



Federico Dragoni (2023), *Watañi lāntaṃ. Khotanese and Tumshuqese Loanwords in Tocharian*, Dr. Ludwig Reichert Verlag, Wiesbaden (Beiträge zur Iranistik, Band 50), ISBN 978-3-7520-0690-2, pp. 332.

Federico Dragoni's *Watañi lāntaṃ* seeks to document, date and investigate the linguistic contact between the Tocharian languages on the one hand and Khotanese and Tumshuqese on the other. Tocharian A and Tocharian B are the conventional names of two closely-related Indo-European languages once spoken in the northern part of the Tarim Basin (present-day Xinjiang, Northwest China). The oldest linguistic neighbours of the Tocharians in the region were the speakers of the easternmost branch of the Iranian languages, once erroneously considered to be the language of the Sakas (Peyrot 2018). This branch is witnessed during the Middle Iranian period by two closely-related languages, Khotanese and Tumshuqese, named after the cities of Khotan and Tumshuq and spoken respectively in the southwestern and northwestern part of the Tarim Basin.

This dense and rich book is a revised version of the PhD dissertation with the same title discussed by the Author in 2022. It is the result of a four-year research endeavour at the Leiden University Centre for Linguistics (LUCL) as part of the project 'Tracking the Tocharians from Europe to China: a linguistic reconstruction', led by M. Peyrot. The overarching objective of this project was to reconstruct the hypothesised migration route of the Tocharians from the Proto-Indo-European homeland in Eastern Europe to the Tarim Basin by tracing the linguistic influence exerted by the various peoples encountered along the way. In this framework, the Author's subproject focussed on the contact between the Tocharians and the Khotanese and Tumshuqese peoples after the Tocharians had reached their historic site in the Ta-

rim Basin. Accordingly, the main goal of the book is to compile a comprehensive corpus of reliable Khotanese and Tumshuqese loanwords in Tocharian A and B.

The core of the book is represented by Chapter 2, which comprises ninety-nine loanword studies and proposes to identify forty-eight cases of borrowing. It is preceded by a brief introduction (Chapter 1), in which the Author presents the aims of his work, provides a brief description of Tocharian, Khotanese and Tumshuqese and a concise overview on the literature concerning their contacts, and finally lays down his criteria to identify a reliable loanword. Chapter 2 is divided in two parts. The first part (§2.1) is organised as a dictionary of borrowed lexical items (more on the structure below). The second part (§2.2) consists of a reference list of reliable (48 items), doubtful (29 items) and rejected loanwords (18 items), plus three more words for which a different Iranian origin is acknowledged. Chapter 3 and 4 are concerned with two main questions: how old is the contact between the Khotanese and Tumshuqese peoples and the Tocharians and what areas of the lexicon were subject to borrowing? In order to address these issues, the Author builds on the new collection of reliable loanwords and attempts to classify them from a chronological and semantic point of view. In Chapter 3, the Author divides the loanwords into four chronological layers, corresponding to different linguistic stages:

- from Proto-Tumshuqese-Khotanese (PTK) into Proto-Tocharian (PT), approximately 1000-500 BCE;
- from Pre-Khotanese (PK) into Pre-Tocharian A and/or Pre-Tocharian B, appr. 500 BCE-400 CE;
- from Old Khotanese (OKh.) into Tocharian A (TA) and/or Tocharian B (TB), from the 5th c. CE onwards;
- from Late Khotanese (LKh.) into TA and/or TB, from the 6th to 7th c. CE onwards.

The first two stages belong to the pre-documentary phase and are distinguished on the assumption that in the earliest stage (PTK → PT) loanwords entered the common ancestor of TA and TB and were inherited by both languages, whereas the second stage (PK → pre-TA, pre-TB) begins after the split of TA and TB. According to the semantic

analysis conducted in Chapter 4, the loanwords entered during the prehistoric phases mostly pertain to the administrative, political and economic spheres, whereas loanwords from Old and Late Khotanese are predominantly medical and Buddhist terms. This qualifies the contact during the historic phase (from 5th c. CE onwards) as mostly taking place through written sources in a learned milieu, as opposed to earlier contact being between travellers and merchants. Chapter 5 further explores the interactions between the two groups by switching sides and discussing a few possible cases of Tocharian loanwords in Khotanese and Tumshuqese, confirming that the dominant direction of borrowing was from Khotanese to Tocharian. As for the fragmentarily attested Tumshuqese, the Author notes that the influence was rather reverse: while he could not find any reliable Tumshuqese loanword in Tocharian (cf. p. 203; p. 39 fn. 50), he acknowledges one Tocharian loanword in Tumshuqese, along with a few words of Tocharian origin that are not integrated into the language and should thus be regarded as foreign words. Finally, Chapter 6 summarises the results and draws conclusions from the book's findings. The book is closed by a set of useful indexes that make it readily consultable for specific queries. Beside the traditional indexes of quoted words and passages (in many ancient and modern languages), this part also includes a list of relevant paragraphs for the discussion of each reliable loanword.

This work immediately qualifies as a fundamental starting point for future investigations on many related subjects of Central Asian studies, from the historical phonology of Khotanese and Tocharian to the broader question of the sociolinguistic dynamics in prehistoric times within the Tarim Basin.

Among the most evident accomplishments of this book is the expansion of the corpus of Khotanese loanwords from fifteen to forty-eight items, making Tocharian the first language for number of lexical items borrowed from Khotanese. Secondly, the identification of an ancient layer of loanwords predating both the split of Tocharian A and B and that of Khotanese and Tumshuqese offers an important argument in favour of an early dating of the first encounter between the Tocharians and the Tumshuqese/Khotanese peoples. This is of

course a crucial point for the ‘Tracking the Tocharians’ project. While none of the languages involved is attested earlier than the 5th century CE, ancient contacts between Tocharians and Iranians have been assumed since K. T. Schmidt (1985) identified an early layer of loanwords from an Old Iranian language into Tocharian. Parallel research carried out within the same project by Ch. Bernard (2023) assigned part of these loanwords to an otherwise unattested Iranian language belonging to the Old Iranian stage, perhaps spoken by inhabitants of the Eastern Eurasian steppe that came into contact with the Tocharians before they entered the Tarim Basin (hence the proposed name ‘Old Steppe Iranian’). The Author’s hypothesis is that another source for loanwords of Old Iranian provenance was the common ancestor of Tumshuqese and Khotanese (Proto-Tumshuqese-Khotanese). The presence of loanwords at this early age would provide a *terminus ante quem* for the contact between Proto-Tumshuqese-Khotanese and Pre-Tocharian speakers: as the split of the two Tocharian languages is dated approximately between 1000 BCE and 500 BCE (Peyrot 2015), the first encounter between the Tocharian and Tumshuqese/Khotanese peoples would have taken place during the first half of the 1st millennium BCE. This would also confirm the presence of Iranians in the Tarim Basin around the year 1000 BCE, as suggested by Peyrot (2018: 275) on archaeological grounds.

The extent to which the book succeeds in demonstrating the thesis of prehistoric contact between Proto-Tumshuqese-Khotanese and Pre-Tocharian speakers deserves closer examination, as the evidence collected is interesting but needs to be treated with some caution. However, the Author’s research into the non-inherited component of the Tocharian lexicon yielded several valuable results beyond the scope of the project from which it stems. As a byproduct of its primary aims, this book offers meticulous analyses of more than a hundred Tocharian words of uncertain etymology and provides new insightful translations to a wealth of obscure passages from both Tocharian and Khotanese texts, often proposing compelling new etymologies for difficult words. Another result of the outmost importance is the contribution to future investigations into the historical phonology of Khotanese, for which this book will serve as a foundational preliminary.

It may be useful to have the philological and linguistic contributions contained in this book all gathered in one place. The Author proposes new readings or translations of the following Tocharian texts:

- THT 108 b9: reading of TB *suwāññe* „*wātatane* as *suwāññe*-*wātanane* ‘(in the tune) of golden Khotan’. TB *suwāññe* is interpreted as an adjectival formation based on Middle Indic *suwaṇṇa*-‘golden’, whereas TB „*wātano*\* is interpreted as the Tocharian rendition of the name of Khotan. (This and the following four entries first appeared in Dragoni 2023);
- A 24 b5, A 163 b2: interpretation of TA *watañi-lāntaṃ* (loc.sg.) as ‘(in the tune) of the king of Khotan’, with TA *watañi* being an *-i* adjective formed on TA *wataṃ*\* (from TB „*wātano*\* ‘Khotan’, cf. previous entry);
- A 71 b3, A 260 b2, THT 1464 b2: interpretation of TA *watañinaṃ* as ‘(in the tune) of Khotan’, loc.sg. of TA *watañi* ‘of Khotan’ (cf. previous entry);
- THT 2688.10: interpretation of TB „*watanams* as ‘of the Khotanese people’ (gen.pl.m.);
- THT 2709.2-3: interpretation of TB „*wātane* as ‘the Khotanese people’ (nom.sg.m.);
- PK AS 8B a2: identification of *keto* as nom.sg. of *keta* ‘estate’;
- PK AS 7H b3-4, THT 31 a2-3; THT 33 b6-7; THT 42 b5: new translation of TB *koto*\* as ‘excrement’ instead of ‘hole, pit’;
- THT 325 a1: new translation of the perlativ TB *kontsaisa* as ‘by mistake’ and tentative interpretation of TB *āmāpi* as ‘by intention’ (but the form remains unclear);
- THT 577 a1: new translation of TB *kompō*\* as ‘plain’ or ‘*yojana* (measure)’;
- THT 91 b6, IOL Toch 5 b2-3: new translation of TB *twār* as ‘moreover’ instead of ‘consequently’;
- THT 554 a6: TB new translation of *paño*\* as ‘binding’;
- A 14 a6-b1, PK AS 8C a3-4, PK AS 9B b5, PK AS 9D b3, THT 29 a8, THT 229 b4, THT 281 b5, THT 364 a5, THT 385 b4, THT 1252 b2, THT 2377.v a2, THT 108 a6-7, PK AS 4B a5 (parallel M 500.1 b4-5), IOL Toch 369 a2, THT 274 b4: new translation of TAB *pānto* as ‘friend’;

- THT 282 b1: new translation of TB *wiñcaññe* as ‘pertaining to a sparrow’;
- Km-034-ZS-L-01 a6, PK AS 17A b1-2, PK NS 399 a3, THT 588 a2, IOL Toch 116 a1, THT 382 a1, THT 1104 a4: new translation of TB *śarko\** as ‘singing, song’;
- YQ I.9 a2, YQ I.9 b3, A 318 a2, A 318 a6, A 126 a6, A 15 b1: new translation of TA *tsärk* as ‘singing, song’ instead of ‘lute’;
- PK AS 8A b7-8: new translation of TB *sumo* as ‘broth’.

New translations of Khotanese texts:

- P 2782.26; P 2781.71: interpretation of LKh. *auska*, *auski* as variant spellings of the adverb *uska* ‘up’ instead of a feminine noun meaning ‘dwelling place’ (as already suggested for IOL Khot S. 6.9 and P 2786.70);
- JS 37r3; P 4099.19-21: new translation of LKh. *kīḍa-* as ‘stake’, interpreted as an Indic loanword, cf. Skt. *kīla-/khīla-* ‘id.’;
- IOL Khot 102/1v2 (*Jivakapustaka*), Fol. 91v2, 91v3: new translation of LKh. *graysāñ-* as ‘to make lament’ instead of ‘to cause to go wild, to cause miscarriage’;
- Hedin 3.15: new translation of LKh. *peri* as loc.sg. of Khot. *pāra-* ‘debt’;
- JS 28r1: new reading of Khot. *samuvā* as *samū* ‘only’ + particle *vā*;
- P 2739.19: tentative interpretation of the sequence *śau rraha: śīyi ttrihe:* as ‘one (portion) of red radish and white radish’;
- Z 4.114 (OKh.), Sudh 406, 414-15, Or.11252/4v (LKh.): new translation of OKh. *hambälke*, LKh. *habvākya-* as ‘without permission’.

New Khotanese etymologies:

- LKh. *aīsta* (*bā*) ‘Boerhavia diffusa’ (attested in various spellings) < OKh. *\*älsta*, an inflected form of a reconstructed *i*-stem noun *\*älsti-*, *h*-less variant of OKh. *hālsti-* ‘spear’ < PIr. *\*Hṛṣti-* (already published in Dragoni 2021). The Author compares the Latin term *lanceolātus* ‘spear-shaped’ to refer to the form of some leaves, but observes that the term may refer to the roots rather than the leaves of the *Boerhavia diffusa*;

- LKh. *maśpa-* ‘road’ (also attested as *maśapa-*, *māśpa-*, *maśma-*, *magpa-*), perhaps connected to a verbal root *\*maś-* < PIr. *\*maj-(a)ya-* ‘to break’;
- OKh. *haḍa-* ‘envoy’ < PIr. *\*Harta-*, pp. of *\*Har<sup>1</sup>* ‘to go to(wards)’ or *\*Har<sup>2</sup>* ‘to set in motion’ (but see the discussion below);
- Khot. *karāśśā-* ‘creeper’, loanword from an Indo-Aryan language, cf. Skt. *karś-* ‘to be lean, thin’ (but cf. Del Tomba (2023: 212 fn. 278) for the reconstruction of an *i*-stem *karāśśi-* rather than an *ā*-stem, and consequently an earlier dating for the borrowing of the word in Tocharian);
- LKh. *gr(r)aysa-* ‘grief, torment’: nominal derivative of a verb *grays\** (cf. the causative LKh. *grays-āñ-*) < PIr. *\*garj-* ‘to lament, weep’;
- LKh. *ttūmgarā-* ‘ginger’ < PIr. *\*tvāna-kara-* ‘strong-maker’, the first member of the compound being the present middle participle of the root *tv-* ‘to be strong’ (already published in Dragoni 2021);
- OKh. *pārgyiñā-* ‘garden’: *-iñā-* derivative from *pārgya\** ‘having leaves’ (cf. OKh. *°vārgya-* ‘id.’) < PIr. *\*parnakika-* (suggested by Mauro Maggi);
- OKh. *bilsamga-*, *bālsamg(h)a-* < OKh. *\*balysa-samga-* ‘Buddha-samgha’, used to translate Skt. *bhikṣusamgha-*;
- Khot. *tcāmph\** ‘to be violent, destroy’ < PIr. *\*cāmf-*, palatal variant of the root *\*čap/f-* ‘to seize, attach, stick, strike’ with nasal infix as in Balochi *čāmp-* ‘to snatch’ and Yaghnobi *čūmf-* ‘to push (to)’;
- Khot. *tcarkā-* ‘play, delight, amusement’: nominal *ka*-derivative of a verb meaning ‘to sing’, from the root PIr. *\*čarH-*, variant of PIr. *\*karH-* ‘to praise, celebrate’, cf. Sariqoli *čīr-* ‘to sing, twitter, chirp’;
- LKh. *šimjā-* name of various trees, originally a Central Asian Wanderwort, whose pre-form may be reconstructed as *\*sinjata-* > OKh. *\*śsinjsata-* > *\*śsinjsaa-* or *\*śsinjsā-*. The Late Khotanese form may be derived through haplology from LKh. *\*šimjīmjā-*, the denominal adjective formed with the fem. suffix *-iṃgyā-*;
- Khot. *jsīr-* ‘to deceive’ < PIr. *\*jar-ya-* from a palatal variant of the root *\*garH-* ‘to greet, call’, although semantically the proposal is not very compelling.

New reconstructed Old Steppe Iranian word:

- OSIr. *\*dzara-* < PIr. *\*jarH-* ‘to hurt, wound, anger (with words)’.
- Loanwords in Tocharian from other Iranian languages:
- TB *armañik* a kind of textile ← Sogdian *rm’nykh* ‘id.’.

Moreover, it is proposed that the name of Khotan was rendered in Bactrian as *\*χο/ωδανο* or *\*χο/ωτανο* or *\*χοαδ/τανο*, as suggested by the Niya Gāndhārī form *khotana-*, which appears to be borrowed from Bactrian. The Author argues that the name OKh. *hvatana-* was borrowed again in later times as Bactr. *οατανο*, attested in the Bactrian documents as a personal name.

As noted above, the core of the book is represented by the dictionary of loanwords in §2.1, on which I shall now focus my attention discussing some points of detail. Each entry in §2.1 contains: 1) a list of the Tocharian and Khotanese/Tumshuqese occurrences; 2) a discussion of the forms and their possible connection; 3) a summary of the results. In the Discussion sections, the Author first addresses any potential philological problem connected to the lexemes under analysis to determine their precise phonological shape and/or meaning, then proceeds to evaluate whether the lexical pair can be considered a case of borrowing. This operation often requires the setting up of proto-forms obtained through internal reconstruction based on what we know of both Tocharian and Khotanese historical phonology. In this way, the Author seeks to identify patterns of phonological adaptation and establish fixed correspondences between Khotanese and Tocharian sounds. This sometimes leads to the Author formulating rules of adaptation based on very small evidence and apply them rather rigidly. While this is sometimes unavoidable when dealing with ancient, fragmentarily attested languages, it is risky to draw too far-reaching conclusions from such hypotheses, however plausible they may be. In the following, I will discuss some entries where, in my opinion, this risk is more concrete.

(In quoting the entries, I follow the numbering system adopted by the Author).

- (5) TB *eñcuwo* TA *añcu\** ‘iron’ ← PTK *\*hens’wanya-* ‘id.’ (cf. OKh. *hīśšana-* ‘id.’).

The connection between the Khotanese and Tocharian words for ‘iron’ was first suggested by the Author in a cowritten paper (Peyrot *et al.* 2022), whose results are only briefly summarised in the book<sup>1</sup>. In the article, it is argued that the Tocharian forms are based on a proto-form *\*henswanya-*, which is in turn traced back to PIr. *\*ham-čwanya-*. In the book (§3.2.1.1), the Author dates this borrowing to the earliest chronological layer (PTK → PT), among other things because of the correspondence PTK *\*e* (OKh. *ī*) ~ TB *e* TA *a*<sup>2</sup>. This implies the preverb PTK *\*hen-* (< PIr. *\*ham*) was adapted as PT *\*en* and then underwent regular changes in TB and TA. However, the same preverb *\*ham-* appears to be lost in another loanword of the same stage: (16) TB *keś* TA *kaś*, which is connected to OKh. *hamkhīs-* ‘to count’ (p. 102). To explain this inconsistency, the Author establishes the following rule: the preverb *\*ham-* was preserved in Tocharian only if it was stressed, otherwise it was dropped. External support for this hypothesis comes from the reconstructed position of the accent in PTK *\*henswanya-* < PIr. *\*ham-čwanya-*. It is argued that the vowel of the prefix was stressed because it could be affected by the *i*-umlaut triggered by the *-y-* in the last syllable (p. 62, point b). Indeed, it is generally accepted that umlaut phenomena in Iranian are confined to stressed syllables<sup>3</sup>.

<sup>1</sup> The borrowing of such an important cultural word is of course relevant to date the first encounter between the Tocharians and the Proto-Tumshuqese-Khotanese, as the first iron finds in the Tarim Basin are dated around the early 1st millennium BCE. In §6.2.2.1, the Author surmises that the Proto-Tumshuqese-Khotanese speakers may be responsible for the introduction of iron in the Tarim Basin and that Pre-Tocharian speakers may have borrowed the term from them together with the new technology. This reasoning corroborates the proposed dating of the first encounter between the two peoples to the first half of the 1st millennium BCE.

<sup>2</sup> Note that the Author reconstructs a short PTK *e* instead of a long *ē* for the Proto-Tumshuqese-Khotanese outcome of umlaut phenomena (cf. p. 209). In this way, he distinguishes it from the outcome of the monophthongisation of PIr. *\*ai* > PTK *ē*. However, both PIr. *\*ai* and PIr. *\*a\_y* resulted in a long *ī* in Old Khotanese. Thus, it seems preferable to posit a long *\*ē* for the Proto-Tumshuqese-Khotanese phase, consistently with its Old Khotanese outcome. Indeed, MAGGI (2024: 171, fn. 23) recently described the *i*-umlaut as a palatalisation originating from «a palatal epenthesis of stressed *\*a* > *\*ai* > *\*ē* > *ī* due to a following *i/ī*». This implies that, historically, the vowels affected by *i*-umlaut became indistinguishable from the monophthongised outcomes of the PIr. *\*ai* diphthongs and both were treated in the same way, ultimately resulting in OKh. *ī*.

<sup>3</sup> Although the topic still needs further investigation, cf. KÜMMEL (2014).

However, this outcome does not align with what is now known about palatalisation and umlaut phenomena in Khotanese. In his account of synchronic umlaut, D. Hitch (1990: 187) identified Khot. *n* /n/ as one of the «umlautable» consonants, i.e. consonants that are palatalised when followed by a palatalisation-triggering suffix (e.g. the loc.sg. ending *-iā*), distinguishing them from the consonants that «absorb» the umlaut and those that are «neutral» to it. Only in the latter case palatalisation is observable on the stressed vowel (e.g. *pajsa-ma* ‘honour’ > loc.sg. *pajsi-ma*), whereas, with palatalisable consonants («umlautable» in Hitch’s use), the palatalising effect is visible on the adjacent consonant but does not affect the stressed vowel (e.g. *uryā-na* ‘garden’ > loc.sg. *uryāñā*). A. Del Tomba (2024) recently showed that the cases of diachronic palatalisation, i.e. caused by suffixes that are no longer productive (e.g. *\*-ya-*), follow the same pattern as that described by Hitch for the synchronic palatalisation (e.g. *kīra* ‘action’ < PIr. *\*kar-ya-*: «neutral» consonant /r/ allowing palatalisation of the stressed vowel vs. *°hvāñā* ‘speech’ < PIr. *\*hwān-ya-*: «umlautable» consonant /n/ being palatalised instead of the stressed vowel). Therefore, returning to PIr. *\*ham-ćwanya-*, we would have expected the final *-ya-* to cause palatalisation of the preceding *-n-* (PTK *\*ham-śwañā-*) rather than *i*-umlaut of the prefix *\*ham-* (PTK *\*henśwanya-*). The development PIr. *\*ham-* > PTK *\*hen-* is described by the Author as a case of trajected umlaut, i.e. an umlaut crossing the second syllable. In the original article (Peyrot *et al.* 2022: 407), two parallels were mentioned for this development, following H. W. Bailey (1979: 352b, 444b, 481b): Khot. *ysīrra-* ‘gold’ < PIr. *\*jaranya-* and LKh. *hījsara-*, loc.sg. of *ha(ṃ)jsarā-* ‘scope, range, place’ < PIr. *\*ham-čarā-*. In the loc.sg. form LKh. *hījsara-*, the palatalisation of the prefix *\*ham-* is allowed by the “neutral” consonant /r/ as the final consonant of the lexical morpheme, a condition that is not satisfied by PIr. *\*ham-ćwanya-*. As for Khot. *ysīrra-*, the *-ī-* is not an instance of trajected umlaut but rather a regular outcome of the process described above: the palatalisation must have occurred after the loss of the intermediate vowel of PIr. *\*jaranya-* which led to the development PIr. *\*rn* > OKh. *rr*, otherwise the expected form would be *\*\*ysarñā* (*-rr-* is another case of «neutral»

consonant, cf. Hitch 1990: 189)<sup>4</sup>. All in all, it appears that the long *-ī-* in OKh. *hīśśana-* can hardly be due to a trajected umlaut, as we would expect the palatalisation to affect the consonant of the stem, and thus it cannot be used to determine the position of the stress on the prefix.

This remark does not invalidate the proposal of a connection between OKh. *hīśśana-* and TB *eñcuwo* TA *añcu\**, which is based on various, rather sophisticated arguments. The Khotanese form is arguably still the best match for the Tocharian forms. However, the reconstruction of the stress on the first syllable of the PTK proto-form can be seriously called into doubt. Therefore, it does not seem advisable to formulate an adaptation rule based solely on this evidence.

The argument just presented against the role of the stress in the adaptation of PIr. *\*ham-* affects two items in Chapter 2. It only slightly affects the derivation of TB *ampoño* ‘rottenness, infection’, for which the Author presents two possible sources: LKh. *\*hambvauña-* ‘infection’ or a Tocharian formation based on the verb TB *ampa-* ‘to rot, decay’, which is in turn derived from LKh. *hambva-* ‘fester’ (pp. 35-36). After establishing the stress rule, he chooses the second option on account of the fact that TB *ampoño* is not stressed on the first syllable as one should expect if it were a loanword from LKh. *\*hambvauña-*. Setting the stress rule aside, both options can be considered as plausible.

The other item deserves more extensive comments:

- (16) TB *keś* TA *kaś* ‘number’ ← PTK pres. inf. *\*ham-xēzi* ‘to count’ (cf. OKh. *hamkhīs-* ‘id.’).

The Author proposes to identify the Tocharian word for ‘number’ as a Proto-Tumshuqese-Khotanese loanword and sets up as a source PTK *\*ham-xēzi* (with stress on the root according to the alleged stress rule), a reconstructed present infinitive form of the verb *hamkhīs-* ‘to count’ (OKh. *hamkhīsä\** < PIr. *\*ham-xaijyai*). The preservation of the initial syllable *\*ham-* in three cases (TB *ampa-*, TB *ampoño* and

<sup>4</sup> As a side note, this example qualifies the palatalisation in Khotanese as a relatively recent process (surely subsequent to the development *\*rn > rr*), rather than one being active already in the Proto-Tumshuqese-Khotanese period.

TB *eñcuwo* TA *añcu\**), when separated from the stress rule discussed above, speaks against a connection between TB *keś* TA *kaś* and a PTK form with *\*ham*. Another issue is the absence of the final vowel in the Tocharian forms. This is the only case of a noun borrowed during the PTK → PT phase that does not show any trace of the Khotanese ending (cf. p. 219)<sup>5</sup>. The Author speculates that the PTK final *-i* of the infinite ending «could have been borrowed into Tocharian as *-ə* after a palatal, cf. the ending TB */-cə/, /-ścə/, /-ñə/* etc. Tocharian *i* was not suitable because it was probably felt as long (< *\*-əy*)». It is assumed that final *-ə* of the PT form *\*keśə* was subsequently dropped in TA and TB. If this is an adaptation pattern attested in loanwords from other languages, one feels the need to be directed to the relevant literature on the subject. Should it be considered a hypothetical statement, it is perhaps more cautious to qualify this connection as doubtful.

Additionally, the reconstruction of the form PTK *\*ham-xēzi* < PIr. *\*ham-xaijyai* with short *\*-i* for the pres. inf. ending entails some problems. Why is the monophthongisation of the PIr. diphthong *\*-ai-* still at the stage *\*-ē-* in internal position but has already progressed to *\*-ī* > *\*-i* in final position? Considering that the Proto-Tumshuqese-Khotanese stage is described as an Old Iranian language spoken around 1000-500 BCE, one does not expect the infinitive form to be so close to the phonological shape attested in Old Khotanese, i.e. *-ā, -i* ([ə], [i]) (Emmerick 1968: 218). Perhaps a later dating for the stage *-i* would be more appropriate (cf. the next entry for a similar argument). The Author's main reason for assigning this loanword to such an early phase is the correspondence TB *e* ~ TA *a*, which points to PT *\*e* and thus to a PTK source with *\*ē* < PIr. *\*-ai-* (the vowel of the root in *\*ham-xēzi*). However, since this leads to a chronological inconsistency with the reconstructed infinitive ending *\*-i*, it is perhaps best to regard the origin of TB *keś* TA *kaś* as still unclear.

(30) TB *parso* TA *pārs* 'letter' ← PTK acc. sg. *\*pṛsu* 'request' (cf. OKh. *puls-* 'to ask').

<sup>5</sup> Apart from (35) TB *wanc\** 'sparrow', which is however not directly attested but reconstructed from the denominative adjective TB *wiñcaññe* based on the treatment of the final vowel in TB *keś* TA *kaś*.

The Author elaborates on a proposal by L. Isebaert (1980: 104) to connect the Tocharian word for ‘letter’ to the Khotanese verbal root *puls-* ‘to ask’. The source for the Tocharian forms is identified in PTK *\*p̄rsu*, a reconstructed noun meaning ‘request’ that may have undergone a semantic change to ‘letter’ within Tocharian. The final *\*-u* in the reconstructed Khotanese form is required to explain the fact that the word is integrated in Tocharian B as a noun with the inflectional pattern nom.sg. *-o*, obl.sg. *-o*. One of the book’s discoveries is that nouns borrowed during the Proto-Tumshuqese-Khotanese, Pre-Khotanese and Old Khotanese stages were predominantly adapted in Tocharian B as nouns with a nom.sg. *-o*. The Author argues that this is an adaptation of the Khotanese acc.sg.m. ending *-u* (§3.4.10, cf. further on this argument). For the matter at hand, since no noun *\*\*pulsā-* ‘question’ is attested in Khotanese, the Author assumes that the word entered Tocharian as an infinitive present with final *-ā*, which was interpreted as the nom. sg. ending *-ä*. This in turn led the Tocharian speakers to «set up an acc. sg. in *-u* that they borrowed as a substantive with nom. sg. *-o*» (p. 138). Such a process appears rather complicated for a borrowing, for a few reasons. First, it gives the impression that the Tocharian speakers deliberately inflected the noun before borrowing it, which appears unlikely: it is hardly plausible that, once it had entered Tocharian, the loanword would be inflected by the Tocharian speakers according to the Khotanese declension. Secondly, to be reanalysed as a nom.sg. *-ä* of the *a*-stems, the ending of the pres. inf. must have been homophonous to it ([ə]). This is of course applicable to Old Khotanese, where the pres. inf. ending is attested as *-ā*, *-i* ([ə], [i]) (cf. above). However, the borrowing is dated to the Proto-Tumshuqese-Khotanese stage, when, as discussed above, the pres. inf. ending (PIr. *\*-ai*) was in all likelihood still clearly differentiated from the nom.sg.m. ending of the *a*-stems (PIr. *\*-ab*), even though their exact phonological shape is hard to determine. As a side note, the assumption that the pres. inf. ending was already *-ə* in PTK conflicts with the reconstruction of PTK *\*ham-xézi* (on which see above).

Perhaps it would be more convenient to move the date of the borrowing to a later stage. The main reason for the dating to the Proto-Tumshuqese-Khotanese phase is that the TB form *parso* /párso/ presupposes a vocalic \*ʔ in the source, which fits nicely with an archaic phase of the language and is suggested by at least another loanword (TB *ṣart-* TA *ṣārttu-* (PT *\*ṣartu-*) ‘incite’ ← PTK past ptc. *\*ṣrtu* ‘id.’). However, the Author also admits the possibility of a source-form PTK *\*pursu*, with «early vocalisation» of \*ʔ during the Proto-Tumshuqese-Khotanese phase and internal PTK *\*-u-* borrowed as TB /ə/ (p. 138, fn. 276). There would be no need to postulate an «early vocalisation» if the borrowing happened during the Pre-Khotanese stage. It is also important to note that the Pre-Khotanese stage is reconstructed as a period spanning over one millennium (500 BCE - 400 CE), during which \*ʔ may have been first vocalised as /ər/, as witnessed by the Tocharian form, before it underwent assimilation to /ur/. As for the final vowel, it may be worth rehabilitating Isebaert’s idea of a substantive PIr. *\*p̄rsā-* ‘question’ (cf. Av. *frasā-* ‘id.’).

(3) TA *ārt\** ‘envoy’ ← PTK acc. sg. *\*(h)ardu* ‘id.’ (cf. OKh. *hada-* ‘id.’).

To establish this loanword, the Author suggests that the noun OKh. *hada-* ‘envoy’ should be connected to the ppp. of the PIr. verb *\*Har<sup>-1</sup>* ‘to go to(wards)’ or *\*Har<sup>-2</sup>* ‘to set in motion’, which he purposefully reconstructs as *\*Harta-* instead of *\*Hṛta-*, as one would expect. According to this reconstruction, PIr. *\*Harta-* then resulted in PTK *\*(h)arda-*, which provides a nice match for the TA word *ārt\** ‘envoy’<sup>6</sup>. The Author’s reasoning for rejecting a form PIr. *\*Hṛta-* is that the vocalic \*ʔ is expected to be still preserved in Proto-Tumshuqese-Khotanese (cf. above), whereas the TA form *ārt\** /’art/ clearly shows that the vocalisation had already taken place. To avoid admitting an early vocalization of \*ʔ in the Proto-Tumshuqese-Khotanese stage, he therefore posits PIr. *\*Harta-* as a starting form. Again, it is perhaps worth considering the hypothesis of a later borrowing. A Pre-Kho-

<sup>6</sup> The initial OKh. *h-* is traced back to a Proto-Iranian laryngeal, following a proposal by KÜMMEL (2018).

tanese form *\*harta-* as a source for TA *ārt\** would have two advantages<sup>7</sup>:

- the vocalisation of *\*r̥* would have happened during the Pre-Khotanese stage, as is expected (p. 39, point 3);
- since the borrowing would have happened after the split between Tocharian A and Tocharian B, there would be no need to reconstruct an inherited TB form *\*ārto* that was later replaced by TB *šīto* ‘envoy’ (p. 39, point 4). The word would have simply entered Pre-TA and regularly result in TA *\*ārt*.

As both a Proto-Tumshuqese-Khotanese and a Pre-Khotanese borrowing are theoretically feasible, I would suggest including this particular case among the loanwords that are generically from the pre-historic stage, i.e. among the items listed in §3.2.2.2.

A similar caution can be extended to entry (9) TB *kāmarto\** TA *kākmart* ‘chief’. The source form is reconstructed as the acc.sg. form PTK *\*kamardu* ‘head’ (cf. OKh. *kamala* ‘id.’), from PIr. *\*kamṛda-* with «early vocalisation of PIr. *\*r̥* > *\*ar*». Perhaps a smoother solution may be to consider *\*kamardu* as a Pre-Khotanese form, as this would make the *ad hoc* assumption of an «early vocalisation» of *\*r̥* unnecessary.

(12) TB *kito\** (*ekita*) ‘help’ ← PK acc.sg. *gīḍu-* ‘id.’ (cf. OKh. *ggīha-* ‘id.’).

The origin of the TB word *ekita*, which is only attested in the phrase *ekita yam-* ‘to help’, has been much debated, and scholars are still not sure as to which part of speech it should be assigned (either a noun or an adverb). The Author successfully identifies a compelling match for this form in the Khotanese word for ‘help’, i.e. OKh. *ggīha-*. This allows for the reconstruction of a PK acc.sg. *gīḍu-* ‘id.’ as a source for the Tocharian form. As for the etymology for OKh. *ggīha-*, the Author accepts Skjærvø’s proposal (2004: 260), who derived it from

<sup>7</sup> I posit a PK form *\*harta-* rejecting the idea that the group PIr. *\*-rt-* had already progressed to *\*-rd-* in Proto-Tumshuqese-Khotanese (cf. p. 39, point 2). The reconstruction of a PTK stage *\*(h)arda-* (which implies PK *\*harda-*) is problematic because the group *\*-rd-* < PIr. *\*-rt-*, *-ṛt-* would be confused with the outcome of the original PIr. *\*-rd-*. The two consonant clusters must have remained distinguished because they resulted in OKh. *-d-* and *-l-* respectively.

PIr. \**gaiḍa-*, the masculine counterpart to PIr. \**gaiḍā-* ‘livestock, cattle’ (cf. Av. *gaēḍā-* ‘living being, world’, OP *gaiḍā-* ‘livestock, flock’) with roughly the same meaning. Phonologically, this etymology works perfectly for OKh. *ggīha-*, but the semantic development from ‘livestock’ to ‘help’ is not as straightforward. The Author proposes the following semantic passages: from the Old Iranian meaning ‘cattle’ to ‘property, subsistence’ (cf. MMP *gyh* ‘property, flock, herds’), to the abstract ‘strength’, which in Khotanese took the sense of ‘military force’, and ultimately that of ‘support, help’. The shift to ‘military force’ would be witnessed by a passage of the Old Khotanese *Suvarṇabhāsottamasūtra* (Suv 1.15), in which *ggīha-* translates Skt. *bala-* ‘strength, military force, army’. However, from the context of the passage it is clear that Skt. *bala-* could be interpreted as ‘auxiliary troops’ (Skjærvø 2004: 98). Thus, rather than attesting the retention of the original Middle Iranian of ‘flock, group’ in Khotanese, it is likely that the word *ggīha-* was simply chosen by the Khotanese translator to match the meaning conveyed by the Sanskrit word in that specific context. Therefore, this parallel only confirms the meaning ‘help’ for the Khotanese form.

An alternative etymology for the Khotanese form may be PIr. \**haka(t)-yaḍ-ya-* ‘help’, a deverbative formation from the verb \**haka(t)-yaḍ-ya-* > OKh. *ggīh-* ‘to help’ (ppp. *gīsta-* < \**haka(t)-yasta-* with extension of the full grade of the present stem; for the root \**yat-* cf. MP *frayād-* ‘to help, assist’, Sogd. *βr’yt-* ‘id.’)<sup>8</sup>. This form has the advantage of providing a perfect semantic match for the Tocharian word. As the source of TB *ekita*, one can reconstruct a PK form \**əgīḍa-* or \**əkīḍa-*, with the regular developments PIr. \**-aya-* > OKh. *-ī-*, PIr. \**-VḍV-* > OKh. *-h-*. It is possible that the Tocharian form bears a trace of the initial syllable of the preverb \**haka(t)-*: TB *e-* could render the reduced vowel \**ə* (from PIr. \**ha-*). To account for the initial *e-* in the Tocharian form, the Author suggested an adverbial formation with the Tocharian prefix *e(n)-* added to a reconstructed obl.sg. *kita\**, which presupposes a nom.sg. form *kito\**. However, the derived TB *ekitatstse* ‘helpful’ and *ekītatsñe* ‘helpfulness, assistance’ do not appear to be built on a noun

<sup>8</sup> For a detailed analysis of this and other Khotanese words with a preverb *gga-* going back to PIr. \**hakat* (cf. Av. *hakat*) see FATTORI & MICHETTI (2024).

*kito\**, but rather to a form like *ekita*. This points to *ekita* being a noun rather than an adverb, as is suggested by the proposed connection with PK \**əgīḏa-* / \**əkīḏa-* ‘help’.

Summing up, the dating of some of the items discussed in §2.1 which the Author assigned to the most ancient layer of loanwords are perhaps better assigned to the later phase (30), or should be considered as generically prehistoric (3, 9). Only in one case, namely (16), I would go as far as to say that the proposed loanword from the Proto-Tumshuqese-Khotanese stage should not be considered as reliable as the book treats it.

These observations do not undermine the general picture presented in the following chapters dealing with the chronological and semantic classification of the loanwords. In Chapter 3, as noted above, the loanwords are distributed into four chronological phases. Each layer of borrowing is described listing the relevant phonological criteria (§3.2). From this classification, a set of phonological correspondences between the (reconstructed) Khotanese and Tocharian forms is extracted for each phase (§3.3), providing the reader with a practical tool to verify the various patterns of adaptation in the different periods. In §3.4 the borrowed nouns are grouped according to the Tocharian inflection class. The Author notes that nouns borrowed during the prehistoric stages of the language (PTK → PT, PK → Pre-TA, Pre-TB) were mostly adapted as members of the *-o/-a* class, whereas the *-o/-ai* class does not feature any loanword from the Proto-Tumshuqese-Khotanese stage. This aligns with the idea that the *-o/-ai* class is secondary (Del Tomba 2023: 160-166). More lists can be found in §3.5: according to part of speech, gender, and borrowing pattern.

Some of the correspondences identified in §3.2 for the PTK → PT phase are worth discussing in some detail. One such correspondences is TB *ś* ← PTK \**č*, which is established based on two proposed loanwords:

- (38) TB *śāmpo\** TA *śāmpām* ‘haughtiness, pride’ ← PTK acc.sg. *čamfu* ‘violence, disturbance’ (cf. OKh. *tcampha-* ‘disturbance, tumult’);
- (39) TB *śarko\** ‘song, singing’ ← PTK acc.sg. \**čarko* ‘song’, TA *tsārk* ‘song, singing’ ← PK \**tsarko* ‘song’ (cf. OKh. *tcar-kā-* ‘play, amusement’).

The Author argues that PIr. \*č was still a palatal sound in Proto-Tumshuqese-Khotanese, since it is reflected by TB ś. He suggests that PTK \*č may have entered in PT as \*č, later resulting in TB ś. The reconstruction of a palatal \*č for the Proto-Tumshuqese-Khotanese phase may appear surprising, since one of the isoglosses shared by Khotanese and Tumshuqese is the depalatalisation \*č > \*ts before non-palatal sounds (cf. Khot. *tcera-*, Tq. *tsārya-* < PIr. \*čārya-). Thus, one would assume the depalatalisation had already occurred in their common ancestor. However, the Author collects interesting evidence supporting this reconstruction, namely the couple of lexemes TB śarko\* ‘song, singing’ TA tsārk ‘id.’. Both terms are connected by the Author to OKh. *tcarkā-* ‘play, delight, amusement’, a noun going back to a PIr. root \*čarH- ‘to praise, celebrate’ (see above). The two Tocharian forms have different initial consonants (TB ś TA ts), pointing to two different dates of borrowing: one before the depalatalisation of \*č (TB śarko\* < PTK \*čarko), one after it (TA tsārk < PK \*tsarko). The only drawback of this hypothesis, as recognised by the Author himself (p. 177), is that the semantics of both Tocharian words is unclear, and the translation ‘song, singing’ should be viewed a working hypothesis.

Another phonological feature attributed to the Proto-Tumshuqese-Khotanese period has more impact on the patterns of morphological adaptation. As already mentioned, most borrowed nouns in Tocharian B have a final -o in the nom.sg. The Author considers it an adaptation of the endings -u, -o of the acc.sg. of the Khotanese *a-* and *ā-* stems respectively (from PIr. \*am and \*ām), and points out a parallel pattern of adaptation in the Middle Iranian loanwords in Armenian (§3.4.10). According to the chronological classification of Chapter 3, loanwords from three different stages were integrated in Tocharian B as nouns with a nom.sg. in -o: Old Khotanese, Pre-Khotanese and Proto-Tumshuqese-Khotanese. While this explanation works neatly for the Old Khotanese loanwords, and can be reasonably extended to the Pre-Khotanese period, I find it harder to apply it to the Proto-Tumshuqese-Khotanese loanwords, an Old Iranian language that is dated back to the first half of the 1st millennium BCE. In order to accept this hypothesis, one needs to assume that the acc.sg. endings PIr. \*am,

\*-ām were already developed into -u, -o, the value that they have in Old Khotanese more than a millennium later. Again, the possibility of a later date of borrowing may be envisaged, or, alternatively, there may have been other reasons for the adaptation as Tocharian nouns with nom.sg. -o.

Finally, the correspondence TB *rt* < PTK \**rd*, as suggested above, may perhaps be extended to the Pre-Khotanese phase.

The last three chapters add the final touches to the general picture, as the Author aims to connect the data collected in the preceding pages to the historical context. In Chapter 4, he notes that the loanwords belong to four semantic macro-areas: medicine (especially from Old Khotanese); administration, politics and economy (since prehistoric times); moral qualities; Buddhist terms (not earlier than the Old Khotanese stage, mostly found in Tocharian A). The presence of Buddhist loanwords from Khotanese into Tocharian A points to a direct contact between these two languages, a further argument to suggest that there existed a Khotanese community in the Tocharian A linguistic area, perhaps in Šorčuq. Chapter 5 has a similar structure to §2.1, but analyses the possible loanwords from Tocharian into Tumshuqese (1 item) and Khotanese (2 items). As the Tocharian loanwords into Khotanese are limited to learned Buddhist terms, it can be surmised that the Tocharian influence on Khotanese was minimal. Finally, Chapter 6 defines the contact situation between Tocharian and Khotanese as «slightly more intense» than the «casual» contact defined by Thomason and Kaufman in their classic work on linguistic contact (Thomason & Kaufman 1988), and as one of *adoption* rather than *imposition* according to the definition given by Haspelmath (2009).

Overall, this book provides the reader interested in the interactions between the Khotanese, the Tumshuqese and the Tocharian peoples within the Tarim Basin with all the relevant data available at present, presented in an insightful way and cleverly organised in order to allow any specific query to be answered quickly. While certain criticisms may be directed towards the manner in which the data are handled, particularly with regard to the phonological characteristics of the Proto-Tumshuqese-Khotanese phase, this weakness is probably

also its strength. Like with any significant development in a discipline, it presents a clear, systematic view, in which each piece of evidence fits neatly into the general picture. In this way, it encourages other scholars to investigate the topic further, introduce different perspectives, try to move pieces of the puzzle around, and perhaps mitigate some clear-cut statement. In short, this book will be an incentive and an obvious starting point for many future studies to come.

Francesca Michetti  
 Dipartimento Asia, Africa e Mediterraneo  
 Università di Napoli L'Orientale  
 francesca.michetti@unior.it

### *References*

- BAILEY, H. W. (1979), *Dictionary of Khotan Saka*, Cambridge, Cambridge University Press.
- BERNARD, CH. (2023), *Like dust on the Silk Road: An investigation of the earliest Iranian loanwords and of possible BMAC borrowings in Tocharian*, Doctoral Diss., Leiden University.
- DEL TOMBA, A. (2023), *The Tocharian Gender System. A Diachronic Study in Nominal Morphology*, Leiden - Boston, Brill.
- (2024), *Action nouns in \*-ia- and a new verbal stem jsīn- 'to kill' in Khotanese*, in «Bulletin of the School of Oriental and African Studies» 87.3, pp. 479-488.
- DRAGONI, F. (2021), *Materia medica Tocharo-hvatanica*, in «Bulletin of the School of Oriental and African Studies» 84.2, pp. 295-319.
- (2023), *"In the Tune of the King of Khotan": The name of Khotan in the Tarim Basin and beyond*, in «Journal of the American Oriental Society» 143.2, pp. 289-308.
- EMMERICK, R. (1968), *Saka Grammatical Studies*, London, Oxford University Press.
- FATTORI, M. & MICHETTI, F. (2024), *On some cognates of Avestan hakaṭ*, in «Indogermanische Forschungen» 129.1, pp. 213-240.
- HASPELMATH, M. (2009), *Lexical borrowing: Concepts and issues*, in HASPELMATH, M. & TADMOR, U. (eds.), *Loanwords in the World's Languages: A Comparative Handbook*, Berlin, Mouton de Gruyter, pp. 35-54.

- HITCH, D. A. (1990), *Old Khotanese synchronic umlaut*, in «Indo-Iranian Journal» 33.3, pp. 177-198.
- ISEBAERT, L. (1980), *De Indo-Iraanse Bestanddelen in de Tocharische Woordenschat*, Doctoral Diss., Leuven University.
- KÜMMEL, M. J. (2014), *Syllable- and word-related developments in earlier Indo-Iranian* (pre-print-version), [https://www.academia.edu/379001/Syllable\\_and\\_word\\_related\\_developments\\_in\\_earlier\\_Indo\\_Iranian](https://www.academia.edu/379001/Syllable_and_word_related_developments_in_earlier_Indo_Iranian).
- (2018), *The survival of laryngeals in Iranian*, in VAN BEEK, L., KLOEKHORST, A., KROONEN, G., PEYROT, M., PRONK, T. & DE VAAN, M. (eds.), *Farnab. Indo-Iranian and Indo-European Studies in Honor of Sasha Lubotsky*, Ann Arbor, Beech Stave Press, pp. 162-172.
- MAGGI, M. (2024), *Khotanese ph- < Iranian \*θ̥- . Late Khotanese phyṣṭa- and related words*, in «Indogermanische Forschungen» 129.1, pp. 165-183.
- PEYROT, M. (2015), *Tocharian language*, *Encyclopædia Iranica*, online edition, <https://www.iranicaonline.org/articles/tocharian-language>.
- (2018), *Tocharian B etswe 'mule' and Eastern East Iranian*, in VAN BEEK, L., KLOEKHORST, A., KROONEN, G., PEYROT, M., PRONK, T. & DE VAAN, M. (eds.), *Farnab. Indo-Iranian and Indo-European Studies in Honor of Sasha Lubotsky*, Ann Arbor, Beech Stave Press, pp. 270-283.
- , DRAGONI, F. & BERNARD, CH. (2022), *The spread of iron in Central Asia: On the etymology of the word for "iron" in Iranian and Tocharian*, in «Bulletin of the School of Oriental and African Studies» 85.3, pp. 403-422.
- SCHMIDT, K. T. (1985), *Zu einigen der ältesten iranischen Lehnwörter im Tocharischen*, in PIEPER, U. & STICKEL, G. (eds.), *Studia Linguistica Diachronica et synchronica*, Berlin - New York - Amsterdam, Mouton de Gruyter, pp. 757-767.
- SKJÆRVØ, P. O. (2004), *This Most Excellent Shine of Gold, King of Kings of Sutras. The Khotanese Suvarṇabhāsottamasūtra*, vol. 2: *Manuscripts, Commentary, Glossary, Indexes*, Cambridge (MA), The Department of Near Eastern Languages and Civilizations, Harvard University.
- THOMASON, S. G. & KAUFMAN, T. (1988), *Language Contact, Creolization and Genetic Linguistics*, Berkeley, University of California Press.

