



Universiteit
Leiden
The Netherlands

Watañi lāntaṃ: Khotanese and Tumshuqese loanwords in Tocharian

Dragoni, F.

Citation

Dragoni, F. (2022, April 13). *Watañi lāntaṃ: Khotanese and Tumshuqese loanwords in Tocharian*. Retrieved from <https://hdl.handle.net/1887/3283437>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3283437>

Note: To cite this publication please use the final published version (if applicable).

5. SUMMARY AND CONCLUSIONS

5.1. SHORT SUMMARY

This study investigated the linguistic contacts between Tocharian A and B and Khotanese and Tumshuqese. The first chapter ('Introduction') located the study in its scientific context and explained the methodology. The second chapter ('Loanword studies') aimed at determining a corpus of reliable Khotanese loanwords in Tocharian. Of 98 analysed items, I classified 48 words as reliable loanwords, 29 as doubtful/less reliable and I reject 19 possible correspondences. Chapter 3 ('Phonological and morphological analysis; determination of the chronology') analysed the corpus of 48 loanwords as determined in ch. 3. It established the main phonological correspondences that govern the adaptation of Khotanese loanwords in Tocharian, it determined an internal chronology (PTK, PK, OKh., LKh.), it analysed the morphological data of the Tocharian substantives and it listed them according to their part of speech and gender. Chapter 4 ('Semantic classification') determined the semantic areas of the loanword corpus and tried to draw some historical conclusions from the material. The current chapter ('Summary and conclusions') recapitulates the most important findings.

5.2. CONCLUSIONS

Most of the conclusions that have emerged from this study are of a linguistic nature. I briefly summarise these in the following, and I will also make an attempt to contextualise my findings chronologically and historically.

5.2.1. A NEW CORPUS OF KHOTANESE LOANWORDS IN TOCHARIAN

The most important conclusion concerns the volume and quantity of language exchange between Khotanese and Tocharian. The discovery of a previously unnoticed group of Khotanese loanwords, documented in this study, has shown that Khotanese exerted much stronger influence on Tocharian than previously imagined. Indeed, according to the scientific literature, the loanwords from Khotanese into Tocharian amounted to no more than 15 items, whereas the items that I classify as assured now total to 48 (cf. §2.2.1.). In many cases, the new interpretation of these Tocharian words on the basis of Khotanese has contributed to a better understanding of the history of the Tocharian words themselves and of the textual passages in which they are attested, which in some cases have received new interpretations (cf. e.g. the case of *pānto* or **wātano**, q.v.).

The newly discovered loanwords have allowed the formation of a new corpus. During this process, some old loanword proposals were rejected (see §2.2.3.). Another group of proposals, on the other hand, was not rejected, but either phonological or semantic

issues did not allow their inclusion into the group of ‘reliable’ loanwords (see §2.2.2.). Thus, the number of analysed Tocharian words amounts to ca. one hundred in total.

The newly formed corpus was subsequently analysed under different lenses. The most important conclusions in this respect are that 1. it is possible to classify the loanwords on chronological grounds and 2. Tocharian has preserved many loanwords from different prehistoric layers of Khotanese (tentatively termed Proto-Tumshuqese-Khotanese and Pre-Khotanese, see §3.2.). Therefore, this corpus is of the utmost importance for the reconstruction of the linguistic history of Khotanese and Tumshuqese, as so far no other language of the area has been shown to contain so many loanwords from historical and prehistorical Khotanese.

5.2.2. THE DIFFERENT LAYERS OF KHOTANESE LOANWORDS IN TOCHARIAN

The most important conclusion concerning the phonological and morphological analysis (§§3.3., 3.4.) is twofold. On the one hand, it has been established that loanwords from PTK, PK and OKh. mostly took the nom. sg. *-o* ending in Tocharian. This is an important distinguishing feature that, together with the correspondence TB */a/* ~ Khot. *a*, allows for the first time a clear distinction from Tocharian borrowings from so-called ‘Old Steppe Iranian’, the Old Iranian language that is the source of the characteristic borrowings with Tocharian *e* for Old Iranian **a*.

It is suggested that the Tocharian ending *-o* is an adaptation of the Khotanese acc. sg. *-u*. On the other hand, it seems that the most frequent Tocharian declension pattern for PTK and PK loanwords, i.e. the prehistoric loanwords, is nom. sg. *-o*, obl. sg. *-a* (the so-called ‘*kantwo*-type’ of Tocharian B nominal inflexion). Loanwords exhibiting this declension pattern are to be exclusively attributed to PTK or PK (see §3.4.). In the following, I provide a summary of the main features of the different layers of borrowings from Khotanese into Tocharian, with an attempt to contextualise these chronologically and historically.

5.2.2.1. Proto-Tumshuqese-Khotanese

Phonology	Possibility to reconstruct the word for Proto-Tocharian. TB <i>rt</i> ← PTK <i>*rd</i> (OKh. <i>ṛ</i>) TB <i>e</i> ← PTK <i>*ē</i>, <i>e</i> (OKh. <i>ī</i>), with <i>*ē</i> < Plr. <i>*ai</i> and <i>*e</i> < Plr. <i>*a_y</i> TB <i>-ñcw-</i> ← PTK <i>*-nśw-</i> (< Plr. <i>*-mćw-</i>) TB <i>/ər/</i> ← PTK <i>*ṛ</i> TB <i>ś</i> ← PTK <i>*č</i> (OKh. <tc> /ts/)
Morphology	The majority of the items shows nom. sg. <i>-o</i> , acc. sg. <i>-a</i> . Two items have nom. sg. <i>-o</i> , obl. sg. <i>-o</i> . No items with nom. sg. <i>-o</i> , obl. sg. <i>-ai</i> .
Semantics	Prevalence of lexemes associated with the administrative, political and economic sphere.
Dating	ca. 1000-500 BCE. The items that can be reconstructed for Proto-

	Tocharian were probably borrowed immediately before the break-up of Proto-Tocharian; the other items may have been borrowed immediately after this date. No precise date can be given for the break-up of Proto-Tocharian, but a date in the range of ca. 1000-500 BCE seems likely.
--	--

An important historical conclusion that may be drawn from the newly discovered material concerns the dating of the first contacts between Tocharian and the ancestor of Khotanese and Tumshuqese and, as a consequence, the dating of the first presence of PTK speakers in the Tarim basin. In fact, the discovery of a group of items that must have been borrowed around the Proto-Tocharian age speaks for the presence of PTK speakers in the Tarim basin long before historical Khotanese. Although this topic still needs thorough study, which exceeds the aims of this thesis, one should note that this was already partially suggested by Peyrot (2018: 275-7), who put forward the hypothesis that the arrival of the Tumshuqese-Khotanese people in the Tarim basin is possibly to be dated around the year 1000 BCE on archaeological grounds. Although more research is needed, the data gathered in this study tend to confirm this hypothesis.

The fact that the lexemes borrowed from PTK reveal a prevalence of items associated with the administrative, political and economic spheres suggests that the ancestors of the historical Khotanese and Tumshuqese people that came into contact with Tocharians were sedentary and possessed a solid hierarchical social structure. Moreover, they were probably engaged in commerce and traveled around in the region. If the hypothesis of the identification of the Ākētālā/Aqtala culture with Proto-Tumshuqese-Khotanese speakers is correct (Peyrot 2018: 275-7, Mallory 2015: 25),³⁸⁵ the oldest items in this group ('envoy', 'chief', 'property, estate', 'number', 'letter') may have been borrowed from Proto-Tumshuqese-Khotanese speaking people inhabiting the urban sites of the this culture in the first half of the first millennium BCE. Due to its position half way between the northern and the southern oases, a good candidate may be the site of Ĵumbulaq Qum, which is one of the most important sites belonging to the Ākētālā/Aqtala culture (Debaine-Francfort and Idriss 2001: 120-136, Peyrot 2018: 275). On the possible western ('Scythian') connections of this site cf. Debaine-Francfort and Idriss (2001: 156-8).

An important argument that speaks in favour of such an early dating of PTK – Tocharian contacts is the Tocharian word for iron, TB *eñcuwo* A *añcu**. In this study, it has been shown that this word was borrowed from PTK (cf. ch. 2. s.v.). Thus, it seems likely that PTK speakers introduced iron in the Tarim basin. Since the first iron finds in Xinjiāng date from the early 1st millennium BCE, it seems reasonable to posit a similar

³⁸⁵ The hypothesis is backed by the alleged western connection (Scythian or Saka) of the Ākētālā/Aqtala culture by contrast with the 'painted pottery' sites (Francfort 2001: 228-9).

date for the first contacts between PTK and Tocharian.³⁸⁶ As a consequence, it is possible that the first PTK speakers entered Xinjiāng around the same time period.

5.2.2.2. Pre-Khotanese (PK)

Phonology	TB <i>i</i> ← PK * <i>ī</i> (PTK * <i>ē</i> , OKh. <i>ī</i> , < PIr. * <i>ai</i>). PTK intervocalic <i>-k-</i> preserved as TB <i>-w-</i> . Loss of intervocalic <i>d</i> . TB <i>u</i> <i>w-</i> ← PK * <i>hw-</i> TA <i>ts-</i> ← PK * <i>ts-</i> (OKh. <i>tc-</i>)
Morphology	The majority of the items shows nom. sg. <i>-o</i> , obl. sg. <i>-a</i> . Two items have nom. sg. <i>-o</i> , obl. sg. <i>-ai</i> .
Semantics	Administrative, political and economic sphere and medical terms.
Dating	ca. 500 BCE – 400 CE. With the exception of TB <i>kātsō</i> (see ch. 2. s.v.), no items can be reconstructed for Proto-Tocharian.

There are unfortunately no elements that allow a precise dating for the beginning of the PK period. Since the *terminus ante quem* for the split of PT is probably 500 BCE, PTK cannot be later than this date. Therefore, it seems reasonable to posit this same date as a possible *terminus post quem* for PK. Thus, the PK period can be situated between 500 BCE and the age of the first Od Khotanese written attestations (5th c. CE). Obviously, it should be stressed that these two dates are to be taken respectively as a broad *terminus post quem* and *ante quem*.

An important phonological feature of this period is TB *i* ← PK **ī* (PTK **ē*, OKh. *ī*, < PIr. **ai*), which characterizes PK against PTK. Cases like TB *pito* and *kito** clearly show *i* < PIr. **ai* against PTK **ē* but cannot be classified as Old Khotanese because of the preserved intervocalic dental TB *-t-* ← PK **-ṭ-* (> OKh. *-h-*). Hence the need for another linguistic stage, distinct from PTK and OKh.

At this stage, words belonging to the administrative, political and economic spheres are as numerous as in borrowings from PTK, but more medical terms were borrowed. It is significant that, probably during the first centuries of the Common Era, the ethnonym of the Khotanese (OKh. *hvatana-*) was borrowed into Tocharian A and B (see s.v. *u**wātano**). The archaic appearance of this PK loanword suggests that Tocharian borrowed the term directly from Pre-Khotanese speakers, not from a later literary source.

5.2.2.3. Old Khotanese (OKh.)

Phonology	Absence of prehistoric features, but nom. sg. ending <i>-o</i> .
-----------	--

³⁸⁶ An in-depth discussion of these problems will be found in Peyrot, Dragoni and Bernard (Forthc.).

Morphology	Prevalence of items with nom. sg. <i>-o</i> , obl. sg. <i>-ai</i> .
Semantics	Mostly medical and Buddhist terms.
Dating	From the 5 th c. CE onwards.

The beginning of the Old Khotanese period coincides with the first Old Khotanese written attestations, dated to the 5th c. CE. It is significant that the oldest extant Khotanese manuscript has been found in Šorčuq, a northern town in which Tocharian A was spoken (Maggi 2004: 184).³⁸⁷ On the presence of a Khotanese speaking religious mission in Tocharian territory, see §4.3.4. Loanwords from Old Khotanese into Tocharian belong mostly rather to the medical and religious (Buddhist) sphere. This may suggest a different type of contact, i.e. mostly 'learned' and based on written texts.

Morphologically, a significant feature is the absence of words showing nom. sg. *-o*, obl. sg. *-a*. The most common pattern seems to be rather nom. sg. *-o*, obl. sg. *-ai*. For a discussion of this problem, see §3.4.3.2.

5.2.2.4. Late Khotanese (LKh.)

Phonology	Absence of prehistoric features and no nom. sg. ending <i>-o</i> .
Semantics	Mostly medical terms.
Dating	From the 6 th -7 th c. CE onwards.

It is difficult to determine chronologically a precise line of demarcation between Old and Late Khotanese, since these are still conventional definitions which do not consider diastratic and diatopic variation. As Old Khotanese was mainly a written religious language, it is possible that an early form of Late Khotanese was spoken during the same period, hence the very cautious dating to the 6th-7th centuries. Xuánzàng's observation that in the area of Khotan OKh. *hvatana*- 'Khotanese' was already pronounced as LKh. *hvaṇna*- (see s.v. „*wātano**" and Emmerick 1987: 42) in the 7th c. CE may back this tentative dating. In §4.3.4.3. I have shown that two manuscript fragments written in Late Khotanese were found in the Kuča area. I have put forward the hypothesis that these findings may be connected with the age of the Four Garrison (7th-8th c. CE), when Kuča, Qarašahr, Khotan and Kašgar were all united under Chinese rule. Thus, the movements of troops may have also favoured the exchange of knowledge between the North and the South of the Tarim basin.

Loanwords from Late Khotanese are way less numerous than those from PTK, PK and OKh. Therefore, it should be stressed that the limited corpus does not allow precise conclusions for the moment. Nevertheless, it can be observed that this group of loanwords does not show the nom. sg. ending *-o* characteristic of the older stages. This may be due to the typical Late Khotanese weakening and loss of final vowels (see §3.4.1.).

³⁸⁷ Since Khotanese loanwords are also found in archaic Tocharian B (cf. e.g. *yolo*), it is not possible to conclude that the contact in the Old Khotanese stage took place only through the Šorčuq area.

The loanwords from Late Khotanese mostly belong to the medical sphere. It is probably significant that a line of one of the two Late Khotanese manuscripts found in the Kuča area (cf. *supra*) may contain fragments of a medical recipe (cf. §4.3.4.3.).

5.2.3. WHAT TYPE OF LINGUISTIC CONTACT TOOK PLACE BETWEEN TOCHARIAN AND KHOTANESE AND TUMSHUQESE?

Before this study, the lexical items borrowed from Khotanese and Tumshuqese amounted according to the scientific literature to no more than 15 lexemes and the majority of them were technical terms. As suggested in §1.4., this could fit a ‘casual’ contact situation, the first category in the borrowing scale elaborated by Thomason and Kaufman (1988: 74-6).³⁸⁸ However, from the analysis of the data gathered in this study, it is clear that the linguistic contact between Tocharian and Khotanese and Tumshuqese should rather be characterized as the initial stage of ‘slightly more intense’ contact, i.e. the second category in Thomason and Kaufman’s (l.c.) borrowing scale.

The fact that the Khotanese and Tumshuqese influence on Tocharian was more intense than previously suspected is shown by different indicators. First, it seems that the direction of borrowing was almost exclusively from Khotanese and Tumshuqese (and their ancestors) into Tocharian. In fact, of the three Khotanese loanwords into Tocharian listed by Tremblay (2005: 44), only OKh. *puka-* ‘cubit’ (← TB *poko** ‘arm’) can be considered certain. OKh. *yaula-* ‘falsehood’ has been explained otherwise (see ch. 2 s.v. *yaula-*) and the Old Khotanese hapax *solāta-*, denoting some kind of animal, is still of uncertain interpretation. In this study (cf. ch. 2 under the treatment of the suffixes *-kke*, *-kka*, *-kko*) I put forward the proposal that the personal name *mukauka-* may also be a Tocharian borrowing. Therefore, the reliable Tocharian borrowings into Khotanese are only 2, against the 48 items of Khotanese and Tumshuqese origin found in Tocharian.

The second indicator concerns the semantics. Even though the majority of the borrowings are content words, there are also traces of function words (see e.g. TB *twār* ← LKh. *tvarä* ‘moreover’) and possibly some suffixes (cf. ch. 2 s.v. *-kke*, *-kka*, *-kko*).³⁸⁹ Moreover, the presence of five verbs among the borrowings (§3.5.5.) is another indicator of more intense language contact, since, at least in synthetic languages, verbs are much more difficult to borrow than nouns (Tadmor 2009: 61-3).

The nature of the examined material clearly suggests that the contact situation can be best described in terms of *adoption* rather than *imposition* (see §1.6.). In fact, no Khotanese or Tumshuqese influence has been detected in the phonology or the syntax of Tocharian, the two areas most affected in an imposition situation (Haspelmath 2009: 50).

Another important conclusion of this study concerns the periodisation of the linguistic contacts between Tocharian and Khotanese and Tumshuqese. Almost twenty

³⁸⁸ See also Thomason (2001: 70-1, 2010: 41).

³⁸⁹ On the borrowability of content words vs. function words cf. Tadmor (2009: 59-60).

years ago, Tremblay (2005: 444) claimed that ‘the language with the most durable influence [on Tocharian] is undoubtedly Khotanese (and its kins), a fact which indicates that Tocharian and Khotanese were already neighbouring in c. 500 BC.’ However, as shown in §1.4., this claim cannot be supported by Tremblay’s data. Nonetheless, the new loanword corpus determined and analysed in this study fully justifies this conclusion. In fact, the new material clearly shows that the majority of the lexemes were borrowed in prehistoric times, mostly from Proto-Tumshuqese-Khotanese and Pre-Khotanese.

The semantic areas affected by prehistoric borrowing concern mostly the administrative, political and economic spheres as well as medicine. This might point to the fact that, in the pre-Buddhist Tarim basin, the ancestors of Khotanese and Tumshuqese were culturally dominant in these domains. In the Old and Late Khotanese stages, Buddhist religious terms and again medical terms were prevalent among the borrowed lexemes. This suggests that Khotanese was an important intermediary in the dissemination of Buddhist knowledge into the Tarim basin (see §4.3.). In this respect, an intriguing result of this study that still awaits a more extensive investigation is the continuity of contact in the medical domain before and after the introduction of ayurvedic knowledge into the Tarim basin (§4.3.1.).