the Sai moved a considerable distance to the south and the Yuè-chih then occupied his lands” (Hulsewé 1979:216). Thus this passage describes the Yuèzhī fleeing for the Xiōngmù and driving off the Sāi towards the south. Since the Sāi had been living north of the Tian Shan mountains, a logical inference is that they moved south into the Tarim Basin and later became the speakers of Khotanese and Tumšuqese. However, a parallel passage is more specific and mentions that the Sāi moved south through the “Suspended Crossing,” a passage through the Pamirs: “When the Ta Yuè-chih [Ta Yuèzhī] turned west, defeated and expelled the king of the Sai, the latter moved south and crossed over the Suspended Crossing; and the Ta Yuè-chih took up residence in his lands” (Hulsewé 1979:144). That their route led south through the Pamirs and past the Tarim Basin is confirmed by the fact that some ended up in Jībīn 剪賓 ‘Kashmir’: “it was in these circumstances that the king of the Sai moved south and established himself as master of Chi-pin [Jībīn]. The Sai tribes split and separated and repeatedly formed several states. To the north-west of Shu-lo [Shūlè], states such as Hsiu-hsūn [Xiǔxūn] and Chiān-tu [Juāndú] are all of the former Sai race” (Hulsewé 1979:104–5).

The latter passage is especially informative. Shūlè 疏勒 ‘Kashgar’ is used as a reference point to locate Saka tribes further to the northwest, but there is no mention of Sakas in Kashgar itself. Likewise, neither Tumšuq nor Khotan is connected with the Sāi, although

---

14In the meaning ‘Saka’, 塞 may be read sāi or sè in standard Mandarin; both readings go back to Middle Chinese sok (William Baxter, p.c.). I will use only “Sāi,” which seems to be more common in the literature.
both places were certainly much better known than Kashmir, Xiūxún 休循 or Juândú捐毒. The Sâi of the Hanshu are therefore definitely not to be identified with the ancestors of the speakers of Khotanese and Tumšuqese. Moreover, these southward migrations of the Sâi in the 2nd century BCE are in any case too late for Khotanese, since no invasion or massive repopulation is recorded for Khotan from the 2nd century onwards (see the quote from Bailey above).

Finally, Bailey (1958:132) proposes an etymology for Juândú捐毒, one of the Saka states, that is daring, but does not seem impossible to me. The Early Middle Chinese reading of this name, evidently based on a pronunciation variant Yuândú, is *jwian-dówk (Pulleyblank 1991:166, 82), Middle Chinese *ywen-dówk (Baxter and Sagart 2014:249 and the related online database), which is close to what Morgenstierne reconstructs for *Idry, *Yidg, the tribe name of the Yidgha: *Indug ≈ *(H)induka- (1938:16). An obvious problem with this etymology is that Yidgha is spoken southwest of Kashgar, not northwest. One would therefore have to assume that they moved south later on, or that a tribe with this name split in two. However, in view of the trajectory of the Sâi, it is conceivable that some of the Pamir languages (with the exception of at least Wakhı) are remnants of this southward migration.

5 The prehistory of Khotanese: Archaeology

If the Sâi cannot be the ancestors of the speakers of Khotanese and Tumšuqese, the common origin of these two languages must be sought further back in time. My hypothesis is that the arrival of Khotanese and Tumšuqese in the Tarim Basin is to be dated earlier than often assumed, and that their predecessors can be identified with an archaeological culture termed Ākètālā/Aqtala.

According to Debaine-Francfort (2001:66a), the Ākètālā/Aqtala culture dates from about 1,000 BCE, lasting apparently at least to 650–550 BCE (Debaine-Francfort 1988:24b). It was located in the west and south of the Tarim Basin and is characterized by gray ceramics and curved stone sickles. The gray ceramics extend to the east from Kashgar (the region where the site Ākètālā/Aqtala itself is located) until Kuča in the north (Kuča itself being different, belonging to an area with painted ceramics) and from Kashgar to Qiēmò/Čärčän in the south (cf. also the map in Idriss and Debaine-Francfort 2001:14). The sickles are found from Kashgar to Ākès¯u/Aqsu and even Kuča in the north and from Kashgar to Mínf¯eng (Niya) in the south.

In her 2001 article, Debaine-Francfort does not discuss possible origins of the Ākètālā/Aqtala culture, but she points out that its typical vessels are parallel to finds in eastern Kazakhstan at Malokrasnojarka, Trušníkovo, and Ust'-Narym, all of which belong to Eastern Andronovo according to Kuz’mina (2007 passim). This is no proof that the Ākètālā/Aqtala culture was Iranian-speaking, but it is a serious option. Mallory interprets the Ākètālā/Aqtala culture as evidence for Iranians in the region too (2015:29; cf. also Francfort 2001:228–9). An important site from the middle of the 1st millennium BCE linked to the Ākètālā/Aqtala culture is Jümblulaq Qum on the Keriya River (Debaine-Francfort and Idriss 2001a:135; Francfort, Idriss, and Zhang 2001:137).

12In French transcription, Djoumboulak Koum. The Chinese name is Yuánshā圆沙.
Figure 1. The find sites of gray (marked with ☰) and painted (marked with ☱) ceramics, based on the map in Idriss and Debaine-Francfort 2001:14. Only the names of the most important sites have been added here. The area of painted ceramics extends east off the map to Yiwú and Hámì.

Even if the Ākètǎlā/Aqtala culture represents an early wave of Iranians, we cannot be sure that this culture was ancestral to the later Khotanese- and Tumšuqese-speaking peoples as long as no solid link can be established with the kingdoms of the early 1st millennium CE, such as Karadong on the Keriya River (Debaine-Francfort, Idriss, and Wang 1994) or Khotan itself. However, the geographical distribution fits the location of Tumšuqese and Khotanese very well, and especially the border between the gray ceramics of the Ākètǎlā/Aqtala culture and the painted ceramics from Kuča onwards to the east coincides perfectly with the border between Tumšuqese and Tocharian B. Moreover, such an early migration into the Tarim Basin would account for the fact that Khotanese-Tumšuqese, probably together with Wakhī, split off relatively early from Proto-Iranian in view of the reflex of PIIR. *ču̯ as Khot. ĺ∙ (see above). The early date of the arrival of this Iranian group in the Tarim Basin does not necessarily mean that the Khotanese-Tumšuqese proto-language would have to be dated around 1000 BCE. I think that such a date is possible in view of the differences between the two languages, but it cannot be excluded that the breakup of the proto-language was later, for instance around 500 BCE. At all events, a date long before 1000 BCE seems unlikely.

In my view, it is too early for a linguistic interpretation of the Ākètǎlā/Aqtala area east of Khotan and the Keriya River. It is to be expected that this area was Iranian-speaking before Middle Indian was introduced there, but the details and consequences of this hypothesis need further study. Nevertheless, we might conceive of the legendary foundation of Khotan in the time of Aśoka as referring to the arrival of Indians in the southern Tarim Basin, possibly speakers of Gândhārī. Their Middle Indian dialect evidently became the language of culture and remained so until the middle of the 1st millennium CE. The arrival of Indians in the Tarim certainly need not have taken place exactly during the reign of Aśoka, but it
is possible that it predated the introduction of Buddhism, as stated in the accounts of the history of Khotan.

6 Khotanese and Sogdian

Another aspect that is relevant to the prehistory of Khotanese is its relationship with Sogdian. Sims-Williams (1989a:169–70) lists a number of common innovations of Sogdian and Khotanese. These mostly concern lexical correspondences, such as Sogd. kntθ ‘city’ ~ Khot. kantbä; Sogd. rθf ‘disease’ ~ Khot. rräba- ‘pain’; Sogd. mθsky ‘palm of the hand’ ~ Khot. nänärna-; Sogd. dr̥w nk’r’k ‘maleficent’ ~ Khot. diramggāra-; Sogd. ñyr’nk’r’k ‘beneficent’ ~ Khot. šāramggāra-; Sogd. ʃm to send’ ~ Khot. baʃem-, < *fra-jāmaya- (Sims-Williams 1983:50); Sogd. prtw ‘time’ ~ Khot. bād-a- (Sims-Williams 1983:49). He further points out the similarity between the 2sg. personal pronoun Sogd. tvwr ~ Khot. thu (Tumš. thu), which according to him both reflect *tuθu or *tuθu < *tuθu < *tuθam (1983:48); the suppletive stems in the demonstrative Sogd. š/-t ~ Khot. ša/-tta-; and the loss of k in the aka-stems and *kar- ‘do’ (Sims-Williams 1990).

A Khotanese-Sogdian subbranch within Iranian is not compatible with the obviously early change of PIIr. *cū to Khotanese ʃ, Tumšuqese ʃ, and Wakhši ʃ, where Sogdian has sp instead (Yaghnobi ᵁp), and it is therefore likely that the above Khotanese-Sogdian isoglosses are due to parallel developments. This is compatible, I think, with the fact that the list includes no early sound changes, the loss of k probably being a relatively late development. Nevertheless, it does seem likely that Khotanese and Tumšuqese took part in the development *ft, *xt > *βt, *γt, virtually the only shared sound change of “East Iranian” (Sims-Williams 1989a:167–8), even though direct evidence for these changes in Khotanese is lacking. The development of *aft to Khot. aud (often od) is likely to have gone through *avt > *avd; here the change of *vt to *vd would have to be a later development than *abu, *au > Khot. ʊ, but this seems unproblematic. In addition, it may be noted that intervocalic *t develops into t (however this was exactly pronounced), e.g. OKhot. mäta ‘mother’ < *māta, while *ft, *xt yielded d, e.g. Khot. loda ‘γ’ < *hafθa, which suggests an intermediary *vd, *yd. Similarly, the sequence *uxt became ṣy, e.g. ᵁyθa- ‘learned’ < *sxɔxta- (with *ax or *aγ > i3 and ida > iy). Probably the development *ft, *xt > *βt, *γt could spread through the early East Iranian dialects because it was in origin just a special phonetic characteristic without phonological consequences that could be reverted (cf. Yaghnobi f, x vs. Sogd. β, γ).

If the parallels between Khotanese and Sogdian adduced by Sims-Williams are not due to a genealogical relationship, they must be due to later, but in all probability still quite early, contacts. Perhaps these contacts have to be seen in the light of the parallels between the vessels of the Ākētalā/Aqtala culture and those of the sites in eastern Kazakhstan mentioned above. Obviously, we also have to reckon with shared influence from a third source, as for instance in the case of Sogd. fīn ‘glory’ and Khot. phāra- ‘fortune, (high) position’, both probably borrowed from a steppe dialect, i.e. “Scythian” *fārāb- (Lubotsky 1998, 2002).

Cf. also ʃtira- ‘bitter’ < *taxra- (Emmerick 1989:214) and ʃim- ‘content’ < *aŋgra- (Bailey 1979:315a).
7 Khotanese and Saka

The speakers of Khotanese and Tumšuqese were different from the people called S¯ai in the Hanshu (see above). Nevertheless, the two languages are widely accepted to be Saka languages. As noted by Emmerick, “The word Saka- is not certainly attested in the Khotanese documents” (Emmerick 1968:2; see also Bailey 1970:68). The Old Khotanese term for ‘Khotan’ and ‘Khotanese’ is hvotana- (Emmerick 1968:1). The Tumšuqese term for Tumšuqese was gyāzdi- (Rong 2009). The first to suggest that Khotanese is a “Saka” language was probably Le Coq, who supposed it “to be, in consequence of geographical and historical considerations, the lost language of the Saka” (1909:318). Nevertheless, Leumann initially termed it “nordarisch” (1912), and it was only with the seminal paper about the language of the ´Sakas by Lüders (1913) that Khotanese became widely seen as “Saka” (e.g. Konow 1935:772).

Lüders’ paper contains several layers of argument. He identified the digraph ⟨ys⟩ in inscriptions from North India and argued that its value is [z]. Since the same remarkable digraph is used with the same value in Khotanese, while it is used in India from the reign of Caṣṭana, well before the earliest Khotanese manuscript of probably the 5th or 6th century, he assumed that it came from India to the Tarim Basin (1913:409). He further argued that the North Indian Śaka era (started 78 CE; Falk 2012) was not a legendary era but a historical dynasty, and that traces of the Iranian language of the Śaka rulers are found in several inscriptions from North India. These conclusions are beyond any doubt.

Lüders then argues that the Iranian language of the Śakas was identical to Khotanese (called “nordarisch” at the time) and that Khotanese was therefore a Saka language. It is this latter conclusion that should perhaps again become part of the scholarly discussion. His main arguments are:

1. The nom. sg. m. of the Śaka language ends in -i (pp. 419–20).

This feature does not necessarily point to Khotanese, because a nom. sg. in -i is also found in the Sogdian light stems (Sims-Williams 1989b:183), and, in fact, the oldest Khotanese nom. sg. is -ā, not -i. This may seem a detail, but since the Śaka language is older than Old Khotanese, it should be noted. It can also be pointed out that Aramaeo-Iranian may attest a nom. sg. in -i in the word krpty (Sims-Williams 1989a:165). The final -y here may simply be the nom. sg. ending instead of being parallel to Old Persian paθi-.

2. The gen. sg. m. ends in -ī ⟨ei⟩ (pp. 423–5).

Indeed, a gen. sg. m. in -i is found in Khotanese, while the Sogdian light stem gen.-dat. sg. ending is to be read -e, not -i (Sims-Williams 1989b:183; pace Lüders 1913:424).

3. The word ḍīnīka is related to Av. daēnā- and shows the sound development ai > i (p. 414).

In the Indian tradition, the word begins with ´s- rather than the expected s-. Nevertheless, it must in the end be the same etymon.

According to Lüders, Caṣṭana’s reign was in the second quarter of the 2nd century CE. Falk now suggests that Caṣṭana was the first Śaka ruler and that his rule began in 78 CE (2012:132a).

This early date is argued for by Maggi (2004) for Zambasta fragment T III S 16.
The monophthongization of ai to i is found in Khotanese but is also widespread in other Middle and Modern Iranian languages; cf. for instance Yidgha (Morgenstierne 1938:98–9).

4. The personal names Āduthuma and Šodasa, Šudasa, Šudisa have ɗ (pp. 414, 420).

Intervocalic retroflex t, d are found in Khotanese, but in Yidgha, for instance, retroflex r derives from *rt probably through *d (Morgenstierne 1938:79). The occurrence of ɗ alone cannot prove that Śaka is a form of Khotanese.

5. The first element of the personal name Ysamotika is zama- ‘earth’ and -ot- reflects the suffix -vant-, -vat- (p. 413).

zama- ‘earth’ is an Iranian word and is also attested in Khot. y-sama-śandāa- ‘world’, but evidently not limited to it. In *zamavant-, *zamavat- the sequence -ava- would in Khotanese certainly contract to o, au, but in many other languages as well.

6. The meaning of the personal name Usavadāta, Usabhadata, Usabhadaata is ‘Ṛṣabhadhāma’. The first element is a borrowing from Skt. rṣabha- and the second element is Iranian dāta- ‘law’ (pp. 413–4).

The element dāta- is too widespread to be of any significance, while the substitution of us-, us- for ṛṛ- is not typical for Khotanese; cf. e.g. irdi ‘miracle; ṛddhi’.

7. The first element of Kharaosta could be a Prakritized from of *ksāra- < *ksāθra- (p. 420).

Khot. ksāra- ‘power, dominion’ is now attested (Bailey 1979:67a), but the interpretation otherwise remains truly uncertain.

8. The element bola-, bora- in several names means ‘gift’, Skt. dāna- (pp. 420–3).

Indeed, Khot. baura- ‘gift’ shows the development of *fra- to ha- that is typically found in Khotanese, and different even from the reflex in Tumšuqese. Should the bora-, bola- of the inscriptions really be the equivalent of Skt. dāna-, then it would point to Khotanese influence.

9. The fact that the Kushan title sāhānu sāhi ‘king of kings’ does not show the expected Kushan (i.e., Bactrian, cf. paovava pao) morphology could be taken to mean that the northern Kṣatrapas were Kushans, not Sakas. However, sāhānu sāhi can be explained as an adaptation of the Kushan title to the morphology of the language of the older Śakas (pp. 425–6).

This assumption is far-fetched, but may be necessary even if Śaka is not identical to Khotanese. In the older Bactrian of the Kushans, the ending -i is oblique, not direct.

In my view, Lüders’ observations are indeed compatible with the assumption that the Śakas spoke a form of Khotanese. However, it is questionable whether the features identified are unique to Khotanese, certainly in view of the fact that the non-Khotanese Sāi migrated as far south as Kashmir according to the Hanshu.
In sum, the evidence that the Śaka language is Khotanese or an earlier form of it is weak. Many of the features are found in other languages as well, and it is known from other sources that non-Khotanese Iranians found their way to northern India. In any case, the large number of Indic elements in Khotanese is no proof “daß das ‘Nordarisch’ sprechende Volk längere Zeit auf indischem Boden saß” (Lüders 1913:413), since there is ample evidence that instead speakers of Middle Indian migrated into the Tarim Basin. Should any of the elements adduced by Lüders really stem from Khotanese, then it would seem much more attractive to me to view them as influences from Khotan in northern India. For instance, in the “Prophecy of the Li country” a joint campaign of Kanishka (reigned 127–50 CE; Falk 2012) and the kings of Khotan and Gu-zan to Saketa is mentioned (Emmerick 1967:47, 94b).

8 The archaic Iranian stratum in Tocharian

If the archaic Iranian stratum in Tocharian cannot be identified as an early form of Khotanese or Tumšuqese-Khotanese, despite the fact that this branch seems to have been present in the Tarim Basin since the early 1st millennium BCE, we have to look elsewhere. We do not have to search far. Not only were the Sāi of the Hanshu originally found to the north of the Tocharian-speaking area, other groups are reported to have lived to their east: the Wūsūn and the Yuèzhí. The Yuèzhí were apparently living in the Hēxī corridor, which is east of the Tocharian towns Kuča and Yānqí/Qarašähār. Although the linguistic identity of the Wūsūn cannot be proved, they may well have been Iranian. This is very likely to be true of the Yuèzhí in any case since they invaded Afghanistan later, and their language is suggested to be a form of Scythian by Sims-Williams (2002). Even though the Chinese sources about all three peoples refer to a period that is probably far too late for such archaic-looking borrowings as *etswe ‘mule’, namely only from the 2nd century BCE onwards, it is clear that there is no need to look for Iranians in the Tarim Basin only. According to Kuz’mina, Andronovo influence in Xinjiang is as early as the 13th–9th centuries BCE (2008:98–107).

Thus from the scenarios mentioned at the end of §2, the most likely in my view is that the western Tarim Basin was populated from an early period onwards by Iranian peoples ancestral to the later speakers of Khotanese and Tumšuqese, while the archaic Iranian stratum in Tocharian derives from an Iranian dialect spoken to the north or east of the Tocharian area.

It is impossible to say with which Iranian group the archaic Iranian stratum must be identified exactly. The words are so archaic that they show hardly any dialect features. Nevertheless, it can be pointed out that Toch. B waipece ‘possessions’ from *hwai-pahya- shows a palatalization of *fy to ce that must be attributed to the Iranian source since in Tocharian *fy develops into fi. This palatalization is not found in the Pamir languages or Tumšuqese-Khotanese, and not in the same form in Sogdian either, but it is found in Ossetic and may be an early dialect feature of the steppe Iranian that we could call “Scythian” (Lubotsky 2002:198).

Abbreviations

References

Michaël Peyrot


Pulleyblank, Edwin G. 1991. Lexicon of Reconstructed Pronunciation in Early Middle Chinese,
Late Middle Chinese, and Early Mandarin. Vancouver: UBC Press.
In Studia Linguistica, Diachronica et Synchronica: Werner Winter sexagenario anno
MCMLXXIII gratis animis ab eius collegis, amicis discipulisque oblata, ed. by Ursula Pieper
Sims-Williams, Nicholas. 1983. “Chotano-Sogdica.” Bulletin of the School of Oriental and
— — —. 1990. “Chotano-Sogdica II: Aspects of the development of nominal morphology in
Khotanese and Sogdian.” In Proceedings of the First European Conference of Iranian Studies,
part 1: Old and Middle Iranian Studies, ed. by Gherardo Gnoli and Antonio Panaino,
documents and inscriptions.” In Indo-Iranian Languages and Peoples, ed. by Nicholas
Tremblay, Xavier. 2005a. “Bildeten die iranischen Sprachen ursprünglich eine genetische
Familie oder einen Sprachbund innerhalb des indo-iranischen Zweiges?” In Sprachkontakt
Wiesbaden: Reichert.
Windfuhr, Gernot. 2009. “Dialectology and topics.” In The Iranian Languages, ed. by Ger
Farnah

Indo-Iranian and Indo-European Studies

in Honor of

Sasha Lubotsky
# Table of Contents

**FARNAH**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>vii</td>
</tr>
<tr>
<td>Bibliography of Sasha Lubotsky</td>
<td>ix</td>
</tr>
<tr>
<td>Ph.D. Students of Sasha Lubotsky</td>
<td>xvi</td>
</tr>
<tr>
<td>List of Contributors</td>
<td>xvii</td>
</tr>
<tr>
<td>Peter C. Bisschop, Vedic Elements in the <em>Pāñcapatasūtra</em></td>
<td>1</td>
</tr>
<tr>
<td>Václav Blažek, The Case of Tocharian ‘silver’: Inherited or Borrowed?</td>
<td>13</td>
</tr>
<tr>
<td>Michiel de Vaan, The Noncanonical Use of Instrumental Plurals in Young Avestan</td>
<td>21</td>
</tr>
<tr>
<td>Desmond Durkin-Meisterernst, Sogdian Plurals in the Vessantara Jātaka</td>
<td>37</td>
</tr>
<tr>
<td>Jost Gippert, A Middle Iranian Word Denoting an Office-Holder</td>
<td>43</td>
</tr>
<tr>
<td>Stephanie W. Jamison, The Vedic Perfect Imperative and the Status of Modal Forms to Tense-Aspect Stems</td>
<td>55</td>
</tr>
<tr>
<td>Michael Janda, Vedisch <em>dhēnā</em>-: Bedeutung und Etymologie</td>
<td>64</td>
</tr>
<tr>
<td>Jay H. Jasanoﬀ, The Phonology of Tocharian B <em>okso</em> ‘ox’</td>
<td>72</td>
</tr>
<tr>
<td>Jared Klein, Syncretism in Indo-European: A Natural History</td>
<td>79</td>
</tr>
<tr>
<td>Alwin Kloeckhorst, The Origin of the Hittite <em>bi</em>-Conjugation</td>
<td>89</td>
</tr>
<tr>
<td>Werner Knobl, Das Demonstrativpronomens <em>ETÁD</em> im Rgveda</td>
<td>107</td>
</tr>
<tr>
<td>Petr Kocharov, A Comment on the Vocalization of Word-initial and Medial Laryngeals in Armenian</td>
<td>123</td>
</tr>
<tr>
<td>Frederik Kortlandt, The Indo-European <em>k</em>-Aorist</td>
<td>137</td>
</tr>
<tr>
<td>Guus Kroonen, Lachmann’s Law, Thurneysen’s Law, and a New Explanation of the PIE <em>no</em>-Participles</td>
<td>143</td>
</tr>
<tr>
<td>Leonid Kulikov, Vedic <em>āhanās</em>- and Its Relatives/Cognates within and outside Indo-Iranian</td>
<td>153</td>
</tr>
<tr>
<td>Author</td>
<td>Title</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Martin Joachim Kümmel</td>
<td>The Survival of Laryngeals in Iranian</td>
</tr>
<tr>
<td>Rosemarie Lühr</td>
<td>Prosody in Indo-European Corpora</td>
</tr>
<tr>
<td>Hrach Martirosyan</td>
<td>Armenian <em>Andpayin oj</em> and Vedic <em>Abhi-Budhyä-</em> ‘Abyssal Serpent’</td>
</tr>
<tr>
<td>Ranko Matasović</td>
<td>Iranian Loanwords in Proto-Slavic: A Fresh Look</td>
</tr>
<tr>
<td>H. Craig Melcher</td>
<td>Semantics and Etymology of Hittite <em>takš</em></td>
</tr>
<tr>
<td>Benédicte Nielsen Whitehead</td>
<td>PIE <em>gʰwɛ₂-ɛy</em>- ‘cow’</td>
</tr>
<tr>
<td>Alan J. Nussbaum</td>
<td>A Dedicatory Thigh: Greek <em>µηρός</em> and <em>µεταρ</em> Once Again</td>
</tr>
<tr>
<td>Norbert Oettinger</td>
<td>Vedisch <em>Vivāsvant</em> und seine avestische Entsprechung</td>
</tr>
<tr>
<td>Birgit Anette Olsen</td>
<td>The Development of Interconsonantal Laryngeals in Indo-Iranian and Old Avestan <em>zäβä ptä</em></td>
</tr>
<tr>
<td>Michaël Peyrot</td>
<td>Tocharian B <em>etswe</em> ‘mule’ and Eastern East Iranian</td>
</tr>
<tr>
<td>Georges-Jean Pinault</td>
<td>New Look at Vedic <em>sáṃ</em></td>
</tr>
<tr>
<td>Tijmen Pronk</td>
<td>Old Church Slavonic <em>(j)utro</em>, Vedic <em>usár</em>- ‘daybreak, morning’</td>
</tr>
<tr>
<td>Velizar Sadovski</td>
<td>Vedic and Avestan Parallels from Ritual Litanies and Liturgical Practices I</td>
</tr>
<tr>
<td>George Starostin</td>
<td>Typological Expectations and Historic Reality: Once Again on the Issue of Lexical Cognates between Indo-European and Uralic</td>
</tr>
<tr>
<td>Lucien van Beek</td>
<td>Greek <em>πέδιλον</em> ‘sandal’ and the Origin of the <em>e</em>-Grade in PIE ‘foot’</td>
</tr>
<tr>
<td>Michael Weiss</td>
<td>Veneti or Venetes? Observations on a Widespread Indo-European Tribal Name</td>
</tr>
</tbody>
</table>

Index Verborum .................................................................................. 359